
18. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V.

# Tagungsband

Herausgeber:
Oliver Burgert, Hochschule Reutlingen
Bernhard Hirt, Universität Tübingen

CURAC 2019

# Tagungsband

18. Jahrestagung der
Deutschen Gesellschaft
für Computer- und
Roboterassistierte
Chirurgie e.V.

19. – 21. September 2019, Reutlingen

**Impressum**

Herausgeber:
Oliver Burgert, Hochschule Reutlingen
Bernhard Hirt, Universität Tübingen

Hochschule Reutlingen
Fakultät Informatik
Alteburgstraße 150
72762 Reutlingen

Redaktion:
Hannah Büchner
Elena Kirsch
Johannes Schuh

Grafik:
Elena Kirsch

ISBN: 978-3-00-063717-9

**Gesellschaft**

Deutsche Gesellschaft für die Computer- und Roboterassistierte Chirurgie e.V.
CURAC Geschäftsstelle
Albstraße 45
70597 Stuttgart

Ansprechpartnerin: Frau Gabriele Schäfer
Tel.: +49 711 76 54 219
E-Mail: geschaeftsstelle@curac.org
www.curac.org

**Vostand**

Präsident:
Prof. Dr.-Ing. Stefan Weber
Informatik, Bern

Vizepräsident für Forschung:
Prof. Dr. med. Hubertus Alfons Ernst Josef Feußner
Chirurgie, München

Vizepräsident für Öffentlichkeitsarbeit:
Prof. Dr. med. Thomas Klenzner
HNO, Düsseldorf

Schriftführer:
Jun.-Prof. Dr. Christian Hansen
Informatik, Magdeburg

Schatzmeister:
Prof. Dr.-Ing. Oliver Burgert
Informatik, Reutlingen

Past-Präsident:
Professor Dr. med. Arya Nabavi MaHM
Neurochirurgie, Hannover

Ehrenpräsident:
Prof. Dr. med. Dr. h. c. mult. Madjid Samii
Neurochirurgie Hannover

Ehrenmitglieder:
Prof. Dr. Ron Kikinis
Briham and Women's Hospital Boston

Prof. Dr. med. Rudolf Fahlbusch
Neurochirurgie Hannover

**Tagungspräsidenten**

Oliver Burgert
Bernhard Hirt

**Vorsitzender des Programmkomitees**
Oliver Burgert

**Programmkomitee**

Oliver Burgert, Reutlingen
Claire Chalopin, Leipzig
Florian Dammann, Bern
Georg Eggers, Heidelberg
Rudolf Fahlbusch, Hannover
Hubertus Feussner, München
Stefan Franke, Leipzig
Wolfgang Freysinger, Innsbruck
Horst Hahn, Bremen
Christian Hansen, Magdeburg
Bernhard Hirt, Tübingen
Rüdiger Hoffmann, Tübingen
Florian Imkamp, Hannover
Lüder Kahrs, Toronto
Sven Kantelhardt, Mainz
Hannes Kenngott, Heidelberg
Ron Kikinis, Bremen
Thomas Klenzner, Düsseldorf
Uwe Kloos, Reutlingen
Christian Kücherer, Reutlingen
Wolfram Lamade, Überlingen
Dirk Lindner, Leipzig
Omid Majdani, Wolfsburg
Jürgen Meixensberger, Leipzig
Arya Nabavi, Hannover
Thomas Neumuth, Leipzig
Tobias Ortmaier, Hannover
Bernhard Preim, Magdeburg
Jörg Raczkowsky, Karlsruhe
Jörg Schipper, Düsseldorf
Alexander Schlaefer, Hamburg
Armin Schneider, München
Patrick Schuler, Ulm
Stefanie Speidel, Dresden
Simon Sündermann, Berlin
Christian Thies, Reutlingen
Gabriela Tullius, Reutlingen
Jürgen Wahrburg, Siegen
Stefan Weber, Bern
Antje Wermter, Tübingen
Rainer Wirtz, Ulm
Thomas Wittenberg, Erlangen
Stefan Zachow, Berlin

**Vorwort**

Liebe Teilnehmerinnen und Teilnehmer der CURAC Jahrestagung,
sehr geehrte Interessentinnen und Interessenten an den wissenschaftlichen Beiträgen,

die Deutsche Gesellschaft für Computer und Roboterassistierte Chirurgie e.V. (CURAC) richtet in diesem Jahr ihre 18. Jahrestagung aus – und wir freuen uns, dass die Hochschule Reutlingen als erste Hochschule für Angewandte Wissenschaften diese ehrenvolle Aufgabe übernehmen darf. Diese Hochschulen haben ein ähnliches Anliegen wie die CURAC-Gesellschaft: Die Translation von wissenschaftlicher Forschung in die Praxis zu fördern und neue Lösungen für drängende Fragestellungen zu entwickeln. Sowohl bei diesen Hochschulen als auch bei der CURAC zeigt sich, dass ein guter interdisziplinärer Austausch, bei dem unklare Fragen offen diskutiert werden und um beste Lösungen gerungen wird, der Schlüssel zum Erfolg sind. Das ist auch der Grund, warum wir für die diesjährige Tagung das Motto „Translation" gewählt haben: Aus der universitären Grundlagenforschung entstehen neue Methoden, die in angewandten Forschungsprojekten gemeinsam mit medizinischen Anwenderinnen und Anwendern und den technischen Fächern auf konkrete medizinische Fragestellungen angewandt werden, und so die Möglichkeiten der modernen Medizin erweitern. Als dritte Säule darf jedoch die industrielle Umsetzbarkeit im größeren Maßstab nicht aus den Augen verloren werden, denn was nützt die beste Technologie, wenn sie nur wenige Male unter erheblichem Mitteleinsatz genutzt werden kann und im Anschluss in einem Keller verschwindet.
Die CURAC Jahrestagung leistet hier seit Jahren hervorragende Beiträge zum interdisziplinären Austausch: Im Anschluss an die Vorträge werden neue Projektideen geboren, Methoden aus der einen Fachdisziplin auf andere übertragen, neue Forschungsteams bilden sich und nicht zuletzt entstehen auch Freundschaften und ein intensiver Austausch der die Fachcommunity trägt. Dass die Sektion für minimal-invasive Computer- und Telematik-assistierte Chirurgie der Deutschen Gesellschaft für Chirurgie (CTAC) seit Jahren die CURAC begleitet und dieses Jahr die Tagung als Rahmen für die Entwicklung eines Strategiepapiers nutzt ist ein weiteres Zeichen für die gute interdisziplinäre Zusammenarbeit.

| | |
|---|---|
| *Prof. Dr.-Ing.*<br>*Oliver Burgert* | *Prof. Dr.-Ing.*<br>*Stefan Weber* |
| Tagungspräsident | CURAC-Präsident |

**Wir bedanken uns bei den Sponsoren der 18. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V.**

The event is approved as COMPLIANT according to the MedTech Europe Code of Ethical Business Practice.

Ethical MedTech
MedTech Europe compliance portal

**Platinsponsor**

KUKA

BEC MEDICAL

**Silbersponsoren**

AXIOS 3D SERVICES

atracsys

**Weitere Sponsoren**

AKTORmed™ ROBOTIC SURGERY

CASCINATION

GADV Softwarelösungen

OR.NET e.V.

Cochlear®

**Weitere Unterstützer**

Hochschule Reutlingen
Reutlingen University

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

CHIRURGISCHE ALLGEMEINE
Konzentration auf das Wesentliche.

# Inhaltsverzeichnis

# Robotik & Navigation

# A novel robotic endoscope guidance system for surgery of the nasolacrimal duct

Schuler PJ[1], Sommer F[1], Boehm F[1], Friedrich DT[1,2], Scheithauer MO[1], Greve J[1], Hoffmann TK[1]

[1]*Department of Otorhinolaryngology, Head and Neck Surgery*

*Ulm University Medical Center, Frauensteige 12, 89075 Ulm, Germany*

[2]*Department of Otorhinolaryngology, Head and Neck Surgery*

*Augsburg University, Sauerbruchstraße 6,86179 Augsburg, Germany*

## SUMMARY

Transnasal surgery of the nasolacrimal duct is a common procedure in otorhinolaryngology. The endoscope holding system by Medineering® allows for hands-free visualization of the surgical field in the nasal cavity. In this clinical study, two patients with lacrimal duct stenosis were treated by a transnasal surgical approach. The endoscope holding system allowed for two-handed surgery which facilitated some steps of the surgical procedure. This may result in shorter operating times or better clinical outcomes.

## INTRODUCTION

Surgery of the nasolacrimal duct is a common procedure in otorhinolaryngology. Indications include tumors, idiopathic reasons and stenosis of the duct due to scarring or infection. In the latter case, the surgical approach to the nasolacrimal duct is determined by the location of the stenosis. If the stenosis is located proximal to the lacrimal sac, surgery is typically performed by accessing the stenosis via the lacrimal canaliculi. If the stenosis is located distal to the lacrimal sac, surgery is typically performed by a combined approach via the lacrimal canaliculi and the nose [1].

The endoscope holding system by Medineering® allows for hands-free visualization of the surgical field in the nasal cavity. The presented clinical study discusses advantages and drawbacks of this surgical assistance tool.

## MATERIAL and METHODS

The setup of the endoscopic system has been published before [2]. It consists of an intelligent mechatronic holding arm with four segments and seven degrees of freedom (**figure 1a)**. It is driven manually by the surgeon and can be locked in any possible position. The maximum load capacity is 2kg. With an overall weight of less than 10kg, the construction is highly versatile and can be easily attached to standard side rails of a surgical table, this in any position, according to the respective procedure or task. Joints of the holding arm are released by touchpads on each segment at any time throughout the procedure. The status of the system (locked/released) is visualized by LED lights. On its distal end, a compact robotic hand with five actuated degrees of freedom performs the movement of the endoscope, driven by five brushless DC motors. Standard 4mm endoscopes are connected to the robotic hand with a specific clip mechanism. 3-dimensional motion and fine adjustment of the endoscope are controlled by the surgeon through a custom foot pedal with joystick (Steute Inc., Loehne, Germany)**.** The pedal also allows to home the endoscope to a basic position with an extra foot button. The surgeon can switch between transitional movement and pivot point rotation of the endoscope.

The system was used for the transnasal steps of a dacrocystorhinoscopy (DCR) in two patients (37 and 57 years) with post-saccal stenosis of the nasolacrimal duct (**figure 1b**).

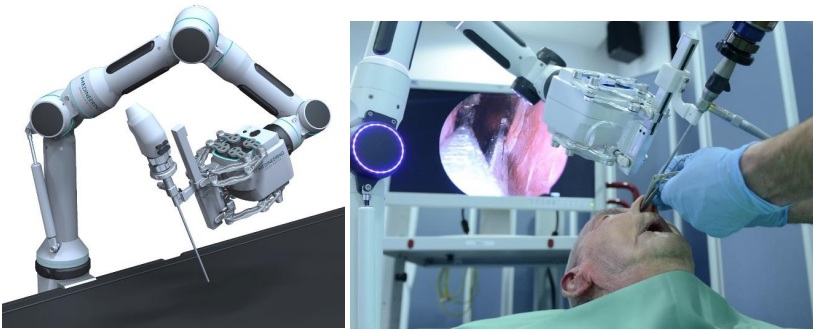


**Figure 1:** (a) Overview of the endoscopic guidance system. (b) Setup of the system for two-handed transnasalsurgery [2].

## RESULTS

Visualization and instrumentation of the surgical field were feasible with the presented setup. Controlling the robotic endoscope guidance system was adequately precise. The surgical DCR procedure contains the following steps:

(I) Transnasal preparation of the nasal mucosa over the lacrimal bone
(II) Transnasal opening of the lacrimal bone with the chisle
(III) Dilatation of the lacrimal canaliculi and introduction of the metal probes
(IV) Transnasal opening of the lacrimal sac with the knife
(V) Fixing the probes in the nasal cavity using clips and sutures

All steps, excluding #III, benefit from the use of the endoscopic system, as two-handed manipulation is advantageous for the surgeon (**figure 2**). The surgical field in the nasal cavity is very limited in its size and easy to reach with the endoscopic system. In addition, the nasal cavity has a large diameter in the proximity of the lacrimal bone. Therefore, interference of the endoscopic system with the rigid tools is less common as compared to other locations, e.g. skull base. Further miniaturization of the system should be pursued, as in some cases access to the nostrils with tools and the endoscope appears to be cumbersome.

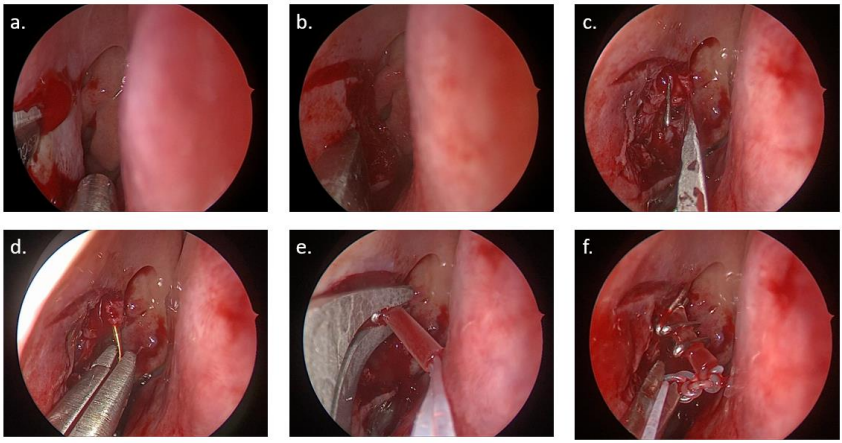



**Figure 2:** Steps of the surgical procedure. (a) lifting of nasal mucosa; (b) opening of the lacrimal bone with chisel; (c) opening the lacrimal sac; (d) relocating the metal probe into the nasal cavity; (e+f) fixing the probes by clips and knots.

## CONCLUSION

Using a robotic endoscope guidance system for transnasal surgery of the nasolacrimal duct is beneficial for the surgeon in a clinical setting. This may result in shorter operating times or better clinical outcomes.

## REFERENCES

1. Trimarchi M. Dacryocystorhinostomy: Evolution of endoscopic techniques after 498 cases. Eur J Ophthalmol 2019; doi: 10.1177/1120672119854582.
2. Friedrich DT, Sommer F, Scheithauer MO, Greve J, Hoffmann TK, Schuler PJ. An Innovate Robotic Endoscope Guidance System for Transnasal Sinus and Skull Base Surgery: Proof of Concept. J Neurol Surg B Skull Base 2017; 78: 466-472.

# Feasibility of Pediatric Robotic Cochlear Implantation in Phantoms

*Gabriela Bom Braga[1], Daniel Schneider[1], Fabian Mueller[1], Jan Hermann[1], Stefan Weber[1], Marco Caversaccio[2]*

[1]*ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*

[2] *Inselspital University of Bern*

Contact: gabriela.braga@artorg.unibe.ch

***Abstract***

*Robotic cochlear implantation is a minimally invasive surgical technique that is already being used in adults, with promising results. Therefore, the feasibility of the application of this surgical procedure in the paediatric population is of interest. With this work, we investigated the feasibility of the robotic surgical approach in patient specific phantom models and test the current workflow in the paediatric population.*

**Keywords**: Cochlear implant, pediatrics, robotic cochlear implantation

## 1 Problem

Newborn hearing screening as defined by the Joint Committee on Infant Hearing, has led to earlier identification of children with significant hearing impairment[1]. Current strategies can routinely detect hearing loss in children under the age of 12 months of age. Studies have shown that early hearing rehabilitation through cochlear implantation (CI) facilitates developmental outcomes in line with normal-hearing peers[2]. Hay-McCutcheon et al. (2008) demonstrated statistical and practical improvement in receptive language development each additional, prospective year a pediatric patient received a cochlear implant[2,3].

The evidence of greater developmental benefit of earlier CI surgery means that CI for patients under the age of 24 months is now routine practice; and potential minor, long-term variations in audiological diagnoses are far outweighed by normative hearing, language and psychosocial development of deaf children with cochlear implants[4].

While the first and second segments of the facial nerve in children are similar to adults in length and width, the mastoid section is more horizontal in adults than during early childhood. More over facial nerve length in the mastoid is different in adults and children, as well as at the landmark positions of the stylomastoid foramen and the chorda tympani nerve[4,5]. Postnatal development of the facial canal is associated with significant changes of the temporal bone and most significant changes take place during the first four years after birth. On the other hand, the inner ear, the tympanic cavity, the mastoid antrum, the ossicles and the tympanic membrane already are adult size in newborns. Furthermore, the mastoid process and the osseous part of the external ear canal start to develop after birth[4,5]. Of note is that the chorda tympani lies on the external surface of the skull, close to the stylomastoid foramen and it is only enclosed in the tympanic bone during the first year. It indicates that the facial canal grows more than the facial nerve in its mastoid portion resulting in the site of the chorda tympani bifurcation shifting upwards with age relative to the stylomastoid foramen. This is an anatomical feature that is of critical importance when planning facial recess access for CI surgery[5]. In addition, infants and young toddler candidates demonstrate greater variability in the orientation of the cochlea within the cranium, which will make greater determination for the insertion process compared to adults, where the cochlea has settled in a population-average orientation[6].

In the majority of cases, mastoidectomy[7] is the most time-consuming and invasive component of cochlear implantation surgery. Therefore, minimally invasive approaches such as a keyhole tunnel that is drilled through the facial recess, have been proposed and are currently evaluated in adults[8,9,10]. So to consider an analogous, minimally invasive approach in children, virtual simulation of planning and conduction of tunnel trajectories for middle ear access through the facial recess at safe distances to the facial nerve have been carried out. These virtual studies suggested general feasibility for a robotic cochlear implantation in children[11]. However, physical drilling

of a robotic tunnel trajectory for direct cochlear access has never been tried in the infant anatomy to date. Hence, the primary aim of this study was to investigate feasibility of a computer-assisted planning and drilling of a robotic middle and inner ear access using pre-existing computed tomography image data sets of pediatric subjects, from which 3D phantoms were created, followed by positioning study and insertion of the electrode array.

## 2 Materials and Methods

With permission of the local institutional review board (KEK 2017-01722), CT image data from 10 subjects meeting inclusion criteria (8-48 months at the time of surgery, CI surgery at our department between 2014 and 2017) were randomly selected, resulting in n=20 sides and included to the study protocol.

### 2.1 Trajectory Planning

Using a commercially available, task-specific software planning system[12] (OTOPLAN®, CAScination AG, Switzerland, Figure 1-A) an experienced otologist identified on the CT data and, 3D reconstructed relevant anatomy including: facial nerve, chorda tympani, incus and malleus, stapes, the external auditory canal, and the round window (Figure 1-B). In addition, for each dataset, a surgical plan was created that contained a 1.8 mm wide trajectory that started at the cortex of the temporal bone, passed through the facial recess[13], and ended at the center of the bony overhang (round window approach). The distances between the trajectory and the segmented surrounding anatomical structures as well as the length of the trajectory were automatically determined by the planning software.

In addition, cochlear parameters such as the orientation of the basal turn and its relation to the cranial base described as round window angle (Figure 1-C) and the angle between the first and second turns of the cochlea (Figure 1-D) were measured (Amira®, Thermo Fisher Scientific).

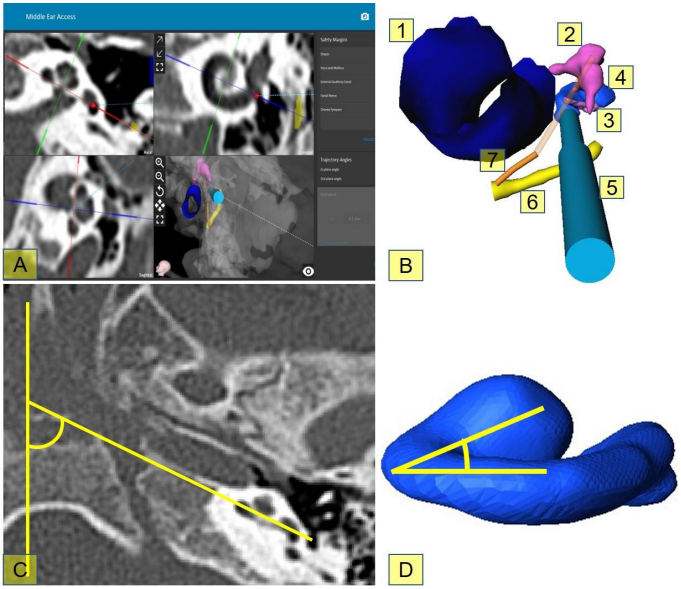



*Figure 1: A- Middle ear access planning of case 3L, using OTOPLAN. B- Illustration of the three-dimensional surface models of the segmented temporal bone structures: (1) external auditory canal, (2) incus and malleus, (3) stapes, (4) cochlea, (5) the trajectory for robotic middle ear access, (6) facial nerve, and (7) chorda tympani nerve; C- basal turn and its relation to the cranial base; D- Cochlear angle, between the first and second turns.*

### 2.2 Assessment of Mastoid thickness for safety of fiducial screw implantation

To investigate the feasibility of placing bone fiducials screws, a sub-study (n=20) on available bone thickness in the mastoid region was conducted. Available datasets were grouped into 4 groups of different ages (group 1: 8-10 months; group 2: 13-14 months; group 3: 24-25 months; group 4: 48-49 months). Available bone thickness in the mastoid region and perpendicular to the mastoid surface was measured (Amira®, Thermo Fisher Scientific) and co-registered to one model per respective group (whereas right sides were mirrored before co-registration) to create an age specific statistical model of the available bone thickness. Available thickness of the mastoid bone at the potential fiducial insertion sites was studied.

**Phantom Preparation**

From the 20 available planned sides, 8 sides representing the bandwidth of the growth process of the temporal bone covering the age range (10, 14, 24 and 48 months) where anatomical changes are occurring and are more prominent, were randomly selected for further creating phantoms and carrying out drilling experiments. From the available clinical CT image data, the temporal bone regions of both sides were extracted, reproduced via 3D printing (Eden260VS, Stratasys, MN, USA) and integrated into plastic replicas of skulls of similar ages (Figure 2-A, Bone Clones Inc., USA). The resulting phantom combined subject individual properties of the mastoid relevant for investigation of robotic drilling and electrode insertion with the macroscopic anatomical properties of pediatric skulls. Phantoms where then filled with silicone for weight emulation and attached to a pediatric CPR training phantom (two models of age 6 and 38 months). Each of the 8 sides were implanted with 4 titanium reference screws (HEARO Fiducial screw, CAScination AG, Switzerland) on the mastoid tip, posterior to MacEwen's triangle (temporal line superiorly tangent to EAC and postero-superior rim of the canal) and parallel to the temporal line and superior and posterior to the spine of Henle. CB-CT images ($0.15 \times 0.15 \times 0.2$ mm$^3$, XCAT XL, Xoran, USA) of the mastoid region were acquired. By means of registration through a mutual information approach (Amira®, Thermo Fisher Scientific), the existing plan (of the previous clinical case) was co-registered to the available CB-CT scan (of the 3D phantom created from that clinical case).

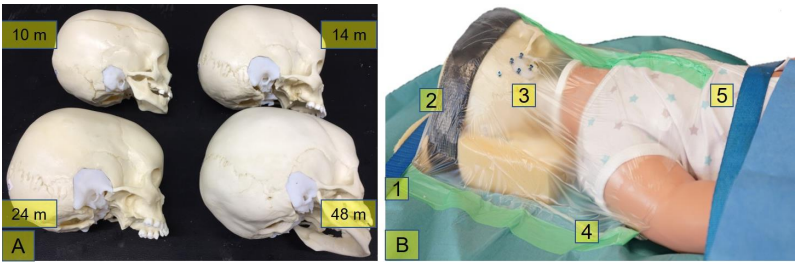



*Figure 2: A- Pediatric age specific mastoid phantoms with skull, labels referring to phantom age. B- 1- Fixation of the horseshoe head support; 2- Headband; 3- Fiducial screws and reference marker screw (more posterior); 4- surgical patient drape head and torso fixation; 5-subject containment strap.*

**Head Support and Fixation**

Effective head fixation is a prerequisite for providing accuracy during the robotic drilling process. Considering the relative and delicate flexibility of a pediatric skull, the use of simple horseshoe shaped cushion together with a head band, both fixated with a SteriDrape® around head and torso (Figure 2-B), was considered to be safe and effective, and to accommodate for potential intubation during an actual surgery.

**Robotic Drilling**

Drilling was carried out using a robotic cochlear implantation system (HEARO, CAScination AG), based on previously reported concepts[7,8,9,13,14,15]. A patient marker base was attached on the parietal border of the squamous bone, approximately 20 millimeters away from the most posterior screw. The phantom was registered to its corresponding planning model using the available fiducial screws[16] (maximum acceptable fiducial registration error FRE: < 0.05 mm). Robotic drilling process commenced (Figure 3-A) using the systems proprietary step spiral drill (1.8 mm) and a previously described protocol (feed forward rate: 0.1 mms$^{-1}$, rotation speed: 1000 min$^{-1}$, incremental pecking depth: 1 mm, Weber et al 2017[7]). Using system specific diamond burr (1.0 mm) access to the cochlea was gained (feed forward rate: 0.02 mms$^{-1}$, rotation speed: 1000 min$^{-1}$). The system limits the axial milling force to 2.0 N through control of the effective feed forward rate. Milling is stopped automatically if a size of the

opening sufficient for electrode insertion has been achieved. In a clinical setting, perforating the fluid-filled hearing cochlea will be avoided.

Upon completion of inner ear access, screws, patient marker, surgical patient drape and head fixation were removed, and the tunnel was cleaned. A system specific titanium guide tube[17] was used as a protective barrier and placed in the tunnel to prevent contact of the electrode with the drilled tunnel and avoid deviations during insertion (Figure 3-B; 3-C). A FLEX28 (MED-EL, Innsbruck, Austria) was then inserted manually into the cochlea through the pre-lubricated guide tube. The insertion was stopped upon the first point of resistance or, when the electrode stopper reached the cochlear opening indicating the full insertion.

Post-operative CB-CT scans images ($0.15 \times 0.15 \times 0.2$ mm$^3$, XCAT XL, Xoran, USA) were acquired and co-registered (using the fiducial screws) with the preoperative CB-CT scans of the phantoms (from prior to the experiment) to investigate available accuracy and precision of the robots geometric drilling performance with respect to the planned drill tunnel.

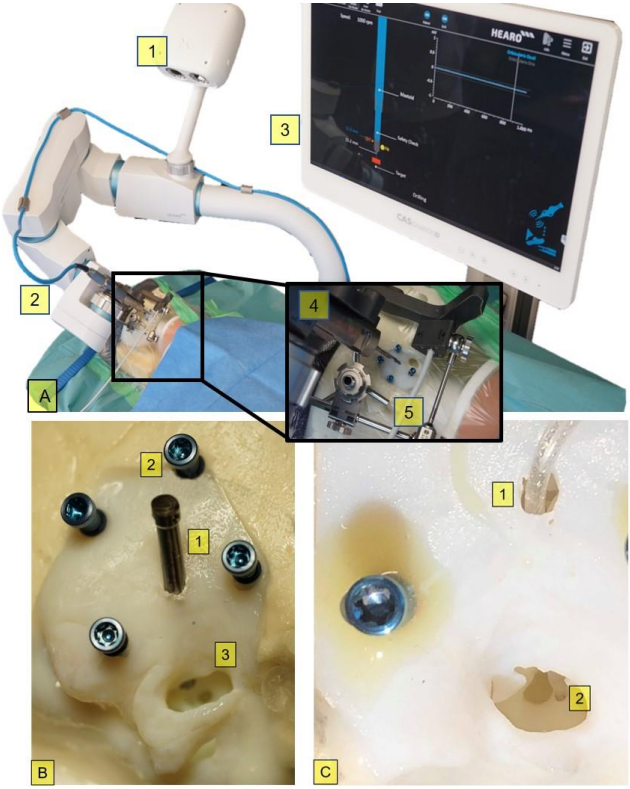



*Figure 3: A- Robotic cochlear implantation system (HEARO®, CAScination AG), (1) Tracking camera (2) Robot (3) user interface (4) Navigated drill bit and (5) Navigated patient reference. B- 1- Insertion tube; 2- Fiducials; 3- External Auditory Canal; C- 1- Electrode going inside the DCA tunnel; 2- Electrode entering the extended round window drilled access.*

# 3 Results

With this study, we were able to demonstrate that it is possible to plan a safe trajectory from the cortical of the temporal bone to the cochlea, maintaining a safe distance from the surrounding structures (17/20 sides) and, in 3 cases (15%) a safe trajectory could not be planned with a 1.8 mm drill tunnel due to a limited distance to the facial nerve (<0.4 mm)[13]. We further studied the behavior of the development of the morphology of the facial recess and, were able to follow its growth progression (Figure 4- A). In addition, the cochlear angle analysis demonstrated that with age, both the round window and cochlear angles decrease, making insertion smoother and diminishing the chance of trauma to the basilar membrane (Figure 4- A).

Investigation of available bone thickness indicated sufficient bone thickness of <4.6 mm for the placement of fiducials screws in general. The placement pattern of the screw fiducials according to the instructions for use of the robotic system (Figure 4- B) could be followed. In age groups 3 and 4, two fiducial screws had to be placed in areas with a thickness of less than 4.6 mm. Retrospectively, CT imaging confirmed potential penetration of those fiducial screws towards the proximal skull wall.

Accuracy and precision of the robotic drilling process (n=8) was found to be 0.15 mm (SD = 0.03) and 0.07 mm (SD = 0.06 mm) at the target and at the level of the facial nerve respectively. Full insertion was achieved in 8 of 8 cases (Figure 5).

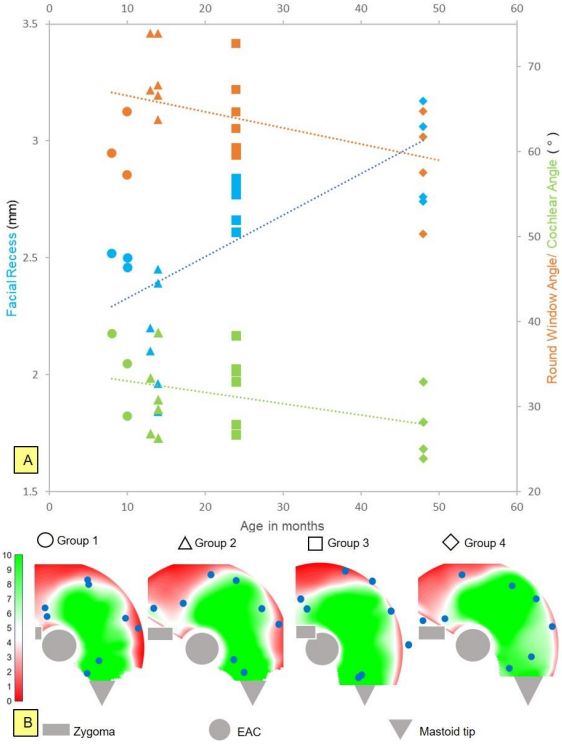



*Figure 4: A- Facial recess width, Cochlear angle and Round Window angle during growth process; B- Average mastoid thickness according to age groups. Green area with bone thickness >10 mm; Red area < 1.0 mm; Blue dots- screw implantation area.*

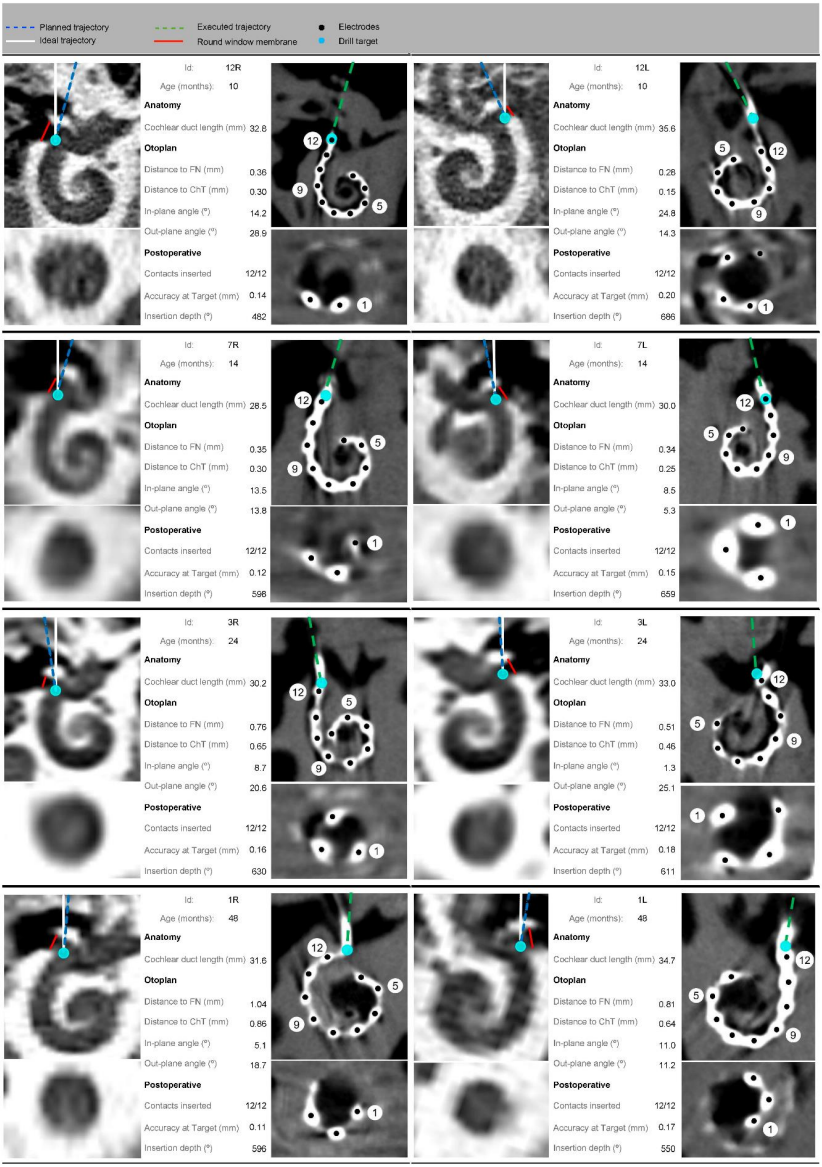



*Figure 5: Drilling and insertion results*

## 4 Discussion

This study highlights the importance of the visual design in positively impacting the user experience of a robotic surgery platform of the ear that must exist in real-time. While the technology behind many novel surgical platforms is highly advanced, some systems like the one used in this study are not fully realized and warrant further development. While unnaturally colored 3D-generated models of anatomy work well in cases in which one must distinguish between different components of the anatomy (for example highlighting a hepatic tumor in green against a naturally colored or grayscale liver), this is not the case in robotic cochlear implantation. In this case, the physical relationship between the structures (needed to program the robotic trajectory) is more important than the need to distinguish between individual anatomical structures. Examples of extremely advanced graphical renderings of anatomy can be found in surgical training/educational platforms like those used in the Touch Surgery app [5] or LevelEx surgical training platforms [6]. Perhaps the same graphical design methods could be applied to platforms for surgical planning and robotic surgery to improve the overall user experience.

## 5 Conclusion

To further validate the proposed procedure in the pediatric population, optimization of the current surgical workflow and adaptation of the surgical materials to the pediatric population should take place in order to increase the safety of the procedure.

## References

1. Holman MA, Carlson LM, Driscoll CLW, et al. Cochlear implantation in children 12 months of age and younger. Otol Neurotol 2013;34:251-258.
2. Svirsky MA, Teoh SW, Neuburger H, Development of language and speech perception in congenitally. profoundly deaf children as a function of age at cochlear implantation. Audiol Neurootol 2004;9:224Y33.
3. Hay-McCutcheon MJ, Kirk KI, Henning SC, et al. Using early language outcomes to predict later language ability in children with cochlear implants. Audiol Neurootol 2008;13:370Y8.
4. James LA, Papsin BC. Cochlear implant surgery at 12 months of age and younger. Laryngoscope. 2004;114:2191-2195.
5. Weiglein AH. Postnatal development of the facial canal an investigation based on cadaver dissections and computed tomography. Surg Radiol Anat 1996;18:115-123.
6. Martinez-Monedero R, Niparko JK, Aygun N. Cochlear coiling pattern and orientation differences in cochlear implant candidates. Otol Neurotol 2011;32:1086-1093.
7. House W.F. Cochlear Implants. Ann Otol Rhinol Laryngol. 1979;85(27): 1–93.
8. Weber S, Gavaghan K, Wimmer W, et al. Instrument flight to the inner ear. Sci Robot 2017;2. eaal4916.
9. Caversaccio M, Gavaghan K, Wimmer W, et al. Robotic cochlear implantation: surgical procedure and first clinical experience. Acta Oto-Laryngologica 2017;137:4. 447-454.
10. Labadie RF, Balachandran R, Noble J, et al. Minimally Invasive Image- Guided Cochlear Implantation Surgery: First report of clinical implementation. The Laryngoscope 2014. 124:1915-1922.
11. Braga GPB, Gebrin E, Balachandran R, et al. Evaluation of the facial nerve recess and cochlea on the temporal bone of stillbirths regarding the percutaneous cochlear implant. Int Arch Otorhinolaryngol 2017; DOI https://doi.org/10.1055/s-0037-1606612.ISSN 1809-9777.
12. Gerber N, Bell B, Gavaghan K, et al. Surgical planning tool for robotically assisted hearing aid implantation. Int J CARS 2014;9:11-20.
13. Ansó J, Dur C, Gavaghan K, et al. A neuromonitoring approach to facial nerve preservation during image-guided robotic cochlear implantation. Otol Neurotol 2015;37:89-98.

# Computer-navigated, stereotactic, percutaneous treatment of abdominal Type II endoleaks

*Nando Mertineit[1], Carlo Tappero[1/2], Gerd Nöldge[1], Frank Mosler[1], Florian Dammann[1], Vladimir Makaloski[3], Iris Baumgartner[4], Johannes Heverhagen[1], Miltiadis Krokidis[1]*

*[1]Universitätsinstitut für Diagnostische, Interventionelle und Pädiatrische Radiologie, Inselspital, University Hospital Bern, Schweiz*
*[2]Interventionelle Radiologie, Kantonsspital Fribourg, Schweiz*
*[3]Universitätsklinik für Herz- und Gefässchirurgie, Inselspital, University Hospital Bern, Schweiz*
*[4]Universitätsklinik für Angiologie, Inselspital, University Hospital Bern, Schweiz*

Contact: nando.mertineit@insel.ch

***Abstract***

Endoleaks are one of the most frequent problems that occur after endovascular aneurysm repair. We developed a CT-guided, computer-navigated, stereotactic percutaneous approach to treat patients with endoleaks without administration of intravasal contrast medium. We used a CT-navigation system to position the percutaneously advanced needle precisely in the endoleak. In all cases we treated a successful placement of the needle in the endoleak nidus was achieved. Only 1 needle repositioning was necessary. Onyx administration in the endoleak was successful in all cases. Intravasal contrast agent administration was not necessary. With this promising new percutaneous, CT-navigated approach we are able treat endoleaks without administration of intravasal contrast agent in a quick and safe procedure.

**Keywords**: Endoleak, Embolisation, CT-navigation

## 1 Problem

Endoleaks are one of the most frequent problems that occur after endovascular aneurysm repair (EVAR) [1]. Many techniques where proposed to treat endoleaks: e.g.: by transarterial embolization, feeder vessel ligation, direct sac puncture or by aneurysm sac plication [2]. It has been shown, that the transarterial access as well as the direct percutaneous puncture of the sac are both safe and efficient techniques to treat type II endoleaks [3] although it is still a topic of debate if and when pts. with type II endoleak need treatment [4, 5]. The success of the embolization depends mainly on the fact whether it is possible to reach the nidus of the endoleak or not [6]. Especially, when lumbar arteries are causing Type II endoleaks, it is often difficult to reach the nidus transarterially and therefore the embolization has a poor success rate in these cases [7]. Many patients with EVAR have an impaired renal function [8], what makes intravasal contrast agent administration difficult.

To treat these patients without applying intravasal contrast agent, we developed a CT-guided, computer-navigated, stereotactic percutaneous approach, to place a needle percutaneously in the center of the nidus for the embolization.

## 2 Material and Methods

Cases were discussed in an interdisciplinary setting. We treated patients for which a treatment was considered necessary due to the enlargement of the aneurysmatic sac and the transarterial treatment was not possible or obviously insufficient. We used a CT-navigation system (CASCination, Bern, Switzerland,) in which we imported previously performed CT images, which showed the endoleak. Diagnostic images were then fused with a nonenhanced planning CT. Hence, it was not necessary to again administer intravenous contrast agent.

Fig 1*: Left: Image Fusion and planning of the needle path / Right: percutaneously placed needle and control after Onyx administration*

Percutaneous access and needle tip placement in the endoleak were planned by means of the software. After the guided puncture with a Chiba-Needle and confirmation of the correct needle tip positioning in the center of the endoleak nidus, the endoleak was embolized with Onyx. The clinical follow-Up was done by contrast enhanced Ultrasound (CEUS).

## 3 Results

In our limited trial we treated 4 patients. The endoleak nidus was successfully punctured in all cases. The needle had to be repositioned in 1 case. Onyx was successfully applied in all cases. The mean intervention time was 68 min (48 – 106 min); the mean time to correct needle placement was 38 min (24 – 65 min). We did not administer intravenous contrast agent, in 1 case contrast medium was administered directly into the endoleak nidus to visualize the exact extend. There was no bleeding from the aneurysm sac. 1 patient developed an extraabdominal, intramuscular hematoma in the pathway of the needle access. There were no other minor or major complications. The mean follow-up for 3 patients was 18.6 months (10-26 months) with a stable extension of the aneurysma (N = 2) or a shrinking aneurysm sac (N = 1) in all of them. One patient had no follow-up, because of the fact, that the intervention was done recently.

## 4 Discussion

Our study is limited due to the very small number of patients and the lack of a complete follow-up. Nevertheless, we could show, that placing a needle percutaneously in the nidus of the endoleak using a 3D-stereotactic-guidance system is a relatively easy, safe and in all cases successful procedure. Because of the short intervention time, it reduces radiation exposure to the patient and the radiologist. On top of this, a short intervention time will obviously reduce procedure and anesthesia related complications. By using the navigation system, we were able to reduce the required CT-Scan controls and necessary needle repositions to a minimum. Because we fused previously acquired diagnostic CT Images with a native planning CT we were able to complete the whole procedure without administration of intravenous contrast agent, which is a benefit for those patients with an impaired renal function.

## 5 Conclusion

This computer-navigated, stereotactic, percutaneous treatment of abdominal endoleaks without intravasal contrast agent administration is a new and an up to our limited experience easy and safe procedure. It promises to be especially an option for those patients with an impaired renal function and specially, when transarterial embolization is not possible or not successful.

## References

[1] Greenhalgh RM et al.; Endovascular versus open repair of abdominal aortic aneurysm; N Engl J Med. 2010 May 20;362(20):1863-71

[2] Moulakakis KG, et. al; Treatment of Type II Endoleak and Aneurysm Expansion after EVAR; Ann Vasc Surg. 2017 Feb; 39:56-66.

[3] Yang RY, et al.; Direct sac puncture versus transarterial embolization of type II endoleaks: An evaluation and comparison of outcomes; Vascular. 2017 Jun; 25(3):227-233.

[4] D'Oria M et al; Natural history, diagnosis and management of type II endoleaks after endovascular aortic repair (EVAR): review and update.; Ann Vasc Surg. 2019 Jul 31. pii: S0890-5096(19)30504-7.

[5] Jones et al; Persistent type 2 endoleak after endovascularrepair of abdominal aortic aneurysm is associated with adverse late outcomes; J Vasc Surg 2007; 46:1-8.

[6] Yu H et al; Comparison of Type II Endoleak Embolizations: Embolization of Endoleak Nidus Only versus Embolization of Endoleak Nidus and Branch Vessels.; J Vasc Interv Radiol. 2017 Feb; 28(2):176-184.

[7] Aziz A.; Outcomes of percutaneous endovascular intervention for type II endoleak with aneurysm expansion; J Vasc Surg. 2012 May; 55(5):1263-7.

[8] Kauvar DS et al; Complication profile, failure to rescue, and mortality following elective endovascular aortic aneurysm repair; Am J Surg. 2017 Aug;214(2):307-311.

# Robotic Positioning for Image-Guided Ultrasound Interventions

*Michael Unger*[1], *Johann Berger*[1], *Lisa Landgraf*[1], *Andreas Melzer*[1]

[1]*Innovation Center Computer Assisted Surgery, University Leipzig, Leipzig, Germany*

Contact: michael.unger@medizin.uni-leipzig.de

***Abstract***

Robotic assistance during clinical interventions provides high precision and performance. Image guidance offers feedback during the procedure in real-time. A system for the robot-assisted positioning of an imaging ultrasound was developed and evaluated for an image-guided biopsy use case. The overall mean targeting error was 1.02 mm ± 0.58 mm. In future applications, any other tool can be placed to enable image-guided procedures.

**Keywords**: robotic surgery, image-guided intervention, focused ultrasound, biopsy

## 1 Problem

Robotic systems offer high precision and repeatability and may improve the clinical workflow, especially in places with restricted access. Robotic systems are used, e.g. during radiation therapy to precisely deliver radiation doses [1]. Focused ultrasound (FUS) is able to generate hyperthermia and mechanical effects in a non-invasive way and may, therefore, prove to be a method to improve radiation therapy (RT) [2,3]. Positioning the ultrasound transducer can be challenging and lacks reproducibility. Robotic arms provide the means of versatile placement of tools in regions that are difficult to reach, especially in a LINAC (linear particle accelerator). Currently available FUS-systems are either integrated into an MRI (Magnetic Resonance Imaging) or used under confocal US guidance [4]. Due to radiation protection regulations, manual positioning of FUS is restricted during radiation therapy. Therefore, new approaches are needed to make FUS available during RT. The aim of this work is to investigate the accuracy and performance of a commercial robotic system (KUKA LBR iiwa) to move different end-effectors to a target region, based on augmented reality anatomical images. In this first step, the tool positioning was evaluated for an ultrasound-guided biopsy in a phantom.

## 2 Material and Methods

To perform flexible positioning of a US imaging probe, a KUKA LBR iiwa 7 R800 (KUKA AG, Augsburg, Germany) robotic setup as described in [5] was used in this study (Fig. 1). A Clarius L7 mobile imaging US device (Clarius Inc., Canada) was attached to the robot and made the imaging data available via a tablet application (Apple iPAD pro 2017). A software module to acquire and process the US images in real time was implemented in C++ on a standard PC using the Listen application programming interface (API) provided by the vendor.

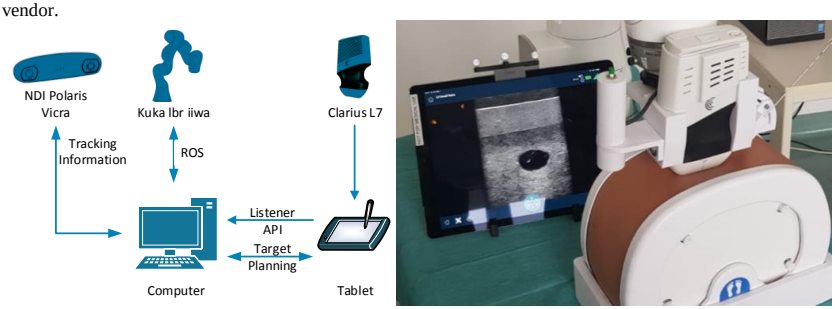



Fig. 1 *Schematic of the experimental setup (left) and actual setup. Robotic arm with attached US device, phantom, and tablet running the US imaging application (right).*

To simulate a target for the robot two lesions in a phantom (Abdominal Triple Modality Phantom, CIRS Inc., USA) were targeted. An augmented reality-based tablet application was used to target the lesions inside the phantom and position the US device attached to the robot on the phantom (Fig. 1). The tablet and therefore the target trajectory was acquired via a tracking system (Polaris Vicra, NDI, Canada). The lesions were segmented using both, pre-operative CT data (Philips iCT 256, Amsterdam, Netherlands) and intraoperative US images. The US images were processed using the computer vision framework OpenCV. The images were smoothed by applying a median blur filter with a kernel size of 15 pixels to reduce noise followed by a threshold segmentation to detect the target. An ellipse was fitted to the detected target to acquire the center position of the lesion as well as the dimensions. The center of the target was used as an input for adjusting the 3D printed needle guide.
We conducted a study to assess the system accuracy. Therefore, the US imaging probe positioned above the target by steering the robot using the augmented reality based application visualizing the lesions inside the phantom. After the target was within the US image, and the parameters (imaging depth, brightness) of the US device were set, the robotic system held the position while the biopsy needle was inserted. The distance between the desired target (center of the segmented lesion) and the reached target (trajectory of the needle) was measured manually.

## 3 Results

Two different lesions were targeted 5 times each with varying positions of the US probe. The depth of the lesion varied between 29.3 mm and 40 mm (center position). The pixel resolution was 102 µm per pixel for the images capturing lesion 1, and 122 µm per pixel for the images capturing lesion 2. An example of the target, the planned and actual needle trajectory is shown in Fig. 2. The overall mean targeting error was 1.02 mm ± 0.58 mm. The mean targeting error of lesion 1 was 0.51 mm ± 0.26 mm, the mean targeting error of lesion 2 was 1.53 mm ± 0.21 mm.

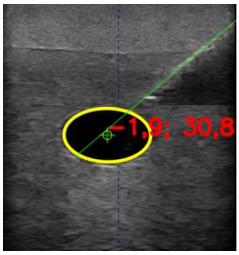



Fig. 2: *Ultrasound image of the phantom. The contour of lesion 1 (black) was automatically extracted (yellow ellipse). The center of the lesion (green cross) was the desired target.*

## 4 Discussion

This study assessed the positioning accuracy of an image-guided robot-assisted biopsy. A mean targeting error of 1.02 mm shows that the system can be used for the precise positioning of a biopsy needle. The actual trajectory of the needle was extracted manually and is therefore subject to errors. But the results provide a basis for future applications including the automatic positioning of a FUS transducer to apply hyperthermia.

**Author Statement and Acknowledgement**
Research funding: The research leading to these results has received funding from Bundesministerium für Bildung and Forschung (BMBF) under grant No.03Z1L511 (SONO-RAY project). Conflict of interest: Authors state no conflict of interest.

## References

[1] Coste-Manière È, Olender D, Kilby W, Schulz RA (2005) Robotic whole body stereotactic radiosurgery: clinical advantages of the Cyberknife integrated system. The International Journal of Medical Robotics and Computer Assisted Surgery 1:28–39. https://doi.org/10.1002/rcs.39

[2] Horsman MR, Overgaard J (2007) Hyperthermia: a Potent Enhancer of Radiotherapy. Clinical Oncology 19:418–426. https://doi.org/10.1016/j.clon.2007.03.015

[3] Maloney E, Hwang JH (2015) Emerging HIFU applications in cancer therapy. International Journal of Hyperthermia 31:302–309. https://doi.org/10.3109/02656736.2014.969789

[4] Peek MCL, Wu F (2018) High-intensity focused ultrasound in the treatment of breast tumours. Ecancermedicalscience 12:. https://doi.org/10.3332/ecancer.2018.794

[5] Berger J, Unger M, Landgraf L, Bieck R, Neumuth T, Melzer A. Assessment of Natural User Interactions for Robot-Assisted Interventions. In: Current Directions in Biomedical Engineering. 2018.

# Ex-vivo evaluation of an ultrasound-based planning and navigation method for non-anatomical liver resections

*Iwan Paolucci[1], Raluca-Maria Sandu[1], Luca Sahli[1], Daniel Candinas[2], Stefan Weber[1], Anja Lachenmayer[2]*

*[1]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*
*[2]Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland*

Contact: iwan.paolucci@artorg.unibe.ch

***Abstract***

*Non-anatomical resections of liver tumors are very challenging as the surgeon cannot use anatomical landmarks on the liver surface or in the ultrasound image for guidance. This makes it difficult to maintain a negative resection margin to achieve an R0 resection and still preserve as much healthy liver tissue as possible, especially for inexperienced surgeons. Even though image-guided surgery systems have been introduced to overcome these challenges, they are rarely used due to their inaccuracy, time-effort and complexity in usage and setup. We have developed a novel approach, which allows to create a surgical plan intra-operatively using navigated ultrasound. In this study, we have evaluated this approach on ex-vivo porcine models. We found that using this technique, an R0 resection could be achieved in 11 out of 12 cases with an average resection margin of 4.1 mm. In conclusion, this approach might represent a useful tool for intra-operative guidance in non-anatomical resection along with conventional ultrasound guidance.*

**Keywords**: Navigation, Liver surgery, Resection, Ultrasound

## 1 Problem

Surgical resection is the current gold standard for curative care of primary and metastatic hepatic tumors. This procedure involves the removal of the part of the liver where the tumor is located. This is typically achieved by removing the segments containing the tumor, so called anatomical resections. The downside of this technique is that it also removes a large part of healthy liver tissue. Recently, non-anatomical resections are becoming more popular, as they spare more healthy liver tissue than anatomical resections with similar oncological outcomes [1]. In a non-anatomical resection, only the part of the liver with the tumor and a safety margin of 5 - 10 mm are resected. This is done to ensure a negative resection margin, which means that no tumor cells are at the boundary of the resection. These non-anatomical resections can be performed repeatedly in case of recurrence. However, compared to anatomical resections, it is more challenging to keep a negative resection margin as anatomical landmarks cannot be used for intra-operative guidance.

In the beginning of a non-anatomical resection, a resection line is drawn onto the liver surface to visualize where the resection shall be started. During the resection process, intra-operative ultrasound is used to confirm a safe distance to the tumor. Finally, once the depth is reached, the distance to the tumor is again confirmed on ultrasound and the tumor is removed. This is a challenging process which depends on the operator's ability and experience with mentally reconstructing the spatial relationships of the ultrasound image and the intra-operative scene. Additionally, the resection margin introduces artifacts and makes it harder to visualize the safety distance to the tumor on ultrasound.

To overcome these challenges, image-guidance systems have been introduced into the surgical workflow [2, 3]. These systems measure the pose of the surgical instruments and display their position on a virtual model of the anatomy. They mainly rely on a registration process to align a preoperative model with the patient's anatomy intraoperatively. This process is time-consuming, complex and error prone which is the main reason why such systems are rarely used [4].

In this study, we have evaluated a different approach, where navigated intra-operative ultrasound data is used to create a virtual model and a surgical plan on the spot. This does not require a separate registration process. With this approach a virtual draft of the surgical plan is created, which serves as a rough guidance map through the procedure. We hypothesize that using such an intra-operative surgical draft allows the surgeon to acquire a

negative resection margin. In contrast to other studies, we have evaluated this technique with respect to the resection margins, which is the primary surgical endpoint in these procedures.

# 2 Material and Methods

The proposed system is based on the CAS-One (CAScination AG, Switzerland) navigation system for liver surgery, which is equipped with the Polaris (Northern Digital, Canada) optical tracking system and the FlexFocus 800 (BK Medical, Denmark) ultrasound system. Marker shields can be attached to any type of cylindrical instrument, such as electrocautery or CUSA, and calibrated with a dedicated calibration device. For these experiments, a customized software, specifically for non-anatomical resections was developed and integrated into the navigation system. The software-workflow for creating such a surgical plan (Figure 1) consists of the following steps:

**I) Surface scanning**

A surface model of the liver is acquired by scanning the liver surface using the intra-operative US probe. A support-vector-machine(SVM)-based image classifier is used to detect whether the US probe touches the liver surface [5]. If this condition evaluates as true, the current position of the US probe is recorded and added to a point cloud. From this point cloud, a surface is reconstructed using the method described by Hoppe et al. [6].

**II) Tumor segmentation**

In order to locate the tumor, a semi-automatic tumor segmentation method based on the "Graph Cuts" algorithm [7] is used. To initialize the tumor segmentation the operator clicks in the center of the tumor and selects an approximate size. If needed, the segmentation result can then be modified by clicking on the area which has to be added or removed from the segmentation mask. This can be done by simply clicking on the part which was over or under-segmented. The software then automatically decides whether this part should be added or removed from the segmentation.

**III) Resection planning**

Once the operator confirms the tumor segmentation, the center and diameter of the tumor are calculated. Then the operator can choose the desired resection margin and shape for this specific case. The software then fits the resection shape into the model with respect to the given shape and distance constraints (Figure 1).

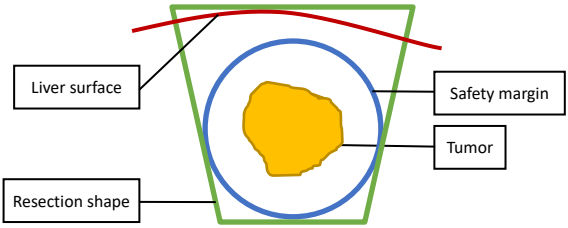



*Figure 1: Concept of planning a resection shape (green) into a local anatomical model of the tumor (yellow) and the surface (red) of the liver with a desired resection margin (blue).*

## 2.1 Experimental Set-up

In all the experiments we used ex-vivo porcine livers with tumor mimics made of agar (1% agar solution). The tumors were randomly placed at depths of 1-2 cm in the liver by injection with a hypodermic needle. These lesions were then visible as hyperechoic regions on the B-Mode US image. The mean size of the tumors was 13.6 mm (± 2.2 mm). To resect the tumors we used a monopolar electrocautery device (Medtronic, Ireland) with a scalpel electrode.

## 2.2 Experimental workflow

During these experiments, a series of 12 tumors in 2 ex-vivo porcine livers were resected. The resection planning workflow consisted of the following steps:

1. Preparing the instruments
2. Scanning the liver surface in the region of the tumor
3. Semi-automatic segmentation of the tumor on an US image plane across the center of the tumor
4. Selecting the desired resection shape and margin
5. Resecting the tumor using only the navigation screen

All resections were planned with a conical resection shape (Figure 1) and 10 mm safety margin, which is the common choice for resections of primary liver cancer.

Based on this resection model, the navigation screen was used to resect the intrahepatic tumor (Figure 2). During the resection procedure, the operator was not allowed to use the intra-operative ultrasound to avoid bias. First, a resection line was drawn on the liver surface using the navigation screen. the operator alternatively resected and double checked the path on the navigation screen. Finally, when the navigation screen indicated sufficient depth below the tumor, the operator cut the resection specimen out of the liver.

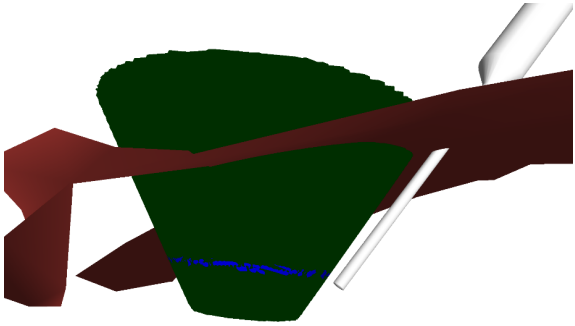


*Figure 2: Visualization during the resection process, which shows the distance and angle to the resection shape, and the depth in the tissue.*

The resection margin was measured on two orthogonal US planes acquired by the intra-operative US scanner. These two planes were acquired such that they show the closest distance from the resection border to the tumor on the side and on the bottom of the specimen. From these images, the minimum and maximum resection margin were measured.

## 3 Results

In 11 out of 12 resections, a negative safety margin and R0 was achieved by only looking at the navigation screen for guidance (Figure 3). In one case the resection margin was 1 mm, which was considered as an R1 resection, especially as the tumor was slightly visible on the resection specimen. Therefore, the R0 resection rate is at 91.7%. Overall the minimum resection margin had a median of 4.1 mm (IQR 3.4 – 5.0 mm) and the maximum margin a median of 12.1 mm (IQR 10.1 – 16.8 mm) (Figure 4).

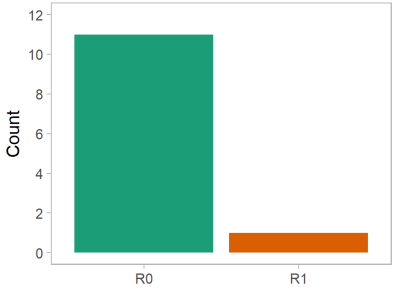



*Figure 3: R0 vs. R1 resections*

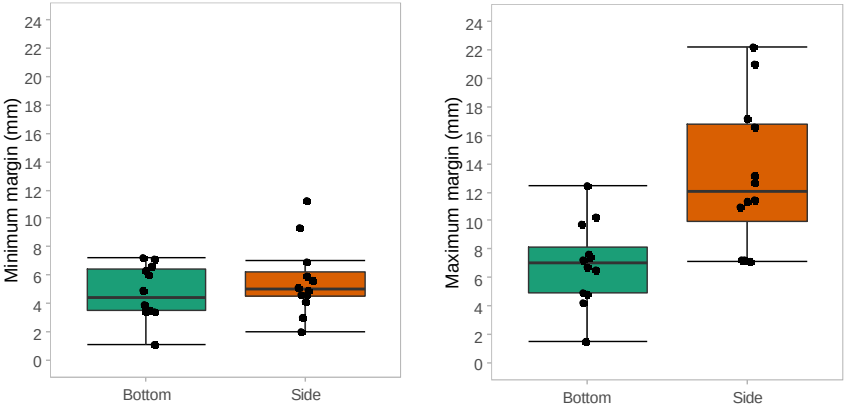



*Figure 4: The minimum (left) and maximum (right) resection margins on the bottom and the side of the resection specimen*

By looking at the location (Bottom vs. Side) of the minimum and maximum margin one can see a difference between them (Table 1). While the minimum margin is not significantly different on the bottom compared to the sides (4.5 vs. 5.0 mm respectively), the difference in the maximum resection margin is significantly smaller on the bottom compared to the side (7.0 vs. 12.1 mm respectively). An extreme example can be seen in Figure 5, where the resection margin on the right side is much larger than on the bottom and the left side.

| | Bottom | Side | p-value |
|---|---|---|---|
| Min. margin | 4.5 [3.5, 6.4] | 5.0 [4.5, 6.2] | 0.583 |
| Max. margin | 7.0 [4.9, 8.1] | 12.1 [10.0, 16.8] | 0.003 |

*Table 1: Results of the resection margins (min and max) as median and IQR.*

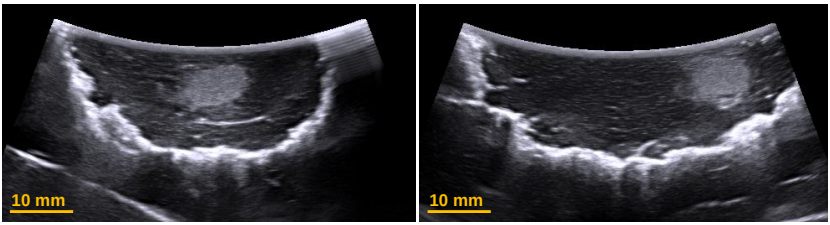



Figure 5: (left) An US image of a resected specimen with a uniform resection margin. (right) An example of an extreme case, where the resection margin is much larger on the left than on the bottom and right side

## 4 Discussion

In this study we evaluated a navigation method for non-anatomical liver resections based solely on navigated intra-operative US on an ex-vivo model. Overall, 12 tumors were resected from ex-vivo porcine liver with a median resection margin of 4.1 mm and an R0 resection rate of 91.7%. On the side of the resection, sufficient resection margin could be achieved in most cased. On the bottom of the resection the margin was rather small, especially when considering a planned resection margin of 10 mm. However, no tumor was resected with a positive margin which means that no residual tumor remained in the parenchyma.

We hypothesize, that deformation and tissue liquid evaporation account for the present inaccuracies. On the one hand, organ deformation occurs during the resection process leading to deformation towards the side where the specimen is pulled. However, this effect is expected to be smaller compared to approaches utilizing preoperatively acquired CT/MRI imaging. During the resection process, the specimen has to be pulled and moved such that the resection line is visible and can be reached with the instrument, which causes large deformations. On the other hand, electrocautery used in these experiments causes tissue liquids to evaporate, which are then not visible anymore in the images with the resection margin.

The R0 resection rate reported in literature ranges from 66.7% - 100% for these kind of resections [1]. However, this data has to be interpreted carefully, as an R0 resection is sometimes not possible due to anatomical constraints (e.g. large vasculature). While not directly comparable, the hypothetical R0 resection rate of 91.7% (ex-vivo) indicates that this tool might be useful for such resection to further increase the R0 resection rate when anatomically possible.

To improve the navigation approach further, we aim to improve the tumor segmentation and visualization during the resection process and include the vascular structures. Currently, the tumor is approximated by a single US slice which can be problematic in tumors with irregular shapes. This could be solved with a multi-slice segmentation method, which combines the results into a 3D mesh. Whereas the vascular structure is less important in these small resections, it is more important for larger resections, where large blood vessels must be clipped. In smaller resections, the blood vessels can be easily coagulated with modern resection devices. To improve the large resections margins that occurred in some cases, we aim to further improve the intraoperative visualization or integrate optical or acoustic feedback, when the surgeon deviates from the planned resection plan.

One limitation in the evaluation in this study is, that the measurement of the resection margin based on two US slices is subject to bias as it depends on the quality of the image. One way to address this limitation is to assess the resection margin in 3D based on a CT scan. Compared to other studies on navigation approaches for liver surgery, we evaluated the end to end accuracy of the whole procedure rather than technical aspects like the pure registration accuracy. Consequently, this includes factors like human error, organ deformation due to US scanning and the resection itself.

## 5 Conclusion

Using intraoperative ultrasound-based anatomical models for navigation in liver surgery might help to achieve a sufficient negative resection margin in non-anatomical resections of liver tumors. This technique might be a useful tool especially for inexperienced surgeons since it augments the conventional intra-operative ultrasound imaging.

## References

1. Moris D, Tsilimigras DI, Kostakis ID, Ntanasis-Stathopoulos I, Shah KN, Felekouras E, Pawlik TM (2018) Anatomic versus non-anatomic resection for hepatocellular carcinoma: A systematic review and meta-analysis. Eur. J. Surg. Oncol. 44:927–938
2. Banz VM, Müller PC, Tinguely P, Inderbitzin D, Ribes D, Peterhans M, Candinas D, Weber S (2016) Intraoperative image-guided navigation system: development and applicability in 65 patients undergoing liver surgery. Langenbeck's Arch Surg 401:495–502 . doi: 10.1007/s00423-016-1417-0
3. Langø T, Vijayan S, Rethy A, Våpenstad C, Solberg OV, Mårvik R, Johnsen G, Hernes TN (2012) Navigated laparoscopic ultrasound in abdominal soft tissue surgery: Technological overview and perspectives. Int. J. Comput. Assist. Radiol. Surg. 7:585–599
4. Kingham TP, Jayaraman S, Clements LW, Scherer MA, Stefansic JD, Jarnagin WR (2013) Evolution of Image-Guided Liver Surgery: Transition from Open to Laparoscopic Procedures. J Gastrointest Surg 17:1274–1282 . doi: 10.1007/s11605-013-2214-5
5. Sahli L, Weber S, Paolucci I Liver surface reconstruction from navigated ultrasound during image-guided liver surgery Acquisition Surface Contact Detection Reconstruction
6. Hoppe H (1995) Surface Reconstruction from Unorganized Points. University of Washington
7. Rother C, Kolmogorov V, Blake A (2004) "GrabCut": interactive foreground extraction using iterated graph cuts. ACM SIGGRAPH 2004 Pap - SIGGRAPH '04 23:309 . doi: 10.1145/1186562.1015720

# Tissue Impedance Spectroscopy to Guide Resection of Brain Tumours: Preliminary experience ex vivo

*Mareike Apelt[1], Gesar Ugen[2], Levin Häni[2], Ekkehard Hewer[3], Andreas Raabe[2], Kathleen Seidel[2], Juan Ansó[1]*

*[1]ARTORG Center for Biomedical Engineering, Universität Bern, Bern, Switzerland*
*[2]Department of Neurosurgery, University Hospital Bern, Bern, Switzerland*
*[3]Department of Pathology, University Hospital Bern, Bern, Switzerland*

contact: mareike.apelt@artorg.unibe.ch

***Abstract***

Visual differentiation of lower grade glioma tissue from normal brain tissue during surgery is difficult even for expert neurosurgeons. Therefore, during tumour removal neurosurgeons rely only on image guidance. It has been proven that higher rates of tumour resection prolong long-term survival of patients. We aim to implement impedance spectroscopy as a potential supportive tool to improve the extent of resection. A simple setup was evaluated and tested on pig brain. White matter showed a higher impedance magnitude than grey matter, higher phase values in the low frequencies and lower phase values in the high frequencies than grey matter. During an initial testing on human biopsy samples, the feasibility to differentiate ex vivo tissue samples (from tumour resections) using impedance spectroscopy was evaluated. Tissues were collected from two patients and impedance spectra differences were found between low-grade glioma, high-grade glioma and healthy brain tissue.

**Keywords**: Impedance Spectroscopy, Glioma, Biopsy, Ex vivo

## 1 Problem

Gliomas are the most common primary brain tumours, with an incidence of 5/100'000 per year (2008-2012, United States of America) [1]. The symptoms depend on the location of the tumour and frequently compass neurological deficits, headaches, epileptic seizures, changes in personality as well as nausea and vomiting due to increased intracranial pressure. Gliomas are tumours which develop from aberrant glial cells, the support cells in the central nervous system, the most common being astrocytoma. The tumour growth is infiltrative and often follows with matter fiber tracts, making them more difficult to resect. The World Health Organization (WHO) grading system describes the malignancy or severity level: WHO grade 1 are usually benign lesions that can be cured by resection. WHO grade 2 (low-grade glioma) notoriously recur due to their infiltrative growth and patients have a median survival of 7-10 years. WHO grade 3 and WHO grade 4 tumours are considered malignant high grade gliomas, with the most frequent type, glioblastoma multiforme, displaying a median survival of 14-18 months. Depending on grade and extent of resection, a chemo- and radiotherapy is considered after maximal surgical resection.

Higher extent of resection has been associated with a longer survival rate [2][3][4][5]. Even small residual tumour volumes decrease overall survival [2]. But at the same time no eloquent area of the brain (motor, speech) should be damaged during surgery, because postoperative deficits are associated with significant disability and lower long-term survival [6]. The surgeons should remove the whole tumour while not damaging any healthy brain tissue. However, there is no visual difference between healthy brain tissue and low-grade glioma tissue. High-grade glioma can be resected with the help of 5-aminolevulinic acid (5-ALA), which enriches in tumour cells and is metabolized into protoporphyrin IX, a fluorescent molecule, which helps resection [5]. Additionally, during a surgery, surgeons can carry out on or more intraoperative MRIs to improve the accuracy of the image guidance procedure.

Impedance spectroscopy can detect differences in tissue architecture (cell density, water content, extracellular matrix content) [7]. Healthy brain and tumours have different tissue architecture. Studies in rats have shown, that brain tumour and healthy brain have a different impedance spectrum [8], when measured with a microelectrode array. Additionally, while collecting data for an improved EEG source model, a study [9] showed, there are different resistivity at 50 kHz in high-grade glioma, low-grade gliomas and healthy cortex.

We aim to integrate impedance spectroscopy into a surgical device to give surgeons real-time feedback to differentiate between tumour and healthy brain. As a first step towards this goal an initial explorative study was

started on human biopsy samples and the simple setup was evaluated on pig brain. A better understanding of the passive electrical properties of brain and tumour tissue will help us to develop an appropriate model that can be translated into a clinical approach to help surgeons striving for more radical resections of gliomas.

# 2 Materials and Methods

## 2.1 Measurement device and verification

A commercial bipolar stimulation probe used for intraoperative neurophysiological monitoring (52-015, Inomed) was connected to an interface box providing a two-electrode measurement set-up. The interface box was in turn connected to an electrical impedance spectroscopy device (ISX-3v2, Sciospec). Data acquisition was performed with a custom-made software interface. A frequency sweep from 100 Hz to 1 MHz (logarithmically spaced, 100 points) and 50 mV voltage amplitude was applied to measure impedance. Precision 1 is the default setting of the impedance measurement system and it corresponds to 160 measured periods per frequency. For the defined measuring bandwidth and resolution, measuring time was 14.524 seconds. The probe was attached to a holding station, which allowed slow and controlled movement of the probe onto a tissue sample.

Prior to start evaluating impedance spectroscopy on tissue sample, the electrical property of the measuring probe was evaluated. The sensitivity field (the volume around the electrode which is measured) was experimentally determined following a similar to Halter et al. [10],[7] see **Figure 1**.

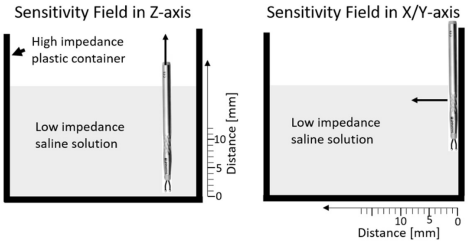



***Figure 1 Experimental Setup*** *to determine the sensitivity field of the bipolar electrode*

A plastic container (high impedance) was filled with 1X phosphate buffered saline solution (PBS) (low impedance). An impedance spectrum was then measured starting at position 0 mm, which corresponds to touching the wall, and retracting in 1 mm steps up to 10 mm, then 15 mm and then 20 mm distance to the wall. This test was performed in x-axis, y-axis and z-axis. The data was analyzed in Matlab (Mathworks). PBS was considered as a baseline, the distance for which the measurement reached baseline corresponds to the largest extent of the sensitivity field in that axis. Signal to noise ratio was evaluated by measuring 50 times in 1X PBS and calculating the percentage of standard deviation versus the mean of the magnitude and phase.

## 2.2 Biopsy

This is a pilot study within an ex vivo (biopsy) trial approved by the ethical commission of the canton of Bern (ID 2018-01158). Tissue samples were extracted during tumour resections in two patients. Patient no 1 was diagnosed with a low-grade glioma (N=1 sample) and patient no 2 with high grade glioma (N=5 samples). Unaltered brain tissue (N=1) was extracted from patient no 1 (overlaying grey matter close to infiltrated white matter).

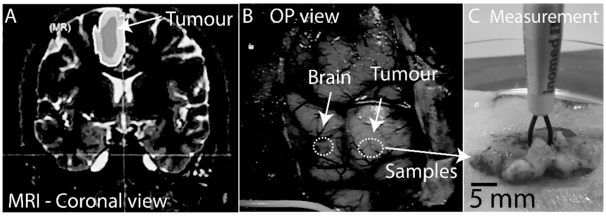



***Figure 2 Surgical views and measurement*** *A: Coronal view of preoperative MRI , darker overlay is the tumour, lighter overlay is the supramarginal resection area, B: Microscope view tumour (right) and unaltered brain tissue (left) are exposed, with no visual difference, C: Measuring probe and biopsy sample during measurement.*

### 2.3 Reproducibility

To evaluate the reproducibility of the measurements, a pig brain (obtained from the butcher the day of slaughter) was measured. To determine which excitation amplitude to use, an amplitude sweep was performed starting at 10 mV going up to 180 mV. An excitation amplitude of 100 mv resulted in sufficient signal to noise ratio and was chosen for subsequent measurements. The brain was cut into coronal slices and grey matter and white matter (**Figure 3**) chosen visually. To determine if there are differences between grey matter and white matter and if those differences are the same across the whole brain, 33 measurement points for grey matter and 33 measurement points for white matter where measured with the measurement setup and an excitation amplitude of 100 mV. The coronal slices of pig brain were approximately 6 cm by 3 cm and 1 cm height (depending on the position in the brain).

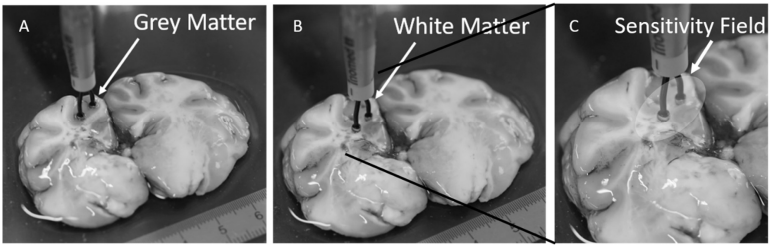



***Figure 3: Measurement setup****: Representative pictures of electrode placement on grey matter (A) and white matter (B). Zoom into the white matter and an approximation of the sensitivity field (C)*

## 3 Results

### 3.1 Verification

**Sensitivity field**

The measured sensitivity field around the electrode was approximately 4 mm by 4 mm by 3 mm (**Figure 4**,A,B).

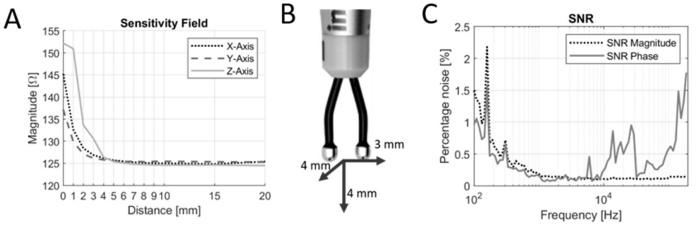



***Figure 4 Verification*** *A: Sensitivity Field – x axis, y axis and z axis, 97 kHz as a representative frequency, B: Sensitivity Field representation, C: Signal to noise ratio percentage of standard deviation to mean signal.*

**SNR**

In PBS, a small inductive effect likely associated to the measuring setup was observed in the high frequencies, therefore the SNR was not represented for the highest frequencies (>190 kHz) (**Figure 4**,C). In the magnitude of the impedance spectrum the noise was higher in the low frequencies (<300 Hz) (max value 2.2 % at 159 Hz), whereas in the middle and high frequencies (>1 kHz), the % of noise was 0.1 to 0.2 % of the signal. In the phase, a mild inductive behavior (phase > 0) was observed in frequencies above 360 kHz, for the frequencies below 360 kHz, the % noise was in the range of 0.08 to 2 % of the signal.

## 3.2 Biopsy

High-grade glioma presented lower magnitude than unaltered brain tissue and low-grade glioma (**Figure 5**, left). High-grade glioma sample showed a wide range of values, due to the heterogeneity of the collected tissue samples. In terms of phase (**Figure 5**, right), high grade and healthy cortex presented differences in high (>200 kHz) and low frequency (< 1kHz) ranges of the spectrum, with crossing point between 10 and 20 kHz. Phase was almost identical between low-grade glioma and healthy brain tissue. Repeated measurements at the same spot on the tissue resulted in mild variations (<2% for most frequencies, 2-5.5% for frequencies <200Hz)

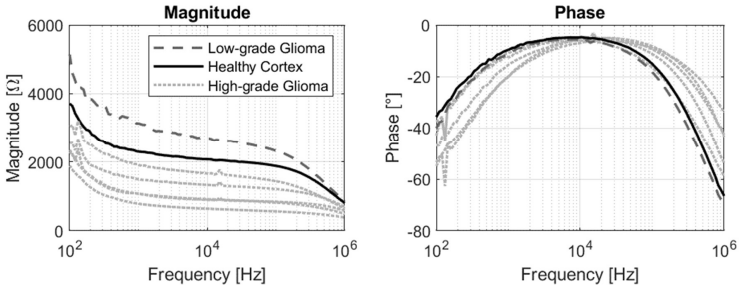



***Figure 5 Bode plot of impedance spectra:*** *High-grade glioma, low-grade glioma and healthy cortex*

## 3.3 Reproducibility

Throughout the whole pig brain, white matter presented higher magnitude than grey matter in all frequencies of the spectrum (**Figure 6**). In the phase, white matter showed a higher phase angle in the lower frequencies and a lower phase angle in the high frequencies with a crossing point approximately from 8 to 15 kHz.

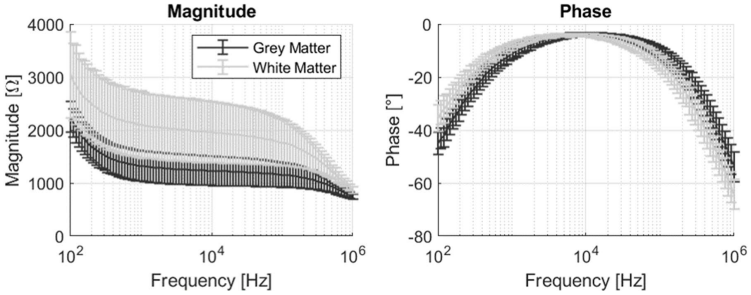



***Figure 6 Bode plot of impedance spectra**: Grey matter versus white matter, magnitude (left), phase (right)*

## 4 Discussion

These preliminary results showed that it is possible to measure tissue impedances spectroscopy with a simple setup. The experimentally determined sensitivity field is small enough to allow us to measure the sample sizes available from tumor resections. The setup measures reliably in PBS with low percentage of noise.

The measurement on the pig brain showed, that there is a measurable difference between grey matter and white matter. We hypothesize that the overlapping standard deviations are caused by the relatively large sensitivity field of the probe compared to the pig brain regions under measurement (see **Figure 3**C). White matter has a higher impedance magnitude than grey matter, because it contains myelin, an insulator which increases the impedance, which is in agreement with [9]. A consistent maximum was found in the phase (intersecting point) which could in the future be used for the development of an algorithm to automatically classify the measured tissues. It also allows for parametrization instead of relying on the absolute values. Observed variations in the measurements, might have been caused by different pressures applied [11] and the dynamically varying hydration levels of the tissue, which influences impedance levels as suggested in [12]. Another variable, which was not controlled, was the area of the electrode contacting the tissue, which may cause some variations.

The initial findings from tumor biopsy samples in humans are promising. It is possible to see a difference between high-grade glioma and healthy brain tissue. Low-grade glioma showed similar measurements as healthy brain tissue. It could be that the healthy tissue sample was partially affected by the neighbouring tumour. In further steps of the study, control tissue samples (healthy) will be acquired from epilepsy patients to exclude infiltration with tumour tissue. We plan to include 10 patients suffering from high-grade glioma, low-grade glioma, and epilepsy each. Biopsy samples will be taken at defined points of the resection, the impedance spectroscopy measured and then sent for histopathological examination. Each biopsy sample measured will be additionally sliced and histologically assessed. This will serve as a ground truth, what exactly we are measuring, what type of tumour and what kind of tissue structure (e.g. cell nuclei) and it will allow us to measure not only tumour but also the infiltration zone and classify the impedance measurements according to the actual tissue type measured. With all the data collected we will work on an algorithm to determine, which frequencies demonstrate the most differentiating character and we would like to use this to then implement a measurement only relying on the most relevant frequencies, thereby making it faster and allowing it to be later used during tumour resection surgery.

## 5 Conclusion

To conclude, this initial result shows the feasibility of the study setup and a potential to differentiate tumour and healthy brain tissue using impedance spectroscopy. The preliminary results are promising, there is a difference in the impedance spectrum between healthy brain, high-grade glioma and low-grade glioma tissue ex vivo.

## Acknowledgement

The authors would like to acknowledge the support of the funding agencies, EANS, Bridge and Innosuisse.

## References

[1] Q. T. Ostrom *et al.*, "CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2008-2012," *Neuro. Oncol.*, vol. 17, no. suppl 4, pp. iv1–iv62, 2015.

[2] J. S. Smith *et al.*, "Role of extent of resection in the long-term outcome of low-grade hemispheric gliomas.," *J. Clin. Oncol.*, vol. 26, no. 8, pp. 1338–1345, 2008.

[3] M. Lacroix *et al.*, "A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival," *J. Neurosurg.*, vol. 95, no. 2, pp. 190–198, 2001.

[4] N. Sanai, M.-Y. Polley, M. W. McDermott, A. T. Parsa, and M. S. Berger, "An extent of resection threshold for newly diagnosed glioblastomas," *J. Neurosurg.*, vol. 115, no. 1, pp. 3–8, Jul. 2011.

[5] W. Stummer, U. Pichlmeier, T. Meinel, O. D. Wiestler, F. Zanella, and H. J. Reulen, "Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial," *Lancet Oncol.*, vol. 7, no. 5, pp. 392–401, 2006.

[6] M. J. McGirt, D. Mukherjee, K. L. Chaichana, K. D. Than, J. D. Weingart, and A. Quinones-Hinojosa, "ASSOCIATION OF SURGICALLY ACQUIRED MOTOR AND LANGUAGE DEFICITS ON OVERALL SURVIVAL AFTER RESECTION OF GLIOBLASTOMA MULTIFORME," *Neurosurgery*, vol. 65, no. 3, pp. 463–470, Sep. 2009.

[7] R. J. Halter, A. Hartov, J. A. Heaney, K. D. Paulsen, and A. R. Schned, "Electrical Impedance Spectroscopy of the Human Prostate," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 7, pp. 1321–1327, Jul. 2007.

[8] H.-G. Jahnke *et al.*, "Impedance spectroscopy—An outstanding method for label-free and real-time discrimination between brain and tumor tissue in vivo," *Biosens. Bioelectron.*, vol. 46, pp. 8–14, 2013.

[9] J. Latikka and H. Eskola, "The Resistivity of Human Brain Tumours In Vivo," *Ann. Biomed. Eng.*, vol. 47, no. 3, pp. 706–713, 2019.

[10] R. D. Butler and R. J. Halter, "Gauging electrical properties of bone with a bioimpedance-sensing drill," *Physiol. Meas.*, vol. 40, no. 1, 2019.

[11] R. H. Smallwood, K. D. Bardhan, D. C. Walker, C. A. González-Correa, and B. H. Brown, "Electrical bioimpedance readings increase with higher pressure applied to the measuring probe," *Physiol. Meas.*, vol. 26, no. 2, pp. S39–S47, 2005.

[12] T. W. Balmer, S. Vesztergom, P. Broekmann, A. Stahel, and P. Büchler, "Characterization of the electrical conductivity of bone and its correlation to osseous structure," *Sci. Rep.*, vol. 8, no. 1, pp. 1–8, 2018.

# Analyzing Natural Language Commands during Laparoscopic Surgery for Robotic Camera Guidance

*P. M. Scheikl*[1], *B. Nigl*[1], *C. Kunz*[1], *F. Mathis-Ullrich*[1]

[1] *Karlsruhe Institute of Technology, Institute for Anthropomatics and Robotics, Karlsruhe, Germany*

Contact: paul.scheikl@kit.edu

***Abstract***

*In this work, we investigate the use of natural language commands for adjusting camera positions during laparoscopic surgery in order to equip a robotic assistance systems with an intuitive voice interface. The integration of natural language commands enables the surgeon to flexibly adapt the camera's position and to provide feedback to a learning assistance system. Based on the transcripts from 20 laparoscopic surgeries the verbal commands were clustered to receive a grounded understanding of the different command types used. These insights then serve as input for supervised classification methods that map natural language commands to actions of the robotic assistance system that guides the camera. The preliminary results show the need for more complex voice command structures than those presently integrated in existing assistance systems.*

**Keywords**: Laparoscopic Surgery, Robot Assisted Surgery, Natural Language Processing

## 1 Problem

In laparoscopic surgery, an optimal and stable view on the operating field as well as smooth and precise camera movements throughout the operation are essential for high quality performance, as the surgeon's field of view is limited to the camera image and is not directly controlled by the surgeon. Especially during long operations, a human-guided camera may lead to fatigue of the assistant, shaky camera images, and errors in camera placement due to decreasing concentration.

The first robotic camera guidance system for holding and manoeuvring rigid endoscopes in minimal invasive surgery, called AESOP, was developed in 1994 by Computer Motion Inc. [1]. The system can be controlled by a voice interface as well as a joystick and foot pedal. Within the voice interface, the speech recognition is activated with the hotword "aesop". The recognised commands are then compared to a predefined grammar and in case of a match a corresponding movement is executed. A similar system was designed in 2005 by Berkelman et al. called VIKY Endoscope Positioner [2]. Both systems allow movements in the four principal directions (up, down, left, and right) as well as zooming in an out while distinguishing between step (fixed distance) and move (continuous) commands. The VIKY-System additionally incorporates a fist implementation of instrument tracking, whereas AESOP only offers speech commands. Another robotic assistance system also proving a voice interface is the Soloassist II [3]. Unlike the other systems the Soloassist's voice control defines five distance modifiers and allows diagonal movements. Furthermore, Soloassist as well as the AESOP-System allow the surgeon to adjust speed settings and to save several keypoints, allowing an easy return to key positions during the surgery. With the KaLAR, developed in 2004, Kim et al. pursued a new approach for the use and control of laparoscopic assistance robots [4]. They developed a surgical model on the example of a cholecystectomy, containing a series of surgical steps (e.g. cutting, suturing and knotting), the surgical instruments used, and a reference camera viewpoint at each stage of the surgery. Depending on the task performed either an instrument tracking mode or a site-keeping mode is activated.

However, most of these approaches require the surgeon to explicitly control the camera at any time during the surgery, offering only low-level command structures to do so and are unable to adapt to the situation or individual needs of the surgeon as a skilled human assistant would.

In this work we therefore analyse the use of natural language as a combined feedback and control interface for a learning robotic camera guidance system. By using natural language within the voice interface, the surgeon is able to use complex control commands similar to communicating with a human assistant.

## 2 Material and Methods

Camera related commands are manually transcribed during 20 laparoscopic surgeries and serve as initial data set for natural language processing. Table 1 shows an overview of the different transcribed surgeries. Based on this data set different natural language processing techniques including topic modelling with Latent Dirichlet Allocation, word embedding, and random forest classification are performed. With the results of different

unsupervised methods it is investigated whether the commands can be clustered in reasonable groups of different command types which then serve as labels for supervised learning methods such as support vector machines and random forest to map natural language commands to actions of the robotic assistance system.

Table 1: *Number of laparoscopic surgeries and the corresponding average number of commands (increases with length and complexity of the procedure) in the data set. Transabdominal preperitoneal is abbreviated as TAPP.*

| | Cholecystectomy | Appendectomy | Inguinal Hernia TAPP | Gastric Bypass | Gastric Sleeve |
|---|---|---|---|---|---|
| No. Surgeries | 8 | 1 | 2 | 1 | 2 |
| No. Commands | 8.5 | 9 | 30 | 27 | 39.5 |

## 3 Results

Transcripts show that existing control structures do not represent the commands actually used, making the clinical need for a more natural voice interface evident. Thus the need for a more natural voice interface seems to exist. Figure 2 shows the relative occurrences of words in the current data set (Table 1), visualised with a word cloud. Figure 1 shows the results of the Latent Dirichlet Allocation with six topic revealing initial insights into the different command types.

| Command | # Occurrences |
|---|---|
| von | 94 |
| gucken | 89 |
| zurück | 75 |
| nach | 50 |
| oben | 48 |
| ran | 41 |

Table 2: *Number of occurences for the most common words based on 13 transcripts (frequent stop words removed).*

```
LDA Model:
Topic #0: 0.068*"Seite" + 0.057*"von_links" + 0.046*"anderen_Seite"
Topic #1: 0.073*"zeigen" + 0.055*"oben" + 0.038*"Bauchdecke"
Topic #2: 0.159*"zurück" + 0.049*"unten" + 0.049*"hol"
Topic #3: 0.091*"dichter_ran" + 0.059*"von_links" + 0.040*"dichter"
Topic #4: 0.082*"von_oben" + 0.050*"von" + 0.034*"kann"
Topic #5: 0.148*"Foto" + 0.119*"gut" + 0.060*"machen"
Topic #6: 0.283*"gucken" + 0.072*"ran" + 0.046*"von_links"
Topic #7: 0.097*"putzen" + 0.050*"sehen" + 0.050*"zeigen"
```

Figure 1: *The Latent Dirichlet Allocation with six Topics (including bi-grams with more than two occurrences.)*

## 4 Discussion and Conclusion

The preliminary results show the associated need for more complex command structures than those presently integrated in existing assistance systems. With complex natural language commands the surgeon is able to flexibly adjust camera positions in an intuitive and natural way, obtaining a stable view on the operating field with smooth and precise camera movements throughout the operation. Furthermore, those complex command structures can be used as verbal feedback for a learning robotic assistance system, enabling it to individually adapt to the surgeons needs.

The next steps include gathering additional data, cluster the observed commands, and define a flexible method to associate newly observed commands with actions of the robotic assistance system.

## 5 References

[1] Yulun Wang, Darrin Uecker, *Speech interface for an automated endoscope system*, US Patent 7,395,249, (2008)

[2] Peter Berkelman, Philippe Cinquin, *Development and testing of a compact endoscope manipulator for minimally invasive surgery*, Computer Aided Surgery, 10/1–13, Taylor & Francis, (2005)

[3] Björn Schuller, Christoph Scheuermann, *Robust Speech Recognition for Human-Robot Interaction in Minimal Invasive Surgery*, Proc. 4th Russian-Bavarian Conference on Bio-Medical Engineering, Moscow, Russia, (2008)

[4] Jonathan Kim, Yun-Ju Lee, *Compact camera assistant robot for minimally invasive surgery: KaLAR*, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 3/2587–2592 (2004)

# Bild- & Signalverarbeitung

# Active Learning mittels Bayesian Deep Networks für die Segmentierung von laparoskopischen Instrumenten

*S. Bodenstedt[1], A. Jenke[1], M. Wagner[2], F. Bechtolsheim[3], B. Müller-Stich[2], M. Distler[3], J. Weitz[3], S. Speidel[1]*

[1] *Translationale Chirurgische Onkologie, Nationales Centrum für Tumorerkrankungen (NCT), Dresden*
[2] *Allgemein-, Viszeral- und Transplantationschirurgie, Universitätsklinikum Heidelberg, Heidelberg*
[3] *Viszeral-, Thorax- und Gefäßchirurgie, Universitätsklinikum Dresden, Dresden*

Kontakt: sebastian.bodenstedt@nct-dresden.de

***Abstract***
*Annotierte Bilddaten sind kritisch für Anwendungen der computergestützten Chirurgie, jedoch ist das Erstellen von Annotationen eine aufwendige Aufbau. Active Learning ermöglicht ein gezieltes Annotation von Bilddaten, in anderen Worten, nur Bilder die neue Informationen beitragen werden annotiert. In diesem Paper stellen wir eine Methode für Active Learning mit Fully Convolutional Neural Networks vor.*

**Keywords**: Segmentierung, Active Learning, Bayesian Deep Learning

## 1 Problemstellung

Ziel der computergestützten Laparoskopie ist es dem Chirurgen zum richtigen Zeitpunkt die richtige Art der Unterstützung, bspw. eine Navigationsunterstützung oder eine robotergestützten Kameraführung, zu bieten. Maßgeblich für die Durchführung einer solche Unterstützung ist ein Verständnis der chirurgischen Szene, bspw. durch das Bestimmen der Positionen von Zielstrukturen, wie Organe oder Tumore, und der Instrumente.

Als Basis hierfür dienen meist Methoden für die semantische Segmentierung, welche jedem Pixel im Bild einer eindeutigen Klasse zuordnen. Der aktuelle Stand der Forschung im Bereich der laparoskopischen Szenensegmentierung sind Fully Convolutional Neural Networks (FCNNs), ähnlich der U-Net Architektur [1]. FCNNs benötigen jedoch eine erhebliche Anzahl an annotierten Trainingsbeispielen um qualitativ hochwertige Ergebnisse zu erzielen. Für diese Beispiele werden meist aufwendige, pixelweise Annotationen benötigt.Gerade im medizinischen Bereich ist man für die Erstellung solcher Annotationen oft auf Experten angewiesen, was das Aufbereiten von größeren Datensätzen erheblich erschwert.

Ein System, das stattdessen aktiv die Expertenmeinung nur für bestimmte Beispiele, z.B. solche mit einer hohen Unsicherheit, anfordern würde, könnte den gesamten Aufwand verringern und das Erstellen großer, annotierter Datensätze vereinfachen. Ein solches System wird Active Learning-System [2] genannt. Während des aktiven Lernens wird zu Beginn ein Anfangsmodell unter Verwendung einer kleinen Menge von annotierten Daten trainiert. Mittels einer Unsicherheitsmetrik wird dann bestimmt welche der Datenpunkte als nächstes annotiert werden sollen. Ein neues Modell wird auf den erweiterten Trainingsdaten [3] trainiert und der Kreislauf beginnt von vorne. Durch das gezielte Annotieren von unsicheren Daten soll so eine hohe Genauigkeit mit einem geringeren Annotationsaufwand erreicht werden.

In [4] haben wir ein Active Learning-System, basierend auf Bayesian Deep Learning, auf die laparoskopischen Phasensegmentierung und Instrumentenpräsenserkennung angewendet. Hier erweitern wir diesen Ansatz auf die semantische Segmentierung und zeigen erste Ergebnisse am Beispiel der Instrumentensegmentierung.

## 2 Material und Methoden

Als Grundlage für unser Active Learning-System dient die TernausNet-Architektur [5] des Sieger der Endoscopic Vision 2017 Sub-challenge für "Robotic Instrument Segmentation"[6]. Wie in [4] beschrieben erweitern wir dieses FCNN mittels Monte-Carlo Dropout in ein Bayessches neuronales Netzwerk, welches imstande ist Unsicherheiten in den Vorhersagen zu modellieren. Da für jeden Pixel im Eingabebild eine Klassifikation durchgeführt wird, können wir pro Pixel die Unsicherheit bestimmen. Hierzu nutzen und vergleichen wir die in [4] beschriebenen Metriken Varianz, Entropie, Mutual Information (MI) und Variationsratio. Um einen Unsicherheitswert pro Bild zu erhalten, mitteln wir über die jeweils berechnete Metrik.

| Anz. | Binär | | | | | Vollständig | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Bilder | Zufall | Var. Ratio | Entropie | MI | Varianz | Zufall | Var. Ratio | Entropie | MI | Varianz |
| 72 | 86% | 91% | 89% | 91% | 90% | 40% | 42% | 40% | 36% | 37% |
| 108 | 87% | 89% | 88% | 90% | 90% | 38% | 45% | 46% | 46% | 44% |
| 144 | 87% | 91% | 90% | 94% | 92% | 40% | 46% | 44% | 40% | 46% |
| 180 | 87% | 93% | 91% | 94% | 92% | 39% | 46% | 47% | 43% | 46% |
| 216 | 92% | 90% | 91% | 93% | 92% | 45% | 49% | 45% | 44% | 48% |
| 252 | 92% | 93% | 94% | 93% | 92% | 43% | 53% | 48% | 45% | 51% |

Tabelle 1: Durchschnittliche F1-Scores auf den Testdaten mit steigender Anzahl an annotierten Trainingsbeispielen für die binäre und die vollständige Segmentierung der laparoskopischen Instrumente.

## 3 Erste Ergebnisse

Dieses Active Learning-System wenden wir auf den Datensatz der Endoscopic Vision 2017 Sub-challenge für "Robotic Instrument Segmentation"[6] an. Dieser Datensatz besteht aus insgesamt 8 Bildsequenzen von robotergestützten Eingriffen am Schweinemodell mit jeweils 225 Bildern. In jedem Bild wurden die Pixel in 5 Klassen unterteilt: Hintergrund, Schaft, Gelenk, Spitze und Ultraschallsonde. Hier betrachten wir zwei Probleme: die binäre Instrumentensegmentierung, bei der die 4 nicht-Hintergrund Klassen zusammengefasst werden und die vollständige Instrumentensegmentierung, bei der alle 5 Klassen betrachtet werden.

Zum Trainieren des FCNN nutzen wir die ersten 6 Bildsequenzen, die übrigen 2 werden zum Testen verwendet. Um eine aktive Annotation zu simulieren, gehen wir davon aus, dass zu Beginn 6 Bilder pro Trainingsbildsequenz äquidistant annotiert sind. Daraufhin trainieren wir unser Netzwerk für 50 Epochen und bestimmen dann aus den übrigen Bildern die 36 Beispiele mit der höchsten Unsicherheit mittels einer der Metriken aus [4]. Beim nächsten Durchlauf nehmen wir diese Beispiele hinzu. Pro Durchlauf bestimmen wir den F1-Score auf den Testdaten und vergleichen die verschiedenen Metriken mit einer zufälligen Datenauswahl (siehe Tabelle 1).

## 4 Diskussion

Im Fall der binären Segmentierung lässt sich kaum ein Unterschied zwischen den Leistungen der aktiven Selektionsmethoden und der zufälligen Datenauswahl erkennen. Alle Durchläufe erreichen ähnlich schnell eine vergleichbare Klassifikationsleistung. Dies führen wir darauf zurück, dass die binäre Abgrenzung der Instrumente vom Hintergrund ein relativ einfaches Problem auf diesem Datensatz ist, da gerade die Hintergründe sich in den verschiedenen Sequenzen sehr ähnlich sind. Bei dem komplexeren Problem der vollständigen Segmentierung sieht man jedoch einen Unterschied zwischen den Methoden. Besonders die Active Learning-Ansätze basierend auf dem Variationsratio und auf der Varianz heben sich von der zufälligen Auswahl ab.

Die erste Evaluation der vorgestellten Methoden liefert schon sehr vielversprechende Ergebnisse. In weiteren Schritten soll der Einsatz auf komplexeren Datensätzen evaluiert werden. Weiterhin sollen Unsicherheitsmetriken die das gesamte Bild betrachten, und nicht nur einzelne Pixel, untersucht werden. Mittels einer prospektiven Evaluation soll die Eignung des Active Learnings als Teil des Annotationsprozesses untersucht werden.

## Literatur

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on MICCAI, pp. 234–241, Springer, 2015.

[2] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, "Active learning with statistical models," JAIR, vol. 4, pp. 129–145, 1996.

[3] Y. Gal, R. Islam, and Z. Ghahramani, "Deep bayesian active learning with image data," in ICML, 2017.

[4] S. Bodenstedt, D. Rivoir, A. Jenke, M. Wagner, M. Breucha, B. Müller-Stich, S. T. Mees, J. Weitz, and S. Speidel, "Active learning using deep bayesian networks for surgical workflow analysis," International Journal of Computer Assisted Radiology and Surgery, vol. 14, pp. 1079–1087, Jun 2019.

[5] A. A. Shvets, A. Rakhlin, A. A. Kalinin, and V. I. Iglovikov, "Automatic instrument segmentation in robot-assisted surgery using deep learning,"

[6] M. Allan, A. Shvets, T. Kurmann, Z. Zhang, R. Duggal, Y.-H. Su, N. Rieke, I. Laina, N. Kalavakonda, S. Bodenstedt, et al., "2017 robotic instrument segmentation challenge," arXiv preprint arXiv:1902.06426, 2019.

# An AI-assisted Labeling Pipeline for Object Detection based on the Bonseyes Platform

*D. Ostler[1], J. Morandell[1], J. Fuchtmann[1], H. Feussner[1,2], D. Wilhelm[1,2]*

[1] *Research Group for Minimally Invasive Interdisciplinary Therapeutical Intervention at the University Hospital rechts der Isar of the TU Munich, Trogerstr. 26, 81675 München, Germany*

[2] *Department for Visceral Surgery at the University Hospital rechts der Isar of the TU Munich, Ismaninger Str. 22, 81675 München, Germany*

Contact: daniel.ostler@tum.de

***Abstract***

*Supervised machine learning for computer vision requires considerable amounts of labeled data. The labeling process is most commonly performed manually by experts in the respective field. To minimize the time requirement and to prevent a bottleneck-effect an AI-assisted labeling pipeline for object detection was developed. The docker-based application comprises an AI-application, an AI-training-pipeline, both based on the Bonseyes platform, and a modification of a pre-existing labeling-tool. The pipeline was tested on surgical instrument detection. The results indicate time reduction when compared to the manual labeling process.*

**Keywords**: Image Labeling, Deep Learning, Computer Vision, Medical Instrument Detection

## 1 Introduction

Deep Learning, a sub-field of machine learning, has proven its superiority over conventional techniques especially for common computer vision problems such as object detection. However, supervised learning requires huge amounts of labeled data, which often causes a bottleneck, since the annotation process is tedious and involves considerable human resources. With this work, we present our solution to streamline the data labeling process for deep learning applications based on the Bonseyes platform [1]. Our data annotation pipeline includes an AI-App, which pre-annotates images and requires human input only for corrections. We demonstrate its usage for the use case of surgical instrument detection.

## 2 Material and Methods

### 2.1 Labeling Pipeline

Our data annotation pipeline consists of multiple dockerized[1] tools within the following structure as shown in Fig.1: The first tool extracts single frames in a defined frequency from a selected video (A). This collection of video frames (image stack (B)) feed the input of a pre-trained AI-App. This App was created using the Bonseyes AI-pipeline (G) [2], which represents a significant part within the assisted labeling pipeline. It facilitates the data pre-processing, training of a DL Network, benchmarking, optimization for deployment and finally the creation of a dependency-free AI-App for inference (C). The resulting AI-App processes the images and predicts object type and corresponding location within each frame. In- and output is facilitated via an HTTP-interface. The resulting predictions (D), together with the images are stored locally and the storage path is forwarded to the labeling tool (E), which is a customized fork of *LabelImg*[2]. This program, through interaction with a user interface, allows the human expert to check the correctness of the proposed annotations, and to adjust them if necessary. The newly generated image stack and annotations are fed back into the Bonseyes Pipeline to further improve the detection accuracy of the AI-App. Optionally, a data augmentation step (F) can be performed, with a customizable set of modifications (e.g. flip, crop, rotate, noise, etc.), to enhance the data set.

[1] https://www.docker.com/
[2] https://github.com/tzutalin/labelImg

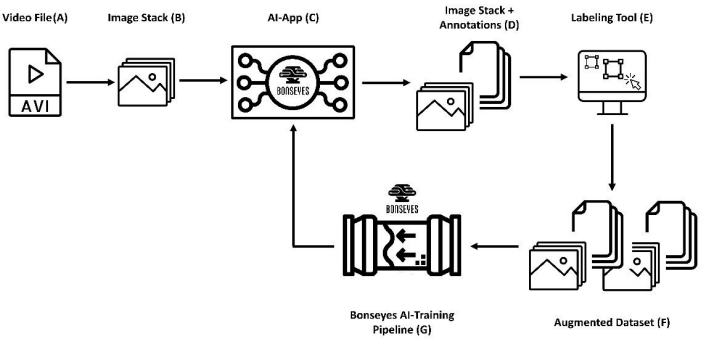



Figure 1*: Schematic overview of the AI-assisted labeling pipeline*

### 2.2 User study

Within a small user study, we evaluated the potential time saving for the annotation task within our assisted labeling pipeline versus the stand-alone usage of the *LabelImg*-tool. We recorded the time needed to annotate 40 images, where each image shows a different amount of objects to be annotated. Only the annotation time within the *LabelImg* tool was recorded; time for setup, pre-processing, loading of data was neglected. The frames were extracted from a central section of a video recording of a laparoscopic cholecystectomy.

## 3 Results

The average annotation time for a single image is 34.8 sec for manual annotation and 28.2 sec for the assisted, pre-labeled annotation with a standard deviation of 12.7 sec and 10.8 sec respectively. The boxplot in Fig. 2 summarizes the results for 40 annotated images.

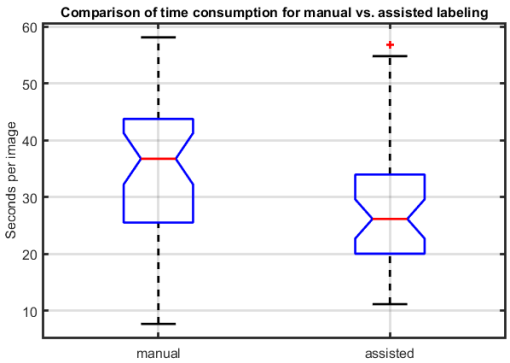



*Figure 2: Boxplot for the time needed to annotate a single image with the manual, stand-alone usage (left) vs. the assisted approach with our pipeline(right). Red lines represent the median: 36.7 sec (left), 26.1 sec (right); notches represent the 95% confidence interval of the median; whiskers represent the 25$^{th}$ and 75$^{th}$ percentiles, respectively; outliers are marked as red cross. Since the notches in the boxplot do not overlap, one can conclude, with 95% confidence, that the true medians do differ.*

## 4 Discussion

The results show an average time reduction by 10.6 sec (28.8%), considering the medians in favor of our AI-assisted approach. The relatively wide distribution (variance of 161 sec² and 116 sec² for manual and assisted approach respectively) are due to the different amounts of objects to be labelled. Users explained outliers in the assisted approach with overlapping bounding boxes, which required moving the larger bounding box first, before the underlying box was exposed for correction.

## 5 Conclusion

We developed a platform independent data annotation pipeline, which is based on the Bonseyes platform. The underlying container technology by Docker ensures a quick and easy setup. This pipeline could be extended towards an online learning solution, which performs a continuous background training process in order to improve detection accuracy. We consider an even further development towards an active learning approach, which will only query the user *informative* samples (e.g. difficult samples). Conclusively, we hypothesize a drastic reduction in human annotation effort with the increasing performance of the underlying AI-App.

## Acknowledgment

This project has received funding from the European Unions Horizon 2020 research and innovation program under grant agreement No 732204 (Bonseyes). This work is supported by the Swiss State Secretariat for Education Research and Innovation (SERI) under contract number 16.0159. The opinions expressed and arguments employed herein do not necessarily reflect the official views of these funding bodies.

## Literature

[1] Llewellynn, Tim, et al. "BONSEYES: platform for open development of systems of artificial intelligence." *Proceedings of the Computing Frontiers Conference*. ACM, 2017.

[2] de Prado, Miguel, et al. "AI Pipeline-bringing AI to you. End-to-end integration of data, algorithms and deployment tools." arXiv preprint arXiv:1901.05049, 2019.

# Hyperspektralbildgebung (HSI) für die intraoperative ortsaufgelöste Quantifizierung des Fettgehalts von Gewebe

*H. Köhler[1], M. Maktabi[1], M. Barberio[2], L. H. Kohler[3], T. Neumuth[1], I. Gockel[4], C. Chalopin[1]*

[1]*Innovation Center Computer Assisted Surgery (ICCAS), Universität Leipzig, Leipzig, Deutschland*
[2] *Institut de Recherche contre les Cancers de l'Appareil Digestive (IRCAD), Straßburg, Frankreich*
[3] *Klinik und Poliklinik für Orthopädie, Unfallchirurgie und Plastische Chirurgie, Universitätsklinikum Leipzig, AöR, Leipzig, Deutschland*
[4] *Klinik und Poliklinik für Viszeral-, Transplantations-, Thorax- und Gefäßchirurgie, Universitätsklinikum Leipzig, AöR, Leipzig, Deutschland*

***Kontakt: hannes.koehler2@medizin.uni-leipzig.de***

***Abstract***

*Die intraoperative Erkennung von Fettgewebe ist bei verschiedenen chirurgischen Eingriffen von großer Bedeutung. Die Machbarkeit der intraoperativen nicht-invasiven Quantifizierung des Fettgehalts von Gewebe mit Hyperspektralbildgebung (HSI) wurde untersucht. Aus HSI-Daten im sichtbaren und nahen Infrarotbereich wurde erstmalig ein ortsaufgelöster Fett-Index (Tissue Lipid Index - TLI) berechnet. Der TLI zeigt deutliche Unterschiede zwischen Fettgewebe (TLI > 40%) und umliegendem Gewebe (TLI < 20%) in den aufgenommen Geweben von Schwein und Mensch. HSI stellt eine vielversprechende Methode zur schnellen und kontaktfreien Lokalisierung und Quantifizierung des Fettgewebes dar.*

**Keywords**: Hyperspektralbildgebung, intraoperative Quantifizierung, Gewebefettgehalt

## 1 Problemstellung

Bei verschiedenen medizinischen Fragestellung spielt das Fettgewebe eine entscheidende Rolle. So wurde etwa der Zusammenhang zwischen Krebs und adipösen Gewebe beschrieben [1]. Adipöses Gewebe wird als Lipid in Adipozyten gespeichert und nach Bedarf freigegeben. Zudem ist adipöses Gewebe ein endokrines Organ und sondert als solches Adipokine ab. Diese unterstützen das Tumorwachstum durch krebsanregende Signalübertragung und durch indirekte Mechanismen wie beispielsweise Angiogenese.

Die Diffuse Reflexionsspektroskopie wurde bereits erfolgreich zur Unterscheidung von gesundem und malignem Gewebe angewendet [2], [3]. Dabei zeigte sich, dass das Verhältnis von Wasser zu Fett ein praktikabler Diskriminierungswert von benignen Gewebe und Tumorgewebe ist [2]. Außerdem kann die Unterscheidung zwischen Fett- und Nervengewebe Anästhesisten bei der Gabe von Anästhetika unterstützen. Dazu wurde an menschlichen Kadavern die Diffuse Reflexionsspektroskopie eingesetzt und verschiedene Klassifikationsmethoden erfolgreich mit einer Sensitivität und Spezifizität von etwa 90 % getestet [4]. Während laparoskopischer Eingriffe konnten die Absorptionsspektren von verschiedenen Geweben wie beispielsweise Fett und Blutgefäßen aufgenommen und deren Unterschiede analysiert werden [5], [6].

Die bisherigen Arbeiten zeigen, dass die intraoperative Erkennung von Fettgewebe eine hilfreiche Unterstützung bei komplexen chirurgischen Eingriffen darstellen könnte. Hyperspektralbildgebung (HSI) ist eine nicht-invasive und kontaktfreie spektroskopische Bildgebungsmethode und erscheint daher besonders geeignet. Ziel dieser Arbeit ist die Untersuchung von HSI für die intraoperative nicht-invasive Quantifizierung des Fettgehalts von Gewebe.

## 2 Material und Methoden

Für die Untersuchung wurden hyperspektralen Daten von flüssigem Schweinefett, Beinscheiben vom Schwein und intraoperativ von menschlichem Unterhautfettgewebe mit dem TIVITA® Tissue-System (Diaspective Vision GmbH, Am Salzhaff, Germany) aufgenommen. Die verwendete HSI-Kamera nimmt 100 Spektralkanäle im sichtbaren und nahen Infrarotbereich von 500 nm bis 1000 nm auf. Die Aufnahmen haben eine räumliche

Auflösung von 640 × 480 Bildpunkten. Eine detaillierte Beschreibung von Aufbau und Messprinzip wurde von Kulcke et al. veröffentlicht [7]. Die Ortsauflösung und die Größe des Bildes sind dabei vom verwendeten Objektiv und dem Messabstand abhängig. Bei makroskopischen Aufnahmen kann eine Ortsauflösung von 0,05 bis 0,5 mm pro Bildpunkt bei resultierenden Bildfeldern von 13 $cm^2$ bis 550 $cm^2$ erreicht werden. Die Beleuchtung während der Aufnahme erfolgt über sechs 20 W Halogenlampen, die direkt in das Kameragehäuse integriert sind. Während der Aufnahme wurden alle externen Lichtquellen ausgeschaltet, um spektrale Artefakte zu vermeiden.

Der Fett-Index (TLI) wird an jedem Bildpunkt der HSI-Aufnahme nach Gleichung (1) berechnet.

$$TLI = \frac{\overline{A''}_{[925-935]}}{\mathrm{a}} \tag{1}$$

Dabei ist $\overline{A''}$ der Mittelwert der zweiten Ableitung der Absorbanz im Wellenlängenbereich von 925 nm bis 935 nm. In diesem Wellenlängenbereich zeigt Fett ein lokales Absorptionsmaximum. Dieser Wert wird durch den Skalierungsfaktor $\mathrm{a} = -0{,}006$ dividiert und der Wertebereich des TLI auf 0 bis 1 beschränkt. In Abbildung *1* sind die Absorbanz und deren zweite Ableitung dargestellt. Die Verwendung der zweiten Ableitung führt zur Eliminierung von systematischen Unterschieden zwischen den aufgenommen Spektren und zur Verstärkung von kleinen Signalanteilen. Das lokale Minimum der zweiten Ableitung der Absorbanz des flüssigen Schweinefetts von 925 nm bis 935 nm wurde als Referenzwert für TLI = 1 verwendet. Anschließend wird die ortsaufgelöste TLI-Matrix mit einen Medianfilter (Blockgröße: 5 Pixel) geglättet.

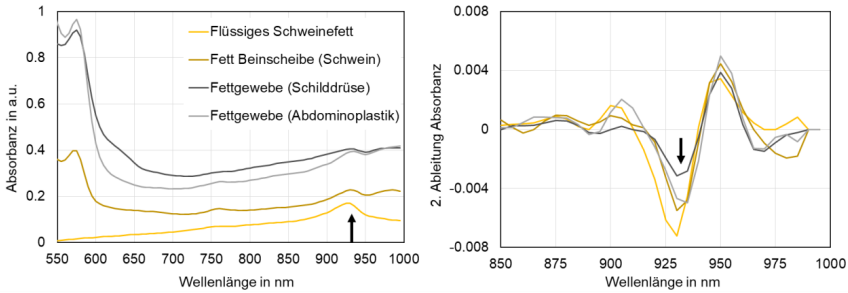



Abbildung 1*: Markierung des Absorptionsmaximums von Fett bei 930 nm in der Absorbanz (links) und zweiten Ableitung der Absorbanz (rechts) der aufgenommenen Proben.*

## 3 Ergebnisse

In Abbildung 2 sind die Ergebnisse der Bestimmung des TLI nach Gleichung (1) und für alle betrachteten Gewebe als Falschfarbendarstellung abgebildet. Wasser und Isopropanol zeigen dabei einen TLI von 0 %, während das flüssige Schweinefett 100 % darstellt. Die Aufnahme der Beinscheibe zeigt eine plausible Verteilung des Fettgewebes zwischen Haut und Muskel, sowie einen mittleren Fettgehalt des Knochenmarks (grüner Bereich in der Bildmitte von Abbildung 2: f). Die markierten Bereiche mit menschlichem Unterhautfettgewebe zeigen einen höheren TLI (> 40%) als das umliegende Gewebe (< 20%). Dabei grenzen sie sich klar gegenüber Haut, Muskel- und Schilddrüsengewebe (Abbildung 2: c, g) sowie der Faszie ab (Abbildung 2: d, h).

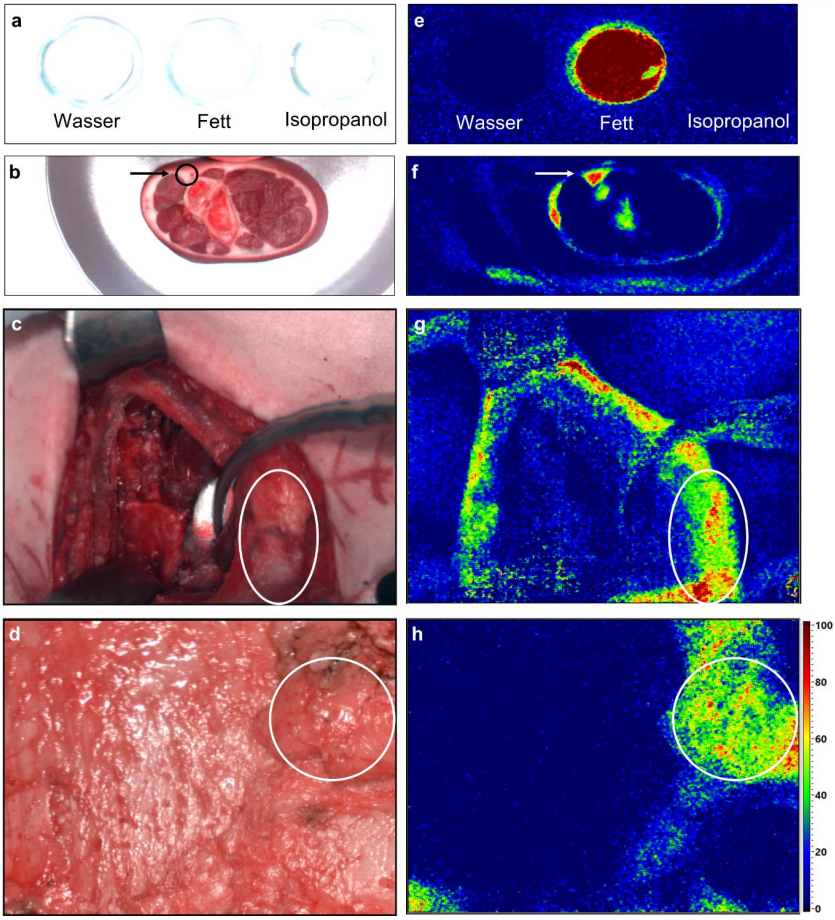



Abbildung 2: *Farbbild (a-d) und Falschfarbendarstellung (e-h) des berechneten Fett-Index (TLI) für Wasser, flüssiges Schweinefett und Isopropanol (a); Beinscheibe vom Schwein (b); Unterhautfettgewebe während Darstellung der Schilddrüse (c) und Abdominoplastik (d). Die Markierungen zeigen die Position des Fettgewebes.*

## 4 Diskussion

Die Arbeit beschreibt erstmalig die Messung und Berechnung eines Fett-Index mit HSI. Im Gegensatz zu bisherigen Arbeiten liefert die beschriebene Methode eine ortsaufgelöste Verteilung des Fettgehalts als Falschfarbendarstellung und quantifizierbare Messwerte. Für die Validierung der Methode sind weitere Untersuchungen an unterschiedlichen Geweben und die histologische Bestätigung des Fettgewebes nötig. Weiterhin sollte der verwendete Wellenlängenbereich auf mögliche zusätzliche Chromophore, die einen störenden Einfluss auf die Messung haben könnten, untersucht werden. Die Beschränkung des Spektralbereichs auf die zur Berechnung verwendeten Wellenlängen würde die Bestimmung des TLI in Echtzeit ermöglichen und eine videofähige intraoperative Überlagerung als Falschfarbendarstellung den chirurgischen Eingriff unterstützen.

## 5 Zusammenfassung

Die genaue Lokalisierung von Fettgewebe ist eine wichtige Fragestellung bei chirurgischen Eingriffen. Es wurde erstmalig gezeigt, dass HSI die intraoperative, kontaktfreie und nicht-invasive Quantifizierung des Fettgehalts von Gewebe ermöglicht. Die beschriebene Methode zeigt einen höheren Fettgehalt bei Fettgewebe (TLI > 40%) im Vergleich zu umliegendem Gewebe (TLI < 20%). Weitere Untersuchungen und eine histologische Validierung müssen diese ersten Ergebnisse bestätigen. Danach kann HSI als intraoperative Methode zur schnellen Bestimmung des Gewebefettgehalts eine wertvolle Unterstützung bei komplexen Eingriffen sein.

## 6 Danksagung und Interessenkonflikte

Die Arbeiten wurden finanziert mit Unterstützung des Bundesministeriums für Bildung und Forschung im Rahmen des Förderprogramms KMU-innovativ: Medizintechnik (13GW0248B).

Die verwendeten Hyperspektralsysteme wurden von Diaspective Vision GmbH entwickelt. H. Köhler ist Angestellter dieser Firma. Langfristig hat Diaspective Vision GmbH wirtschaftliches Interesse an der Entwicklung von Hyperspektralsystemen für den klinischen Einsatz. Alle weiteren Autoren geben an, keine Interessenskonflikte zu haben.

## Referenzen

[1] E. Lengyel, L. Makowski, J. DiGiovanni, und M. G. Kolonin, „Cancer as a Matter of Fat: The Crosstalk between Adipose Tissue and Tumors“, *Trends in Cancer*, Bd. 4, Nr. 5, S. 374–384, Mai 2018.

[2] L. L. de Boer *u. a.*, „Fat/water ratios measured with diffuse reflectance spectroscopy to detect breast tumor boundaries“, *Breast Cancer Res Treat*, Bd. 152, Nr. 3, S. 509–518, Aug. 2015.

[3] N. Reistad, J. Nilsson, O. V. Timmermand, C. Sturesson, und S. Andersson-Engels, „Diffuse reflectance spectroscopy of liver tissue“, in *Biophotonics South America*, 2015, Bd. 9531, S. 95314E.

[4] B. H. W. Hendriks *u. a.*, „Nerve detection with optical spectroscopy for regional anesthesia procedures“, *J Transl Med*, Bd. 13, Nr. 1, S. 380, Dez. 2015.

[5] R. M. Schols *u. a.*, „Differentiation between nerve and adipose tissue using wide-band (350–1,830 nm) in vivo diffuse reflectance spectroscopy“, *Lasers in Surgery and Medicine*, Bd. 46, Nr. 7, S. 538–545, 2014.

[6] R. M. Schols, P. Dunias, F. P. Wieringa, und L. P. S. Stassen, „Multispectral characterization of tissues encountered during laparoscopic colorectal surgery“, *Medical Engineering & Physics*, Bd. 35, Nr. 7, S. 1044–1050, Juli 2013.

[7] A. Kulcke, A. Holmer, P. Wahl, F. Siemers, T. Wild, und G. Daeschlein, „A compact hyperspectral camera for measurement of perfusion parameters in medicine“, *Biomed Tech (Berl)*, März 2018.

# Fast Volumetric Auto-Segmentation of Head CT Images in Emergency Situations for Ventricular Punctures

*Christian Kunz*[1], *Pit Henrich*[1], *Paul Maria Scheikl*[1], *Michal Hlaváč*[3], *Max Schneider*[3], *Heinz Wörn*[1], *Björn Hein*[1,2] *and Franziska Mathis-Ullrich*[1]

[1] *Karlsruhe Institute of Technology, Institute for Anthropomatics and Robotics — Intelligent Process Automation and Robotics, Karlsruhe, Germany*
[2] *Karlsruhe University of Applied Sciences, Karlsruhe, Germany*
[3] *Department of Neurosurgery, University of Ulm, Günzburg, Germany*

Contact: christian.kunz@kit.edu

***Abstract***

*The puncture of the cranial ventricular system inside the human brain is a routine procedure in neurosurgery, which is often performed in emergency situations. It must be done with high accuracy to avoid damage to the brain. One way to assist a surgeon in order to minimize puncture failure is to provide augmented reality (AR)-based assistance during the procedure by superimposing the virtual ventricular system over the patient's head. As part of the assistive visualization we require the extraction of the cranial ventricular system from head computed tomography (CT) scans. In this work we propose a fast and automated algorithm to perform a volumetric segmentation of the ventricular system from CT scans of the brain. We investigated a segmentation pipeline that is structured as follows: 1. Safe area creation, 2. Noise removal, 3. Voxel selection and 4. Cluster selection. In comparison to other approaches, our system has a fast runtime. On average, our proposed method requires less than two seconds for the whole segmentation. Results show a good segmentation accuracy, which is robust against many pathologies. Only the structures of interest are segmented.*

**Keywords**: CT Image Segmentation, Volumetric Segmentation, Ventricular Puncture, Computer Assisted Surgery, Surgical Assistance, Augmented reality

## 1 Introduction

The ventricular system of the brain is located approximately 4 cm deep inside the human head and is characterized by an elongated shape with a width of approximately 1 to 1.5 cm (see Figure 1). The puncture is performed through a burr hole trepanation at the Kocher point on the skull. It is a common entry point for the ventricular puncture. It is chosen in a way that brain damage is minimized. The puncture of the cranial ventricular system has to be performed with high accuracy to avoid damage to the delicate brain tissue and to minimize brain function losses as a misplaced puncture can result in serious brain injuries.

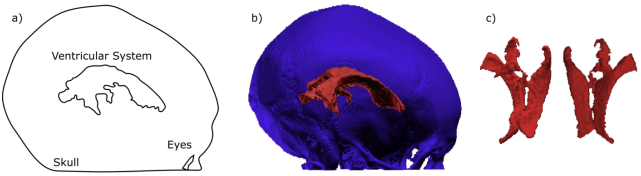



Figure 1: The extracted ventricular system is visualized in red. The patients skull is depicted in blue.

The puncture of the ventricular system is often performed in emergency situations such as intracranial hemorrhage, traumatic brain injury or stroke. The surgical process involves a mandatory computed tomography (CT) scan of the patient's head. If an overpressure (hydrocephalus) is detected inside the ventricular system, the puncture often has to be performed immediately. In the standard of care, the first puncture attempt is only successful in 60% of cases [1]. As the surgeon is not able to see the anatomical structures inside the brain, puncturing is guided by visible and palpable landmarks of surface anatomy. Previously, the authors have proposed an augmented reality (AR)-based system to overcome this difficulty and to assist a surgeon during this procedure [2]. Our system superimposes a 3D model of the ventricular system over the patient's head, enabling the surgeon to see the internal structures of the patient's brain. The surgeon is provided with a detailed and

accurate visualization of the position, shape and size of the target structure to be punctured. In order to visualize an individual patient's anatomy correctly, it is required to segment head CT images to extract the ventricular system of the patient. Time is a decisive factor in this emergency operation. Accordingly, the entire process must be designed in such a way that all required data is available as quickly as possible for optimal treatment of the patient. Due to its great shape variance and fine anatomical structures, the ventricular system is challenging to segment. In addition, head CT scans provide a relatively low dynamic range and an associated unfavourable signal-to-noise ratio of soft tissue.

Segmentation algorithms can be very time-consuming. Various approaches have been presented in the past. Nguyen et al. propose a clustering-based method that requires about three minutes for head segmentation [4]. In order to elaborate regions of interest in the volumetric data, transfer functions are applied to the volume data. However, no detailed evaluation is given. The procedure uses parallelization with Nvidia CUDA. The work of Lan [3] deals with the optimization of transfer functions to better separate adjacent regions. The paper shows, among others, the segmentation of the ventricular system. However, parts of the surrounding brain are also included in the segmentation. Furthermore, there are user-centered approaches where the surgeon interactively creates a segmentation to his needs [5],[6].

In this work, the authors propose an algorithm that provides a fast and automated process for volumetric segmentation of the ventricular system from CT image data.

## 2 Material and Methods

The individual layers of a CT head scan are discrete and can be defined by an intensity matrix $G \in \mathbb{R}^{n \times m}$, where $g_{i,j} = I(i,j)$ represents the intensity at a specific pixel. The CT-layers form a stack of 2D images, which can be interpreted as a 3D volume model. As CT scanners only provide intensity values, images are assumed to be presented in grey-scale.

### 2.1 Volumetric segmentation of the ventricular system

The proposed segmentation pipeline for the ventricular system consists of four steps, as illustrated in Figure 2. First, we create a safe area to remove regions that lie near the skull. Subsequently, the noise is removed. Then we identify the voxels of interest. Finally, appropriate clusters are selected.

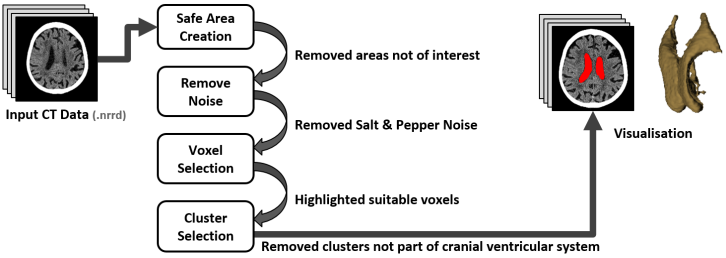



Figure 2: Segmentation pipeline.

The segmentation is performed on volumetric models, stored as '.nrrd' files, which have a height (n), width (m) and depth (k). The coordinate domain is defined in Equation 1.

$$\mathcal{V}_{domain} = \{(x,y,z)^\intercal \in [1,n] \times [1,m] \times [1,k] \subset \mathbb{N}^3\} \tag{1}$$

Further we define the mapping function in Equation 2 for the content of the region, in our case the CT scan intensity values.

$$Map : \mathbb{N}^3 \to \mathbb{Z} \tag{2}$$

We can now define the region $\mathcal{V}$ as the graph of the function $Map$ over $\mathcal{V}_{domain}$ in Equation 3.

$$\mathcal{V} = \{(v, Map(v))^\intercal = (v_{position}, v_{value})^\intercal \mid v \in \mathcal{V}_{domain}\} \subset \mathbb{N}^3 \times \mathbb{Z} \tag{3}$$

Our approach does not need any alignment of the CT-data. It should be noted that when used for intraoperative navigation later, the model needs to be registered to the patient.

**Safe area creation:** Without a safe area creation, our algorithm finds many wrongfully included clusters near skeletal structures, because large bodies of fluid can accumulate between the brain tissue and the skull (subarachnoid space), for example caused by brain degenerative diseases. To remove areas close to bone in the image, we grow bone structures using a bone growing operator defined in Equation 4. As it is assumed that the ventricular system is centred within the human brain, the operator removes sections not of interest that are within a certain vicinity of skeletal structures. This approach allows a high degree of parallelism and the additional step greatly improves results. Here we use a local operator that also considers the neighborhood $N_{s,r}$ of a voxel. We define the neighbourhood $N_{s,r}$ as a set of all voxels around the selected voxel $s$ with radius $r$, where $r$ is the maximum distance between $s$ and surrounding voxels. Different norms can be used to calculate the distance, such as the infinity or euclidean norm.

$$Bone_r^{local} : \mathcal{V} \to \mathbb{Z}, Bone_r^{local}(s) = \begin{cases} 900 & |\{n \in N_{s,r} \mid n_{value} \geq 900\}| \geq 1 \\ s_{value} & else \end{cases} \tag{4}$$

We choose a radius of $r = 1$ voxels and perform $\frac{max(n,m,k)}{30}$ iterations, where $n, m$ and $k$ are the number of voxels along each dimension, thus, only taking into account the directly neighbouring voxels. Using a divisor of 30 was chosen empirically by utilizing knowledge of the position of the cranial ventricular system within the human brain as a basis. Figure 3 a) shows an example of a generated safe zone through the bone growing operator.

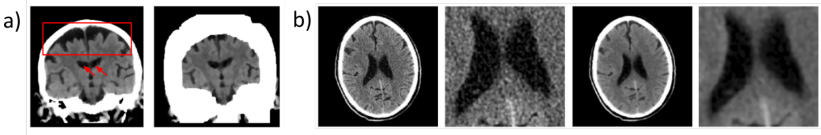



Figure 3: a) Left: Subarachnoid spaces overlying the brain convexity close to bone with a similar density as the ventricular system, that are causing issues during later stages, Right: Areas close to bone removed, decreased size of problematic area. b) The left two pictures show the original image, the right two images the filtered ones using the bilateral filter.

**Noise removal:** A conventional bilateral filter is used on the image data as shown in Figure 3 b) to smooth the image and remove noise during preprocessing.

**Voxel selection:** After preprocessing, the voxels with desired Hounsfield Units (HU) are selected. This involves binarization to create a mask with the voxels of interest. Equation 5 shows our binarization function applied to each voxel element $s$ in $\mathcal{V}$. Generally, two different operators can be used to perform the binarization: A point based operator and a local operator. The point based operator focuses on the selected voxel only, not taking into account surrounding voxels. We are using a point based operator to segment the skull. Here we although use a local operator that also considers the neighborhood $N_{s,r}$ of a voxel. A voxel is only selected if a defined percentage ($\alpha$) of the voxels' neighbourhood, including the voxel itself, have values within a predefined range. The ventricular system consists mainly of cerebrospinal fluid, which is represented by the similar HU as water in a range between 0 and 14. Thus, we choose r = 2 voxels and $\alpha = 0.5$, which effectively means that a voxel is selected only if 50% of the voxels within radius 2 have a HU between 0 and 14.

$$VS_{min,max,\alpha,r}^{local} : \mathcal{V} \to \{0,1\}, VS_{min,max,\alpha,r}^{local}(s) = \begin{cases} 1 & |\{n \in N_{s,r} \mid min \leq n_{value} \leq max\}| \geq \alpha |N_{s,r}| \\ 0 & else \end{cases} \tag{5}$$

**Cluster selection:** In the last step we select the clusters that represent the ventricular system using Equation 6. It is assumed that the ventricular system is the largest homogeneous region filled with cerebrospinal fluid inside the human brain after creation of the safe zone. First we determine the set $C$ of discrete clusters in the selected region. We define a cluster as a contiguous set of voxels that have all been assigned the value 1 in

the mask. All clusters $C_{cvs}$ are selected that are at least 50% ($\beta$) of the size of the biggest cluster $M_{max}$.

$$C_{cvs} = \{M \in C \mid |M| \geq \beta |M_{max}|\} \tag{6}$$

The reason to select more than just the biggest cluster is, that the ventricular system consists of multiple ventricles, i.e. left and right lateral ventricles, which do not necessarily appear to be connected in the CT data due to noise and inaccuracies in the scan. In Figure 4 a) all detected clusters are visualized, b) and c) show clusters of the ventricular system.

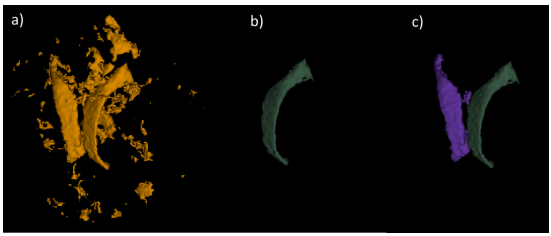



Figure 4: Left: Example $M \in C$, Middle: $M_{max}$, Right: $C_{cvs}$ with $\beta = 0.5$

### 2.2 Volumetric segmentation of the skull and the skin

In addition to the segmentation of the ventricular system, the proposed approach is also applied to segmentations of other anatomical structures of the human head, such as the skin and the skull. Similar operators are derived for the segmentation. For bone segmentation we are using a local operator using HUs in a range from 300 to 1900. For the segmentation of the skin voxels near air are selected with HUs from -750 to 200. Segmentations of the skin and skull are shown in Figure 5.

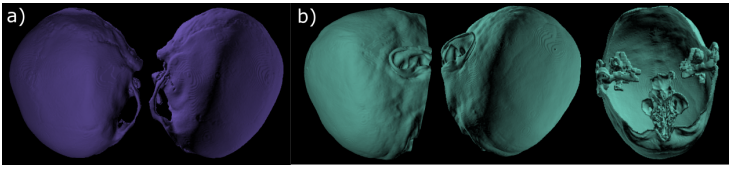



Figure 5: a) Segmented skull from a CT head scan. b) Segmented skin.

## 3 Results

Volumetric segmentations were performed on 26 head CT scans from humans, consisting of 15 data sets without pathology (a) and 11 with pathology (b). The extracted ventricular systems are presented in Figure 6 a) and b). We measured the runtime of the extraction of the ventricular system, the skin and the skull in all 26 data sets to quantitatively analyze our approach. The reading and writing of the .nrrd volume in raw format and with gzip compression was evaluated. Table 1 shows the results. All measurements were performed on a workstation with the following configuration: Intel Core i7 8700k, Nvidia GeForce 1070Ti, 32GB DDR4 RAM, M.2 PCIe SSD.

| | Raw | | | GZIP | | |
|---|---|---|---|---|---|---|
| | Ventricle System | Skull | Skin | Ventricle System | Skull | Skin |
| Average (s) | 1.940 | 1.430 | 1.331 | 2.689 | 2.159 | 2.150 |
| Standard Deviation (s) | 0.158 | 0.098 | 0.094 | 0.217 | 0.130 | 0.146 |

Table 1: Runtime behaviour of our approach

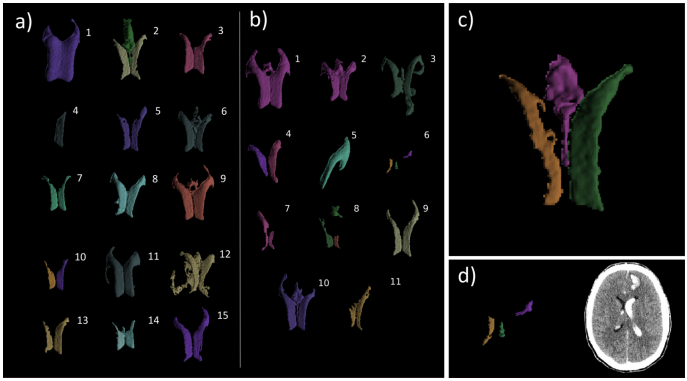



Figure 6: a) Example of extracted cranial ventricular systems without pathology, b) Ventricular systems with pathology (1,2: Cerebral atrophy, 3: Blood reabsorbed post bleeding, status post endovascular coiling of an aneurysm, 4: Deformation due to hemorrhage, 5: Prominent midline shift, 6: Hydrocephalus intraventricular blood clot, status post endovascular coil, 7: Midline shift, 8,9,10: Slim ventricles, 11: Slim ventricles due to chronic subdural haematoma c) The extracted ventricular system a-4 with $\beta = 0.5$ d) Scan b-6 of patient with a cranial ventricular system infused with blood.

To evaluate the quality of the segmentation, the results were shown to an expert neurosurgeon, who interprets head CT scans on a daily basis. For this purpose, the segmentation results were loaded into MITK (The Medical Interaction Toolkit - https://www.mitk.org) so that a layer-by-layer assessment was possible. The created mask of the ventricular system was overlayed with the original scan. Thus, a pixel-wise assessment of the created mask in regard to the original scan was possible. In addition, the 3D representation was examined. The extraction for each CT scan was rated according to four questions.

1. **How good is the quality of the segmentation on a scale from 1-5.** Result: The quality of the segmentation was rated with 4.5/5.
2. **Are all relevant areas included?** Result: As seen in the results in Figure 6 a) and b) the left and right lateral ventricles were extracted in most cases, with no pathology and no bigger deviation in the patients anatomy.
3. **How is the quality of surface detail of the extracted region?** Result: The medical expert rated some of the visualized surfaces of the ventricles as too irregular. This is due to the low resolution of CT-scans. If surface models are generated from the volumetric segmentation smoothing algorithms can be applied.
4. **What problems exist?** Result: Some ventricles, especially those with pathology, are sometimes not properly segmented and some parts are missing. Most problems occur with blood clots.

## 4 Discussion

As shown in the evaluation, our approach is able to extract the ventricular system, the skull and the skin from head CT scans. Runtime measurements show a fast execution on the data. The extraction of raw data is about 30% faster than with gzip compression. Storing the results uncompressed leads to a larger size of the files, e.g. gzip: 50.1 MB, raw: 101.5 MB. It is not possible to make a general statement about the file sizes, as this characteristic strongly depends on the scan data. Since the runtime of both approaches has the same order of magnitude, the advantage of a smaller data storage outweighs the slightly faster method and compression of data is recommended. Several pathologies, such as (hyperdens) intraventricular blood clots or subdural collections with consecutive brain shift and deformation of the ventricular system, may lead to erroneous segmentation of the ventricular system. The ventricular systems without pathology (Figure 6 a) were segmented successfully, except the ventricle a-4. As seen enlarged in Figure 6 c) the left ventricle is much smaller than the correctly

extracted right one. We used a threshold $\beta = 0.5$ in Equation 6. If $\beta$ is parameterized with 0.25 the cluster is selected correctly. In this case, the algorithm must be parameterized differently by an expert. Most faulty segmentations occur when ventricles have infused blood, as seen in Figure 6 d) as an example. As blood has a different HU, our approach is not able to extract areas infused with blood or other foreign substances without user adjustment. Currently, the quantitative analysis of our approach is limited due to the absence of a ground truth. To provide this, fully manually segmented CT scans of all 26 data sets are required to be provided by a medical expert, who has the necessary domain knowledge. The annotation of medical data remains a major challenge. During evaluation the representation of the surface was experienced as to irregular, presumably due to the scan quality, which lies at 1mm per scan value. Thus, we implemented additional smoothing methods to receive a better visualization experience. Additionally we customized our binarization equation to segment the skin and the skull. In preliminary experiments, the algorithm shows promising results for segmentation of the skin and the skull. In the future, we plan to perform a quantitative evaluation of our approach and work on the robustness against pathologies, especially infused blood as this showed challenging outcomes. One approach is to expand the range of HUs of the binarization function to include such regions or to segment these regions as separate clusters. One possible area of application is the drug treatment of blood clots in the ventricular system.

## 5 Conclusion

In this work we have demonstrated an algorithmic approach for fast volumetric segmentation of the ventricular system, skull and skin from CT head images with the most challenging task being the segmentation of the ventricular system. We have proposed a segmentation pipeline to correctly perform this task. The method consists of the creation of a safe region, pre-processing through filters, a binarization function and clustering. Our binarization function uses a local operator and takes the Hounsfield Unit for cerebrospinal fluid and the neighbourhood of a voxel into account. Our approach can also be used to segment the skin and the skull. The fast segmentation speed of the proposed method is advantageous compared to existing segmentation methodologies, which are often time consuming, especially in emergency situations speed is an important factor. While the proposed algorithm works fast, an expert rated the segmentation result with a rating of 4.5/5. In most cases our algorithm is able to extract the lateral ventricles, that are necessary for the ventricular puncture.

## 6 Acknowledgement

This work was funded by the Federal Ministry of Education and Research within the project 'HoloMed - Context-sensitive support of a surgeon in the operating room by Augmented Reality'.

## 7 References

[1] Paul M Foreman, Philipp Hendrix, Christoph J Griessenauer, Philip GR Schmalz, and Mark R Harrigan. External ventricular drain placement in the intensive care unit versus operating room: evaluation of complications and accuracy. Clinical neurology and neurosurgery, 128:94–100, 2015.

[2] Christian Kunz, Michal Hlavac, Max Schneider, David Puljiz, Steffen Peikert, and Bjoern Hein. A system for augmented reality guided ventricular puncture using a hololens: Design, implementation and initial evaluation [in press]. In 17. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (CURAC), Leipzig, Germany, 13 - 15 September 2018, pages 132–137, 2018.

[3] Shouren Lan, Lisheng Wang, Yipeng Song, Yu-ping Wang, Liping Yao, Kun Sun, Bin Xia, and Zongben Xu. Improving separability of structures with similar attributes in 2d transfer function design. IEEE transactions on visualization and computer graphics, 23(5):1546–1560, 2016.

[4] Binh P Nguyen, Wei-Liang Tay, Chee-Kong Chui, and Sim-Heng Ong. A clustering-based system to automate transfer function design for medical image visualization. The Visual Computer, 28(2):181–191, 2012.

[5] F-Y Tzeng, Eric B Lum, and K-L Ma. An intelligent system approach to higher-dimensional classification of volume data. IEEE Transactions on visualization and computer graphics, 11(3):273–284, 2005.

[6] Yingcai Wu and Huamin Qu. Interactive transfer function design based on editing direct volume rendered images. IEEE Transactions on Visualization and Computer Graphics, 13(5):1027–1040, 2007.

# IVUS-Simulation for Improving Segmentation Performance of Neural Networks via Data Augmentation

*F. Sommer*[1], *L. Bargsten*[1], *A. Schlaefer*[1]

[1] *Hamburg University of Technology, Institute of Medical Technology, Hamburg, Germany*

Contact: lennart.bargsten@tuhh.de

***Abstract***

*Convolutional neural networks (CNNs) produce promising results when applied to a wide range of medical imaging tasks including the segmentation of tissue structures like the artery lumen and wall layers in intravascular ultrasound (IVUS) image data. However, large annotated datasets are needed for training to achieve sufficient performances. To increase the dataset size, data augmentation techniques like random image transformations are commonly used. In this work, we propose a new systematic approach to generate artificial IVUS image data with the ultrasound simulation software Field II in order to perform data augmentation. A simulation model was systematically tuned to a clinical data set based on the Fréchet Inception Distance (FID). We show that the segmentation performance of a state-of-the-art CNN with U-Net architecture improves when pre-trained with our synthetic IVUS data.*

**Keywords**: IVUS simulation, segmentation, data augmentation

## 1 Problem

Cardiovascular diseases like atherosclerosis are the leading cause of death globally [1]. In assessing the severity of atherosclerosis, intravascular ultrasound (IVUS) is widely used complementary to angiography as it provides additional quantitative and qualitative information on plaque in the vessel wall [2, 3]. Automatic segmentation of the image components based on deep learning has the potential to facilitate tracking the progress of the disease and reduce treatment time [2]. A challenge in developing deep learning algorithms for clinical applications is the need for a large amount of annotated image data. To overcome this problem, data augmentation is widely used [4]. The aim of this work is to develop a systematic approach to generate realistic data from IVUS simulations with Field II [5] that is feasible for pre-training a neural network. Although previous work on IVUS simulations has been presented by [6, 7, 8], the focus of our work is to systematically evaluate the improvement of image quality, leading to valuable results for our deep learning application.

## 2 Material and Methods

For the ultrasound simulation, the Matlab program Field II, developed by Jensen et al. [5], was used. It is based on the spatial impulse response of an ultrasound pulse emitted by a user-defined aperture. The tissue is modelled by a discrete distribution of point-scatterers with varying back-scattering amplitudes defining the echogenicity of the specific tissue.

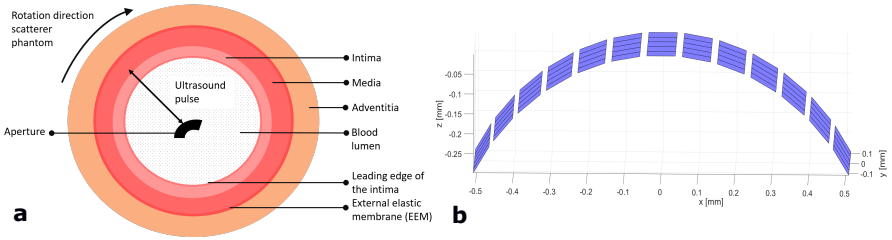



Figure 1: **a:** Schematic representation of the cross-section of an artery. In the simulation, the aperture is placed at the coordinate center and the artery phantom rotates around it. **b:** Aperture used for the simulation in Field II.

### 2.1 Simulation setup

Our simulation model is based on a phased multi-element transducer and therefore consists of a curved aperture representing the active sector, and an artery phantom rotating around it in increments of one degree (see Figure 1). To reduce side lobe artifacts [9] and computation time, a 30 degree sector of the scatterer phantom is used at an instant. The transducer dimensions were initialized according to the Philips Eagle Eye Platinum transducer [10] and pre-tuned with respect to the point spread function. An excitation frequency of 20 MHz as implemented in the Philips transducer, and a sampling frequency of 500 MHz were used.

The basic setup of the artery phantom was chosen to mimic the typical appearance of IVUS images: The trilaminar look made up of the blood lumen with the characteristic speckle noise, and the arterial wall separated by two bright acoustical interfaces: the leading edge of the intima and the external elastic membrane (EEM) separating media and adventitia (see Figure 1).

In our model, the layers are elliptically shaped with thicknesses varying within a typical range according to [3, 11]. A total number of 10000 scatterers was used. As proposed by [6], the amplitudes for each layer are based on the tissue-specific differential backscattering cross-section (DBC) taken from [6] while the scatterer density was kept constant, a choice that was also made to simplify future deep learning applications.

### 2.2 Parameter tuning

In order to enable systematic adjustments in transducer, scatterer phantom and image processing parameters, the quality of the simulation results was quantitatively measured for each choice of parameters with a sample of 1000 images. The metric used was the Fréchet Inception Distance (FID) [12], a tool widely applied in deep learning to evaluate the structural similarity of two sets of images. Here, smaller values indicate higher similarity. As clinical reference data, we used a publicly available data set including 435 annotated images provided by [13]. These images were captured with a frequency of 20 MHz and have spatial dimensions of 384×384 pixels.

The scatterer amplitudes as well as the number of transducer elements and excitation function were tuned successively with regard to the FID score. While tuning scatterer and transducer parameters, the image post-processing was applied with constant settings that gave the image a subjectively realistic appearance to limit the influence of inter-dependence. Then, the definition range for logarithmic contrast enhancement and the Gaussian smoothing filter were tuned in parallel. Finally, calcified plaque was added to the scatterer model and a shadow artifact was superimposed on the scatterer amplitudes.

### 2.3 Evaluation of segmentation performance

With the final parameter settings, a set of 1000 images was generated, half of which had calcified plaque. Segmentation masks of the same size as the images encoding the regions blood lumen and intima/media analogue to the annotations of the publicly available data set were created with the scatterer data. The generated image data was then used to pre-train a state-of-the-art neural network with U-Net architecture for segmentation [14]. The network consisted of residual blocks in the down- and upsampling path. In each of the three downsampling blocks, the spatial sizes of the input images were halved while the numbers of feature maps were doubled. The upsampling blocks operated vice versa. The network inputs had a size of 10×384×384, where the first dimension denotes the batch size and the second and third dimension the image width and height respectively.

The pre-trained network was then fine-tuned with the clinical IVUS images. Afterwards, the results of the pre-trained networks were compared to the performance of a network with the same architecture that had been trained with clinical reference data only to evaluate the effect on the segmentation performance. We repeated the training procedure with different numbers of utilized clinical IVUS images to systematically assess the improvement of the segmentation performance by means of the synthetic data.

## 3 Results

This section outlines the process of parameter tuning and states the according FID-scores. Then, the segmentation results of neural networks with and without pre-training are given.

### 3.1 Parameter tuning

In the first step, the artery phantom was tuned with respect to the distribution of the scatterer amplitudes. An initial transducer setup of ten active elements and an excitation duration of three periods of a sine function was used. Following [6]'s approach, the blood and all three layers of the arterial wall were modeled separately. As

Table 1: *FID-scores for different numbers of transducer excitation periods and active elements*

| **FID-score** | **8 elements** | **10 elements** | **13 elements** |
|---|---|---|---|
| **2 periods** | 126.19 | 109.63 | 105.04 |
| **3 periods** | 99.29 | **91.89** | 94.22 |
| **4 periods** | 113.84 | 115.08 | 125.30 |

Table 2: *Best FID-scores for each stage of parameter settings.*

| **Setting** | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **FID-score** | 108.83 | 101.61 | 91.89 | 83.11 | 74.71 | **74.08** |

proposed by [5], the scatterer amplitudes were normally distributed with zero mean and the variance taken to be the DBC-value. As reflected by the high and therefore poor FID-score of 108.83 (see Table 2, Setting 1), this model has several drawbacks: Firstly, this model appears to be a good model for speckle in the blood lumen but it fails to depict the dense speckle pattern in the artery wall. Since the scatterer density is kept constant, a zero-mean amplitude offers too few degrees of freedom to model both at the same time. To overcome this issue, the mean value of the scatterer amplitudes in the arterial wall was changed to the DBC-value as proposed by [6] leading to an FID-score of 101.61 (see Table 2, Setting 2).

Yet, this model cannot account for the predominant scattering at the echodense edges described above. Moreover, it ignores the fact that in IVUS images the intima and media merge into each other. Therefore, these two layers were combined for the next model such that the arterial wall model was reduced to 2 layers and a spatial-dependent factor in shape of a $Chi^2$-function over the radial axis was superimposed on the scatterer amplitudes of both layers separately. To model the varying decay of brightness towards the edge of the image in the reference set, the $Chi^2$-function with 2.5 degrees of freedom was implemented with varying definition range, while the mean brightness was kept constant. Thereby, the FID score was reduced to 91.89 (see Table 2, Setting 3).

With this set of scatterer parameters, the transducer settings were varied. Table 1 shows that the best FID score remained the one obtained with the transducer settings that were used in the previous steps: ten active elements and an excitation pulse of three periods of a sine function filtered through a hanning window.

Then, the scatterer-model was fine-tuned by implementing a constant factor on the amplitudes in the three regions. The best result with an FID-score of 83.11 (see Table 2, Setting 4) was achieved with a factor of 0.25 applied to the inner wall layer leaving the other amplitudes unchanged.

The image post-processing was then optimized with regard to logarithmic compression and subsequent Gaussian smoothing. The best FID-value of 74.71 was found for logarithmic compression to a definition range of $[1, 35]$ and Gaussian filtering with $\sigma = 1.8$. The settings used for previous tuning were compression to $[1, 15]$ and $\sigma = 1.2$.

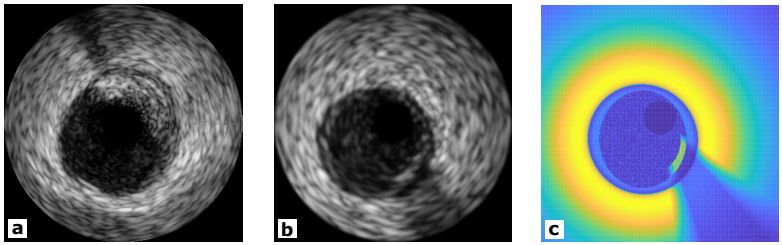



Figure 2: **a:** Clinical IVUS image with calcified plaque at the top, leading to a shadowing artifact. **b:** Simulated IVUS image with a calcified plaque at the bottom. This image was synthesized with the simulation settings yielding the highest FID score. **c:** Distribution of backscattering amplitudes related to **b**. Brighter colors indicate higher amplitudes.

Table 3: Segmentation performances of the trained neural networks reagrading Dice coefficients. Training was performed using different dataset sizes with and without pre-training by means of simulated IVUS data.

| **Training set size** | **Dice coefficient** [%] | | | |
|---|---|---|---|---|
| | Lumen | | Intima/Media | |
| | pre-training | no pre-training | pre-training | no pre-training |
| 50 | $\mathbf{92.23 \pm 1.51}$ | $88.37 \pm 1.83$ | $\mathbf{80.17 \pm 2.02}$ | $75.75 \pm 0.90$ |
| 100 | $\mathbf{94.36 \pm 0.34}$ | $92.12 \pm 0.75$ | $\mathbf{82.31 \pm 0.35}$ | $78.65 \pm 1.05$ |
| 150 | $\mathbf{94.68 \pm 0.36}$ | $92.83 \pm 0.41$ | $\mathbf{83.87 \pm 0.78}$ | $80.48 \pm 0.98$ |
| 200 | $\mathbf{94.60 \pm 0.39}$ | $93.48 \pm 0.77$ | $\mathbf{84.45 \pm 0.95}$ | $82.39 \pm 0.59$ |

Table 4: Segmentation performances of the trained neural networks regarding modified Hausdorff distances. Training was performed using different dataset sizes with and without pre-training by means of simulated IVUS data.

| **Training set size** | **Modified Hausdorff Distance [pixels]** | | | |
|---|---|---|---|---|
| | Lumen | | Intima/Media | |
| | pre-training | no pre-training | pre-training | no pre-training |
| 50 | $\mathbf{3.11 \pm 1.52}$ | $6.52 \pm 3.76$ | $\mathbf{2.65 \pm 0.57}$ | $3.55 \pm 1.28$ |
| 100 | $\mathbf{1.18 \pm 0.29}$ | $2.93 \pm 1.36$ | $\mathbf{1.71 \pm 0.16}$ | $2.84 \pm 0.53$ |
| 150 | $\mathbf{1.12 \pm 0.39}$ | $2.18 \pm 0.95$ | $\mathbf{1.47 \pm 0.14}$ | $2.44 \pm 0.13$ |
| 200 | $\mathbf{1.85 \pm 0.91}$ | $2.74 \pm 1.72$ | $\mathbf{1.24 \pm 0.15}$ | $2.19 \pm 0.20$ |

Lastly, calcified plaque was incorporated by increasing the scatterer amplitudes in part of the blood lumen by a factor chosen according to the best FID-value, and a shadow artifact was added. The corresponding FID is 74.08 (see Table 2, Setting 6).

Figure 2 shows a clinical reference image (a), a simulated image generated with settings which led to the highest FID-score (b) and a two-dimensional representation of the corresponding distribution of backscattering amplitudes (c).

### 3.2 Evaluation of segmentation performance

Table 3 and Table 4 show the segmentation performances of neural networks achieved with and without pre-training by means of Dice coefficient and modified Hausdorff distance [15]. The number of used clinical images was systematically increased from 50 up to 200. Evaluation was performed with five-fold cross-validation. The simulated data was only used for pre-training, not for evaluation.

Both tables present a significant increase in segmentation performance when the neural networks are pre-trained with our simulated data, even when utilizing 200 clinical images. Generally, the improvement increases with decreasing number of used clinical images.

## 4 Discussion

As can be seen from Table 2, by systematic tuning of scatterer, transducer and image processing parameters, the similarity of the artificially generated image data and the clinical data could be improved in each stage. Altogether, starting from a literature based scatterer setting for ultrasound in general, an improvement from the initial FID-score of 108.83 down to 74.08 was achieved, which is a decrease of 31.93 %. The striking influence of the transducer settings as depicted in Table 1 can be explained by the fact that the excitation duration limits the axial resolution whereas the lateral resolution is influenced by the width of the aperture [16]. Also worth mentioning is the improvement of FID as a result of image post-processing. For future work, we therefore suggest to investigate the potential of image post-processing by means of domain adaption strategies [17].

When comparing the images to the reference data visually, the lumen appears too bright and blurry with increased compression and smoothing, whereas the adventitia, the larger and therefore dominant part with respect to the FID, corresponds well to the reference data. This shows that the scatterer model still lacks to model all anatomical parts properly. By loosening the condition of an equal scatterer distribution, the outcome

could potentially be improved. Furthermore, our model is limited due to the elliptical shape of lumen and arterial wall and the lack of certain kinds of artifacts e.g. due to bifurcations. Furthermore, some important physical effects such as attenuation and multiple scattering are neglected in the simulation process of Field II and could only partially be incorporated by variation of scatterer amplitudes.

Nevertheless, we were able to present a new systematical approach to IVUS simulation with Field II that was proven to enhance the overall segmentation performance of a state-of-the-art U-Net, observable in both, the Dice coefficient and the modified Hausdorff distance. Due to the implementation with stochastic influence in many parameters, it is possible to automatically generate image data with great variation which could be of good practical value for clinical applications. Furthermore, the systematic approach allows for adjustment to new image data from different IVUS systems.

## 5 Conclusion

In this work, we have presented a new systematic approach to generate realistic artificial IVUS image data with Field II. We were able to show that the overall segmentation performance of a state-of-the-art neural network could be improved by pre-training with our synthetic data. Following this approach, data augmentation for any set of clinical IVUS data can be performed. Our results can therefore help to overcome the lack of annotated clinical data and facilitate the development of deep learning applications for the catheter laboratory.

## Acknowledgment

This work was partially funded by the European Regional Development Fund (ERDF) and by the Free and Hanseatic City of Hamburg.

## References

[1] Mendis, S., et al.: (2011) Global atlas on cardiovascular disease prevention and control. World Health Organization

[2] Zhao, F., et al. (2015) Computer vision techniques for transcatheter intervention. IEEE Journal of Translational Engineering in Health and Medicine **3**, 1–31

[3] Tuzcu., E.M., et al. (2010) Coronary intravascular ultrasound: a closer view. Heart **96**(16), 13–18

[4] Krizhevsky, A., et al. (2012) Imagenet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems, vol. 1, pp. 1097–1105

[5] Jensen, J.A., et al. (1992) Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **39**(2), 262–267

[6] Rosales, M., et al. (2005) A basic model for ivus image simulation. In: J.S. Suri, et al. (eds.) Handbook of Biomedical Image Analysis, vol. 1, chap. 1, pp. 1–55. Springer

[7] Cardoso, F.M., et al. (1992) Realistic ivus image generation in different antraluminal pressures. Ultrasound in medicine and biology **39**(2), 262–267

[8] Groot, S., et al. (2006) Ivus simulation based on histopathology. Computers in Cardiology pp. 681–684

[9] Bom, N., et al. (1972) An ultrasonic intracardiac scanner. Ultrasonics **2**(2), 72–76

[10] Philips: Eagle Eye Platinum Specifications (2019 (accessed June 24, 2019)). URL `https://www.usa.philips.com/healthcare/product/HC85900P/eagle-eye-platinum-digital-ivus-catheter/specifications`

[11] Perry, R., et al. (2013) Coronary artery wall thickness of the left anterior descending artery using high resolution transthoracic echocardiographynormal range of values. Echocardiography **30**(7), 759–764

[12] Heusel, M., et al. (2017) Gans trained by a two time-scale update rule converge to a local nash equilibrium. In: 31st conference on neural information processing system, pp. 6629–6640

[13] Balocco, S., et al. (2014) Standardized evaluation methodology and reference database for evaluating ivus image segmentation. Computerized medical imaging and graphics **38**(2), 70–90

[14] Ronneberger, O., et al. (2015) U-net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - Part III, pp. 234–241

[15] Dubuisson, M.P., et al. (1994) A modified hausdorff distance for object matching. In: Proceedings of 12th International Conference on Pattern Recognition, vol. 1, pp. 566–568

[16] Azhari, H.: (2010) Basics of Biomedical Ultrasound for Engineers. John Wiley & Sons

[17] Zhu, J.Y., et al. (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2242–2251

# Comparison of Similarity Measurements and Optimizers for Intraoperative Registration of 2D C-arm Images with Preoperative CT Data in Computer-Assisted Spine Surgery

*J. Alvarez-Gomez[1], H. Roth[1], J. Wahrburg[1]*

*[1] Institute of Control Engineering and Center for Sensor Systems (ZESS), University of Siegen, Siegen, Germany.*

Contact: julio.agomez@uni-siegen.de

***Abstract***

*In this paper, we compare twelve similarity measurements and optimizers for 2D/3D registration in our computer-assisted spine surgery application. We use an iterative procedure to find the registration of X-ray images, respecting to a DICOM volume from a CT scan. For each registration result, we compute the averaged RMS error of four points taken on the phantom used for the CT scan and X-rays. We registered pairs of five anteroposterior- and lateral-projection X-ray images with each similarity measurement and optimizer. We computed a set of 3600 registrations that we analyzed to find the most suitable similarity measurement and optimizer for our application. Using the entire set of results, we found Best Neighbors is the most suitable optimizer with an error of 2.38mm. Using the results of the five best-performing optimizers, we found mutual information and gradient correlation give the most precise registrations with 2.29mm and 2.36mm of error, respectively.*

**Keywords**: 2D/3D registration, intraoperative registration, similarity measurements, optimizers, optimization algorithms, spine surgery

## 1 Introduction

2D to 3D registration algorithms have been studied intensively based on intensity- and stochastic-base methods. Additionally, these algorithms mostly use iterative approaches. In our research on spine surgery and implantation of pedicle screws, we use preoperative CT-Scans taken for planning. Further, we register them using intraoperative C-arm images. Our registration uses as well an iterative approach, which can be accurate, but it depends on, how close is the given initial pose to the actual registration pose [1]. Iterative approaches also give different results depending on the body segment to be registered, the used image measurement, and the optimizer [2][3][4]. In another work, we are developing a convolutional neural network that predicts an initial pose, which is close to the actual registration pose. Therefore, we need to define a registration procedure with the proper similarity measurement and optimizer for our specific application.

There are several similarity measurements studied in the literature, such as sum of square differences, cross-correlation, and mutual information. We implemented twelve of them, and we wanted to determine which one is the most suitable measurement to use in our application. Additionally, we implemented five popular optimizers, among them best neighbors and gradient descent, with the idea of selecting the optimizer that gives the most accurate result.

In our facilities, we have a spine phantom and its CT-Scan in DICOM format. We also captured some X-ray images of the phantom, i.e., five with anteroposterior projection (AP) and another five with lateral projection (LAT). Each of the ten images is different in the position of the X-ray detector respecting to the phantom. Our registration process is performed using a combination of one AP and one LAT image. Therefore, it was possible to make 25 registration procedures for each image measurement with a specific optimizer.

Using our planning software, we could select four landmarks on the CT-Scan of the phantom, which later we could also identify on the physical phantom. The measurements on the phantom were carried out using a rigid body (RB) with a calibrated tip point, and an optical tracking system (OTS). This process gives an equivalence between the real world and the DICOM volume. When we execute a registration, it gives us a transformation matrix between the DICOM volume and the real world, so the previous relation between points is used to evaluate the registration result quantitatively. Consequently, the points in the real world are transformed into the DICOM frame and compared with the landmarks' positions taken with the planning software. We calculated the RMS error of each

landmark and then averaged them to get an average RMS error of the four landmarks. We used this measurement as a coefficient that indicates the quality of every registration and can compare one registration with another.

Then we started finding the best optimizer by computing a decision matrix that compares the average RMS error of each optimizer. Finally, using the average errors of the best optimizers, we determined the best similarity measurement with the help of a second decision matrix. In the end, we obtained the most suitable similarity measurement and optimizer for our application in computer-assisted spine surgery.

# 2 Materials and Methods

## 2.1 Hardware

In our laboratory, we have an L1-L5 and sacrum phantom, from whom we have its CT-scans. In our facilities, it is available a Ziehm Vario 3D C-arm, which has 130° of rotation range and gives X-ray images with a resolution of 568x568 pixels with a density of 0.365mm/pixel. Additionally, we use an NDI Vega stereo OTS for computing measurements over the phantom. We utilize RBs with calibrated pivoted tooltip for taking such measurements with the OTS, and they are reported for the manufacturer to give sub-millimeter precision.

## 2.2 Software Tools

We work with the MeVisLab framework as the leading software of our application, which allows creating functions graphically using modules but also gives the chance to create new modules using C++. We implemented there our planning software, which allows us to send instructions to our robot and to visualize DICOM sets. We can navigate on these 3D sets, know coordinates of desired points, and plan the location of implants. It is worth to mention that taken points, on a DICOM volume, have coordinates respecting to the DICOM reference frame. Additionally, we run on MeVisLab our registration process, which has, as one of the main components, the digitally reconstructed radiography (DRR) module. It generates an X-ray-like image from a DICOM set depending on other parameters like position and orientation of the DICOM respecting to the render element. Those coordinates are crucial for the registration process, and this native MeVisLab module gives us a resolution of 1° in rotational degrees of freedom (DOF) and 1mm in translational DOF. Furthermore, we implemented a module for calculating similarity measurements and another one for carrying out the optimization.

## 2.3 Implemented Similarity Measurement Methods

During the registration process, there is a constant search for finding the DRR image, within a set, that looks the most alike to the X-ray image we want to register. Such a decision-making process requires a measurement that indicates how similar a pair of images are, and that is what the similarity measurements precisely do. They give a value, within a certain range, defined for each specific measurement method, and the larger the value, the more similar both images are. It is worth to clarify that for some of them, the smaller the value, the more similar both images are. Bearing that in mind, we implemented twelve of the most known intensity image measurement in the literature [5][6]. We coded some measurements that only use the intensity values like sum of square differences (SSD), sum of absolute differences (SAD), normalized cross-correlation (NCC), and normalized absolute cross-correlation (NACC) [7]. Also, we incorporate some intensity methods that also use Spatial Information, i.e., pattern intensity (PI), gradient correlation (CR), gradient difference(GD), sum of local normalized correlation (SLNC), and variance-weighted sum of local normalized correlation (VWSLNC) [7]. Finally, we implemented histogram-based methods such as mutual information (MI), normalized mutual information (NMI) and correlation ratio (CR) [9]. Altogether, they are twelve similarity measurements to test as the most reliable for our application.

## 2.4 Implemented Optimizers

Evaluating a set of images indicates which is the best image of the set. After choosing the best image, the following action triggers a movement, which generates images better than the previous set. In other words, we are looking for the maximum (the minimum for SAD and SSD) of the similarity measurement. This process is known in the literature as optimization of a cost function. We implemented two of the most common optimizers in the 2D/3D registration application, i.e., best neighbors (BN) and gradient descent (GRDE) [10].

Additionally, we decided to use three well-known optimizers of the machine learning fields, which are AdaDelta (AD), AdaGrad (AG), and Adams [11]. Moreover, we saw when we started our experiments that despite best neighbors is the slowest of the optimizers, it gives the most accurate results. Therefore, we implemented some combination of optimizers, finishing with gradient descent. That means it was possible that after carrying out a

registration with some optimizers, an additional registration was automatically executed using the result of the first registration as the initial pose for the second one using best neighbors. Acting in that way, we implemented GRDE + BN, AD + BN, AG + BN, and Adams + BN. We also implemented combinations of three optimizers finishing in GRDE + BN. From this cascade of three registrations with different optimizers, we obtained AD + GRDE + BN, AG + GRDE + BN, and Adams + GRDE + BN. All in all, we implemented twelve optimizers to be tested as the most suitable for spine surgery application.

### 2.5 Registration Process

Our registration process was carried out using an iterative process, which is one of the most common procedure for 2D/3D registration [1]. Using an iterative methodology, the more images used for executing the registration, the most accurate the result is, but most radiation the patient must be exposed to [5]. Therefore, we decide to use one AP image and one LAT image for our registration process. The first step for performing a registration is to give the algorithm an initial pose, which can be manually or automatically selected [12] [13]. What is essential for the initial pose is to be no more than ±5° and ±10mm away in each axis respecting to the actual pose [13]. Beforehand, we knew the registration pose of our phantoms, so we could guarantee that our initial poses were in the above mentioned range. As a second step, the registration process starts to evaluate the positive and negative contiguous values of each of the DOF. This evaluation, which is executed using a similarity measurement, takes twelve different measurements, i.e., two by each DOF, one in the positive adjacent value of the current position and the other in the adjacent negative value. In the third step, an optimizer is used to compute the next pose based on the taken set of values, and then another similarity measurement is computed and compared with the value of the previous pose. If the new pose has a better measurement value, the complete step two is repeated using the new pose as starting pose otherwise the previous pose is defined as the optimal value of the cost function and defined as the registration pose.

For finding which optimizer and similarity measurement are most suitable for our application, we took five LAT and five AP X-rays for our phantom. Because we require one AP and one LAT image for the registration, we could make a combination of the AP and LAT images having in total 25 combinations of them. For each of these combinations, we executed the registration process using the combination of the 12 optimizers and the 12 similarity measurement. That means, we ran 144 registrations for each combination of AP and LAT X-rays. All in all, we performed 3600 registrations. Depending on the similarity measurement, an iteration takes from 7s up to 20s. The number of iteration is proportional to the selected optimizer and initial pose. Within our set, BN made an average of 12 iterations per registration, GD 4 iterations, and AD 5 iterations. That means one registration can take as short as 28s, but longer as 240s.

### 2.6 Evaluation of Registration Results

With the help of our RB with pivot and OTS, we selected four landmarks on the phantom and saved the measurements, which are referenced respecting to the OTS reference frame. Then we selected the same four landmarks using our planning software running the DICOM set of the model. These measurements, taken respecting to the DICOM reference frame, were also saved.

A registration pose for our application is giving the transformation matrix from the DICOM to the OTS reference frame. Using this relation, we can transform the four saved measurements of the landmarks from the OTS to the DICOM reference frame and compare the results with the measures taken with the planning software. We computed the RMS error of each of the four points for each performed registration. Here we built a decision table, which is the essential element for any decision-making method [14]. However, the decision tables were created based on the component we wanted to determine, i.e., optimizer or similarity measurement. Then we selected the option that generates the smallest average RMS error.

## 3 Results

### 3.1 Finding Suitable Registration Optimizer

For this selection, we build the decision table with the optimizers as the rows and the columns having the average of the RMS error of each landmark as Table 1 shows. Additionally, Figure 1 shows the total average RMS error of each optimizer. They show that best neighbors give the best average RMS error of about 2.38mm followed by AD+BN and AD+GRDE+BN with 2.55mm, and almost indistinguishable by AG+GRDE+BN and AG+BN with 2.56mm and 2.58mm respectively. From here, we can conclude that the best optimizer for our application will be best neighbors. BN is not the most popular choice in 2D/3D registration because it takes longer to reach the

maximum of the cost function and does not self-adjust the step size. However, it makes a lot of sense in our case because the minimum movement allowed by our DRR module, i.e., other optimizers different from BN will try to obtain a DRR image in between the minimum allowed coordinate, yet they will obtain the same previous image, which leads to improper results.

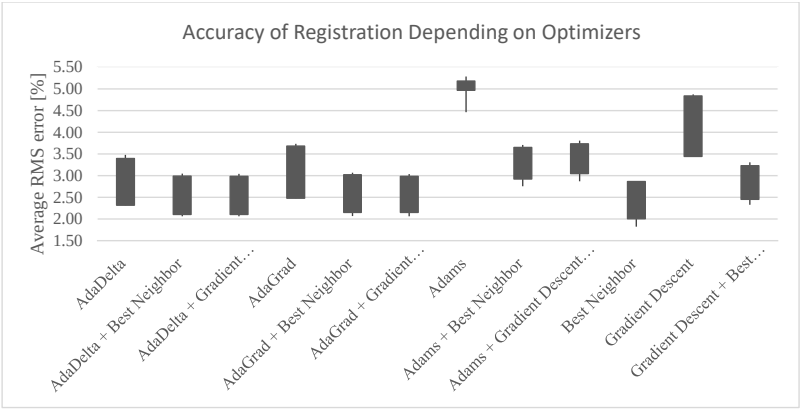



Figure 1*: Average RMS error of each optimizer*

Table 1*: Decision Table for selecting optimizers*

| Optimizer | RMS error Landmak1 [mm] | RMS error Landmak2 [mm] | RMS error Landmak3 [mm] | RMS error Landmak4 [mm] | Averaged RMS error [mm] |
|---|---|---|---|---|---|
| AD | 3.39 | 3.48 | 2.41 | 2.32 | 2.90 |
| AD + BN | 2.99 | 3.05 | 2.07 | 2.11 | 2.55 |
| AD + GRDE + BN | 2.98 | 3.04 | 2.06 | 2.10 | 2.55 |
| AG | 3.68 | 3.74 | 2.63 | 2.48 | 3.13 |
| AG + BN | 3.02 | 3.07 | 2.07 | 2.15 | 2.58 |
| AG + GRDE + BN | 2.99 | 3.04 | 2.06 | 2.15 | 2.56 |
| Adams | 5.18 | 5.28 | 4.46 | 4.96 | 4.97 |
| Adams + BN | 3.65 | 3.71 | 2.76 | 2.92 | 3.26 |
| Adams + GRDE + BN | 3.73 | 3.81 | 2.87 | 3.05 | 3.36 |
| BN | 2.86 | 2.84 | 1.83 | 2.01 | 2.38 |
| GD | 4.84 | 4.87 | 3.71 | 3.44 | 4.22 |
| GD+ BN | 3.23 | 3.31 | 2.33 | 2.46 | 2.83 |

## 3.2 Finding Suitable Similarity Measurement

In this case, we build a decision table with the similarity measurements as the columns, and again, the columns having the average of the RMS error of each landmark as Table 2 shows. However, we did not calculate the averaged error of each landmark over all the optimizers instead used the best five found in section 3.1. At first, we considered to calculate the averaged error using only the results of BN, but it would reduce each average to 25 samples. While using the best five optimizers, we will still have 125 samples per similarity measurement. Considering that the averaged RMS errors of these optimizers are the smallest, we expect the results, of selecting

the similarity measurement with this subset, give us the similarity measurement with the lowest error. Additionally, Figure 2 shows the total average RMS error of each similarity measurement.

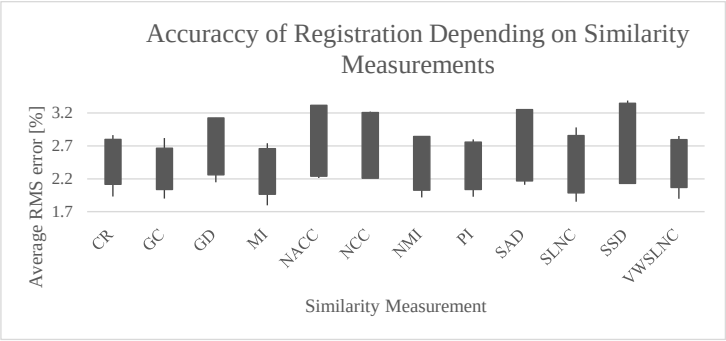



Figure 2*: Average RMS error of each similarity measurement*

Table 2*: Decision Table for selecting Similarity Measurements*

| Similarity Measurement | RMS error Landmak1 [mm] | RMS error Landmak2 [mm] | RMS error Landmak3 [mm] | RMS error Landmak4 [mm] | Averaged RMS error [mm] |
|---|---|---|---|---|---|
| CR | 2.80 | 2.86 | 1.93 | 2.12 | 2.43 |
| GC | 2.66 | 2.82 | 1.90 | 2.04 | 2.36 |
| GD | 3.12 | 3.08 | 2.15 | 2.26 | 2.65 |
| MI | 2.66 | 2.74 | 1.80 | 1.97 | 2.29 |
| NACC | 3.32 | 3.25 | 2.21 | 2.24 | 2.75 |
| NCC | 3.21 | 3.22 | 2.24 | 2.21 | 2.72 |
| NMI | 2.84 | 2.83 | 1.92 | 2.03 | 2.41 |
| PI | 2.76 | 2.80 | 1.93 | 2.04 | 2.38 |
| SAD | 3.25 | 3.26 | 2.11 | 2.17 | 2.70 |
| SLNC | 2.86 | 2.98 | 1.85 | 1.99 | 2.42 |
| SSD | 3.35 | 3.39 | 2.28 | 2.13 | 2.79 |
| VWSLNC | 2.80 | 2.85 | 1.90 | 2.07 | 2.40 |

We can see that MI is the similarity measurement leading to the most accurate registrations. In our test results, we obtained 2.29mm of averaged RMS error followed by GC with 2.36mm. This result was not a surprise because some previous researches have shown the effectivity of these similarity measurements for 2D/3D registration applications [15] [16] [17].

## 4 Discussion and Conclusions

We use an iterative methodology for intraoperative registration of C-arm images with CT Data in computer-assisted spine surgery. Determining the next pose to evaluate requires a DRR module to generate a radiograph in the desired pose from a DICOM volume, a similarity measurement to compare the actual X-ray being registered with the DRR image, and an optimizer. There are several similarity measurement and optimizers in literature, which are most suitable than others depending on the application and used modalities [2][3][4]. We have tested twelve similarity measurements and twelve optimizers with the help of five AP and five LAT images. We have run 3600 registrations using a different combination of pair of AP and LAT images, and the entire combination of

optimizers and similarity measurements in each pair of images. We found the registration error with the help of four landmarks that we measured in both phantom and DICOM volume.

We found that Best Neighbors gives the most accurate results in our application, although other researches do not commonly use it. The resolution of the Mevislab DRR module is 1° for rotations and 1mm for translations. If we had a DRR module with better resolution, we assume that the average RMS error of the entire dataset would improve, and also the selection of the most suitable optimizer for our computer-assisted surgery application. Additionally, Using the results of the best five optimizers, we found that mutual information and gradient correlation gives the most accurate registration errors. This result is not a surprise since other researches are showing good results while using these similarity measurements [15] [16] [17].

## Acknowledgment

Part of this work is funded by the German Federal Ministry of Education and Research (KMU-innovativ: contract number 13GW0175B).

## References

[1] S. Miao, J. Liao, et al. *Toward Accurate and Robust 2-D/3-D Registration of Implant Models to Single-Plane Fluoroscopy, in Augmented Reality Environments for Medical Imaging Computer-Assisted Interventions*, (2013)

[2] J. Pluim, M. Viergever, *Mutual-Information-Based Registration of Medical Images: A Survey*, IEEE Transactions on Medical Imaging, Vol. 22, No. 8, (2003)

[3] Y. Otake, A. Wang, *Robust 3D–2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation*, Phys Med Biold. (2013), doi: 10.1088/0031-9155/58/23/8535

[4] J. Kim, S. Lin, *Comparison of Similarity Measures for Rigid-body CT/Dual X-ray Image Registrations*, Technology in Cancer Research and Treatment, Vol. 6, No. 4, (2007)

[5] G. Penney, J. Weese, et. Al, *A Comparison of Similarity Measures for Use in 2-D–3-D Medical Image Registration*, IEEE Transactions on Medical Imaging, Vol. 17, No. 4 (1998)

[6] D. Russakoff, T. Rohlfing, Evaluation of Intensity-Based 2D-3D Spine Image Registration Using Clinical Gold-Standard Data, Biomedical Image Registration (WBIR 2003)

[7] J. Hajnal, D. Hill, D. Hawkes, *Medical Image Registration*, CRC Press (2001)

[8] D. Knaan, L. Joskowicz, Effective Intensity-Based 2D/3D Rigid Registration Between Fluoroscopic X-Ray and CT, Medical Image Computing and Computer-Assisted Intervention, DOI: 10.1007/978-3-540-39899-8_44, (2003)

[9] P. Markelj, D. Tomazevic, et al, *A Review of 3D/2D Registration Methods for Image-Guided Interventions*, Medical Image Analysis, vol. 16, 642-661 (2012)

[10] D. Russakoff, T. Rohlfing, et al, *Evaluation of Intensity-Based 2D-3D Spine Image Registration using Clinical Gold-Standard Data*, (WBIR 2003)

[11] L. Berrada, A. Zisserman, M. Kumar, Trusting SVM for Piecewise Linear CNNs, arXiv:1611.02185v5 (2016)

[12] A. Varnavas, T. Carrell, G. Penney, *Fully Automated initialisation of 2D-3D Image Registration*, IEEE 10th International Symposium on Biomedical Imaging, San Francisco, CA, (2013), pp. 568-571.

[13] J. Alvarez-Gomez, H. Roth, J. Wahrburg, Intraoperative Registration of 2D C-arm Images with Preoperative CT Data in Computer Assisted Spine Surgery: Motivation to Use Convolutional Neural Networks for Initial Pose Generator, (BMT 2019)

[14] D. Sabaei, J Erkoyuncu, et al, *A Review of Multi-Criteria Decision Making Methods for Enhanced Maintenance Delivery*, Procedia CIRP 37 (2015) 30 -35 (Decision Map)

[15] L. Zollei, E. Grimson, et al, *2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators*, CVPR (2001). (MI)

[16] C. Gendrin, H. Furtado, et al, *Monitoring tumor motion by real time 2D/3D registration during radiotherapy*, Radiotherapy and Oncology 102 (2012) 274–280 (MI)

[17] J. Schmid, C. Chenes, *Segmentation of X-ray Images by 3D-2D Registration based on Multibody Physics*, ACCV (2014), DOI:10.1007/978-3-319-16808-1_45 (GC)

# A Collision-Aware Articulated Statistical Shape Model of the Human Spine

*R. Joachimsky,*[1] *L. Ma,*[2] *C. Icking,*[2] *S. Zachow*[1]

[1] *Zuse Institute Berlin (ZIB), Germany*
[2] *FernUniversität in Hagen, Germany*

Contact: {joachimsky, zachow}@zib.de, {lihong.ma, christian.icking}@fernuni-hagen.de

***Abstract***

*Statistical Shape Models (SSMs) are a proven means for model-based 3D anatomy reconstruction from medical image data. In orthopaedics and biomechanics, SSMs are increasingly employed to individualize measurement data or to create individualized anatomical models to which implants can be adapted to or functional tests can be performed on. For modeling and analysis of articulated structures, so called articulated SSMs (aSSMs) have been developed. However, a missing feature of aSSMs is the consideration of collisions in the course of individual fitting and articulation. The aim of our work was to develop aSSMs that handle collisions between components correctly. That way it becomes possible to adjust shape and articulation in view of a physically and geometrically plausible individualization. To be able to apply collision-aware aSSMs in simulation and optimisation, our approach is based on an efficient collision detection method employing Graphics Processing Units (GPUs).*

**Keywords**: Spine Modeling, Spinal Kinematics, GPU-Accelerated Collision Detection, Sperical Shells

## 1 Motivation

Statistical Shape Models (SSMs) are a well established means of representing a large variety of shapes in a very compact manner, allowing analysis and modeling of variation in shape, e.g. for computer-assisted diagnosis and therapy planning [1,2]. SSMs are also an effective tool for model-based 3D anatomy reconstruction from medical image data [3, 4], since they represent *a priori* statistical knowledge of the shape of an anatomical structure of interest as well as its range of variation. SSM-based anatomy reconstruction typically leads to anatomically plausible results even in cases where measurements are disturbed by low Signal to Noise Ratio (SNR), artefacts, occlusion or any other reason for corruption or partial incompleteness [5–7].

In orthopaedics and biomechanics, SSMs are increasingly employed to individualize measurement data or to create individualized anatomical models to which implants can be adapted to or functional tests can be performed on. Among others, articulated structures, like hip, knee, or shoulder are analyzed in this way with respect to joint kinematics for an early identification of altered conditions [8]. For modeling and analysis of articulated structures, so called articulated SSMs (aSSMs) have been developed [9–11]. Such aSSMs do model ensembles of anatomical structures being coupled via joints. That way the variation in shape of all components is statistically modelled keeping joint centers and/or axes in an anatomically and physiologically consistent manner. Additionally, kinematic degrees of freedom of joint components can be parameterized independently for functional analysis. In this way, articulated structures can be individually analyzed and compared with population data to provide decision support for diagnosis and therapy planning. However, a missing feature of aSSMs is the consideration of collisions in the course of individual fitting and articulation.

The aim of our work is to develop an aSSM of the human spine that handles collisions between components correctly. That way it becomes possible to adjust shape and articulation in view of a physically and geometrically plausible individualization. The clinical application in mind is the adaptation of an aSSm of the human spine to medical image data such as 2D fluoroscopy while preventing its vertebrae from overlapping each other. Klinder et al. [4] proposed a two-stage method to individualize an SSM of the spine according to measured data. At first, a complete spine model is adapted to features in a given medical image in order to represent the spatial relationship in the data set. In a second step, individual SSMs of single vertebrae are adjusted to fit the respective vertebral shape to the image data. With such an approach both, pose and shape of the spine can be considered, although in separate and isolated steps. Boisvert et al. [12, 13] used an equivalent approach to analyze and characterize 3D spine deformities that occur in scoliosis with the help of the resulting principal deformation. Within our work we will demonstrate a novel approach for collision-aware modeling of articulated structures (i.e. statistical 3D shape models of vertebrae) in view of an individualization of the human spine. Our approach is based on an efficient collision detection method employing Graphics Processing Units (GPUs).

# 2 Materials and Methods

Building and employing a collision-aware aSSM of the human spine involves (i) deriving an SSM, respectively many SSMs for every vertebra from medical image data, (ii) defining a suitable joint model for each pair of vertebrae, and (iii) implementing an efficient collision handling approach. A proper handling of collisions will ensure a plausible shape and pose of the complete spine model during articulation. Since collision-aware aSSMs are to be used for model fitting and simulation, the large number of computations involved demands for an efficient collision handling scheme, as in each optimization step the collision status needs to be queried. In the following all three tasks are described in further detail.

## 2.1 Statistical Shape Models for individual Vertebrae

Shape Analysis is applied to Magnetic Resonance Image (MRI) data [14] of the human spine to derive an SSM per vertebra. In our study 31 MRI data sets have been selected for processing. Principal Component Analysis was applied to 3D reconstructed, uniformly parameterized, and aligned geometries of 31 vertebrae to derive a mean shape as well as 30 principal components or shape modes. Mean shape and variation in shape captured in such an SSM is illustrated in fig. 1 for an L1-vertebra.

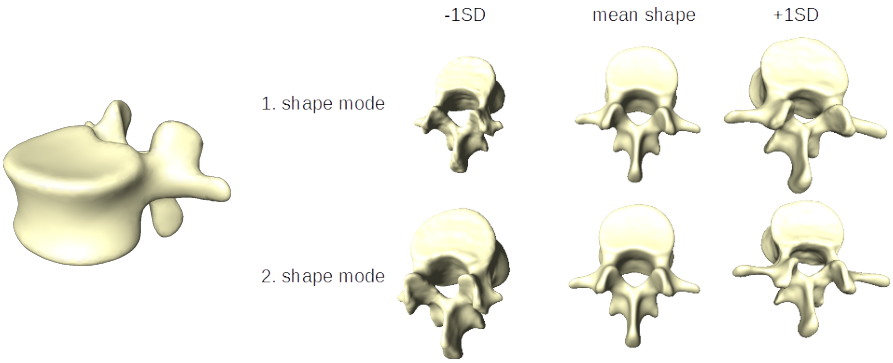



Figure 1: *Left: Mean shape of an L1-vertebra from anterolateral view. Right: Shape variation captured in an SSM of an L1-vertebra. First two modes of variation are illustrated for positive and negative standard deviation around the mean shape.*

## 2.2 Vertebral Articulation Model

Any two adjacent vertebrae in the spine together provide the basic function of motion of the spine and result in a kinematic chain. Two adjacent vertebrae are therefore called a Functional Spinal Unit (FSU). On a closer look, joint function of an FSU is constituted from three independent joints. These three joints include two facet joints and the intervertebral disc (IVD). In facet joints, motion is constrained when inferior and superior articular processes of an FSU exhibit contact. The IVD dominates in contributing to joint function. This is underlined by the fact that the center of rotation (COR) of an FSU is located within an IVD.

The location of the COR is based on values from literature: In biomechanical spine modling it is an established common choice to position the COR on the mid sagittal plane of two vertebrae [15]. On that plane it is typically located mid-disc [16]. Hence, we also chose the mid-dic position as a preset COR in our joint model. Since the IVD itself is not explicitly modeled, the COR location needs to be constructed. Its position is given by the center point of the line between the center of the upper end plate of the lower vertebra and the center of the lower end plate of the upper vertebra (fig. 2, left). A fixed COR is chosen as this puts focus on the collision detection part of the pipeline without loosing generality. Nonetheless, the proposed approach could also be applied to a non-fixed but shifting COR. Our vertebral articulation model is abstracted and simplified to purely rotational Degrees of Freedom (DoF) and does not consider translational DoF (fig. 2, right). However, additional DoF can be easily added on demand.

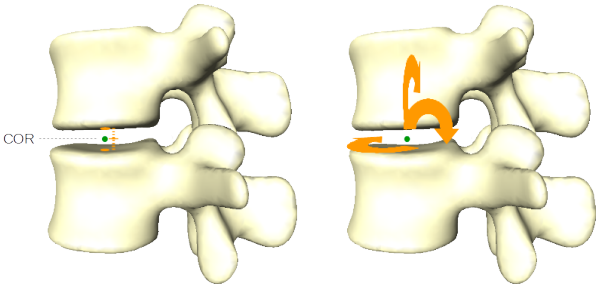



Figure 2: *Left: Mid-disc location of COR is obtained by the center between upper and lower end plate. Right: Joint Model allowing three rotational degrees of freedom.*

### 2.3 Collision Detection

An important question when modeling articulation of a pair of vertebrae is, how an articulation is constrained by contact? To address this question, potentially colliding areas need to be identified during motion and in the case of contact this motion must be stopped or redirected. For triangulated geometric models, triangles have to be tested for collision against each other. In a naive approach all triangles of both models are mutually tested resulting in a complexity of $O(n^2)$ and a large number of unnecessary collision tests. Various methods to detect and to handle collisions have been proposed in literature. These methods introduce spatial data structures to speed up computation. A well known representative is the octree, which hierarchically separates the space into eight disjunct cubes [17]. However, octrees or related approaches do not reflect the special setting and conditions of anatomical joints. Anatomical joints show large irregular formed contact regions in close proximity. This results in a need to traverse the underlying hierarchical data structures to its leafs in a large number of cases. Furthermore, rotation as a major component of motion is not well represented within an octree. These factors limit the practicability of standard approaches.

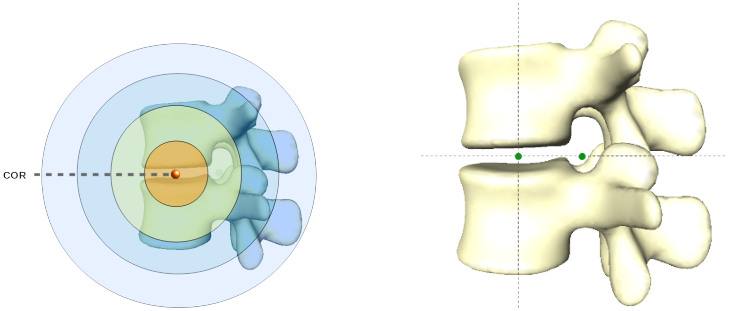
That's why, an approach based on spatial subdivision into multiple spheres is chosen. Within this approach, triangles are clustered based on their location in a set of concentric spheres. This in turn allows classifying candidates for collision testing without the need of updating data structures under rotation. The following steps are involved in the collision detection pipeline: (a) Sampling (b) Spherical Shells (c) Shell Filter (d) Elementary Tests (e) Collision Resolution.

As human joints typically exhibit natural contact areas, these regions need to be preferably checked for collisions, whereas collision tests in other regions are less likely as the constrained motion will prevent any contact. In a pre-processing step, a **sampling** procedure identifies and stores regions that might exhibit contact interaction. Additionally, connected regions that are likely to exhibit contact are determined and matched from sampled contact regions. That way, not all regions prone to exhibit contact are mutually tested for collision, but only regions that really interact during motion will be tested.

Sampling and computation of potential contact regions are followed by a spatial subdivision of the involved FSU to further reduce the search space. For that purpose, Spherical Shells are built around the common COR [18]. **Spherical Shells** are constructed by taking the difference of concentric spheres. These Spherical Shells allow for rotational invariant spatial hashing, if the COR is not translated [19]. With a proper choice of the sphere layout triangles of both upper and lower vertebra of an FSU are assigned to be members of a certain shell. An FSU subdivided by Spherical Shells is illustrated in fig. 3a.

Only triangles of upper and lower vertebra of an FSU that share a common shell have to be checked against each other. Triangles residing on an edge of a shell are tested for all of the shells they are associated with. That way, spatial subdivision is used for pruning unnecessary collision tests. The number of shells is restricted to be a power of two. Hence, this constraint allows to use binary search to assign triangles to their Spherical Shells. In order to build the Spherical Shells, Euclidian distances from the COR to every vertex of both upper and lower vertebra of an FSU are computed.

The **Shell Filter** is an additional pruning step separating each Spherical Shell into eight segments by defining three orthogonal planes (see fig. 3b). For every vertex its position in relation to all planes is classified by evaluating the sign of the inner product. A major advantage is that the Shell Filter is completely independent from Spherical Shell construction. Therefore both steps can be computed in parallel. Furthermore, the Shell Filter has to be updated only, if a vertebra undergoes any change in terms of shape and pose.




(a) *Spherical Shell constructed around the Center of Rotation (CoR) of a Functional Spinal Unit (FSU). Only triangles lying in a common Spherical Shell will be tested against each other.*

(b) *Location of Shell Filter planes. Center of mass of the FSU and midpoint between both end plates are used to define the planes. The third plane is orthogonal to both planes and not illustrated. The planes are chosen to prune a high number of triangles tests.*

Figure 3: *Search space for collision detection is decomposed by (a) Spherical Shells and (b) Shell Filter.*

Spherical Shells and Shell Filter classify vertex candidates according to their location within the Spherical Shells and plane sectors. The classified final subsets of collision candidates are handed over to the elementary primitive tests. These tests yield exact contact information. The well established method proposed by Moeller is used in its optimized version [20]. The steps involved in collision detection are illustrated in fig. 4.

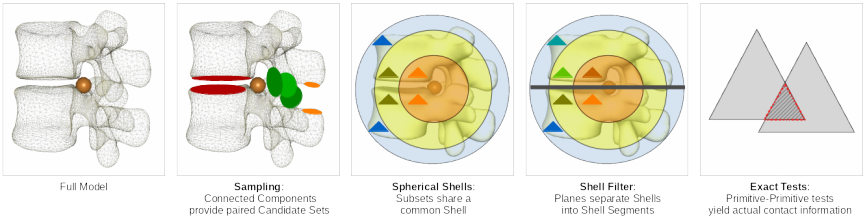



Figure 4: *Steps involved in collision detection pipeline (COR is relocated for a more compact presentation).*

Detected collisions are resolved by constraining further motion and searching the closest position between the current colliding position and the former legal position. In fitting the proposed model to medical image data or other measured data the shape model optimization only requires the information, whether current shape and pose are valid or not. The collision detection is implemented on the GPU using CUDA [21]. Using the GPU throughout the pipeline reduces latency-prone data transfer. An example for an identified collision in an FSU model is illustrated in fig. 5.

## 3 Results

The proposed approach is implemented within the scientific visualization framework Amira ZIB Edition [22]. Mean run time is evaluated for 1000 runs of random shape and pose. Run time evaluation was executed on a system composed of an AMD Ryzen 5 2600x six-core CPU of 4.25 GHz each, 16 Gb of RAM and an NVIDIA GeForce GTX 1050 GPU with 2 Gb of RAM. Evaluated methods include a trivial approach of mutual testing of all triangles on the CPU in a standard (CPU) and parallelized version (PCPU) and on the GPU, and the proposed approach with and without sampling. To examine run time behaviour for different model resolutions three SSMs of different sizes in terms of vertex and triangle count have been created (see fig. 6). Measured run times are given in table 1.

Table 1: *Run time (in ms) of proposed method for three model sizes (see fig. 6).*

| | Model L | Model XL | Model XXL |
|---|---|---|---|
| trivial CPU | 1,257 | 69,799 | 1,108,431 |
| trivial PCPU | 204 | 11,416 | 187,734 |
| trivial GPU | 41 | 1,845 | 26,710 |
| Spherical Shells (16) | 15 | 558 | 6,005 |
| Spherical Shells (16) + Sampling | 3 | 97 | 1,035 |

As expected, run times benefited from parallelized computation. The run time of the trivial CPU-based approach is reduced in parallel CPU and GPU implementation. The Spherical Shell approach further reduced run time, and involving the complete pipeline resulted in 3 ms (L), 97 ms (XL), and 1,025 ms (XXL) run time.

## 4 Discussion

The classification of a vertex regarding the Spherical Shells or planes of the Shell Filter does not influence the classification of other vertices, which supports parallelism and thus fits the design consideration to parallelize computation using the GPU. The proposed approach provides improved run time results even in rather large models. Run times might be further improved considering non-uniform thickness of Spherical Shells. Vertebral models represent rather small anatomies and exhibit planar areas that are very well described by few triangles. The three model sizes are illustrated in their triangular mesh representations in fig. 6. Measured run times suggest that models with a size of about 30,000 triangles can be used with our approach close to real time. That way, shape model optimization could be enhanced by our method without largely increasing performance.

## 5 Conclusion

An efficient collision-aware articulated statistical shape model of an FSU is presented. The model incorporates the two major aspects of variability: Shape of a single vertebra and the relative positioning of neighboring vertebrae. Both aspects have been modeled by an aSSM. Relative positioning of neighboring vertebrae has been constrained by introducing an efficient GPU-based collision detection approach adapted to articulated motion. The SSM is based on thirty-one different spine MR images and thus incorporates a wide variation in shape. From this variability the relationship of shape and range of motion of an FSU can be explored using the developed interactive testing and visualization environment. In future work several aspects could be addressed. Most importantly, the joint model constituting the aSSM will be improved and IVDs and ligaments will be considered within the FSU model allowing for better kinematic simulation. However, the simplified joint model is appropriate in order to focus on the collision detection aspect of the proposed work without loosing generality. Furthermore, focussing on specific medical applications the SSM will be extended to better serve specific pathological use cases like scoliosis.

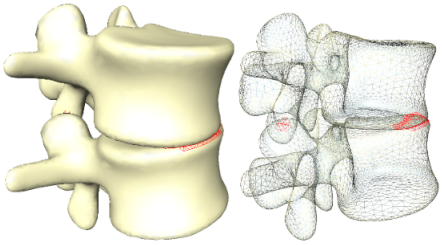
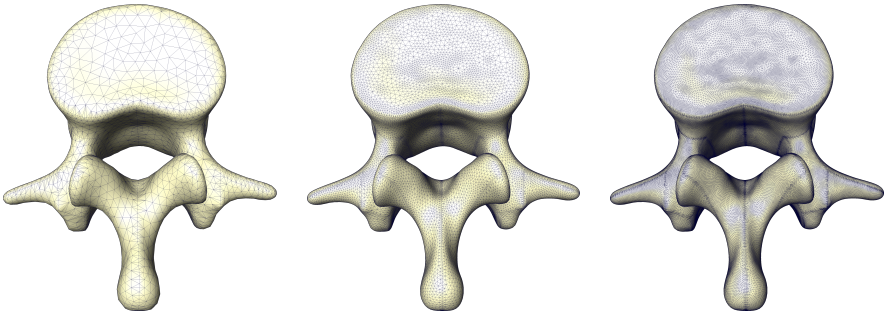


Figure 5: *Colliding regions of an FSU model have been detected and visualized.*



(a) *Model L with ≈ 8k triangles* (b) *Model XL with ≈ 58k triangles* (c) *Model XXL ≈ 232k triangles*

Figure 6: *Vertebral SSMs in three different triangle resolutions. Model L already discretizes the vertebral shape sufficiently well as no geometric irregularities can be observed between the models.*

## 6 Acknowledgements

Parts of this work were funded by the German Federal Ministry of Education and Research (BMBF), grant no. 01EC1406E (TOKMIS). The authors would like to thank Heiko Ramm from 1000shapes GmbH, Berlin, for contributing to the SSM generation. Furthermore, the author would like to thank the incredible team of the Therapy Planning research group at Zuse Institute Berlin, especially Felix Ambellan and Alexander Tack, for valuable discussions, companionship and pleasant work atmosphere.

## References

[1] F. Ambellan, H. Lamecker, Chr. von Tycowicz, S. Zachow: *Statistical Shape Models - Understanding and Mastering Variation in Anatomy.* To appear in Advances in Experimental Medicine and Biology - Biomed. Visualisation (2019)

[2] H. Lamecker, S. Zachow: *Statistical Shape Modeling of Musculoskeletal Structures and Its Applications.* In Comp. Radiology for Orthopaedic Interventions, Lecture Notes in Comp. Vision and Biomechanics, 1–23 (2016)

[3] F. Ambellan, A. Tack, M. Ehlke, S. Zachow: *Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the osteoarthritis initiative.* Medical Image Analysis, 52, 109–118 (2019)

[4] T. Klinder, R. Wolz, C. Lorenz, A. Franz, J. Ostermann: *Spine Segmentation Using Articulated Shape Models.* Medical Image Computing and Computer-Assisted Intervention, Springer LNCS 5241, 227–234 (2008)

[5] F. Bernard, L. Salamanca, J. Thunberg, A. Tack, D. Jentsch, H. Lamecker, S. Zachow, F. Hertel, J. Goncalves, P. Gemmar: *Shape-aware Surface Reconstruction from Sparse 3D Point-Clouds.* Med. Imag. Anal., 38, 77–89 (2017)

[6] I. Vidal-Migallon, H. Ramm, H. Lamecker: *Reconstruction of Partial Liver Shapes based on a Statistical 3D Shape Model.* Proc. of Symposium on Statistical Shape Models & Applications (SHAPE), 2015

[7] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, N. Ayache: *Articulated Spine Models for 3D Reconstruction from Partial Radiographic Data.* IEEE Transactions on Biomedical Engineering 55:11, 2565–2574 (2008)

[8] K. Mills, B. A. Hettinga, M. B. Pohl, R. Ferber: *Between-limb kinematic asymmetry during gait in unilateral and bilateral mild to moderate knee osteoarthritis.* Archives of Phys. Med. and Rehab., 94:11, 2241–2247 (2013)

[9] D. Kainmüller, H. Lamecker, S. Zachow, H.C. Hege: *An Articulated Statistical Shape Model for Accurate Hip Joint Segmentation.* Proc. of IEEE Engineering in Medicine and Biology Conference (EMBC), 6345–6351 (2009)

[10] M. Bindernagel, D Kainmüller, H. Seim, H. Lamecker, S. Zachow, H.C. Hege: *An Articulated Statistical Shape Model of the Human Knee.* Proc. Bildverarb. für die Medizin 2011 (BVM), Informatik aktuell, Springer, 59–63 (2011)

[11] S. Balestra, S. Schumann, J. Heverhagen, L. Nolte, G. Zheng: *Articulated Statistical Shape Model-Based 2D-3D Reconstruction of a Hip Joint.* Information Processing in Computer-Assisted Interventions, Springer LNCS 8498, 128–137 (2014)

[12] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, N. Ayache: *Geometric variability of the scoliotic spine using statistics on articulated shape models.* IEEE Trans. Med. Imaging 27, 557-568 (2008)

[13] J. Boisvert, X. Pennec, H. Labelle, F. Cheriet, N. Ayache: *Principal Spine Shape Deformation Modes Using Riemannian Geometry and Articulated Models.* Articulated Motion and Deformable Objects, Springer LNCS 4069, 346–355 (2006)

[14] SpineWeb, `http://spineweb.digitalimaginggroup.ca`, 10 06 2018.

[15] I. A. Stokes, M. Gardner-Morse: *Lumbar spine maximum efforts and muscle recruitment patterns predicted by a model with multijoint muscles and joints with stiffness.* Journal of biomechanics, 28:2, 173–186 (1995)

[16] A. Abouhossein, B. Weisse S. J. Ferguson: *Quantifying the centre of rotation pattern in a multi-body model of the lumbar spine.* Computer methods in biomechanics and biomedical engineering, 16:2, 1362–1373 (2013)

[17] D. Meagher: *Geometric modeling using octree encoding.* COMPUT VISION GRAPH, 19.2, 129–147 (1982)

[18] S. Krishnan, P. Amol M.C. Lin: *Spherical shell: A higher order bounding volume for fast proximity queries.* (1997)

[19] A. Maciel, R. Boulic, D. Thalmann: *Efficient collision detection within deforming spherical sliding contact.* IEEE Transactions on Visualization and Computer Graphics, 13:3, 518–529 (2007)

[20] T. Möller: *A fast triangle-triangle intersection test.* Journal of Graphics Tools, 2:2, 25–30 (1997)

[21] J. Nickolls, I. Buck, M. Garland, K. Skadron: *Scalable Parallel Programming with CUDA.* Queue 6, 40–53 (2008)

[22] Amira ZIB Edition, `https://amira.zib.de/`, 01 07 19

# Modellbildung, Simulation & Systeme

# Design of a dynamic user interface for IEEE 11073 SDC interoperability testing

*C. Georgi[1], M. Rockstroh[1], S. Franke[1], T.Neumuth[1]*

*[1]Universität Leipzig, Innovation Center Computer Assisted Surgery, Leipzig, Germany*

Contact: christoph.georgi@iccas.de

***Abstract***

The realization of cross-vendor interoperability in the operating room is an enabling technology but also poses challenges in design, development, and workflow integration. The available interactive test tools for interoperable medical devices are of limited use for testing in clinically relevant scenarios. The tools yet require a solid knowledge of the underlying communication and the technical representation of functionalities. When testing the integration in more complex scenarios, the tools fall short in the ability to represent the device's technical environment and in realistic user interaction.

The emerging IEEE 11073 SDC standard family for medical device interoperability was analyzed concerning core requirements for a universal user interface to ease the development and test of new medical devices. Based on the analysis, visualization and interaction strategies are proposed and a dynamic, universal user interface is designed and implemented. When combined with software simulators of device ensembles to integrate with, it may ease interoperability testing processes.

**Keywords**: Medical device interoperability, IEEE 11073 SDC, human-machine interaction

## 1 Problem

Cross-vendor interoperability of medical devices in the operating is considered an enabling technology. Research initiatives, such as MDPnP [1] and OR.NET [2], have proved the feasibility [3] and have demonstrated the potential of integrated operating rooms (ORs) [4]. However, these concepts have not been implemented in the daily surgical routine yet. The cross-vendor interoperability poses new challenges in the design, development, regulatory approval, and workflow integration of medical devices. Besides the core device functionalities, technical interfaces need to be developed and tested. New strategies for human-device interaction and integration into surgical workflows need to be designed as well. The overall aim is to enable the dynamic combination of medical devices and clinical IT systems regardless of vendor-specific properties. This plug-and-play functionality, which is a major advantage of openly integrated OR systems, results in communication partners being partly unknown at the design time of medical devices. A large set of devices in different configurations needs to be considered during development. This is an enormous challenge for the vendors and might impede the implementation of the interoperability concepts. Hence, new interactive tools are needed to facilitate the development and integration process.

Cross-vendor connectivity based on open standards, especially the IEEE 11073 SDC standards family, require new strategies for testing. Hence, a three-step testing, which consists of testing for conformity, interoperability, and integration, was proposed [5]. Whilst conformity is mainly focused on testing the fulfillment of requirements posed by the underlying communication standards, interoperability and integration require complex scenarios and multiple device ensembles that can serve as technical ecosystems for testing. Especially small- and medium-sized enterprises (SMEs) may use software simulations of devices to cover the large variety of communication partners for their devices. However, the user interaction is then not covered by physical user interfaces but needs to be provided by software tools. So far, central control interfaces [7] and workstations for the OR [8], [9] were designed with a specific focus on the safe application during surgery. In this context, *User Interface Profiles* have been developed to address safety and usability in the OR by means of the provision of device-wise additional constraints for user interfaces. However, the indispensable risk management for clinical use limits the capabilities for interacting with novel devices that are still under development.

The yet available test tools that we are aware of, for example the *DPWS Explorer*[1] or the *Device Modeler* and *Swiss Army Knife* [6], have a strong technical focus. They are designed for an in depth analysis of the communication of a message-to-message basis. Hence, these tools are of limited use for testing the integration on complex, clinically relevant scenarios. The available development tools yet require a solid knowledge of the underlying integration principles and the technical modeling of the device functionalities. When testing the integration of a new device into changing device ensembles, the tools fall short in the ability to represent the whole device network for realistic user interactions. Thus, a network-based, dynamic graphical user interface is needed to interact with new devices under development as well as with eventually simulated devices in varying integration settings and test scenarios.

# 2 Material and Methods

We aim to develop a dynamic user interface, *UcSDC* (*Universal consumer for SDC*), for scenario-based integration testing of new, interoperable medical devices. The emerging IEEE 11073 SDC standard family for medical device interoperability provides the foundational OR integration capabilities. Especially the machine-interpretable description of device capabilities serves as the starting point for the run-time user interface generation. We performed an analysis of the SDC communication capabilities which revealed the core requirements for an universal user interface for interoperability testing. Based on the analysis, visualization strategies are proposed.

## 2.1 IEEE 11073 SDC for interoperability of medical devices

This section provides a brief summary of the concepts of the IEEE 11073 SDC standard family. In the context of this work, we focus on the machine-interpretable description of device capabilities. Details on the technical aspects of device communication via IEEE 11073 SDC can be found in [10]. *Medical Devices Profile for Web Services* (MDPWS; IEEE 11073-20702) extends *Devices Profile for Web Services* (DPWS) and enables basic data transmission. It provides discovery, messaging, basic device description, and eventing for asynchronous notification. For medical applications, dual channel transmission and capabilities to define required safety-relevant contextual information for remote control requests were defined among others [11]. The general service-oriented architecture (SOA), in IEEE 11073-20701 refined to a *service-oriented medical device architecture* (SOMDA) [10], distinguishes between service providers and service consumers. The providers offer information and functionalities in the SDC network whereas service consumers use the information or orchestrate the available services for clinical purposes. Medical devices usually act primarily as service providers sharing for example measurements, parameters, or control functionalities. However, in complex, realistic integration scenarios, the might also consume services to establish a cooperative behavior. To be able to do so in dynamic, cross-vendor scenarios, consumers rely on the machine-interpretable description of the providers. The description of devices is addressed in the *Domain Information and Service Model* (DIM; IEEE11073-10207) which is derived from the IEEE 11073-10201 DIM. The *Medical Device Information Base* (MDIB) combines a rather static description of the device capabilities (*Device Description*) with a dynamic representation of its states at a given point in time (*Device State*). Described entities can be referenced by a *handle*, which is unique in a *Device Description*.

A run-time creation of a universal user interface uses the device description for the structuring while the device state is used for the actual data content. The basic structure of a device description is a four-layer tree hierarchy. The *Medical Device System* (MDS) represents the overall device and may contain multiple *Virtual Medical Devices* (VMDs), which represent subsystems. In each VMD, *Channels* are logically grouping *Metrics*, which contain the actual information, for example (physiological) measurements, calculations, or settings. On all four layers of structuring from MDS to *Metrics*, a safety classification is provided. The standard supports four levels (IEC 62304): *Informational*, *Class A* (no injury or damage to health possible), *Class B* (nonserious injury possible), or *Class C* (death or serious injury possible). In addition to metrics and operations, medical devices may implement various types of contexts. These represent the overall operational context of the device (patient, location, workflow, etc.) and are thus associated directly with the MDS. The context information are represented in complex objects with multiple attributes. By means of that, order data may be distributed among the medical devices [12]. The IEEE 11073 SDC standard also defines several additional components, such as an alert system, or batteries. However, they are yet not considered in user interface conceptualization.

## 2.2 Technical scenario and application context

As a test tool, the intended user interface needs to operate with any integrated device ensemble. It acts as a service consumer within a network of medical devices acting as providers. These providers can be unknown at design

[1] http://ws4d.org/dpws-explorer/

time and dynamically join and leave the network at run time. The user interface cannot rely on predefined layouts for specific device classes but needs to be able to generate them from the providers' device descriptions. The *UcSDC* may be used a as test consumer for a single device under development. However, the full advantage may be drawn in scenario-based testing of interoperability with other devices included. The communication partners may be simulators or real devices depending on the test setup. Hence, the integration of the user interface into a test platform with additional simulation and analysis capabilities should be supported.

Technically, the *UcSDC* repeatedly performs a discovery of available devices (providers) and is as well reacting to *Hello* and *Bye* message (IEEE 11073-20702). Whenever a new device has been discovered, the MDIB is retrieved and events are subscribed. Afterwards, the user interface will be automatically generated based on the device description and updated according to asynchronous events on changes of the device state.

## 2.3 Structured presentation of information

The information provided by the medical devices need to be displayed in a structured manner. The IEEE 11073 standards are very limited in their capabilities to express semantic relationships between information entities. Cross-device relationships cannot be described and semantics is provided entity-based using IEEE 11073-1010X term codes [13]. Hence, the structuring of information can only rely on the four-layered structure of the device description. On the top level, a list of devices is provided where each device is equivalent to an MDS. If there is more than one subsystem to the device, multiple VMDs are included in the device description. Since these subsystems may be removed or recombined during run time, their structure need to be reflected in the user interface as well. Additionally, the VMDs may serve as a scope for operations and alerts. To be able to communicate this scope to the user, considering the VMDs as structuring elements in the user interface is inevitable in case of multiple VMDs being present. We propose the use of tabs as user interface elements for VMDs to preserve clarity, since there might be a lot of subsystems combined to one MDS.

The information entities (metrics) are logically grouped with channels. Using these in interfaces to group information entities is obvious; however, channels have a very limited amount of properties besides a semantic term code (IEEE 11073-1010X) which can be used to create a label for the grouping user interface element (*LogicGroup*). The safety classification of the channel may a valuable information in some interoperability testing scenarios. Therefore, it is unobtrusively represented by different background colors for the label of the *LogicGroup*. For the metrics within the group, a label is generated from the term code as well. Additionally, the type of metric (setting, calculation, measurement, etc.) can be relevant during test scenarios. We propose the use of icons to graphically represent the type. The actual value, and, if present, a unit, are displayed based on the current device state. In IEEE 11073 SDC, the state of a metric includes a validity of the metric value. Values that are marked invalid by the provider, should be displayed differently, for instance grayed-out.

In addition, context information (patient, location, workflow, etc.) is usually set only once at the beginning of a surgery. Hence, we proposed an additional user interface element on the level of VMDs (tab) to represent contexts, if present. The structure and at least partly the semantics of the data are defined in IEEE 11073 SDC, which allows to provide specifically designed user interface elements to represent the context information. The same strategy could be applied to additional elements, for instance the representation of batteries, with specified properties.

## 2.4 Interactions

There are various ways of interaction with a medical device that need to be represented in the user interface. Although the provision of control functionalities in not mandatory, considering the use case of testing in integration scenarios, one can expect that most devices will also provide some sort of remote control. We consider three basic types of interaction: set a parameter, invoke an operation, and manipulate an alert presents or limit condition.

Medical devices may provide set operations for metrics, which allow consumers to change the value of a metric remotely. String metrics are the most general type of metrics providing no explicit limitations on the sequence of characters or range. Thus, the user interface represents those metrics with a text entry and an additional button to actually trigger the set operation. For enumeration string metrics, a set of allowed values is provided in the device description. Depending of the number of allowed values, either a set of exclusive radio buttons or a drop-down selection may be used. The radio buttons provide a direct overview on the available options and allow for a one-click setting. However, they are space-consuming; hence a drop-down selection with a cancel option is the preferred user interface element for more than three options. Metrics with numeric values should be handled according to their properties. If a technical or physical range and a resolution are provided in the device description then a slider would be appropriate otherwise a text entry for numeric values must be used.

Another common form of interaction with medical devices is the invocation of operations. In IEEE 11073 SDC, these activate operations are primarily designed to represent simple triggers to change a device state. They may be

associated with either the MDS or a VMD. Their semantics are represented with IEEE 11073-1010X term codes and an optional target. The target may reference channels and metrics among others. Hence, the operation target should determine the positioning of the corresponding button. The operation is either displayed on top of the *LogicGroup* in case of a channel as operation target or next to the targeted metric. By means of that, the operation is implicitly communicated to the user. Increase and decrease operations can be identified by their term codes and should be handled by the user interface explicitly. *UcSDC* automatically creates specific buttons left and right to the metric value accordingly. For operations without a given target, an additional tab is added. Activate operations may have additional parameters, which may be set by the user through a popup where the presentation of parameters follows the presentation strategies of metrics.

### 2.5 Semantics and custom visualizations

The semantics of the described provider capabilities (metrics, units, operations, etc.) heavily rely on the nomenclature definitions (IEEE 11073-1010X) [10]. Their technical representation (usually numeric codes) may be transformed to a human-readable label using the *Rosetta Terminology Mapping*[2] provided online by the *National Institute of Standards and Technology*. The new application of the nomenclature to surgery, however, demands extensions since needed term codes are yet not defined [13]. Developers and vendors may use non-standardized term codes especially in the development of novel devices where standardization is yet not done. Considering the use case of testing, there may also be no internet connection available or desired for the testing network. Hence, we prefer the use of a local mapping from term codes to human-readable labels, which includes the standardized terms and may easily be extended by the developers according to their needs. Nonetheless, the mapping may be incomplete and terms may be missing during run-time. In those cases, displaying the term code and/or the *handle* is used as a fallback strategy. In addition, a term code may occur for multiple metrics which in turn may lead to false conclusions of the user. Note that the *handle* is not necessarily useful because the actual string should be meaningless according to the IEEE 11073 SDC.

Besides semantics issues, the automatic generation of an user interface based on the medical device description suffers from several other limitations. The most striking limitation is the lack of ordering information for the user interface elements. While ordering is useless in machine-machine communication, it is essential to human-machine interaction. A suitable ordering may be derived from two major factors: semantic linkage between elements and frequency of use. However, these data is not available. To overcome the limitations and increase the flexibility, we introduced optional custom visualization hints. They are provided device-wise structured in XML and mainly allow to regroup, to reorder, and to rename metrics and operations. The hints also enable multiple visualizations of the same metric, alternative labels, precisions, and units, as well as new associations of operations to channels and metrics. Additionally, *UcSDC* supports the visualization of a numeric metric, which is usually a measurement, as a graph. This eases the analysis of time-dependent properties and trends of metrics in test scenarios.

## 3 Results

We are implementing a prototype using a web-based user interface, which enables access from any device in the corresponding network, including mobile devices (see Figure 1).

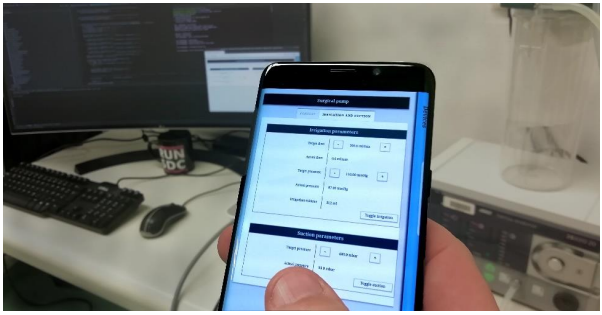


Figure 1*: Photograph of the UcSDC web-based user interface on a mobile device controlling a surgical pump in a test setting.*

[2] https://rtmms.nist.gov/rtmms/

The SDC communication was implemented using the open source library *SDCLib/C*[3] provided by *SurgiTAIX AG*. The medical devices found in the network are listed on a toggleable overlay to the right of the user interface. The layout of the user interface is adapted to the available display size mainly by scaling and rearranging the *LogicGroups* in the selected VMD tab. The way of presentation of a medical device can be switched between automatically generated layout and custom visualization. The current state of implementation was technically tested with a large variety of simulated and real medical devices.


Figure 2*: Screenshots of the UcSDC user interface of a surgical fraise with an irrigation pump on a desktop system using the automatic user interface generation (left) and the custom visualization hints (right).*

Example interfaces are shown in Figure 2 for a surgical fraise which includes an irrigation pump. The user interface is reorganized by custom visualization hints (Figure 2, right). For instance, the technical maximum of the motor revolution speed is integrated into the user selected speed limit setting. The actual revolution speed of the motor is displayed over time in an interactive diagram below the default metric display.

## 4 Discussion

The IEEE 11073 SDC standard provides complex mechanisms for the description of device capabilities, including term codes for semantics, grouping in VMDs and channels, and safety classifications. Nonetheless, the automatic generation of user interfaces from these descriptions can only achieve limited usability. For instance, non-standardized term codes are common for devices with novel functions in development but these codes cannot be mapped to human-readable labels automatically. Another aspect that impairs usability is the limited capabilities of grouping and ordering information entities. For intraoperative surgical workstations, this has already been discussed in the context of user interface profiles [14]. Also, there is ongoing research concerning the definition of key purposes and of device specializations (device classes) [15], which is yet not considered but may improve the automatic user interface generation.

For scenario-based interoperability testing, the introduced custom visualization capabilities of *UcSDC* address the limitations of automatically generated interfaces. However, they could furthermore be specifically adapted to the intended test scenario. Metrics and operations may be grouped and ordered across devices based on the prior knowledge of the planned scenarios. In this case, scenarios would provide the custom visualization hints for functional grouping instead of those potentially provided by the involved devices and/or simulators, which may result in an improved usability.

Currently, *UcSDC* lacks some features that are particularly useful for scenario-based interoperability testing. The IEEE 11073 SDC standard defines several additional components, such as an alert system, or batteries. However, they are yet not included in the implementation. Finally, a logging should be integrated to support the documentation and reproducibility of scenarios and tests. For this purpose, the *UcSDC* should be able to persist all technical data, such as metric changes or alert presence, as well as any user interaction performed. The yet available tools provide a logging on the message content level, which is most relevant for the development of communication stacks. However, this work is concerned with later device development phases. One can assume that the first level of testing (conformity) is already performed. Hence, the logging of the *UcSDC* should focus on higher level properties and processes. With these features added, the *UcSDC* will be a valuable tool for scenario-based testing of interoperability of medical devices. It might be used stand-alone or integrated into an overarching test platform.

[3] https://github.com/surgitaix/sdclib

# 5 Conclusion

The presented dynamic user interface is designed to ease the interoperability testing of networked medical devices. It is suitable for later development phases, especially for testing the integration into medical device ecosystems. The design decisions that have been made are a compromise between intuitive, efficient, and safe user interaction and the flexibility required for use with devices that are still under development. When combined with software simulators of the targeted medical device ensembles to integrate with, it may ease the interoperability development and testing processes.

# Acknowledgement

This work was partly funded by the German Ministry of Research and Education (BMBF), in the project MoVE (03VNE1036).

# References

[1] D. Arney, J. Plourde, and J. M. Goldman, "OpenICE medical device interoperability platform overview and requirement analysis," *Biomedical Engineering / Biomedizinische Technik*, vol. 63, no. 1, pp. 39–47, Feb. 2018.

[2] M. Czaplik, V. Voigt, H. Kenngott, H. Clusmann, R. Hoffmann, and A. Will, "Why OR.NET? Requirements and perspectives from a medical user's, clinical operator's and device manufacturer's points of view," *Biomedical Engineering / Biomedizinische Technik*, vol. 63, no. 1, pp. 5–10, Feb. 2018.

[3] M. Rockstroh *et al.*, "From SOMDA to application – integration strategies in the OR.NET demonstration sites," *Biomedical Engineering / Biomedizinische Technik*, vol. 63, no. 1, pp. 69–80, Feb. 2018.

[4] M. Rockstroh *et al.*, "OR.NET: Multi-perspective qualitative evaluation of an integrated operating room based on IEEE 11073 SDC," *International Journal of Computer Assisted Radiology and Surgery*, May 2017.

[5] A. Janß *et al.*, "Extended device profiles and testing procedures for the approval process of integrated medical devices using the IEEE 11073 communication standard," *Biomedical Engineering / Biomedizinische Technik*, vol. 63, no. 1, pp. 95–103, Feb. 2018.

[6] M. Leucker, M. Schmitz, and D. à Tellinghusen, "Runtime Verification for Interconnected Medical Devices," in *Leveraging Applications of Formal Methods, Verification and Validation: Discussion, Dissemination, Applications*, vol. 9953, T. Margaria and B. Steffen, Eds. Cham: Springer International Publishing, 2016, pp. 380–387.

[7] S. Franke, M. Rockstroh, and T. Neumuth, "Context-awareness for control consoles in integrated operating rooms," *Current Directions in Biomedical Engineering*, vol. 4, no. 1, pp. 291–295, Sep. 2018.

[8] M. Köny *et al.*, "Getting Anesthesia Online: The smartOR Network," *International Journal on Advances in Internet Technology Volume 5, Number 3 & 4, 2012*, 2012.

[9] J. Benzko *et al.*, "Modular user interface design for integrated surgical workplaces," *Biomedical Engineering / Biomedizinische Technik*, vol. 61, no. 2, Jan. 2016.

[10] M. Kasparick *et al.*, "OR.NET: a service-oriented architecture for safe and dynamic medical device interoperability," *bmte*, vol. 63, no. 1, p. 11, 2018.

[11] M. Kasparick, S. Schlichting, F. Golatowski, and D. Timmermann, "New IEEE 11073 standards for interoperable, networked point-of-care Medical Devices," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 1721–1724.

[12] B. Andersen *et al.*, "Connecting the clinical IT infrastructure to a service-oriented architecture of medical devices," *Biomedical Engineering / Biomedizinische Technik*, vol. 63, no. 1, pp. 57–68, Feb. 2018.

[13] B. Andersen, M. Kasparick, F. Golatowski, and J. Ingenerf, "Extending the IEEE 11073-1010X nomenclature for the modelling of surgical devices," in *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, 2016, pp. 244–247.

[14] A. Janß, J. Benzko, P. Merz, J. Dell'Anna, and K. Radermacher, "Development of Medical Device UI-Profiles for Reliable and Safe Human-Machine-Interaction in the Integrated Operating Room of the Future," in *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics 2014*, Krakow, 2014, pp. 1855–1860.

[15] B. Andersen *et al.*, "Point-of-care medical devices and systems interoperability: A mapping of ICE and FHIR," presented at the IEEE Conference on Standards for Communications and Networking (CSCN 2016), 2016, pp. 1–5.

# Systemmodellierung eines Homemonitoringsystems für Herzschrittmacher

*Lena Essig*[1], *Christian Kücherer*[1]

[1] *Hochschule Reutlingen, Fakultät Informatik, Reutlingen, Deutschland*

Kontakt: lena.essig@student.reutlingen-university.de — christian.kuecherer@reutlingen-university.de

***Abstract***

*Telemetrie und Homemonitoring werden bereits in vielen Ge-sundheitsbereichen erfolgreich genutzt. Moderne Herzschrittmacher ermöglichen durch telemetrische Datenübertragung das Homemonitoring aktueller Gesundheits- und Zustandsdaten durch PatientInnen und ÄrztInnen. Für die Weiterentwicklung existierenden Produkte ist ein grundlegendes Verständnis der Anforderungen an und des Aufbaus solcher Systeme notwendig. Bisher existieren herstellerunabhängigen Betrachtungen dieser noch nicht. Durch die Verwendung von SysML als semiformale Notationssprache wird das System Herzschrittmacher und Homemonitoring modelliert. Die Anforderungen an ein solches System lassen sich aus bestehenden Produkten ableiten. Die vorliegende Arbeit beschreibt die Systemarchitektur solcher Systeme, anhand derer die Anbindung an Informationssysteme über das Homemonitoringsystem und die dadurch umgesetzten Funktionen gezeigt werden.*

**Keywords**: Herzschrittmacher, Telemetrie, Homemonitoring, Systemmodellierung, SysML, SYSMOD

## 1 Problemstellung

Nach der Implantation eines Herzschrittmachers müssen regelmäßige Kontrolluntersuchungen bei ÄrtzInnen durchgeführt werden, um Daten aus dem Gerät auszulesen, um beispielsweise den Batteriestand oder die korrekte Funktionsausführung zu prüfen. Das Auslesen der Daten erfolgt drahtlos über eine interne Telemetrieantenne, welche am implantierten Herzschrittmacher befestigt ist. Seit der Entwicklung von Homemonitoringsystemen in den 2000'er Jahren ist eine transtelephonische Übertragung der Daten möglich, wodurch Kontrollen in der Klinik reduziert werden können. Bisher existieren keine Beschreibungen des grundlegenden Aufbaus solcher Systeme in semiformaler Notation wie der Systems Modeling Language (SysML) der OMG [14] und einer systematischen Beschreibung der Anforderungen an solche Systeme. In dieser Arbeit ist der Aufbau von Herzschrittmachern und Homemonitoringsystemen mit Hilfe von SysML modelliert, wodurch ein grundlegendes Verständnis des Systemaufbaus vermittelt wird, und für Weiterentwicklungen hinsichtlich Funktionalität, Sicherheit und Datenschutz genutzt werden kann.

## 2 Material und Methoden

### 2.1 Herzschrittmacher und Homemonitoringsysteme

Die Ursachen für Implantationen eines Herzschrittmachers sind vielfältig, und umfassen beispielsweise Bradykardien mit der Gefahr von Schwindel und Ohnmacht, sowie Fehlfunktionen des Sinusknoten mit der Folge von Herzrhythmusstörungen [5]. Abhängig von der Art der Herzerkrankung besitzt ein Herzschrittmacher eine bis drei Sonden, welche in den Herzkammern und -vorhöfen positioniert werden. Mithilfe der Sonden werden die elektrischen Herzaktivitäten abgetastet (Sensing), und bei Bedarf wird durch sie eine elektrische Stimulation an den Herzmuskel abgegeben. Ein Herzschrittmacher verfügt über einen internen Speicher, in dem Informationen über Implantataktivitäten wie abgegeben Impulse oder auffällige System- und Herzaktivitäten gespeichert werden [3]. Bei regelmäßigen Kontrollen wird dieser Speicher mithilfe der im Herzschrittmacher integrierten Telemetrieantenne ausgelesen, und durch ÄrztInnen ausgewertet. Die Übertragung der Daten erfolgt dabei mittels Funkverbindung [10]. Während Kontrollterminen bei ÄrztInnen können notwendige Umprogrammierungen der Herzschrittmacherfunktionen durchgeführt werden. Homemonitoringsysteme ermöglichen einen Datenaustausch ohne persönliche Vorstellung bei ÄrztInnen und dem Zugriff auf die Daten durch die PatientInnen selbst.

### 2.2 SysML und SYSMOD Prozess

Die Anforderungen und Systemmodelle wurden dem durch Weilkiens beschriebene SYSMOD-Prozess [18] erstellt. Dieser Prozess gibt einen Rahmen vor, um von Systemzielen über Anforderungen ein Systemdesign zu

erstellen. Die Systemmodellierung wurde mit SysML durchgeführt, welche eine auf der Unified Modeling Language (UML) basierte standardisierte Modellierungssprache ist, und diese durch Sprachelemente erweitert. Das in dieser Arbeit erstellte Systemmodell wurde auf Grundlage von Funktionalitäten bestehender Systeme erstellt.

# 3 Ergebnisse

## 3.1 Systemziele

Die Idee des Systems ist es, Daten eines implantierten Herzschrittmachers in ein Informationssystem zu übertragen, um sie von dort einsehen und analysieren zu können. Für PatientInnen soll es möglich sein die Daten außerhalb einer Klinik an das Informationssystem zu übertragen. Tabelle 1 zeigt die Ziele des Systems. Dazu gehört, dass PatientInnen entlastet werden, da sie seltener zu Routineuntersuchungen ins Krankenhaus müssen und somit Anfahrtszeiten und Aufenthalte im Krankenhaus reduziert werden [16]. Obwohl sich PatientInnen seltener in einem Krankenhaus persönlich vorstellen müssen (Z1), können ihnen durch die manuellen oder automatischen Datenübertragungen ein Gefühl von Sicherheit (Z2) vermittelt werden, da ÄrztInnen automatisch über Auffälligkeiten und Systemstörungen benachrichtigt werden [9]. Gleichzeitig sind PatientInnen durch die ortsunabhängige Datenübertragung unabhängiger und ungebundener (Z3), wodurch auch spontanes Verreisen möglich ist [1]. Auch für ÄrztInnen bietet das System eine Erleichterung, da die Kontrollen der Herzschrittmacherdaten schneller gehen (Z4) [16], und sie unabhängiger sind, da sie an weniger Termine gebunden sind, und eine Kontrolle der Daten von zu Hause oder unterwegs möglich ist (Z5) [17].

Tabelle 1: *Systemziele von Homemonitoringsysteme für Herzschrittmacher*

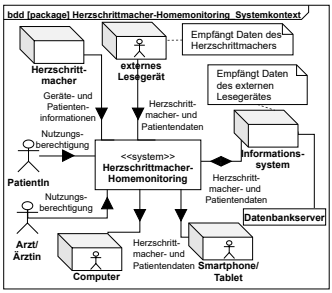
| ID | Systemziel | Beschreibung |
|---|---|---|
| Z1 | Entlastung von PatientInnen | Das System schickt Herzschrittmacherdaten an ÄrztInnen, ohne dass PatientIn vor Ort sein muss. Dies führt zu weniger Klinikbesuchen. |
| Z2 | Sicherheit von PatientInnen | Regelmäßige und automatische Übertragung der Daten und Benachrichtigung der ÄrtzInnen bei Systemstörungen ermöglicht eine schnelle Störungsbeseitigung. |
| Z3 | Unabhängigkeit von PatientInnen | PatientInnen wird durch das System eine ortsunabhängige Kontrolle der Herzschrittmacherdaten geboten, weshalb sie insgesamt unabhängiger sind. |
| Z4 | Zeitaktuelle Kontrolle der Herzschrittmacher | Das System ermöglicht ÄrztInnen eine schnelle Kontrolle der Herzschrittmacherdaten und daher schnelles Handeln bei aktuellen Problemen. |
| Z5 | Unabhängigkeit und Entlastung von ÄrztInnen | Die terminunabhängige Kontrolle der Herzschrittmacherdaten ermöglicht ÄrztInnen eine terminungebundene Befundung. |

## 3.2 Basisarchitektur

Bei der Betrachtung von aktuell am Markt befindlichen Produkten, welche die in Tab. 1 beschriebenen Systemziele umsetzen, kann folgender prinzipieller Systemaufbau abgeleitet werden. Die Basisarchitektur ist nach SYSMOD ein eigenes Diagramm, allerdings aus Platzgründen nun in Kombination mit dem Systemkontext dargestellt. Das Herzschrittmacher-Homemonitoringsystem lässt sich durch die in Abb. 1a gezeigte Basisarchitektur abstrahieren. Hauptbestandteile des Systems sind neben dem Herzschrittmacher mit integrierter Telemetrieantenne ein externes Lesegerät, ein Informationssystem und eine mobile Anwendung oder eine Webanwendung. Ein externes Lesegerät ermöglicht im Homemonitoringsystem eine Kommunikation zum unidirektionalen Datenaustausch mit der Telemetrieantenne mittels Funkverbindung, wobei im klinischen Bereich eine bidirektionale Verbindung aufgebaut werden kann. Für medizinische Implantate wird dabei der Frequenzbereich von 401 bis 406 MHz genutzt, welches in der Medical Implant Communication Service (MISC) Standard [7] festgelegt ist. Die Daten des Herzschrittmachers werden dann mithilfe des Mobilfunks über festgelegte Provider [1, 3] an ein herstellerspezifisches Informationssystem übertragen. Die dort gespeicherten Daten können über eine mobile Anwendung oder eine Webanwendung abgerufen werden.

## 3.3 Stakeholder

Stakeholder sind Personen, Rollen oder Organisationen, die sowohl das Systemergebnis als auch den Projektverlauf bei der Entwicklung eines Produkts beeinflussen können [15]. Für das System lassen sich die in Tab. 1b gezeigten Stakeholder identifizieren. Sowohl PatientInnen (SH1) als auch ÄrztInnen (SH2) wollen das System




(a) Basisarchitektur und Systemkontext

| ID | Rolle | Beschreibung |
|---|---|---|
| SH1 | PatientIn | NutzerIn des Systems. Interessiert an verständlicher Darstellung der Gesundheitsdaten. |
| SH2 | Arzt/ Ärztin | NutzerIn des Systems. Möchte aktuelle Daten im Detail einsehen und den Verlauf kontrollieren. |
| SH3 | IT Klinikmit-arbeiterIn | Inbetriebnahme, Konfiguration und Wartung des Systems. |
| SH4 | Herzschritt-macher-hersteller | Auftraggeber des Systems. Stellt Schnittstellenbeschreibung zur Verfügung, ist für administrative Prozesse verantwortlich. |

(b) Stakeholder

Abbildung 1: Basisarchitektur und Stakeholder des Herzschrittmacher-Homemonitoringsystem

zur Überwachung und Auswertung der Gesundheitsdaten verwenden. Beide stehen im direkten Bezug zum System und sind daher Akteure des Systems. Die Stakeholder IT-KlinikmitarbeiterIn (SH3) sind mit der Wartung des Systems betraut und dafür verantwortlich, dass es zu keinen Verbindungsunterbrüchen kommt bzw. diese zu beseitigen. MitarbeiterInnen von Herzschrittmacherherstellern (SH4) haben ebenfalls die Aufgabe technische Probleme zu beseitigen, die Datenübertragungen beeinflussen oder verhindern. Diese Stakeholder sind ausschlaggebend für die Definition und Umsetzung von technischen Anforderungen. Sie werden im Weiteren nicht näher betrachtet.

### 3.4 Anforderungen

Dieser Abschnitt zeigt die aus der Literatur und den Produkten MyCareLink Patientenmonitor (Medtronic) sowie CardioMessenger (BIOTRONIK) extrahierten Anforderungen. **Funktionale Anforderungen** beschreiben Funktionsmerkmale des Systems [15]. Die durch die Akteure *PatientIn* und *Arzt/ Ärztin* definierten Anforderungen sind in Abb. 2 dargestellt. Die telemetrische Datenübertragung durch das System ist der Hauptvorteil im Gegensatz zur konventionellen Auswertung der Daten während einer ärztlichen Sprechstunde. Daher wird eine orts- und zeitunabhängige Datenübertragung gefordert (REQ1.1). Für diese Übertragung der im Herzschrittmacher gespeicherten Daten an das externe Lesegerät ist eine Funkverbindung notwendig. Damit die manuelle Übertragung bei ausreichend guter Verbindung gestartet werden kann, soll die Verbindungsqualität zwischen Herzschrittmacher und externem Lesegerät geprüft und dem Nutzer angezeigt werden können (REQ1.2). Auch während dem Verbindungsaufbau soll dem Nutzer akustische oder visuelle Rückmeldung gegeben werden, so-

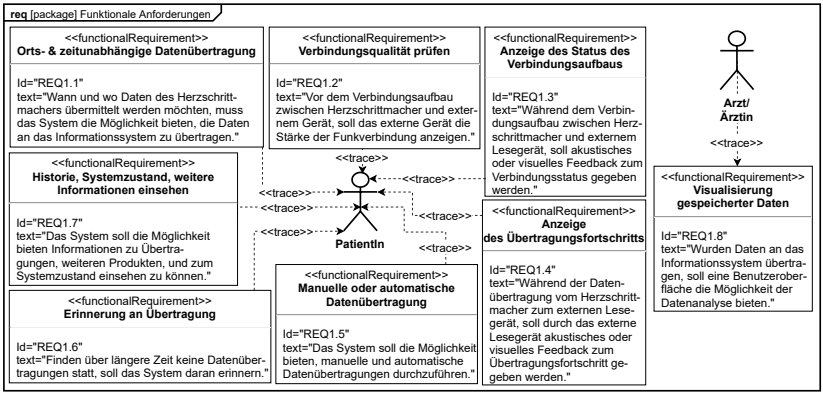



Abbildung 2: *Funktionale Anforderungen an das System*

dass ein erfolgreicher oder nicht erfolgreicher Verbindungsaufbau erkannt wird (REQ1.3). Um den Status der Übertragung der Daten verfolgen zu können, soll während der Datenübertragung ein Fortschrittsbalken den Stand der Übertragung visualisieren (REQ1.4). Neben einer manuellen Übertragung der Daten soll das System zusätzlich die Möglichkeit bieten, die Daten automatisch in regelmäßigen Abständen vom Herzschrittmacher über das externe Lesegerät in das Informationssystem zu übertragen (REQ1.5).
Bei der Verwendung der manuellen Übertragungstechnik soll eine Erinnerung an die Übertragung verhindern, dass über einen längeren Zeitraum keine Herzschrittmacherdaten an das Informationssystem übermittelt werden (REQ1.6). Erinnerungen an bevorstehende Übertragungen und das Einsehen weiterer Informationen, wie beispielsweise Produktinformationen oder Angaben zum aktuellen Systemzustand, sollen den PatientInnen durch das System ermöglicht werden (REQ1.7). Durch ÄrztInnen werden die Anforderungen gestellt, die Daten über eine Benutzeroberfläche visualisieren zu können, um sie zu analysieren und zu bewerten (REQ1.8).
**Nicht-funktionale Anforderungen** (NFAs) beschreiben die qualitativen Merkmale oder Randbedingungen, welche die Entwicklung betreffen [15]. In der Domäne der Medizinprodukte spielen Qualitätsanforderungen eine zentrale Rolle (vgl. Medizinproduktegesetz Abschnitt 2, 4, 5), was sich in den speziell für Medizinprodukte verfügbaren Normen ISO 13485 [8] und ISO 14971 [2] widerspiegelt. Sie beschreiben wesentliche Merkmale hinsichtlich Sicherheit, Verfügbarkeit, Funktionstüchtigkeit, und Stabilität, sowie die Notwendigkeit für ausreichende Qualitäts- und Risikomanagement-Prozesse bei der Herstellung und kontinuierlichen Wartung des Medizinproduktes, wozu bspw. die Aktualisierungen des externen Lesegeräts durch Updates zählt. Da die übertragenen Daten patientenbezogen sind, müssen diese vertraulich behandelt werden. Dazu gehört die Notwendigkeit der Verschlüsselung der Daten und ein beschränkter Zugriff zu diesen nur für autorisierte Personen. Neben den Anforderungen an den Entwicklungs- und Wartungsprozess können die in Abb. 3 dargestellten NFAs abgeleitet werden. Daten werden auf einem Datenbankserver gespeichert, der über regelmäßige Datensicherungen (REQ2.1) verfügen muss. Die regelmäßige Datenübertragung darf zu keiner wesentlichen Laufzeitverkürzung der Batterie führen (REQ2.2). Damit das Lesegerät von unterwegs aus genutzt werden kann, benötigt es zusätzlich eine Stromversorgung durch Batterien (REQ2.3). Durch die ISO 9241 (REQ2.4) kommen für die Zertifizierung weitere Gebrauchstauglichkeitsanforderungen hinzu. Dazu gehört die einfache und intuitive Bedienung des Systems (REQ2.5), die durch gut erkennbare und eindeutige Symbole, der einfachen Navigation durch die Anwendung und dem effektiven, effizienten und zufriedenstellenden Erreichen des Ziels realisiert werden kann.

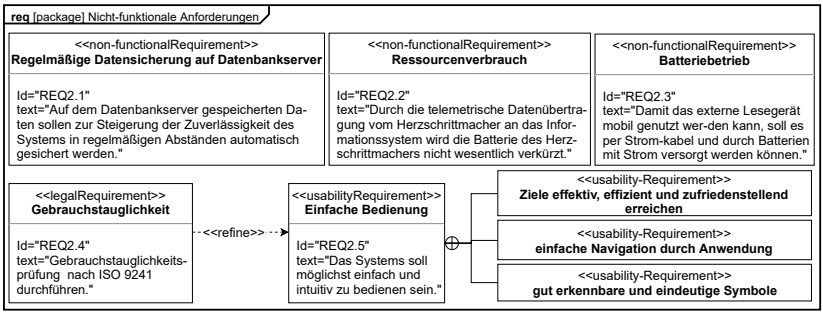



Abbildung 3: *Nicht-funktionale Anforderungen an das System*

## 3.5 Systemkontext und -grenze

Mithilfe des Systemkontext wird die Umgebung des Systems definiert. Den prinzipiellen Aufbau des Systems zeigt Abb. 1a. Die Akteure *PatientIn* und *Arzt/ Ärztin* verwenden Benutzersysteme wie *Computer* und *Smartphone/ Tablet* um mit dem System zu interagieren. Der *Herzschrittmacher*, welcher selbst ein eigenständiges externes System darstellt, liefert Geräte- und Patienteninformationen. Diese Informationen werden mithilfe des Benutzersystems *externes Lesegerät* ausgelesen und übertragen. Das *Informationssystem* mit einem *Datenbankserver* ermöglicht die Speicherung und Abfrage der Daten, sowie die Kommunikation zwischen PatientInnen und ÄrztInnen.

### 3.6 Anwendungsfälle

Anwendungsfälle sind fachlich motivierte Nutzungsszenarien des Systems durch die Akteure, die für das Homemonitoringsystem in Abb. 4 gezeigt sind. Der zentrale Anwendungsfall ist das manuelle und automatisierte Auslesen und Übertragen der im Herzschrittmacher gespeicherten Daten [1, 3], welche auch von unterwegs erfolgen soll [12]. Das Abrufen von Informationen erfolgt über die in Abb. 1a genannten Benutzersysteme Webseite und mobile Anwendung [11, 13, 4]. Weiter konkretisiert bedeutet dies (i) das Abrufen von Informationen vergangener oder bevorstehender Übertragungen, wie beispielsweise Bestätigungen der letzten automatischen oder manuellen Übertragungen [11, 13, 4], (ii) die Anzeige von Informationen über den implantierten Herzschrittmacher und das genutzte Lesegerät und eventuelle weitere Produkte des Herstellers [13], und (iii) den Erhalt von Systeminformationen, wie die Bezeichnung des Herzschrittmachers und des externen Lesegerätes, die aktuelle Softwareversionen des externen Lesegerätes und, wenn vorhanden, Informationen zum Update des externen Lesegerätes erhalten [13]. ÄrztInnen nutzen eine Webseite oder eine mobile Anwendung, um patientenbezogene Daten abzurufen, zu analysieren und auszuwerten. Dazu gehört (i) der Abruf von Daten über Herzaktivitäten der PatientInnen zur Analyse auf Basis eines intrakardialen Elektrokardiogramms [10], und (ii) der Abruf von Daten über die Funktionstüchtigkeit, der programmierten Funktionen und des Batteriestandes [10]. Bei automatischen Datenübertragungen an das Informationssystem legt der Arzt/ die Ärztin das Zeitintervall dieser über eine Webseite oder einer mobilen Anwendung fest.

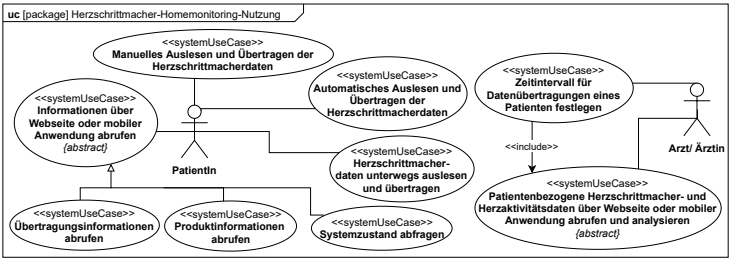
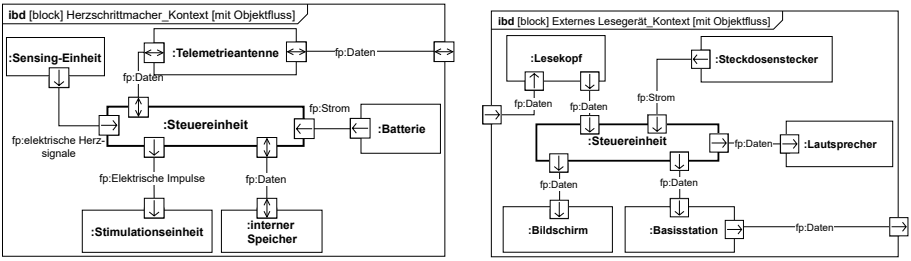



Abbildung 4: *Anwendungsfälle der PatientInnen und ÄrztInnen*

### 3.7 Systemaufbau

Der vereinfachte Aufbau der Subsysteme *Herzschrittmacher* und *externes Lesegerät* des Herzschrittmacher-Homemonitoringsystems werden durch interne Blockdiagramme beschrieben. Die *flow properties (fp)* der Verbindungen beschreiben die Eigenschaften der Elemente, welche zwischen den Blöcken fließen. Welche Daten fließen, hängt vom Verwendungskontext ab.

Ein Herzschrittmacher besteht aus den in Abb. 5a gezeigten Systembausteinen, zu welchen eine Batterie als Energiequelle gehört. Die Systembausteine *Sensing-Einheit*, welche zur Abtastung von elektrischen Signalen am Herzmuskel dient, und *Stimulationseinheit*, welche zur Stimulation des Herzmuskels durch elektrische Impulse dient, sind physisch in einer Sonde vereinigt. Die Steuerung dieser Bausteine wird in der zentralen *Steuereinheit* durchgeführt. Der *interne Speicher* des Herzschrittmachers dient zur Speicherung von Herz- und Geräteaktivitäten sowie der Programmierung des Systems. Er kann mithilfe der Telemetrieantenne beschrieben und ausgelesen werden. Voraussetzung für die Kommunikation zwischen Telemetrieantenne und Umgebung ist die Kompatibilität der verwendeten Ports. Die bidirektionale Kommunikation ist aktuell nur in klinischen Untersuchungen möglich. Das interne Blockdiagramm des externen Lesegerätes ist in Abb. 5b zu sehen. Die Modellierung dieses Subsystems hängt von der Art des Lesegerät für das Homemonitoring ab. Es gibt ein- und mehrteilige Systeme, wobei das hier vorgestellte System aus mehreren Hardwarekomponenten aufgebaut ist. Mithilfe der Telemetrieantenne aus Abb. 5a und dem Systembaustein *Lesekopf* kann eine Verbindung zwischen Herzschrittmacher und externem Lesegerät hergestellt werden. Dabei werden Daten des Herzschrittmachers über eine Funkverbindung zwischen Telemetrieantenne und Lesekopf übertragen [10, 6]. Zweite Hardwarekomponente ist die Basisstation, welche unidirektional mit der *Steuereinheit* und der Systemgrenze verbunden ist. Vom Lesegerät empfangene Daten werden per Funk an die Basisstation übermittelt, und von dort über das Mobilfunknetz eines herstellerspezifischen Anbieters [1, 3] an ein Informationssystem übertragen [17]. Für visuelles und akustisches Feedback dienen die Systemkomponenten Bildschirm und Lautsprecher [3, 12]. Als Energiequelle dient ein Steckdosenstecker für das Stromnetz, oder Batterien in anderen Systemausführungen.




(a) Internes Blockdiagramm eines Herzschrittmachers (b) Internes Blockdiagramm eines externen Lesegerätes

Abbildung 5: Systemaufbau von Herzschrittmacher und externem Lesegerät

# 4 Zusammenfassung

In dieser Arbeit konnte die herstellerunabhängige Modellierung eines Homemonitoringsystems für Herzschrittmacher mithilfe des SYSMOD-Prozesses und der SysML gezeigt werden. Für die Modellierung des im System stattfindenden Datenaustausches wurden zunächst die Hauptkomponenten *Herzschrittmacher*, *externes Lesegerät* und *Informationssystem* identifiziert. Um die Datenübertragung zwischen Herzschrittmacher und Informationssystem detailliert beschreiben zu können, wurden der Herzschrittmacher und das externe Lesegerät zusätzlich durch ein internes Blockdiagramm modelliert. Mithilfe der dadurch ersichtlichen Verbindungen zu den jeweiligen Systemgrenzen konnten die Funkverbindung zwischen Herzschrittmacher und externem Lesegerät sowie die Mobilfunkverbindung zwischen externem Lesegerät und Informationssystem aufgezeigt werden. Durch die Modellierung kann ein grundlegendes Verständnis der Funktionsweise und des Aufbaus der Systeme übermittelt werden, welches für Weiterentwicklungen solcher notwendig ist.

# 5 Referenzen


[1] Informationsbroschüre- medtronic mycarelink -patientenmonitor, 2013.

[2] ISO 14971. Medical devices-application of risk management to medical devices, 2007.

[3] Biotronik. Biotronik home monitoring. `https://www.biotronik.com/de-de/products/home-monitoring`, 2019. 2019-10-06.

[4] Biotronik. Biotronik patient app. `https://www.biotronik.com/de-de/patients/patientapp`, 2019. 2019-12-06.

[5] Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Herzschrittmacher-implantation - indikatoren 2016.

[6] Craig M. Housworth and Nancy P. Pool. Implantable medical device system with communication link to home appliances, 2005. US Patent Application No. US 2005/0187593 A1.

[7] Mohd Noor Islam and Mehmet R. Yuce. Review of medical implant communication system (mics) band and network. ICT Express, 2(4):188 – 194, 2016. Special Issue on Emerging Technologies for Medical Diagnostics.

[8] EN ISO. 13485: 2003. Medical Device: Quality Management Systems. Requirements for the Regulatory Process, 2003.

[9] Jörg Kiefer. Medtronic Pressemitteilung- Medtronic, hat die CE- Zulassung für die automatische Version des Medtronic CareLink® Networks mit der Conexus$^{TM}$ Wireless Telemetry Funkübertragungstechik erhalten. Düsseldorf, 2007.

[10] Rüdiger Kramme. Medizintechnik. Springer-Verlag Berlin Heidelberg, 2016.

[11] Medtronic. Medtronic mycarelink connect. `https://www.mycarelinkconnect.com/PatientInfo`, 2019. 2019-12-06.

[12] Medtronic. Medtronic mycarelink smart patient monitor. `https://www.medtronic.com/us-en/patients/treatments-therapies/remote-monitoring/available-monitors/mycarelink-smart-patient.html`, 2019. 2019-12-06.

[13] MedtronicCardiac. Medtronic mycarelink connect$^{TM}$. `https://www.youtube.com/watch?v=w2bI2vfOWaU`, 2016. 2019-12-06.

[14] OMG. OMG Systems Modeling Language (OMG SysML), Version 1.4, 2015.

[15] Klaus Pohl and Chris Rupp. Basiswissen Requirements Engineering: Aus-und Weiterbildung nach IREB-Standard zum Certified Professional for Requirements Engineering Foundation Level. dpunkt. verlag, 2015.

[16] M. J. Pekka Raatikainen et al. Remote monitoring of implantable cardioverter defibrillator patients: a safe, time-saving, and cost-effective means for follow-up. In EP Europace, volume 10, pages 1145–1151. 2008. Journal Article.

[17] Mark H. Schoenfeld et al. Remote monitoring of implantable cardioverter defibrillators: A prospective analysis. In Pacing and clinical electrophysiology, volume 27, pages 757–763. 2004.

[18] Tim Weilkiens. Systems Engineering mit SysML/UML. dpunkt.verlag Heidelberg, 3 edition, 2014.

# Can basic (bio-)mechanical characteristics define tumor tissue in head and neck squamous cell carcinomas?

*R. Grässlin[1], A.M. Seitz[2], M.O. Scheithauer[1], T.K. Hoffmann[1], S. Weißinger[3], P. Möller[3], L. Dürselen[2], P.J. Schuler[1]*

*[1] Department of Otorhinolaryngology, Head and Neck Surgery*

*Ulm University Medical Center, Frauensteige 12, 89075 Ulm, Germany*

*[2]Department of Biomechanics, Ulm University*

*Helmholtzstraße 14, 89081 Ulm, Germany*

*[3]Department of Pathology, Ulm University*

*Albert-Einstein-Allee 23, 89081 Ulm, Germany*

contact: rene.graesslin@uniklinik-ulm.de

**Abstract:**

The characterization of tumor tissue is gaining more and more importance. The molecular typification catalogue is expanded by the advancement in immunotherapy and the search for suitable biomarkers for individualized therapy. However, tumors can also be characterized by clinical characteristics such as their palpable nature. Most head and neck tumors present themselves as solid lesions. Therefore, the question arises whether a biomechanical characterization model for head-neck carcinomas can be created. For this purpose, an experiment was established to compare the biomechanical characteristics of tumor tissue originating from head and neck squamous cell carcinomas and healthy mucosa (n=16). A significant difference in tissue stiffness could be shown. The tactile sensation of malignant neoplasia, portrayed in solid nodes, could be confirmed. Further study approaches could lead to the development of further tumor profiles based on "classical" mechanical properties. The goal is a fast and favorable basic evaluation of suspicious clinical features.

**Keywords:** head and neck, squamous cell carcinoma, biomechanical characteristics, stiffness

## 1. Problem:

In the age of individualized tumor therapy tissue characterization reaches new levels of importance. A recent example is the inclusion of the HPV-status in the 8th edition of the AJCC Cancer Staging Manual for classifying oropharyngeal cancers due to their significantly better prognosis [5][11]. Furthermore, the success of immunotherapy, majorly based on the PD-1/PD-L1 and CTLA-4 pathways, was honored with the 2018 Nobel Prize [17]. These developments require biomarkers for a targeted therapy because each tumor presents a unique set of molecular attributes. Still, a malignant lesion provides certain visual and haptic features which are often the reason for patients to go to the doctor in the first place. In the head neck, tumors are frequently detected in the physical examination including inspection and more important palpation. Suspect lesions can often be discovered as stiff or rigid nodes, especially in cervical lymph nodes metastasis as well as reachable interior sites such as the oral cavity and parts of the oropharynx. There have been developments to create specialized imaging systems to determine rigid areas [8][9]. Not only is tissue stiffness a clinical sign which helps the diagnosis but its association with the risk of malignant transformation has been shown in the breast and liver [18][3]. The regulatory role of mechanical properties may just be as important as the biochemical interactions between and within cells. The tissue's stiffness determines its homeostasis and can, if thrown off balance, as in an event of injury, result in disease

progression, even without alteration of the biochemical composition [10][4]. For this study the focus was set on the basic biomechanical differences between tumor tissue of head and neck squamous cell carcinomas and healthy mucosa surrounding said tumors. The aim was to evaluate whether the clinical sensation the physician receives from touching tumor lesions can be measured.

## 2. Material & Methods:

In this study, tissue from head and neck squamous cell carcinomas was analyzed. Samples were taken by an 6mm punch biopsy needle (**Fig. 1**) To ensure consistency only large tumors, mostly T3 or T4, were chosen. The punch biopsy was performed in the midst of the primary tumor after it had been surgical resected. If possible, an additional sample was taken from the border area. Lymph node metastasis were not included. Yet another biopsy was taken from the surrounding mucosa with a clear distance from the main tumor. All samples were stored in NaCl and afterwards immediately transported to the biomechanical laboratory to be tested the very same day.

**Fig. 1** *Biopsy Punch, 6mm, by Stiefel*

The 6mm diameter biopsy samples were tested using a custom-made test setup (Fig. 1), which was integrated into a standard-materials testing machine (Z010, Zwick GmbH, Ulm, Germany). After an initial preloading phase the samples were exposed to stress-relaxation tests under confined compression conditions at a compression level of 2 N [13]. The strain level was held constant for 10 minutes to ensure that the equilibrium state was reached. Then, the aggregate modulus (HA), hydraulic permeability (k) and the equilibrium modulus (Eeq) were determined solving a one-dimensional diffusion equation [7][14] using an iterative least square algorithm:

$$\sigma_t = \sigma_{t\to\infty} + 2 * H_A * \varepsilon_i * e^{(-(\frac{\pi}{h_0})^2 * H_A * k * t)} \quad \textbf{(Eq. 1)}$$

where $\sigma_t$ is the time-dependent strain, $\sigma_{t\to\infty}$ is the equilibrium stress, and h0 the initial sample height.

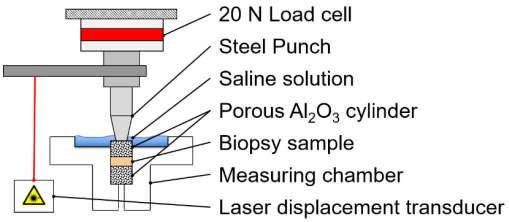
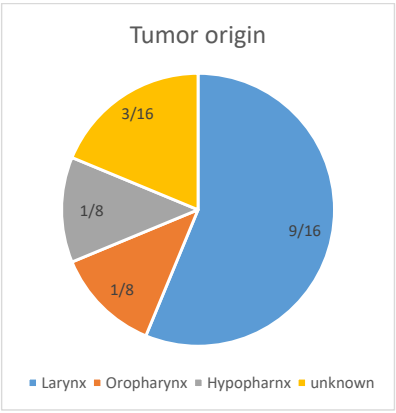



**Fig. 2** Customized test setup to perform a confined compression stress relaxation test to identify the biomechanical material properties of the biopsy samples.

Gaussian distribution of the results data (Ha, k and Eeq) was checked using the Shapiro-Wilk test. Normally distributed data were then examined using Student´s t-test (SPSS) and $p \leq 0.05$ was considered statistically significant.

### 3. Results:

A total of 16 biopsy pairs were analyzed. All of which were squamous cell carcinomas: 9/16 (56,25%) from the larynx, 2/16 (12,5%) from the oropharynx, 2/16 (12,5%) from the hypopharynx and 3/16 (18,75%) of unknown origin (**Fig. 3**). Mostly large primary tumors were included in the testing (T2 1/16 (6,25%), T3 5/16 (31,25%), T4 7/16 (43,75%), unknown 3/16 (18,75%) (**Fig. 4**). 8/16 (50%) cases revealed a tumor first diagnosis, in 5/16 (31,25%) a recurrence, 3/16 (18,75%) were unknown. The median age was 69 (53 – 75). In two cases it was possible to gain 3 samples, one from the center, one from the border area of the tumor and one from healthy mucosa.



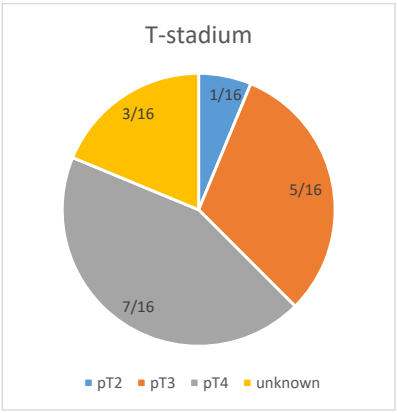
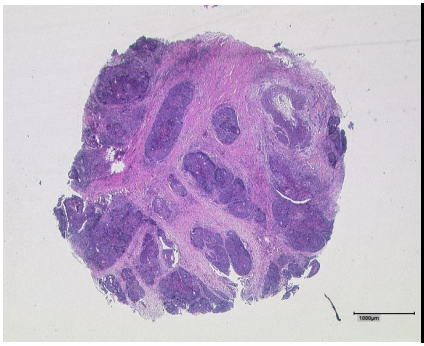




**Fig.3** Primary tumor locations

**Fig. 4** pT-Stadium of the resected tumors

The histological diagnosis was obtained from the surgically R0-resected tumor specimen, not the punch biopsies. As another control the biopsies themselves were histologically examined by a professional pathologist. In **Fig. 5** a complete section of the punch biopsy is shown, **Fig. 6** shows an enhancement of a tumor fraction. It could be shown that 87,5% (14/16) of the biopsies revealed a squamous cell carcinoma. The histological proof of a carcinoma in the tumor specimen was obtained in all cases. The median percentage of malignant cells within the biopsies was 35 (5 – 80), the median percentage of necrosis was 1 (0 – 25).



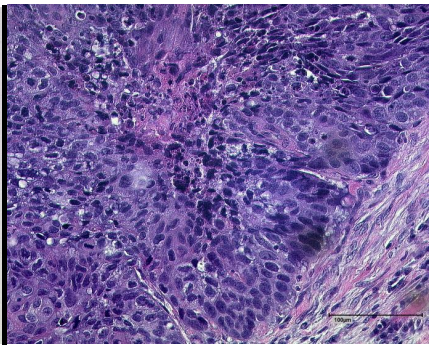


**Fig. 5** Punch biopsy, complete section, HE

**Fig. 6** Tumor tissue, close-up

Aggregate modulus (Ha) calculated with the diffusion equation (**Eq. 1**) of the healthy tissue was 36.1% higher (p = 0.04) than those of the tumor tissue (**Fig. 7**). Although the permeability of the tumor tissue was 71.4% higher than those of the healthy tissue the results were not statistically different (p = 0.18). The equilibrium modulus (Eeq) was 40.0% higher and therefore statistically higher for the tumor tissue (p = 0.04) when compared to the healthy tissue.

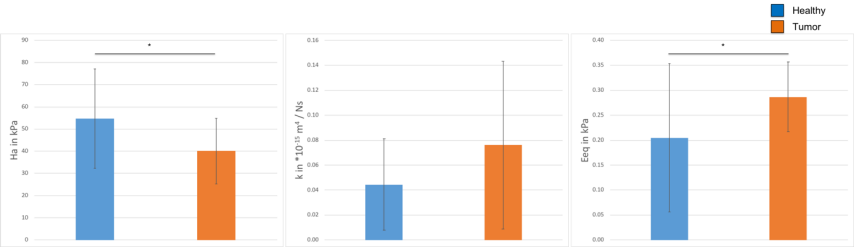



**Fig. 7** Aggregate modulus (HA) permeability (k), and equilibrium modulus (Eeq) values of healthy and tumor tissue samples (Mean values ± SD; n = 16; *p<0.05).

## 4. Discussion:

The importance of biomechanical aspects in tumor development is beginning to become more and more important. These aspects are based on the tissue composition of the cancer itself. Besides malignant transformed cells the tumor microenvironment includes non-malignant cells, proteins and blood vessels [1]. All components within the tumor interact and affect each other. Therefore, tumors can be considered as functioning organs [2]. It is now shown that increased stiffness can often lead to a malignant transformation as well as metastasis and can already be found in premalignant tissue [12][6]. But in relation to healthy mucosa the structural components within cancer tissue which stabilize and define the “organ” may be less resistant. This thought can be underlined by the higher permeability shown in our results. An increased deformability of the cancer cells themselves could be an explanation [15]. Considering this option, the extracellular matrix must play a leading role in the creation of tissue stiffness [4]. Still, there is yet another special feature when it comes to cancer tissue. One lesion may be tough by touch but can also have soft parts on the inside due to necrotic parts within the lesion. Necrosis is considered as a poor prognostic factor [16]. It is possible for necrosis to reduce tissue stiffness. Our results show that the amount of tumor cells and the percentage of necrosis within a tumorous lesion can vary widely. The hypothesis that samples with a higher percentage of necrosis display more elasticity could not be proved directly. But the samples with biopsies from the border area and the center showed a clear difference in stiffness, the peripheral biopsies being significantly stiffer. This represents an often-observed phenomenon of tumor tissue with a hard surrounding and a soft inside due to necrosis. The alteration of the stiffness in necrotic parts of the tumor can increase invasiveness and account for an advanced disease [19]. The small sample size with a widespread variation of tumor cells and necrosis as well as the non-standardized biopsies may be limiting factors. Still, the equilibrium modulus suggests that there is in fact a significant difference in tissue stiffness when comparing cancerous lesions and healthy tissue and therefore a tissue definition based on biomechanical characteristics can be possible.

## 5. Conclusion:

Biomechanical characterization of different tissue types is possible. Cancer as an inhomogeneous tissue has the possibility to present multiple different phenotypes. Therefore, larger sample sizes and further studies are necessary to create a biomechanical catalogue to distinguish different types.

**6. References:**


1. Ansell S M, Vonderheide R H: Cellular composition of the tumor microenvironment. Am Soc Clin Oncol Educ book Am Soc Clin Oncol Annu Meet (2013)
2. Bissell M J, Radisky D: Putting tumours in context. Nat Rev Cancer 1: S. 46 (2001)
3. Boyd N F, Li Q, Melnichouk O, Huszti E, Martin L J, Gunasekara A, Mawdsley G, Yaffe M J, Minkin S: Evidence That Breast Tissue Stiffness Is Associated with Risk of Breast Cancer. PLoS One 9: S. 1–8 (2014)
4. Cox T R, Erler J T: Remodeling and homeostasis of the extracellular matrix : implications for fibrotic diseases and cancer. Dis Model Mech S. 165–178 (2011)
5. Doescher J, Veit J A, Hoffmann T K: [The 8th edition of the AJCC Cancer Staging Manual : Updates in otorhinolaryngology, head and neck surgery]. HNO 65: S. 956–961 (2017)
6. Emon B, Bauer J, Jain Y, Jung B, Saif T: Biophysics of Tumor Microenvironment and Cancer Metastasis - A Mini Review. Comput Struct Biotechnol J 16: S. 279–287 (2018)
7. Frank E H, Grodzinsky A J: Cartilage electromechanics—I. Electrokinetic transduction and the effects of electrolyte pH and ionic strength. J Biomech 20: S. 615–627 (1987)
8. Garra B S: Imaging and estimation of tissue elasticity by ultrasound. Ultrasound Q 23: S. 255–268 (2007)
9. Glaser K J, Felmlee J P, Manduca A, Kannan Mariappan Y, Ehman R L: Stiffness-weighted magnetic resonance imaging. Magn Reson Med 55: S. 59–67 (2006)
10. Handorf A M, Zhou Y, Halanski M A, Li W: Tissue Stiffness Dictates Development , Homeostasis , and Disease Progression. Organogenesis 11: S. 1–15 (2015)
11. Lai K, Killingsworth M, Matthews S, Caixeiro N, Evangelista C, Wu X, Wykes J, Samakeh A, Forstner D, Niles N, Hong A, Lee C S: Differences in survival outcome between oropharyngeal and oral cavity squamous cell carcinoma in relation to HPV status. J oral Pathol Med Off Publ Int Assoc Oral Pathol Am Acad Oral Pathol 46: S. 574–582 (2017)
12. Levental K R, Yu H, Kass L, Lakins J N, Erler J T, Fong S F T, Csiszar K, Giaccia A, Yamauchi M, Gasser D L, Weaver V M: Matrix Crosslinking Forces Tumor Progression by Enhancing Integrin signaling. Cell 139: S. 891–906 (2009)
13. Martin Seitz A, Galbusera F, Krais C, Ignatius A, Durselen L: Stress-relaxation response of human menisci under confined compression conditions. J Mech Behav Biomed Mater 26: S. 68–80 (2013)
14. Mow V C, Kuei S C, Lai W M, Armstrong C G: Biphasic Creep and Stress Relaxation of Articular Cartilage in Compression: Theory and Experiments. J Biomech Eng 102: S. 73–84 (1980)
15. Remmerbach T W, Wottawah F, Dietrich J, Lincoln B, Wittekind C, Guck J: Oral cancer diagnosis by mechanical phenotyping. Cancer Res 69: S. 1728–1732 (2009)
16. Richards C H, Mohammed Z, Qayyum T, Horgan P G, McMillan D C: The prognostic value of histological tumor necrosis in solid organ malignant disease: a systematic review. Future Oncol 7: S. 1223–1235 (2011)
17. Rotte A, D'Orazi G, Bhandaru M: Nobel committee honors tumor immunologists. , *Journal of experimental & clinical cancer research : CR*, 37: . England, S. 262 (2018)
18. Singh S, Fujii L L, Murad M H, Wang Z, Asrani S K, Ehman R L, Kamath P S, Talwalkar J A: Liver Stiffness Is Associated With Risk of Decompensation, Liver Cancer, and Death in Patients With Chronic Liver Diseases: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 11: S. 1573-e89 (2014)
19. Swaminathan V, Mythreye K, Tim O'Brien E, Berchuck A, Blobe G C, Superfine R: Mechanical Stiffness grades metastatic potential in patient tumor cells and in cancer cell lines. Cancer Res 71: S. 5075–5080 (2011)

# An Online Model Checking Approach to Soft-Tissue Detection for Rupture Prediction

*Sven-Thomas Antoni*[1*], *Sascha Lehmann*[2*], *Sibylle Schupp*[2], *Alexander Schlaefer*[1]

[1] *Institute of Medical Technology, Hamburg University of Technology, Hamburg, Germany*
[2] *Institute for Software Systems, Hamburg University of Technology, Hamburg, Germany*
[*] *Authors contributed equally*

Contact: s.lehmann@tuhh.de

***Abstract***

Robotic needle insertion based on haptic feedback can be imprecise and error-prone, especially for sudden force changes in case of ruptures. To predict rupture events early during tissue deformation, knowledge is required about the type and characteristics of the tissues involved. Several approaches to this exist and increase system complexity by including additional sensors or imaging modalities. We introduce a new approach based on formal *model checking*, which allows us to identify tissue by a directed search through the state space of a needle insertion model. Using force data measured at the needle shaft during cutting motion, our method identifies the most probable tissue iteratively at run-time, based on *a priori* information of possible tissues. In a case study of needle insertions into gelatin phantoms with varying gelatin-water ratios, our approach allowed 90.7% correct identifications and may thus be considered to identify tissue during robotic needle insertion.

**Keywords**: Robotic needle insertion, tissue rupture, online model checking, run-time verification, tissue discrimination

## 1 Problem

Precise positioning of needles with a robot is a very challenging task, and systems need to compensate for deformation and have to avoid ruptures. Here, forces build up during deformation and rapidly unload into the tissue once a tissue-specific rupture force is exceeded. This results in uncontrollable crack extension [1] that can complicate precise needle placement and can cause additional damage to deeper tissue. The later aspect is especially dangerous if the deeper tissue, for example the bladder during prostate biopsy, should not be damaged. Ruptures can be counteracted twofold, by adding a safety margin towards sensitive tissue or by using some techniques to avoid ruptures, i.e., rotating the needle [2]. A constantly rotating needle will cause additional stress to the tissue and complicate the navigation.

One way to evaluate ruptures is measuring the forces acting on the needle. If this is done externally, the measured force is likely to be a combination of several smaller forces from different sources. Okamura et al. [3] define a model that splits the needle insertion force into a cutting force, a friction force, and a stiffness force that is due to the elastic properties of the tissue and its capsule. The cutting and stiffness forces occur at the needle tip, and the friction force at its shaft. For the rupture, the friction force plays a secondary role, and cutting forces are only occurring post-rupture. Hence, the stiffness force is best examined to predict ruptures. While using external measurements from the needle shaft, the stiffness force may be superimposed by the ever increasing friction force. Force measurements at the needle tip would allow direct measurement of the current cutting or stiffness force. Different approaches to measure forces have been proposed, i.e., miniaturized force sensors [4], interferometer based [5], using fiber Bragg grating sensors [6], or using optical coherence tomography (OCT) [7]. Unfortunately, the elastic properties and hence the stiffness forces of tissues vary significantly between tissues, individuals, and also over time. A force that will rupture one tissue may not rupture another tissue. As a result, if a rupture should be predicted, the possible tissues in front of the needle need to be known beforehand.

While direct imaging could be used to locate the needle and identify the surrounding tissue, this complicates the setup, increases costs and for some modalities leads to exposure to ionizing radiation (for example CT and X-Ray fluoroscopy). Instead, a number of methods have been proposed that use the needle itself as a sensor to discriminate between different tissues during runtime. Probably the most popular approach is using the needle for impedance spectroscopy with first products based on this technique currently entering the market [8]. Other methods include needle-based fluorescence spectroscopy [9], or using optical coherence tomography (OCT) from the needle tip to either discriminate between tissues directly [10] or to investigate the tissue's cellularity [11]. One major drawback of all these approaches is that they require additional equipment.

| Identifier | Ratio | $a_{\text{friction}}$ [N/mm] | $c_{\text{friction}}$ [N] | $c_{\text{cutting}}$ [N] | Young's modulus [kPa] |
|---|---|---|---|---|---|
| *GP-4* | 1:4 | $0.038 \pm 0.001$ | $0.042 \pm 0.05$ | $0.886 \pm 0.04$ | $42.57 \pm 3.18$ |
| *GP-6* | 1:6 | $0.029 \pm 0.001$ | $0.096 \pm 0.03$ | $0.293 \pm 0.02$ | $34.93 \pm 2.25$ |
| *GP-8* | 1:8 | $0.017 \pm 0.001$ | $0.071 \pm 0.03$ | $0.238 \pm 0.03$ | $32.33 \pm 2.09$ |
| *GP-10* | 1:10 | $0.014 \pm 0.001$ | $0.046 \pm 0.03$ | $0.141 \pm 0.02$ | $24.34 \pm 0.97$ |
| *GP-12* | 1:12 | $0.013 \pm 0.001$ | $0.026 \pm 0.04$ | $0.116 \pm 0.01$ | $17.94 \pm 0.94$ |

Table 1: *Mechanical properties of different gelatin-water ratios for our custom made 14 gauge needle at 1mm/s.*

We propose that by knowing the forces that occur along the path of the needle through the different tissues, possible adjacent tissues could be identified. As the forces at the shaft or the tip are important measurements for rupture avoidance, they would need to be measured anyway.

Our approach uses Online Model Checking (OMC), a model-based verification approach from the field of formal methods, which was recently also applied in the medical field [12]. One advantage of OMC is that it can be based on very simplified mechanical models, which are varied and verified over time. OMC requires no traditional training and, instead, is fed with a-priori information of the physical tissue properties. By comparing the measured forces with predictions of different tissues, the most likely candidates can be identified. This way, OMC is an attractive alternative to machine learning approaches. While machine learning in theory could be used to solve the problem at hand, typically extensive amounts of data are required for training. Additionally, interpreting the result becomes harder if the actual decision mechanism is hidden as it usually is with most machine learning methods.

In this first, preliminary study, we measured shaft and tip forces during needle insertion of different gelatin phantoms. We will show that using our approach, the tissue can be correctly identified in 90.7% of all cases.

# 2 Material and Methods

## 2.1 Needle Insertion Model

For this work we use the needle insertion model of Okamura et al. [3] to calculate the expected external forces.

$$f_{\text{needle}}(x) = f_{\text{stiffness}}(x) + f_{\text{friction}}(x) + f_{\text{cutting}}(x) \tag{1}$$

Using only constant insertion speeds, we can use a *Coulomb & viscous friction* model

$$f_{\text{friction}}(x) = a_{\text{friction}} x + c_{\text{friction}} \tag{2}$$

with a viscous slope factor $a_{\text{friction}} > 0$ that expresses the increase in applied friction per depth unit and a Coulombic friction value offset $c_{\text{friction}}$. While ruptures are an important part of our overall approach, we focus on the tissue discrimination here for which the stiffness force $f_{\text{stiffness}}(x)$ is not relevant. The cutting force is assumed to be constant per tissue, so $f_{\text{cutting}}(x) = c_{\text{cutting}}$.

## 2.2 Gelatin Phantom Data

**Gelatin** We use Balistic 3 gelatin (GELITA AG, Germany) to create gelatin phantoms. The gelatin allows very accurate and reproducible phantoms with a Bloom variance of below 2%. Glycerol is added to slow down deterioration and titan dioxide is used to improve the optical properties of the gelatin. We combine two gelatins with different mechanical properties in a single beaker. This allows us to compare the performance of our method at different tissue depths. A plastic membrane is inserted between the two gelatins to avoid mixing. See Figure 1D for an example.

**Needle** We use a custom made 14 gauge (1.98mm) needle, see Fig. 1B. An optical fiber is included in the needle and is used to measure the deformation of a deformable epoxy layer at the needle tip, compare Fig. 1C.

**Mechanical Properties** Many of the mechanical properties of gelatin measured during needle insertion depend on the needle. From previous experiments we know the key attributes for different gelatin-water concentrations with the previously introduced needle, see Table 1.

The gelatin phantoms are chosen to represent Young's modulus values of typical cancerous (*GP-4*, *GP-6*, *GP-8*) and non-cancerous (*GP-10*, *GP-12*) tissues in the medial-mid prostate [13].

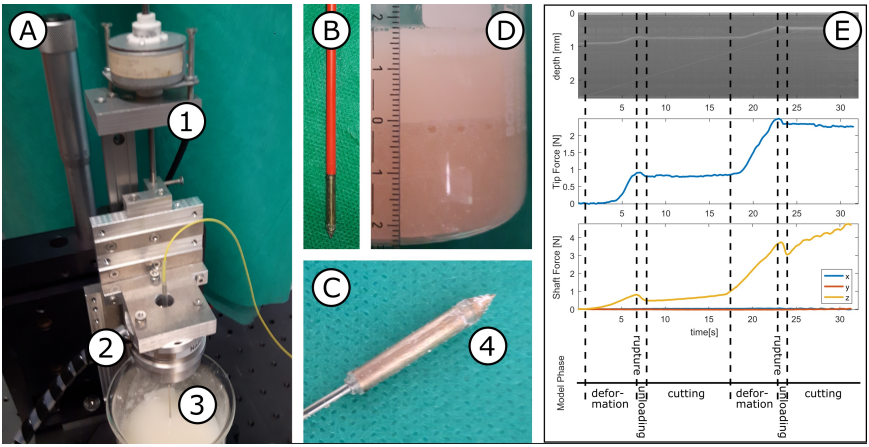



Figure 1: *The setup (A) with stepper motor controlled linear stage (1), force torque sensor (2) and needle (3 and B). At the needle tip (C) a small compressible epoxy layer (4) is used to measure force. The gelatin (D) consists of two layers divided by a thin plastic membrane. An example recording (E) is presented consisting of OCT A-Scan, tip force, shaft force and labeled phases (cf. Fig. 2)*

**Setup** The setup to measure the forces is displayed in Fig. 1A. The needle is attached to a force torque sensor (Nano43, ATI Industrial Automation Inc., USA), which is moved using a linear stage that is actuated by a stepper motor. The needle is connected to the OCT system (Telesto, Thorlabs GmbH, Germany) by an optical fiber. The stepper motor inserts the needle at a constant speed of 1mm/s.

## 2.3 Material

We recorded 45 needle insertions (see Figure 1E for an example), and extracted the forces during the cutting phases for the analysis. Of the resulting 90 cutting phases, we sorted out cutting phases that were less than 2.5mm deep. A total of 75 cutting phases remained of which 20, 14, 11, 11 and 19 belonged to *GP-4*, *GP-6*, *GP-8*, *GP-10* and *GP-12*, respectively.

## 2.4 Online Model Checking

Model checking approaches use a mathematical-logical description of a system and a requirement specification, and derive statements on whether the specified properties hold for the given system or not, providing a concrete counterexample in the latter case. For our application, we rely on *timed automata (TA)* [14] as model formalism for the system description, which is based on the concept of finite state machines, extended to infinite-state cases via (continuous) data variables and real-valued clocks. Possible requirements to be checked on such models include that a given condition always holds on all model paths (*safety*), that a property is satisfied at some point on one of the paths (*reachability*), or that a property is always eventually satisfied (*liveness*). For many real-world scenarios, the basic model checking approach reaches its practical limits due to the large state spaces induced by the underlying system models, which would have to be traversed and analyzed exhaustively. *Online Model Checking (OMC)* in particular targets this performance problem by only checking relevant sections of the state space iteratively at run-time instead of verifying the model as a whole in advance. For that purpose, a time horizon is set, which defines the model-internal clock time up to which model states should be verified. As back-end for our online verification and synchronization framework, we use the model checker *Uppaal* [15], which uses models based on the timed automata formalism, and the TCTL logic for timed requirement specification.

The goal of our work is to use the property-based model traversal of model checking to identify paths through the state space - spanned by multiple parameter values reflecting the potential types of tissue layers - which sufficiently fit the force data measured at the needle. To achieve that goal, we repeatedly perform a synchronization step of the model and the force data observed at the underlying needle system, where we try

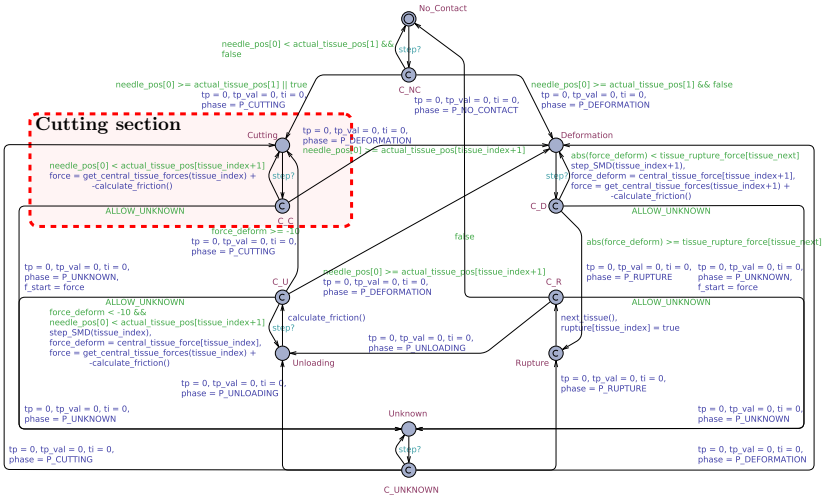



Figure 2: *The main model component of the needle insertion process, divided into phases of* `No_Contact`*,* `Cutting`*,* `Deformation`*,* `Rupture`*, and* `Unloading`*. The important cutting section is marked by a dashed box.*

to determine a valid alignment of the model and the observed data, and to thus identify the most likely model path based on the given data evidence.

## 2.5 Repeated Synchronization

To keep the model synchronized with our observations, we transform the force observation data into a new model component synchronized (via synchronization channels in Uppaal) with the original system model [16]. While the original model allows taking an arbitrary path through its state space, the extended model now only allows paths to reach as far as the simulated model data matches the timed force observation data. This way, we can guide the execution through the model along desired force data points, and discard paths that will not fit the data - and should thus not be considered for tissue identification - at an early stage. A final location `Synced`, indicating the successful identification of the tissue, is entered in the new model component on a path as soon as the end of the observation data trace is reached without prior mismatching data. The property `E<> Synced` (a reachability property for the `Synced` location formulated in the temporal logic TCTL) then allows searching for the existence of such path. The concrete parameters assigned at the beginning of the identified path finally tell us the current tissue.

## 2.6 Needle Automata System

For this case study, we use a network of parallely composed timed automata (NTA). Generally speaking, the base model network consists of 4 different components: The initialization model `Init_T`, a stepper model `TimeStep_T`, the needle model `Needle_T`, and the insertion phase model `InsertionPhase_T`. `Init_T` allows us to set the system variables and clock constraints to their required starting values, and `TimeStep_T` allows the simulation to step through the time domain discretely. `Needle_T` represents the motion of the needle, which in our case is moved with a constant velocity of 1mm/s. Finally, `InsertionPhase_T` models the tissue force changes during the four phases `Cutting`, `Deformation`, `Rupture`, and `Unloading`, and is shown in Fig. 2. The boolean-valued expressions on the edges represent guards, which only allow the traversal of the edge when the condition holds. Furthermore, the assignments are evaluated as soon as the edge is eventually taken.

For this work, we only use parts of the NTA, namely the *cutting* phase. Here, the needle cuts through a given tissue $T_1$, and a cutting force specific to the needle, needle velocity, and tissue can be measured alongside a linearly increasing friction force.

| | | Identified as | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | *GP-4* | *GP-6* | *GP-8* | *GP-10* | *GP-12* | *None* | |
| **Actual Tissue** | *GP-4* | 6 | 9 | 1 | 0 | 0 | 4 | 45 |
| | *GP-6* | 4 | 7 | 3 | 0 | 0 | 0 | |
| | *GP-8* | 0 | 1 | 10 | 0 | 0 | 0 | |
| | *GP-10* | 0 | 0 | 2 | 8 | 1 | 0 | 30 |
| | *GP-12* | 0 | 0 | 1 | 12 | 6 | 0 | |

Table 2: *The identification results for different types of gelatin phantoms (GP)*

## 3 Results

The result of all 75 insertions are presented in Table 2. Out of 45 insertions into gelatin representing cancerous tissue (GP-4, GP-6, GP-8), 41 are identified as cancerous and 4 as outlying (*None*). The outliers (i.e., non-matched data traces) are present in *GP-4*. The remaining 30 insertions into gelatin represent non-cancerous tissue (GP-10, GP-12). Here, 27 were discriminated correctly. Processing times were at least 3mm/s. Hence, analysis of our system moving at 1mm/s is realtime capable.

## 4 Discussion

The presented data shows that almost all matchings for each tissue identified either the correct tissue type, or similar ones w.r.t. their mechanical properties. The effect of the stiffness on the friction is non-linear. Hence, the friction torque $a_{\text{friction}}$ between *GP-8*, *GP-10*, and *GP-12* becomes very similar. As a result, the only misclassifications regarding cancerous or non-cancerous tissue are between these three, affecting only 3 out of 75 insertions.

While the different gelatin phantom were shown to be generally distinguishable, the actual outcome of the experiments depends on the overall distinctness of possible paths. In the current state of the implementation, if multiple tissue types would potentially fit, our tool decides for a similar tissue which was discovered first by the model. This decision may lead to cases in which a different gelatin concentration is chosen, even though the correct concentration was identifiable as well. Nonetheless, our results indicate that these different gelatin phantom concentrations in 68 out of 71 cases belonged to the same tissue type. A-priori information on which types of tissue are likely to occur after another specific tissue may help to further reduce the amount of potential cases and thus the risk of overlapping data.

Due to the state space restrictions imposed by the early discarding of unmatched model paths, the approach scales well for more complex models as well. Here, the time for path traversals depends neither on the number of existing paths nor the overall model depth anymore, but only on the set of relevant paths and the desired temporal search scope, which is typically bounded. For that, providing enough a-priori information is crucial to sufficiently limit the state space to only a few expectable scenarios. Fortunately, such data is easily acquired.

In terms of the applicability during run-time, it is important to consider the division of the analysis approach into both an *offline* and an *online* component. Based on experimental training data, the parameter values for a set of distinct experiment progressions can be obtained offline, which we can then locate in the corresponding system model and define as reference cases. During application, our approach then searches the surroundings of such reference cases, which are valid paths in the model, for an online classification of the actually observed data. Therefore, the computational effort during run-time can be significantly reduced.

## 5 Conclusion

In this paper, we presented an online method for the run-time identification of tissue types during surgical needle insertions. We established the data workflow involved, from data acquisition to repeated model adaptions and verification queries. We could show that 90.7% of measured tissue forces were correctly classified by our approach and 5.3% were classified as outlying. Overall, this means that only 3 traces were classified wrong. As all these were from the normal tissue class, the sensitivity of our approach on this limited data set is perfect, i.e., all cancerous tissue was identified, with a very low specificity of 4%.

The tissue identification has been done in a formal and comprehensible way that - in contrast to modern machine learning approaches - allows a physician to assess the decision of the system. Different from machine learning techniques, the amount of training data could be kept low, and instead, tested and evaluated a-priori information was incorporated into the system.

## References


[1] M. Mahvash and P. E. Dupont. "Mechanics of dynamic needle insertion into a biological material". In: *IEEE Transactions on Bio-Medical Engineering* 57.4 (2010), pp. 934–943.

[2] N. Hungr et al. "Design of an ultrasound-guided robotic brachytherapy needle-insertion system". In: *31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2009), pp. 250–253.

[3] A. M. Okamura, C. Simone, and M. D. O'Leary. "Force modeling for needle insertion into soft tissue". In: *IEEE Transactions on Bio-Medical Engineering* 51.10 (2004), pp. 1707–1716.

[4] C. Hatzfeld et al. "A Miniaturized Force Sensor for Needle Tip Force Measurements". In: *Biomedical Engineering* 62.s1 (2017).

[5] S. Beekmans et al. "Fiber-Optic Fabry-Pérot Interferometers for Axial Force Sensing on the Tip of a Needle". In: *Sensors* 17.1 (2016).

[6] S. Kumar et al. "Detecting stages of needle penetration into tissues through force estimation at needle tip using fiber Bragg grating sensors". In: *Journal of Biomedical Optics* 21.12 (2016).

[7] N. Gessert et al. "Needle Tip Force Estimation Using an OCT Fiber and a Fused convGRU-CNN Architecture". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. 2018, pp. 222–229.

[8] S. Halonen et al. "Real-Time Bioimpedance-Based Biopsy Needle Can Identify Tissue Type with High Spatial Accuracy". In: *Annals of Biomedical Engineering* 47.3 (2019), pp. 836–851.

[9] P. N. Werahera et al. "Human feasibility study of fluorescence spectroscopy guided optical biopsy needle for prostate cancer diagnosis". In: *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2015), pp. 7358–7361.

[10] A. Patel et al. "MP34-14 Investigating the feasibility of optical coherence tomography to identify prostate cancer – an ex-vivo study". In: *Journal of Urology* 195.4S (2016).

[11] N. Iftimia et al. "Investigation of tissue cellularity at the tip of the core biopsy needle with optical coherence tomography". In: *Biomedical Optics Express* 9.2 (2018), pp. 694–704.

[12] S.-T. Antoni et al. "Online model checking for monitoring surrogate-based respiratory motion tracking in radiation therapy". In: *International Journal of Computer Assisted Radiology and Surgery* 11.11 (2016), pp. 2085–2096.

[13] B.-M. Ahn et al. "Mechanical property characterization of prostate cancer using a minimally motorized indenter in an ex vivo indentation experiment". In: *Urology* 76.4 (2010), pp. 1007–1011.

[14] R. Alur and D. L. Dill. "A theory of timed automata". In: *Theoretical Computer Science* 126.2 (1994), pp. 183 –235.

[15] G. Behrmann, A. David, and K. G. Larsen. "A Tutorial on Uppaal". In: *Formal Methods for the Design of Real-Time Systems: International School on Formal Methods for the Design of Computer, Communication, and Software Systems*. 2004, pp. 200–236.

[16] S. Lehmann and S. Schupp. "Trace Synchronization and Efficient State Space Reconstruction for Timed Automata in Uppaal". *Submitted*. 2019.

# Model-based Hearing Diagnosis of Middle Ear Condition Using Inverse Fuzzy Arithmetic and Artificial Neuronal Network

*B. Sackmann[1,2], B. Priwitzer[1], M. Lauxmann[1]*

*[1] Reutlingen University, Alteburgstrasse 150, 72762 Reutlingen, Germany*
*[2] Institute of Engineering and Computational Mechanics, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany*

Contact: Benjamin.Sackmann@Reutlingen-University.de

***Abstract***

*Due to the large interindividual variances and the poor optical accessibility of the ear, the specificity of hearing diagnostics today is severely restricted to a certain clinical picture and quantitative assessment. Often only a yes or no decision is possible, which depends strongly on the subjective assessment of the ENT physician. A novel approach, in which objectively obtainable, non-invasive audiometric measurements are evaluated using a numerical middle ear model, makes it possible to make the hidden middle ear properties visible and quantifiable. The central topic of this paper is a novel parameter identification algorithm that combines inverse fuzzy arithmetic with an artificial neural network in order to achieve a coherent diagnostic overall picture in the comparison of model and measurement. Its usage is shown at a pathological pattern called malleus fixation where the upper ligament of the malleus is pathologically stiffened.*

**Keywords**: middle-ear model, model-based ENT diagnostic, wideband tympanometry, fuzzy arithmetic, artificial neural network

## 1 Problem

Current clinical diagnostics of the hearing capability of a patient is based on a number of different subjective and objective audiometric tests. Subjective tests require attention and the active reaction of the patient to sound and comprise amongst others pure-tone audiometry and speech audiometry. These tests are fairly accurate, but test the complete hearing path, consisting of three stages: middle ear, inner ear, and neural path. Comparing with bone-conduction, a less accurate diagnosis of the air conduction (i.e. middle-ear) is possible. Current clinical practice cannot distinguish, with any degree of certainty, the multiple pathologies that produce conductive hearing loss in patients with an intact tympanic membrane and a well-aerated middle ear without exploratory surgery [1]. Further, subjective tests are impossible to perform in new-borns or difficult and with clearly reduced accuracy to perform in young children, but also in elderly or seriously ill adults.

None of the today clinically used objective audiometric methods provides a frequency-specific loss of a particular stage of the hearing path (i.e. middle ear/inner ear/neural path) in isolation, and only a minority provide a direct, quantitative measure of a frequency-specific loss at all. Rather, most methods provide a norm range of performance against which the test data are compared, and are given in a measure, which cannot be directly converted to a loss figure. Thus, today, the medical specialist is concluding his diagnosis of the state of the hearing path based on a weighted judgement of the typically five to eight different tests. In addition, he has to draw his conclusion keeping in mind the highly diverse limitations of these techniques. Thus the most desirable and concise description would be a frequency-, level-, and stage-specific loss (or gain) figure.

Impedance audiometry is today the most important objective audiometric diagnostic test. In the past decades its performance and validity range was improved by expanding the frequency range from single tones (226/1000 Hz) to wideband measurements (0.2 – 8 kHz) and analysing impedance data in amplitude and phase, rather than just comparing the impedance (compliance) curve of a single tone with its normal range. The benefit of wideband impedance measurements was shown by several authors who observed distinctive frequency-specific patterns for various middle ear pathologies like otosclerosis, ossicular discontinuity, tympanic membrane perforation and otitis media with effusion ([2], [1]) in comparison to the statistical range of normal hearing. However, so far, the effects of middle ear pathologies on wideband impedance measurements are not completely understood and there are neither simple objective criteria to be able to distinguish between several pathologies with any certainty, nor a quantitative estimation of middle-ear transmission loss to the inner ear.

Lately, the classical tympanometry and wideband impedance measurements were combined in the so called wideband impedance tympanometry (WBT), which essentially measures the wideband acoustic input impedance of the hearing path (typically 0.2 – 8 kHz) under different static ear canal pressures (-300 – 200 daPa). WBT is not only very practical and efficient, since tympanometry and impedance measurements are performed simultaneously, but also allows to evaluate the wideband impedance at balanced static middle ear pressure (tympanic peak pressure), which leads to better repeatability and comparability of wideband impedance measurements between individuals. Since the static ear canal pressure is varied, the nonlinear stiffening properties of the middle ear (mainly tympanic membrane and joints) are reflected in a systematic way in the measurements which can potentially be used for differential diagnosis as shown in [1] with regard to disarticulation of the incudo-stapedial joint and malleus fixation. Clear advantage that WBT shares with classical tympanometry is stage specificity for the middle ear. Compared to laser vibrometry, WBT has the clear advantage of economic instrumentation and a quick and easy measurement procedure.

Several numerical middle ear models are known in hearing research today, but they do not yet contribute to individual, patient-specific quantitative diagnostics. The basic structure of the middle ear and the function of the individual middle ear elements (eardrum, bones, ligaments, joints, etc.) are the same for all people and can be represented in a mathematical model with the help of the finite element method in a firmly defined structure. The individually different properties of the individual anatomical middle ear structures, such as ligament stiffness or inertia of the ossicles, can be described in the model by variable numerical parameters. Their selection determines the individually different response of the ear model to a specific excitation signal. Our finite element (FE) model of the ear presented in [3], [4] and [5] is based on micro-CT data and static as well as dynamic measurements.

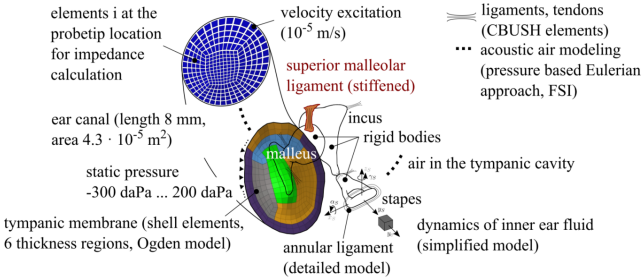



Figure 1: *FE model of the middle ear used for the simulation of wideband tympanometry and the model-based diagnosis. The model has an acoustic ear canal, a tympanic cavity, an elastic tympanic membrane, and an ossicular chain. Furthermore, the superior malleolar ligament is highlighted, since it is later stiffened to simulate a pathological pattern called malleus fixation to demonstrate the model-based diagnostic approach.*

Focusing on the middle ear, this paper suggests a novel approach to arrive at a frequency-, level- and stage-specific diagnosis of middle ear hearing loss by the evaluation of the most advanced audiometric measurement technique, the wideband impedance tympanometry (WBT), within the scope of an identification process to determine individual middle ear model parameters using inverse fuzzy-arithmetic and artificial neural network. By comparing the model response with objective diagnostic measurements on a patient, a patient-specific parameter set can be determined using parameter identification algorithms. Since the model parameters are directly linked to anatomy-specific physical properties of the ear, this process quantifies the individual middle ear properties hidden behind the eardrum and assigns them to a localizable specific anatomical property. For the parameter identification, it is advantageous to have as much information as possible about the individual ear, which is why the model-based approach benefits from the use of extensive objective measurement data and combines these into a coherent overall picture. The diagnosis can finally be made based on objectively determined, quantified and locally resolved middle ear properties, which promises a considerably higher specificity for common middle ear diseases compared to the conventional diagnostic approach [6]. The question therefore arises whether it is possible to move from the conventional comparison with standard curves to an objective, quantitative and personalized diagnosis of the middle ear by means of a model-based evaluation of objective measurements.

The complexity of the anatomical structure requires a large number of model parameters, which cannot be unambiguously determined today by the mere application of standard parameter identification algorithms. The paper therefore suggests a hybrid solution strategy combining inverse fuzzy arithmetic and artificial neural network. Its usage is shown at a pathological case called malleus fixation where the superior ligament of the malleus is pathologically stiffened due to an ossification process.

# 2 Material and Methods

In the following firstly the FE model of the ear and the parameter identification procedure including the inverse fuzzy arithmetic approach and the artificial neural network are presented.

## 2.1 FE Model of the Human Middle Ear

The FE model of the middle ear is shown in Fig. 1. The air in the ear canal and tympanic cavity is modelled with acoustic finite elements. The aditus ad antrum and the air in the mastoid are modelled as a one mass oscillator, which reproduces the oscillating air between the tympanic cavity and the mastoid, and corresponds to a Helmholtz resonator phenomenon. The fluid-structural coupling between the tympanic membrane (TM) and the adjacent air is implemented using a pressure-based Eulerian approach. The ear canal wall is considered rigid by fixing all degrees of freedom. The ossicles are modelled as rigid bodies and characterized by their mass and inertia, because the ossicles deformation can be neglected up to 10 kHz. Whereas malleus and incus have all 6 degrees of freedom, the stapes is further constrained, allowing only a translational piston motion along the y-axis and two rotational (rocking) motions around the x- and z-axis of the stapes coordinate system. The ligaments, tendons and joints are represented by passive spatial spring-damper elements.

The geometry of the tympanic membrane is reconstructed from micro-CT data. Six regions are defined with constant thickness, which were derived from the characteristic relative thickness distribution. When preloaded, the TM stiffens due to large deformation and material nonlinearity. In the model, shell elements are used and the nonlinear material characteristics are described by a stress-dependent tangent modulus $d\sigma/d\epsilon = a\sigma + b$ with the stiffening characteristic a and the Young's modulus b in the unloaded case. In the model, a is derived from uniaxial tensile tests of TM stripes and b is chosen according to measurements of the middle ear's transfer behaviour.

The annular ligament is modelled according to measurements from [4], who derived the stiffness characteristics of the annular ligament on human temporal bones. Therefore, the stapes annular ligament system was isolated by dissecting the chain. At several points on the stapes footplate the applied forces were increased step by step and the spatial displacement of the footplate was measured using a laser vibrometer. The derived inhomogeneous stiffness distribution of the annular ligament was modelled by eight translational and two rotational progressive nonlinear springs, distributed along the circumference of the stapes footplate.

We simulated a complete calcification of the superior malleolar ligament (SML) as a test case for the model-based diagnosis, where the ligament is grown together with the tympanum. In that case, the stiffening effect is high, because rotational degrees of freedom are severely affected. Therefore, in our simulation, we used a stiffening factor of 10000 for the SML for all translational and rotational stiffness values. In comparison, the normal stiffness value for the SML in its principal direction in our model is 700 N/m, which corresponds to a Young's Modulus of 9 MPa. The simulation of WBT is done in two steps. First a non-linear large deformation static analysis is conducted with pressure levels from -300 daPa to 200 daPa. The static pressure is applied in steps of 50 daPa on the tympanic membrane. Second, based on the static analysis a dynamic frequency response analysis is carried out in the frequency range from 0.2 to 8 kHz, with a uniform velocity excitation in the ear canal of $10^{-5}$ m/s at a distance of 8 mm from the umbo, see [8]. The energy absorbance (EA) is calculated from the velocity excitation and the resulting pressure averaged over the excitation plane. This is done for every static ear canal pressure level to calculate the 3D WBT diagrams.

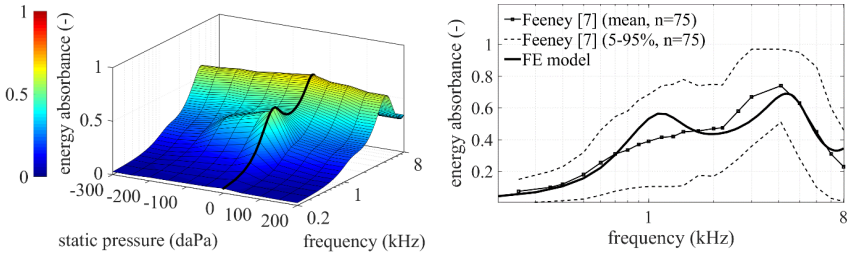



***Figure 2**: Simulated EA-pressure-frequency curve of a mean normal ear (left). The black line shows EA at static pressure of 0 daPa and is compared in the right figure to published measurement data on normal hearing patients.*

### 2.2 Inverse Fuzzy Arithmetic

The parameter identification by means of inverse fuzzy arithmetic represents an optimization problem with the model parameters as design variables. Its goal is to minimize the deviation between the values predicted by the model and the measured values. During modeling, assumptions and simplifications lead to errors that cause uncertainties in the statements made by the model. On the measurement side, the uncertainties arise mainly from the complexity and poor accessibility of the middle ear, which do not allow all necessary model states to be observed.

Both together lead to an ambiguity or uncertainty in the parameter identification, since different parameterizations lead to similarly plausible model responses. The application of inverse fuzzy arithmetic now makes it possible to quantify the uncertainties in the identification of the model parameters using fuzzy numbers. The parameter set with the smallest overall uncertainty in the identified parameters (smallest parameter scatter) is selected as the most plausible parameterization. In contrast to purely stochastic methods, fuzzy arithmetic allows a simple integration of expert knowledge, which is indispensable for the definition of parameter clusters. In addition, measurement uncertainties can be taken into account without having to have precise knowledge of the stochastic distribution functions.

The propagation of uncertainties described by fuzzy sets is based on fuzzy arithmetic which is fundamentally based on the extension principle by Zadeh [9]. The direct problem is described by a mapping of a fuzzy input set of the uncertain model parameters to a fuzzy valued output set. The transformation method introduced by [10] and used in this study is a numerical approximation of the extension principle and prevents the overestimation of uncertainties in the fuzzy output. The transformation method uses a so called α-cut approach, where the fuzzy sets are subdivided by α-cuts into several subsets with equally spaced membership levels and transformed into arrays that define possible parameter combinations. The evaluation may be done with crisp numbers and therefore the transformation can be applied on black-box models. Afterwards the generated output arrays are retransformed and the fuzzy-valued output is recomposed. A detailed description of this procedure can be found in [10]. This approach further enables sensitivity studies through an efficient calculation of influence measures of the uncertain model parameters on the system response. The inverse problem is to identify the input fuzzy sets $\mathbf{x}$ that generate a prescribed output. To reach a conservative coverage of the output, the uncertainty of the estimated sets needs to be maximized and the distance to the output set using the estimated input sets needs to be minimized. Using the α-cut approach, the input-output mapping is approximated by a linear mapping $\mathbf{F}$ and by introducing the tolerances $\mathbf{m}$ of the prescribed output, the quadratic program for the identification problem can be described as

$$\min_{\mathbf{x}} \ \mathbf{x}^{\mathrm{T}}\mathbf{F}^{\mathrm{T}}\mathbf{F}\mathbf{x} - 2\,\mathbf{m}^{\mathrm{T}}\mathbf{F}\mathbf{x} \quad \text{subject to} \ \mathbf{F}\mathbf{x} \leq \mathbf{m}.$$

### 2.3 Artificial Neural Network (ANN)

However, the inverse fuzzy approach needs to solve the direct problem first and therefore requires an initial guess for the fuzzy sets in advance, which in turn may influence the result. Furthermore, a costly optimization problem needs to be solved and this approach is additionally restricted to the transformation method. In the context of this paper, a novel concept is investigated, where an artificial neural network (ANN) determines the initial parameter set for the direct problem. These networks are capable to recognize patterns or characteristics based on graphical data like WBT contour plots. Apart from using standard multilayer perceptrons with two hidden layers, convolutional networks will be used. Convolutional networks are well known from deep learning and offer the advantage that preprocessing of the data is not necessary as the network itself can extract meaningful features from the data. Therefore, a two-step algorithm is proposed: The ANN assigns the individual ear to a rough class of middle ear characteristics associated with a certain initial parameter set by an initial assessment of the diagnostic measurements. Based on that, an inverse fuzzy arithmetic is used to make a fine adjustment and to determine the most plausible parameter set. The training and cross validation for the ANN is done with virtual patient WBT measurements obtained by simulations with the ear model. Interindividual variations are considered by a random selection of the parameter set from a predefined parameter domain. The performance of the ANN is tested at three categories "normal hearing", "pathology I: malleus fixation" and "pathology II: disarticulation".

## 3 Results

Fig. 3 shows three examples of simulated WBT diagrams out of the mentioned three ear categories. The differences between the categories are most evident in the shape of the 3D hill at the first resonance and result from an increased stiffness of the SML for pathology I and a reduced stiffness of the incudo-stapedial joint for pathology II, each by an average factor of 10000. The differences within the categories are due to a superposition of

significantly smaller stochastic parameter fluctuations. In this way a total of 100 data sets were randomly generated. In order to classify these data sets we used a four-layer perceptron with resilient backpropagation. The three hidden layers had 50, 50, and 10 neurons. The activation function at the hidden layers was a hyperbolic tangent and at the output layer, consisting of three neurons (one for each class) the softmax function, a smoothed version of the threshold function. The error measure was the cross entropy. The training goal was usually reached in less than 15 iterations. The ANN was trained on 40 randomly selected data sets and was then able to assign the remaining 60 data sets with almost 98% to the correct category. Details about the performance are given in the confusion matrix for the 60 test data sets in Fig. 3.

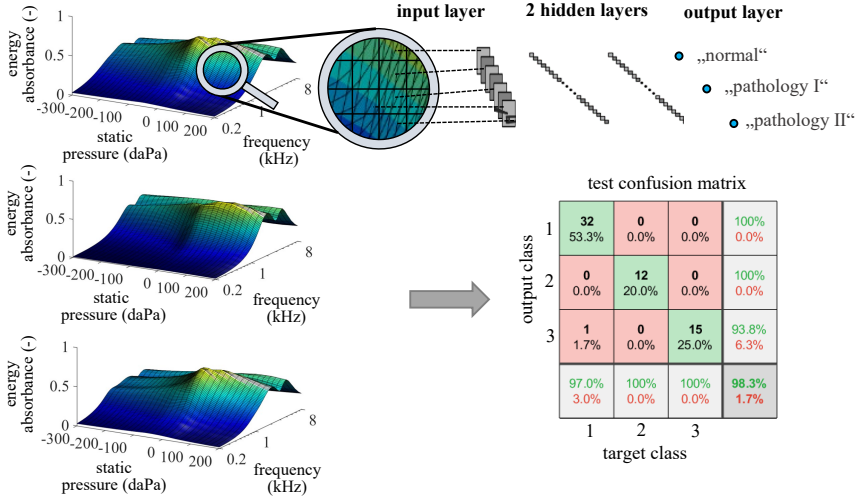



Figure 3: *Simulated WBT 3D-diagrams of a normal ear (top), pathological ear type I (middle) and type II (bottom). Testing on 60 simulated data sets out of the three categories resulted in the shown confusion matrix.*

After the rough classification with the ANN the fine-tuning of the parameters is carried out with inverse fuzzy arithmetic. Exemplarily it was done for the ear class “pathology I”, an ossified superior malleolar ligament. Since the fixation of the superior malleolar ligament (SML) was simulated with a stiffening factor of 10000, see section 2.1, the inverse fuzzy arithmetic is supposed to identify a fuzzy set including the nominal parameter value. Fig. 4 shows the membership function of the stiffening factor together with the nominal value. The narrow bound of the fuzzy set in Fig. 4 denotes that the stiffening parameter could be determined with a high degree of confidence. The shape of the identified fuzzy set consists of nested sets and therefore can be asymmetric. The nominal value of the parameter read meets nearly exactly the parameter value for a membership of one from the identified fuzzy set.

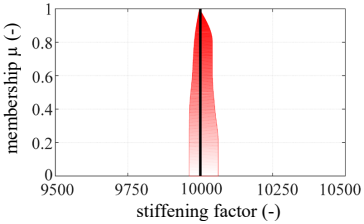



Figure 4: *Stiffening factor of the superior malleolar ligament derived with the FE model and inverse fuzzy. The nominal value is shown by a vertical line.*

## 4 Discussion

In contrast to the current clinical diagnosis, in which global middle ear characteristics are compared with normal ranges, the new parameter identification algorithm can resolve and quantify middle ear characteristics locally. In

clinical practice, this can be used to determine stage-specific hearing loss and localize pathologies to better plan middle ear surgery. In addition, a patient-specific model can be used for a tailored development of prostheses and hearing aids. In order to be suitable for clinical diagnosis, first, the model must be able to describe all anatomical changes through pathologies and inter-individual variations. Secondly, the model parameters need to be sensitive to specific pathological parameter changes to reliably predict them. The cause-effect-relationships between model parameters and ear characteristics can be identified by fuzzy arithmetic, while the inverse fuzzy arithmetic allows us to estimate the most likely parameter set. The reliability of the diagnosis may be valued by the fuzzy approach, as the worst-case intervals of the identified fuzzy numbers and fuzzy output can be used as a measure of the model validity. Furthermore, fuzzy arithmetic is especially suited to dealing with uncertainties introduced by inaccurate (mathematical) modeling due to lack of knowledge, simplification and idealization in modeling. Concerning the initial guess for the parameter set providing a basis for fuzzy arithmetic using ANN, we plan to leave out the detour determining a pathology class to the data set. We want to train a convolutional network for the regression task of delivering directly a parameter set. However, it is only possible to prove the validity of the prediction of model parameters by the parameter identification procedure by measurements on temporal bones. Therefore, laser vibrometer measurements of ossicles motions and WBT measurements should be carried out.

## 5 Conclusion

In this paper, we have introduced a new model-based approach based on the example of a FE ear model, which was used to simulate the WBT of an ear with malleus fixation. Through this example, we have shown the benefits of a model-based quantitative evaluation. The consideration of uncertainties due to model simplifications and incomplete knowledge of parameter distributions is done using fuzzy arithmetic and the parameter identification, with the help of inverse fuzzy and an artificial neural network. By considering these approaches, our example has given insight into the great potential of the model-based approach in providing us with a more reliable patient-specific quantitative evaluation of WBT measurements that may be extended to other diagnostic methods.

## Acknowledgements

This work has been partially funded by Volkswagen Foundation and a scholarship of the Ministry of Science, Research and Art Baden-Württemberg (MWK). This support is gratefully acknowledged.

## References

[1] G. R. Merchant, S. N. Merchant, J. J. Rosowski, H. H. Nakajima. *Controlled exploration of the effects of conductive hearing loss on wideband acoustic immittance in human cadaveric preparations.* Hearing Research, 341:19-30, 2016.

[2] S. E. Voss, G. R. Merchant, N. J. Horton. *Effects of middle-ear disorders on power reflectance measured in cadaveric ear canals.* Ear and Hearing, 33(2), 195, 2012.

[3] S. Ihrle, M. Lauxmann, A. Eiber, P. Eberhard. *Nonlinear modelling of the middle ear as an elastic multibody system - applying model order reduction to acousto-structural coupled systems.* Journal of Computational and Applied Mathematics, 246, 18–26, 2013.

[4] M. Lauxmann, A. Eiber, F. Haag, S. Ihrle. *Nonlinear stiffness characteristics of the annular ligament.* The Journal of the Acoustical Society of America, 136, 1756–1767, 2014.

[5] S. Ihrle, A. Eiber, P. Eberhard. *Modeling of the incudo-malleolar joint within a biomechanical model of the human ear.* Multibody System Dynamics, 39(4), 291-310, 2017.

[6] B. Sackmann, E. Dalhoff, M. Lauxmann. *Model-based hearing diagnostics based on wideband tympanometry measurements utilizing fuzzy arithmetic*. Hearing Research, 378, 126-138, 2019.

[7] M. P. Feeney, I. L. Grant, L. P. Marryott. *Wideband energy reflectance measurements in adults with middle-ear disorders.* Journal of Speech, Language, and Hearing Research 46, 901–911, 2003.

[8] B. Sackmann, B. Warnholtz, J. H. Sim, D. Burovikhin, E. Dalhoff, P. Eberhard, M. Lauxmann. *Investigation of tympanic membrane influences on middle-ear impedance measurements and simulations.* Computational Methods in Applied Sciences, 53, 3-10, 2020.

[9] L. Zadeh. *Fuzzy sets as a basis for a theory of possibility.* Fuzzy Sets Syst. 100 (Suppl. 1), 9-34, 1999.

[10] M. Hanss. *Applied Fuzzy Arithmetic - an Introduction with Engineering Applications*. ISBN 978-3-540-27317-2, Springer, Berlin, Heidelberg, 2005.

# Parameter identification of a human stapedial annular ligament model in the context of a model-based hearing diagnosis of the human middle ear

*D. Burovikhin[1], B. Sackmann[1], M. Schär[2], J. H. Sim[2], M. Lauxmann[1]*

[1] *Reutlingen Research Institute, Reutlingen University, Reutlingen, Germany*
[2] *Otology and Biomechanics of Hearing, University Hospital Zürich, Zürich, Switzerland*

Contact: Dmitrii.Burovikhin@Reutlingen-University.de

***Abstract***

*This study describes a non-contact measuring and parameter identification procedure designed to evaluate inhomogeneous stiffness and damping characteristics of the annular ligament in the physiological amplitude and frequency range without the application of large static external forces that can cause unnatural displacements of the stapes. To verify the procedure, measurements were first conducted on a steel beam. Then, measurements on an individual human cadaveric temporal bone sample were performed. The estimated results support the inhomogeneous stiffness and damping distribution of the annular ligament and are in a good agreement with the multiphoton microscopy results which show that the posterior-inferior corner of the stapes footplate is the stiffest region of the annular ligament. This method can potentially help to establish a correlation between stiffness and damping characteristics of the annular ligament and inertia properties of the stapes and, thus, help to reduce the number of independent parameters in the model-based hearing diagnosis.*

**Keywords**: middle ear, model-based hearing diagnosis, annular ligament model, parameter identification procedure

## 1 Problem

The model-based hearing diagnosis of the human middle ear is an innovative approach in modern medical diagnostics. Due to the complexity of the middle ear model and uncertainties that arise from the model and measurements, ambiguities occur during parameter identification. To avoid that, the model parameters are correlated to each other by means of anatomically conditioned links which reduces the number of independent parameters. For example, the area ratio between the eardrum and stapes footplate and the lever ratios in the ossicular chain show relatively small fluctuations from one person to another, despite the fact that the absolute anatomical dimensions vary significantly. Therefore, the individual anatomical properties and, thus, the associated model parameters cannot vary their values arbitrarily, but rather correlate to each other and can be combined into characteristic parameter clusters, which can significantly reduce the number of independent parameters. As another example, the first middle ear resonant frequency of the stapes transfer function shows that interindividual variances in the range of 800 to 1100 Hz are much smaller than expected from the variations in material properties and anatomical dimensions of individual ligaments and ossicles. This leads to the conclusion that the stiffness properties of the ligaments and joints depend on the masses and inertia properties of the ossicles. It remains challenging, however, to quantify these relationships. This study proposes a measuring and parameter identification procedure designed to evaluate stiffness and damping characteristics of the annular ligament which can potentially help to find a correlation between stiffness and damping characteristics of the annular ligament and inertia properties of the stapes and reduce the number of independent parameters in the model-based hearing diagnosis.

The stapes is a bone found in the middle ear of humans and other mammals which is involved in the conduction of sound vibrations to the inner ear, see Fig. 1. The stapedial annular ligament (AL) is a ring of fibrous tissue that connects the base of the stapes, its footplate, to the oval window of the inner ear. The anatomical dimensions of the AL support the hypothesis of an inhomogeneous stiffness distribution. In [3] it has been revealed that the cross-section of the AL is posteriorly narrower and thicker, resulting in a higher stiffness on the posterior side, and in [1] it is stated that the properties of the AL largely determine the transfer characteristic of the middle ear in the lower frequency range.

A number of studies on the topic of the stiffness characteristics of the human stapedial AL were conducted in the past, but all of them were focused on determining the stiffness properties of the AL in the quasi-static frequency range by applying large, mostly static, external forces to the stapes, which caused displacements beyond the physiological range and, as a consequence, introduced artefacts that hid underlying physical properties of the AL [1, 2, 4, 7]. Moreover, so far the inhomogeneous stiffness and damping properties of the AL have not been investigated in the dynamic frequency range.

The measurement and parameter identification procedure described here offers a non-contact way to determine the stiffness and damping characteristics of the AL within the physiological amplitude and frequency range by using the inertia forces of the stapes itself instead of large external forces. The inertia properties of the stapes were determined with the help of a micro-CT scanner. Based on that the inhomogeneous stiffness and damping characteristics of the AL

were estimated within the physiological frequency range. We verified the procedure by conducting measurements on a theoretically well-known mechanical structure similar to that of the stapes-AL system, where the AL was represented by a steel beam and the stapes by a small steel cube. The introduced system identification procedure can potentially help to find correlations between the inertia properties of the stapes and the stiffness and damping properties of the AL among different human cadaveric temporal bone samples.

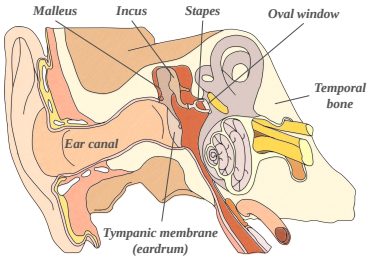



Figure 1: *Anatomical structure of the human ear.*

## 2 Material and Methods

The stapes and the peripheral bone (PB) can be modelled as two rigid bodies connected by the AL modelled as a spring/damper combination, as shown schematically in Fig 2. The stapes is represented by its generalized mass matrix **M** which is referred to the centre of its footplate and obtained by scanning the stapes and the plastic plate, to which the temporal bone sample is glued, with a micro-CT scanner [8]. The plastic plate serves as a reference for the scanner. The mass of the AL can be neglected. The stapes is subjected to the base excitation from the PB mounted on a shaker table. The spatial motion of the stapes and the PB is measured by a 3D laser Doppler vibrometer (3D-LDV) at multiple points and the generalized coordinates for each rigid body are calculated with reference to the centre of the stapes footplate. The stapes-AL system is theoretically unknown, which makes it difficult to confirm the validity of our system

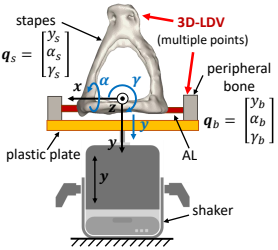



Figure 2: *Stapes - measurement setup.*

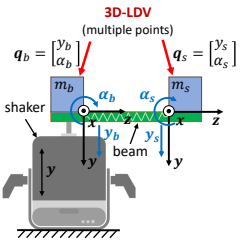



Figure 3: *Beam - measurement setup.*

identification procedure. Therefore, it was first verified on a theoretically well-known mechanical structure consisting of two $10 \times 10$ mm steel cubes with masses $m_b$ and $m_s$ attached to a steel beam. Each cube, together with the part of the beam underneath it, is modelled as a rigid body. One of these rigid bodies represents the PB and is mounted on the shaker table. The other represents the stapes, which is located at the free end of the beam, and is characterized by its generalized matrix **M**. The part of the beam between the cubes represents the AL. Its mass can be neglected compared to the mass of the rigid bodies, and thus, it is modelled only with regard to its stiffness and damping properties. The spatial motion of the cubes is measured by the 3D-LDV system at multiple measurement points and their generalized coordinates are calculated with reference to the origin of their axes.

Both the stapes-AL system and the beam system represent a structure subjected to base excitation whose dynamic equation of motion is given by

$$\mathbf{M} \cdot \ddot{\mathbf{q}}_s(t) + \mathbf{D} \cdot \dot{\mathbf{q}}_s(t) + \mathbf{K} \cdot \mathbf{q}_s(t) = \mathbf{D} \cdot \dot{\mathbf{q}}_b(t) + \mathbf{K} \cdot \mathbf{q}_b(t), \tag{1}$$

with the generalized displacement vectors of the structure $\mathbf{q}_s(t)$ and the base $\mathbf{q}_b(t)$, generalized symmetric viscoelastic damping matrix $\mathbf{D}$, and the generalized symmetric stiffness matrix $\mathbf{K}$.

The assumption of harmonic base excitation at $k$ number of frequencies $\omega_j$ and steady state response yields

$$\omega_j^2 \, \mathbf{M} \cdot \mathbf{q}_s(\omega_j) = (\mathbf{K} + i\omega_j \, \mathbf{D}) \cdot \Delta\mathbf{q}(\omega_j) \; , \quad j = 1,2,...,k, \tag{2}$$

with the relative displacement vector $\Delta\mathbf{q}(\omega_j) = \mathbf{q}_s(\omega_j) - \mathbf{q}_b(\omega_j)$. Having in mind that $\Delta\mathbf{q}(\omega_j) = \Delta\mathbf{q}'(\omega_j) + i\Delta\mathbf{q}''(\omega_j)$ and $\mathbf{q}_s(\omega_j) = \mathbf{q}_s'(\omega_j) + i\mathbf{q}_s''(\omega_j)$, the amplitude and phase of each measured point at the $j$-th excitation frequency $\omega_j$ are transformed to the real and imaginary part of the steady-state response

$$\begin{aligned} \omega_j^2 \, \mathbf{M} \cdot \mathbf{q}_s'(\omega_j) &= \mathbf{K} \cdot \Delta\mathbf{q}'(\omega_j) - \omega_j \, \mathbf{D} \cdot \Delta\mathbf{q}''(\omega_j) \; , \\ \omega_j^2 \, \mathbf{M} \cdot \mathbf{q}_s''(\omega_j) &= \mathbf{K} \cdot \Delta\mathbf{q}''(\omega_j) + \omega_j \, \mathbf{D} \cdot \Delta\mathbf{q}'(\omega_j). \end{aligned} \tag{3}$$

Equations 3 can be written as

$$\mathbf{A}_j \cdot \mathbf{p} = \mathbf{b}_j \; , \quad j = 1,2,...,k, \tag{4}$$

with a design parameter vector $\mathbf{p}$ containing the unknown stiffness and damping coefficients of matrices $\mathbf{K}$ and $\mathbf{D}$. For k number of frequencies Eq. 4 yields an overdetermined system of linear equations, see Eq. 5, that can be solved with the least squares method by calculating the pseudo inverse $\mathbf{A}^I$ of matrix $\mathbf{A}$ to avoid singularity in matrix $\mathbf{A}$ caused by measurement noise

$$\underbrace{\begin{bmatrix} \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{A}_j \end{bmatrix}}_{\mathbf{A}} \cdot \mathbf{p} = \underbrace{\begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_j \end{bmatrix}}_{\mathbf{b}} \implies \mathbf{p} = \mathbf{A}^I \cdot \mathbf{b}. \tag{5}$$

The stapes-AL system is more complex than that of the beam, partly, because it is more damped, and partly because of larger measurement errors caused by misalignment and measurement noise. Because of that solving Eq. 5 for the stapes-AL system using an unconstrained least squares solution leads to large estimation errors, since, in this case, a local minimum found by LSQ algorithm can become the global one. Instead, we can iteratively minimize the function $|\mathbf{b} - \mathbf{A} \cdot \mathbf{p}|^2$ by solving a constrained optimization problem with the help of MATLAB's *fmincon* programming solver. From [6] the signs of the stiffness coefficients and their plausible range can be determined and the stiffness and damping coefficients can be constrained in a range where the system is stable. The stability of the system is determined by the eigenvalues criterion meaning that the real part of the eigenvalues of the system estimated with these coefficients should be negative in order for the system to be stable.

# 3 Results

## 3.1 The Beam Experiment Results

The motion of the beam system has two degrees of freedom (DOF), one translational $y_s$ along $y$-axis and one rotational $\alpha_s$ around $x$-axis, see Fig. 3. From the differential equation of the cantilever beam, the theoretical stiffness matrix of the beam can be derived and is given by

$$\frac{1}{EI} \begin{bmatrix} \frac{l^3}{3} & -\frac{l^2}{2} \\ -\frac{l^2}{2} & l \end{bmatrix} \cdot \begin{bmatrix} F \\ M \end{bmatrix} = \begin{bmatrix} y_s \\ \alpha_s \end{bmatrix} \implies \mathbf{K}_{theory} = \begin{bmatrix} c_{1,th} & c_{2,th} \\ c_{2,th} & c_{3,th} \end{bmatrix} = EI \begin{bmatrix} \frac{12}{l^3} & \frac{6}{l^2} \\ \frac{6}{l^2} & \frac{4}{l} \end{bmatrix} \tag{6}$$

with force $F$ and rotational moment $M$ applied at the free end of the beam. For the beam parameters, see Table 1.

Table 1: *Beam parameters*

| Parameter | Value | Unit |
|---|---|---|
| Young's modulus, $E$ | 210000 | $\frac{N}{mm^2}$ |
| Moment of inertia, $I = \frac{bh^3}{12}$ | 0.44287 | $mm^4$ |
| Height of the beam, $h$ | 0.81 | $mm$ |
| Width of the beam, $b$ | 10 | $mm$ |
| Length of the beam, $l$ | 48 | $mm$ |

For a beam with two DOF, Eq. 2 can be written as

$$\omega_j^2 \mathbf{M} \cdot \underbrace{\begin{bmatrix} y_s(\omega_j) \\ \alpha_s(\omega_j) \end{bmatrix}}_{\mathbf{q}_s(\omega_j)} = \left( \underbrace{\begin{bmatrix} c_1 & c_2 \\ c_2 & c_3 \end{bmatrix}}_{\mathbf{K}} + i\omega_j \underbrace{\begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix}}_{\mathbf{D}} \right) \cdot \underbrace{\begin{bmatrix} y_s(\omega_j) - y_b(\omega_j) + l\alpha_b(\omega_j) \\ \alpha_s(\omega_j) - \alpha_b(\omega_j) \end{bmatrix}}_{\Delta\mathbf{q}(\omega_j)}. \tag{7}$$

Equation 7 can be written in the form of Eq. 5 and solved for the unknown stiffness and damping coefficients. The generalized mass matrix $\mathbf{M}$ is calculated from the geometric parameters given in Table 1. To verify the results, the motion of the free end of the beam is simulated using estimated $\mathbf{K}_{est}$ and $\mathbf{D}_{est}$ matrices by calculating the complex frequency response of a structure subjected to base excitation which is given by

$$\mathbf{H}_j = (-\omega_j^2 \mathbf{M} + i\omega_j \mathbf{D}_{est} + \mathbf{K}_{est})^{-1} \cdot (i\omega_j \mathbf{D}_{est} + \mathbf{K}_{est}), \quad j = 1,2,...,k, \tag{8}$$

with the $n \times n$ frequency response matrix $\mathbf{H}_j$ for each frequency $j$, where $n$ is the number of DOF, which in the case of the beam equals two.

Figure 4 shows a comparison between the measured and simulated motions of the free end $y_s$ and $\alpha_s$ related to the measured $y$-motion of the base $y_{b\,mes}$ in order to normalize the results. This is done because $y_{b\,mes}$ is the dominant motion, since the shaker table excites the structure primarily in the $y$-direction. The solid lines show the measured motion and the dashed lines show the estimated motion. To quantitatively compare the estimated and theoretical stiffness coefficients, a deviation error in % is calculated

$$\text{deviation} = \begin{bmatrix} (c_{1,est} - c_{1,th})/c_{1,th} \\ (c_{2,est} - c_{2,th})/c_{2,th} \\ (c_{3,est} - c_{3,th})/c_{3,th} \end{bmatrix} 100\% = \begin{bmatrix} 3.17 \\ 1.28 \\ 0.38 \end{bmatrix} \%. \tag{9}$$

A misalignment of the centre of mass of the free end of 100 $\mu$m in the $z$-direction can already explain the deviation error of about 3% in estimating the stiffness values as well as an error of about 1% in estimating the resonance frequencies of the beam, see Fig. 4. This shows the validity of the proposed system identification procedure.

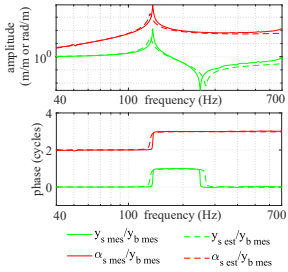



Figure 4: *Beam - comparison between measured and estimated (simulated) data.*

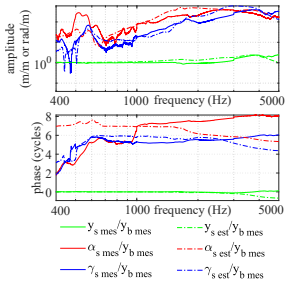



Figure 5: *Stapes - comparison between measured and estimated (simulated) data.*

## 3.2 The Stapes Experiment Results

It is generally accepted that the motion of the stapes has three DOF, one translational piston-like motion $y_s$ along the $y$-axis, and two rotational motions $\alpha_s$ and $\gamma_s$ around the $x$-axis and $z$-axis, respectively, as shown in [5].

For a stapes system with three DOF, Eq. 2 can be written as

$$\omega_j^2 \mathbf{M} \cdot \underbrace{\begin{bmatrix} y_s(\omega_j) \\ \alpha_s(\omega_j) \\ \gamma_s(\omega_j) \end{bmatrix}}_{\mathbf{q}_s(\omega_j)} = \left( \underbrace{\begin{bmatrix} c_1 & c_2 & c_3 \\ c_2 & c_4 & c_5 \\ c_3 & c_5 & c_6 \end{bmatrix}}_{\mathbf{K}} + i\omega_j \underbrace{\begin{bmatrix} d_1 & d_2 & d_3 \\ d_2 & d_4 & d_5 \\ d_3 & d_5 & d_6 \end{bmatrix}}_{\mathbf{D}} \right) \cdot \underbrace{\begin{bmatrix} y_s(\omega_j) - y_b(\omega_j) \\ \alpha_s(\omega_j) - \alpha_b(\omega_j) \\ \gamma_s(\omega_j) - \gamma_b(\omega_j) \end{bmatrix}}_{\Delta\mathbf{q}(\omega_j)}. \tag{10}$$

As in the case of the beam, Eq. 10 can be written in the form of Eq. 5 and solved for the unknown stiffness and damping coefficients. The generalized mass matrix of the stapes $\mathbf{M}$, obtained with the help of a micro-CT scanner, reads

$$\mathbf{M} = \begin{bmatrix} 3.2 \cdot 10^{-6}\, kg & -434.88 \cdot 10^{-12}\, kg \cdot m & 143.04 \cdot 10^{-12}\, kg \cdot m \\ -434.88 \cdot 10^{-12}\, kg \cdot m & 9.77 \cdot 10^{-12}\, kg \cdot m^2 & -29.7 \cdot 10^{-15}\, kg \cdot m^2 \\ 143.04 \cdot 10^{-12}\, kg \cdot m & -29.7 \cdot 10^{-15}\, kg \cdot m^2 & 11.14 \cdot 10^{-12}\, kg \cdot m^2 \end{bmatrix}.$$

Just like in the beam experiment, to verify the results, the frequency response is simulated for $\mathbf{K}_{est}$ and $\mathbf{D}_{est}$, see Eq. 8. Figure 5 shows a comparison between the measured and simulated motions of the stapes $y_s$, $\alpha_s$ and $\gamma_s$ related to the measured $y$-motion of the base $y_{b\,mes}$. The solid lines show the measured motion and the dashed lines show the estimated motion. The estimated stiffness and damping matrices read

$$\mathbf{K}_{est} = \begin{bmatrix} 1722\,\frac{N}{m} & 0.414\,\frac{N}{rad} & -0.48\,\frac{N}{rad} \\ 0.414\,N & 1.37\cdot10^{-3}\,\frac{Nm}{rad} & -6.41\cdot10^{-5}\,\frac{Nm}{rad} \\ -0.48\,N & -6.41\cdot10^{-5}\,\frac{N\cdot m}{rad} & 3.79\cdot10^{-3}\,\frac{N\cdot m}{rad} \end{bmatrix} \text{ and}$$

$$\mathbf{D}_{est} = \begin{bmatrix} 22.89\cdot10^{-3}\,\frac{N\cdot s}{m} & 61.42\cdot10^{-5}\,\frac{N\cdot s}{rad} & -36.79\cdot10^{-5}\,\frac{N\cdot s}{rad} \\ 61.42\cdot10^{-5}\,N\cdot s & 30.15\cdot10^{-9}\,\frac{N\cdot s\cdot m}{rad} & -67.56\cdot10^{-13}\,\frac{N\cdot s\cdot m}{rad} \\ -36.79\cdot10^{-5}\,N\cdot s & -67.56\cdot10^{-13}\,\frac{N\cdot s\cdot m}{rad} & 56.46\cdot10^{-9}\,\frac{N\cdot s\cdot m}{rad} \end{bmatrix}.$$

# 4 Discussion

To further evaluate stiffness characteristics of the AL, a principal axes transformation of the estimated stiffness matrix $\mathbf{K}_{est}$ is performed which yields a diagonal stiffness matrix that reads

$$\mathbf{K}_{est,prin} = \begin{bmatrix} 1722\,\frac{N}{m} & 0 & 0 \\ 0 & 1.26\cdot10^{-3}\,\frac{Nm}{rad} & 0 \\ 0 & 0 & 3.65\cdot10^{-3}\,\frac{N\cdot m}{rad} \end{bmatrix}.$$

According to [7] the mean stiffness value in the $y$-direction is 1238 $\frac{N}{m}$ with a standard deviation of 973 $\frac{N}{m}$. The estimated stiffness value of 1722 $\frac{N}{m}$ in the $y$-direction found in the upper left corner of $\mathbf{K}_{est,prin}$ fits well in that range. The principle axes are located in the posterior-inferior corner of the stapes footplate indicating that it is the stiffest region of the AL, see Fig. 6. If a force is applied to the stapes at the origin of the principle axes in the y-direction, the stapes will be moved in the y-direction only.

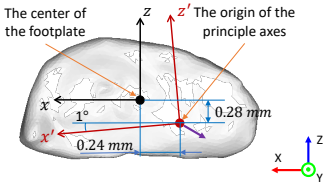



Figure 6: *The origin of the principal axes.*

It was also shown in [6] that applying a centric force orientated in the lateral direction (the y-direction) to the center of the stapes footplate moves the stapes head laterally and in the posterior-inferior direction, which means that the posterior-inferior corner of the stapes footplate is the stiffest region of the AL. This type of spatial motion supports the inhomogeneous stiffness distribution of the AL and corresponds to the signs of the stiffness coefficients estimated here. The same results were observed in [3], where the morphometry of an intact human AL along the full boundary of the annular rim was conducted via multiphoton microscopy revealing cross sections of the AL resembling a sandwich-composite structure. For this structure, the main effect on the variation in bending stiffness originates from thickness variations of the core layer, and not the variation in thickness of the face layers. While the total thickness and core thickness are highest around the posterior and anterior poles, the AL is found to be narrowest in the posterior-inferior region, and widest in the anterior region.

Since a viscoelastic damping model was used in this study, the eigenvectors are complex. In order to interpret them, the absolute values of the eigenvector components were calculated. To derive the motion of the stapes head in the $x$-, $y$-, and $z$-directions, $\alpha$- and $\gamma$-components of the eigenvectors can be multiplied by the height of the stapes (approximately 3 mm). Based on this information it can be concluded that in the first normal mode the $\alpha$-motion is dominant, in the second the $\gamma$-motion is dominant, and in the third the piston-like $y$-motion is dominant. The modal parameters of the estimated system can be found in Table 2.

Within the frequency range, where our system was estimated (400 Hz to 5 kHz), the maximum displacement amplitude in the $y$-direction is 148 nm, which is lower than the expected displacement amplitude of 150 nm measured with acoustic excitation at 100 dB SPL as shown in [5]. This means that the displacement amplitudes used in this study fit within the physiological range.

Table 2: *Modal parameters of the estimated system*

| Parameters | 1st normal mode | 2nd normal mode | 3rd normal mode |
|---|---|---|---|
| Resonant frequencies | $f_1 = 1753\,Hz$ | $f_2 = 2783\,Hz$ | $f_3 = 3862\,Hz$ |
| Eigenvectors | $\begin{bmatrix} \lvert\phi_{1y}\rvert \\ \lvert\phi_{1\alpha}\rvert \\ \lvert\phi_{1\gamma}\rvert \end{bmatrix} = \begin{bmatrix} 33.5\,nm \\ 90\,\mu rad \\ 4.6\,\mu rad \end{bmatrix}$ | $\begin{bmatrix} \lvert\phi_{2y}\rvert \\ \lvert\phi_{2\alpha}\rvert \\ \lvert\phi_{2\gamma}\rvert \end{bmatrix} = \begin{bmatrix} 33.4\,nm \\ 9.2\,\mu rad \\ 56\,\mu rad \end{bmatrix}$ | $\begin{bmatrix} \lvert\phi_{3y}\rvert \\ \lvert\phi_{3\alpha}\rvert \\ \lvert\phi_{3\gamma}\rvert \end{bmatrix} = \begin{bmatrix} 164\,nm \\ 25\,\mu rad \\ 32\,\mu rad \end{bmatrix}$ |
| Damping ratios | $\zeta_1 = 0.12$ | $\zeta_2 = 0.14$ | $\zeta_3 = 0.18$ |

There are several reasons why the measured and estimated curves in Fig. 5 do not match precisely. We assume a three DOF system, however, due to the shaker excitation, which is qualitatively different to an acoustic excitation used in [5], in-plane motions, especially *x*- and *y*-translations, can become too large to be neglected and may cause estimation errors. Thus, a further investigation of a stapes-AL system with 5 and 6 DOF may provide more accurate results. In addition, the inaccuracies in alignment of the stapes reference frame with respect to the plastic plate and to the 3D-LDV system lead to further evaluation errors. In order to reduce those uncertainties, the mentioned error sources should be systematically investigated.

## 5 Conclusion

In this study a non-contact measuring and parameter identification procedure for evaluating inhomogeneous stiffness and damping characteristics of the AL is introduced. What makes this procedure unique is that it estimates the stiffness and damping coefficients in the physiological amplitude and frequency range without the application of large static external forces that can cause unnatural displacements of the stapes. To verify the procedure, measurements were first conducted on a steel beam. Then measurements on an individual human cadaveric temporal bone sample were performed. The estimated results support the inhomogeneous stiffness and damping distribution of the AL and are in a good agreement with the multiphoton microscopy results obtained in [3] where it was concluded that the posterior-inferior corner of the stapes footplate is the stiffest region of the AL. The procedure itself, however, requires further investigation with regard to the mentioned error sources, but it can potentially help to establish a correlation between stiffness and damping characteristics of the annular ligament and inertia properties of the stapes and, thus, help to reduce the number of independent parameters in the model-based hearing diagnosis.

## References

[1] von Békésy G, Experiments in hearing. McGraw-Hill, New York, pp. 1514-1523 (1960)

[2] Gan RZ, Yang F, Zhang X, Nakmali D, Mechanical properties of stapedial annular ligament. Medical Engineering & Physics, vol 33, pp. 330-339 (2011)

[3] Schär M, Dobrev I, Chatzimichalis M, Röösli C, Sim JH, Multiphoton imaging for morphometry of the sandwich-beam structure of the human stapedial annular ligament. Hearing Research, doi: 10.1016/j.heares.2018.11.011 (2019)

[4] Helmholtz H, Die Mechanik der Gehörknöchelchen und des Trommelfells. Max Cohen & Sohn, Bonn, pp. 34-37, (1869)

[5] Lauxmann M, Sim JH, Chatzimichailis M, Heckeler C, Ihrle S, Huber A, and Eiber A, In-plane motions of the stapes in human ears. Journal of the Acoustical Society of America, vol 132, pp. 3280-3291, (2012)

[6] Lauxmann M, Eiber A, Haag F, Ihrle S, Nonlinear stiffness characteristics of the annular ligament. Journal of the Acoustical Society of America, vol 136, pp. 1756-1767 (2014)

[7] Waller T-S, Dynamic measurement of the circular stapes ligamentum using electrostatic forces. Ph.D. thesis, Faculty of Medicine of the Bavarian Julius Maximilian University of Würzburg, pp. 28-29 (2002)

[8] Sim JH, Puria S, Steele CR, Calculation of the inertial properties of the malleus-incus complex from micro-CT imaging. Journal of Mechanics of Materials and Structures, vol 2 (8), pp. 1515-1524 (2007)

# Computersimulation patientenindividueller Lendenwirbelsäulen am Beispiel degenerativer Spondylolisthesen

*Kosterhon M.[1], MüllerA.[2], Damm N. [2], Ringel F.[1], Gruber K.[2], Kantelhardt S.R.[1]*
*1Universitätsmedizin Mainz, Klinik und Poliklinik für Neurochirurgie, Mainz, Deutschland*
*2Universität Koblenz Landau, MTI Mittelrhein - Institut für Medizintechnik und Informationsverarbeitung, Koblenz, Deutschland*

Kontakt: mikoster@uni-mainz.de

***Abstract***

*Für die Planung von Wirbelsäulenoperationen stehen in der Regel lediglich diagnostische Mittel wie CT, MRT und Röntgenaufnahmen zur Verfügung, die dieses bewegliche Organ in einer Momentaufnahme abbilden. Dynamische Untersuchungsmethoden, die funktionelle Parameter und die Biomechanik berücksichtigen, könnten entscheidende Verbesserungen darstellen.*

*Es wurde ein biomechanisches Computersimulationsmodell erstellt, welches individuell auf die Wirbelsäulen einzelner Patienten angepasst werden kann.*

*Dieses Modell wurde auf Röntgenfunktionsaufnahmen (in Neutralstellung, Inklination und Reklination) angepasst und die fehlenden Zwischenschritte durch Simulation ergänzt. So entstand ein flüssiger, biomechanischer Bewegungsablauf eines individuellen Patienten.*

*Die Technik wurde erstmals bei 9 Patienten mit degenerativer Spinalkanalstenose und Wirbelgleiten klinisch eingesetzt um den Bewegungsablauf einer Inklination nachzusimulieren.*

*Dabei wurden die Auswirkungen zweier operativer Dekompressionsmethoden (Interlaminäre Fensterung und Laminektomie) hinsichtlich ihres destabilisierenden Einflusses auf die Bewegung untersucht.*

*Dabei zeigte sich, dass die invasivere Laminektomie im Mittel eine mehr als doppelt so große destabilisierende Wirkung hinsichtlich des Wirbelgleitens aufweist, als die gewebeschonendere interlaminären Fensterung.*

**Keywords**: Computersimulation, Biomechanik, Wirbelsäule, Spondylolisthesis, Spinalkanalstenose, Operationsplanung, individuelle Patientenversorgung,

## 1 Problemstellung

Rückenschmerzen zählen zu den häufigsten Erkrankungen des Menschen [1-3]. Insbesondere mit zunehmendem Alter sind hierfür zahlreiche degenerative Prozesse verantwortlich. Neben der konservativen Therapie insbesondere für unspezifische Rückenschmerzen, stellt für spezifische Rückenschmerzen mit organischem Korrelat die operative Behandlung einen wichtigen Pfeiler der Therapie dar [4]. Bei der Entscheidungsfindung für ein geeignetes operatives Verfahren stehen Medizinern allerdings nur eingeschränkte diagnostische Mittel wie CT, MRT und Röntgenaufnahmen zur Verfügung, die das dynamisch komplexe Verhalten der Wirbelsäule in einer Momentaufnahme statisch abbilden. Dynamische Untersuchungsmethoden, die funktionelle Parameter und die Biomechanik mit einbeziehen, könnten hier entscheidende Verbesserungen darstellen.

Computersimulationsmodelle könnten diese diagnostische Lücke schließen und helfen, biomechanisch besser begründbare Therapieentscheidungen zu treffen und funktionelle Parameter mit in die Entscheidungsfindung einzubeziehen [5].

Ein Beispiel stellt die Behandlung der degenerativen Spinalkanalstenose (Wirbelkanaleinengung) mit begleitender Spondylolisthesis (Wirbelgleiten) der Lendenwirbelsäule (LWS) dar. Durch Degeneration und Lockerung der Bandscheiben und Bänder kommt es hier zu einem vermehrten Ventralversatz eines Wirbelkörpers über den anderen, wodurch der Spinalkanal im Querschnitt eingeengt wird. Als Gegenreaktion hypertrophieren zusätzlich andere Haltestrukturen, wie die Ligamenta flava und verstärken so die Stenose im Spinalkanal [4]. Die Folge sind eine Kompression der nervalen Strukturen und starke Schmerzen. Die operative Therapie dieser Erkrankung besteht aus der Dekompression des Spinalkanals mit teilweisem oder ganzem Entfernen der begrenzenden Strukturen wie Knochen und Lig. flava. Erreicht werden kann dies durch eine Interlaminäre Fensterung (ILF) oder Laminektomie (LAM). Der Therapieeffekt wird jedoch insbesondere bei der invasiveren Methode der LAM mit

einer zunehmenden Destabilisierung der Wirbelsäule erkauft. Diese Destabilisierung wiederum kann erneute Schmerzen verursachen [6] und als Folge muss entweder direkt oder sekundär eine Stabilisierung mittels einer Spondylodese (Schrauben-Stab-System, das zwei oder mehr Wirbelkörper rigide miteinander verbindet) durchgeführt werden. In der einschlägigen Fachliteratur wird daher immer wieder die Notwendigkeit einer solchen zusätzlichen Spondylodese diskutiert [7, 8]. Es scheint, dass ein gewisser Teil der Patienten von einer zusätzlichen Stabilisierung profitiert [8]. Zusätzlich verkompliziert wird die Entscheidung dadurch, dass nicht nur individuelle biomechanische Eigenschaften, sondern offensichtlich auch der Grad der Destabilisierung durch die gewählte Technik der Dekompression einen Einfluss hat. Insbesondere für die Technik der ILF gibt es noch keine eindeutigen Daten, in wie fern der Eingriff die Wirbelsäule destabilisiert [9]. Bei der Frage, ob nur eine reine Dekompression oder auch eine zusätzliche Stabilisierung erfolgen sollte bzw. bei der Auswahl, welche der Patienten von einer zusätzlichen Stabilisierung profitieren würden und welche nicht, wären Entscheidungshilfen, die individuelle biomechanische Voraussetzungen berücksichtigen, also eine große Hilfe für die behandelnden Chirurgen.

Die patientenindividuelle, biomechanische Simulation von Bewegungsabläufen könnte hier helfen eine fundiertere Entscheidung zu treffen, indem zum einen die individuelle Beweglichkeit der Wirbelsäule per se, zum anderen der Einfluss der gewählten Dekompressionstechnik in diesem individuellen Fall, vorab simuliert werden können.

Ziel dieses Projektes ist es, ein Computersimulationsmodell der LWS zu entwickeln, das sich unter Verwendung der in der Klinik verfügbaren Diagnostika, wie CT, MRT und Röntgen(funktions)aufnahmen auf den individuellen Patienten anpassen lässt. Anschließend wird der Bewegungsablauf einer Inklination oder Reklination nachsimuliert. Als Eingabeparameter dienen hier die im klinischen Alltag häufig angefertigten Röntgenfunktionsaufnahmen, die die LWS im Zustand Neutral, In- und Reklination zeigen. Aus strahlenschutztechnischen Gründen, werden hier nur die jeweiligen Endstellungen abgebildet, sodass eine Beurteilung des Bewegungsablaufes nur schwer möglich ist. Die Simulation soll die zwischenliegenden Schritte biomechanisch sinnvoll interpolieren und so die gängigen statischen Untersuchungsmethoden um die dynamische Komponente erweitern, ohne dass eine zusätzlich für den Patienten schädliche Bildgebung erfolgen muss.

Exemplarisch wurden hierfür retrospektiv 9 Patienten mit einer degenerativen Spinalkanalstenose und begleitender Spondylolisthesis, die im Zeitraum von 01.01.2017 bis 30.07.2018 in der Klinik für Neurochirurgie der Universitätsmedizin Mainz behandelt wurden, simuliert. Neben dem Ist-Zustand wurden auch die Operationen via ILF und LAM simuliert und deren Auswirkung auf die Stabilität in Form von zusätzlichem Ventralgleiten und Winkeländerungen der Wirbelkörper bei Inklinationsbewegung untersucht.

# 2 Material und Methoden

## 2.1 Aufbau des biomechanischen Grundmodells

Es wurde ein Multi-Body-Simulation (MBS) – Modell der Lendenwirbelsäule mit der Simulationssoftware Simpack (Dassault Systèmes Deutschland GmbH) erstellt. Hierzu wurden die Wirbelkörper eines anatomischen Grundmodells der LWS (Sawbones, Washington, USA) in einem Computertomographen dreidimensional erfasst und mittels der Visualisierungssoftware Amira (ThermoFischer scientific, Massachusetts, USA) segmentiert und als Drahtgittermodelle (stl-Format) in Simpack importiert.

Die Bänder wurden als Kraftelemente zwischen anatomisch definierten Landmarken auf den Wirbelkörpern realisiert. Die entsprechenden Kraft-Dehnungs-Kennlinien der einzelnen Bänder wurden aus biomechanischen Messungen aus der Literatur entnommen [10].

Die Bandscheiben wurden über viskoelastische Elemente definiert, die 6 Freiheitsgrade der Bewegung zulassen. Die entsprechenden Kraft-Dehnungs-Kennlinien wurden aus biomechanischen Belastungsexperimenten der Bandscheiben LWK2/LWK3 (Lendenwirbelkörper) entnommen [11].

Die Facettengelenke wurden über eine Kontaktmodellierung realisiert. Hierzu wurden auf jeder Facettengelenksfläche neun Punkte definiert und eine Regressionsebene berechnet.

Das Modell wurde anschließend mit einer axialen Kraft von 500N (entspricht etwa einem Oberkörpergewicht von 50kg) belastet. Anschließend wurden Drehmomente, verteilt auf die einzelnen Wirbelkörper angewandt, so dass das Modell in eine In- oder Reklinationsbewegung übergeht.

## 2.2 Anpassung des Modells auf individuelle Patientenwirbelsäulen

Um das Modell auf die individuelle Wirbelsäulengeometrien anpassen zu können, wurden LWS-CT-Scans von Patienten in Amira dreidimensional segmentiert. Auf den Oberflächen jedes Wirbelkörpers wurden anschließend

manuell die anatomisch definierten Bandansatzpunkte markiert (insgesamt 47 Punkte pro LWS-Modell). Durch einen skriptgesteuerten Transfer wurden die Koordinaten dieser Punkte inklusive der neuen Wirbelkörperoberflächen in das bestehende Simulationsmodell in Simpack übertragen, so dass die individuelle LWS-Geometrie nun für weitere Berechnungen verwendet werden konnte.

## 2.3 Erfassung der Endpositionen der LWS in In- und Reklination aus Röntgenfunktionsaufnahmen

Um die Endstellungen der LWS in In- und Reklination dreidimensional erfassen zu können, mussten diese Positionen aus den entsprechenden zweidimensionalen Röntgenfunktionsaufnahmen extrahiert werden.

Hierzu wurden die Röntgenbilder in Amira zusammen mit den segmentierten 3D-Wirbelkörpern (aus CT-Scans) eines Patienten importiert. Anschließend wurden die Wirbelkörper in der Sagittalebene auf die neuen im Röntgenbild dargestellten Positionen und Rotationen manuell angepasst. Dieser Vorgang wurde sowohl für die Neutral-, als auch die In- und Reklinationsstellung durchgeführt, so dass die drei dreidimensionalen Stellungen der LWS resultierten. Das Os sacrum diente hierbei jeweils als Fixpunkt, der in allen Aufnahmen dieselbe Position beibehielt (siehe Abb. 1, A1-A3).

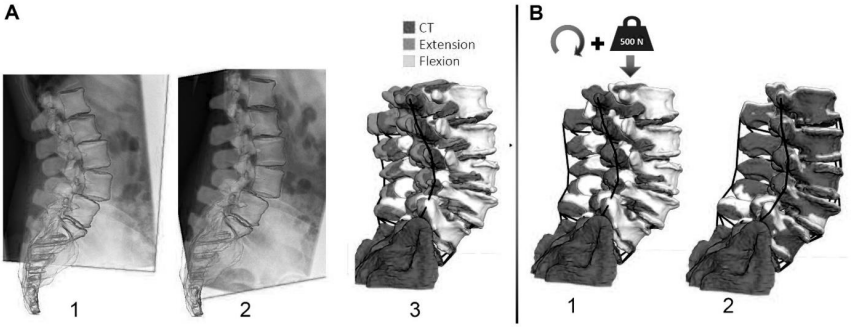



Abbildung 1*: **A1&A2:** Auf die Re- und Inklination im Röntgenbild angepasste Stellung der 3D-LWS; **A3:** Vergleich der extrahierten 3D-LWS-Stellungen mit dem Ausgangs-CT; **B1:** Ausgangsstellung des Simulationsmodells vor Belastung mit Drehmomenten an allen Wirbeln und 500N Gewichtskraft und zu erreichende Zielstellung in Flexion (violett); **B2:** Am Ende der Simulation hat das Modell die Zielstellung der Flexion erreicht.*

## 2.4 Simulation einer Inklinationsbewegung am Beispiel von Patienten mit einer degenerativen Spinalkanalstenose und begleitender Spondylolisthesis

### 2.4.1 Patientenkollektiv

Es wurden retrospektiv 9 Patienten mit einer degenerativen Spinalkanalstenose und begleitender Spondylolisthesis, die im Zeitraum von 01.01.2017 bis 30.07.2018 in der Klinik für Neurochirurgie der Universitätsmedizin Mainz operativ behandelt wurden, aus dem Krankenhausinformationssystem (KIS) identifiziert. Einschlusskriterien waren eine monosegmentale Spinalkanalstenose der LWS mit zusätzlichem Ventralversatz der Wirbelkörper. Es mussten sowohl ein präoperatives CT, als auch präoperative Röntgenfunktionsaufnahmen in Neutral-, In- und Reklinationsstellung vorliegen. Ausschlusskriterien waren eine in den Röntgenfunktionsaufnahmen radiologisch nachgewiesene Instabilität, Voroperationen, Frakturen oder multisegmentale Stenosen. Das mittlere Patientenalter betrug 69,7 Jahre (55 – 87 Jahre). Das Verhältnis von Männern zu Frauen betrug 4:5.

### 2.4.2 Simulation der Inklinationsbewegung

Nach Anfertigen der individuellen Simulationsmodelle wurden alle Wirbelsäulen zunächst mit einer axialen Kraft von 500N angreifend auf die Deckplatte von L1 belastet, um das Oberkörpergewicht zu simulieren. Anschließend wurden Drehmomente an die einzelnen Wirbelkörper angelegt, um die aus den Röntgenfunktionsaufnahmen ermittelten Endstellungen bestmöglich zu erreichen (siehe Abb. 1, B1 und B2). Hierzu war in den meisten Fällen

ein durchschnittliches Drehmoment von 10 Nm auf LWK 1 und 100 Nm auf LWK 5, in einigen Fällen auch zusätzlich ein Drehmoment auf LWK 3 und 4 notwendig.

### 2.5 Simulation nach virtuell durchgeführter Dekompression über ILF und LAM

Zusätzlich zum präoperativen Ausgangszustand wurden die operativen Eingriffe der ILF und LAM an den jeweiligen Patientenwirbelsäulen untersucht. Hierzu wurden im Fall der ILF die Lig. flava auf der entsprechenden Höhe aus der Simulation entfernt (dies entspricht in etwa dem operativen Vorgehen der Dekompression über eine einseitige ILF und Resektion der Bänder auf der ipsilateralen sowie mittels sog. Undercutting auf der kontralateralen Seite).

Im Fall der LAM wurden die Lig. flava, die intraspinösen und supraspinösen Bänder auf Segmenthöhe aus der Simulation entfernt.

Für beide Fälle wurde anschließend die Inklinationsbewegung wiederholt simuliert und die Auswirkungen einer Destabilisierung berechnet. Hierzu wurden als Parameter die zusätzlich auftretende Ventralverschiebung der Wirbelkörper sowie die zusätzlich auftretende Winkeländerung am Ende des Bewegungsablaufs bestimmt.

## 3 Ergebnisse

### 3.1 Grundlegende Simulationsgenauigkeit der MBS im Vergleich zu anderen Simulationsmethoden und biomechanischen Messungen

Um die generelle Zuverlässigkeit der MBS Simulation an der LWS einschätzen zu können, wurde das Basismodell mit einer Gewichtskraft von 500N belastet und anschließend der hieraus resultierende intradiskale Druck in den einzelnen Bandscheiben berechnet.

Diese Ergebnisse wurde mit Daten aus der Literatur und selbigem Versuchsaufbau sowohl in Simulationsmodellen (finite element simulation (FE), Rohlmann) als auch in biomechanischen Messungen (Anderson, Sato und Wilke) verglichen (siehe Abb. 2).

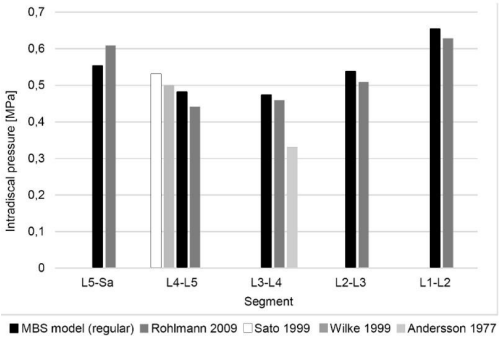



Abbildung 2*: Intradiskaler Druck in den verschiedenen LWS-Segmenten bei Belastung mit 500N. Vergleich des MBS-Modells mit Literaturdaten vergleichbarer Simulationen oder Messungen. (Quellen: Rohlmann [12], Sato [13], Wilke [14], Andersson [15])*

Hier zeigt sich, dass die schnelle MBS-Simulation bezüglich des intradiskalen Drucks Werte in der derselben Größenordnung liefert, wie die wesentlich zeitaufwändigere FE-Simulation. Zusätzlich zeigt sich eine gute Übereinstimmung mit den Messungen aus biomechanischen Experimenten.

### 3.2 Erreichen der aus den Röntgenfunktionsaufnahmen extrahierten Endstellungen in Inklination bei intakter Wirbelsäule

Es wurde für alle 9 Wirbelsäulen (ohne Veränderung an Bandstrukturen) jeweils der Abstand jedes Wirbels am Ende der Inklinationsbewegung im Vergleich zu dem entsprechenden Wirbel aus den Funktionsaufnahmen

gemessen, der mit der angelegten Drehmomentverteilung zwischen LWK 1 und 5 am besten erreicht werden konnte. Insgesamt ergab sich ein durchschnittlicher Abstand von: LWK1 5,3 mm; LWK2 4 mm; LWK3 2,9 mm; LWK4 3,2 mm; LWK5 2,7 mm.

Die nach oben zunehmende Ungenauigkeit ist dem größeren zurückgelegten Weg dieser Wirbelkörper bei der Inklinationsbewegung geschuldet.

### 3.3 Vergleich der destabilisierenden Wirkung einer ILF mit der einer LAM

Die potentiell destabilisierenden Wirkungen einer ILF und LAM wurden an allen 9 Patientenmodellen untersucht. Hierbei wurden sowohl die Abstände, als auch die Drehwinkel in Sagittalebene der Wirbelkörper in Inklinationsstellung für ILF und LAM mit denen der intakten Patientenwirbelsäulen aus 3.2 verglichen. Die Kräfte und Drehmomente wurden hierbei aus den in 3.2. ermittelten Werten für jeden Fall beibehalten.

Es zeigte sich in Bezug auf die prozentuale Änderung der Abstände (entspricht einem vermehrten Ventralgleiten der Wirbelkörper) eine deutliche Zunahme nach durchgeführter LAM im Vergleich zu einer ILF. Kam es beispielsweise nach ILF bei LWK 3 zu einer Zunahme des Ventralgleitens von ca. 2% so ergab sich eine Zunahme der Ventralbewegung nach LAM bei LWK 3 um ca. 6,5%. (siehe Abb. 3)

Bezüglich der Winkel ergaben sich im Durchschnitt bei der ILF Abweichungen von: 10,5 % (LWK1), 4% (LWK2), 9,3 % (LWK3), 6,7% (LWK4) und 1 % (LWK5). Bei der LAM dagegen zeigen sich durchschnittliche Winkelabweichungen von: 15,6 % (LWK1), 10 % (LWK2), 13,4 % (LWK3), 8,7 % (LWK4) und 3,6 % (LWK5).

Diese Ergebnisse sprechen für eine deutlichere destabilisierende Wirkung der LAM gegenüber der ILF. Dieser Umstand entspricht der klinischen Erfahrung von Wirbelsäulenchirurgen, konnte nun im vorliegenden Simulationsmodell aber erstmals auch quantitativ nachvollzogen werden.

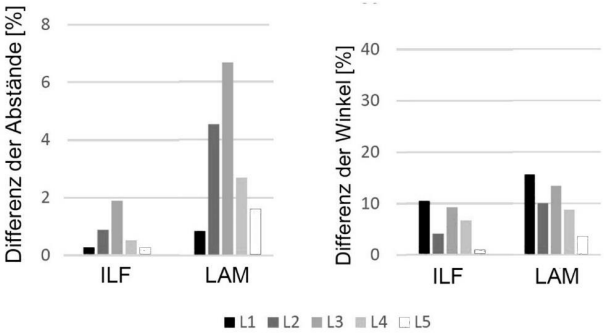



Abbildung 3*: **Links:** zeigt die durchschnittlichen Abstandsänderungen in Prozent, die in Inklinationsstellung der LWS bei ILF oder LAM im Vergleich zur intakten LWS auftreten. **Rechts:** zeigt die durchschnittlichen Winkeländerungen in Prozent, die in Inklinationsstellung der LWS bei ILF oder LAM im Vergleich zur intakten LWS auftreten.*

## 4 Diskussion

In der vorliegenden Arbeit ist es beispielhaft gelungen, anhand routinemäßig erhobener klinischer Daten, bei einer typischen Patientengruppe, die Effekte verschiedener Operationsmethoden individuell vorherzusagen. Dies ist von Bedeutung, um vorher zu sagen, welche der Patienten von einem zusätzlichen operativen Schritt, wie der Spondylodese, mit einer hohen Wahrscheinlichkeit profitieren werden und welche nicht.

Wie bereits eingangs erwähnt, finden funktionelle Parameter derzeit nur über die persönliche Erfahrung der Chirurgen Eingang in die Operationsplanung. Daher ist jedes System, welches eine Annäherung an die reellen biomechanischen Verhältnisse erlaubt, eine Verbesserung, auch wenn, wie z.B. bei der hier verwendeten Simulation, noch deutliche Diskrepanzen zur Wirklichkeit bestehen. Dennoch scheint die Computersimulation eine einzigartige Chance zu bieten, bereits vor der Operation funktionelle Parameter und deren Veränderung durch verschiedene operative Schritte im Vorhinein quantitativ beurteilen zu können.

Einer der offensichtlichen Kritikpunkte sind die Defizite in der Individualisierung des Systems. Derzeit sind nur die Dimensionen und Positionen der Festkörper, namentlich der im CT dargestellten Wirbelkörper, berücksichtigt. Dies liegt offensichtlich daran, dass lediglich die Daten verwendet werden können, die in der Klinik bereits routinemäßig erhoben werden. Die übrigen Parameter, insbesondere die Steifigkeit der Bänder, muss aus biomechanischen Versuchen ergänzt werden, die natürlich nicht individuell sind. Dennoch stellt alleine die Verwendung der richtigen Größendimensionen und relativen Positionen der Band- und Muskelansatzpunkte eine wesentliche Verbesserung zu nicht individualisierten Simulationen dar, da diese relativen Abstände als Hebel wirken und die jeweiligen Kräfte potenzieren, wohingegen die individuelle Stärke eines Bandes etc. in erster Linie von seiner Dicke abhängt und somit die biomechanisch gemessenen Werte eine gewisse Übertragung auf andere Individuen erlauben.

In der derzeitigen Version der Simulation sind außerdem noch keine Muskeln integriert. Dies ist ein komplexes Thema, da nicht nur die neuronale Kontrolle derselben, sondern auch deren Trainingszustand hoch variabel sind. Allerdings sind die Muskelkräfte eben dadurch, dass sie dem Trainingszustand unterliegen, auch nicht mit der Operation statisch verändert, sondern können durch geeignete Maßnahmen (Physiotherapie etc.) auch postoperativ noch beeinflusst werden, woraus sich ergibt, dass sie bei der Operationsplanung von geringerer Bedeutung sind, als die starren Einheiten, die in der Regel nicht reversibel verändert werden.

Hier hat sich die Verwendung einer Multi-Body-Simulations-Software bewährt, weil gut neue Daten (wie die derzeit in Arbeit befindliche Simulation von Muskeln) in das Modell zu integrieren sind. Ein weiterer Vorteil ist die vergleichsweise kurze Rechenzeit. Da das System gedacht ist, um in einem im wesentlichen iterativen Vorgehen einen Operationsplan zu optimieren, ist dies notwendig, um im individuellen Fall rasch ein klares Ergebnis zu erhalten.

## 5 Zusammenfassung

Es konnte gezeigt werden, dass sich die Endstellungen aus den Röntgenfunktionsaufnahmen mit hoher Übereinstimmung insbesondere im unteren LWS-Bereich nachsimulieren ließen. In Bezug auf das gewählte Operationsverfahren zeigte der invasivere Eingriff der LAM im Mittel eine mehr als doppelt so große destabilisierende Wirkung hinsichtlich eines vermehrten Ventralgleitens und Winkeländerungen verglichen mit der Technik der ILF. Hierdurch konnte die klinische Beobachtung der potentiell destabilisierenden LAM (heute nur noch durchgeführt mit begleitender Implantation einer stabilisierenden Spondylodese) im vorliegenden Computersimulationsmodell bestätigt werden.

## Referenzen

[1] Heliovaara, M., Sievers, K., Impivaara, O., Maatela, J., Knekt, P., Makela, M. and Aromaa, A. Descriptive epidemiology and public health aspects of low back pain. *Annals of medicine*, 21, 5 (Oct 1989), 327-333.
[2] Raspe, H. Rückenschmerzen. Gesundheitsberichterstattung des Bundes. *Robert Koch-Institut*, Heft 532012).
[3] Schmidt, C. O., Raspe, H., Pfingsten, M., Hasenbring, M., Basler, H. D., Eich, W. and Kohlmann, T. Back pain in the German adult population: prevalence, severity, and sociodemographic correlates in a multiregional survey. *Spine*, 32, 18 (Aug 15 2007), 2005-2011.
[4] Czabanka, M., Thomé, C., Ringel, F., Meyer, B., Eicker, S.-O., Rohde, V., Stoffel, M. and Vajkoczy, P. Operative Versorgung degenerativer Erkrankungen der Lendenwirbelsäule. *Der Nervenarzt*, 89, 6 (June 01 2018), 639-647.
[5] Dreischarf, M., Shirazi-Adl, A., Arjmand, N., Rohlmann, A. and Schmidt, H. Estimation of loads on human lumbar spine: A review of in vivo and computational model studies. *J Biomech*, 49, 6 (Apr 11 2016), 833-845.
[6] Panjabi, M. M. Clinical spinal instability and low back pain. *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, 13, 4 (Aug 2003), 371-379.
[7] Forsth, P., Olafsson, G., Carlsson, T., Frost, A., Borgstrom, F., Fritzell, P., Ohagen, P., Michaelsson, K. and Sanden, B. A Randomized, Controlled Trial of Fusion Surgery for Lumbar Spinal Stenosis. *The New England journal of medicine*, 374, 15 (Apr 14 2016), 1413-1423.
[8] Liang, H. F., Liu, S. H., Chen, Z. X. and Fei, Q. M. Decompression plus fusion versus decompression alone for degenerative lumbar spondylolisthesis: a systematic review and meta-analysis. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, 26, 12 (Dec 2017), 3084-3095.
[9] Mobbs, R. J., Li, J., Sivabalan, P., Raley, D. and Rao, P. J. Outcomes after decompressive laminectomy for lumbar spinal stenosis: comparison between minimally invasive unilateral laminectomy for bilateral decompression and open laminectomy: clinical article. *Journal of neurosurgery. Spine*, 21, 2 (Aug 2014), 179-186.

[10] Shirazi-Adl, A., Ahmed, A. M. and Shrivastava, S. C. Mechanical response of a lumbar motion segment in axial torque alone and combined with compression. *Spine*, 11, 9 (Nov 1986), 914-927.
[11] Schmoelz, W., Erhart, S., Unger, S. and Disch, A. C. Biomechanical evaluation of a posterior non-fusion instrumentation of the lumbar spine. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, 21, 5 (May 2012), 939-945.
[12] Rohlmann, A., Zander, T., Rao, M. and Bergmann, G. Applying a follower load delivers realistic results for simulating standing. *J Biomech*, 42, 10 (Jul 22 2009), 1520-1526.
[13] Sato, K., Kikuchi, S. and Yonezawa, T. In vivo intradiscal pressure measurement in healthy individuals and in patients with ongoing back problems. *Spine*, 24, 23 (Dec 01 1999), 2468-2474.
[14] Wilke, H. J., Neef, P., Caimi, M., Hoogland, T. and Claes, L. E. New in vivo measurements of pressures in the intervertebral disc in daily life. *Spine*, 24, 8 (Apr 15 1999), 755-762.
[15] Andersson, G. B., Ortengren, R. and Nachemson, A. Intradiskal pressure, intra-abdominal pressure and myoelectric back muscle activity related to posture and loading. *Clinical orthopaedics and related research*, 129 (Nov-Dec 1977), 156-164.

# Production of Intracranial Dynamic Aneurysm Models for Neurosurgical Applications and Future Directions

*Fredrick Johnson Joseph[1], David Bervini[2], Andreas Raabe[2], Stefan Weber[1]*

*[1]ARTORG center, University of Bern, Bern, Switzerland*
*[2]Department of Neurosurgery, Inselspital Bern, Bern, Switzerland*

contact: fredrick.joseph@artorg.unibe.ch

***Abstract***

A large portion of intracranial aneurysm (ICA) is found complex to be treated due to their unique anatomy and limited degrees of freedom under the microsurgical field. Experts with several years of treating experience performing ICA clipping find it difficult to visualize the image dataset as a 3D object pre and intraoperatively. At the same time, medical residents are also not exposed to realistic training due to the intricacy in reproducing microvascular structures. In recent times, advancements in additive manufacturing have paved a way to produce dynamic life-like ICA with patency and blood flow mimicking a real patient. This paper discusses the approaches followed to produce realistic ICA models for dynamic clipping-training stations. Such a model will also lead to explore various possibilities to improve the surgeon's learning curve, establish better clipping strategies, understand the clinical workflow during medical residency, pre-operative practice, develop robotic manipulators and guidance for accurate ICA clipping interventions.

**keywords**: intracranial-aneurysm, micro-neurosurgery, training, simulator

## 1 Problem

Approximately 0.6 to 6% of the population has an intracranial aneurysm and the rupture rate is about 0.7% per year [1]. These are responsible for 80% of all spontaneous subarachnoid hemorrhages (SAH) and the mortality risk is nearly 60% at 6 months for such cases [2]. Patients with a high risk of rupture are selected for treatment using clipping by neurosurgery or coiling through the endovascular procedure. However recent studies discussed the trade-off in patient selection method and better clinical significance on followup when ICA is clipped [3]. Successful clipping is greatly surgeon dependent, subjective and can lead to complications intraoperatively or post-operative. Some of these problems can be mitigated by efficient realistic training at the very early stages of medical education and preoperatively [4]. Though general surgeries and interventions can be educated using standard phantoms and cadavers, intracranial aneurysm treatment training is unique and difficult to mimic patient anatomy. Several researchers from the past have reported production of the intracranial aneurysm without dynamic blood flow in the model during neurosurgical training. The blood flow through the ICA lumen significantly impacts an aneurysm to rupture and appropriate clip positioning during the procedure. Simultaneously, bloodstream injected with optical tracers like ICG is essential to quantify the clipping quality during training. These vital behaviors are missing in the current state of the art models for training purposes. Thus, the production and training in lifelike Dynamic ICA are necessary when clip treatment training needs to be patient-specific.

Latest trends in additive manufacturing have driven potential methods for rapid production of patient-specific aneurysm models in a few hours. A novel method was found for producing realistic intracranial aneurysm model on a true scale (1:1) from the patient data. These models were tested with a closed-loop flow mechanism mimicking the cardiac pulsation to determine the usefulness of such models for surgical training applications. Incorporating such vascular anatomy for the education approaches can report better clip placement, reduction of surgical cost and time by improving the surgeon's fidelity and skills.

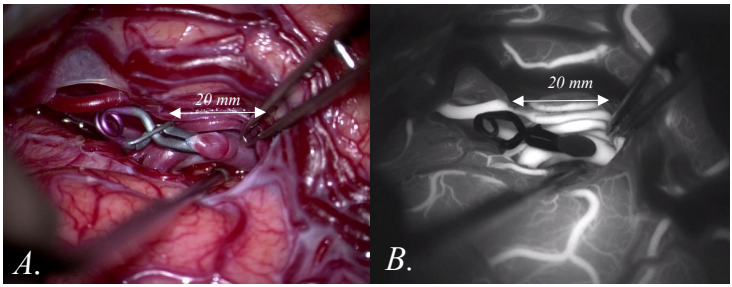



Figure 1*: A. Image representation of a clipping procedure on the intracranial aneurysm under the microscopic surgical field, B. Illustration of post-clipping investigation for testing aneurysm sac patency using IR optical microscope with ICG marker*

## 2 Material and Methods:

Intracranial vascular structures from the pre-operative diagnostic imaging like DSA (Digital Subtraction Angiography-most useful for model creation), CT and MRI were segmented using Amira. The segmented model is processed using software like Solid works for preparing a 3D printable form as shown in Figure 2.A-C. The specially prepared material mixture of Acrylonitrile-Butadiene-Sterol-Copolymer and Polyvinyl alcohol is used for 3D model building using Ultimaker 3D printer with a resolution of 0.8 mm. The models are preserved for several stages of treatments using benzol vapor and coated with shape-retaining materials like standard wax. The form is then dried up for a few minutes before special differential coating processes using a material mixture of polyurethane and silicone gel. The prepared models are treated using chemical-based material disposition of the coated elements for a few minutes and prepared for testing under a flow loop circulation. Special properties like patency, neck diameter, and wall thickness are tested as shown in figure 2.D-F, otherwise, the produced model is discarded from the training application in order to achieve a realistic life-like feeling.

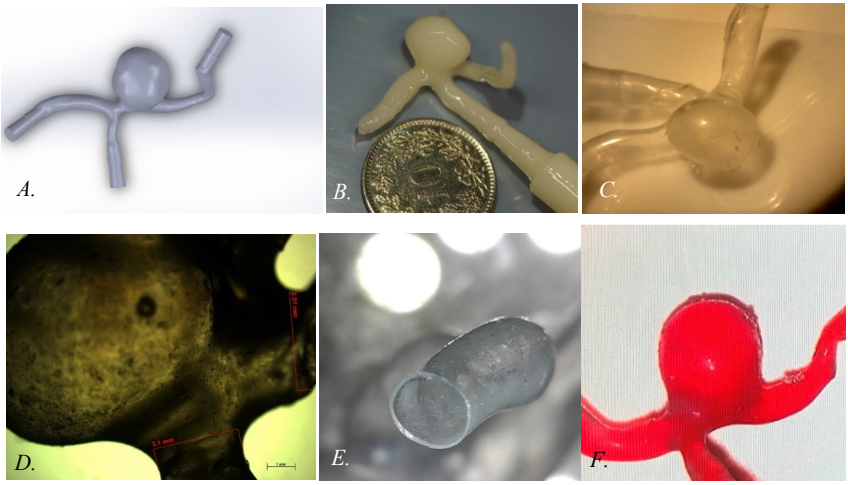



Figure 2*: A. 3D model of the segmented intracranial aneurysm, B. 3D produced model, C. Processed optically transparent 3D fabricated model of aneurysm with parent vessel structures, D. Wall patency, and neck flow testing under a microscope, E. Microscopic Image representation of the parent vessel wall F. 3D produced aneurysm model tested for dynamic flow under cardiac mimicking pulsatile flow system with artificial blood.*

## 3 Results

Specially produced 3D intracranial aneurysm models were tested for the patency and the flow behavior in a standard skull phantom to evaluate the quality of intraluminal flow channels. The models were connected under a specially designed cardiac pulse simulating dynamic flow system carrying artificial blood infused with ICG marker for examination. Figure 3 shows the aneurysm sac imaged under a microscopic field using IR and color optical microscopic cameras. Surgical manipulators and clipping applicators with a suitable clip were used to close the aneurysm neck. Observations were made to record the aneurysm neck closure, post-clipping procedure without rupturing the aneurysm or any parent and neighboring vessels.

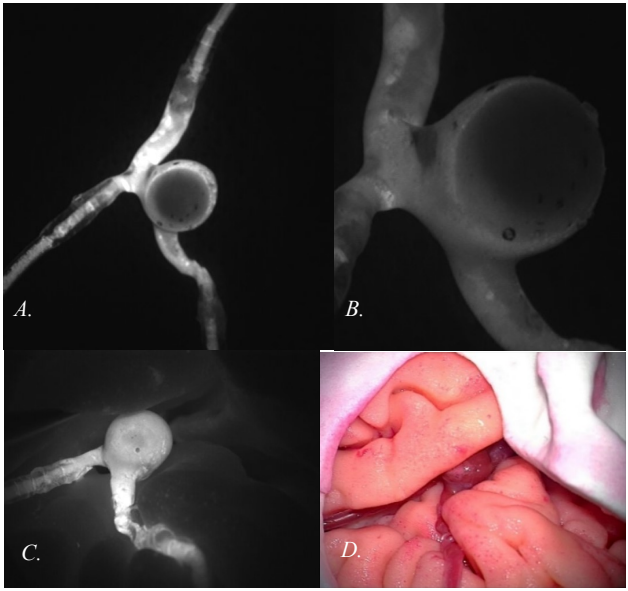



Figure 3*: A. 3D produced model tested for patency under IR surgical microscope with the dynamic flow loop with ICG marker, B. Microscopic image focused on the region of interest to check for flow through the aneurysm sac, C. 3D model placed on the modelled brain phantom at the respective anatomical position of the patient and tested using IR camera, D. Image representation of an aneurysm model placed in the surgical training brain phantom model.*

## 4 Discussion

Anatomical equivalence between the physical 3D fabricated models with respect to the vessel thickness, stiffness, roughness to hold clip during manipulation, replicating reality is yet to be studied. Production methods have also resulted in large failure due to the delicate microvascular structures; optimization of the methods can result in higher production ratio leading to quick fabrication for pre-surgical training. Standard solution or a tool combined with the radiological image guidance is under construction for solving the major problems in positioning vascular structures in the accurate anatomical position of the soft brain phantom. Mechanical positioner for all seven most common clipping aneurysms can improve the effectiveness in evaluating the surgeon's skill in navigating the aneurysm from the time point of post-craniotomy to post-clipping investigation. This method can also help in understanding more precisely the standards methods to manipulate the vessels for determining the most suitable clip to close the aneurysm neck for building a clipping strategy before the intervention.

## 5 Conclusion

Rapid production of patient-specific and dynamic realistic 3D models will lead to a large understanding of the clinical point of view for a medical trainee or experienced surgeon to improve one's fidelity for such complex surgical interventions. The production process described enlightens the possibilities to integrate the models in a compact bench-top training module. Studies and assessment of interventional capabilities in such models can lead to exploring further indefinite evidence in procedures with a limited degree of freedom. Understanding the subjective variability of clipping procedure and tricks among successful surgeons from such models can help junior surgeons to carry over the treatment legacy. Related information can also result in the preparation of next-generation medical devices and instruments aiding the surgeons for therapy guidance as well for the development of semi/automatic robotic manipulators for intracranial clipping applications. At the same time, pre-operating training can reduce the adverse effects among the patients treated and can emphasize patient safety who refuses for such procedures.

## References

[1] G. J. E. Rinkel, M. Djibuti, A. Algra, and J. van Gijn, "Prevalence and Risk of Rupture of Intracranial Aneurysms," *Stroke*, vol. 29, no. 1, pp. 251–256, 2011.

[2] T. Steiner, S. Juvela, A. Unterberg, C. Jung, M. Forsting, and G. Rinkel, "European stroke organization guidelines for the management of intracranial aneurysms and subarachnoid haemorrhage," *Cerebrovasc. Dis.*, vol. 35, no. 2, pp. 93–112, 2013.

[3] A. Molyneux, "International Subarachnoid Aneurysm Trial ( ISAT ) of neurosurgical clipping versus e ... International Subarachnoid Aneurysm Trial ( ISAT ) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms : a ra," vol. 6736, no. 02, pp. 1–2, 2013.

[4] A. Alaraj *et al.*, "Virtual reality training in neurosurgery: Review of current status and future applications," *Surg. Neurol. Int.*, vol. 2, no. 1, p. 52, 2011.

# Interventionelle Techniken & Bildgebung

# Quantitative volumetric assessment of percutaneous image-guided microwave ablations for colorectal liver metastases

*Raluca-Maria Sandu[1], Iwan Paolucci[1], Pascale Tinguely[2], Daniel Candinas[2], Jacob Freedman[3], Stefan Weber[1]*

*[1]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*
*[2]Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland*
*[3]Danderyd Hospital, Karolinska Institute, Department of Clinical Sciences, Stockholm, Sweden*

Contact: raluca.sandu@artorg.unibe.ch

***Abstract***

*Thermal ablations are becoming an increasingly common alternative treatment to surgery in patients with primary and secondary hepatic tumors. In image-guided percutaneous ablations of liver tumors, the technical post-interventional success of ablation treatment is defined by achieving complete tumor destruction and an ablation margin of at least 5 mm. Therefore, it is crucial that proper ablation verification is ensured intra-operatively or immediately post-operatively. In the current setting, the ablation-tumor coverage is visually evaluated by an interventional radiologist who compares the differences between the pre- and post-ablation images. To address this limitation, we have developed an image analysis pipeline consisting of a set of quantitative methods for evaluating the volumetric coverage of the ablation, which could be added to image-guidance systems. In this work, we present the quantitative ablation evaluation method and preliminary results after applying it to a retrospective cohort of 100 patients that were treated with image-guided percutaneous ablations for colorectal liver metastasis.*

**Keywords**: Image-guidance, ablation, liver metastases, computer-assisted, image processing, quantification

## 1 Problem

Thermal ablation of colorectal liver metastases (CRLM) is a minimally invasive treatment alternative to surgical resection of which is the current gold standard for curative care. Not only is ablation considered for patients not amenable for surgical resection (< 20%), but also increasingly for resectable tumors even with a curative intent. Percutaneous ablation is a minimally invasive technique generally performed under image-guidance based on CT, MRI or ultrasound. Image based navigation leads to a simple, fast and accurate placement of the ablation probe into the liver tumor. The patients undergoing an ablation benefit from a drastically improved quality of life, as they go home the next day rather than staying in the hospital for a week. Despite the advantages provided by an ablation procedure, it has not yet taken over resection due to the local tumor progression (LTP) rates, which in these cases reach up to 48% [1]–[5]. LTP is generally associated with poor survival prognosis. One of the major limitations of this technique is achieving complete coverage of the tumor including a 5 mm margin, which has been shown to decrease the risk of LTP by 46% [1]–[5].

In hepatic resections it is straight-forward to evaluate the resection margin by simply cutting the specimen intra-operatively and even by histological assessment. In contrast to this, the ablated part remains in the liver and can only be visualized by non-invasive imaging like CT or MRI. In the clinical routine, the ablation is visually assessed by the expert eye of the radiologist, by displaying the pre- and post-ablation CT scans side-by-side or by overlaying them on top of each other. The ablation coverage and the margin between the tumor and the ablation is visually estimated and this presents several drawbacks. Firstly, visual assessment is a highly subjective technique which is limited by repeatability. Secondly, this is a 2D approach, where only the maximum and minimum diameter of the ablation zone are estimated. However, the ablation margin should be measured all around the tumor, but this is a challenging and time-consuming task in the 3D space of a CT or MRI image. All in all, these problems introduce large inaccuracies in measuring the ablation coverage and its margin, which might be a risk factor for LTP.

Currently there are very few software platforms available on the market for quantitatively assessing the ablation coverage post-operatively [1], [5], [6]. Unfortunately, the post-operative evaluation excludes the possibility of re-ablation during the same procedure and thus completely eradicating the tumor. More importantly, the sample size of these studies is rather small with a possible inhomogeneous study population [1], [5], [6].

To address the current limitations in the area of quantitative ablation assessment, we developed an image analysis pipeline to quantify the ablation margin with different metrics based on overlap of surface distances. We are currently evaluating these quantitative metrics on a retrospective multi-centric study of 100 patients with CRLM (clinicaltrials.gov: NCT02642185). In this work the image-analysis pipeline and some of the very early results from the multicentric study are presented.

# 2 Materials and Methods

A customized software for semi-automatic segmentation of tumors and ablations and subsequent quantification has been developed and integrated into a navigation system for percutaneous ablations [7]. The tumor and ablation volumes can be segmented manually or by employing "Fast-Marching" semi-automatic algorithm. The seeding point for the tumor and the ablation volumes are taken from the target point of the ablation probe. The binary masks are then saved as DICOM files along with the other CT scans. The entire analysis pipeline is shown in Figure 1, and based on these segmentations the following radiomics are computed:

a) Surface distance

The surface distance was calculated as the Euclidean distances between each surface voxel of the tumor and ablation volumes, also known as Maurer distance [8]. The calculated distances are displayed in a histogram marking the percentage of surface voxels covered by each range. A traffic light color scheme was applied to the histogram bins to mark the safety margin convention (green > 5mm, orange 0 – 5 mm, red <0 mm).

b) Volume coverage ratio and residual tumor

The volume coverage ratio (VCR) was computed as an adapted DICE coefficient between the tumor and the ablation volume. Additionally, the volume of the residual tumor is computed as the subtraction of the ablation from the tumor volume.

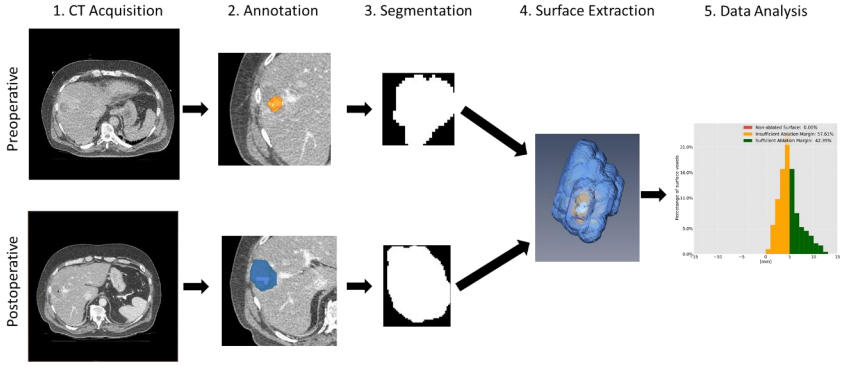



*Figure 1 Quantitative ablation evaluation pipeline from acquisition of the pre-/post-operative images to the visualization of the results*

## 2.1 Retrospective evaluation on CRLM cohort

This analysis is part of the MAVERRIC trial (Microwave Ablation Versus Resection for Resectable Colorectal Liver Metastasis) which includes the analysis of the ablation volumes. This trial aims to prove that a strategy of first line local ablation of colorectal liver metastases with microwaves is not inferior to liver resections in terms of survival rates at three years with secondary endpoints being survival at five and ten years. The study included a cohort of 100 patients with tumors < 31 mm in diameter and 1-5 metastases, treated with microwave ablation under CT-image guidance. In this cohort a total of 173 lesions were treated. From each patient, the tumor and the ablation from the interventional CT scans were extracted using a semi-automatic segmentation tool [7]. The segmentations were validated by an experienced radiologist. From these datasets, 65 lesions where excluded due to missing ablation validation scans, which is mainly due to the constraints in applicable contrast agent.

# 3 Results

Out of the 173 lesions in this study 108 can be included in this analysis pipeline. From 108 lesions eligible for analysis, we have currently analyzed 2 out of 173 (2 patients with a single lesion each). Figure 2 depicts these two examples, consisting of sufficiently (right) and insufficiently (left) volumetric ablation coverage. The histograms show the percentage of distances plotted with respect to the surface voxels covered. The tumor with insufficient margin (left) had an LTP whereas the tumor with sufficient margin (right) had no LTP at 6 months.

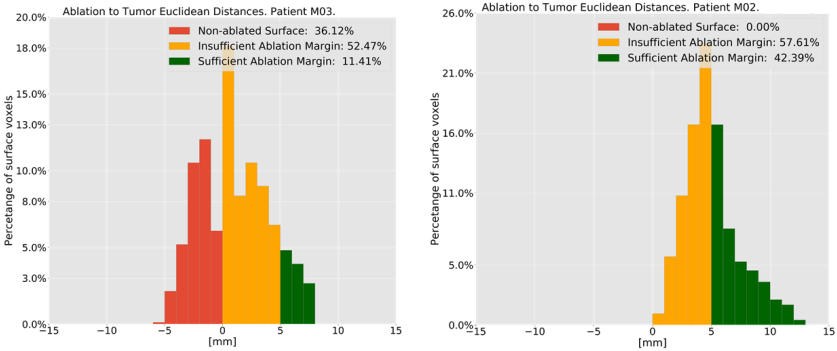



*Figure 2: Histograms of the surface distances between the tumor and the ablation volumes*

Figure 3 shows a 3D visualization of the volume overlap with the tumor (orange) and the ablation zone (blue). In the image on the left, one can see the residual tumor with a volume of 1.08 ml. That means that the volume coverage ratio [7] is only 31.68% , whereas in the right picture from Figure 3 there is no residual tumor visible.

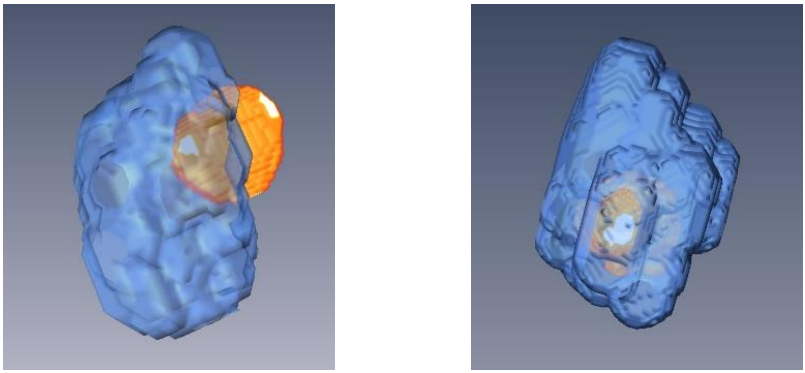


*Figure 3: Volumetric representation of residual tumor (corresponding to Figure 2)*

# 4 Discussion

The quantitative evaluation pipeline (Figure 1) for assessing ablation coverage might be a useful tool to provide a quantitative measurement over the entire 3D space. As shown by previous studies it is crucial to cover the tumor with a minimum of 5 mm margin [1]–[5]. Therefore, this tool will be especially useful when applied intra-operatively, so the radiologist can either re-ablate the residual tumor or call the patient for an earlier follow-up.

The ablation margin histograms (Figure 2) reflect the coverage depicted by their 3D model representation. We hypothesize that the percentage of tumor covered within a certain ablation margin could correlate with the local tumor progression rate. To test this hypothesis, we will correlate the results from the ablation metric with the local tumor progression at 6-months and 1-year follow-up for the whole study population once the rest of the follow-up data is available. However, this analysis needs to include more variables than just these radiomics to control for other confounding factors. It is important to note that the results presented here are preliminary and reflect just the first 2 cases that were fully analyzed.

So far, we have identified 108 out of 173 complete datasets that could be analyzed with this method. At this point, some special cases (52 out of 108) had to be excluded from the analysis because the distance metric does not account for them. One case for which we need an adjusted metric is subcapsular lesions, i.e. lesions that are within 5 mm of the border of the liver. These lesions need an adjusted ablation margin metric which considers that lesions close to the surface cannot have a margin larger than 5 mm. Another case where an adjusted metric is needed is when ablation zones from several ablations merge together. The exceptionally large ablation volume created in this instance might skew the ablation margin metrics towards a false positive trend.

Overall, the major advantages of this volumetric quantitative analysis pipeline are its speed, automated steps and most importantly the fact that it has already been integrated into an existing image-guided navigation system. The last point just makes it easier for the method to be deployed intra-operatively, becoming part of the clinical routine if the ablation margin metric is validated and ultimately decreasing the risk of local tumor progression. If we manage to achieve this, there will be a paradigm shift from resection towards ablation as a first line of treatment for liver tumors.

## 5 Conclusion

To conclude, this work proposes a quantitative ablation evaluation for measuring the coverage of the ablation volumes. The method could enable a precise measurement of ablation margins and volumes that could be utilized to identify patients with increased risk of local tumor progression. Eventually, this method might provide an important intra-operative feedback that would allow precise re-ablations in the same treatment session.

## References

[1] E. A. Kaye *et al.*, “Volumetric 3D assessment of ablation zones after thermal ablation of colorectal liver metastases to improve prediction of local tumor progression,” *Eur. Radiol.*, vol. 29, no. 5, pp. 2698–2705, 2019.

[2] X. Wang *et al.*, “Margin size is an independent predictor of local tumor progression after ablation of colon cancer liver metastases,” *Cardiovasc. Intervent. Radiol.*, vol. 36, no. 1, pp. 166–175, 2013.

[3] W. Shady *et al.*, “Percutaneous Microwave versus Radiofrequency Ablation of Colorectal Liver Metastases: Ablation with Clear Margins (A0) Provides the Best Local Tumor Control,” *J. Vasc. Interv. Radiol.*, vol. 29, no. 2, pp. 268-275.e1, 2018.

[4] W. Teng *et al.*, “Insufficient ablative margin determined by early computed tomography may predict the recurrence of hepatocellular carcinoma after radiofrequency ablation,” *Liver Cancer*, vol. 4, no. 1, pp. 26–38, 2015.

[5] M. Solbiati *et al.*, “A novel software platform for volumetric assessment of ablation completeness,” *Int. J. Hyperth.*, vol. 0, no. 0, pp. 1–7, 2019.

[6] C. Rieder *et al.*, “Software-assisted post-interventional assessment of radiofrequency ablation,” *Med. Imaging 2014 Image-Guided Proced. Robot. Interv. Model.*, vol. 9036, p. 903604, 2014.

[7] R. Hrabuska, R. Sandu, I. Paulocci, K. Gerber, and S. Weber, “A Framework For The Quantitative Assessment of Image-guided Percutaneous Ablation of Hepatic Lesions,” in *CURAC-17th Annual Meeting of the German Society for Computer- and Robot-Assisted Surgery, 2018*, pp. 1–5.

[8] C. R. Maurer, Rensheng Qi, and V. Raghavan, “A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 265–270, Feb. 2003.

# Thermal ablation volumes with configurable shapes

*Iwan Paolucci[1], Jan Hermann[1], Raluca Sandu[1], Daniel Candinas[2], Pascale Tinguely[2], Stefan Weber[1]*

*[1]ARTORG Center for Biomedical Engineering Research,*
*[2]Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital*
*University of Bern, Bern, Switzerland*

Contact: iwan.paolucci@artorg.unibe.ch

***Abstract***

*"Technology-Pitch": Stereotactic percutaneous ablation is a rapidly advancing modality for treatment of tumors in soft solid organs such as the liver, kidney and lung. Guidance is typically created through physical constraints through passive or active guiding devices. Most prominently, robotic arms have been proposed to automatically align with a desired trajectory. However, their technical, logistical and regulatory costs prevent utilization for tasks of low complexity. Robotic technology can only supersede it delivers "super-human" performance and can achieve outcomes far beyond that of a human.*

*"Clinical Pitch": Stereotactic percutaneous ablation is a fast advancing modality for treatment of tumors in soft solid organs such as the liver, kidney and lung. Ablation needles typically deliver ellipsoidal shaped ablation volumes varying in size as a function of ablation time and power. As a result, irregularly shaped tumors can only be ablated by delivering volumes significantly larger than the target itself and by use of several needle applicators.*

*In this study, we have investigated the possibility of creating ablation volumes with configurable shapes through synchronous modulation of ablation power and applicator position. More specifically, in an ex-vivo porcine liver model, ablations very carried out by dynamically modulating the energy level depending on position and time.*

**Keywords**: Interventional robotics, thermal ablation, image-guidance, configurable ablation shapes

## 1 Problem

Each year, there are about 2'000'000 cases of lung, 850'000 cases of liver and 400'000 cases of kidney cancer worldwide[1]. Out of these patients, only about 20% - 30% are candidates for surgical resection, which is considered gold standard. Ablation of tumors of less than 30 mm becomes more and more standard for these cases. Even reimbursement systems show slow adaptation and provide a business model for stereotactic ablation procedures with cost efficiency for the hospital.

Among stereotactic ablation guidance systems, there are several robotic devices for needle guidance in ablation procedures [2–4]. However, these are merely motorized articulated arms for alignment of ablation needles, providing a 6-dimensional alignment of a trajectory with respect to the target. Therefore, the surgeon still has to deliver the needle by hand and the robot passively guides it. Due to the increased cost and complexity in setup devices which only provide an alignment function are unlikely to be used on a regular basis. This is especially highlighted when considering that the improvement in needle placement accuracy is not better than using stereotactic systems which are much cheaper. Therefore, it is also unlikely that the added cost will be reimbursed which will hinder the introduction of robotic technology on a larger scale. Therefore, we started to investigate the use-cases where the use of robotic technology can provide a treatment which is beyond the capability of a human.

One such application are tumors greater than 30 mm or with irregular shape, which are problematic to treat in a reproducible manner. In most recent studies, tumor size of > 30 mm has been reported to be a risk factor for recurrence [5]. Irregularly shaped tumors have to be ablated with a large margin because most ablation systems are limited to ellipsoidal ablation shapes, which involves ablating large parts of healthy tissue. Especially in the light of the drastically increased quality of life for patients who undergo an ablation vs a resection – they only stay in the hospital for one night instead of two weeks [5] – we aim to develop the procedure and the technology further to accommodate tumors irregularly shaped or larger than 30 mm. One idea to achieve this would be to introduce the concept of ablations that are dynamically modulated in space, time and energy level (Figure 1) to achieve an ablation shape customized to the tumor. In this exploratory study, we investigated the possibility of achieving a superposition of ablation zones by modulating the energy level depending on the position along a trajectory using robotic technology in an ex-vivo study.

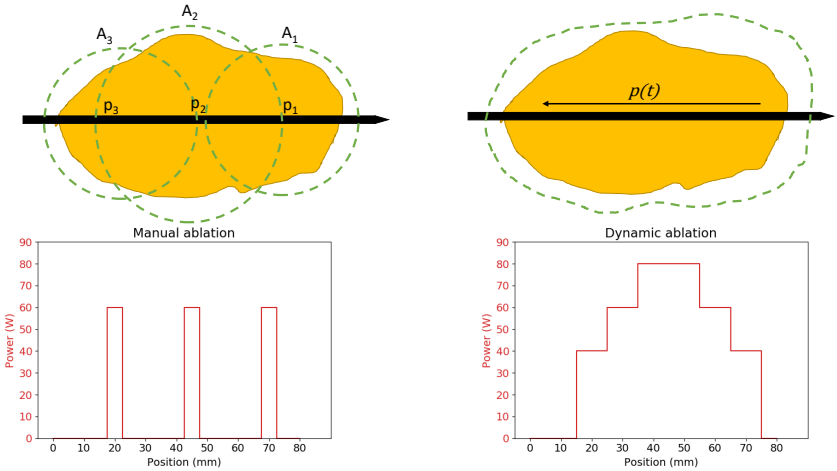



*Figure 1: (left) Concept of manual superposition of ablation shapes by multiple ablations vs. (right) superposition by dynamically modulating the energy depending on time and position*

## 2 Materials and Methods

Based on preliminary work, we have built an interventional robotic platform (**Fehler! Verweisquelle konnte nicht gefunden werden.**) consisting of a dedicated surgical robotic system [6] with an ablation needle (Solero, AngioDynamics, USA) combined with a stereotactic navigation system (CAS-One Vario, CAScination AG, Switzerland). Using this device, we are able to plan an ablation trajectory on a CT image using the stereotactic system. The system then sends the desired trajectory to the robotic device. The patient and instrument position are periodically sent to the robot for visual servoing and correction of the trajectory. With this demonstrator we can combine different means of robotic needle delivery with subsequent control of energy delivery of an ablation system.

In general, the ablation model is defined as the thermal energy induced into the tissue at each voxel as the integral of the energy delivery function and a factor depending on the tissue properties:

$$\vartheta = \int f(t,p,e) \times tp(p)\ dt$$

where $f$ is a function of time ($t$), position ($p$) and the applied energy ($e$) by the ablation device as variable parameters. The tissue properties $tp$ are specific to the tissue at the location of the ablation center.

*Figure 2: Experimental setup with the robotic device, the ablation system and ex-vivo bovine liver*

### 2.1 Experimental evaluation

In an ex-vivo experiment, we evaluated whether a superposition of ablation zones can be achieved by synchronously moving the ablation needle and modulating the energy level. Therefore, we created ablation profiles where the applied energy depends on time and location (Figure 3 bottom). In conventional ablations, location and power are constant over the whole time-period of the ablation and are applied only at one location at a time. These dynamic ablation profiles were then applied to ex-vivo bovine livers using the interventional robotic platform (Figure 2), which moves the ablation needle according to the planned profile. The energy was set manually on the ablation device according to the profile using the provided timer on the screen. We also set the velocity constant, such that the needle travels the specified path during the desired time. The velocity at which the ablation needle was moved during these experiments was between 0.13 and 0.33 mm/s. For simplicity of the setup, we placed the ablation needle manually in the liver specimen and did not use the placement function of the robot. Overall, we performed 8 ablations where we tried the following profile types:

a) Subsequent stationary ablations at two predefined positions (n = 1)
b) Constant energy along a path (n = 2)
c) Increasing and decreasing energy over time along a path (n = 5)

The specific parameters (energy, time and position) were chosen differently for each ablation profile to get a wide variety of shapes. After the ablation profiles were applied, we dissected the specimens along the needle trajectory to visualize the resulting ablation zones. The primary endpoint of this study was the visual assessment of the difference between ablation profiles. Therefore, for (a) we expected to get two disconnected spherical ablation zones, for (b) a long cylindrical ablation and for (c) an elongated ellipsoid.

## 3 Results

In study, we have performed 8 ablations in ex-vivo bovine liver with different ablation profiles, which are shown in (Figure 3). All ablations yielded different shapes, with (a) being the most obvious. The profiles with a constant energy over a path (b) produced an almost rectangular ablation shape, and the increasing/decreasing energy (c) yielded a large elongated ellipsoid. In Figure 3 right, one can also see the heat conducted towards a nearby blood vessel where the needle didn't pass through.

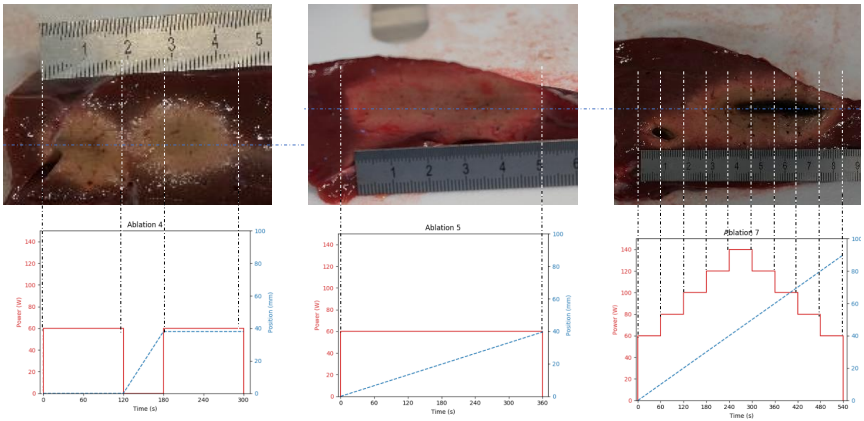



*Figure 3: (a) constant power at two specific locations (b) constant power along a 50 mm long trajectory with constant velocity (c) increasing/decreasing energy along a 90 mm long trajectory*

# 4 Discussion

In this experiment we found, that it is possible to acquire superpositions of ablation zones while moving the ablation needle. In other words, the tissue is heated sufficiently while the needle is moving. Using this approach, we were able to create large and elongated ablation shapes, which would not be possible with static ablations. Combined with multi needle ablations, this could become a practice where the use of a robotic device, would provide a treatment benefit beyond the capabilities of a human.

In general, the produced ablation shapes turned out as expected. However, we expected the cylindrical shapes to have a larger diameter, especially those with a higher energy. It seems that in these cases the needle has to be moved much slower to also heat the tissue further away from the needle. In contrast, the ablation with increasing/decreasing energy levels turned out larger than expected. Especially the central part of the ablation was burned and evaporated, which was unexpected as the needle was constantly moved. This also shows that more experiments are required to understand this process better.

One limitation of this study is the evaluation of the shape by dissection of the specimen. While this is commonly done in other studies, it introduces an error as it depends on how well the specimen is cut in the middle of the trajectory. This can be especially problematic when trying to analyze irregular shapes. Unfortunately, ablation zones in ex-vivo tissues are not visible on CT scans, as it is not possible to inject contrast agent. Therefore, we are searching for other imaging modalities to visualize the ablation zone non-invasively.

As this was an initial exploratory work, we aim to conduct reproducibility studies to verify our findings of this preliminary findings. In particular we will focus on the difference in the ablation shapes depending on the profile. To eliminate the human error due to manual ablation energy setting, we will also integrate the ablation device with our interventional robotic platform to set the ablation energy automatically by the robot. In the next phases, in-vivo animal trials should be carried out, to verify that these dynamic ablations are able to induce sufficient thermal energy in living and perfused tissue.

# 5 Conclusion

Dynamic energy delivery during ablation procedures using robotic technology could achieve reproducible large or irregular ablations customized to the tumor shape. However, more investigation is needed to confirm these findings, as this was a rather exploratory analysis.

# References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68:394–424 . doi: 10.3322/caac.21492
2. Heerink WJ, Ruiter SJS, Pennings JP, Lansdorp B, Vliegenthart R, Oudkerk M, de Jong KP (2019) Robotic versus Freehand Needle Positioning in CT-guided Ablation of Liver Tumors: A Randomized Controlled Trial. Radiology 290:826–832 . doi: 10.1148/radiol.2018181698
3. Beyer LP, Pregler B, Niessen C, Dollinger M, Graf BM, Müller M, Schlitt HJ, Stroszczynski C, Wiggermann P (2016) Robot-assisted microwave thermoablation of liver tumors: a single-center experience. Int J Comput Assist Radiol Surg 11:253–259 . doi: 10.1007/s11548-015-1286-y
4. Ben-David E, Shochat M, Roth I, Nissenbaum I, Sosna J, Goldberg SN (2018) Evaluation of a CT-Guided Robotic System for Precise Percutaneous Needle Insertion. J Vasc Interv Radiol 29:1440–1446 . doi: 10.1016/j.jvir.2018.01.002
5. Beermann M, Lindeberg J, Engstrand J, Galmén K, Karlgren S, Stillström D, Nilsson H, Harbut P, Freedman J (2019) 1000 consecutive ablation sessions in the era of computer assisted image guidance – Lessons learned. Eur J Radiol Open 6:1–8 . doi: 10.1016/j.ejro.2018.11.002
6. Weber S, Gavaghan K, Wimmer W, Williamson T, Gerber N, Anso J, Bell B, Feldmann A, Rathgeb C, Matulic M, Stebinger M, Schneider D, Mantokoudis G, Scheidegger O, Wagner F, Kompis M, Caversaccio M (2017) Instrument flight to the inner ear. Sci Robot. doi: 10.1126/scirobotics.aal4916

# Comparison of Background Removal Approaches in X-ray fluoroscopy for Detection of Cerebral Stent Markers

*N. Chabi*[1], *O. Beuing*[2], *B. Preim*[1], *S. Saalfeld*[2]

[1] *OvG-University, Department of Simulation and Graphics, Magdeburg, Germany*
[2] *OvG-University, Institute of Neuroradiology, Magdeburg, Germany*

Contact: negar.chabi@ovgu.de

***Abstract***

*Stents are mesh tube devices designed for the treatment of endovascular diseases such as aneurysms. Cerebral stents are very small and hardly visible in the acquired X-ray images. To improve stent visibility, they are equipped with radio-opaque markers. Because of the low contrast to noise ratio (CNR) and the very small size of opaque markers, pre-processing is desirable. We provide a comparison of different background subtraction approaches in X-ray fluoroscopy images for the stenting procedure. We compare three algorithms including a method based on digital subtraction angiography (DSA), bottom hat transform and our linear subtraction approach. The algorithm is validated on eight clinical datasets. Then performance evaluation was done through four different criteria including CNR, local contrast, contrast improvement ratio (CIR) and detection ratio (DR). Unlike our method simplicity, based on the visual and numerical assessment, concluded that it has better performance compared to other techniques.*

**Keywords**: background removal, X-ray fluoroscopy, radio-opaque markers, intracranial stent

## 1 Problem

Accurate intracranial stent deployment is a challenging task which requires highly skilled physicians and well suited X-ray images. In order to spare the patient, low-dose X-ray fluoroscopy is usually acquired during these interventions, which is often accompanied by poor stent visibility [1]. In order to improve the visualization, stents are equipped with radio-opaque markers which are better visible in X-ray and eventually help the clinicians to guide the stents more precisely through the vessels [2]. Because of the small size of the markers, their visualization is also hampered by other background structures. Figure 1 depicts an original X-ray image with stent, markers, guide wire and its corresponding background subtracted image.

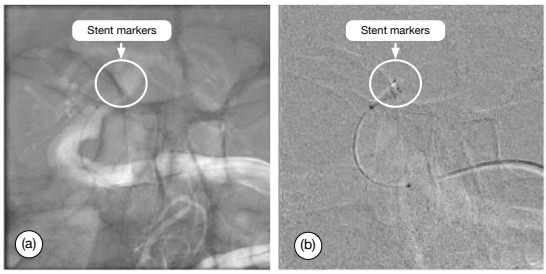



Figure 1: *(a) The original X-ray image from a clinical dataset. (b) Background subtracted image with improved stent marker visualization*

One of the most widely used approaches for background removal is *digital subtraction angiography* (DSA), which is conceptually defined for visualizing blood vessels. Here, an image obtained after contrast agent injection is subtracted from pre-contrast image based on *Beer-Lambert law*[3] which results in removing distracting tissues and renders a clear visualization of blood vessels [4].

Manthey et al. [5] extended DSA for virtual enhancement of opaque markers and employed a blob detection technique to detect the stent markers. They provided the results for some phantom datasets as well as a clinical dataset. However, the major issue here is the lack of complete registration in the real clinical dataset.


Figure 2: *Schematic illustration of the proposed technique; (a) Original image, (b) Mean image, (c) Proposed method (Linear subtracted result)*

Another pre-processing technique which was employed by Bismuth et al. [6] is the bottom hat transform, which extracts dark objects from a light background. It removes the image background and preserves dark objects that are smaller than the specified structural element. In this work, we propose a technique which mostly emphasizes the dynamic parts, including a stent, its markers, guide wire, and catheters while neglecting the static background regions. Static regions are the constant background parts where elimination of these parts help to enhance visualization of the markers and guide wire. This is achieved by subtracting the average of image sequence from the original image itself. Our main purpose is to remove the static parts in the image and to keep the dynamic parts of the image sequence.

## 2 Material and Methods

Our proposed method plus two different background subtraction approaches are explained in Section 2.1- 2.3. Then, the details associated with medical image datasets used for our experiment are illustrated in Section 2.4.

### 2.1 Linear Subtraction algorithm

We aim at enhancing marker visualization by emphasizing parts of the image which are affected by motion. For this purpose, static parts of the image are removed. To do so, the image sequence was subtracted from the mean image of the sequence. If $A(x, y, s)$ and $B(x, y, s)$ denote the original image and enhanced image respectively; $(x, y)$ refers to image coordinates, $s$ and $S$ indicate the slice number and the total number of slices respectively; then the enhancement is defined as:

$$B\left(x, y, s\right) = A\left(x, y, s\right) - \frac{\sum_{i=1}^{S} A\left(x, y, s\right)}{S} \tag{1}$$

A schematic illustration of the proposed technique is shown in Figure 2.

### 2.2 Digital Subtraction Angiography (DSA)

DSA is a fluoroscopy technique which provides a clear visualization of vessels. There are two images, one before contrast agent injection and one after when the contrast media reaches the vessel. DSA images are obtained via logarithmic subtraction of pre-contrast image from acquired subsequent images after contrast media administration. Logarithmic subtraction actually removes static parts of the image. An X-ray image is a result of signal attenuation after passing through the object, where the equation based on *Beer-Lambert law* is defined as:

$$I = I_0 e^{-\mu D} \tag{2}$$

$I$: is the obtained intensity
$I_0$: is the initial signal intensity
$D$: is the full thickness of the object
$\mu$: is the attenuation coefficient Hence, before entering the devices, the equation is as follows:

$$I_A = I_0 e^{-(\mu_A D_A)} \tag{3}$$

Where $D_A$ is the full thickness of the objects and $\mu_A$ represents the absorption coefficient for the objects, that radiation passes through. While in the presence of the devices we have:

$$I_B = I_0 e^{(-\mu_A * (D_A - D_B) + \mu_B D_B)} \tag{4}$$

Where $D_B$ and $\mu_B$ represent the absorption coefficient of the device and its respective thickness. After applying the logarithm on these two equations:

$$\ln(I_A) - \ln(I_B) = D_B * (\mu_B - \mu_A) \tag{5}$$

This equation shows that the pixel value in the subtracted image is different from zero if only the change has occurred in the original image. Hence the subtraction is carried out in the same way as DSA. Based on this concept, DSA not only is useful for vessel visualization in angiography, but also can be extended to the other object detection tasks in X-ray.

### 2.3 Bottom Hat Transform

Bottom hat transform, also called black top hat transform, is defined as the difference between the input image and its closing [7]. It removes the background while maintaining objects which are smaller than the predefined structural element and darker than the surrounding area (See Eq. 6).

$$BottomHat(A) = (A \cdot B) - A \tag{6}$$

$A$ and $B$ refer to the input image and structural element, respectively. $A \cdot B$ refers to closing operation between $A$ and $B$. In this study, a disk-shaped structuring element with a 5-pixel radius size was empirically selected.

### 2.4 Image Acquisition

Eight clinical X-ray angiography datasets were used in this study. Images are acquired at the Institute for Neuroradiology, Magdeburg, Germany with an Artis Q, Siemens Healthineers, Forchheim, Germany. Image size is $512 \times 512$. Frame rate per second is 30. Cerebral Stent placement is done under the guidance of X-ray fluoroscopy which provides clinicians with a real-time video of different devices including a catheter, guide wire, and stent. Different types of stents based on aneurysm type were used. For example, for the first dataset, a DERIVO flow diverter (ACANDIS GmbH & Co. KG, Pforzheim, Germany) was used.

## 3 Results

Performance comparison is based on different metrics including CNR, local contrast measure, CIR and finally the number of detected markers. comparison results are provided in Tables 1- 5. Visual comparisons are shown in Figures 3 and 4.

### 3.1 Contrast to Noise Ratio (CNR)

CNR is the ratio of contrast to the amount of noise, which corrupted the image. Contrast refers to the signal intensity difference between a specified region with its nearby background. CNR is an outstanding measure for quality evaluation of the low contrast region. CNR is defined in Eq. 7 according to [8]:

$$\frac{I_{ROI} - I_{BG}}{\sqrt{\frac{SD_{ROI}^2 + SD_{BG}^2}{2}}}, \tag{7}$$

$I_{ROI}$: is the mean value of pixel intensity in the region of interest.
$I_{BG}$: is the mean value of intensity for the specified background region.
$SD_{ROI}$ and $SD_{BG}$ stand for the standard deviation of ROI and background respectively.
We tested an ROI with window size of $11 \times 11$ pixels (see Table 1) and $5 \times 5$ pixels (see Table 2). In the following Tables 1- 5, DS, I, BH and our refer to the dataset number, the original dataset, bottom hat transform and our proposed method respectively.

Table 1: *CNR by* $11 \times 11$ *window*

| DS | I | DSA | BH | Our |
|---|---|---|---|---|
| 1 | **1.5782** | 0.7835 | 1.3837 | 0.5251 |
| 2 | 1.8442 | **2.3793** | 1.7631 | 1.1958 |
| 3 | **4.3825** | 1.1323 | 1.0747 | 0.7062 |
| 4 | **2.5426** | 1.9577 | 0.9950 | 1.0030 |
| 5 | **5.2949** | 2.1703 | 1.4155 | 1.2337 |
| 6 | 1.8220 | **3.2990** | 1.4829 | 1.1889 |
| 7 | 0.3534 | 0.6164 | **0.8167** | 0.5769 |
| 8 | 3.5542 | **5.6542** | 0.8267 | 0.6833 |

Table 2: *CNR by* $5 \times 5$ *window*

| DS | I | DSA | BH | Our |
|---|---|---|---|---|
| 1 | **4.5880** | 1.6818 | 3.5008 | 1.5775 |
| 2 | 1.8745 | **4.2059** | 3.7786 | 3.1082 |
| 3 | **6.6124** | 2.227 | 2.2960 | 1.6035 |
| 4 | **4.2943** | 3.9793 | 2.4806 | 2.3967 |
| 5 | **8.1762** | 3.6678 | 3.9988 | 3.1171 |
| 6 | **6.1554** | 5.3914 | 5.2616 | 4.9123 |
| 7 | 1.4670 | 1.2321 | **2.7513** | 2.4503 |
| 8 | 5.1544 | **8.0550** | 2.9523 | 1.3999 |

### 3.2 Local Contrast

The other algorithm which is used to measure local contrast is a window-based algorithm, where there are an inner window and an outer window, the center of the windows are located at $(x, y)$ position [9]. It is defined through the following equation:

$$c\,(x, y) = \frac{|s - p|}{|s + p|} \tag{8}$$

$s$ and $p$ are the average value of intensity for inner and outer window, respectively. The window size is selected empirically and based on the marker size, the inner window and the outer window size is assumed 7 and 15 pixels respectively. The results are shown in Table 3.

Table 3: *Local contrast*

| DS | I | DSA | BH | Our |
|---|---|---|---|---|
| 1 | 0.0628 | 0.0784 | 0.0578 | **0.0786** |
| 2 | **0.0655** | 0.0139 | 0.0601 | 0.0643 |
| 3 | **0.0201** | 0.0104 | 0.0080 | 0.0175 |
| 4 | 0.0362 | 0.0194 | **0.6381** | 0.0381 |
| 5 | 0.0348 | 0.0200 | 0.0427 | **0.0452** |
| 6 | 0.1022 | 0.0471 | **0.2694** | 0.1329 |
| 7 | 0.0587 | 0.0199 | 0.0973 | **0.1488** |
| 8 | **0.1149** | 0.0477 | 0.0669 | 0.0195 |

Table 4: *CIR for different approaches*

| DS | DSA | BH | Our |
|---|---|---|---|
| 1 | **0.7909** | 0.4108 | 0.5567 |
| 2 | **0.8805** | 0.2681 | 0.3638 |
| 3 | 0.6286 | **0.7621** | 0.5289 |
| 4 | 0.6573 | **4.4794** | 0.6998 |
| 5 | **0.6483** | 0.4582 | 0.5047 |
| 6 | 0.7163 | **1.2756** | 0.5248 |
| 7 | 0.7611 | 0.8440 | **1.3008** |
| 8 | 0.7572 | 0.6443 | **0.9079** |

### 3.3 Contrast Improvement Ratio (CIR)

CIR is the difference between local contrast of the original image and the background subtracted images in the region of interest. It is calculated via the following equation:

$$CIR = \frac{\sum_{(x,y)\in R} \left(c(x, y) - c'(x, y)\right)^2}{\sum_{(x,y)\in R} c(x, y)^2} \tag{9}$$

Where $R$ is the region of interest, $c$ and $c'$ are the local contrast for initial image and the image after background removal respectively [9]. Results for CIR are presented in Table 4. ROI is assumed the same as Section 3.2

### 3.4 Number of detected markers

A marker detection algorithm [5] was used in order to detect markers for each dataset which conceptually demonstrated that which algorithm is more successful in providing high-quality images where the markers are easier to be specified. The results were acquired on the clinical datasets and are described in terms of the following detection ratio:

$$DR = \frac{Number\ of\ Detected\ Markers}{Number\ of\ Existing\ Markers} \tag{10}$$

Due to low CNR and weak visibility of the stent and the stent markers, the number of detected markers for the third dataset is not acceptable enough. The obtained result for one of the slices are shown in Figure 3. This ratio was calculated for all datasets and is presented in Table 5. Finally, all the approaches are compared visually in Figure 4.Where, Figure 4(a) shows an image from a clinical image sequence with an intracranial stent within the vessel. In this original image, stent and stent markers are almost invisible. To provide a

Table 5: Results after using marker detection technique on the images obtained by 3 different background removal approaches.

| DS | DSA | BH | Our |
|---|---|---|---|
| 1 | 50.00% | **85.00%** | **85.00%** |
| 2 | 7.50% | **62.50%** | **62.50%** |
| 4 | 7.50% | 32.50% | **90.00%** |
| 5 | 7.50% | **92.50%** | 70.00% |
| 6 | **75.00%** | 50.00% | 65.00% |
| 7 | 30.00% | 60.00% | **80.00%** |
| 8 | 10.00% | 22.50% | **37.50%** |

better visualization of the stent placement position, this area is focused. Figure 4(b) and (c) show the result of employing DSA and bottom hat transform on Figure 4(a) respectively. The result of our proposed approach is presented in Figure 4(d).

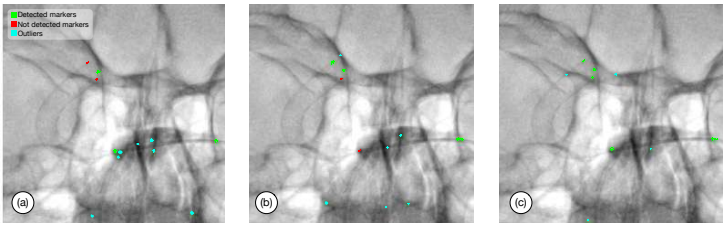



Figure 3: *Results after using the marker detection technique on the images obtained by three different background removal approaches; (a) DSA, (b) Bottom hat, (c) Proposed method*

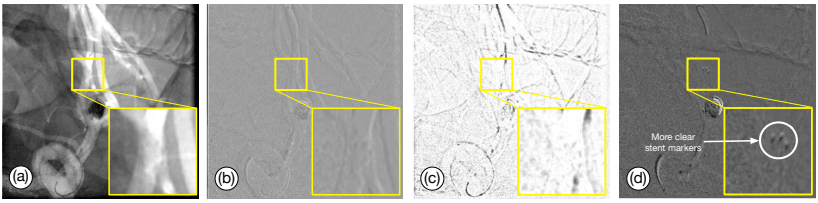



Figure 4: *Results after applying the marker detection technique on background-subtracted images. Detected markers for the different background subtraction techniques are overlaid on the original image. Results for (a) DSA, (b) Bottom hat and (c) Proposed method.*

## 4 Discussion

The main advantage of our proposed approach is its simplicity, since it does not require any additional prerequisite steps. Therefore, it needs lower computational cost as well as it will be free from the errors which are resulted from not complete registration. Whereas, DSA is highly dependent on complete registration.

One of the problems with the bottom hat transform is its dependency on parameters such as structural element size and shape which are required to be adjusted by the target marker size Besides, it does not completely remove the background objects, whereas the proposed method presents a uniform and almost weakened background.

Further assessment based on the results presented in Table 1 and Table 2, demonstrated that CNR relies on the selected window size as well as the dataset. This ratio is lower for $11 \times 11$ compared to $5 \times 5$ window size.

CNR has almost the highest value for the original image itself for different window sizes and after that DSA has the second place.

Although our proposed method does not provide a good result based on CNR, it shows acceptable performance through other quantitative and visual comparisons. As can be seen in Table 3, our method provides better contrast in the proximity of stent markers for the datasets 1,5 and 7, for dataset 4 and 6, increases the local contrast and for the rest of datasets, the original image itself has the highest local contrast.

The CIR metric implies the effectiveness of each technique in improving the local contrast. According to the results in Table 4, it can be concluded that different approaches have almost the same performance in improving contrast.

DR criteria illustrate each method's effectiveness in producing an image with better detectable markers. The higher visibility of the stent markers supports the recognition by the blob detection algorithm [5]. The results presented in Table 5 and Figure 3 validates the higher performance of the proposed method. Based on results shown in Figure 3 and Figure 4, it is clear that linear subtraction outperforms other approaches in providing an image with a higher visibility of the stent markers.

## 5 Conclusion

DSA imaging does not have a good performance in marker enhancement. Its performance highly relies on the success of registration, which is a challenging task. Although the dark top hat transform removes the background, background objects are still visible. Its performance relies on the specified size of structural element. On the other hand, our proposed method seems to omit the static part of the image while preserving the dynamic part of the image including stent, guide wire and its markers. The presented results show that our method performs well enough compared to the other approaches. CNR results emphasizes the importance of noise estimation before and after marker enhancement, therefore, for the future study, It would be interesting to study on noise removal approaches in order to increase CNR.

## 6 Acknowledgement

This work is partly funded by the German Research Foundation (SA3461/2-1) and the Federal Ministry of Education and Research within the Forschungscampus *STIMULATE* (13GW0095A) as well as partly conducted within the *International Graduate School MEMoRIAL* at Otto von Guericke University (OVGU) Magdeburg, Germany, which is kindly supported by the *European Structural and Investment Funds* (ESF) under the programme *"Sachsen-Anhalt WISSENSCHAFT Internationalisierung"* (project no. ZS/2016/08/80646).

## References

[1] Min Seok Lee, Chul Hee Park, and Moon Gi Kang. *Edge enhancement algorithm for low-dose X-ray fluoroscopic imaging.* Computer methods and programs in biomedicine, 152:45–52, 2017.

[2] J Wiskirchen, R Venugopalan, AD Holton, C König, U Kramer, J Trübenbach, G Tepe, CD Claussen, and Duda St H. *Radiopaque markers in endovascular stents-benefit and potential hazards.* In RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren, volume 175, pages 484–488, 2003.

[3] DF Swinehart. *The beer-lambert law.* Journal of chemical education, 39(7):333, 1962.

[4] William R Brody. *Digital subtraction angiography.* IEEE Transactions on Nuclear Science, 29(3):1176–1180, 1982.

[5] Samuel Manthey, Thomas Hoffmann, G. Cattaneo, Oliver Beuing, Bernhard Preim, and Sylvia Saalfeld. *Virtual enhancement of marker X-ray visibility for cerebral stents and flow diverters.* In 16. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie, Hannover, 2017.

[6] Vincent Bismuth, Régis Vaillant, François Funck, Niels Guillard, and Laurent Najman. *A comprehensive study of stent visualization enhancement in X-ray images by image processing means.* Medical image analysis, 15(4):565–576, 2011.

[7] Pierre Soille. *Morphological image analysis applied to crop field mapping.* Image and Vision computing, 18(13):1025–1032, 2000.

[8] R Edward Hendrick. *Breast MRI: fundamentals and technical aspects*. Springer Science & Business Media, 2007.

[9] Hamid Hassanpour, Najmeh Samadiani, and SM Mahdi Salehi. *Using morphological transforms to enhance the contrast of medical images.* The Egyptian Journal of Radiology and Nuclear Medicine, 46(2):481–489, 2015.

# Machbarkeit CNN-basierter Erzeugung von Kandidatenstrahlen für Radiochirurgie der Prostata

*S. Gerlach*[1], *M. Schlüter*[1], *C. Fürweger*[2] *A. Schlaefer*[1]

[1] *Technische Universität Hamburg, Institut für Medizintechnische Systeme, Hamburg, Deutschland*
[2] *Universität zu Köln, Klinik für Stereotaxie und Funktionelle Neurochirurgie, Köln, Deutschland*

Kontakt: stefan.gerlach@tuhh.de

***Abstract***

*Bei der Radiochirurgie wird ein Roboterarm verwendet, um Dosisabgabe aus nahezu beliebig vielen Richtungen zu ermöglichen. Allerdings ist wegen dieser Flexibilität die Behandlungsplanung eine anspruchsvolle Aufgabe. Üblicherweise wird eine Heuristik auf Grundlage randomisierter Kandidatenstrahlen verwendet, um die Anzahl der tatsächlich betrachteten Einstrahlrichtungen zu begrenzen. Im Gegensatz dazu schlagen wir die Verwendung eines Convolutional Neural Networks vor, um Kandidatenstrahlen auf Basis von Behandlungsplänen anderer Patienten zu generieren. Unsere Ergebnisse zeigen, dass dieser Ansatz nur halb so viele Kandidatenstrahlen für eine vergleichbare Planqualität benötigt.*

**Keywords**: Strahlentherapie Behandlungsplanung, Maschinelles Lernen, Strahlengenerierung

## 1 Problemstellung

Bei der Radiochirurgie mit dem CyberKnife wird Dosis mittels fokussierter Strahlen aus vielen Richtungen abgegeben. Die Strahlen erzeugt ein Linearbeschleuniger, der an einem Roboterarm angebracht ist. Dadurch können gute Konformität mit dem geplanten Zielvolumen (PTV) und ein steiler Dosisgradient am Rand des PTVs erreicht werden. Bei der Behandlungsplanung werden dann die optimalen Strahlen bestimmt. Dazu hat sich eine Heuristik auf Grundlage randomisiert verteilter Kandidatenstrahlen als robust erwiesen. Die Auswahl und Gewichtung der Strahlen wird gewöhnlich durch das Lösen eines linearen Optimierungsproblems realisiert [1]. Allerdings steigt die Optimierungszeit überproportional zu der Anzahl an Kandidatenstrahlen. Darüber hinaus existieren Ähnlichkeiten zwischen den Organstrukturen verschiedener Patienten. Daher schlagen wir ein Convolutional Neural Network (CNN) vor, um den Einfluss von Strahlen auf die Dosisabgabe vorherzusagen und basierend darauf Kandidatenstrahlen auszuwählen. Diese Methode vergleichen wir mit dem Erzeugen von randomisiert verteilten Kandidatenstrahlen.

## 2 Material und Methoden

Computertomographie (CT) Scans der Patienten bilden die Grundlage für die Behandlungsplanung. In den Scans sind das PTV und besonders zu schützende Organstrukturen (OARs) konturiert. Wir erweitern den üblichen Ablauf des Generierens eines Behandlungsplans, indem wir randomisiert verteilte Strahlen mithilfe des CNNs bewerten und auf Grundlage dieser Bewertung auswählen. Ein üblicher Ansatz, um den Einfluss von Strahlen auf die Dosisabgabe bei der Strahlentherapie zu beschreiben, ist die Projektionen der betrachteten Organstrukturen (VOIs) aus der Sicht des korrespondierenden Strahlenursprungs (Strahlenknoten) auf eine Ebene senkrecht zum Zentralstrahl [2]. Eine ähnliche Methode wurde bei einem Case-Based-Reasoning-Ansatz verwendet [3], bei dem Ähnlichkeit zwischen Fällen unter anderem über Projektionen von VOIs inferiert wurde. Vergleichbar zu diesem Ansatz führen wir eine Projektionsebene senkrecht zu der Linie vom Strahlenknoten zum PTV-Schwerpunkt ein. Auf diese Ebene projizieren wir den zu bewertenden Strahl und die VOIs. Als Intensitätswerte der Projektion werden die minimale und maximale radiologische Tiefe verwendet, welche zusätzliche Information über die Form und Lage der VOIs bieten. Die radiologische Tiefe beschreibt den Einfluss, den die Abschwächung des bereits durchlaufenen Gewebes auf die Dosisabgabe hat. Für die Behandlungsplanung betrachten wir die Prostata als PTV, sowie Rektum und Blase als OARs. Abb. 1 zeigt die Projektion des Strahls und jeweils ein Beispiel für die Projektion der minimalen radiologischen Tiefe der VOIs.

Diese Projektionen werden anschließen in der Kanaldimension konkateniert, in der sonst beispielsweise die Farbwerte von RGB-Bildern abgebildet werden. Für Evaluierung und Training nutzen wir Daten von zehn Patienten, von denen vier für das Bestimmen der Hyperparameter mit 4-facher Kreuzvalidierung und sechs zum Evaluieren mit 6-facher Kreuzvalidierung eingeteilt werden bei konstanter Anzahl von 15 Trainingsepochen.

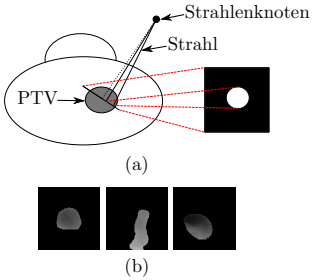



Abbildung 1: *Projektion eines Strahls (a) und der minimalen radiologischen Tiefe von Prostata, Rektum und Blase (b, von links nach rechts)*

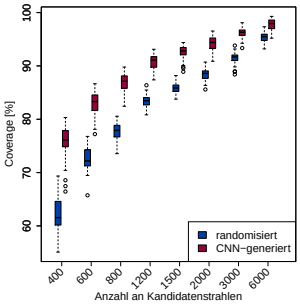



Abbildung 2: *Vergleich der Coverage für randomisiert verteilte und CNN-generierte Kandidatenstrahlen*

Für das Trainingsset wurden pro Patient 30 Behandlungspläne mit jeweils 6000 randomisierten Kandidatenstrahlen optimiert. Das Optimierungsziel war die Maximierung der Coverage des PTVs, welche den Anteil des PTVs beschreibt, der mindestens die verschriebene Dosis erhält. Als Label für einzelne Strahlen wird das jeweils durch das Optimieren zugeteilte Gewicht relativ zu dem maximal pro Strahl erlaubten Gewicht definiert. Für die Strahlenauswahl wird dann das prädizierte relative Gewicht als Annahmewahrscheinlichkeit eines neuen zufälligen Strahls in die Menge der Kandidatenstrahlen verwendet. Anschließend werden die Kandidatenstrahlen durch das Lösen des Optimierungsproblems gewichtet.

## 3 Ergebnisse

Wir werten Behandlungspläne für unterschiedliche Anzahlen an Kandidatenstrahlen aus und vergleichen randomisiert generierte Kandidatenstrahlen mit CNN-generierten Kandidatenstrahlen. Außerdem wird die Berechnung der Behandlungspläne mit jeweils zehn Strahlenmengen wiederholt, um die statistische Signifikanz der Ergebnisse zu erhöhen. Abb. 2 zeigt die berechnete Coverage. Für die gleiche mittlere Coverage werden beim CNN Ansatz nur etwa halb so viele Kandidatenstrahlen wie bei der randomisierten Generierung benötigt. Umgekehrt ist die erreichbare Coverage bei gleicher Größe des Optimierungsproblems deutlich größer. Beispielsweise lässt sich auch bei 6000 Kandidatenstrahlen die Coverage von 95,35 % auf 97,67 % verbessern.

## 4 Diskussion und Zusammenfassung

Das Vorhersagen des Einflusses von Kandidatenstrahlen auf die Dosisabgabe mithilfe von CNNs ist machbar und kann dazu genutzt werden, um bei vergleichbarer Planqualität die Anzahl der Kandidatenstrahlen gegenüber der etablierten Methode der Strahlengenerierung um die Hälfte zu reduzieren.

## Danksagung

Diese Arbeit wurde teilweise von der DFG (Förderkennzeichen SCHL 1844/3 – 1) gefördert.

## Referenzen

[1] Schlaefer, Schweikard A, *Stepwise multi-criteria optimization for robotic radiosurgery*, Medical Physics, 35(5), 2094-2103 (2008)

[2] Kalet IJ, Austin-Seymour MM, *The use of medical images in planning and delivery of radiation therapy*, Journal of the American Medical Informatics Association, 4(5), 327-339 (1997)

[3] Schlaefer A, Dieterich S, *Feasibility of case-based beam generation for robotic radiosurgery*, Artificial Intelligence in Medicine, 52(2), 67-75 (2011)

# Combination adapter with switchable collimator for gamma-ultrasound guided surgery of sentinel lymph nodes

*A. Pashazadeh[1], T. Karkhanis[1], M. Kalmar[1], A. Boese[1], M. Friebe[1]*

*[1]Chair for image guided procedures, Otto-von-Guericke University, Magdeburg, Germany*

Contact: ali.pashazadeh@ovgu.de

***Abstract***

*Sentinel lymph node (SLN) biopsy is an intraoperative technique to determine the status of a primary tumor. The procedure is based on the use of an external gamma probe, which generates an audio signal based on the direction and intensity of gamma radiation to guide the surgeon to the location of the SLNs. This method, however, does not provide any anatomical information and is also unable to distinguish the SNL in case of several lymph nodes that are close to each other. In the current study, a demonstrator model of an adaptor containing a gamma probe, with a switchable collimator, is presented to be attached to an ultrasound (US) probe and with that allowing combined gamma-US scanning of the SLNs. Our experimental results showed that the quality of the US image acquired through the CB is comparable with that of normal direct ultrasonography. Through the use of the CB, the distance of the gamma detector increases, but the ability to switch the working mode of the probe between high sensitivity and high resolution improved the overall performance.*

**Keywords**: gamma probe, US probe, sentinel lymph node, collimator, high sensitivity, high resolution

## 1. Introduction

Sentinel lymph node studies provide valuable information regarding the metastasis of the primary tumor, and a basis for the treatment planning of the patient. The procedure is based on the injection of a gamma-emitting pharmaceutical close to the primary tumor and to subsequently scan the injection area using a gamma scanner. Dual-modality SPECT/CT imaging has shown to improve the outcome of the procedure because of the excellent anatomical localization of the SLNs. Inside an operating room, however, the localization of the SLNs is performed using a 1D gamma probe, which generates a sound signal to guide the surgeon. With that, the directional information and guidance of this intra-operative procedure are based on audio, not on imaging information, and additionally suffers from the lack of anatomical information. To obtain intraoperative anatomical information, we combined a gamma probe with ultrasound (US) technology [1]. The basis for this integration is the use of a sound reflecting plate placed in an angled position inside an adaptor to direct and overlay the US wave with the scanning direction of the gamma probe. This configuration allows for the gamma probe to be placed on the adaptor aligned to the scanning plane. While technically feasible, this configuration leads to degrading the performance of the gamma probe mainly due to an extended source-to-detector distance. A gamma probe with high sensitivity (HS) collimator may allow to collect more photons and compensate part of the sensitivity loss. However, this degrades the spatial resolution of the probe. On the other hand, a high resolution (HR) collimator may compensate for resolution loss but at the cost of significantly reduced sensitivity. To improve both sensitivity and spatial resolution of the gamma probe when used with the CB, the adaptor design was developed with a two-mode collimator included. This collimator is a lead plate, containing two circular openings of different diameters. The study aimed to assess the effect of the two-mode collimator in improving the overall performance of the gamma probe. The signal to noise ratio of the US imaging is likely affected negatively as well because of the longer sound travel distance, additional sound barriers, and the associated signal loss due to that. Therefore, the quality of US imaging through the CB was also evaluated.

## 2. Materials and Methods

The CB box was designed in CAD software (SolidWorks 2018, Dassault Systèmes) and printed using a 3D printing machine (Form 2, Formlabs Inc., USA). A 58×18×0.5 $mm^3$ aluminum plate was placed in the base part of the CB with an angle of 60°. This configuration allowed the US wave to be directed perpendicularly to the scanning surface while the US probe is used in 30° angle. It was then filled with gelatin to provide a traveling medium for

the US wave. The upper part contained a cylindrical hole to place the gamma detector, under which, a 42×11×3 $mm^3$ lead plate was placed as a two-mode collimator. Two circular openings with diameters of 2 mm and 5 mm and a separation distance of 15-mm were drilled on the lead plate to form HS and HR front-collimator for the detector. To test the quality of US imaging through the CB, we used a linear US probe (GE Logiq e, GE Healthcare, Wauwatosa, WI, USA) to scan a custom-made gelatine US phantom containing four grapes. The probe was used at 9 MHz frequency and scanned the phantom through the CB. Then the experiment was repeated only using the US probe. A gamma probe (HiSens, Crystal Photonics, Berlin, Germany) was used to scan a 33 kBq Co-57 radioactive point-source placed under a 7mm PMMA plate. The gamma detector, inside the CB, scanned the radioactive source laterally from -20 mm to +20 mm off-axis of the source, at 5-mm steps. The on-axis measured count (x=0) was reported as the sensitivity of the probe. Full-width-half-maximum of the curve of the collected data was calculated and reported as the spatial resolution of the probe.

## 3. Results

The general concept of the CB and the obtained US images of the phantom are presented in Fig. 1. All the targets displayed in the normal US image are also visible in the US image taken through the CB. Relative SNR of the US image obtained with the CB was 93.15% of the normal US image, obtained without the CB. At the upper part of the US image obtained by the CB, bubbles inside the gelatin are seen. Also, because of the extra path that the US wave travels inside the gelatin of the CB, grapes are seen deeper in the image taken by the CB. Performance test results of the gamma probe, when used in HS and HR modes, are given in table 1.

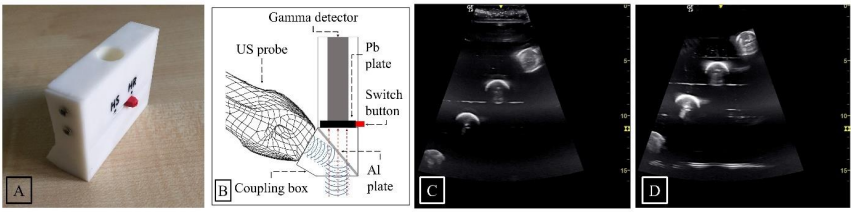



*Figure 1: The CB (A), concept illustration (B), and US images with (C) and without (D) the CB*

*Table 1: Sensitivity and spatial resolution of the gamma probe inside the CB in HS and HR modes*

| Performance parameter of the gamma probe | HS mode | HR mode |
|---|---|---|
| Sensitivity (cps/kBq) | 10.97 | 5.18 |
| Spatial resolution (mm) | 21.26 | 11.59 |

## 4. Discussion

The complementary features of radionuclide scanning and US imaging for SLN localization motivated us to develop an adaptor to integrate a gamma probe with US technology. Despite the expected negative effect of the CB on the normal function of the gamma probe, the two-mode collimator could somewhat compensate for the loss of sensitivity and spatial resolution. Using this collimator, the probe is first set to HS mode to scan the target area in a short time; then it is set at HR mode to improve the accuracy of SLN localization. The current study proved that this two-mode collimator is almost twice as efficient as a single mode collimator when the gamma probe scanned the source through the CB. Quality of US imaging through the CB was comparable with that of normal ultrasonography, and the observed loss of 6.85% in US SNR is most likely not relevant for this application.

## References

[1] A. Pashazadeh, O. Grosser, M. C. Kreissl, D. Kupitz, A. Boese, A. Illanes, M. Friebe, *A concept to combine an ultrasound probe with a gamma probe for accurate localization of sentinel lymph nodes: a feasibility study for future developments*, European Journal of Nuclear Medicine and Molecular Imaging (under review)

# Vergleichsstudie von objektbasiertem Hochgeschwindigkeits-Tracking der Glottis

*S. Ihler[1] *, J. Seifert[1] *, M.-H. Laves[1], L. A. Kahrs[2,3], M. Ptok[4], T. Ortmaier[1]*

[1] *Leibniz Universität Hannover, Institut für Mechatronische Systeme, Hannover, Deutschland*
[2] *Centre for Image Guided Innovation & Therapeutic Intervention, SickKids, Toronto, Kanada*
[3] *Dept. of Mathematical and Computational Sciences, University of Toronto Mississauga, Toronto, Kanada*
[4] *Medizinische Hochschule Hannover, Klinik für Phoniatrie und Pädaudiologie, Hannover, Deutschland*
* *Beide Autoren haben zu gleichen Teilen zu diesem Skript beigetragen*

Kontakt: sontje.ihler@imes.uni-hannover.de

***Abstract*** *Beim bildbasierten Tracking der Glottis kommt es unter Anderem durch Schluckbewegung zu Gewebedeformation, mit temporärer Verdeckung von laryngealen Strukturen. Punktbasierte Tracking-Methoden sind deshalb nur eingeschränkt geeignet. Wir untersuchen, ob generische Objekt-Tracker geeignet sind, die Glottis als visuelle Einheit zu tracken. Ziel ist Robustheit gegenüber partieller Verdeckung bei gleichzeitig hohen Prozessierungsraten. Ein schnelles und genaues Feedback ist essentiell für die Patientensicherheit. Ein sehr gutes Ergebnis auf unserer Testsequenz liefert der MOSSE-Tracker mit einer hohen, mittleren Genauigkeit, hohen Robustheit sowie sehr geringen, mittleren Rechenzeit (0,003 s ≙ 333,3 Hz). Dies ermöglicht ein robustes Tracking auch bei partieller Verdeckung durch Schluckbewegung mit niedrigen, rechenzeitbedingten Latenzen.*

**Keywords**: Objekt-Tracking, Larynx, Computerassistierte Chirurgie

## 1 Problemstellung

Die Strukturen des Kehlkopfes, insbesondere die Stimmlippen, sind u. A. durch die Respiration, spontane Schluckakte und den Schluckreflex bei wachen, aber auch bei sedierten Patientinnen und Patienten in ständiger Bewegung. Resezierende mikrochirurgische oder laserchirurgische Eingriffe, die immer einen optimalen Erhalt gesunden Gewebes anstreben, werden daher bevorzugt in Narkose durchgeführt. Die narkoseinduzierte Ruhigstellung des Zielgewebes hat jedoch den Nachteil, dass während oder direkt nach der Resektion eine Funktionsprüfung, d.h. eine Inspektion der Stimmlippenschwingungen, nicht möglich ist. Die Zielvorgabe für neue, zu entwickelnde Interventionen beinhaltet somit eine geweberespektierende Operation in Lokalanästhesie mit der Möglichkeit der intraoperativen Funktionskontrolle. Eine mögliche Lösung bietet eine bildgestützte Laserchirurgie, die in der Lage ist, Bewegungen mit hoher Präzision und Geschwindigkeit zu erkennen und kompensieren.

Klassischerweise werden in der Medizintechnik punktbasierte Tracking-Ansätze für die bildbasierte Strukturverfolgung eingesetzt. Durch Schluckbewegung kommt es jedoch zu partieller/temporärer Verdeckung von Strukturen, weshalb diese Verfahren für die Anwendung im Larynx nur begrenzt geeignet sind. Ein hier sinnvoller Ansatz ist ein globaleres Tracking durch Segmentierung [1,2]. Aktuelle Umsetzungen sind jedoch nicht auf hohe Geschwindigkeit ausgelegt. In diesem Beitrag soll deshalb untersucht werden, ob generische Objekt-Tracker geeignet sind, die Glottis als visuelle Einheit zu verfolgen und dabei eine hohe Robustheit und Geschwindigkeit gewährleisten, um auch bei nicht vermeidbaren Bewegungen sicher tracken zu können. Diese könnten dann mit punktgenauen Trackern kombiniert werden um Robustheit bei Verdeckungen zu ermöglichen.

## 2 Material und Methoden

Um die Robustheit von Objekt-Tracker in laryngealer Anwendung zu überprüfen, haben wir sieben Objekt-Tracker auf eine endoskopische Aufnahme des Larynx angewandt. Die Testsequenz ist ein Ausschnitt eines öffentlich zugängiges Videos [3], weist eine Länge von 440 Bildern (Frames) und eine regelmäßige Deformation der Glottis durch Schlucken auf. Die untersuchten Objekt-Tracker sind BOOSTING [4], Multiple Instance Learning (MIL) [5], der Kernelized Correlation Filter (KCF) [6], MedianFlow [7], Tracking, Learning and Detection (TLD) [8], Minimum Output Sum of Squared Error (MOSSE) [9] und der Discriminative Correlation Filter (CSRT) [10]. Diese basieren alle auf einem Kernel-Tracking . Die Tracker sind in der Bibliothek OpenCV 3.4.1 bereits implementiert, was eine gute Verfügbarkeit gewährleistet. Alle Tracker erreichen eine Prozessierungsrate $> 10$ fps auf der CPU (hier: Intel i5-4670K 3,40GHz). Für die Auswertung wurde die Lage der Glottis im ersten, sowie in jedem zehnten Frame der Testsequenz manuell bestimmt. Die Genauigkeit wird über den Intersection over Union (IoU) ermittelt, welcher sich aus der Schnittmenge geteilt durch die Vereinigungsmenge des Tracking-Ergebnisses und der manuellen Referenzlage berechnet. Um auf die Robustheit zu schließen, wurde anschließend eine Langzeitbetrachtung der Genauigkeit über 450 Frames durchgeführt.

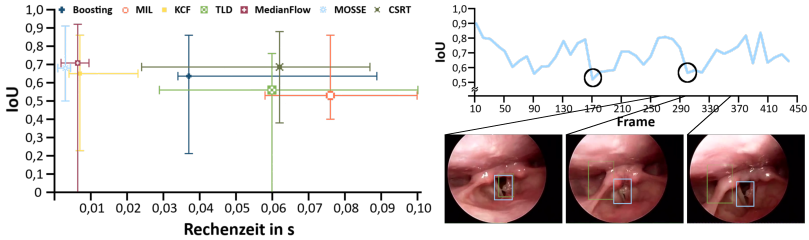



Abbildung 1: *Tracking-Ergebnisse. Links: Mittlere Genauigkeit (IoU) und Rechenzeit über Testsequenz. Die Antennen zeigen Grenzen gemessener Werte. Rechts: Sequentielle Genauigkeit des MOSSE-Trackers (hellblau) im Verlauf der Testsequenz. Kreise kennzeichnen Frames mit maximaler Deformation. (TLD-Tracker in hellgrün)*

## 3 Ergebnisse

Abb. 1 links zeigt die mittlere Genauigkeit aller Objekt-Tracker auf der Testsequenz gegenüber der Rechenzeit. Je höher der IoU, desto genauer ist das Tracking-Ergebnis. Wir stellen fest, dass der MOSSE-Tracker eine sehr geringe mittlere Rechenzeit (0,003 s $\hat{=}$ 333,3 Hz) und hohe mittlere Genauigkeit (IoU: 0,69) mit vergleichsweise geringer Streuung aufweist. Abb. 1 rechts zeigt die Robustheit des Mosse-Trackers. Unten ist ein qualitativer Tracking-Vergleich des TLD- und des MOSSE-Trackers für die Frames 270, 298 und 360 gegenübergestellt. Die obere Abbildung zeigt das Tracking-Genauigkeit des MOSSE-Trackers über den Verlauf der Testsequenz. Es wird deutlich, dass sich der MOSSE-Tracker nach jedem Einbruch der Genauigkeit wieder reinitialisiert, auch nach einer Deformation der Glottis (bspw. Frames 170 und 298).

## 4 Zusammenfassung und Ausblick

Erste Ergbenisse zeigen, dass es möglich ist, die Lage der Glottis mittels generischer Objekt-Tracker über einen längeren Zeitraum zu verfolgen. Auch bei Deformation konnte stabiles Tracking erreicht werden. Der MOSSE-Tracker hat sich auf Grund seiner hohen Genauigkeit und Rechengeschwindigkeit als besonders geeignet herausgestellt, die Glottis lässt sich mit einer Bildrate von 333 Hz verfolgen bzw. rechenzeitbedingte Latenzen können auf ca. 0,003 s begrenzt werden. Schnelles Feedback zur Bewegungsänderung ist essentiell für die Patientensicherheit. Eine ausführlichere Untersuchung auf mehr Bilddaten ist sinnvoll. Bei weiterhin guten Ergebnissen, könnte die Fusionierung mit einem punktgenauen Trackingverfahren folgen - mit dem Ziel die Robustheit durch Wissen über den globalen Kontext, somit partieller/temporärer Verdeckungen, zu erhöhen.

## 5 Referenzen

[1] Andrade-Miranda, G. et al., *An automatic method to detect and track the glottal gap from high speed videoendoscopic images*, BioMedical Engineering OnLine **14(1)** S.100 (2015)
[2] Laves, MH. et al., *A dataset of laryngeal endoscopic images with comparative study on convolution neural network-based semantic segmentation*, IJCARS, (2019)
[3] Bedarkar, S., *vocal cord benign lesion*, https://www.youtube.com/watch?v=25n9nFrWUOQ
[4] Grabner, H. et al., *Real-Time Tracking via Online Boosting* BMVC, S.47-56 (2006)
[5] Babenko, B. et al., *Visual tracking with online Multiple Instance Learning*, CVPR, S.983-990 (2009)
[6] Henriques, J. F. et al., *High-Speed Tracking with Kernelized Correlation Filters*, TPAMI **37(3)** S. 583-596 (2015)
[7] Kalal, Z. et al., *Forward-Backward Error: Automatic Detection of Tracking Failures*, ICPR, S.2756-2759 (2010)
[8] Kalal, Z. et al., *Tracking-Learning-Detection*, TPAMI **43(7)** S. 1409-1422 (2012)
[9] Bolme, D. S. et al., *Visual object tracking using adaptive correlation filters*, CVPR, S.2544-2550 (2010)
[10] Lukezic, A. et al., *Discriminative Correlation Filter Tracker with Channel and Spatial Reliability*, IJCV **126(7)** S.671-688 (2018)

**Diese Arbeit wurde als Teil des EFRE OPhonLas Projekt von der Europäischen Union finanziert.**

# Treatment outcome validation tool for radiofrequency ablations of spinal metastases

*G. Hille*[1], *M. Becker*[2], *S. Saalfeld*[1], *K. Tönnies*[1]

[1] *Department of Simulation and Graphics, University of Magdeburg, Magdeburg, Germany*
[2] *Department of Neuroradiology, University Hospital of Magdeburg, Magdeburg, Germany*

Kontakt: georg.hille@ovgu.de

***Abstract***

*This paper presents an outcome validation tool for radiofrequency ablation of spinal metastases. For this purpose, an automatic image registration of pre- and post-interventional spine MR images was implemented, followed by the calculation of multiple validation measures, such as tumor coverage or distance between the necrosis zone and surrounding risk structures. We tested our tool in 30 clinical patient cases and achieved a mean registration accuracy of 1.57 mm, while the computational time never exceeded 15 s. The results of our treatment outcome validation tool showed that no tumor was fully covered by the necrosis zone and the mean size of the residual tumor was 44.1 % w.r.t. the pre-interventional size of the metastasis. An accurate and reliable outcome validation tool provides the physician with various quantitative and reproducible information as well as visualisations to assess the treatment outcome. Furthermore, it can give valuable feedback for the improvement of therapy and simulation procedures by matching predictions with the actual outcome.*

**Keywords**: Outcome validation, Radiofrequency ablation, Image Registration

## 1 Introduction

Due to the enhancement of medical treatment and diagnostic procedures, life expectancy has increased steadily over the last decades. However, this lifetime gain promotes also age-related diseases like cancer and cancer induced malicious metastases. Beside liver and lung, bones are the third most structure affected by metastases and thereof up to two thirds are located in the spine [1]. Radiofrequency ablation (RFA) is a minimally invasive procedure for mostly palliative treament of such painful spinal metastases. For this purpose, electrodes heat up local tissues by high frequency current phase changes, leading to temperatures above 70 °C and therefore, coagulation necrosis. A significant reduction of the tumour mass with relief from spinal cord and nerve root compression is the primary goal of such RFA treatments. To assess the quality of those tumor and metastasis ablations, it is essential to compare pre-interventional and follow-up scans and evaluate the congruence of both the ablation zone and the former metastasis via distance and volume overlap measures. Unfortunately, this validation task is not straightforward. Currently it is done by seperately examining and mentally matching the pre- and post-interventional images, without any computer-assisted image fusion or superimposed segmentation masks. This means, however, that the current procedure does not support objective, quantifiable and reproducible conclusions about the RFA outcome. Moreover, often only small contrasts between necrosis and residual tumour tissue as well as inflammation processes due to the coagulation effects hamper the assessment of tumor coverage. An appropriate computer-assisted validation tool can help to remedy this lack in clinical practice and support the physician in his mandatory assessment of the applied RFA treatment. Some studies present such tools, but almost all were applied to RFAs of hepatic tumors [2, 3] and computer tomography (CT) images [4, 5, 6, 7]. A main feature of these tools are different visualisations as qualitative treatment outcome validation [6, 7], some of them additionally provide the physician with quantitative measures [4, 5]. To our best knowledge, so far no approach for RF ablations of spinal metastases have been developed, even though Dresky et al. [8] stated tumor coverage measures, while optimizing generator modulation and electrode positions for RFA of spine metastases. Unfortunately, they did not describe any registration procedure nor stated the accuracy, which is crucial for such a validation tool to be reliable and trustful, as all of the following calculations of treatment outcome are based on these results.

In this work, we present a tool for visual and quantitative assessment of treatment outcome by a fully automatic image registration approach for pre- and post-interventional magnetic resonance (MR) images of patients, who underwent RFAs of spinal metastases. Following a segmentation of the metastasis and the corresponding necrosis zones, we automatically register both image volumes and provide the physician with relevant validation measures including tumor coverage, residual tumor volume, volume of necrotisized non-cancerous tissue and the distance between the centers of pre-RFA metastasis and post-RFA necrosis zone. Furthermore, we also enable the inclusion of risk structures, e.g. the spinal canal, in the validation, since the minimum distance between the necrosis zone and the risk structure is of particular interest.

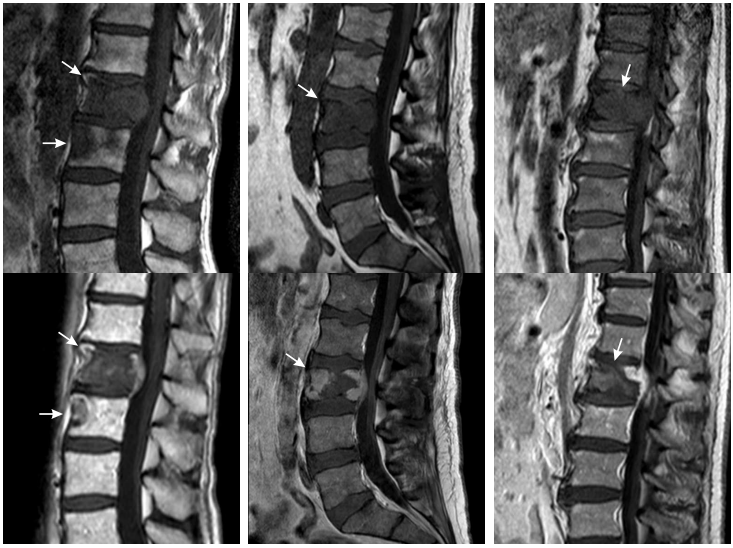


Figure 1: Original pre- (upper row) and post-interventional (lower row) MR images on which our method has been applied. Exemplarily shown are $T_1$-weighted (pre-RFA) and contrast-enhanced $T_1$-weighted (follow-up) scans of patient cases 7, 12 and 22 (from left to right) with corresponding metastases and necrosis zones (arrows).

## 2 Materials and Methods

30 patients who underwent RFAs of spinal metastases were chosen retrospectively. For diagnostic and treatment planning purposes spine MR imaging was performed, containing sagittal and axial native $T_1$- and $T_2$-weighted sequences as well as contrast-enhanced $T_1$-weighted images. Follow-up MR scans contained the same imaging sequences. We assembled our evaluation set consisting of pre-interventional $T_1$ and follow-up contrast-enhanced $T_1$ sequences, since the segmentation of metastasis, respectively necrosis, was manually performed in these particular images of each patient by a neuroradiologist (see Fig. 1). The in-plane image resolution ranged from 0.47 mm to 1.25 mm (average 0.63 mm) and the slice spacing was 3.30 mm for all scans.

The segmentation tasks can be performed within our presented tool and with the following step we can implement a fully automatic registration algorithm. Our method was initialized by a superposition of the centers of mass of both segmentation masks. Subsequently, a multi-resolution image-based rigid registration approach precisely registers both corresponding image volumes combining mutual information (MI) and a Quasi-Newton optimizer. MI bining was set to 32 and depending on the original image resolution we used 4 to 6 multi-resolution levels, since this fits our data the best.

Following the image fusion step, different measures were automatically calculated, which are essential for treatment outcome validation (see Fig. 2). Thus, the residual tumor size is an important indicator, defined by all voxels belonging to the pre-interventional tumor mask and which are not included in the contoured necrosis zone after image registration. A percentage index $T_R$ was computed with respect to the original tumor size. Furthermore, surrounding non-cancerous tissue should be spared as much as possible, which was also considered as a validation index $N_{NC}$. It is defined as the fraction of necrotisized non-cancerous tissue with regard to the total ablation volume. An distance $d_C$ between the centers of mass of metastasis and necrosis zone was computed, providing a measure to assess RFA applicator placement and the resulting ablation zone. Finally, a minimal distance $d_R$ between the necrosis zone and any adjacent risk structure was computed. For qualitative validation purposes, a slice-wise visualisation including superimposed segmentations of the metastasis and the corresponding necrosis and a 3D rendered scene is provided. Additionally, segmentations of different risk structures can be integrated and visualized, e.g., to determine the minimum distance between the necrosis zone and the spinal canal.

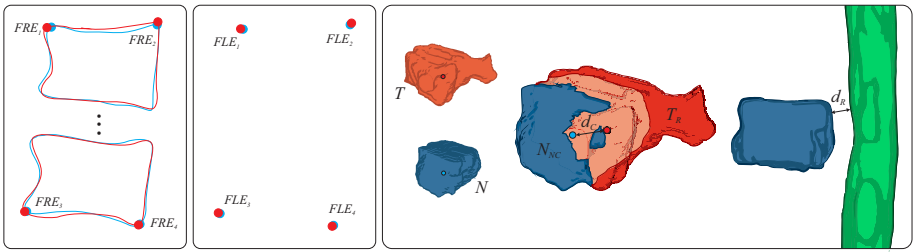



Figure 2: For the purpose of registration accuracy assessment two different types of registration errors were computed. The FRE measured the distance between corresponding fiducial pairs within the target volume after the image-based registration. The FLE is the minimal possible registration error introduced by localisation errors of the user due to the difficulty of identifying identical anatomical locations in different image volumes. The treatment outcome validation primarily based on the residual tumor volume $T_R$, the fraction of necrotisized non-cancerous tissue $N_{NC}$, $d_R$ as the minimal distance between the necrosis zone and a risk structure (e.g. spinal canal) as well as $d_C$ as the distance between the centers of mass of the metastasis and necrosis after the registration process.

# 3 Results

## 3.1 Image registration

Since image registration may introduce registration errors, we examined the accuracy for each patient case via landmarks manually placed on the vertebrae adjacent to the metastasized vertebra (see Fig. 2). Thus, a field expert marked four fiducials within the sagittal plane of symmetry of the vertebra, two on the cranial endplate of a vertebra above and two on the caudal endplate below the metastasized vertebra. For this purpose, we interpolated the slice spacing to provide a quasi-isotropic spatial resolution in order to reduce the fiducial localization error (FLE) to a minimum. Registering both fiducial sets of each patient directly via Horn's quaternion-based method resulted in a FLE of average $0.72 \pm 0.15$ mm, meaning this is the minimum possible registration error that can be achieved while measuring accuracy with the given fiducials due to user localization errors. The average fiducial registration error (FRE), i.e. the accuracy of our image-based registration approach was computed between each pair of corresponding fiducials per patient via Euclidean distances and was overall mean $1.57 \pm 0.32$ mm. The registration procedure never exceeded 15 s.

## 3.2 Outcome Validation

Table 1 lists the results obtained with the outcome validation tool for each patient. The metastasis size of our clinical dataset ranged from 2.85 $cm^3$ to 52.88 $cm^3$ with a mean of $24.84 \pm 13.28$ $cm^3$. In less than 15 % of our patient cases, the necrosis zone was larger than the corresponding metastasis, resulting in a mean size of $17.06 \pm 8.69$ $cm^3$ (min. 3.83 $cm^3$, max. 42.11 $cm^3$). Since the primary goal of these palliative interventions was not a total tumor coverage with an additional safety margin, but rather tumor control and mass reduction, there was no patient case with an entirely covered tumor (the closest to this were 98.1 % coverage of patient 22, see Table 1). On average $44.1 \pm 20.82$ % of the original tumorous tissue was not covered by the indicated necrosis zone in the follow-up scans and therefore, likely to remain untreated. In addition, on average $23.26 \pm 24.87$ % of the necrosis zone voxels were not part of the metastasis mask and therefore, can be labeled as necrotisized non-cancerous tissue. The minimum distance $d_R$ between necrosis zone and the spinal canal as an adjacent risk structure was on average $3.58 \pm 1.86$ mm and there was no overlap between the two structures in any patient case. The distance between the centers of mass of each metastasis and corresponding necrosis was on average $3.05 \pm 2.84$ mm with a mininal distance of 0.17 mm (patient 6) and a maximal distance of 13.46 mm (patient 13).

# 4 Discussion

A pre- and post-interventional image registration with preceding segmentation procedures can be a valuable tool for efficient and objective treatment outcome validation. Figure 3 shows the user interface of our presented

Table 1: *Quantitative validation of RFAs of spinal metastases. ID - patient ID, T - tumor size, N - necrosis zone size, $T_R$ - residual tumor w.r.t. the pre-RFA size, $N_{NC}$ - necrotization of non-cancerous tissue w.r.t. the total ablation zone, $d_R$ - minimal distance between necrosis zone and a risk structure, $d_C$ - weighted-centroid distance between metastasis and necrosis zone*

| ID | T [$cm^3$] | N [$cm^3$] | $T_R$ [%] | $N_{NC}$ [%] | $d_R$ [mm] | $d_C$ [mm] | ID | T [$cm^3$] | N [$cm^3$] | $T_R$ [%] | $N_{NC}$ [%] | $d_R$ [mm] | $d_C$ [mm] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 22.51 | 14.91 | 45.33 | 17.46 | 5.07 | 1.52 | 16 | 11.55 | 3.83 | 70.50 | 11.14 | 6.15 | 1.99 |
| 2 | 10.19 | 9.85 | 68.37 | 67.26 | 3.99 | 4.11 | 17 | 16.24 | 13.41 | 33.21 | 19.13 | 4.88 | 2.12 |
| 3 | 30.92 | 16.77 | 50.95 | 9.57 | 1.06 | 2.05 | 18 | 30.78 | 19.23 | 41.70 | 6.69 | 4.05 | 0.02 |
| 4 | 32.84 | 22.40 | 35.41 | 5.32 | 0.25 | 5.96 | 19 | 2.90 | 14.02 | 26.41 | 84.76 | 5.26 | 3.95 |
| 5 | 17.61 | 13.33 | 27.38 | 4.05 | 3.16 | 3.57 | 20 | 4.59 | 2.91 | 42.06 | 8.75 | 0.64 | 2.82 |
| 6 | 42.30 | 18.03 | 79.39 | 51.65 | 3.77 | 0.17 | 21 | 52.88 | 11.92 | 78.57 | 4.91 | 8.91 | 3.37 |
| 7 | 2.85 | 9.26 | 13.66 | 73.47 | 3.78 | 3.36 | 22 | 13.41 | 18.84 | 1.87 | 30.19 | 2.27 | 6.64 |
| 8 | 27.34 | 29.58 | 18.85 | 25.00 | 3.83 | 1.86 | 23 | 20.62 | 11.41 | 47.01 | 4.27 | 3.25 | 0.60 |
| 9 | 46.19 | 16.29 | 70.10 | 15.24 | 5.58 | 1.16 | 24 | 38.61 | 28.40 | 29.09 | 3.59 | 1.77 | 2.35 |
| 10 | 25.88 | 42.11 | 35.96 | 60.65 | 3.32 | 0.82 | 25 | 30.37 | 23.28 | 27.63 | 5.59 | 0.90 | 1.85 |
| 11 | 38.90 | 17.29 | 56.29 | 1.67 | 3.14 | 0.43 | 26 | 28.72 | 17.69 | 40.44 | 3.30 | 4.65 | 1.65 |
| 12 | 34.65 | 19.66 | 45.85 | 4.58 | 3.72 | 1.70 | 27 | 13.55 | 12.72 | 13.50 | 7.82 | 4.16 | 9.56 |
| 13 | 18.48 | 9.01 | 53.76 | 5.21 | 2.17 | 13.46 | 28 | 11.46 | 5.16 | 73.31 | 40.78 | 1.86 | 4.68 |
| 14 | 31.73 | 30.59 | 56.78 | 55.17 | 3.81 | 1.17 | 29 | 30.41 | 26.95 | 20.34 | 10.10 | 5.22 | 1.67 |
| 15 | 14.26 | 9.41 | 69.04 | 53.08 | 6.15 | 3.75 | 30 | 42.55 | 23.57 | 48.68 | 7.36 | 1.63 | 3.20 |

tool, combining automatically generated quantitative measures with intuitive visualizations, i.e., either slice-wise image viewer with superimposed segmentation masks or corresponding 3D volume renderings. This supports the physician with a straightforward and intuitive representation of the spatial relationships of the individual structures, which would be considerably more challenging if both images were viewed separately as it is currently practise. Futhermore, measures derived from the overlap of pre-interventional metastasis and post-RFA necrosis zone quantify the treatment outcome and enable a comprehensible and more objective validation. However, it must also be mentioned that the segmentations of the neuroradiologists represent a certain bias affecting the following treatment validation. Since the patients in our study underwent RFAs of spinal metastases for palliative treatment, there were no tumors classifiable as fully-treated and the mean percentage of residual tumor tissue was between 1.9 and 79.4 %. This is primarily due to the fact that the spatial proximity of a metastasis to surrounding risk structures, such as the spinal cord or nerve roots, complicates the intervention process and reduces the possibility of a complete tumor coverage due to required safety margins towards above mentioned structures (e.g. temperatures of around or above 45 °C have been stated to be cytotoxic to the spinal cord [9]). Our evaluation showed, that there was no intersection between necrosis zone and spinal canal in any patient case, meaning that the sensitive spinal cord tissue in particular was protected from necrotizing effects, which is a necessary requirement for these interventions. Furthermore, tissues like the vertebral rim with its periost are more heat-resistant than e.g. bone marrow, leading to unpredictable temperature distributions to a certain extent. Similar effects in RFAs of hepatic tumors arising from the liver blood flow and the associated vascular cooling effect [10]. The center of mass distance between the metastasis and necrosis is a valuable measure for validating and optimizing applicator placement and potential distribution of multipolar RFAs. Therefore, the measures defined in this paper may play an important role in therapy planning and simulation of RFAs in addition to the outcome validation. The retrospective comparison of the RFA simulation with the actual treatment outcome for instance, enables a successive adaptation and improvement of future simulations for more precise and reliable outcome predictions.

While almost all of the previous studies focused on hepatic tumor coverage and CT images, we found only one study dealing with RFAs of spine metastases [8]. In contrast to those studies, we have also included surrounding risk structures in our validation tool, since in RFAs of vertebral metastases the spinal canal in particular must be protected from necrotizing effects and is therefore an important indicator for treatment outcome. Regarding RFAs of spinal metastases, Dresky et al. [8] stated a tumor coverage of 91.8 % and 85.3 %, as well as 46.4 % and 24.1 % necrotization of non-metastasized tissue w.r.t. the total ablation zone. In comparison with our results, on average a substantially larger fraction of the metastases was ablated, however the proportion of necrotisized non-tumorous tissue is similar. However, due to the comparatively small dataset (n = 2) and the lack of any information about the registration approach and its accuracy, significance of the results and the conclusions is limited.

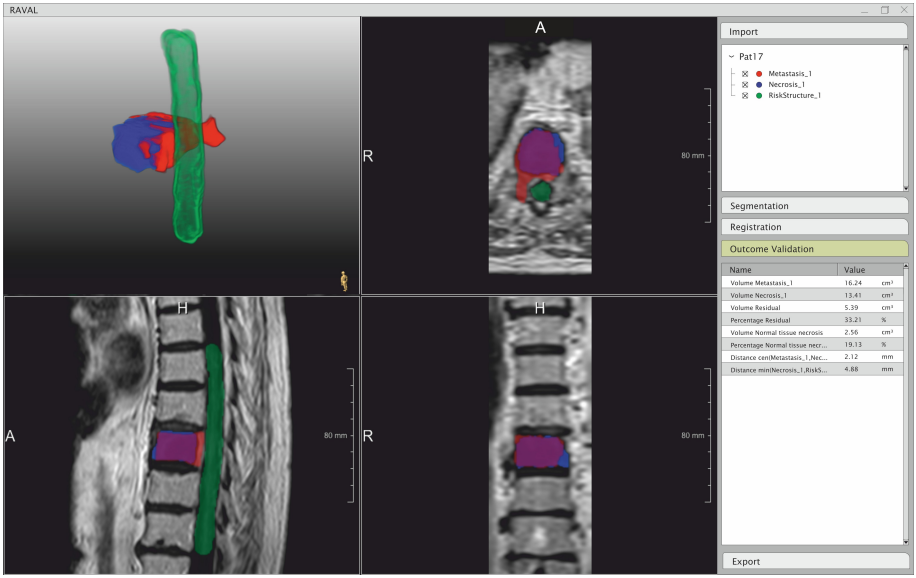



Figure 3: Graphical user interface of the presented tool with both, 2D and 3D visualizations, as well as quantitative measures for a patient individual outcome validation (exemplarily performed for patient 17).

## 5 Conclusion

We presented a tool to provide the physician with qualitative and quantitative information to support treatment outcome validation of RFAs of spinal metastases. For this purpose, we combined a fast and accurate image registration of pre- and post-interventional MR scans with a series of quantitative measures and intuitive visualizations for an objective and reliable assessment of the interventional results.
Our approach offers different links to future work, e.g. the implementation into a RFA therapy planning and simulation framework to include ablation zone predictions in the validation process. Moreover, automatic segmentations of the target structures as well as of surrounding risk structures like the spinal cord could speed up the validation process and save valuable time in clinical practice.

## References

[1] D. A. Wong, V. L. Fornasier, and I. MacNab, "Spinal metastases: the obvious, the occult, and the impostors.," *Spine*, vol. 15, pp. 1–4, 1990.

[2] Y. Makino, Y. Imai, T. Igura, S. Kogita, Y. Sawai, K. Fukuda, T. Iwamoto, J. Okabe, M. Takamura, N. Fujita, *et al.*, "Feasibility of extracted-overlay fusion imaging for intraoperative treatment evaluation of radiofrequency ablation for hepatocellular carcinoma," *Liver Cancer*, vol. 5, no. 4, pp. 269–279, 2016.

[3] E. S. McCreedy, R. Cheng, P. F. Hemler, A. Viswanathan, B. J. Wood, and M. J. McAuliffe, "Radio frequency ablation registration, segmentation, and fusion tool," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 3, pp. 490–496, 2006.

[4] K. Passera, S. Selvaggi, D. Scaramuzza, F. Garbagnati, D. Vergnaghi, and L. Mainardi, "Radiofrequency ablation of liver tumors: quantitative assessment of tumor coverage through ct image processing," *BMC Medical Imaging*, vol. 13, p. 3, 2013.

[5] C. O. Laura, K. Drechsler, S. Wesarg, and R. Bale, "Accurate physics-based registration for the outcome validation of minimal invasive interventions and open liver surgeries," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 2, pp. 362–371, 2016.

[6] C. Rieder, S. Wirtz, J. Strehlow, S. Zidowitz, P. Bruners, P. Isfort, A. H. Mahnken, and H.-O. Peitgen, "Automatic alignment of pre-and post-interventional liver ct images for assessment of radiofrequency ablation," in *Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 8316, p. 83163E, International Society for Optics and Photonics, 2012.

[7] C. Fujioka, J. Horiguchi, M. Ishifuro, H. Kakizawa, M. Kiguchi, N. Matsuura, M. Hieda, T. Tachikake, F. Alam, T. Furukawa, *et al.*, "A feasibility study: evaluation of radiofrequency ablation therapy to hepatocellular carcinoma using image registration of preoperative and postoperative ct," *Academic Radiology*, vol. 13, no. 8, pp. 986–994, 2006.

[8] C. von Dresky, H. Ballhausen, M. Becker, S. Serowy, M. Skalej, and C. Rieder, "Optimization of generator modulation and electrode positions for the radiofrequency ablation of spine metastases," in *Proceedings of the 14. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (CURAC)*, pp. 175–180, 2018.

[9] T. Yamane, A. Tateishi, S. Cho, S. Manabe, M. Yamanashi, A. Dezawa, H. Yasukouchi, and K. Ishioka, "The effects of hyperthermia on the spinal cord.," *Spine*, vol. 17, no. 11, pp. 1386–1391, 1992.

[10] T. Kröger, T. Pätz, I. Altrogge, A. Schenk, K. Lehmann, B. Frericks, J.-P. Ritz, H.-O. Peitgen, and T. Preusser, "Fast estimation of the vascular cooling in RFA based on numerical simulation," *The Open Biomedical Engineering Journal*, vol. 4, p. 16, 2010.

# Digitale Mammographie

Qualitätsanforderungen an ein tomosynthesefähiges Mammographie-Gerät

*Elisabeth Suzanne Reiter, Christian Kücherer*

[1]*Hochschule Reutlingen, Fakultät Informatik, Reutlingen, Germany*

Kontakt: Elisabeth_Suzanne.Reiter@Student.Reutlingen-University.DE, christian.kuecherer@reutlingen-university.de

***Abstract***

*Mammographie-Geräte werden in der Diagnostik von Mammakarzinomen eingesetzt. Die ursprüngliche Technik wurde in den letzten Jahren von analogen Röntgenfilmen zu digital integrierten Systemen weiterentwickelt. Durch die Tomosynthese, bei der in einem Schnittbildverfahren mehrere Schichten des Organismus untersucht werden können, können auch überlagerte Strukturen sichtbar gemacht werden. Um als adäquate Grundlage zur Diagnostik von malignen Tumoren dienen zu können, müssen einige qualitative Anforderung erfüllt werden. Bisher gibt es wenig Literatur, die Anforderungen und den Aufbau solcher Geräte systematisch beschreiben. Im Rahmen dieser Arbeit werden diese auf Basis der Literatur und bestehender Systeme die qualitativen Anforderungen identifiziert. Der prinzipielle Aufbau solcher Systeme wird anhand der einzelnen Systembausteinen in der semiformalen Notationssprache SysML gezeigt. Die grundlegende Funktionsweise eines tomosynthesefähigen Mammographie-Gerätes wird in dieser Arbeit zusammenfassend und anhand der einzelnen Systembausteine beschrieben. Diese Arbeit dient der Vermittlung eines umfassenden Verständnisses für die digitale Mammographie, um als Grundlage für die Dokumentation von qualitativen Anforderungen dienen zu können.*

**Keywords**: **Digitale Mammographie, Tomosynthese, Röntgen**, **Anforderungen, SysML, SYSMOD, Modellierung, Qualitätsanforderungen.**

## 1 Problemstellung

Für die 2D Mammographie existieren etablierte Prüfprotokollen und Standards, sowohl auf nationaler als auch internationaler Ebene. Dazu zählt die DIN 6868, die DIN EN 61223 sowie das European Protocol for the Quality Control of the Physical and Technical Aspects of Mammography Screening (EPQC) [3].Für die tomosynthesefähige Mammographie-Geräte werden oft nur Empfehlungen ausgesprochen, wodurch die konkrete Formulierung von Anforderungen erschwert wird [7]. Die Hersteller von Mammographie-Geräten wie GE oder Hologic beziehen sich auf unterschiedliche Patente und nutzen verschiedene Materialen und Techniken [5]. Um eine best-practise Variante des Mammographie-Gerätes zu beschreiben, fehlt jedoch eine übersichtliche Zusammenstellung der funktionalen und nicht-funktionalen Anforderungen. Um Anforderungen vollständig und umfassend formulieren zu können, ist es notwendig das Einsatzgebiet und den Untersuchungsablauf näher zu betrachten. Weiter muss ein umfassendes Verständnis des technischen Aufbaus und dem gewünschten Ergebnis bestehen. Dieses Verständnis kann durch ein Systemmodell vermittelt werden. Für die Mammographie existiert derzeit noch kein umfassendes Systemmodell. Aus diesem Grund wird eine systemische Modellierung durchgeführt um die bestehenden Anforderungen der 2D Mammographie und die aus Empfehlungen abgeleiteten Anforderungen der Tomosynthese einzelnen Systembausteinen und Anwendungsfällen zuzuordnen. Um die Vollständigkeit und die Konsistent der Anforderungen zu gewährleisten wird der SYSMOD Prozess nach Weilkins genutzt [13].

## 2 Material und Methoden

Die digitale Vollfeldmammographie, engl. full-field digital mammography (FFDM) ist derzeit der Standard im Screening-Programm für Brustkrebs. Durch die technischen Entwicklungen in den letzten Jahren konnte der Röntgenfilm durch digitale Systeme ersetzt werden. Dabei kommen Systeme zum Einsatz, die entweder mit Speicherleuchtstofffolien oder integrierten Detektoren als Empfänger der Röntgenstrahlung arbeiten. Durch die Digitalisierung der Signalübertragung konnte zwar die Bildqualität gerade bei dichtem Brustgewebe verbessert werden, dennoch können übereinanderliegende Strukturen nicht konkret dargestellt werden [2, 7]. Dieses

Problem wird durch die digitale Tomosynthese, engl. digital breast tomosynthesis (DBT) teilweise gelöst. Sie ist seit 2008 in Deutschland zugelassen und wird regelhaft im Screening Programm eingesetzt [14].

Durch die oben beschriebenen technischen Entwicklungen werden die Geräte komplexer. Um komplexe und verteile Systeme zu entwickeln existieren Werkzeuge die das System beschreiben. Die grafische Modellierungssprache OMG Systems Modeling Language (OMG SysML) ist eine Erweiterung der Unified Modeling Language (UML) und dient dazu den Aufbau und das Verhalten eines Systems darzustellen. Der Modellierungsprozess SYSMOD ist ein methodisches Vorgehen welches zunächst Systemziel, Systemidee und Basisarchitektur definiert um draus Anforderungen ableiten zu können [13].

# 3 Ergebnisse

## 3.1 Basisarchitektur

In diesem Abschnitt werden die Hauptbestandteile eines Mammographie-Gerätes mit integrierter Tomosynthese auf Basis des Patentes EP1816965B1 der Firma Hologic Inc näher beschrieben [5].

Da nur die Brust geröntgt wird, hängt der Röntgenstrahler an einem höhenverstellbaren Röntgenarm, welcher über einen Drehachse in zwei verschiedene Positionen verstellt werden kann. Der Röntgenarm wird an einem Standfuß befestigt. Während der Untersuchung wird die Brust zwischen zwei Kompressionsplatten fixiert. Da die Patientin bei der Untersuchung möglichst nahe am Gerät steht muss, um eine vollständige Abbildung der Brust zu erhalten, haben viele Geräte noch ein Gesichtsschild. Am Mammographie-Gerät befindet sich eine digitale Anzeige der Kompressionskraft. Die Röntgenstrahlen treffen auf eine integrierte Detektor Platte, wo diese in elektrische Signale umgewandelt werden. Die elektronischen Signale werden dann an einen Rechner übermittelt, wo die Bildrekonstruktion erfolgt [14].Um das Gerät zu bedienen gibt es zum einen ein Bedienfeld welches entweder seitlich am Gerät angebracht ist oder sich an einem externen Bedienpult befindet. Die Einstellungen des Bedienpults werden von der Systemsteuerung umgesetzt, wie die Drehung des Röntgenarms oder die Verschiebung der Kompressionsplatten Um die Kompressionsplatten zu steuern, werden Fußpedale eingesetzt, damit der/die medizinisch technische Röntgenassistent (MTRA) während der Einstellung neben der Patientin stehen kann [1].

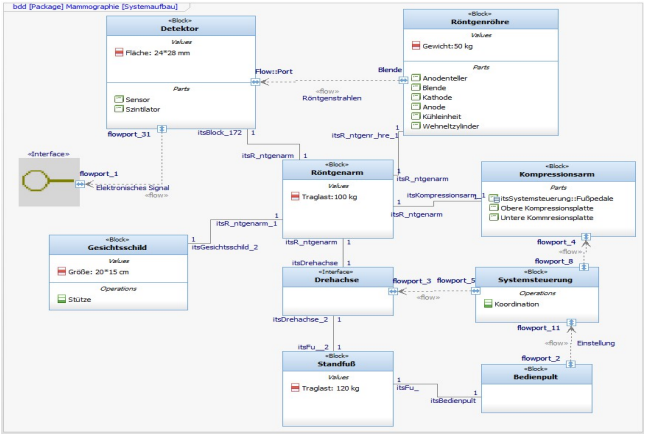



Abbildung 1: *Der Systemaufbau eines digitalen Mammographiegeräts als Blockdefinitionsdiagramm*

## 3.2 Röntgenröhre

In einem geschlossenen und unter Vakuum stehenden Glaskolben befinden sich die beiden Hauptbestandteile der Röntgenröhre: die Kathode und die Anode. Die Kathode wird mithilfe einer Hochspannung auf 2600°C erhitzt wobei Elektroden freigesetzt werden. Die Elektronen werden mit Hilfe eines elektronischen Feldes, das zwischen Kathode und Anode besteht, beschleunigt. Gegenüber der Kathode befindet sich ein Wehneltzylinder, um die beschleunigten Elektronen zu bündeln. Die Elektronen treffen auf die Anode, die aus einem Plättchen aus

Wolfram besteht, an der sie abgebremst werden. Hierbei wird die dabei entstehende Energie in Wärme und in Röntgenstrahlung umgewandelt. Die Wärme wird über ein integriertes Kühlsystem aus Öl wieder reduziert. Das Vakuum in dem Glaskolben verhindert, dass die Elektroden zwischen der Kathode und der Anode mit anderen Teilchen zusammenstoßen und dabei an Energie verlieren. Die Stelle, an der die beschleunigten Elektronen auf die Anode auftreffen, wird als Brennfleck bezeichnet [5, 8]. Für eine Mammographie-Aufnahme benötigt man eine Röntgenröhre mit sehr kleinem Brennfleck, da die Bildschärfe sehr stark von der Größe des Brennfleckes beeinflusst wird [4, 8]. Da die Röntgenstrahlung von der Anode aus in alle Richtungen verteilt wird, muss das Gehäuse um die Röntgenröhre aus einem Strahlenschützenden Material bestehen. Nur über das Austrittsfenster sollen die Röntgenstrahlen gezielt auf das Untersuchungsobjekt gelenkt werden. Um Streustrahlungen zu verhindern, befindet sich an dem Austrittsfenster ein System aus mehreren Blenden, an denen auch die Strahlendosis gemessen wird [8].

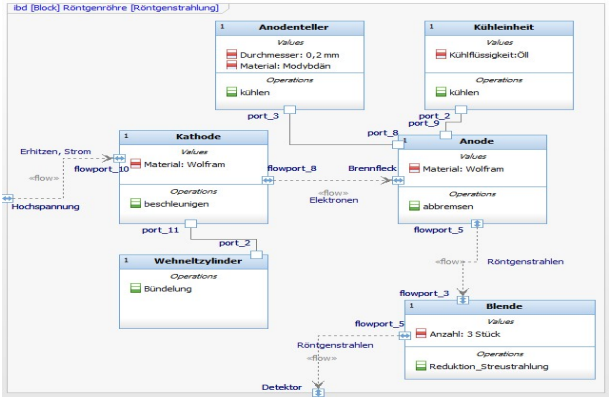



Abbildung 2*: Internes Blockdiagramm Röntgenröhre*

## 3.3 Anforderungen

Für die Patientin ist die Durchführung einer Mammographie oftmals belastend. Dies kann auf mehrere Ursachen zurückgeführt werden. Die Strahlenbelastung während der Untersuchung wird von vielen Patienten als potenziell gefährlich bzw. schädlich eingestuft. Des Weiteren ist die Mammographie zwar ein nicht-invasives Verfahren, dennoch ist die Kompression der Brust schmerzhaft. Das Warten auf das Untersuchungsergebnis ist für viele Patientinnen eine psychische Belastung [6]. Durch die neuen Technologien der DBT kann die Untersuchung angenehmer und kürzer gestaltet werden. Im Folgenden werden einzelne Aspekte für eine angenehme und sichere Untersuchung näher beschrieben. Die konkreten Zielwerte der Anforderungen leiten sich aus den Daten des Herstellervergleiches von et al. Hondel-Adametz ab. Auf dieser Basis werden Referenzwerte für die Eigenschaften des Detektors, der Röntgenstrahlung und der Anzahl der Aufnahmen definiert [4].

## 3.4 Strahlenbelastung

Bei der DBT wird die Patientin ionisierender Strahlung ausgesetzt, welche zu einer maligen Entartung der Zellen führen könnte. Im Vergleich zu einer Film-Folien-Mammographie konnte die Gesamtdosis bei der digitalen Mammographie reduziert werden, da keine Wiederholungsaufnahmen wegen Fehlbelichtungen oder Vergrößerungsaufnahmen angefertigt werden müssen [14]. Da bei der DBT jedoch mehrere Einzelaufnahmen benötigt werden wird mit Niedrigdosisaufnahmen gearbeitet. Der konkrete Grenzwert der Parenchymdosis ist abhängig von der Brustdicke. Mit steigender Objektdicke steigt auch die Dosis. Aus diesem Grund ist eine Kompression der Brust erforderlich [4]. Der mittlere Grenzwert der Parenchymdosis bei einer Brustdicke von 45 mm für eine einzelne Aufnahme liegt bei 2 mGy. Das entspricht ca. der doppelten Dosis einer normalen Mammographie ohne Tomosynthese, liegt aber immer noch weit unter dem Grenzwert des Bundesamts für Strahlenschutz für die Oberflächendosis von 10 mGy [10].

Ziel soll es sein, trotz geringer Strahlendosis eine gute Bildqualität zu erhalten. Die Strahlendosis ist abhängig von der Brustdicke, der Anzahl der Aufnahmen, des Scanwinkels, dem Vorhandensein eines Streustrahlenrasters, sowie dem Detektortyp und der Anode des Röntgenstrahlers [4]. Noch konnte keine optimale Parenchymdosis für die DBT festgelegt werden [11].

## 3.5 Kompression der Brust

Während des Scanvorgangs wird die Brust komprimiert um die Dicke der Brust zu verringern und das Gewebe aufzuspreizen. Dadurch soll erreicht werden, dass übereinanderliegende Strukturen getrennt werden. Die Belichtungszeit und die damit verbundene Bewegungsunschärfe, kann durch die Kompression der Brust reduziert werden. Außerdem wird dadurch die Brust während des Scanvorgans fixiert und es kommt zu keinen Bewegungsartefakten [3, 4].

Die Kompression der Brust wird jedoch von vielen Patienten als unangenehm bzw. schmerzhaft empfunden. Um die Belastung für die Patientin zu minimieren, ist es notwendig die Form und Flexibilität der Kompressionsplatten, engl. *Paddel*, anzupassen. Die Oberfläche der Paddel sollte der Form der Brust nachempfunden sein und sich beweglich an die Brust anpassen können. Des Weiteren soll die Oberfläche aus einem angenehm warmen Material bestehen [4].

Die MTRA stellt zu Beginn der Untersuchung den Druck der Kompressionsplatten manuell ein. Durch eine Fernbedienung soll die Patientin den Druck je nach Schmerzempfinden selbst nachregulieren können. Dabei hat eine Studie erwiesen, dass die meisten Frauen durch diese eigenständige Regulation eine noch größer Kompression tolerieren [1].

Sobald der Röntgenvorgang abgeschlossen ist, soll die Kompression der Brust automatisch gelöst werden, um die Patientin zu entlasten [4].

Um die Dauer der Brustkompression zu verringern, muss die Scanzeit reduziert werden. Die Dauer des Scanprozesses ist wiederum abhängig von der Größe des Scanwinkels, der Anzahl der Aufnahmen, der Röhrenbewegung und der Röhrengeschwindigkeit [11].

## 3.6 Untersuchungszeit

Die Dauer Vorbereitung sowie die die Scan-Zeit sollte so kurz wie möglich gehalten werden. Laut einer Studie von der Kassenärztlichen Bundesvereinigung ist eine unvollständige Darstellung der Brust die Hauptursache für mangelhafte Aufnahmen. Aus diesem Grund muss die MTRA die Patientin optimal am Gerät positionieren können. Da bei der Untersuchung meistens zwei Richtungen gescannt werden, erfolgt eine Verschiebung der Paddel von kraniokadal nach mediolateral oblique [3]. Die Einstellung der Scan Position und der Höhe der Lagerungsplatte soll manuell erfolgen. Die vorgeformten Paddel sollen über ein Klicksystem angebracht werden können. Es konnte festgestellt werden, dass die reine Scandauer bei einer kontinuierlichen Röhrenbewegung kürzer ist, als bei einer Step and Shot Variante, wo die Röhre sich während der Aufnahme nicht bewegt. Durch einen kleinen Scan-Winkel und einer schnelle Röhrengeschwindigkeit kann die Scandauer ebenfalls reduziert werden [4].

## 3.7 Röntgenstrahlung

Um die Gewebestruktur der Brust abzubilden wird bei einer herkömmlichen 2-D Mammographie eine weiche Röntgenstrahlung verwendet mit einer Röntgenröhrenspannung von 25 bis 36 kV [9]. Aufgrund der schrägen Einstrahlung wird bei der DBT eine etwas härtere Strahlung benötigt [4]. Dies kann durch eine höhere Röntgenröhrenspannung als auch durch eine höhere Filterung erreicht werden [9].

Da sich bei der DBT der Röntgenstrahl um die Brust bewegt, ist es notwendig, vor der Aufnahme eine Belichtungsmessung durchzuführen. Dabei soll es möglich sein, die Strahlendosis abhängig vom Aufnahmewinkel anzupassen, um sicherzustellen, dass auf der Detektorfläche eine gleichmäßige Quantenkonzentration besteht. Derzeit ist ein derartig vollständiges dynamisches Vorgehen nicht möglich [12]. Allerdings werden Es werden Vorschussaufnahmen angefertigt, entweder in der ersten Winkel Aufnahme oder bei 0°, um anschließend die Strahlendosis kalibrieren zu können [4].

## 3.8 Anzahl der Aufnahmen

Auch die Anzahl der Aufnahmen sind je nach Hersteller unterschiedlich. Die Bildqualität wird durch eine große Anzahl an Bilder verbessert. Dies bedeutet aber auch eine höhere Strahlenbelastung für die Patientin. Die anschließende Rekonstruktion der Bilder zu einem Schnittbild benötigt bei größerer Anzahl an Aufnahmen mehr Zeit. Gleiches gilt für den benötigten Speicherplatz, je nach Bildauflösung und Rekontruktionsmethode ist die Datei mehrere Gigabyte groß. Für die Speicherung und Archivierung dieser Bilder wird daher ein Picture Archiving and Communication System (PACS) benötigt [4, 15].

Auch die Verteilung der Aufnahmen während des Scanvorganges kann variieren. Werden die meisten Bilder in der senkrechten Position zur Brust angefertigt, so kann ggf. auf eine 2D Aufnahme der Brust verzichtet werden. Je kleiner der Abstand zwischen den einzelnen Aufnahmen ist, desto kleiner ist auch der Parallaxenfehler [12]. Ein Parallaxenfehler ist ein Beobachtungsfehler, der entsteh, wenn hintereinander liegende Gegenstände aus zwei Blickrichtungen betrachtet werden [4].

### 3.9 Detektorplatte

Für eine gute Bildqualität spielt auch die Beschaffenheit des Detektors eine Rolle. Dabei haben der Quantenwirkungsgrad, das Signal-Rausch-Verhältnis sowie die Modulationsübertragungsfunktion des Detektors Einfluss auf die Bildqualität.

Bei dichtem Gewebe der Brust ist eine gute Kontrastdarstellung relevant für die Bildqualität. Durch die Modulationsübertragungsfunktion (MFT) wird die Stärke der Kontrastisierung eines Objektes beschrieben. Je kleiner die Objekte sind, wie z.B. Microverkalkungen, desto höher muss die Kontrastwiedergabe, auch Modulation genannt, sein [12]. Ziel ist eine möglichst hohe Modulationsübertragungsfunktion [4].

Neben der MFT spielt auch das Signale-Rausch-Verhältnis eine Rolle bei der Erkennung von kleinen Objekten. Man unterscheidet zwischen Quantenrauschauen und Systemrauschen. Das Quantenrauschen kann vor allem dann entstehen, wenn lediglich eine geringe Strahlung vom Detektor erfasst wird. Das Systemrauschen ergibt sich aus dem Strukturrauschen des Detektors und dem Elektronikrauschen bei der Bildverarbeitung. Ziel ist ein möglichst hohes Signal-Rausch-Verhältnis [12].

Bei der DBT hat jede Aufnahme nur eine sehr geringe Strahlendosis. Der Detektor muss also auch bei einer niedrigen Strahlendosis ein rauscharmes Signal zurückgeben.Mit der detective quantum efficiency (DQE) wird der Quantenwirkungsgrad gemessen. Bei den aktuell verfügbaren Flächendetektoren liegt die DQE bei ca. 50-70%. Die DQE lässt sich aus der Modulationsübertragungsfunktion und dem Signal-Rausch-Verhältnis berechnen. Je höher die DQE ist, desto weniger Strahlendosis wird benötigt um trotzdem ein qualitativ hochwertiges Bild zu erhalten [12].

## 4 Diskussion

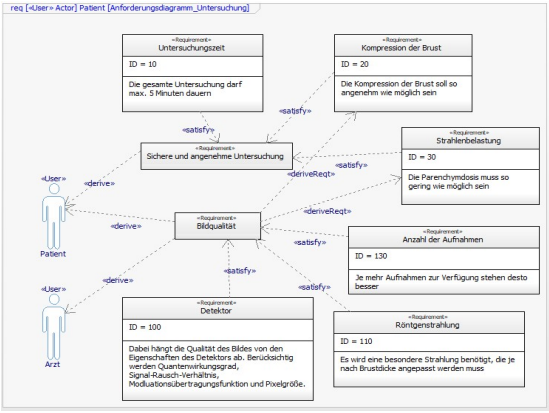



Abbildung 3*: Modellierung der Anforderungen und ihre Abhängigkeiten*

Die Anforderungen sind nicht widerspruchsfrei. Beispielsweise gibt es eine direkte Abhängigkeit zwischen Strahlenbelastung und Bildqualität. Je höher die Strahlendosis ist desto höher ist die DQE der Detektorfläche. Dadurch kann eine bessere Bildqualität erzielt werden. Jedoch kann die Strahlendosis nicht beliebig hoch gesetzt sein. Es müssen die Grenzwerte der Strahlenschutzverordnung eingehalten werden [10].

Für eine optimale 3D Rekonstruktion mit möglichst hoher Tiefenauflösung wird eine große Anzahl an Bilder benötigt. In Folge dessen muss daher die Einzeldosis reduziert werden, worunter die Qualität der einzelnen Aufnahmen leidet [10]. Bisher konnte noch durch keine Studie die optimalen technischen Parameter identifizieren, weshalb es bei den Hersteller gerade bei der Anzahl der Aufnahmen und dem damit verbunden

Scan-Winkel große Unterschiede gibt [7]. Für die 2D Mammographie existieren etablierte Prüfprotokollen und Standards, sowohl auf nationaler als auch internationaler Ebene. Für die DBT werden oft nur Empfehlungen ausgesprochen, was es schwer macht, konkrete Anforderungen zu definieren [7].

# 5 Zusammenfassung

Die zentrale Fragestellung der Hausarbeit ist die Identifikation der Anforderungen an die DBT. Bei der Modellierung der Basisarchitektur wurde das Mammographie-Gerät in seine einzelnen Systembausteine zerlegt. Dabei konnte die SysML-Notation die Beziehungen der einzelnen Bausteine zueinander verdeutlichen. Die Bausteine wurden als Blöcke definiert und konnten dann über ein Internes Blockdiagramm näher beschrieben werden. Dieses Vorgehen war bei den Systembestandteilen Detektor und Röntgenröhre vorteilhaft, um die nachfolgenden Anforderungen besser nachvollziehen und zuordnen zu können. Der SYSMOD Prozess eignet sich somit in jeder Hinsicht zur Identifikation von Anforderungen, da das System als Ganzes betrachtet wird und somit die Gefahr reduziert werden kann, Anforderungen zu übersehen.

# 6 References


[1] Christina Adelheid Mehling. 2006. *Schmerz- und Angsterleben bei Mammographie mit Fremd- und Eigenkompression.* Universität Würzburg.

[2] Julian Hägele, Telja Pursche, J. Barkhausen, and F. Schäfer. 2015. Digitale Tomosynthese der Brust. *Radiologie up2date* 15, 03, 239–258. DOI: https://doi.org/10.1055/s-0034-1393054.

[3] Karin Hellerhoff. 2010. Digitale Brusttomosynthese : Technische Grundlagen, aktuelle klinische Relevanz und Perspektiven für die Zukunft. *Der Radiologe* 50, 11, 991–998. DOI: https://doi.org/10.1007/s00117-010-2008-7.

[4] Michaela Hondl-Adametz. 2018. Digitale Brusttomosynthese (DBT) – ein großer Schritt in der Mammografie. *Radiopraxis* 11, 02, 83–96. DOI: https://doi.org/10.1055/s-0043-123733.

[5] Ken Defreitas,Tom Farbizio,Baori Ren,Chris RuthIan, Shaw Andrew Smith, Jay Stein. Integriertes multimodus-mammographie-/tomosynthese-röntgensystem. Patent No. EP1816965B1.

[6] Heinz Otto. 2002. *Mammadiagnostik für MTRA und Ärzte*. Springer, Berlin, Heidelberg

[7] Alejandro Rodríguez-Ruiz, Maria Castillo, Julia Garayoa, and Mathieu Chevalier. 2016. Evaluation of the technical performance of three different commercial digital breast tomosynthesis systems in the clinical environment. DOI: https://doi.org/10.1016/j.ejmp.2016.05.001.

[8] Wolfgang Schlegel, Christian P. Karger, and Oliver Jäkel, Eds. 2018. *Medizinische Physik. Grundlagen - Bildgebung - Therapie - Technik* (1. Auflage 2018). Springer Berlin; Springer Spektrum, Berlin.

[9] Hans Schöfer, D. Kotsianos, S. Wirth, S. Britsch, and M. Reiser. 2005. Qualität digitaler Mammogramme. Entwicklung und Einsatz eines Phantoms zur optimierten Sicherung. *Der Radiologe* 45, 3, 278–285. DOI: https://doi.org/10.1007/s00117-005-1180-7.

[10] Rüdiger Schulz-Wendtland, Nicolas Becker, K. Bock, K. Anders, and W. Bautz. 2007. Mammographie-screening. *Der Radiologe* 47, 4, 359-69. DOI: https://doi.org/10.1007/s00117-007-1490-z.

[11] Rüdiger Schulz-Wendtland, K.-P. Hermann, and M. Uder. 2010. Digitale Tomosynthese der Brust. *Radiologie up2date* 10, 03, 195–205. DOI: https://doi.org/10.1055/s-0030-1255574.

[12] Friedrich Semturs, E. Sturm, R. Gruber, and Thomas H. Helbich. 2010. Physikalische Aspekte der verschiedenen Tomosynthesesysteme. *Der Radiologe* 50, 11, 982–990. DOI: https://doi.org/10.1007/S00117-010-2012-y.

[13] Tim Weilkiens. 2014. *Systems engineering mit SysML/UML. Anforderungen, analyse, architektur* (3., überarbeitete und aktualisierte Auflage (Online-ausg.)). Dpunkt.verlag, Heidelberg, Germany.

[14] Susanne. Wienbeck and U. Fischer. 2015. Digitale Mammografie und ihre Weiterentwicklungen. *Radiologie up2date* 15, 03, 215–234. DOI: https://doi.org/10.1055/s-0034-1393053.

[15] Alfred Winter. et al. 2011. *Health information systems. Architectures and strategies* (2nd ed.). Health informatics. Springer, London.

# Workflow & Prozesse

# Comparison of intraoperative angiography in a hybrid operating room (OR) setup with indocyanine green videoangiography in the treatment of cerebrovascular lesions in regard to efficacy, workflow, risk benefit and cost – a prospective analysis

G. Durner[1], S. Wahler[1], M. Braun[1], T. Kapapa[1], C. R. Wirtz[1], R. König[1], A. Pala[1]

[1]*Neurosurgery department, BKH Günzburg, Günzburg, Germany*

Kontakt: gregor.durner@bkh-guenzburg.de

## 1 Introduction

Due to its image quality, digital subtraction angiography (DSA) has become the gold standard for evaluating the result of neurovascular procedures. Intraoperative use (ioDSA) allows for instant success control and intraoperative interventions. However, the value and efficiency of these hybrid operating theaters has been controversially discussed.

In a prospective study, we examined the hybrid operating room at Ulm University equipped with an Artis Zeego robotic c-arm regarding cost, efficiency and workflow. We furthermore compared it with other methods like microdoppler and indocyanine green (ICG) angiography.

## 2 Methods

We examined all neurovascular patients between 2012 and 2015, treated in the hybrid operating theater in a risk-benefit analysis. After using microdoppler and ICG for best operative result, every patient received an additional ioDSA to look for remnants or unfavorable clip placement which might lead to a change of operating strategy or result.

Furthermore, a workflow-analysis was conducted on randomly selected cases. Here a detailed examination of the individual operating steps, staff positioning, costs, technical errors or complications was performed.

## 3 Results

54 patients were enrolled in the risk-benefit analysis, 22 in the workflow analysis. The average duration of a cerebrovascular operation was 4h 58min. 2min 35s accounted for ICG angiography, 46min 4s for ioDSA. Technical errors occurred in 40,9% of angiographies and mainly concerned screen reflections and steering difficulties. Adverse events occurred during one of 52 ioDSAs, resulting in an occlusion of the left distal medial cerebral artery (MCA).

In the risk-benefit analysis, ioDSA was able to detect a perfusion rest in 2 out of 44 cases (4,7%) of aneurysm surgery, which could not be displayed with ICG before. In arterio-venous-malformation (AVM) surgery, one of 11 examined patients (7,7%) showed a remnant in ioDSA and additional resection was necessary.

The average cost of ioDSA in Ulm University can be estimated with 1843,00€ (not including costs of materials used).

## 4 Conclusion

According to our results, ioDSA associated complications in a hybrid OR are low. Relevant findings in ioDSA can avoid additional intervention by displaying pathologies which are undetectable by other imaging modalities. ICG videoangiography is equally safe but offers better availability and lower costs.

# Intuitive Orchestrierung kontext-adaptiver chirurgischer Assistenzsysteme

*L. Bernhard[1], D.Ostler[1], D. Wilhelm[1,2], H. Feußner[1,2]*

*[1]Klinikum rechts der Isar, Forschungsgruppe für minimal-invasive interdisziplinäre therapeutische Intervention, München, Deutschland*
*[2] Klinikum rechts der Isar, Klinik und Poliklinik für Chirurgie, München, Deutschland*

Kontakt: lukas.bernhard@tum.de

***Abstract***

*Moderne Operationssäle bieten eine kontinuierlich steigende Zahl von medizinischen Assistenzsystemen mit immer umfassenderer Funktionalität. Um der zwangsläufig ebenfalls zunehmenden Komplexität bei der Bedienung zu begegnen, zielen kontext-sensitive Assistenzsysteme darauf ab, den chirurgischen Workflow automatisiert zu erkennen und das OP-Team (teil-)autonom mit der richtigen Aktion zur richtigen Zeit zu unterstützen. Die in dieser Arbeit vorgestellte Workflow-Assistenz-Architektur sieht eine grafische Benutzeroberfläche vor, die den Chirurgen bzw. die Chirurgin bei der Orchestrierung aller für einen bestimmten chirurgischen Eingriff relevanten Assistenzsysteme unterstützt. Die Implementierungen wurden im Rahmen einer Vorstudie gemeinsam mit ChirurgInnen evaluiert. Sowohl die Bewertungsergebnisse verschiedener Usability- und User-Experience-Aspekte als auch die allgemeine Resonanz der ProbandInnen weisen darauf hin, dass eine flexible Konfiguration des Assistenzverhaltens durch den Chirurgen bzw. die Chirurgin praktikabel und erstrebenswert ist.*

**Keywords**: Kognitiver OP-Saal, Kontext-adaptive Systeme, Workflow-Assistenz, Surgical Data Science

## 1 Problemstellung

In modernen Operationssälen kommt eine Vielzahl von Medizinprodukten und Assistenzsystemen zum Einsatz, die eine immer umfassendere Funktionalität und höhere intraoperative Flexibilität bieten [1, 2]. Jedoch steigt – bedingt durch die Fortschritte in der Leistungsfähigkeit – auch zwangsläufig die Komplexität bei der Bedienung und somit die Anzahl der potenziellen Fehlerquellen und Unwägbarkeiten [1, 3]. Die Folgen sind eine verringerte Patientensicherheit aufgrund von Fehlbedienungen [4, 5], eine nicht vollständige Ausschöpfung des diagnostischen und therapeutischen Potenzials [1] sowie eine suboptimale Auslastung des Operationssaals durch Verzögerungen beim Geräteeinsatz. Um diesen Problemen zu begegnen, zielen chirurgische Workflow-Erkennungssysteme darauf ab, die aktuelle Situation im OP während der Operation automatisch zu erfassen, indem sie Informationsflüsse aus der technischen Infrastruktur wie Videosignale, Sensorsignale und Gerätezustände analysieren. Dies ist eine unverzichtbare Voraussetzung für kontextabhängig agierende Assistenzsysteme, die das OP-Team mit der richtigen Aktion zur richtigen Zeit unterstützen. Während in der Vergangenheit eine Vielzahl von Ansätzen zur Workflow-Erkennung vorgeschlagen wurden (etwa [6, 7, 8]), konzentrieren sich nur wenige Arbeiten auf die kontextsensitive Orchestrierung von chirurgischen Assistenzsystemen und Medizinprodukten. Das in dieser Arbeit vorgestellte Workflow-Assistenzframework soll deshalb eine intuitive Konfiguration des Workflow-Assistenzverhaltens direkt durch den Chirurgen bzw. die Chirurgin selbst ermöglichen.

## 2 Material und Methoden

Die vorgeschlagene Workflow-Assistenzarchitektur besteht aus zwei Hauptkomponenten: das Erkennungsmodul und das Assistenzmodul. Während einer laufenden Operation liefert das Erkennungsmodul Kontextinformationen über die aktuelle intraoperative Situation. Das Assistenzmodul verarbeitet diese Informationen und adaptiert den Zustand der technischen Infrastruktur automatisiert an die aktuellen Bedürfnisse des OP-Teams. Das exakte Verhalten des Assistenzmoduls kann präoperativ durch die Definition von Regeln festgelegt werden. Jede Regel bildet einen bestimmten Kontext bzw. eine bestimmte Situation auf eine gewünschte unterstützende Aktion ab. Das Konzept sieht explizit vor, dass die Definition dieser Regeln vom Chirurgen bzw. der Chirurgin selbst vorgenommen wird, da er über das notwendige Fachwissen über den chirurgischen Eingriff verfügt. Darüber

hinaus wird es hierdurch ermöglicht, Präferenzen des einzelnen Chirurgen hinsichtlich Art und Umfang der Workflow-Unterstützung zu berücksichtigen und das Assistenzverhalten fallspezifisch zu rekonfigurieren.

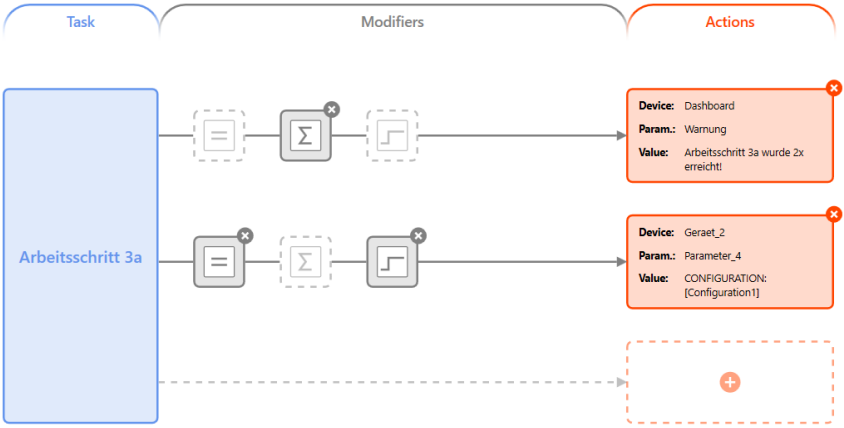



Abbildung 1*: Benutzeroberfläche zur Definition des intra-operativen Workflow-Assistenzverhaltens*

Um den Chirurgen bzw. die Chirurgin bei der Modellierung des gewünschten Workflow-Assistenzverhaltens zu unterstützen, wurde im Rahmen dieser Arbeit eine intuitive grafische Benutzeroberfläche entwickelt (siehe Abbildung 1), die die Definition von Workflow-Assistenzregeln ermöglicht, ohne dabei technisches Hintergrundwissen, wie etwa Programmier- oder Struktursprachen, vorauszusetzen. Die Benutzerfreundlichkeit der Software wurde im Rahmen einer Vorstudie mit 6 erfahrenen ChirurgInnen und 5 DoktorandInnen mit medizintechnischer Ausbildung evaluiert.

## 3 Ergebnisse

Mit Hilfe eines mehrdimensionalen Bewertungssystems wurden Aspekte aus dem Bereich *Usability* (Durchschaubarkeit, Effizienz, Steuerbarkeit) und *User-Experience* (Originalität, Stimulation, Attraktivität) bewertet. Nach einer kurzen Einführung in die Benutzeroberfläche wurden die TeilnehmerInnen dazu aufgefordert, ein in Prosaform vorgegebenes Workflow-Verhalten zu konfigurieren. Anschließend wurden die TeilnehmerInnen gebeten, die Software im Hinblick auf die oben genannten Aspekte in einem Intervall von -3 (am schlechtesten) bis 3 (am besten) zu bewerten. Die Ergebnisse sind in Tabelle 1 dargestellt.

Tabelle 1*: Ergebnisse der Nutzerstudie*

| | Mittelwert (-3 bis 3) | Konfidenz | Konfidenzintervall |
|---|---|---|---|
| Attraktivität | 1,758 | 0,337 | 1,421 – 2,094 |
| Durchschaubarkeit | 1,795 | 0,361 | 1,435 – 2,156 |
| Effizienz | 1,773 | 0,404 | 1,368 – 2,177 |
| Steuerbarkeit | 1,818 | 0,368 | 1,450 – 2,187 |
| Stimulation | 1,409 | 0,487 | 0,922 – 1,896 |
| Originalität | 1,455 | 0,556 | 0,899 – 2,011 |

## 4 Diskussion

Im Zuge der Studie konnte demonstriert werden, dass die ProbandInnen nach einer kurzen Einführung dazu in der Lage waren, die Benutzeroberfläche selbstständig und zügig zu bedienen. Die Bewertung des Nutzererlebnisses durch die ProbandInnen ergab Durchschnittswerte im mittleren positiven Bereich für die einzelnen Bewertungsaspekte. Insbesondere die Aspekte Durchschaubarkeit, Effizienz und Steuerbarkeit, die speziell für die Bewertung der intuitiven Bedienbarkeit relevant sind, erzielten besonders gute Ergebnisse. Im Gespräch mit den ProbandInnen stießen sowohl das Gesamtkonzept als auch die Implementierung der Benutzeroberfläche auf Zustimmung. Die Workflow-Assistenztechnologie im Allgemeinen sowie das im Rahmen dieser Arbeit entwickelte konfigurierbare System wurden als sinnvoll und für die Zukunft wünschenswert erachtet. Die Bedienung der Benutzeroberfläche wurde insgesamt als klar und effizient wahrgenommen, sobald das Grundkonzept verstanden ist. Die Ergebnisse und die Resonanz der ProbandInnen weisen darauf hin, dass eine flexible Definition des Verhaltens der Workflow-Unterstützung durch den Chirurgen bzw. die Chirurgin realistisch und wünschenswert ist. Es ist jedoch notwendig, umfangreichere Studie durchzuführen, um die Ergebnisse mit einer größeren Teilnehmerzahl zu bestätigen und neue Hinweise zur Optimierung der Regelsprache und ihrer visuellen Darstellung zu finden. Ein zentraler Aspekt ist es, die Beschreibungskraft der Regelsprache weiter zu erhöhen, um komplexere Orchestrierungssituationen abbilden zu können. Das resultierende Verhalten der Workflow-Assistenz muss hierbei jedoch weiterhin transparent bleibt.

## 5 Zusammenfassung

Der vorgestellte Ansatz erlaubt eine Orchestrierung von kontext-adaptiven chirurgischen Assistenzsystemen mithilfe einer intuitiv und effizient zu bedienenden Benutzeroberfläche. Das Konzept sieht explizit vor, dass die Modellierung des Assistenzverhaltens von den Chirurgen selbst vorgenommen wird, um individuelle Präferenzen abbilden zu können und eine fallspezifische Rekonfiguration des Systems zu erlauben. Im Rahmen einer Vorstudie wurden diverse Usability- und User-Experience-Aspekte des Systems evaluiert. Die Ergebnisse fielen durchweg positiv aus und weisen auf eine gute Akzeptanz des Konzeptes bzw. der Implementierung hin.

## Referenzen


[1] H. Kenngott, M. Wagner, A. Preukschas, B. Müller-Stich: *Der intelligente Operationssaal*: *Vom passiven Gerätepark zum mitdenkenden, kognitiven Assistenten*, Der Chirurg; Zeitschrift für alle Gebiete der operativen Medizin, 87 (Nr. 12), S. 1033–1038 (2016)

[2] M. Rockstroh, S. Franke, M. Hofer, A. Will, M. Kasparick, B. Andersen, T. Neumuth: *OR.NET: multi-perspective qualitative evaluation of an integrated operating room based on IEEE 11073 SDC*, International journal of computer assisted radiology and surgery, 12 (Nr. 8), S. 1461–1469 (2017)

[3] A. Guédon, L. Wauben, M. Overvelde, J. Blok, M. van der Elst, J. Dankelman, J. van den Dobbelsteen: *Safety status system for operating room devices*, Technology and health care: official journal of the European Society for Engineering and Medicine, 22 (Nr. 6), S. 795–803 (2014)

[4] Y. Hu, A. Arriaga, E. Roth, S. Peyre, K. Corso, R. Swanson, R. Osteen, P. Schmitt, A. Bader, M. Zinner, C. Greenberg: *Protecting patients from an unsafe system: the etiology and recovery of intraoperative deviations in care*, Annals of surgery, 256 (Nr. 2), S. 203–210 (2012)

[5] C. Christian, M. Gustafson, E. Roth, T. Sheridan, T. Gandhi, K. Dwyer, M. Zinner, M. Dierks, *A prospective study of patient safety in the operating room*, Surgery, 139 (Nr. 2), S. 159–173 (2006)

[6] R. Stauder, A. Okur, N. Navab, *Detecting and analyzing the surgical workflow to aid human and robotic scrub nurses*, The Hamlyn Symposium on Medical Robotics, 91, 91-92 (2014)

[7] S. Bodenstedt; D. Rivoir; A. Jenke; M. Wagner; M. Breucha; B. Müller-Stich, *Active Learning using Deep Bayesian Networks for Surgical Workflow Analysis*, arXiv preprint arXiv:1811.03382v2, (2018)

[8] C. Nwoye, D. Mutter, J. Marescaux, N. Padoy, *Weakly supervised convolutional LSTM approach for tool tracking in laparoscopic videos*, arXiv preprint arXiv:1812.01366, (2018)

# Entwicklung einer adaptiven Situationserkennung zur Informationserfassung und -anzeige im intraoperativen Bereich

*D. Junger[1], S. Frommer[1], C. Ryniak[1], O. Burgert[1]*

*[1]Hochschule Reutlingen, Fakultät Informatik, Forschungsgruppe Computer Assisted Medicine (CaMed), Reutlingen, Deutschland*

Kontakt: denise.junger@reutlingen-university.de

***Abstract***

*Zur Unterstützung des Operateurs wird eine patientennahe Informationsanzeige entwickelt, die kontextrelevante Informationen entsprechend der aktuellen Situation bereitstellen kann. Hierfür soll eine Situationserkennung konzipiert werden, die auf unterschiedliche intraoperative Prozesse übertragen werden kann. Ziel der adaptiven Situationserkennung ist das Erkennen spezifischer Situationen durch intraoperative Informationen unterschiedlicher Datenquellen im Operationssaal. Innerhalb der Datenerhebung und -analyse wurden Anwendungsfälle für die Situationserkennung definiert sowie chirurgische Prozessmodelle erstellt, die intraoperative Ereignisse abbilden. Auf Basis dieser Informationen wurde ein Konzept entworfen, das sich zunächst auf die Erkennung abstrakter generalisierter Phasen, unabhängig vom Eingriff, fokussiert und sich Schritt für Schritt auf granulare Prozessschritte spezifizieren lässt. Diese Flexibilität soll die Übertragbarkeit des Konzepts auf intraoperative Prozesse ermöglichen und den Operateur dadurch gezielt mit kontextrelevanten Informationen unterstützen. Das Konzept wird in zukünftigen Schritten weiterentwickelt.*

**Keywords**: Situationserkennung, Phasenerkennung, Situationsbewusstsein, Adaptivität

## 1 Problemstellung

Innerhalb des anwendungsorientierten Forschungsprojekts *OR-Pad - Nutzung von portablen Informationsanzeigen im Operationssaal* wird an der Hochschule Reutlingen ein High-Fidelity-Prototyp entwickelt, der die Unterstützung des Informationsflusses im perioperativen Bereich zum Ziel hat. Eine sterile Informationsanzeige soll dem Operateur klinisch relevante Informationen in unmittelbarer Nähe zum Situs bereitstellen und Informationen für die OP-Dokumentation erfassen. Durch die begrenzten Möglichkeiten des Operateurs intraoperativ mit dem System zu interagieren, sollen kontextrelevante Informationen (z.B. präoperative Bilder) basierend auf dem Operationsverlauf und dem aktuellen Prozessschritt automatisiert bereitgestellt werden. Die Erfassung intraoperativer Informationen soll dabei automatisiert werden, um den Interaktionsaufwand zu reduzieren.

Zur Realisierung der Erfassung und Anzeige kontextrelevanter Informationen, muss sich das System der aktuellen Situation im Operationssaal bewusst sein. Hierfür muss der chirurgische Workflow durch die Nutzung von Daten verfügbarer Signalquellen auf verschiedenen Granularitäten identifiziert werden [1]. Ziel ist es, ein Situationsbewusstsein durch eine Situationserkennung bereitzustellen, die die exakte OP-Phase mithilfe der gesammelten Informationen über Instrumente, Personen, anatomische Strukturen etc. erkennt. Zur Generierung von Informationen eignen sich beispielsweise Technologien wie Barcodes oder RFID [2]. Basierend auf der erkannten Situation wird es möglich sein, Informationen automatisiert und situationsbedingt anzuzeigen und so das OP-Team und im Speziellen den Operateur zu unterstützen.

Für spezifische lineare Eingriffe existieren bereits Lösungen, durch die intraoperative Phasen erkannt werden können, z.B. auf Basis von Videodaten [1, 3] oder mithilfe eines Fußschalters [4]. Allerdings lassen sich diese Ansätze nicht direkt auf andere komplexere Eingriffe übertragen. Die Systeme sind oftmals für einen bestimmten Eingriff ausgelegt, werden speziell für diesen trainiert und können schlecht mit Prozess- und Technologieänderungen umgehen. Eine einheitliche Lösung, die sich auf unterschiedliche Prozesse und Granularitäten des intraoperativen Bereichs adaptieren lässt und dabei robust die aktuelle Situation im Operationssaal erkennt, konnte bisher nicht umgesetzt werden.

Im Rahmen des *OR-Pad* Projekts soll ein Konzept für eine Situationserkennung entwickelt werden, die generalisiert auf einen Großteil aller Eingriffe angewandt werden kann, vorausgesetzt es stehen die benötigten Sensortechnologien zur Verfügung. Ziel der Situationserkennung ist hierbei das Erkennen spezifischer Situationen

im aktuellen Kontext, die durch Parameterkombinationen unterschiedlicher Datenquellen, wie Instrumente oder Personen, definiert werden. Je nach Abstraktion bzw. Granularität sollen dadurch unterschiedlich detaillierte Phasen oder Schritte erkannt werden, um dem Operateur und seinem Team kontextrelevante Informationen bereitzustellen. Das Konzept fokussiert sich hierbei zunächst auf die Erkennung von abstrakten Phasen und sieht eine Spezifizierung in den weiterführenden Entwicklungsphasen vor.

# 2 Material und Methoden

Für die Konzeption der Situationserkennung wurden Prozesse verschiedener Eingriffe in der Urologie und Orthopädie des Universitätsklinikum Tübingen in Hospitationen beobachtet. Die Prozessschritte des Operateurs, definiert durch die benutzten Instrumente und Materialien, wurden von Hautschnitt bis Naht händisch dokumentiert. Die gewonnenen Daten wurden anschließend in Bezug auf die Situation, den Prozessablauf und Ereignisse analysiert. Auf dieser Grundlage wurden Anwendungsszenarien einer Situationserkennung definiert und intraoperative Abläufe in Prozessmodellen modelliert.

## 2.1 Anwendungsszenarien

Mithilfe einer Analyse der durchgeführten Hospitationen konnten drei Szenarien identifiziert werden, die die Reduzierung des Interaktionsaufwands mit dem *OR-Pad* durch den Einsatz einer Situationserkennung ermöglichen. Im ersten Szenario „Anzeigen des zeitlichen Fortschritts der Operation“ werden, neben Informationen zur geplanten Dauer, die bisher benötigte Zeit sowie Abweichungen im Zeitverlauf durch die Situationserkennung erkannt und bereitgestellt. Zur Schätzung der verbleibenden Zeit des Eingriffs und einer möglichen Verzögerung nutzt das Situationsbewusstsein, neben dem Wissen zum aktuellen Prozessschritt, auch Vergleiche zu ähnlichen bereits durchgeführten Prozessen. Daraufhin wird sowohl der OP-Planer als auch das OP-Team informiert.

Durch die Situationserkennung werden im zweiten Szenario „Bereitstellen von Informationen“ aktuell relevante patienten- und eingriffsspezifische Informationen, wie beispielsweise ein präoperativ aufgenommenes Röntgenbild beim Freilegen einer Struktur oder die Planung des chirurgischen Eingriffs, angezeigt. Hierfür werden ebenfalls durch Kontextinformationen Situationen erkannt, die mit fallspezifischen Informationen assoziiert sind. Diese Informationen werden daraufhin für das OP-Team ohne Interaktionsaufwand angezeigt.

Ein zusätzlicher Mehrwert entsteht durch das dritte Szenario „Erstellung von OP-Dokumentations-Einträgen“. Das durch aktuelle Prozessschritte und Vergleiche mit ähnlichen Prozessen ermöglichte Situationsbewusstsein wird hierbei genutzt, um spezifische Situationen während dem Eingriff automatisiert zur OP-Dokumentation hinzuzufügen. Dabei liegt der Fokus der Situationserkennung auf der Erfassung von relevanten Zeitpunkten, wie beispielsweise dem Hautschnitt, dem Entfernen eines Organs oder dem Einsetzen eines Implantats.

## 2.2 Prozessmodelle

Basierend auf den erfassten Prozessschritten aus den Hospitationen wurden Prozessmodelle in der Modellierungssprache Business Process Model and Notation (BPMN) erstellt. Diese bilden die intraoperativen Abläufe der verschiedenen Eingriffe ab. Subprozesse der feingranularen Modelle wurden im nächsten Schritt zu Phasen zusammengefasst. Identifizierte Ereignisse, die spezifische Situationen während dem Eingriff widerspiegeln (z.B. Mikroskop-Nutzung nur in einer bestimmten Phase, häufige Positionsänderungen etc.), stellen in den Modellen Möglichkeiten zur Erkennung von Phasen dar. Abbildung *1* zeigt einen Ausschnitt einer Klauenzehenkorrektur, der sowohl den Ablauf des Eingriffs als auch Ereignisse, wie z.B. Positionsänderungen oder Röntgenaufnahmen, enthält.

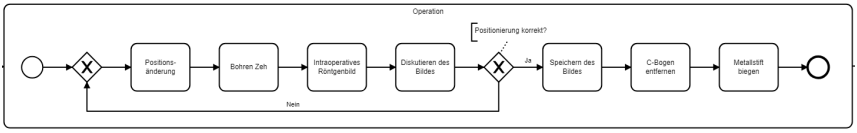



Abbildung 1: *Ausschnitt einer Klauenzehenkorrektur, modelliert in BPMN*

Die Modellierung der Prozesse, die im Hintergrund für eine Situationserkennung benötigt werden, setzt voraus, dass Eingriffe derselben Art unabhängig ihrer Einflussfaktoren über ein Modell abgebildet werden können. Dafür müssen die Modelle flexibel genug sein, um auch in Ausnahmesituationen wie Komplikationen, spontanen Änderungen o.Ä. angewandt werden zu können. Da ein Eingriff nicht nur strukturierte Phasen besitzt, ist es nötig,

alternative Modellierungstechniken für komplizierte, unstrukturierte und variable Prozesse zu nutzen. Für die Darstellung von strukturierten und unstrukturierten Abläufen sowie komplexen Entscheidungen eignet sich deshalb eine Kombination aus BPMN, Case Management Model and Notation (CMMN) und Decision Model and Notation (DMN) zur Darstellung chirurgischer Prozesse [5].

Durch die Kombination aus BPMN, CMMN und DMN ist es möglich, flexible Workflowmodelle von Eingriffen zu erstellen, deren Struktur sowie implementierte Regeln sich auch auf andere intraoperative Prozesse anpassen lassen. Dadurch sollen ebenfalls definierte Situationen und deren Parameterkombinationen übertragen und genutzt werden können. Ist die Adaptivität intraoperativer Prozesse gegeben, kann die Erkennung spezifischer Situationen unabhängig vom Eingriff gewährleistet werden.

# 3 Ergebnisse

Auf Basis der Datenerhebung und -analyse wurde ein Konzept für die zu realisierende Situationserkennung entworfen, das die benötigte Adaptivität erwarten lässt. Dieses soll die Verwendung für unterschiedliche Eingriffe in Kombination mit verschiedenen Systemen ermöglichen. Das erstellte Konzept nutzt hierfür einen adaptiven Ansatz, der zunächst eine abstrakte Phasenerkennung ermöglichen soll, aber weitere Granularitätsstufen und somit eine Spezifizierung vorsieht.

## 3.1 Aufbau der Situationserkennung

Abbildung *2* zeigt die Kommunikation der geplanten Situationserkennung mit Geräten im Operationssaal sowie einem Workflow Management System (WFMS). OP-Geräte und -Systeme generieren Rohdaten, wie z.B. Geräteparameter. Diese von verschiedenen Quellen im Operationssaal erworbenen prozessrelevanten Informationen werden von einer Schnittstelle innerhalb der Situationserkennung gesammelt, die damit eine flexible Anbindung an im Operationssaal verfügbare Geräte und Systeme gewährleistet. Die erfassten Rohdaten werden daraufhin verarbeitet und interpretiert. Hierfür werden diese an die Dateninterpretationseinheit weitergeleitet, die mit Datenspeichern kommuniziert. Die lesbaren Daten werden daraufhin durch die Situationsinterpretationseinheit weiterverarbeitet, um die Situation in Bezug auf den Kontext zu interpretieren. Die sich daraus ergebenen Informationen werden vom WFMS inklusive der Workflow Engine entsprechend dem Kontext genutzt. Die erkannten Situationen dienen dem Situationsbewusstsein und der Steuerung des Prozessablaufs entlang eines hinterlegten Workflowmodells. Das Modell bildet hierbei den stattfindenden intraoperativen Prozess ab und enthält Informationen zu diesem. Der aktuelle Prozessschritt kann daraufhin wieder als Input in die Situationserkennung miteinfließen.

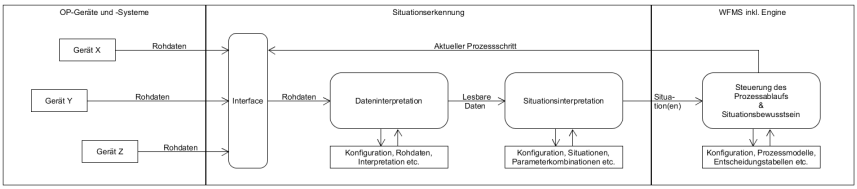



Abbildung 2: *Kommunikation der Situationserkennung mit OP-Geräten und -Systemen sowie einem WFMS inklusive Engine*

Die Situationserkennung mit seinem Situationsbewusstsein bildet die Grundlage zur situationsbedingten Steuerung von Systemen wie dem *OR-Pad*. Sobald eine spezifische Situation erkannt wird, kann über den jeweiligen Prozessschritt auf Informationen zur automatisierten Ausführung von Aktivitäten zugegriffen werden.

## 3.2 Konzept der Situationserkennung

Um eine adaptive Situationserkennung zu realisieren, sollen die für das Situationsbewusstsein benötigten Situationen und deren Parameterkombinationen einheitlich für alle Eingriffe definiert und angewandt werden. Dadurch soll ermöglicht werden, Aktionen, wie z.B. das Entfernen von Gewebe, in unterschiedlichen Eingriffen durch dieselben Parameterkombinationen als diese Situation identifizieren zu können, unabhängig davon wie die Unterprozesse im Detail zwischen Arten, Methoden etc. variieren. Die einheitliche Nutzung definierter Situationen und deren Parameter soll die Übertragung der Situationserkennung auf neue Eingriffe möglich machen. Zudem kann die Situationserkennung bei jedem Eingriff lernen, sodass von diesem Wissen auch alle folgenden Eingriffe profitieren können.

Diese Übertragbarkeit setzt voraus, dass Phasen und Schritte von modellierten Eingriffen zum Ausführen durch eine Workflow Engine in Beziehung zueinander gesetzt werden. Um einheitliche Begrifflichkeiten innerhalb der unterschiedlichen Modelle zu realisieren, können Ontologien, die eine Menge von Begrifflichkeiten und zwischen ihnen bestehende Beziehungen darstellen, genutzt werden. Ontologien und semantische Technologien gelten hierbei als Schlüssel für standardisierte chirurgische Prozessmodelle [6]. Zur Realisierung einheitlicher Begrifflichkeiten wurden zunächst beobachtete Prozessschritte in abstrakten Phasen klassifiziert (z.B. „Schnitt“, „Gewebeentfernung“, „Materialnutzung“, „Kontrolle“ oder „Naht“). Diese sollen intern mit der entsprechenden Situation und somit Parameterkombination referenziert werden, z.B. über Entscheidungstabellen mit DMN.

Für die Erkennung der abstrakten Phasen soll eine generalisierte Phasenerkennung realisiert werden, die in intraoperativen Prozessen einheitlich definierte Phasen unabhängig von detaillierteren Informationen erkennt. Im Beispiel der „Materialnutzung“ wird hierbei nicht zwischen dem Einsetzen eines Wirbelkörpers, einer Platte oder anderen Aktionen unterschieden. Dadurch können die Phasen auf andere Prozesse übertragen werden. Die Nutzung von einheitlichen Phasen (siehe Abbildung *3*) ermöglicht es zudem, Prozesse besser miteinander in Beziehung zu setzen und auf ein Minimum an Prozessschritten zu reduzieren.

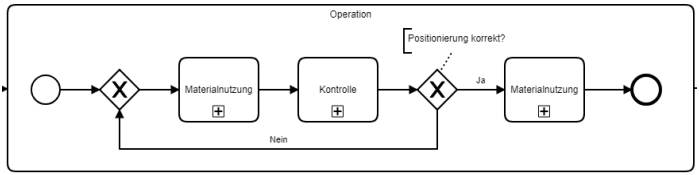



Abbildung 3: *Ausschnitt der Klauenzehenkorrektur aus Abschnitt 2.2, modelliert in abstrakten Phasen*

Ist eine spezifischere Erkennung bei einem Eingriff erwünscht, um in bestimmten Situationen unterschiedliche Informationen anzuzeigen, kann das Konzept erweitert werden. Hierfür werden innerhalb der einheitlichen Phasen detaillierte Situationen im Workflowmodell abgebildet, die daraufhin ebenfalls durch die Situationserkennung erkannt werden sollen. Dadurch kann der Operateur gezielter unterstützt werden.

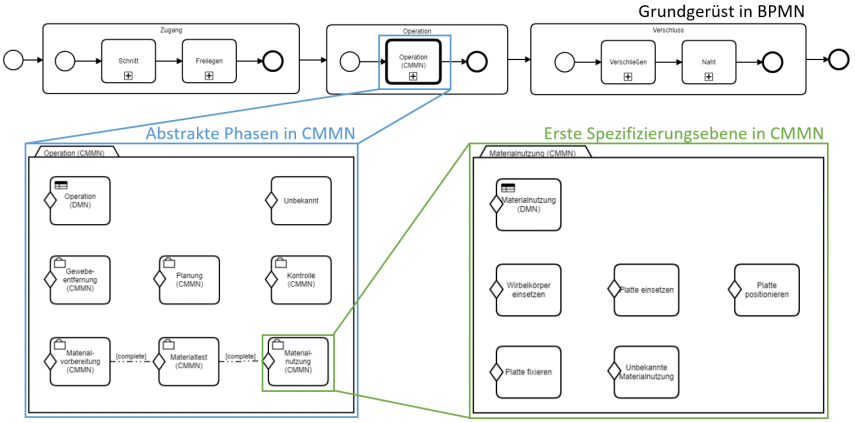



Abbildung 4: *Konzept der Situationserkennung in unterschiedlichen Ebenen*

Für die Umsetzung der Workflowmodelle bildet BPMN, wie in Abbildung *4* dargestellt, das *Grundgerüst* des Modells und gibt die grobe Struktur des Eingriffs vor. Die darunter definierten *abstrakten Phasen* werden in CMMN dargestellt. Falls eine Spezifizierung benötigt wird, können diese wiederum spezifischere CMMN Modelle enthalten. Diese bilden die *erste Spezifizierungsebene*. DMN dient innerhalb der Modelle als Entscheidungsunterstützung. Je nach Anwendungsfall reicht die abstrakte Phasenerkennung aus oder es wird eine Erkennung auf Basis der ersten, zweiten oder x-ten Spezifizierungsebene angestrebt. Je feingranularer, desto mehr Informationen werden innerhalb der Phasen modelliert. Dadurch kann zwischen unterschiedlichen Gewebearten, Materialien, Planungsinformationen etc. unterschieden werden. Beim Erkennen einer Phase wird dabei sowohl die

übergeordnete Phase als auch, wenn möglich, die detailliertere Phase erkannt. Das Konzept lässt sich beliebig auf weitere Unterphasen und Schritte erweitern.

Je abstrakter die Prozesse (geringe Granularität), desto einheitlicher kann deren Erkennung durch eine Situationserkennung realisiert werden. Eine Umsetzung auf der abstrakten Ebene ist deshalb zunächst sinnvoll, um grobe Abläufe eingriffsunspezifisch erkennen und damit unterstützen zu können. Zur Realisierung der Situationserkennung soll deshalb als erster Schritt die abstrakte Phasenerkennung ermöglicht werden. Dadurch wird die Unterstützung des Operateurs zunächst allgemein gehalten und eine zuverlässige Phasenerkennung realisiert, die auf alle Eingriffe anwendbar ist. Hierfür sollen übergeordnete Situationen und Parameterkombinationen für die Hauptphasen definiert und auf Workflowmodelle übertragen werden. Das dadurch entstehende Situationsbewusstsein soll, unter anderem, für die Unterstützung in den Anwendungsfällen aus Abschnitt 2.1 genutzt werden.

Nachdem eine Phasenerkennung auf abstrakter Ebene erfolgreich umgesetzt wurde, wird eine Optimierung bzw. Präzisierung von dieser angestrebt, bei der Spezifizierungsebenen miteinbezogen werden. Hierfür sollen Situationen und Parameterkombinationen, die spezifischere Prozessschritte beschreiben, integriert werden. Die Situationserkennung kann dadurch entsprechend der Spezifizierungsebenen Schritt für Schritt hinsichtlich der Präzision der Erkennungen verfeinert werden.

Zur Umsetzung des Konzepts wird eine Anbindung an ein WFMS inklusive Engine benötigt (siehe Abbildung *2*). Hierfür wurde bereits die für das vorgestellte Konzept erforderliche Infrastruktur umgesetzt und prototypisch erprobt [7]. Das WFMS aus [7] ermöglicht die automatische Ausführung von Prozessen über die Camunda BPM Workflow Engine. Eine Middleware erlaubt es zudem, verschiedene Systeme mit der Workflow Engine zu verbinden, um mit dieser zu interagieren. Grafische Benutzeroberflächen ermöglichen hierbei die Darstellung von Informationen sowie die Steuerung der Prozesse. Zur Reduzierung der Interaktion ist die Anbindung einer Situationserkennung vorgesehen.

# 4 Diskussion

Die zur Umsetzung des vorgestellten Konzepts erforderliche Infrastruktur soll als Grundlage zur Entwicklung der adaptiven Situationserkennung dienen. Mit dieser soll die Informationserfassung und -anzeige im intraoperativen Bereich unterstützt werden.

## 4.1 Adaptivität

Durch das flexibel erweiterbare Konzept der Situationserkennung soll es zunächst möglich sein, Situationen auf Basis von generalisierten und einheitlichen Phasen im intraoperativen Bereich zu erkennen. Daraufhin kann das Konzept auf Anwendungsfälle spezifiziert werden. Im Gegensatz zu Lösungen, die sich auf bestimmte Eingriffe beschränken, soll das Konzept auf jeden Eingriff übertragbar sein. Diese Adaptivität bezieht sich auf alle Komponenten der Situationserkennung und betrifft demnach den Datenerwerb prozessrelevanter Informationen von verschiedenen Quellen im OP, die Verarbeitung und Interpretation der erworbenen Daten sowie die Automatisierung von individuell konfigurierten Aktivitäten zur Unterstützung durch ein WFMS.

Die Basis der Adaptivität bildet hierbei das zugrundeliegende Workflowmodell, das Flexibilität durch CMMN innerhalb der unstrukturierten Phasen bietet, wohingegen strukturierte Phasen in BPMN benötigte Abhängigkeiten beinhalten. Durch eine einheitliche Struktur und konsistente Begrifflichkeiten kann an richtiger Stelle ausreichend Flexibilität ermöglicht werden, ohne dabei die Grundstruktur der Hauptphasen zu beeinflussen. Informationen aus unterschiedlichen Quellen können je nach Kontext über, für spezifische Situationen definierte, Parameterkombinationen interpretiert und mit Prozessschritten im Workflowmodell assoziiert werden. Das Modell verweist daraufhin auf die Aktionen, die ausgeführt werden sollen, sodass je nach Phase Informationen erfasst oder angezeigt werden. Weitere automatisierte Aktivitäten, wie z.B. das Senden von Nachrichten, können ebenfalls implementiert werden.

## 4.2 Mehrwert

Das vorgestellte Konzept zur Situationserkennung für das *OR-Pad* und andere Systeme ermöglicht es, Phasen zu erkennen. Entsprechend der erkannten Phase können daraufhin bestimmte Aktionen, die für den Anwendungsfall individuell konfigurierbar sind, automatisiert ausgeführt werden. Die Definition, in welcher Phase was passieren soll, wird abhängig von Unterphasen, der Operation, dem Operateur und auch dem Klinikum sein.

Im Falle des *OR-Pad* können kontextrelevante Informationen wie z.B. präoperative Röntgenbilder zur Unterstützung des Operateurs und seines Teams angezeigt werden. Der Mehrwert ist hierbei abhängig von der

Operation. Zudem können die Restdauer sowie Verzögerungen eingeschätzt werden. Jegliche Zeitpunkte, die eine Phase im Workflowmodell darstellen (z.B. Hautschnitt), können in die OP-Dokumentation aufgenommen werden. Eine Reduzierung auf die wichtigsten Ereignisse (z.B. Schnitt, Organentfernung, Materialnutzung, Naht) kann sinnvoll sein. Diese Automatisierung unterstützt den Operateur durch die Reduzierung der Interaktion, um kontextrelevante Informationen anzuzeigen, sowie Kommunikation, um Informationen über zeitliche Aspekte zu erhalten. Zudem wird die Erstellung von OP-Dokumentations-Einträgen, die andernfalls manuell getätigt werden müssten, unterstützt.

## 5 Zusammenfassung

Basierend auf der Datenerhebung und -analyse wurden Anwendungsfälle definiert, Prozessmodelle erstellt und ein Konzept für eine adaptive Situationserkennung entworfen. Das Konzept basiert auf der Erkennung von Phasen in unterschiedlicher Granularität durch Informationen aus einer Vielzahl an Datenquellen. Der Fokus liegt hierbei zunächst auf einer Phasenerkennung, die generalisierte Prozessschritte unabhängig des Eingriffs durch einheitliche Situationen und Parameterkombinationen erkennt und den Operateur daraufhin mit kontextrelevanten Informationen unterstützt. Durch die einheitliche Basis lässt sich das Situationsbewusstsein auf andere intraoperative Prozesse übertragen. Die Lösung kann zudem für Eingriffe spezifiziert werden, um den Operateur und sein Team gezielter unterstützen zu können. Die Weiterentwicklung sowie Evaluation sind zukünftige Schritte zur Optimierung des Konzepts.

## Danksagung

Diese Forschung wird durch Fördermittel vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg sowie vom Europäischen Fonds für regionale Entwicklung unterstützt.

Wir danken allen Klinikern für die durchgeführten Hospitationen und Unterstützung bei der Datenerhebung zur Definition der Anwendungsszenarien und Erstellung der Prozessmodelle.

## Referenzen

[1] Dergachyova O, Bouget D, Huaulmé A, Morandi X, Jannin P, *Automatic Data-Driven Real-Time Segmentation and Recognition of Surgical Workflow*, in: Int J Comput Assist Radiol Surg, 2016 Jun, 11(6), S. 1081-1089, DOI: 10.1007/s11548-016-1371-x (2016).

[2] Kranzfelder M, Schneider A, Gillen S, Feussner H, *New technologies for information retrieval to achieve situational awareness and higher patient safety in the surgical operating room: the MRI institutional approach and review of the literature*, in: Surg Endosc, 2011 Mar, 25(3), S. 696-705, DOI: 10.1007/s00464-010-1239-z (2011).

[3] Lalys F, Riffaud L, Morandi X, Jannin P, *Automatic phases recognition in pituitary surgeries by microscope images classification*, IPCAI 2010: International Conference on Information Processing in Computer-Assisted Interventions, Geneve, Switzerland, S. 34-44, Springer-Verlag Berlin, Heidelberg (2010).

[4] Feige K, Gollnick I, Schmitz P, Strauss G, *The application of surgical procedure manager (SPM): first experience with FESS*, in: Eur Arch Otorhinolaryngol, 2017 Sep, 274(9), S. 3407-3416, DOI: https://doi.org/10.1007/s00405-017-4658-9 (2017).

[5] Wiemuth M, Junger D, Leitritz M A, Neumann J, Neumuth T, Burgert O, *Application fields for the new Object Management Group (OMG) Standards Case Management Model and Notation (CMMN) and Decision Management Notation (DMN) in the perioperative field*, in: Int J Comput Assist Radiol Surg, 2017 Aug, 12(8), S. 1439-1449, DOI: 10.1007/s11548-017-1608-3 (2017).

[6] Gibaud B, Forestier G, Feldmann C, Ferrigno G, Gonçalves P, Haidegger T, Julliard C, Katić D, Kenngott H, Maier-Hein L, März K, de Momi E, Nagy D Á, Nakawala H, Neumann J, Neumuth T, Rojas Balderrama J, Speidel S, Wagner M, Jannin P, *Toward a standard ontology of surgical process models*, in: Int J Comput Assist Radiol Surg, 2018 Sep, 13(9), S. 1397-1408, DOI: 10.1007/s11548-018-1824-5 (2018).

[7] Wiemuth M, Burgert O, *A workflow management system for the OR based on the OMG standards BPMN, CMMN, and DMN*, SPIE Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling, San Diego, United States, Proc 10951, S. 79, DOI: https://doi.org/10.1117/12.2512724 (2019).

# Ergonomic Assessment of Operating Room Setups for Orthopedic Reconstructive Surgery

*J. Neumann[1], C. Angrick[1], A. Roth[2], T. Neumuth[1]*

*[1]Innovation Center Computer Assisted Surgery (ICCAS), Leipzig University, Leipzig, Germany*
*[2]Department of Orthopaedic, Trauma and Plastic Surgery, Division of Endoprothetic Joint Surgery and General Orthopaedics, University of Leipzig Medical Center, Leipzig, Germany*
***Contact: Juliane.Neumann@iccas.de***

***Abstract***

Orthopedic reconstructive surgery is characterized by a high risk of work-related musculoskeletal injuries such as back and neck pain, shoulder tendonitis and arthritis, due to poor ergonomics in the operating room (OR). Although the work environment has a significant impact on the ergonomics, the internal OR layout has not yet been analyzed from an ergonomic perspective. The aim of this study is to evaluate and optimize the ergonomic situation of OR setups in Total Hip Arthroplasty and Total Knee Arthroplasty surgeries. For this purpose, an ergonomic assessment has been performed with the OWAS method. The evaluation results showed ergonomic critical positions mostly for the surgeon, the scrub nurse and the assistants regarding bent or twisted back and head postures as well as static positions and prolonged standing during surgery. To address this problem, requirements for the design of ergonomically improved OR setups were presented.

**Keywords**: Operating Room Layout, OR Ergonomics, OWAS, Orthopedic Reconstructive Surgery

## 1 Problem

Orthopedic reconstructive surgery is characterized by repetitive surgical activities with high physical demand. The exertion of force needed e.g. for joint dislocation, prolonged standing, as well as static positions and static muscle load, lead to an increased risk of musculoskeletal injuries of the whole operating room team and especially of the surgeon [1],[2]. The most common work-related injuries are back pain, shoulder tendonitis, arthritis and lumbar disc herniation, which often need to be treated with a surgical intervention itself [3]. Previous studies have shown, that poor ergonomics of the surgical work environment, as well as abnormal body and neck positions during long periods of the surgery, could cause the remarkably large number of injuries in this high-risk surgical profession [1],[3]. To improve the ergonomics in the operating room (OR), ergonomic evaluation was already done for different surgical disciplines, such as general surgery [4], ENT [4] and minimal-invasive surgery [5],[6] as well as for different occupational groups, like surgeons [7] and scrub nurses [8].

Although the work environment has a significant impact on the ergonomics, the internal OR layout has not yet been analyzed from an ergonomic perspective. The aim of this study is to evaluate and optimize the ergonomic situation of OR setups in Total Hip Arthroplasty (THA) and Total Knee Arthroplasty (TKA) surgeries. For this purpose, the body positions of the operating room staff were ergonomically analyzed with the ultimate goal to improve the internal OR layout considering positions of the OR table, instrument tables (mayo stands, instrument/equipment stands, solutions stands, etc.), displays, medical devices (e.g. c-arm and anesthesia equipment), as well as the position of the OR staff during surgery. Therefore, the body positions of the OR team during THA and TKA surgery were analyzed with the *Ovako Working Posture Analyzing System* (OWAS), a method for ergonomic assessment. Based on the evaluation results, ergonomic requirements for optimized OR setups were defined and subsequently ergonomically improved OR setups could be designed. The objective is to optimize the operating room processes efficiently and ergonomically to create an amended work environment with safe body postures and movements during orthopedic reconstructive surgery. In this way, the number of physical complaints among the OR personnel, such as back and neck pain should be reduced.

## 2 Material and Methods

### 2.1 OR setups for THA and TKA surgery

THA and TKA surgeries are among the most performed surgeries worldwide. In THA the hip joint is replaced by a prosthetic implant to treat arthritic pain or hip fractures. During TKA the knee joint is replaced to relieve debilitating pain or osteoarthritis. For ergonomic assessment, the body postures of the whole OR team were observed for one THA and one TKA surgery at the University of Leipzig Medical Center, Division of Endoprothetic Joint Surgery and General Orthopaedics in 2016. Thereby, the commonly used OR setup for left

side TKA (Figure 1, left side) and THA surgery (Figure 1, right side) were recorded and modeled in a graphical 3D simulation environment. The corresponding right-side setups are mirrored along the operating table. The ground plan of the orthopedic OR was digitized. Measurements and heights of furnishing, e.g. OR table, instrument tables, medical devices, OR display, etc., were acquired for ergonomic assessment.

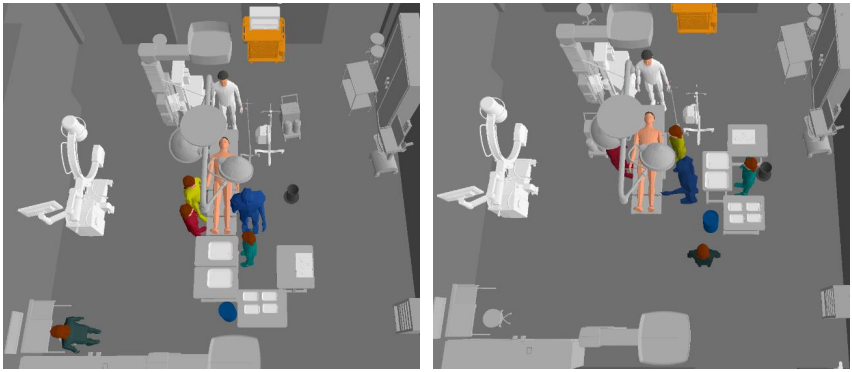


Figure 1 - *TKA setup (left-side) and THA setup (right-side) with OR and instrument tables, medical devices and positions of the OR personnel – surgeon (blue), scrub nurse (green), 1st assistant (yellow), 2nd assistant (red), circulator (grey) and anesthesiologist (white).*

## 2.2 Ergonomic assessment

In order to reduce risk factors for musculoskeletal injuries and improving the ergonomics in the OR environment, the working positions of the OR team were analyzed quantitatively. Ergonomic assessment techniques, such as OWAS [9], RULA [10] or REBA [11] have been developed, for the analysis of stress caused by different body postures in the industrial domain. In this study, the OWAS technique was utilized due to the simple coding scheme for live observation. Additionally, OWAS has already been applied to the operating room for different surgeries, e.g. evaluation of ergonomics in minimal-invasive interventions [6], [7].

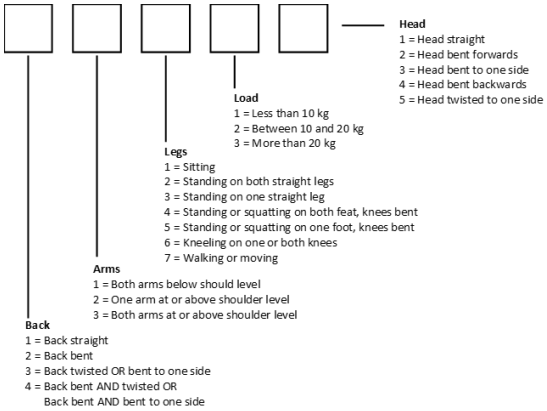



Figure 2 - *OWAS Code with an additional digit for the head position according to [8]*

The *Ovako Working Posture Analyzing System* (OWAS) was developed by Karhu et al. in 1977 [9] and is used to classify working postures of the back, arms, legs and optionally the head to determine, if the current position is acceptable under ergonomic considerations. During the OWAS assessment, an observer records the working postures of the actors in the form of a coded value once in regular intervals of 30 or 60 seconds. The original OWAS code consists of the four digits for the posture of the back, the arms, the legs, and the carried load (Figure 2). Since 7 different postures are represented for the back, 4 for the arms, and 3 for the legs, it results in 84 standardized basic body postures and 252 working postures, if the workload is also considered. The position of

the head is also included in the assessment, which increases the number of possible working positions to 1.260.

Subsequently, for each predefined posture category, the percentage of time in the total duration of the activity was calculated. Thereby, an action category can be derived by comparing the results to benchmark tables (available in e.g. [12]), which evaluate the ergonomics of the working posture depending on their temporal proportion on the total duration of the surgery. Based on the benchmark tables, an action category (AC) can be derived. The action categories could indicate a change in the working environment, if necessary. According to [9], there are four different action categories:

1 (green): common working postures, which do not need any change
2 (yellow): working postures must be considered during the next regular check
3 (orange): working postures must be considered and changed in the near future
4 (red): working postures, which need immediate change

In addition to the assessment of single postures, posture combinations of the back, arm, and leg postures and the workload were calculated according to the benchmark table in Figure 3.

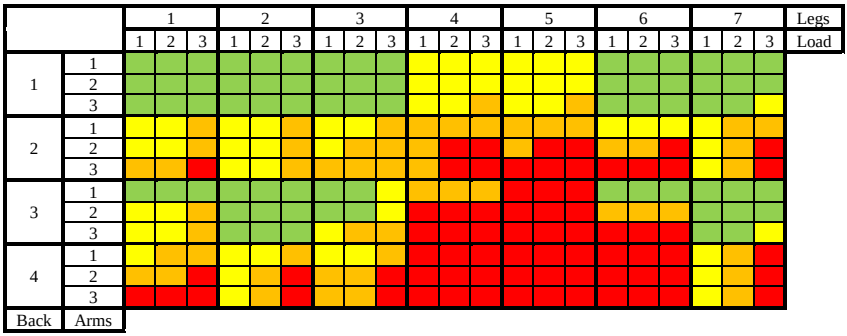



Figure 3 – *Classification of posture combinations to action categories (AC) 1-4 (1=green, 2=yellow, 3=orange, 4=red) [12]*

The OWAS assessment was performed for the surgeon, assistants, scrub nurse and the circulator in THA and TKA surgery. The postures were encoded once in an interval of 60 seconds by a human observer. A single posture assessment of back, arms, legs, and the head, as well as posture combinations of arms, legs, back and workload, were calculated and allocated to the appropriate action category. Finally, the *Lundqvist index L* [13] was calculated according to the *percentages of the individual action categories* (*PAC*) [12]. This is a cumulative load index, which is calculated as follows:

$$L = 1 * PAC1 + 2 * PAC2 + 3 * PAC3 + 4 * PAC4 \; [100; 400]$$

The indices were calculated for the single postures and posture combinations. Higher values indicate a higher workload. To obtain only one index, the total Lundqvist index, a mean value of the single posture and posture combinations, was calculated. In this way, the workloads of different persons can be compared with each other.

### 2.3 Requirements analysis for ergonomically improved OR setups

Due to the lack of literature in the objective determination of how an optimal OR setup could be defined for different intervention types and surgical disciplines, a non-deterministic list of requirements was identified with the help of process analysis and clinical expert interviews. For this purpose, 22 orthopedic procedures were observed in respect to their functional, spatial and ergonomic requirements. Additionally, a workshop with 3 senior surgeons and 5 scrub nurses was conducted to define the requirements of an optimal OR setup.

## 3 Results

### 3.1 OWAS assessment of the TKA Setup

*Single postures*

In Figure 4 the results of the single posture assessment of the OR team as well as the calculation of the total Lundqvist index are presented for TKA surgery. The graphs show also the color-coded assignment of the body postures to one of the four action categories (AC). Thereby, the surgeon was observed with the most ergonomic deficits due to a mostly bent or twisted back posture, which is often necessary for the instrument handover in the

TKA setup. In the setup, the scrub nurse is positioned on the left side of the surgeon, which requires a theoretical 90° rotation of the surgeon to grab an instrument. Instead of performing a whole 90° rotation, in practice, the surgeon and the scrub nurse bent and twist their backs during handover. Therefore, the calculation of the Lundqvist index was higher for the surgeon compared to the other OR team members. Except for the circulator, the OR personnel stands continuously during surgery, which resulted in the AC2 leg classification. The overall assessment of the Lundqvist index showed that the value is relatively high for the surgeon and ergonomics should be improved.

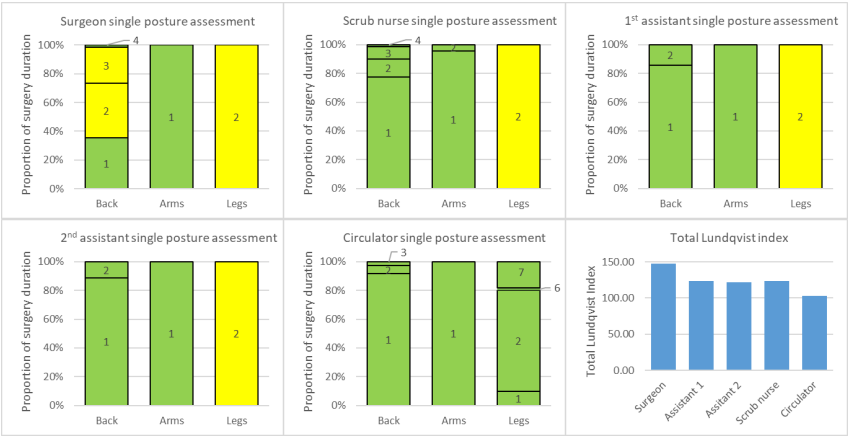



Figure 4 – *Single posture assessment of the OR team with coded body postures and classification into AC 1 -4 based on the proportions of the total surgery duration in [%] and Total Lundqvist index for TKA surgery*

*Posture combinations*

In Figure 5 (left side) the results of posture combinations including the workload are depicted for all OR team members. The assessment indicates that the working positions of the assistants, the scrub nurse, and the circulator are acceptable, while the working position of the surgeon should be considered for improvement in the near future.


Figure 5 - *Distribution of action classes (AC) of the body posture combinations of arms, legs and the back (left) and the head posture (right) of the OR personnel for TKA setup*

The head posture classification of the OR staff is presented in Figure 5 (right side). The surgeons' head was bent in 83%, the assistants' in 82% and the scrub nurses' in 61% of the surgery, which was caused by the height of the OR table and is barely avoidable. Although, this indicates a huge physical strain and leads to a critical ergonomic position. Permanently bent forward and twisted head positions should be avoided due to potential damage to the cervical spine [14]. Additionally, the head of the surgeon and the scrub nurse were twisted to one side, which is caused by the positioning of the instrument tables and happened mainly during instrument handover.

## 3.2 OWAS assessment of the THA Setup

*Single postures*

The results of the single posture assessment and the Lundqvist index for THA surgery are presented in Figure 6. Especially, the 1st assistant was observed with a bent back position most of the surgery, which also results in a higher Lundqvist index. This was mostly caused by a position change between the surgeon and the assistant during

the insertion of the femur implant, which led to a better working position of the surgeon but also in ergonomic deficits of the scrub nurse and the 1st assistant who had limited access to the operating area. For instrument handover, the surgeon theoretically performs a 180° rotation which results in a bent and twisted back position in clinical practice. Additionally, the scrub nurse needed to bent forward over the instrument tables for handover.

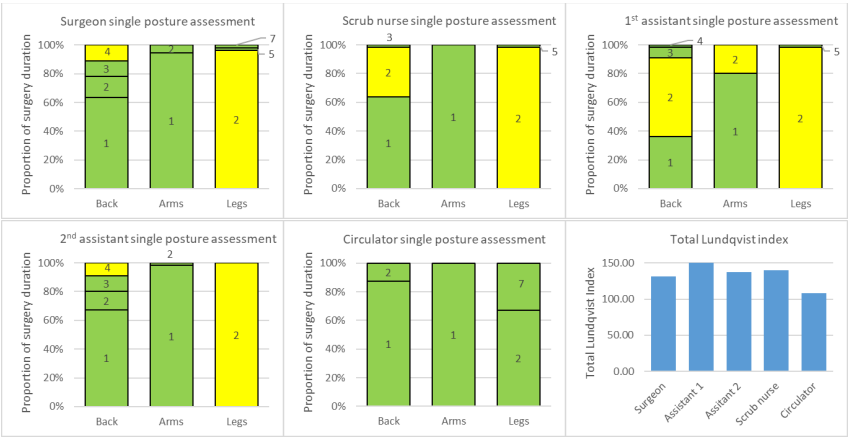



Figure 6 - *Single posture assessment of the OR team with coded body postures and classification into AC 1 -4 based on the proportions of the total surgery duration in [%] and Total Lundqvist index for THA surgery*

*Posture combinations*

The results of the posture combination assessment are presented in Figure 7 (left side). Due to a threat of potential health damage, the classification of AC 3 and AC 4 resulted in a need for action with regard to the back posture of the 1st assistant and the scrub nurse. The OR layout should be optimized to reduce rotational movement during instrument handovers. Additionally, the circulator was observed lifting heavy loads, e.g. during fetching heavy supplies (implants, materials), which results in a classification of AC3 in 4% of the time.

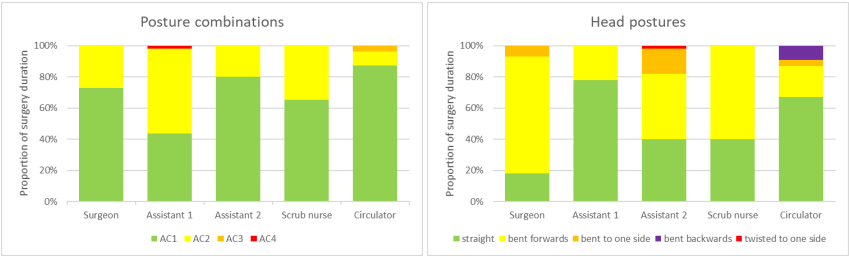



Figure 7 - *Distribution of action classes (AC) of the body posture combinations of arms, legs and the back (left) and the head postures (right) of the OR personnel for THA setup*

The head posture classification of the THA surgery is presented in Figure 7 (right side). Especially, the heads of the surgeon and scrub nurse were bent forward in 75% and in 60% of the surgery time, respectively, which is highly critical under ergonomic considerations. The head of the 2nd assistant was bent forward in 42%, bent to one site in 16% and twisted in 2% of the surgery, which indicates poor visibility of the operating area. In addition, the circulator had been bent the head backward in 9% of the time to see the information on the working place display without turning the back, which should be prevented in the future.

## 3.3 Design of ergonomically improved OR setups

Nowadays, the OR layout, including the position of OR table, instrument tables, and the OR personnel, is defined by the preference of the surgeon or the scrub nurse. Even for the same intervention type in the same institution, the OR layout varies widely. Although the OR layout is affecting the surgical processes, the ergonomics as well as team collaboration substantially, there is no research available on how an optimal OR setup should be designed.

From an ergonomic point of view the following requirements should be considered during OR setup design:

- If possible, all OR team members should have a direct view of the operating area, which prevents twisted or bent head and back postures.
- The instrument tables should be positioned in the direct vicinity of the surgeon and in a U-shape in front of the scrub nurse. This minimizes rotational movement and improves ergonomic back postures.
- The height and position of the OR table and instrument tables should be adapted to the individual needs of the OR personnel to prevent an immensely bent forward head posture.
- All relevant displays in the OR should be seen by the personnel without bending or twisting the head or back.

# 4 Discussion

Although an ergonomic assessment of different OR setups considering OR table, instrument table, and personnel positions have not done before, the results of the OWAS assessment confirm the findings of previous studies regarding the poor ergonomics in the OR [1], [3]. During orthopedic reconstructive surgery, the OR personnel is exposed to severe physical stress and body postures with ergonomic deficits, which may lead to neck or back pain, musculoskeletal injuries or cardiovascular problems. Especially, the surgeon has a high Lundqvist index, which is mainly caused by repetitive movements and static positions with bent head and neck positions for a direct view of the operating area. Orthopedic interventions are lengthy surgeries, which also include static positions with prolonged standing. Except for the circulator, who is able to walk freely in the OR and has the possibility to sit down, all OR team members have a high Lundqvist index for the leg postures. There is only small room for optimization. Nonetheless, the use of compression stockings, anti-fatigue mats or standing aids would be a possibility to improve the ergonomic situation and decrease edema of the lower extremity [3]. Additionally, leg and muscle exercises such as standing on one leg and circling the ankle could relief the physical stress on the lower extremities. The head position is rarely influenced by the OR layout but by the height and position of the OR and instrument tables. If possible, the operating area should be in the direct view of all OR team members, which prevents twisted or bent head and back postures. In addition, every possibility should be utilized to adapt height-adjustable instrument and OR tables to the individual height as well as to compensate the body size differences between the persons through platforms. Furthermore, all relevant OR displays should be positioned in a way, that all team members have a direct view of the displayed information.

Even though the options to improve ergonomics in the OR are very limited, the assessment of the TKA and THA setups indicates a need for ergonomic reconsideration of the OR setup and a necessary change in the table layout. For this purpose, requirements for ergonomic improvement of the OR setup have been defined to reduce rotational movement during instrument handover and improve head and back postures of the OR team. Future work consists of the design, implementation, and evaluation of improved OR layout for reconstructive orthopedic surgery.

# 5 Conclusion

For the first time, an ergonomic assessment of different OR setups considering the positions of OR tables, medical equipment, and the OR personnel has been performed for orthopedic reconstructive surgery. The overarching goal is the design of an efficiently and ergonomically improved OR environment. For this purpose, the ergonomic assessment was done with the OWAS method for THA and TKA setups. The evaluation results showed ergonomic critical positions mostly for the surgeon, the scrub nurse and the assistants regarding bent or twisted back and head postures as well as static positions and prolonged standing during surgery. To address this problem, requirements for the design of ergonomically improved OR setups were presented.

### Acknowledgment

This work was funded by the German Federal Ministry of Education and Research (BMBF). The authors would like to thank the staff of the Department of Orthopaedic, Trauma and Plastic Surgery, Division of Endoprothetic Joint Surgery and General Orthopaedics at the University of Leipzig Medical Center for their kind support.

# References

[1] J.D. Lester, S. Hsu, C.S. Ahmad, *Occupational hazards facing orthopedic surgeons*, Am. J. Orthop. Belle Mead NJ, 41(3), 132–139, 2012.

[2] S.M. Mirbod, H. Yoshida, K. Miyamoto, K. Miyashita, R. Inaba, H. Iwata, *Subjective complaints in orthopedists and general surgeons*, Int. Arch. Occup. Environ. Health, 67(3),179–186, 1995.

[3] S.M. Alqahtani, M.M. Alzahrani, M. Tanzer, *Adult Reconstructive Surgery: A High-Risk Profession for Work-Related Injuries*, J. Arthroplasty, 31(6),1194–1198, 2016.

[4] I.J. Kant, L.C. de Jong, M. van Rijssen-Moll, P.J. Borm, *A survey of static and dynamic work postures of operating room staff*, Int. Arch. Occup. Environ. Health, 63(6), 423–428, 1992.

[5] L. Shan, B. Shan, D. Graham, A. Saxena, *Total hip replacement: a systematic review and meta-analysis on mid-term quality of life*, Osteoarthritis Cartilage, 22(3), 389–406, 2014.
[6] K. Radermacher, C.V. Pichler, G.Rau, *Aspects of ergonomics in Minimal-Invasive Surgery — Analysis and approaches*, *14th IEEE EMBS*, 4, 1564–1565, 1992.
[7] R. Berguer, G.T. Rab, H. Abu-Ghaida, A. Alarcon, J. Chung, *A comparison of surgeons' posture during laparoscopic and open surgical procedures*, Surg. Endosc., 11(2), 139–142, 1997.
[8] J.A. Engels, J.A. Landeweerd Y. Kant, *An OWAS-based analysis of nurses' working postures*, Ergonomics, 37(5), 909–919, 1994.
[9] O. Karhu, P. Kansi, I. Kuorinka, *Correcting working postures in industry: A practical method for analysis*, Appl. Ergon., 8(4), 199–201,1977.
[10] L. McAtamney E.N. Corlett, *RULA: a survey method for the investigation of work-related upper limb disorders*, Appl. Ergon., 24(2), 91–99, 1993.
[11] S. Hignett und L. McAtamney, *Rapid Entire Body Assessment (REBA)*, Appl. Ergon., 31(2), 201–205, 2000.
[12] T. Gudehus, Entwicklung eines Verfahrens zur ergonomischen Bewertung von Montagetätigkeiten durch Motion-Capturing, ISBN: 978-3899586503, Kassel University Press, 2009.
[13] P. Lundqvist, *Working Environment in Farm Buildings*. 1988.
[14] K.K. Hansraj, *Assessment of stresses in the cervical spine caused by posture and position of the head*, Surg. Technol. Int., 25, 277–279, 2014.

# MRI-Guided Liver Tumor Ablation - A Workflow Design Prototype

*J. Alpers*[1,2], *B. Hensen*[2,3], *F. Wacker*[2,3], *C. Rieder*[4], *C. Hansen*[1,2]

[1] *Faculty of Computer Science, Otto-von-Guericke University, Magdeburg, Germany*
[2] *Research Campus STIMULATE, Otto-von-Guericke University, Magdeburg, Germany*
[3] *Hannover Medical School, Department of Diagnostic and Interventional Radiology, Hannover, Germany*
[4] *Fraunhofer MEVIS, Bremen, Germany*

Contact: jalpers@isg.cs.ovgu.de

***Abstract***

*Thermal ablation procedures such as radiofrequency ablation have become a clinically accepted treatment method for liver tumors. Using image guidance like magnetic resonance imaging (MRI), a needle-shaped instrument is navigated to the target position aiming at a complete destruction of the focal liver malignancies. Planning the intervention, navigating the instrument to the target position, and monitoring the ablation of the malignant tissue are currently three separated steps during MRI guided interventions. This is hampering the clinical workflow and results in an unnecessary amount of additional work for the clinicians due to data export and import or mental data transfer.*
*We introduce a prototype for a combined workflow design for interventional radiology using MRI guidance. The prototype is able to guide the radiologist during planning, navigation and monitoring the ablation without manual data transfer. Evaluation with four clinical experts shows a strongly positive trend to the combined workflow design.*

**Keywords**: Workflow, Navigation, Monitoring, Radiofrequency Ablation, MRI Guidance

## 1 Problem

In the recent years needle-based percutaneous liver tumor ablations, such as radiofrequency ablations, have become a promising alternative to surgical resection [1]. During these interventions a needle-shaped instrument has to be placed inside the malignant tissue. After reaching the target structure, energy is induced in the tip of the instrument leading to a temperature rise in the surrounding tissue. A temperature of $> 60°C$ results in an immediate cell death and coagulation necrosis. Correct planning of the therapy, accurate positioning of the instrument, and real-time monitoring during the ablation are very crucial for a successful treatment. Insufficient execution of any of the previously mentioned steps may lead to malignant tissue outside of the coagulation necrosis and therefore a failed therapy [2]. Accurate planning of these interventions has been mostly covered in the past years [3], [4]. Nevertheless, the transfer of the planned data into the intervention room still needs a lot of attention. One of the reasons may be the hampering of the fast data transfer due to heterogeneous software and hardware systems in the intervention room. In this case the planning data has to be manually exported and imported into the navigation software, if available. This time consuming task results in additional work for the radiologist and the assistants. The instrument placement may be guided by fast MRI sequences to support the radiologist [5]. Using this image modality, several navigation systems have been developed to provide intra-interventional information to the performing radiologist [6]–[8]. Rothgang et al.[9] tried to improve the workflow during navigation but did not provide a proper monitoring during the ablation of the tumor. MR thermometry has been used to display the temperature of the tissue surrounding the instrument tip during ablation [10]. Unfortunately, these sequences have to be planned very carefully in advance, which results in a lot of additional work for the radiologist and the assistants again. In addition, the acquisition of MR images during the ablation is usually omitted because of known interferences between the radiofrequency generator and MR scanner [11], [12]. Nonetheless, a missing monitoring of the coagulation necrosis may lead to either an insufficient or excessive ablation. Former results in a failed therapy, whereas the latter results in a successful treatment but also a higher patient trauma due to the additional coagulated tissue.

In this work, we propose a prototype for a workflow design for MRI guided thermal tumor ablation combining planning, navigation, and monitoring of the therapy, which is based on the SAFIR (Software for Interventional Radiology) toolbox [13]. This workflow is aiming at reducing the mental workload of the performing radiologists by eliminating the transfer steps between the three steps of planning, navigation and monitoring. To the best of our knowledge, there is no system available yet combining these three steps for MRI guided interventions without a manual transfer of the data from one step to another.

The outline of this paper is as follows: Section 2 describes the planning, navigation, and monitoring parts of the presented prototype as well as the evaluation of the proposed workflow with the help of four clinical experts. Section 3 shows the results of the evaluation. Section 4 shows a critical discussion of the results showing advantages and disadvantages of the proposed prototype and Section 5 concludes this paper with a short summary and a look ahead.

# 2 Material and Methods

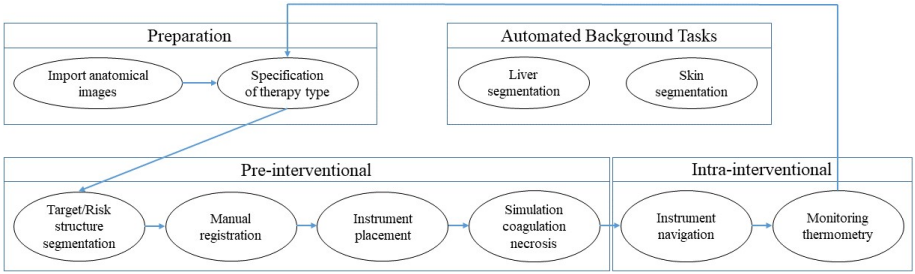



Figure 1: *Sketch of the basic workflow. Preparation takes place before the actual intervention. Automated background tasks take place during therapy creation. Pre-interventional describes the planning phase. Navigation and monitoring cover the intra-interventional workflow phases.*

In general, the proposed workflow consists of four different steps. The preparation of the intervention describes the import of already existing DICOM images via file system or DICOM node, e.g. previously acquired planning data sets or other anatomical data. Afterwards, the therapy will be created specifying the type of intervention, which will be performed. Depending on the therapy type created, automated background tasks will perform common preprocessing of the data like liver segmentation or segmentation of the patient skin mask, which may be used as an initial input for further image processing methods during the navigation or monitoring. The pre-interventional planning is part of the sequential workflow starting with the segmentation of target as well as risk structures and the manual co-registration of different anatomical data sets if necessary. The instrument placement allows for choosing an optimal trajectory path to reach a selected target structure followed by a simulation of the coagulation necrosis to verify if the tumor can be completely ablated with the chosen target point. During intervention, continuous MR images are acquired to guide the clinicians during instrument insertion. After reaching the desired target point, a thermometry sequence allows for monitoring of the tissue temperature during ablation. A graphical illustration of the workflow can be found in Figure 1.

## 2.1 Planning

The planning phase of our workflow consists of an automated preprocessing and four manual steps: segmentation, manual registration, instrument placement and simulation. The preprocessing covers the initial segmentation of the patient skin mask and the liver. The patients skin mask can be visualized during instrument insertion to aid the performing radiologist during the task of finding the correct incision point. The liver mask, in addition, may be used for additional image processing in the future as an initial region of interest, e.g. for instrument detection during the navigation phase. The tumor and additional risk structures such as vessels are extracted using semi-automatic segmentation methods. An accurate segmentation of risk structures allows the radiologist to keep track of them during navigation and therefore bypass them without damaging them. In case the patient is re-positioned after the initial planning, image acquisition or the target organ has moved due to respiration the following intra-interventional image has to be manually co-registered to the planning data set. During path planning, an optimal access path and trajectory for the instrument insertion has to be found. To determine an accurate path, various virtual instrument models can be manually placed and aligned inside the planning data set. The trajectory itself is defined by the center of the instruments ablation zone and the incision point. The planning of the instrument position is crucial to avoid impenetrable structures such as bones or other critical risk structures such as the lung or vessels. During path planning, the performing radiologist defines a target point to completely ablate the tumor. To simplify the placement, ellipsoidal ablation zones

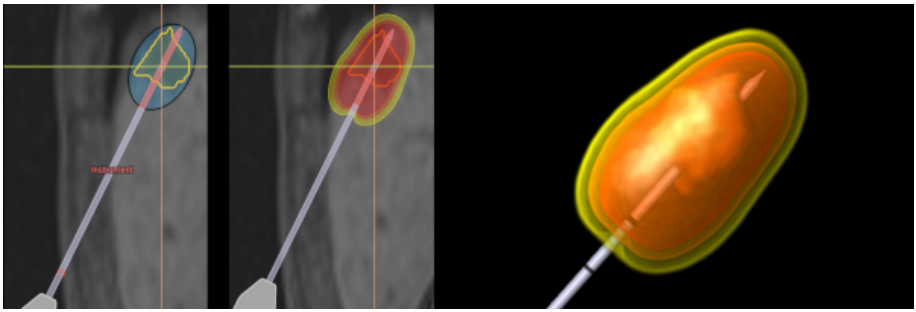


Figure 2: *Instrument placement. Left: Instrument including ellipsoidal ablation zone covering the whole target structure (tumor). Middle: Patient-specific numerical simulation of the thermal field. Right: 3D representation of the thermal field for better observation.*

as specified by the vendors are displayed around the instrument tip. After instrument placement, a numerical simulation of the ablation zone allows for a patient-specific display of the coagulation necrosis. If the simulation shows an incomplete overlay of the tumor the clinician may look for a better position or adjust the time of the therapy. An illustration of the instrument placement can be seen in Figure 2.

## 2.2 Navigation and Monitoring

After planning the intervention, the instrument has to be placed at the specified target position. Especially if more than one instrument has to be placed in a target structure the risk of excessive thermal coagulation is increasing due to the overlapping ablation zones. To guide the performing radiologist during insertion MRI guidance is used to verify the current position of the instrument inside the human body. The SIEMENS Scanner Remote Control (SRC) interface was used to offers full access to the MR scanner including readout of the image data stream and modifying the present sequences. Our current workflow offers three different image plane positions during navigation showing three different time steps of the current instrument position and the surrounding tissue (see Figure 3). The central image plane is aligned with the current instrument position and orientation to provide a good visibility of the instrument intersection along the main axis. To take care of out of plane angulation a ventral and dorsal slice are also visualized for better spatial perception. The update rate is set to 500ms for each slide independently offering a resolution of 128 x 128 pixel and a pixel spacing of 2.5mm / 2.5mm in X and Y direction. Using the SRC interface we are able to adjust these planes according to the needs of the radiologist at any time during the instrument insertion.

After instrument placement, the ablation of the tumor should be monitored to avoid an insufficient or excessive coagulation necrosis. MR thermometry is a known method to get information about the temperature inside a voxel of the dataset. To achieve this, a phase image is acquired prior to the intervention. From this reference image the following intra-interventional phase images are subtracted during the ablation. This results in a phase difference over the time, which is linear to the temperature of the tissue. To improve the accuracy, several reference images may be acquired and averaged to reduce e.g. motion artifacts. Currently, the image plane position and orientation to acquire the phase images have to be planned very carefully in advance, which is very time consuming and exhaustive. Due to the SRC interface integration we are able to modify the image plane position and orientation automatically during intervention providing a better support for the radiologist by aligning the position and orientation with the instrument. The generated phase difference images are presented as an overlay on top of the corresponding slice in the planning data set (see Figure 4).

## 2.3 Evaluation

The presented workflow prototype was evaluated with the help of four medical experts who are at least familiar with needle-based percutaneous interventions. For evaluation purpose the workflow was simulated using real intervention data from one patient with hepatocellular carcinoma and liver cirrhosis. The patient was previously treated with ablation and transarterial chemoembolisation but showed tumor recurrence at two sites of the liver. The shown recurrence was treated under MRI guidance because of excellent visibility of the tumor and the possibility of a risk structure avoiding trajectory (lung). Regarding the monitoring we did not have the

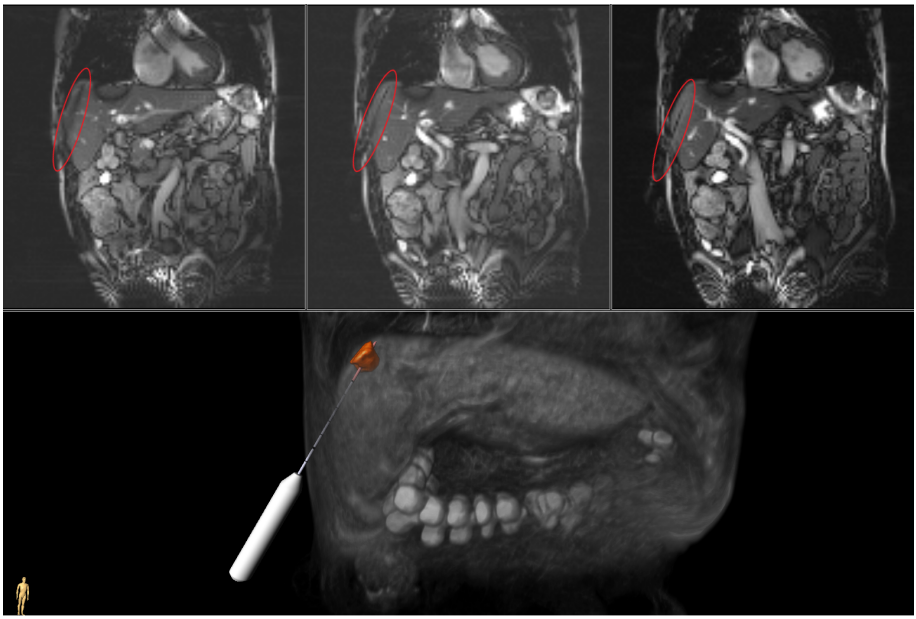


Figure 3: *Mosaic view during the navigation. Red circles indicate instrument position in different time steps. Top left: Pre-slice. Top middle: Image plane corresponding to the instrument orientation and position. Top right: Post-slice. Bottom: 3D-View showing the planning data set to evaluate the instrument position.*

chance to use real thermometry data due to insufficient quality of those. To solve this issue, the thermometry overlay was simulated and approximated to the ablation zone specified by the instrument vendor to offer a whole workflow during evaluation. To reduce the bias caused by the user interface the workflow was presented as a short video showing all steps for planning, navigation and monitoring. In addition, two questionnaires were created and handed out to the experts. The first questionnaire was handed out before the workflow video was shown, asking about the degree of expertise regarding needle-based interventions, the expertise regarding MRI guided interventions in particular and the advantages and disadvantages of the state of the art workflow. In addition, we asked for the rough number of interventions performed in the last year. After filling out the questionnaire, the participants watched the video. Afterwards, they were handed out the second questionnaire asking about the advantages and disadvantages of the proposed workflow, particular wishes for better support and guidance and other comments or suggestions to improve the workflow. All four experts rated their experience regarding needle-based interventions at least as "high" or "expert" as stated in the questionnaire. In average they performed roughly 250 interventions in the past year supporting their stated experience. Regarding MRI guided interventions, in particular, the experience ranged from "low", "medium" and "high" to one participant who stated that he was an "expert" in this field performing roughly 50 interventions in the last year.

## 3 Results

Regarding the current workflow for MRI guided interventions one of the main advantages stated was the absence of contrast agent during planning (visibility of small lesions) and monitoring (good visibility of the ablation zone). The real-time instrument guidance using fast MRI sequences and the absence of radiation were also stated as positive aspects. In addition, the availability of thermometry sequences for ablation monitoring and the free angulation of imaging planes (especially useful for lesions in the hepatic dome) are also two major advantages. Nonetheless, the limited space and therefore limited active interaction by the interventionist and the accessibility of the patient are a big drawback. It was also stated that the manipulation of the imaging

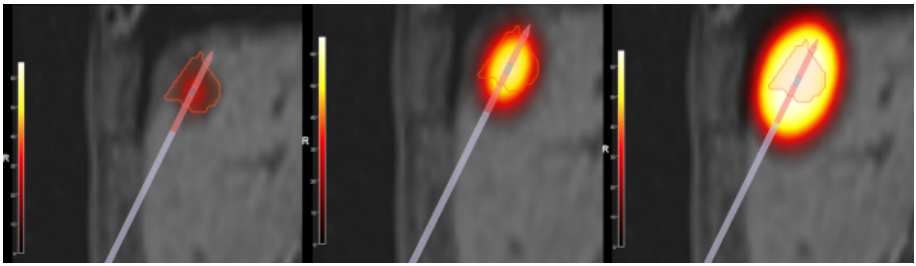


Figure 4: *Different time steps during the simulated thermometry monitoring of the intervention. Left: Beginning of the therapy. Middle: Half way into the therapy. Right: End of the therapy.*

planes only from the host computer is troublesome. Due to this, the performed interventions tend to be very time consuming and require an experienced technician and a very good communication. Furthermore only a few devices are available for use in the MR room.

Our proposed workflow was evaluated regarding possible benefits and downsides as well. First of all, the participants stated that the shown workflow will possibly lead to faster interventions because all intermediate steps are wrapped in one software. In addition, the procedure planning may be easier during path planning. The biggest advantage was said to be the possibility of safer interventions using the thermometry monitoring. The participants assume that the ablation zone can be controlled better leading to a more reliable assessment of the ablation outcome. Two out of the four experts did not see any disadvantages of the proposed workflow whereas the other two where concerned about the respiration motion during instrument placement and the monitoring of the ablation zone. The actual size of the ablation zone depends on several factors (fat content of the liver, previous embolization, fibrosis and others). In addition, it was stated that the time for the data transfer may result in a delay, e.g. during live navigation of the instrument.

## 4 Discussion

As mentioned previously, the time for performing an intervention may be reduced during our workflow, which would be a benefit for both, the performing radiologist and the patient. The all in one workflow offers a faster intervention and a help during planning and instrument placement. Because this was one of the main goals of the proposed workflow this outcome is not surprising. The other aim was to provide a good accessibility for an accurate tumor monitoring. This was as well said to be a very positive factor of the workflow design. Nonetheless, there are still some parts left for better observation. Especially the real-time monitoring using thermometry sequences needs to be treated with caution. Currently, only 2D thermometry images are available in our setup, which may give a good observation of the coagulation necrosis in the current image plane but not with respect to the whole target structure. This problem needs to be addressed in the future. One solution might be a pseudo 3D thermometry by acquiring several 2D phase images rotated along the current instrument axis and interpolating the intermediate image planes. To achieve this, we need to make sure that the image acquisition is sufficiently fast enough. Additionally, expiration motion may cause problems during instrument insertion. We currently assume a breath holding state but it would be more suitable to either track the motion during breathing or integrate deformation models into our software to compensate this movement. Another possibility would be to extract the liver contour in the intra-interventional images and compare them to the corresponding liver intersection in the planning data. In this case we might give the clinician an information about the current breath state of the patient allowing for choosing the perfect moment for instrument insertion. To also guide the radiologist during instrument insertion a fast segmentation of the instrument in the intra-interventional images should be integrated into the workflow. The elongation of the trajectory and the orthogonal visualization along this trajectory were suggestions during the evaluation of the workflow, which may help the radiologists even more during insertion.

Even though there are still many parts to take care of, the proposed workflow was rated positively by the experts including good suggestions for further improvement. In general, the disadvantages seem to be less compared to the current clinical workflow for MRI guided interventions.

## 5 Conclusion

In this work, we introduce a new workflow design for needle-based MRI guided percutaneous liver tumor ablations combining the planning, navigation and monitoring phase in a capsulated workflow. We used the SAFIR toolbox for planning of the intervention in combination with the SIEMENS SRC interface for real-time MR scanner access to manipulate the running sequences during navigation and monitoring.

We evaluated our design with the help of four clinical experts who are highly experienced regarding needle-based interventions in general and moderately experienced with those interventions under MRI guidance. Evaluation shows a strongly positive trend towards our new workflow design in comparison to the current workflow. Nonetheless a few drawbacks were stated, which will be part of our future research.

However this workflow design is robust platform for our future development. This may include the development of suitable instrument detection algorithms, visualization techniques for navigation guidance and a pseudo 3D thermometry approach for better ablation zone assessment.

## 6 Acknowledgement

The work of this paper is partly funded by the Federal Ministry of Education and Research within the Forschungscampus STIMULATE under grant number 13GW0095A.

## 7 References

[1] J. Hof, M. Wertenbroek, P. Peeters, *et al.*, "Outcomes after resection and/or radiofrequency ablation for recurrence after treatment of colorectal liver metastases," *British Journal of Surgery*, vol. 103, no. 8, pp. 1055–1062, 2016.

[2] P. L. Pereira, "Actual role of radiofrequency ablation of liver metastases," *European radiology*, vol. 17, no. 8, pp. 2062–2070, 2007.

[3] C. Rieder, M. Schwier, A. Weihusen, *et al.*, "Visualization of risk structures for interactive planning of image guided radiofrequency ablation of liver tumors," in *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, International Society for Optics and Photonics, vol. 7261, 2009, p. 726 134.

[4] R. Khlebnikov, B. Kainz, J. Muehl, *et al.*, "Crepuscular rays for tumor accessibility planning," *IEEE transactions on visualization and computer graphics*, vol. 17, no. 12, pp. 2163–2172, 2011.

[5] S. Clasen and P. L. Pereira, "Magnetic resonance guidance for radiofrequency ablation of liver tumors," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 2, pp. 421–433, 2008.

[6] M. M. Arnolli, N. C. Hanumara, M. Franken, *et al.*, "An overview of systems for ct-and mri-guided percutaneous needle placement in the thorax and abdomen," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 11, no. 4, pp. 458–475, 2015.

[7] S. Clasen, A. Boss, D. Schmidt, *et al.*, "Magnetic resonance imaging for hepatic radiofrequency ablation," *European journal of radiology*, vol. 59, no. 2, pp. 140–148, 2006.

[8] T. Maeda, J. Hong, K. Konishi, *et al.*, "Tumor ablation therapy of liver cancers with an open magnetic resonance imaging-based navigation system," *Surgical endoscopy*, vol. 23, no. 5, p. 1048, 2009.

[9] E. Rothgang, W. D. Gilson, F. Wacker, *et al.*, "Rapid freehand mr-guided percutaneous needle interventions: An image-based approach to improve workflow and feasibility," *Journal of Magnetic Resonance Imaging*, vol. 37, no. 5, pp. 1202–1212, 2013.

[10] U. Kägebein, O. Speck, F. Wacker, *et al.*, "Motion correction in proton resonance frequency–based thermometry in the liver," *Topics in Magnetic Resonance Imaging*, vol. 27, no. 1, pp. 53–61, 2018.

[11] A. H. Mahnken, A. Buecker, E. Spuentrup, *et al.*, "Mr-guided radiofrequency ablation of hepatic malignancies at 1.5 t: Initial results," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 19, no. 3, pp. 342–348, 2004.

[12] G. Gaffke, B. Gebauer, F. D. Knollmann, *et al.*, "Use of semiflexible applicators for radiofrequency ablation of liver tumors," *Cardiovascular and interventional radiology*, vol. 29, no. 2, pp. 270–275, 2006.

[13] A. Weihusen, F. Ritter, P. L. Pereira, *et al.*, "Towards a workflow oriented software assistance for the radiofrequency ablation.," in *GI Jahrestagung (1)*, 2006, pp. 507–513.

# Intraoperative Data Acquisition through Visual Sensing of Surgical Workflow

*Christian Marzi*[1], *Jörg Raczkowsky*[1], *Franziska Mathis-Ullrich*[1]

[1] *Karlsruhe Institute of Technology, Institute for Anthropomatics and Robotics - Health Robotics and Automation, Karlsruhe, Germany*

Contact: christian.marzi@kit.edu

***Abstract***

*A multi-camera visual sensing setup for intraoperative data acquisition inside the operation room is presented. The system is designed to allow acquisition of visual information during surgery, which can be used to analyze context information for surgical assistance or robotic systems. Based on a previously introduced system for data acquisition, a new modular approach for a completely integrated and portable system is presented. Through additional functionality, such as disinfectability, cooling of enclosed components and integration into surgical workflow, it can be applied in surgical research projects, e.g. for the acquisition of staff movement data.*

**Keywords**: Research Setup, Intraoperative, RGBD-Camera, Calibration, Medical Robotics

## 1 Problem

For research of new medical technologies, acquisition of data from real surgeries is necessary. While working with simulations, models and phantoms can be sufficient in early studies, these modalities can often only provide limited realism. For instance, workflow data in surgery can be studied by simulating a surgery with actors or medical students in a lab environment, which however will always differ from real surgery, especially in the case of adverse events. Leaving the lab and acquiring data during a real surgery however, carries a lot of challenges. Due to simplicity, most research setups are developed and applied in a lab environment. Directly transferring these setups to an operation room (OR) is rarely possible as besides the technical usability also the medical safety and usability needs to be regarded. Technical requirements demand cooling of components, broadband data interfaces, calculation power, and data storage. Additionally, assurance of patient safety is complicated due to issues regarding hygiene, influence on the laminar airflow inside the OR and introduction of foreign particles. Cables and space taken can be a tripping hazard for medical staff and affect the workflow. Additionally, especially multi-camera setups require room and time for calibration procedures, which both are often limited. While lab setups can often be mounted to ceilings or walls, in research in the OR this is not always feasible or too costly, further complicating their use. This paper presents an approach, how an established lab setup can be extended to allow acquisition of RGBD-Streams and derived workflow data in running surgery.

## 2 Material and Methods

This work is based on the research setup OP:Sense [1,2], which has been demonstrated for acquisition and fusion of various surgical data, in particular in combination with surgical robotics. Its sensing system is used to acquire contextual information and track the surgical workflow to provide context sensitivity to surgical robotic systems. To further research and validate workflow analysis, the setup is extended to acquire visual information of the OR during a running surgery. Therefore, the improved sensing system is equipped with an RGBD-imaging setup, capable of recording visual data and track medical staff through four RGBD-cameras covering the whole OR and reducing occlusions. The presented system consists of four Kinect v2 (Microsoft, Redmond, Washington), each using a preprocessing PC to calculate tracking data. Via a 2.5 to 10 GBit LAN these PCs are connected to a central PC fusing and recording the data centrally outside of the OR. The previously developed software framework of OP:Sense uses the Robot Operating System (ROS) as middleware and can be directly applied to this setup. The original calibration method relies on a marker based tracking system. In order to simplify the setup, current work involves the development of a calibration method, which does not rely on a common reference (marker based system) but calibrates the cameras relative to each other using an disinfectable chessboard. To clearly understand the demands and handling of the system, an initial prototype devoid of electrical components is disinfected and placed inside an empty OR. Subsequently, handling and usability is evaluated to improve the finalization of the design. The final setup is equipped with PC, ventilation and sensors, so it can acquire recordings for analysis of the surgical workflow, as well as provide contextual information for robotic surgery.

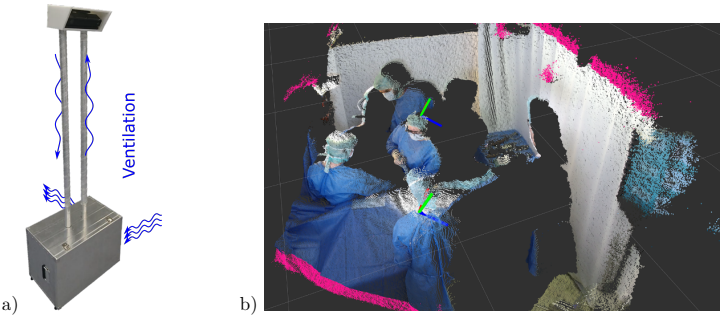



Figure 1: *a) The ready to use mobile system and ventilation flow. b) Exemplary data recorded by the setup. Position and pose of staff can be tracked and recorded.*

## 3 Results and Discussion

Figure 1 a) shows one camera stand of the proposed setup. Aluminum boxes at the base provide stability, room for a preprocessing PC and ventilation control. Height-adjustable feet are installed rather than wheels to provide stability. Additionally, hygiene is increased as the system needs to be carried into the OR, thus no dirt is gathered up during transport.To prevent the introduction of foreign particles, camera and PC need to be enclosed. A ventilation system with the exhaust pointing towards the back controls airflow through PC and camera (see Figure 1 a) ). Thus, as few particles as possible are whirled up and the laminar airflow above the patient is not disturbed. Imaging tests through the front window of the camera box showed mineral glass to be more suitable than acrylic glass due to its better transmission properties for infrared light. Nevertheless, the depth data is affected by increased noise than without glass. A future alternative could be the sealed integration of the camera front in a windowless box. The aluminum surface of the stand is smooth and has successfully been tested to be resistant against common disinfectants. The mechanical prototype proved its usability in the OR, as it was easy to set up. No tools or additional equipment were necessary. Due to its compact design the setup can be handled by a single person. An intuitive and fast calibration method was designed: A disinfectable calibration pattern (chessboard) is printed on aluminum. During the calibration process, it is moved in the common field of view of at least two cameras. A neighboring pair of cameras samples 50 frames with the chessboard visible in each camera's field of view. The relative calibrations matrices are calculated and averaged. An audio feedback marks if enough samples for two cameras have been sampled. Thus, a calibration of all four cameras can be performed fast and without risk of affecting OR hygiene by the need to access the cameras or PCs. Figure 1 b) shows an example of the recorded data.

## 4 Conclusion

The realized setup is now applicable for data acquisition in surgery. The setup was tested and optimized regarding its usability and handling. It is comfortably installed and calibrated, fitting the needs for effective integration in the OR and its processes. For future work this setup will be applied in research on surgical assistance systems and robotics. Further development of this setup may include integration of further sensors, such as marker based tracking or integration to a realtime assistance system as it is able to stream online workflow data of an ongoing surgical procedure.

## 5 References

[1] Nicolai, P., et al *The OP:Sense Surgical Robotics Platform: First Feasibility Studies and Current Research.*, International Journal of Computer Assisted Radiology and Surgery, **8(1)**, 136-137(2013).

[2] Bihlmaier, A., et al *ROS-Based Cognitive Surgical Robotics*, Robot Operating System (ROS) - The Complete Reference, 317-342, 978-3-319-26052-5, Springer, Cham, (2015)

# HeiChole – Heidelberger Cholezystektomie-Datensatz für die chirurgische Workflow- und Skill-Analyse

*A. Kisilenko[1], S. Bodenstedt[2], D. Tran[1], B. Kinny-Köster[1], F. Nickel[1], S. Speidel[2], H.G. Kenngott[1], B.P. Müller[1], M. Wagner[1]*

*[1]Universitätsklinikum Heidelberg, Klinik für Allgemein-, Viszeral- und Transplantationschirurgie, Heidelberg, Deutschland*
*[2]Nationales Centrum für Tumorerkrankungen, Abteilung für Translationale Chirurgische Onkologie, Dresden, Deutschland*

Kontakt: martin.wagner@med.uni-heidelberg.de

**Keywords:** Kognitive Assistenzsysteme, Chirurgische Workflowanalyse, Laparoskopische Cholezystektomie, Chirurgisches Skill Assessment, Open Data

## 1 Problemstellung

Die laparoskopische Chirurgie geht für den Patienten mit zahlreichen Vorteilen einher, jedoch fordert sie von den Chirurgen größeres Geschick und eine längere Einarbeitungszeit. Zudem erschwert der Ärztemangel die kontinuierliche Präsenz von Experten. Zum Zweck der Risikominimierung für die Patienten und Unterstützung des Operationsteams wird daher vermehrt an der Modellierung, Analyse und Automation chirurgischer Workflows, beispielsweise durch kognitive Assistenzroboter, geforscht. Ein signifikantes Problem stellt dabei das Fehlen offen zugänglicher und qualitativ hochwertiger klinischer Datensätze als Grundlage für maschinelle Lernverfahren dar. So ist der Cholec80-Datensatz [1] mit 80 Videos sehr groß, enthält allerdings nur eine Phasen- und Instrumentenannotation. Zur Bewältigung dieser Schwierigkeit, stellen wir einen neuen Datensatz laparoskopischer Cholecystektomien zur Verfügung, der zusätzlich Annotationen chirurgischer Aktivitäten und des chirurgischen Skills enthält. Unser Ziel ist eine offene, standardisierte Validierungsreferenz für maschinelle Lernverfahren zur Verfügung zu stellen.

## 2 Material und Methoden

### 2.1 Datensammlung

Der DatenSatz enthält laparoskopische Cholecystektomien aus drei chirurgischen Zentren. In zwei Zentren wurden je n=15 Videos aufgezeichnet, in einem dritten Zentrum n=3 Videos, insgesamt also n=33 Videos. Zur Anonymisierung der Videos, wurden Szenen außerhalb des Bauchraums, zum Beispiel beim Herausziehen der Kamera zum Zweck ihrer Reinigung, geschwärzt und nicht personenbeziehbare Namen vergeben (Chole-1, Chole-2 etc.).

### 2.2 Datenannotation

In einem zweiten Schritt wurden die Videodaten von medizinischen Experten mithilfe des Annotationsprogramms Anvil [2] annotiert. Dabei wurden zunächst Informationen zum Workflow, also Operationsphasen, Aktivitäten des Chirurgen und Präsenz der Instrumente annotiert.
Für die Operationsphasen wurden analog dem Cholec80-Datensatz jedem Einzelbild eine von sieben folgenden chirurgischen Phasen zugeordnet:

- Vorbereitung
- Dissektion des Calot'schen Dreiecks
- Clipping und Durchtrennung der Gefäße
- Dissektion der Gallenblase (vom Leberbett)
- Bergung der Gallenblase
- Säuberung und Blutstillung
- Retraktion der Gallenblase

Für die Aktivitäten des Chirurgen wurde jeweils die Aktivität, die durchführende Hand des Chirurgen oder Assistenten und das jeweilige durchführende Instrument annotiert (z.B. die linke Hand des

Chirurgen führt die Greif- und Halteaktivität durch mit der Darmfasszange). Folgende Aktivitäten wurden annotiert:

- Greifen
- Halten
- Schneiden
- Clippen

Bei der Innstrumentenpräsenz galt ein Instrument als sichtbar, sobald seine charakteristische Instrumentenspitze im Bild zu sehen war.
Für die Standardisierung und Reproduzierbarkeit der Annotation wurden sowohl für die Phasen-, Aktivitäts- als auch Instrumentenannotation explizite Regeln formuliert.
Zur Bewertung des chirurgischen Skill wurde die Annotation um eine Bewertung mittels Global Operative Assessment of Laparoscopic Skills (GOALS) [3] ergänzt.

Um die Reliabilität der Annotation zu erhöhen, wurden die Operationsphasen von drei medizinischen Experten und der chirurgische Skill von zwei medizinischen Experten unabhängig voneinander annotiert. Sämtliche Abweichungen wurden diskutiert und im Konsens aufgelöst.

2.3 Datenveröffentlichung

Die Daten werden im Rahmen der Endoscopic Vision Challenge 2019 [4] zur Evaluation maschineller Lernverfahren für die chirurgische Workflowerkennung durch unabhängige Forschungsgruppen veröffentlicht. Dabei wurden bereits 24 Videos (je 12 aus Zentrum 1 und 2) als Trainingsdaten veröffentlicht. Die restlichen neun Videos (je drei aus Zentrum 1, 2 und 3) werden als Testdaten genutzt werden.

## 3 Bisherige Ergebnisse

Die 24 Operationsvideos für die Trainingsdatensätze umfassen insgesamt 15,3 Stunden Video mit 184 Phasenübergängen. Jede Phase dauerte durchschnittlich 5,5 Minuten, wobei die längste Phase mit durchschnittlich 17,2 Minuten die „Dissektion des Calot'schen Dreiecks" war. Die zweitlängste Phase war „Dissektion der Gallenblase (vom Leberbett)" mit durchschnittlich 6,7 Minuten. Am längsten war das Instrument „Darmfasszange" zu sehen mit einer durchschnittlichen Dauer von 17,4 Minuten pro Video, am zweitlängsten das Instrument „elektrisches Häkchen" mit 17 Minuten und am kürzesten der „Stapler" mit 0,03 Minuten. Die Greifaktivität fand insgesamt 2027 Mal statt, wobei circa 474 Mal keine Halteaktivität folgte.

## 4 Diskussion

Der vorgelegte Datensatz ermöglicht die Validierung von Algorithmen der Erkennung chirurgischer Aktivitäten wie „Greifen" und „Halten" in klinischen Videos. Zusätzlich können die Workflowinformationen mit der Annotation des chirurgischen Skill verknüpft werden.
Hierbei ist es wichtig, standardisiert und reproduzierbar zu annotieren, weshalb Annotationsregeln festgelegt wurden, die mit der Veröffentlichung des Datensatzes ebenfalls veröffentlicht werden. Zudem wurde die Annotation von Operationsphase und chirurgischem Skill von mehreren medizinischen Experten durchgeführt, um die Inter-Observer-Variabilität zu untersuchen.

## Referenzen

[1] Twinanda, A. P., Shehata, S., Mutter, D., Marescaux, J., de Mathelin, M. und Padoy, N. (2017). EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos. IEEE Trans Med Imaging *36*, 86-97, doi: 10.1109/tmi.2016.2593957.

[2] Anvil – The Video Annotation Research Tool – Home, URL: http://www.anvil-software.org/# [as of May 31, 2019]

[3] Vassiliou, M. C., Feldman, L. S., Andrew, C. G., Bergman, S., Leffondre, K., Stanbridge, D. und Fried, G. M. (2005). A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg *190*, 107-113, doi: 10.1016/j.amjsurg.2005.04.004.

[4] EndoVisSub2019-Workflow - Home (o. J.)., URL: https://endovissub-workflowandskill.grand-challenge.org/EndoVisSub-WorkflowAndSkill/ [as of May 31, 2019]

# Session of Excellence

# Towards Deep Learning-Based EEG Electrode Detection Using Automatically Generated Labels

*N. Gessert[1], M. Gromniak[1], M. Bengs[1], L. Matthäus[2], A. Schlaefer[1]*

[1] *Institute of Medical Technology, Hamburg University of Technology, Hamburg, Germany*
[2] *eemagine Medical Imaging Solutions GmbH, Berlin, Germany*

Contact: nils.gessert@tuhh.de

***Abstract***

*Electroencephalography (EEG) allows for source measurement of electrical brain activity. Particularly for inverse localization, the electrode positions on the scalp need to be known. Often, systems such as optical digitizing scanners are used for accurate localization with a stylus. However, the approach is time-consuming as each electrode needs to be scanned manually and the scanning systems are expensive. We propose using an RGBD camera to directly track electrodes in the images using deep learning methods. Studying and evaluating deep learning methods requires large amounts of labeled data. To overcome the time-consuming data annotation, we generate a large number of ground-truth labels using a robotic setup. We demonstrate that deep learning-based electrode detection is feasible with a mean absolute error of* $5.69 \pm 6.10\,\mathrm{mm}$ *and that our annotation scheme provides a useful environment for studying deep learning methods for electrode detection.*

**Keywords**: Deep Learning, CNN, Electrode Detection, Generated Labels

## 1 Problem

Electroencephalography (EEG) is a method that allows for measuring electrical brain activity, e.g., to assess patients' motor function impairment or monitor progress in patients' recovery process [1]. For accurate brain current estimation based on the measured signals on the scalp, knowledge of the electrodes' location is required [2].

A typical method for electrode placement is the 10-20 system [3] or its refined variants [4] where the positions are determined based on anatomical landmarks. Identification of anatomical landmarks relies on visual inspection and palpation by the practitioner which is error-prone. Instead, using accurate localization systems, e.g., using optical digitizing scanners with a stylus [5] or MRI-based localization have been proposed [6]. Often, these systems are expensive and recording all electrodes' location is time-consuming. Therefore, photogrammetric methods have been proposed where a single [7] or multiple cameras [8] are used to localize the electrodes on the head. These methods are advantageous as cheap cameras can be used for accurate localization. Recent methods often rely on depth (time-of-flight) and/or multiple RGB images for reconstruction of the 3D electrode positions [9].

Previous photogrammetric methods come with two major drawbacks. First, computer vision techniques for 3D reconstruction and electrode detection rely on handcrafted features and algorithms which are often limited to the specific scenarios they were engineered for and the algorithms often come with long execution times. Second, previous approaches usually assume a fixed head location. Both hinder application in mobile and changing environments such as ambulances where head movement is inevitable and fast detection is needed. Thus, fast algorithms that deal with large head pose variation are required. In recent years, deep learning methods have shown remarkable performance for a variety of computer vision tasks such as real-time object detection [10] and head pose estimation [11]. In this paper we study the feasibility of electrode detection using convolutional neural networks (CNNs). To facilitate and study this approach, large amounts of annotated data are required. Therefore, we propose a setup using a robot with a head phantom attached to the robot's endeffector. An EEG electrode cap is placed on the head phantom. Then, an RGBD camera acquires images of the head phantom which is moved to different positions and orientations. The electrodes are first labeled in a single image. Then, the electrode locations are transformed to each head pose using a hand-eye calibration and the initial markings. We study whether these automatically generated labels can be learned by a CNN which directly predicts the electrode locations from the images.

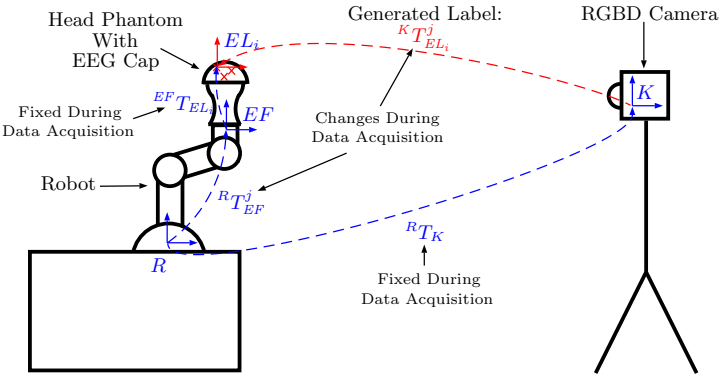



Figure 1: *The experimental setup we use for data acquisition and label generation. Red crosses represent electrodes on the EEG cap.*

# 2 Material and Methods

## 2.1 Experimental Setup

Our experimental setup for data acquisition is shown in Figure 1. First, we perform a hand-eye calibration between the robot (UR3, Universal Robots) and the camera (Kinect V2, Microsoft) using a checkerboard mounted to the robot. Camera poses of the checkerboard are obtained with OpenCV [12] and the calibration transformations are obtained with QR24 [13] using 50 robot and camera poses. Then, the head phantom wearing the electrode cap (waveguard touch, eemagine) is mounted to the robot. We now move the robot into different endeffector positions and orientations while continuously acquiring RGB and depth images while also logging the endeffector poses.

## 2.2 Automatic Data Annotation

After acquiring a set of images, we map the RGB images to the depth sensor's coordinate frame using calibrations provided by the manufacturer. The RGB images now have the same resolution as the depth images ($524 \times 424$). Next, we annotate all electrodes in a single image. As the RGB image was transformed to the depth sensor's coordinate frame, we can now obtain the 3D coordinates from the 2D depth image and its corresponding point cloud using the 2D pixel locations from the RGB image. Using this 3D position and assuming identity orientation we obtain the poses ${}^{K}T_{EL_i}$ of all $N$ electrodes $i \in N$. Using the hand-eye calibration between robot and camera ${}^{R}T_{K}$ and the current endeffector pose ${}^{R}T_{EF}$ we obtain the electrode pose with respect to the robot endeffector:

$$ {}^{EF}T_{EL_i} = ({}^{R}T_{EF})^{-1}\, {}^{R}T_{K}\, {}^{K}T_{EL_i} \tag{1} $$

Next, we can automatically obtain the electrode poses for all other robot poses $j$ in the dataset:

$$ {}^{K}T^{j}_{EL_i} = ({}^{R}T_{K})^{-1}\, {}^{R}T^{j}_{EF}\, {}^{EF}T_{EL_i} \tag{2} $$

The position of the pose ${}^{K}T^{j}_{EL_i}$ is now used as a 3D label for each electrode position for each RGBD image in the dataset. Besides the 3D labels, we also consider image-level (pixel) labels by projecting the 3D points back into the RGB images. Here, we perform nearest-neighbor matching, i.e., we assign the electrodes' transformed 3D location to the closest point in the point cloud. Then, we project this point back on the RGB image. While the 3D labels are ultimately used for EEG, the pixel labels can be useful for purely image-based algorithms. The automatically generated labels will likely be affected by calibration errors. Thus, we also compare to a more accurate ground-truth by manually labeling a small set of images. To ensure consistent labels, the annotator selects the center pixel of each electrode.

| | | Generated Labels | | | Manual Labels | | |
|---|---|---|---|---|---|---|---|
| | | MAE | rMAE ($10^{-3}$) | aCC | MAE | rMAE ($10^{-3}$) | aCC |
| 2D Labels | Gen. Labels | - | - | - | $1.17 \pm 0.99$ | $38 \pm 2$ | 0.999 |
| | DN RGB | $\mathbf{0.71 \pm 0.59}$ | $25 \pm 2$ | 0.999 | $\mathbf{1.30 \pm 1.5}$ | $43 \pm 5$ | 0.998 |
| | DN RGBD | $0.73 \pm 0.61$ | $25 \pm 2$ | 0.999 | $1.31 \pm 1.5$ | $43 \pm 5$ | 0.998 |
| | DN D | $1.50 \pm 1.3$ | $53 \pm 5$ | 0.998 | $1.91 \pm 1.8$ | $64 \pm 6$ | 0.997 |
| | SR RGB | $0.72 \pm 0.64$ | $25 \pm 2$ | 0.999 | $1.31 \pm 1.5$ | $44 \pm 5$ | 0.998 |
| | SR RGBD | $0.72 \pm 0.65$ | $25 \pm 2$ | 0.999 | $1.32 \pm 1.5$ | $44 \pm 5$ | 0.998 |
| | SR D | $1.56 \pm 1.3$ | $55 \pm 5$ | 0.998 | $2.00 \pm 1.8$ | $67 \pm 6$ | 0.997 |
| 3D Labels | Gen. Labels | - | - | - | $4.76 \pm 5.1$ | $51 \pm 2$ | 0.985 |
| | DN RGB | $3.54 \pm 2.9$ | $62 \pm 5$ | 0.997 | $6.24 \pm 6.4$ | $131 \pm 19$ | 0.981 |
| | DN RGBD | $3.82 \pm 3.3$ | $65 \pm 6$ | 0.996 | $6.35 \pm 6.3$ | $132 \pm 19$ | 0.982 |
| | DN D | $5.43 \pm 4.5$ | $92 \pm 8$ | 0.994 | $7.24 \pm 6.6$ | $146 \pm 20$ | 0.979 |
| | SR RGB | $3.19 \pm 2.7$ | $56 \pm 5$ | 0.997 | $6.00 \pm 6.1$ | $127 \pm 19$ | 0.983 |
| | SR RGBD | $\mathbf{3.13 \pm 2.7}$ | $56 \pm 5$ | 0.997 | $\mathbf{5.69 \pm 6.1}$ | $123 \pm 19$ | 0.983 |
| | SR D | $5.43 \pm 4.8$ | $92 \pm 8$ | 0.994 | $7.18 \pm 6.7$ | $143 \pm 19$ | 0.980 |

Table 1: Results for all experiments with Densenet121 (DN) and SE-Resnext50 (SR). We consider the MAE, the rMAE, and aCC. For 3D poses, the MAE is given in mm, for 2D poses, the MAE is given in pixels. rMAE and aCC are relative metrics without unit. Generated labels refer to our automatic annotations with our robotic setup. Manual labels refer to manual annotation by a human.

### 2.3 Deep Learning Models

We employ two state-of-the-art CNNs, Densenet121 [14] and SE-Resnext50 [15]. The input to the network is an RGB image, the depth image or a full RGBD image. The images are cropped to the relevant region around the robot workspace based on the extent of the ground-truth annotations. Including a margin, this results in a network input size of $270 \times 254$ pixels. All models are pretrained on ImageNet to overcome relatively small dataset sizes. Using an EEG cap with $N = 8$ electrodes, the model output is of size $N \times 3$ for 3D point prediction and $N \times 2$ for 2D pixel location prediction. As we solve a regression problem, the loss is the mean squared error, minimized using the Adam algorithm with an initial learning rate of $l_r = 10^{-5}$ and a batch size of $b = 10$. We train for 300 epochs and halve the learning after 50 epochs each. For implementation we use PyTorch [16]. Training, evaluation and inference time measurement is performed on an NVIDIA GTX1080 TI.

In terms of evaluation metrics we follow [17] and use the mean absolute error (MAE) as an absolute metric, either in pixels or mm for 2D and 3D positions, respectively. To compare 2D and 3D labels, we consider the relative MAE (rMAE) which is the absolute error divided by the targets' standard deviation. The metric does not have a unit as it is relative. Last, we consider the average correlation coefficient (aCC) between predictions and targets as a relative metrics. Values close to 1 indicate that a regression task was generally learned well.

Note that our fixed-size CNN output always forces the CNN to make a prediction for all electrode locations, even when they are not visible. Also, our automatic labeling strategy can provide annotations for learning even if some electrode locations are not visible as they are still transformed to their corresponding 3D location. Thus, our model is given the capability to obtain robustness towards partial electrode occlusion.

## 3 Results

To evaluate our setup we generate a set of 3000 images for training and validation and we manually annotate 150 images for testing. The positions cover a range of 500 mm × 500 mm × 300 mm which corresponds to $200 \times 200$ pixels in 2D images.

Using 40 pairs of poses for calibration and 10 pairs for evaluation, the hand-eye calibration between the robot and the camera results in a position error of $3.910 \pm 0.007$ mm and a rotation error of $0.780 \pm 0.004°$.

Quantitative results are shown in Table 1. We provide errors with respect to both the generated labels and the manually annotated labels. In general, our automatically generated labels are close to the real labels. Comparing Densenet121 and SE-Resnext50, both models perform similar with Densenet121 showing the best performance on 2D labels and SE-Resnext50 showing the best performance on 3D labels. With respect to color channels, using RGB and RGBD images performs similar. Both our pixel labels and the real-world 3D coordinates are learned well by the CNNs with aCCs close to 1. For pixel labels, the relative metrics indicate

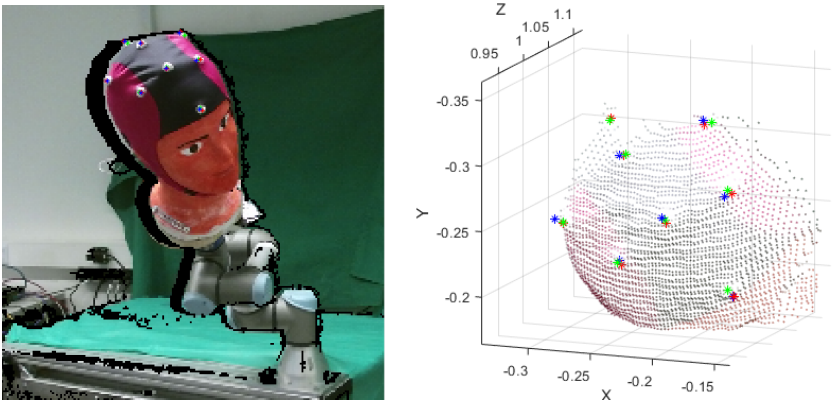



Figure 2: *Left, a cropped RGB image with 2D electrode locations is shown. The black pixels are artifacts caused by the transformation of the RGB image to the depth sensor coordinate frame. Right, a point cloud with a zoom on the EEG cap with 3D electrode locations is shown. Green marks indicate the manually annotated electrode locations. Red marks indicate the automatically generated labels. Blue marks indicate the model predictions. Note that the 2D and 3D locations were predicted by different models.*

a higher performance than for real-world 3D coordinates. Inference times are 34 ms for Densenet121 and 58 ms for SE-Resnext50. Training times are 3 h and 5 h for Densenet121 and SE-Resnext50, respectively.

Qualitative results are shown in Figure 2. We show an RGB image and a point cloud of the head with the EEG cap and the manually annotated electrode locations, the generated labels and the predicted locations. Qualitatively, the predicted electrode locations are close to the manually annotated labels. Also, note that our approach is able to provide a reasonable prediction although one of the electrodes at the back of the head is only partially visible.

## 4 Discussion

In this paper we address deep learning-based electrode detection using 2D camera images. This approach is particular promising as CNNs can provide fast predictions and they can be adjusted to versatile environments without requiring manual feature handcrafting. However, their main drawback is the large amount of annotated data that is usually required. We address this issue with a robotic setup for automatic data and label acquisition. We evaluate the approach by using different types of input images, labels and CNN architectures.

In general, our automatic label generation framework works well although the setup is affected by calibration errors between the robot and the camera. The generated labels closely match the more accurate manual ground-truth with an MAE of $4.76 \pm 5.10$ mm and $1.17 \pm 0.99$ pixels while our labels cover a range of approximately 500 mm × 500 mm × 300 mm and 200 × 200 pixels. Also the point cloud plots in Figure 2 demonstrate that the actual electrode locations are well matched. Overall, the deep learning models approximate the ground-truth well, although there is a large variation in the target positions. Notably, there is a performance difference between using the generated and the manual labels for evaluation. This reflects the calibration errors in the setup which mainly cause the difference between generated and manual labels. Thus, the error between the generated labels and the manual labels can be seen as an upper bound for model performance.

At the same time, predictions are fast with a range of 34 ms to 58 ms which indicates real-time capability. Other photogrammetric methods typically require seconds up to minutes for detection [9].

Using either pixel or real-world 3D coordinates works well while predicting 2D labels appears to be easier with an average aCC of 0.998 compared to 0.981. Intuitively, deriving 3D coordinates from a 2D image is more difficult and thus the results match expectations. In terms of application, the 3D coordinates are more relevant as the overall goal is to obtain the electrode locations with respect to a 3D head coordinate frame. Adding head coordinates for deriving a head coordinate frame is straight forward with our approach and could be addressed in future work.

In terms of CNN models, the performance with respect to the actual labels is very similar as both models achieve aCCs close to 1. SE-Resnext50 performs slightly better with respect to the 3D labels while Densenet121 shows the best performance for 2D pixel labels. Notably, the task of predicting 3D real-world coordinates appears to be more difficult as the rMAE and aCC are generally lower for this task. Considering that SE-Resnext50 has more parameters than Densenet121, the additional capacity might be beneficial for solving the more difficult problem. However, the slight increase in performance is bought with a substantial increase in inference time which needs to be carefully traded off for application.

For the different types of input modalities the performance with respect to the manual labels is very similar. Adding the depth channel to the RGB images does not appear to be beneficial in our setup. This appears to indicate that depth information is not helpful, however, our current setup utilizes a particular head shape and EEG cap model. When generalizing to different head sizes, shapes or EEG caps, depth information will be more important as it is difficult to differentiate between a smaller head, close to the camera and a larger head, further away. Our automatic acquisition and labeling approach is well suited for covering more diverse scenarios, e.g., with different head models which can be addressed in future work.

Overall, we demonstrate that predicting electrode locations works well and enables studying deep learning methods for EEG electrode detection. The setup allows for considering variations such as head shape and size or different EEG caps. However, a clear drawback of our method is the fact that it is limited to the use of head phantoms for automatic label generation. Thus, future work needs to incorporate our approach in real-world settings. This could be facilitated by applying a pre-segmentation of the electrode cap which would make the approach independent of the underlying head. Another approach would be to employ transfer learning techniques [18] or few-shot learning [19] where a CNN that is pretrained with our setup is adapted to the real-world scenarios with a few new images.

## 5 Conclusion

We study deep learning-based EEG electrode detection from camera images. To facilitate the approach, we propose a robotic setup for automatic data and label generation. This allows for quick generation of arbitrarily-sized datasets including ground-truth annotations. We demonstrate that CNNs are able to detect the electrodes using either RGB or depth images. Furthermore, our automatically generated labels closely match a more accurate, manual ground-truth annotation. Thus, our setup allows for developing and studying deep learning-based electrode detection approaches. Future work could study more extended scenarios, e.g., with different head shapes and sizes.

## 6 Acknowledgements

This work was partially funded by AiF research grant number ZF4026302CR7.

## References

[1] Otten, P., Kim, J., Son, S. (2015) A framework to automate assessment of upper-limb motor function impairment: A feasibility study. Sensors **15**(8), 20,097–20,114

[2] Plummer, C., Harvey, A.S., Cook, M. (2008) Eeg source localization in focal epilepsy: where are we now? Epilepsia **49**(2), 201–218

[3] Jasper, H. (1958) Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr Clin Neurophysiol **10**, 370–375

[4] Jurcak, V., Tsuzuki, D., Dan, I. (2007) 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. Neuroimage **34**(4), 1600–1611

[5] Towle, V.L., Bolaños, J., Suarez, D., Tan, K., Grzeszczuk, R., Levin, D.N., Cakmur, R., Frank, S.A., Spire, J.P. (1993) The spatial location of eeg electrodes: locating the best-fitting sphere relative to cortical anatomy. Electroencephalography and clinical neurophysiology **86**(1), 1–6

[6] Brinkmann, B.H., O'Brien, T.J., Dresner, M.A., Lagerlund, T.D., Sharbrough, F.W., Robb, R.A. (1998) Scalp-recorded eeg localization in mri volume data. Brain topography **10**(4), 245–253

[7] Qian, S., Sheng, Y. (2011) A single camera photogrammetry system for multi-angle fast localization of eeg electrodes. Annals of biomedical engineering **39**(11), 2844

[8] Reis, P.M., Lochmann, M. (2015) Using a motion capture system for spatial localization of eeg electrodes. Frontiers in neuroscience **9**, 130

[9] Clausner, T., Dalal, S.S., Crespo-García, M. (2017) Photogrammetry-based head digitization for rapid and accurate localization of eeg electrodes and meg fiducial markers using a single digital slr camera. Frontiers in neuroscience **11**, 264

[10] Ren, S., He, K., Girshick, R., Sun, J. (2015) Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems, pp. 91–99

[11] Ranjan, R., Patel, V.M., Chellappa, R. (2019) Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **41**(1), 121–135

[12] Bradski, G. (2000) The OpenCV Library. Dr. Dobb's Journal of Software Tools

[13] Ernst, F., Richter, L., Matthäus, L., Martens, V., Bruder, R., Schlaefer, A., Schweikard, A. (2012) Non-orthogonal tool/flange and robot/world calibration. The International Journal of Medical Robotics and Computer Assisted Surgery **8**(4), 407–420

[14] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708

[15] Hu, J., Shen, L., Sun, G. (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132–7141

[16] Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A. (2017) Automatic differentiation in PyTorch. In: NIPS Autodiff Workshop

[17] Borchani, H., Varando, G., Bielza, C., Larrañaga, P. (2015) A survey on multi-output regression. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **5**(5), 216–233

[18] Oquab, M., Bottou, L., Laptev, I., Sivic, J. (2014) Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1717–1724

[19] Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., Hospedales, T.M. (2018) Learning to compare: Relation network for few-shot learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1199–1208

# Localization and Classification of Teeth in Cone Beam CT using Convolutional Neural Networks

*M. Neumann*[1], *O. Hellwich*[2], *S. Zachow*[1]

[1] *Zuse Institute Berlin (ZIB), Germany*
[2] *Berlin University of Technology (TUB), Germany*

Contact: {mario.neumann, zachow}@zib.de, olaf.hellwich@tu-berlin.de

***Abstract***

*In dentistry, software-based medical image analysis and visualization provide efficient and accurate diagnostic and therapy planning capabilities. We present an approach for the automatic recognition of tooth types and positions in digital volume tomography (DVT). By using deep learning techniques in combination with dimension reduction through non-planar reformatting of the jaw anatomy, DVT data can be efficiently processed and teeth reliably recognized and classified, even in the presence of imaging artefacts, missing or dislocated teeth. We evaluated our approach, which is based on 2D Convolutional Neural Networks (CNNs), on 118 manually annotated cases of clinical DVT datasets. Our proposed method correctly classifies teeth with an accuracy of 94% within a limit of 2 mm distance to ground truth labels.*

**Keywords**: Dental Imaging, Tooth Classification, Dimension Reduction, Image Reformatting

## 1 Motivation

Determining the dental status is a common task in dentistry, oral surgery, and even in forensics. A continuous monitoring of a patient's dental status is typically reflected in a patient's dental record. For a standardized classification, unambiguous tooth classification schemes have been established, like the FDI two-digit notation used by the Fédération Dentaire Internationale (FDI) and the World Health Organization [1]. However, the process of tooth classification still is a manual respectively a visual task, which is time consuming and in certain cases even not that clear. With medical imaging techniques, a dental state can also be derived from image data, like panoramic X-rays, Digital Volume Tomography (DVT) respectively dental Cone Beam Computed Tomography (CBCT), or even intraoral optical scans. A fully automated and reliable tooth classification and localization based on medical image data, thus is a desired tool for trustworthy and efficient diagnosis and treatment planning. On the one hand, dental practitioners would benefit from an automatic classification of the teeth since this can speed up the documentation process and reduce potential human errors. On the other hand, it could be used to evaluate tooth positions and relative distances on a larger scale (i.e. cohort studies) to provide statistical data on dentition that might be useful for orthodontics. Moreover, automatically predicted locations of correctly classified teeth or tooth types are a prerequisite for more elaborate image analysis tasks.

Detection of teeth in medical images has already been approached in the form of segmentation or identification and classification in various kinds of medical image data, like panoramic X-rays [2], periapical radiographs [3], DVT [4, 5], or regular CT [6]. Gao et al. [7] segmented individual teeth in CT data using a level set based approach, however, without any tooth classification. Hosntalab et al. [8] segmented teeth in CT and used wavelet-Fourier descriptor features for a CNN based classification of tooth types (incisors, canines, premolars, and molars). The ordering of teeth and therefore the tooth numbers according to the FDI tooth classification scheme were analyzed by Nguyen et al. [9] by deploying a graph based algorithm to separate the complete dentition region into a chain of separate tooth regions that are passed to an SVM classifier. However, these methods struggle when dealing with imaging artefacts or variety in dentition.

Recent methods employing deep learning techniques became popular. Learning methods can either be applied to panoramic X-rays or 3D tomograms. Miki et al. [4] used manual crops of individual teeth with a CNN classifier to classify teeth in axial CT slices. In a multi-class segmentation approach of Ezhov et al. [5], information on tooth position and tooth number is given. Although this segmentation approach already provides very useful information it is computationally demanding due to the chosen 3D convolutions. In addition, such a volumetric approach requires a large amount of labeled training data that has to be generated. With our approach we try to tackle this task in a more efficient way by extracting the essential information from the volumetric DVT data and applying 2D CNNs to it.

## 2 Materials and Methods

First own results on automated localization and classification of teeth in medical image data (DVT, CBCT) already yielded promising results [9]. However, the chosen approach revealed problems with missing and/or dislocated teeth and imaging artifacts due to metal shielding caused by dental fillings and metal crowns. In the jawbone itself a clear classification of the dental roots is not possible and the dental crowns often cannot be reliably reconstructed from tomographic data. Implants (bone-anchored teeth or dental bridges) lead to image degradation (cf. Fig. 1, left) complicating image analysis. In addition, there is a very high variability in the dentition (Fig. 1, right), which is difficult to describe using an explicit model.

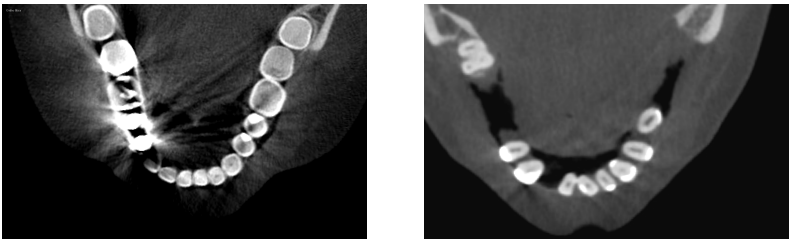


Figure 1: Exemplary axial DVT images: Metal fillings (left) and missing teeth (right) are included in our study.

Our hypothesis is that machine learning allows a more robust, fully automated classification of the dentition status from tomographic image data than explicit model-guided algorithms can achieve. Instead of training algorithms on a complete volume of the dentition region, convolutional neural networks (CNNs) will be trained on reformatted 2D image data of the dentition region that are generated from pre-segmented tomograms. In this way, larger amounts of training data can be processed while keeping the memory requirements manageable.

For the development of our method, 158 data sets were selected from a database of more than 3,000 dental CBCTs ($\leq 0.3$ mm voxel size) based on the anatomical structure that is covered by the field of view, the scan quality, and the presence of implants. Images with very low quality as well as images showing highly pathological cases were discarded. On the basis of the selected data our first study has been carried out.

### 2.1 2D Images of the Alveolar Ridge for CNN Training

The method presented herein uses our automated procedures to segment the jawbone from tomographic data (DVT, CBCT) [10, 11, 12]. The teeth are not included in the Statistical Shape Model (SSM), which is used for model-based segmentation, as it is difficult to statistically describe and geometrically represent a highly varying dentition. Instead, the bone surface that is penetrated by the teeth (the alveolar ridge) is segmented and closed. The dentition region can then be directly inferred from the SSM right after segmentation (Fig. 2, left). The tomographic image data are mapped onto the surface to visualize the tooth interfaces (Fig. 2, center). This representation already allows a clear visual identification of all teeth in a single view. In a subsequent step, the dentition region is slightly enlarged and extracted from the reconstructed bone surface in order to become independently examinable (Fig. 2, right). We will call this visualization an Alveolar Ridge Image (ARI).

Figure 2: Left: Segmentation of the mandibular bone using a statistical shape model, which is decomposed by construction into semantic patches (dentition region in green). Right: Mapping of the tomographic image data onto the surface and extraction of the dentition patch.

The assumption is, that the extracted region (Fig. 2, right) is much less susceptible to disturbances by metal fillings, fixed braces or the like and that we can make a reliable classification based on the image data around this region. The 2D image information extracted in this way (i.e. the ARI) is then made available for training neural networks by projecting the curved surface consistently and as distortion-free as possible into a plane [13]. The projection is performed according to Wang et al. [14] operating on the given triangle mesh of the dentition patch. It minimizes the distortion of the triangles under projection and thus also the distortion of the image data itself (Fig. 3, left). In result, we obtain a flattened Alveolar Ridge Image (fARI).

To further reduce the image size and to minimize the background, a subsequent reformatting of the fARI is performed. This reformatting is equivalent to straightening the dental arch. It is achieved by applying piecewise rotations on mini patches along the centerline of the fARI ((Fig. 3, top right). For each mini patch, the rotation is given by the rotation of the closest segment of the center line into the horizontal axis. The entire reformatting arranges all teeth in a linear pattern. The image is finally sampled into a rectangular image of $1024 \times 256$ pixels and the background is padded with black pixels (Fig. 4). The resulting image will be called straightened Alveolar Ridge Image (sARI) in the remainder of this paper.

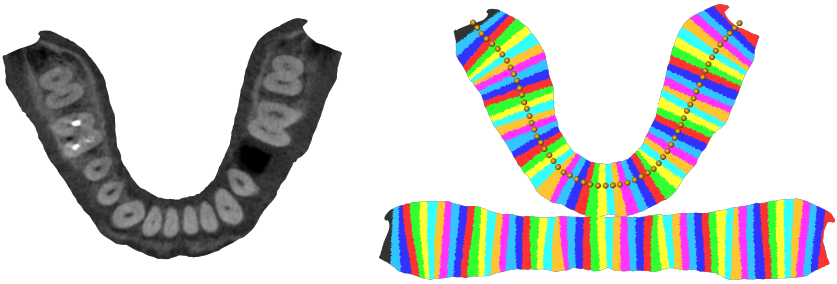


Figure 3: The extracted dentition region is flattened with minimal distortion, thus preserving the appearance of the teeth within the resulting fARI (left). The flattened dentition region is then again straightened with minimal distortion (right) to obtain an sARI (cf. Fig. 4).

Figure 4: Straightened Alveolar Ridge Image (sARI) of the lower jaw.

In addition to the sARI, an Artificial Panoramic Radiogram (APR) is generated from the pre-segmented tomographic data to enrich the image information and hence the training data (Fig. 5). The APR is generated by shifting a line segment perpendicular to the dentition patch along its centerline, hence spanning a so called ruled surface (Fig.5, left). The tomographic data around that surface are integrated, normalized and projected onto it in normal direction. The textured ruled surface is finally flattened yielding the APR (Fig.5, right).

Both, the curved ruled surface and the ARI intersect in 3D space, thus having image information in common. In principle, any number of images can be generated that way by rotating the APR projection around the dentition region to increase the number of training images. The entire reformatting and image generation process is fully automated and has been implemented in Amira ZIB-Edition such that more pre-segmented image data can be easily included in the training data and the process can be consistently applied to new image data. Due to the fact that all reformatting can be inverted in a lossless manner a consistent relationship between 3D and 2D coordinates is assured. That way CNNs can be trained with 2D image data, while being able to transfer classification and localization results back into 3D space. APR and sARI are finally resampled to $64 \times 256$ pixels serving as input for training CNNs.

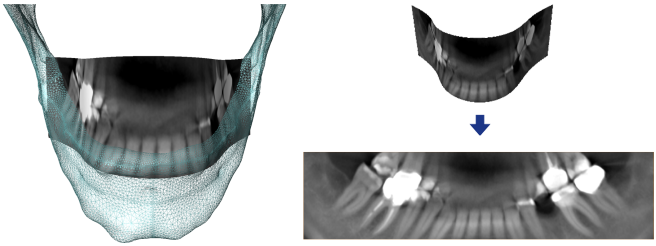


Figure 5: Generation of an Artificial Panoramic Radiogram (APR) for the mandibular dentition.

## 2.2 CNN based Localization and Classification of Teeth

Since we were interested in both, localization *and* classification of teeth (in particular, the classification of individual teeth and not just tooth types), we investigated two methods taking both, sARI and APR as input. This allows us to also consider interdependencies between teeth.

**Coordinate Regression:** In a first investigation, a simple 2D CNN with $16 \times 3$ output channels was employed. The channels correspond to the X and Y coordinates in the normalized sARI extended by a Boolean flag indicating existing or missing teeth. The network was trained by optimizing the L1 and L2 distance over all true positive tooth entries in the corresponding rows combined with a binary probability loss of the existence flag. Despite optimization of the network parameters and different loss functions the localization accuracy was not really satisfying. This could eventually be improved by iteratively cropping the image around the prediction and then refining the location on the cropped image. However, since far outliers would not contain the tooth in the cropped image anymore, we instead investigated a heatmap approach for better localization accuracy.

**Heatmap Regression:** There are two advantages of using heatmap instead of direct coordinate regression. (1) Heatmaps are a solution when dealing with multiple or missing instances. The number of objects per class is not limited by the architecture since it is possible to consider multiple locations in one heatmap. Any number of labels can be provided within a heatmap acting as a ground truth heatmap with the labels representing the probability of an object at the respective location [15]. (2) The network actually learns an image to image mapping that is potentially less complex and even more intuitive, because the spatial arrangement of landmarks within the heatmap directly correspond to the input image. To create the heatmaps, we first generated empty images $h_{n,c}$ for every class $c \in (1, ..., C)$ and all cases $n \in (1, ..., N)$, where N is the number of cases and C the granularity used for the classification problem, i.e. 16 for distinguishing all 16 teeth individually. The U-Net architecture has been utilized to regress the heatmaps in the same resolution as the input images. We used the original U-Net architecture [16] extended by batch normalization [17] after each 2D convolution and used zero padded convolutions to keep the heatmap size equal to the size of the training images. The network can now be optimized using a pixel-wise cross-entropy loss function. At the deepest convolution stage, we employed a wider convolution filter to explicitly enable the network to not only recognize the local appearance of the tooth-bone interfaces but to evaluate the high level features regarding the neighbourhood information of the teeth. The kernel is constructed, such that the height is untouched but the width along the tooth row covers half of the width of the feature map resulting in a $3 \times 16$ convolution filter.

**3D Back Projection:** Since the 2D images are generated by intensities that are transported via the surface mesh, we can determine each predicted position within the respective triangle and calculate its barycentric coordinates. The 3D position is given by applying the barycentric coordinate to the corresponding unflattened triangle in the 3D segmentation of the dentition region.

## 2.3 Data Augmentation

In order to increase the amount of training data, we performed some simple data augmentation steps. For instance, we mirrored all images on the Y-axes such that the order of the dentition was flipped. A rotation of the shape of a single tooth in the images would have been desirable but was not easily possible because we did not have the segmentation of the tooth shape and rotating an estimated bounding box would heavily introduce artefacts and destroy the general appearance of an image. Instead, we rotated the whole image by randomly selected angles within $[-5, 5]$ degrees.

A further augmentation addresses the combinatorial problem of possibly missing teeth. Therefore, we

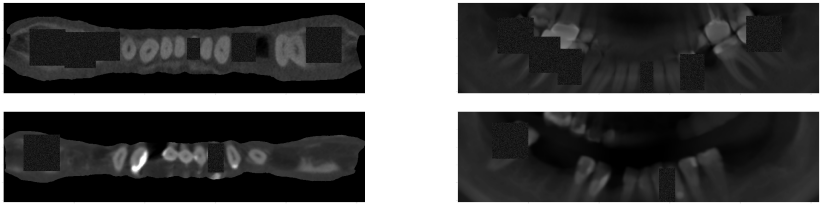


Figure 6: Data augmentation by image rotations and artificial removal of teeth for both, sARI and APR.

artificially mimiced missing teeth within the complete dental arch in a very simple manner. How many and which teeth were removed was in each case defined by a randomly generated binary mask for all 16 positions. At each position of the designated tooth that was to be removed, a dark rectangular pattern was placed with its center at the ground truth location and appropriately sized for the respective tooth type. In addition, the corresponding label was removed from the ground truth data.

## 3 Results

Our approaches were evaluated on 118 labeled cases with a 6-fold cross validation strategy leaving out 20 datasets for validation at each iteration. The heatmap training was stopped after 50 epochs and the coordinate regression model after 400 epochs, as in both cases no notable improvement was observed. We also did not observe any indicators for overfitting. A visual comparison of the results for both approaches in 2D and in 3D is shown in Figure 7. A quantitative evaluation of the prediction accuracy and the deviation of predicted locations is shown in Figure 8. Some more examples on the performance of the heatmap approach are shown in Figure 9. Incorrect results were: An ambiguous prediction of a premolar (Fig. 9 a) and the resulting misclassification of adjacent teeth (Fig. 9 b) as well as a dental screw that was mistakingly predicted as a tooth (Fig. 9 c).

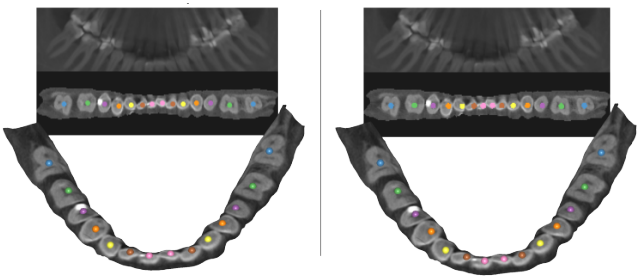


Figure 7: *Prediction of tooth locations and types in 2D and in 3D for heatmap regression (left) and coordinate regression (right). In the example, both approaches do classify all teeth correctly. A better performance in localization by the heatmap approach can be seen on the patient's right canine and premolars.*

When trained without data augmentation, our CNN approach showed a drop of 10% in effective accuracy (withing 2 mm bounds) for both regression models. Removing the APR as input channel and training only on the sARI, lead to a drop of only 0.5% for the heatmap model, while the coordinate model showed a drop of 4%, which is less than we expected.

## 4 Discussion

Our U-Net with heatmap regression model achieves good results by directly transforming our reformatted input images into a tooth location probability map. Although the coordinate regression turned out to be the weaker approach, combining both, landmark representations and classification information within one network to produce both outputs in a multi-task learning fashion has potential. The strength of our approach probably

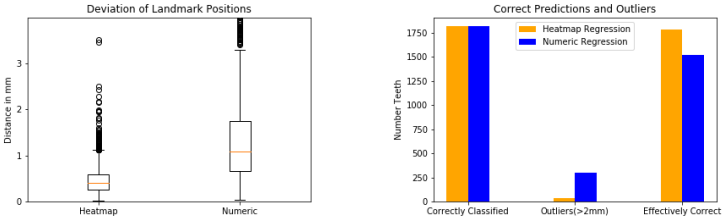



Figure 8: *C*omparison of 3D predictions for heatmap and coordinate regression. Depending on the threshold that is selected to interpret a landmark as outlier, the coordinate regression performs significantly worse which is shown by the 'Effectively Correct' entries that subtract the outliers from the number of correct predictions.

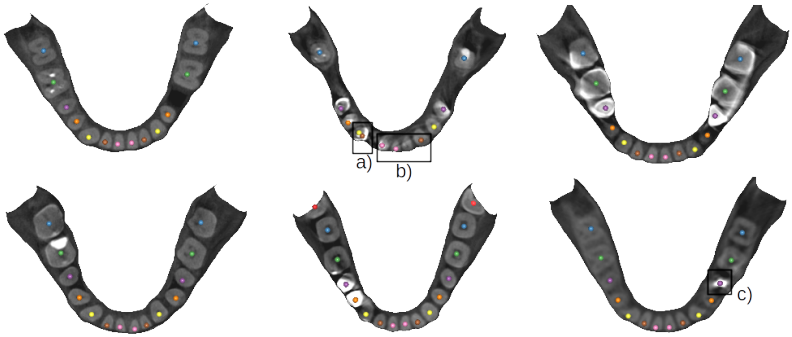



Figure 9: *Prediction results for the U-Net heatmap model for some validation examples.*

lies in the way how the training data are generated. Reformatting image data in such a way that the relevant information is maximized while the amount of data that needs to be processed within training is minimized opens opportunities for challenging tasks. The additional use of the APR did not prove as useful as expected. Its position and orientation was defined on the SSM patches for the dentiton region, which obviously was no good choice, since the correspondence of the teeth along the x-dimension between sARI and APR is not consistent. Using the center line of the fARI instead and the same bending operation as for the sARI could fix this issue and result in a better usage of the additional features. An alternative would be to construct a more realistic APR by detecting the occlusal plane. Compared to the method of Miki et al. [4], our approach has a slightly better accuracy (94 % vs. 88 %), although, a direct comparison is inconclusive because the results were not obtained on the same data. Our original vision was to achieve similar results as Tuzoff et al. [2] who obtained $\approx 98\%$ with their approach on panoramic X-rays based on 1,574 labeled data sets. Besides also training our network on more data, we see two opportunities to keep up: Either we also apply their heuristic for regularization, or we improve our APR and directly apply their approach on DVT data. The use of transfer learning might an additional option to improve the overall performance of our approach. One very important issue is data augmentation. This could be dramatically improved by a more realistic representation of missing teeth. The idea is to employ GANs for that purpose in our future work.

## 5 Conclusion

Within this work a supervised learning approach using 2D CNNs for the localization and classification of teeth in dental CBCT data has been presented, which robustly handles missing teeth, dental fillings, implants, and artefacts while providing useful results. Data pre-processing, dimension reduction by image reformatting, and

deep learning have been efficiently combined into a processing pipeline. Full tooth enumeration or classification of tooth types and sectors can be trained and successfully predicted in clinical data.

## 6 Acknowledgements

The authors would like to thank Heiko Ramm from 1000shapes GmbH Berlin, for his assistance in data provision and pre-processing, the SSM based segmentation, as well as the generation of the APRs. Furthermore, the author would like to thank Felix Ambellan of the Therapy Planning research group at ZIB for his implementation as well as his support with the application of the quasi isometric projection method. Our thanks also go to Alex Tack and Moritz Ehlke for their support in machine learning issues. Finally many thanks to all other members of the team that supported me during my master thesis [18] at ZIB.

## References

[1] International Organization for Standardization: *Designation system for teeth and areas of the oral cavity.* Dentistry, ISO 3950:2016 (2016)

[2] D. Tuzoff, L. Tuzova, M. Kharachenko: *Report on tooth detection and numbering in panoramic radiographs using CNNs.* 1st Conference on Medical Imaging with Deep Learning (2018)

[3] K. Zhang, J. Wu, H. Chen, P. Lyu : *An effective teeth recognition method using label tree with cascade network structure* . Computerized Medical Imaging and Graphics 68, 61–70 (2018)

[4] Y. Miki, C. Muramatsu, T. Hayashi, X. Zhou, T. Hara, A. Katsumata, H. Fujita: *Classification of teeth in cone-beam CT using deep convolutional neural network* . Journ. Computers in Biology and Medicine, 80, 24–29 (2017)

[5] M. Ezhov, A. Zakirov, M. Gusarev: *Coarse-to-fine volumetric segmentation of teeth in Cone-Beam CT.* arXiv preprint arXiv:1810.10293 (2018)

[6] Y. Gan, Z. Xia, J. Xiong, G. Li, Q.Zhao: *Tooth and Alveolar Bone Segmentation From Dental Computed Tomography Images*

[7] H. Gao, O. Chae: *Individual tooth segmentation from CT images using level set method with shape and intensity prior.* Pattern Recognition Journal, vol. 43:7 2406–2417 (2010)

[8] M. Hosntalab, A. Reza, A. Ali, S. Gholamreza: *Classification and numbering of teeth in multi-slice CT images using wavelet-Fourier descriptor.* Int. Journ. of Computer Assisted Radiology and Surgery, vol 5:3 237–249 (2010)

[9] TD. Nguyen, H. Lamecker, D. Kainmueller, S. Zachow: *Automatic Detection and Classification of Teeth in CT Data.* Proc. of Medical Image Computing and Computer Assisted Intervention (MICCAI), (Ayache et al., eds.), Springer LNCS 7510, 609–616 (2012)

[10] H. Lamecker, S. Zachow, A. Wittmers, B. Weber, H.-C. Hege, B. Elsholtz, M. Stiller: *Automatic Segmentation of Mandibles in Low-Dose CT-Data.* Int. Journal for Computer Assisted Radiology and Surgery, Springer, 1(1), Suppl. 7, 393–395 (2006)

[11] D. Kainmller, H. Lamecker, H. Seim, M. Zinser, S. Zachow: *Automatic Extraction of Mandibular Nerve and Bone from Cone-Beam CT Data.* Proc. of Medical Image Computing and Computer Assisted Intervention, (Yang et al., eds.), Springer LNCS 5762, 76–83 (2009)

[12] D. Kainmller, H. Lamecker, H. Seim, S. Zachow: *Multi-object Segmentation of Head Bones.* MIDAS Journal, 1–11 (2009)

[13] J. Kreiser, M. Meuschke, G. Mistelbauer, B. Preim, T. Ropinski: A Survey of Flattening-Based Medical Visualization Techniques. Computer Graphics Forum 37(3):597-624 (2018)

[14] Y. Wang, B. Liu, Y. Tong: *Linear surface reconstruction from discrete fundamental forms on triangle meshes.* Computer Graphics Forum 31(8):2277-2287 (2012)

[15] C. Payer, D. Stern, H. Bischof, and M. Urschler: Regressing Heatmaps for Multiple Landmark Localization Using CNNs, Springer LNCS 9901, 230–238 (2016)

[16] O. Ronneberger, F. Fisher, T. Brox: *U-Net: Convolutional Networks for Biomedical Image Segmentation.* Medical Image Computing and Computer-Assisted Intervention (Navab et al., eds.), Springer LNCS 9351, 234–241 (2015)

[17] S. Ioffe, C. Szegedy: *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* Proc. of 32nd International Conference on Machine Learning (ICML), arXiv:1502.03167 (2015)

[18] M. Neumann: *Localization and Classification of Teeth in Cone Beam CT using 2D CNNs.* Master thesis, Berlin University of Technology (2019)

# Towards robotic embedding of cochlear implants in the temporal bone

*J. Hermann[1], J. Ledergerber[1], M. Wigger[1], D. Schneider[1], F. Müller[1], S. Weber[1], G. O'Toole Bom Braga[1]*

*[1]ARTORG Center for Biomedical Engineering Research, University of Bern*

Contact: jan.hermann@artorg.unibe.ch

***Abstract***

The fixation of cochlear implant devices and their electrode is vital for long-term functionality. Today in robotic cochlear implantation, there is no standardized procedure for implant embedding in the temporal bone and dealing with excess electrode length. In this work, a robotic embedding approach for cochlear implants is proposed, and investigated in two experiments. In the first, the robotic system and suitable milling parameters were validated. It was shown that the robotic system can mill electrode channels in cortical bovine femur bone with an accuracy of 0.2 mm ± 0.1 mm and sufficient smoothness for successful electrode embedding. The second experiment in an upper-body phantom validated the proposed approach and clinical workflow. It showed in four example cases with different channel designs and implant bed positions that it was possible to mill and embed the cochlear implant, and to preoperatively plan the excess electrode lead length with sufficient accuracy.

**Keywords**: cochlear implant, robotic embedding, electrode channel milling

## 1 Introduction

Cochlear implantation is an otologic microsurgical procedure for people with profound sensorineural hearing loss. To restore hearing in these patients, an electronic device is implanted in the lateral skull base with an electrode array placed in the scala tympani of the cochlea. A study on long-term complications after cochlear implantation found that in 156 of 6204 patients a problem with the electrode occurred (e.g. electric failure, slip-out of the cochlea, exposure) that required revision surgery [1]. According to this study, most electrode slip-outs were avoided after the introduction of the modified split-bridge technique [1], [2], which proposes anchoring the electrode at the edge of the mastoidectomy cavity after stowing the excess electrode within.

Robotic cochlear implantation (RCI) is hypothesized to improve residual hearing preservation and improved audiological outcomes [3]. During RCI, a tunnel for the electrode is drilled based on an image-guided approach with a high-precision stereotactic surgical robot. This tunnel provides middle ear access directly from the mastoid surface, negating the need for a mastoidectomy. Therefore, a technique analogue to the split-bridge method has not yet been employed. To avoid electrode slip-outs and other failures, this study proposes a robotically milled electrode lead channel on the surface of the temporal bone. The channel will secure the excess lead length, accommodate lead length uncertainty due to incomplete insertion, and minimize lead curvature along the channel and the surface-to-tunnel interface to decrease the risk of electrode breakage.

In this study, two separate experiments were conducted. The first experiment is focused on the robotic manipulator, the milling parameters and the resulting accuracy in bovine cortical bone, while the second experiment targets the feasibility of the proposed approach and clinical workflow in an upper-body phantom.

## 2 Materials und Methods

### 2.1 The approach for robotic embedding of the cochlear implant

The robotic cochlear implantation surgery is planned preoperatively with a dedicated planning software [4], which generates models of the temporal bone, the cochlea and other necessary anatomical structures. As soon as the drill tunnel trajectory is placed, all necessary information for the robotic embedding are known. The model of the temporal bone defines the milling surface, while the drill tunnel and cochlear duct length define the excess electrode length that needs to be stowed in the temporal bone region. As the total length of the electrode of the chosen cochlear implant is known, the planned excess electrode length can be calculated.

The clinical workflow of robotic cochlear implantation is as follows: First, fiducial screws for patient-to-image registration are placed in the mastoid region. Then, a CT image is taken, upon which the robotic surgery is planned. Subsequently, the robotic system is set up, the patient tracking marker is fixed in the recommended position, and the direct cochlear access tunnel is drilled. At this part of the workflow, the robotic embedding of the cochlear implant has to be performed due to the two following reasons. Firstly, after the cochlea is opened for electrode insertion, no milling can take place since debris and blood could lead to a contamination of the cochlea, which in turn could lead to the loss of residual hearing [5]. Secondly, milling cannot be performed later due to the proximity to the already inserted electrode in the drill tunnel. Therefore, the implant bed and the channel have to be created before the electrode is inserted into the cochlea. Since the insertion depth during the surgery can differ from the planning, the channel needs to be able to accommodate this uncertainty. As a consequence, we propose the introduction of widenings in the path design (see Figure 1: Left). The widening are calculated based on a negative tolerance (i.e. the electrode is inserted deeper into the cochlea as planned) and positive tolerance (i.e. only a partial insertion was achieved).

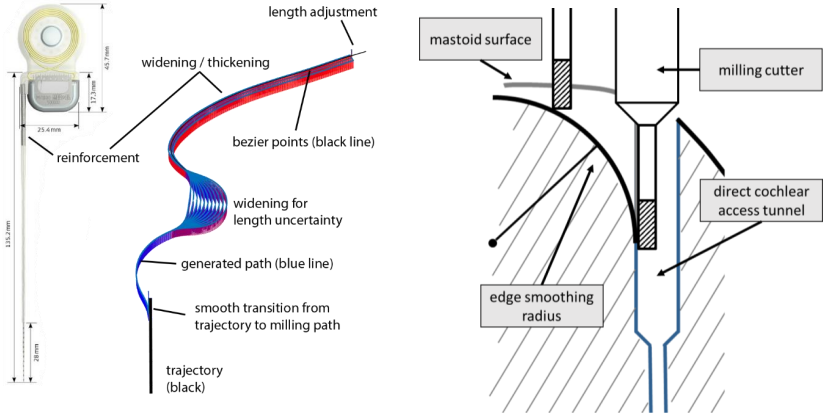



Figure 1: *Left: path design with widening, thickening* [6]. *Right: radius in channel path for smooth transition from direct cochlear access tunnel to channel in temporal bone surface*

Ideally, a planning software would guide the surgeon to place the implant according to the manufacturer's guidelines, and then draw the electrode channel on the surface of the mastoid while restricting its length and automatically introducing widenings at favorable locations.

To increase the safety for the electrode and decrease the risk of electrode failure, the curvature of the channel should be minimized and the channel width carefully chosen and dimensionally tolerated. Any acute angles of the embedded electrode should be avoided due to the higher likeliness of breakage through fatigue because of micro-movements over a period of time. Special attention needs to be paid to the tunnel-to-surface interface, because a sharp angle is usually created between the direct cochlear access tunnel and the mastoid surface [7]. Here, we propose rounding off this edge with the milling cutter during the execution of the channel milling (see Figure 1: Right).

For optimal stability and safety of the electrode, the channel is created slightly narrower than the electrode itself, creating a press-fit. For the example device the channel diameter needs to be enlarged to accommodate the reinforcement of the electrode at the attachment to the implant housing (see Figure 1).

The planning of the implant bed is accomplished with a self-developed software that allows the placement of an implant in the anatomy, and the creation of a milling path that the robotic manipulator can then follow (see Figure 2: Right). The milling path is created by slicing the implant model in a parametrized thickness, and subsequent offsetting and joining of the resulting contours. The algorithm then builds the path such that the milling cutter will begin in the middle of a slice and work its way outwards. This planning method follows the proposed approach by Federspil et al. [8].

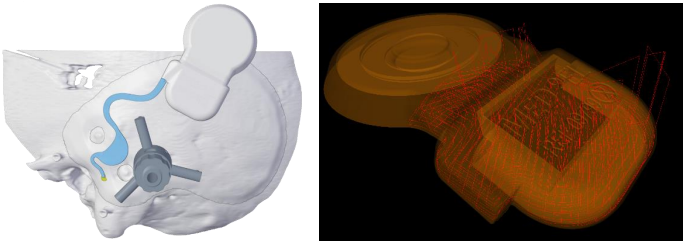


Figure 2: *Left: Visualization of the robotic embedding during cochlear implantation. Right: Implant bed milling path (red lines) overlaid on the cochlear implant in the self-developed software*

After the planning is complete, the milling paths are executed with the stereotactically-guided robotic manipulator [3] from our research group. To allow milling, the end-effector motor and gearbox were replaced with a high-speed motor without gearheads (ECX SPEED 13L, maxon, Switzerland) to enable spindle speeds up to 70'000 RPM. Additionally, the robot's motion control algorithm was adjusted with a cascaded proportional and derivative velocity control through the visual tracking and a proportional force control above 4 N up to a threshold of 10 N, where the system reduced its forward velocity to zero.

The experiments following up to a patient study in clinics have been separated into three stages which are roughly: determination and verification of milling parameters in bovine bone, feasibility of the clinical workflow in phantoms, and verification of the approach in cadaveric heads (n = 8). This work will present results from the first and second experiments, the cadaveric study will be conducted in the near future.

## 2.2 Experiment 1: Robotic milling in bovine bone

Milling experiments were performed on bovine bone specimens to test the overall performance of the robotic milling system and to validate a set of milling parameters. We hypothesize that straight electrode lead channels can be milled with an accuracy better than 0.5 mm with a 1.2 mm cylindrical milling cutter in bovine cortical bone. Additionally, the validity of the channel smoothness for embedding of the electrode will be evaluated. In all our experiments, the cochlear implant SYNCHRONY (MEDEL, Austria) was used, whose electrode diameter is roughly 1.3 mm. The bone thickness is around 5 mm [9] in the region surrounding the mastoid where the implant should be embedded according to the manufacturer's guidelines. Therefore, a robotic accuracy of 0.5 mm will enable a large portion of the population to be treated with this approach.

The milling experiments were performed on fresh bovine femurs in the cortical layer. First, three fiducial screws were implanted in suitable region on the cortical bone. The fiducial screws served as reference points to create a flat surface, within which several lines with different milling parameters could be milled, while recording forces on the robot end-effector (see Figure 4). The milling parameters to be validated are spindle speed, cutting depth, forward velocity, milling cutter-to-surface angle, and force limits.

The accuracy of the milled channels was evaluated with postoperative cone-beam computed tomography (CBCT) scans that were obtained from the university hospital in Bern (Inselspital, Switzerland) with an isotropic resolution of 0.07 mm. The accuracy and channel smoothness was measured using the segmentation software Amira (Thermo Fisher Scientific, USA) and Blender (Blender Foundation, Netherlands). The registration of planned channels and postoperative segmented channels was performed through the fiducials and exhibits a target localization error of 0.1 mm [4].The milling accuracy was determined by measuring the errors as the distance from a 3d-model of the planned channel to the 3d-model of the segmented channels (see Figure 4: Right). The errors were divided into depth and lateral errors. The channel smoothness was determined to be the precision of the robot system, i.e. the deviation errors from a straight line.

## 2.3 Experiment 2: Approach and clinical workflow in upper-body phantom

The experiments to validate the proposed approach and clinical workflow were performed in an upper-body phantom (see Figure 3). The phantom was created from a reanimation dummy, and modified with a skull replica where the temporal bone region can be easily replaced with 3d-printed templates. Additionally, the phantom has a silicon skin, where already an incision typical for RCI has been executed.

The planning of the robotic embedding of the cochlear implant in the temporal bone surface is divided into two sections: the implant bed for the inductive coil housing and the channel for the electrode array. We hypothesize that it is feasible to embed the cochlear implant housing and electrode in the milled cavities. This includes the

ability of the channel to accommodate the insertion depth uncertainty. Additionally, it is investigated if the workspace of the robotic system limited by kinematics and visual tracking is sufficient for the proposed approach. Another question is if path designs can easily be found that will not coincide with fiducial screw placement and the fixation of the tracking marker on the patient's skull. Also, the management of incision size and skin are evaluated. To investigate the feasibility of the proposed approach and its clinical viability, different path designs are planned on purpose, the positioning of the phantom is varied and the tracking marker of the robotic system is mounted in different locations.

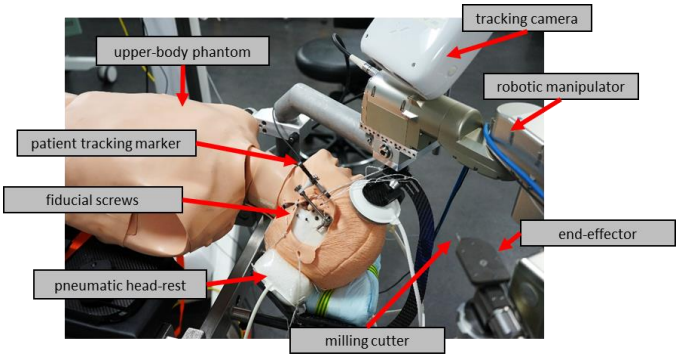



Figure 3: *Upper-body phantom with additively manufactured and replaceable temporal bone templates for the feasibility evaluation of the current approach and the clinical workflow*

Firstly, the cochlear implant is placed virtually on the temporal bone surface in the graphics software Blender. A path is drawn on the surface of the temporal bone model from the housing of the cochlear implant to the planned direct cochlear access tunnel (see Figure 2). This is repeated or adjusted until the drawn path length is slightly longer than the planned electrode length that needs to be embedded in the temporal bone. An algorithm written in MatLab (MathWorks, USA) will then use this drawn path, smoothen it, introduce the widenings and the tunnel-to-surface interface and cut the path to the desired length.

Secondly, the virtually placed cochlear implant model is loaded into the self-developed software to create a milling path for the implant bed. Thirdly, the milling path is executed with the robotic system. And lastly, an attempt is made to place the cochlear implant housing in the implant bed and the electrode into the drill tunnel and channel and its success or failure is reported. The electrode is then placed in the minimal and maximal length positions in the widenings, at which the embedded electrode length is measured and compared with the planned length.

# 3 Results

## 3.1 Experiment 1: Robotic milling in bovine bone

The straight channels in the fresh bovine femur bone were milled with a 1.2 mm diameter cylindrical 3-fluted hardened-steel milling cutter, with [-20, 0, 20] degrees angular inclination to the surface normal, with spindle speeds of [30, 45, 60] kRPM at depths of [1.5, 2.0, 2.5] mm with a forward velocity setpoint of 2 mm/s.

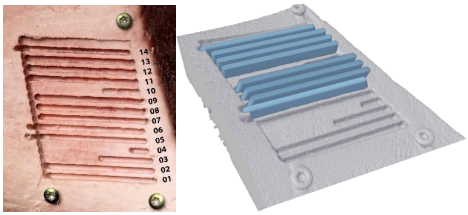


Figure 4: *Left: The milled straight electrode channels in fresh bovine femur bones. The lines are milled on a previously milled surface still visible in the right half of the picture. Right: The comparison of the post-experimental CT (gray) with the planned channels (blue).*

The depth accuracy in n = 30 channels evaluated in 0.07 mm isotropic CBCT was 0.2 mm ± 0.1 mm with a maximal depth error of 0.4 mm. The lateral accuracy was measured to be 0.2 mm ± 0.1 mm with a maximal lateral error of 0.4 mm. The maximal error usually occurred at the beginning or the end of the channel. The results are similar or equal when only results with one angular inclination of the milling cutter to the surface normal are considered, same as with the spindle speeds or depths. The channel smoothness was measured to be 0.2 mm.

With a forward velocity setpoint of 2 mm/s, the recorded forces with a spindle speed of 30 kRPM were measured to be 7 N on average with a maximum of 8.5 N while having an actual mean forward velocity of 0.7 mm/s. At 45 kRPM, both the average and maximal forces were lower at 4.2 and 6.9 N respectively. At a higher spindle speed of 60 kRPM, the average and maximal forces increased to 4.6 and 8.7 N respectively.

During the experiments, the hardened-steel milling cutter was replaced periodically such as not to obscure the force measurements with the tool wear.

### 3.2 Experiment 2: Approach and clinical workflow in upper-body phantom

In this experiment, n = 4 templates have already been planned, milled and analyzed. In all 4 cases, the implant bed and electrode could be successfully embedded in the milled implant bed and channel. Therefore, for this setup, the workspace of the robotic system was sufficient, and the patient marker could be left fixated in the recommended region for robotic cochlear implantation (i.e. no reattaching). The milling of the channels took less than 3 minutes, while the milling of the implant bed took less than 15 minutes.

Table 1*: Comparison of planned and measured excess electrode length that was embedded in the milled channel. The measurements were taken in the minimal and maximal length positions (see Figure 3: Right).*

| | **Minimal channel length (mm)** | | | **Maximal channel length (mm)** | | |
|---|---|---|---|---|---|---|
| **No.** | **Planned** | **Measured** | **Diff.** | **Planned** | **Measured** | **Diff.** |
| **0** | 68.6 | 68.2 | 0.5 | 77.3 | 76.3 | -1.0 |
| **1** | 67.3 | 68.2 | -0.9 | 76.4 | 77.0 | 0.6 |
| **2** | 66.4 | 66.5 | -0.1 | 76.5 | 75.8 | -0.7 |
| **3** | 66.9 | 67.0 | -0.1 | 75.5 | 76.5 | 1.0 |

Table 1 shows the difference between the planned stowed electrode length and the measured length. For all cases, the measured minimal, and the maximal length all lie within 1 mm of the planned lengths.

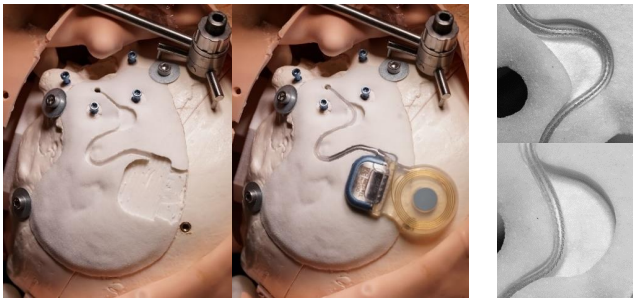


Figure 5: *Left: The resulting milled channel and implant bed for case 0 in the additively manufactured template without and with inserted cochlear implant device. Right: The electrode at maximal and minimal insertion depth in the widening of the channel.*

## 4 Discussion

The results of the experiments demonstrate the feasibility of robotically milling an electrode channel for cochlear implantation. Analogously to the split-bridge technique, the proposed electrode channel should provide fixation of the electrode in the mastoid and the inner ear. Furthermore, the channel should prohibit breakage of the electrode or its silicone sheath through minimized curvature and embedded topology.

The experiments in fresh bovine femur bone proved that the robotic system is able to mill electrode channels with a 1.2 mm milling cutter in inclinations of at least 20 degrees, at depths of at least up to 2.5 mm with a forward velocity setpoint of 2 mm/s. The robotic system was demonstrated to be able to mill straight channels with sufficient accuracy and smoothness for successful electrode embedding. The channels were milled with an accuracy of 0.2 mm ± 0.1 mm and with a maximal error of 0.4 mm. The channel smoothness error was measured to be 0.2 mm, which should be small enough not to pose any danger to the electrode.

In the phantom experiments, the electrode channel was planned preoperatively, and robotically milled in artificial temporal bones. For all experiments, paths could be found where the robot manipulator's end-effector would not be in danger of collision (e.g. with the patient tracking marker, or the fiducial screws), or occlusion, and all kinematic positions could be reached. The measured channel length matched the measured excess lead length in all cases with errors below 1 mm, and insertion depth variations could be managed with the appropriate widening of the channels.

The implant bed in these experiments are milled perpendicular to the surface of the phantom, which in a real patient is not desirable since it would require a large incision. A better approach would be to mill the implant bed with a spherical burr at an acute angle to the surface while holding up the skin, much like surgeons do manually. Alas, with the robotic system used, it proved as of yet too cumbersome due to restrictions in robot kinematics and visual tracking.

During the milling of the implant bed and the channel, the temperatures of the surrounding bone should not exceed 47°C degrees, such as not to cause thermal osteonecrosis [10]. Future works might include experimental validation of the milling parameters also in respect to temperature, but they were disregarded here because of the relatively shallow and thin channels, and the possibility to increase irrigation and milling cutter cooling.

## 5 Conclusion

The proposed approach for robotic embedding of cochlear implants in the temporal bone through implant bed and channel milling was validated. In bovine cortical bone it was shown that the robotic system achieves sufficient accuracy and channel smoothness for the application. In the phantom study it was shown that the proposed approach of preoperative planning and robotic milling after a direct cochlear access is feasible in the clinical workflow. These experiments provide the groundwork for a validation study in cadaveric head specimens.

## References

[1] J. Ikeya, A. Kawano, N. Nishiyama, S. Kawaguchi, A. Hagiwara, and M. Suzuki, "Auris Nasus Larynx Long-term complications after cochlear implantation," *Auris Nasus Larynx*, vol. 40, no. 6, pp. 525–529, 2013.

[2] A. Kawano, H. Chiba, K. Ueda, N. Nishiyama, and M. Suzuki, "The modified split bridge method to prevent electrode slip-out.," *Adv. Otorhinolaryngol.*, vol. 57, pp. 84–6, 2000.

[3] S. Weber *et al.*, "Instrument flight to the inner ear.," *Sci. Robot.*, vol. 2, no. 4, p. eaal4916, Mar. 2017.

[4] N. Gerber, B. Bell, K. Gavaghan, C. Weisstanner, M. D. Caversaccio, and S. Weber, "Surgical planning tool for robotically assisted hearing aid implantation.," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 9, no. 1, pp. 11–20, Jan. 2014.

[5] A. Khater and M. El-Anwar, "Methods of Hearing Preservation during Cochlear Implantation," *Int. Arch. Otorhinolaryngol.*, vol. 21, no. 03, pp. 297–301, Jul. 2017.

[6] MEDEL, "Mi1200 SYNCHRONY Surgical Guideline," 2019.

[7] W. G. Morrel, A. D. L. Jayawardena, S. M Amberg, B. M. Dawant, J. H. Noble, and R. F. Labadie, "Revision surgery following minimally invasive image-guided cochlear implantation.," *Laryngoscope*, vol. 129, no. 6, pp. 1458–1461, Jun. 2019.

[8] P. A. Federspil, U. W. Geisthoff, D. Henrich, and P. K. Plinkert, "Development of the first force-controlled robot for otoneurosurgery.," *Laryngoscope*, vol. 113, no. 3, pp. 465–71, 2003.

[9] J. Guignard, A. Arnold, C. Weisstanner, M. Caversaccio, and C. Stieger, "A Bone-Thickness Map as a Guide for Bone-Anchored Port Implantation Surgery in the Temporal Bone.," *Mater. (Basel, Switzerland)*, vol. 6, no. 11, pp. 5291–5301, Nov. 2013.

[10] G. Augustin, S. Davila, K. Mihoci, T. Udiljak, D. S. Vedrina, and A. Antabak, "Thermal osteonecrosis and bone drilling parameters revisited.," *Arch. Orthop. Trauma Surg.*, vol. 128, no. 1, pp. 71–7, Jan. 2008.

# Towards MR-Safe Endovascular Robotics

*D. Kundrat[1], G. Dagnino[1], M. E. M. K. Abdelaziz[1], T. M. Y. Kwok[2], W. Chi[1], C. Riga[2], G.-Z. Yang[1]*

[1] *The Hamlyn Centre for Robotic Surgery, Imperial College, London, UK*
[2] *Faculty of Medicine, Department of Surgery and Cancer, Imperial College London, UK*

Contacts:
d.kundrat@imperial.ac.uk
g.dagnino@imperial.ac.uk

***Abstract***

*Cardiovascular diseases are the most common cause of death worldwide. Endovascular interventions are considered as Gold standard in disease treatment. Remotely operated robotic devices are employed for minimally invasive approaches. The main advantages are: shorter recovery times, improvement of clinical skills (precision, stability), and facilitation of the procedure. However, benefits are compensated by absence of haptic feedback and high doses of radiation to the patient. This contribution describes a novel robotic master-slave concept that targets to join haptic feedback on the master, an intuitive user interface, and the capability of safe deployment to all clinically relevant imaging modalities within a familiar clinical workflow. Hence, the slave device can manipulate conventional instrumentation under fluoroscopic or MRI guidance. A phantom cannulation study evaluated the performance, reliability, and the procedural protocol. Secondly, a complementary study assessed and validated MR compatibility. Results strongly motivate the applicability and prospective clinical translation.*

**Keywords**: endovascular intervention, pneumatics, mri

## 1 Problem

Cardiovascular diseases impact global health and account for one third of world-wide deaths and almost 50% of deaths in Europe [1]. In general, heart and blood vessels are affected, and disease progress can manifest in cardiac arrests or strokes if no prior treatment was indicated. Endovascular interventions under image guidance have become the Gold standard and involve the deployment and manipulation of thin and flexible instruments, e.g. guidewires and catheters, to the targeted area in the vessel. State-of-the-art instrumentation enables different treatment options, such as embolization, stenting, or ablation. This minimally invasive approach reduces recovery times with improved outcome for the patient and likewise decreases procedural times [2].

Robotic technology was introduced for procedural assistance with improvement of precision and usability for the clinical operator and reduction of radiation exposure for both the patient and the operator. Various platforms were proposed in endovascular research [3]. Commercial devices are the Magellan and Sensei X2 platforms (AurisHealth, Redwood city, CA, USA), the Amigo platform (Catheter Precision, Mt. Olive, NJ, USA), the R-one robot (Robocath, Rouen, France), or the CorPath GRX platform (Corindus, Waltham, MA, USA). These devices enable electro-mechanical instrument manipulation in up six degrees of freedom (DoF) and generally employ conventional joysticks or 3D input devices as human machine interfaces (HMI).

However, proposed devices show the following major limitations: 1) restricted use of standard instrumentation, 2) usability neglects human motion patterns and kinaesthetic feedback, 3) collaborative use and intra-procedural transition between robotic and manual manipulation not implemented, and 4) fluoroscopy (ionizing radiation) still considered as Gold standard.

This work proposes the design and evaluation of a robotic platform that addresses current limitations. The device comprises a versatile master-slave setup and navigation framework. Unlike previous platforms, the robot can be safely integrated and used in MR environments thanks to pneumatic actuation and additive fabrication. Furthermore, an intuitive HMI is described which mimics the human motion pattern. Motions are mapped to an MR-safe slave robot which manipulates guidewires and catheters. The platform is evaluated in experimental studies under different imaging modalities and the results strongly pave the way for clinical translation with improved outcome for patients and clinical operators.

## 2 Material and Methods

This section describes the design of the robotic platform design, the control architecture, and the set-up for experimental device evaluation.

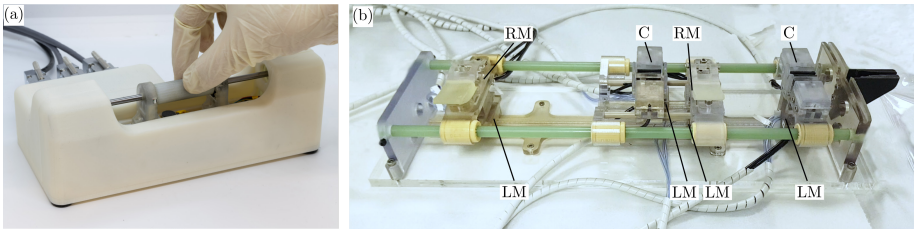



Figure 1: *Versatile endovascular robotic platform for teleoperation of catheters and guidewires with different imaging modalities: (a) haptic master device and (b) pneumatically actuated MR-safe slave robot with 4 DoF for instrument manipulation. Acronyms are defined as follows: linear motor (LM), rotary motor (RM), and clamp (C).*

## 2.1 Robotic Platform

The design of the robotic teleoperation platform implements a master-slave configuration. A haptic master is considered as human machine interface as depicted in Figure 1a. The device design is inspired by previous work [4] and mimics the clinically established routine of handling guidewires and catheters, i.e. grasping, push/pull or rotary motion. It further enables rotary and linear user input with force/torque feedback based on linear and rotary motors. Optical and piezo-resistive sensors implement contactless detection of grip condition and user intention. After linear displacement of the user handle to advance or retract the selected instrument, the integrated linear motor homes the handle automatically. Hence, mimicking the human motion pattern, arbitrary strokes can be commanded. Front buttons enable the operator to intuitively change to the manipulation of the desired instrument, i.e. catheter or guidewire. Housing illumination can be modulated to indicate the current system status.

The MR-safe slave robot in Figure 1b is implemented as bespoke 4 DoF pneumatically actuated instrument driver that mechanically maps the human motion pattern, i.e. grasping of an instrument and subsequent manipulation. Translation is achieved with pneumatic linear stepper motors; rotation is implemented with pneumatic rotary motors [5]. Linear motors show step widths of 0.3 mm and rotary motors angular steps of 10°. The maximum actuation frequency for safe and robust motion with additively manufactured components is approximately 30 Hz. The common pressure level of the pneumatic source is set to 0.3 MPa.

Customised pneumatic clamps are designed to transmit forces to both instruments in the translation phase.

All slave components are selected or fabricated with regard to MR compatibility. The robotic slave comprises multiple platforms for docking of instruments and manipulation. Pneumatic motors are controlled by a valve manifold that is located outside the theatre. The pneumatic controller interfaces high level robot control. A supplemental *video* highlights and summarises the main functions of the proposed robotic platform.

## 2.2 Control Architecture

Core components of the control architecture comprise a navigation framework and the robotic controller. This enables the operator to be in continuous control of the procedure and the navigation framework outputs haptic feedback for surgical guidance with improved accuracy and safety.

The navigation system was described in depth in [2] and comprises two core functions: 1) Intra-operative visualisation of the surgical imaging data to the operator and 2) generation of dynamic active constraints for haptic feedback rendered to the master based on visual information extracted from live imaging data. In brief, instruments (catheter, guidewire) were tracked in the imaging live stream for the angiography system (2d fluoroscopy) for determination of tip positions. Simultaneously, the corresponding vessel wall was detected and tracked to obtain a distance metric. The latter is combined with current motor currents to update viscous friction models for both axes that are rendered to the operator through the master device. In detail, friction was decreased to a minimum if the tip of the instrument is located in the centre of the vessel. The modelled friction increases if the detected distance between instrument and wall decreases. Ambiguous instrument configurations, such as the tip is centred in the vessel and pointing towards anterior or posterior direction, will be addressed in future work with advanced tracking and imaging procedures.

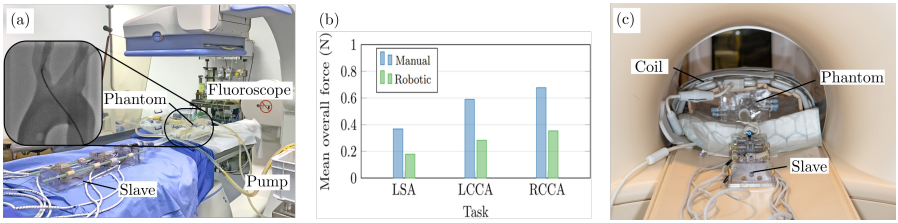



Figure 2: *Experimental set-ups and results of device evaluation: (a) robotic performance analysis in angiographic suite. Overlay shows X-ray imaging of employed aortic arch phantom. (b) Results of study under fluoroscopic guidance. (c) Experimental set-up of the device in MR environment.*

### 2.3 Experimental Set-Up

System feasibility and robotic task performance were assessed in two experimental studies. Firstly, experimental trials under fluoroscopic guidance were conducted in a research angiography theatre (Innova 4100 IQ, GE Healthcare, Barrington, IL, USA) and targeted the simulation of clinical manoeuvres with manual and robotic manipulation, such as cannulation of different arteries. A soft-silicone phantom of the normal adult aortic arch was employed for the cannulation task and coupling to a multi-axial force sensor enabled measurement of interaction forces between instrumentation and model. Slave robot and phantom were located underneath an X-ray imaging unit to mimic an endovascular intervention with arterial access from the femoral artery (see Figure 2 a. Vascular flow dynamics were simulated with a pulsatile pump. Subsequent to a familiarisation phase of approximately 6 min, subjects were asked to complete cannulation of three different arteries with increasing demands of manipulation difficulty and randomised order of manual and robotic sequences.

The task was composed of three phases: 1) traversing the descending aorta, 2) navigation within the aortic arch, and 3) cannulation of the specific artery. Targets were the left subclavian (LSA), the left common carotid (LCCA), and the right common carotid (RCCA) arteries. The manipulation difficulty increases with previous order. Each task was performed multiple times ($N = 4$) per condition and target under fluoroscopic guidance. Task metrics comprised cannulation and fluoroscopy times, mean overall, and mean maximum forces.

Additionally, a preliminary MRI study was conducted to demonstrate that the proposed system is MR-safe and performs in the environment without introduction of imaging artefacts. A phantom of an abdominal aorta (Elastrat Sarl, Geneva, Switzerland) was filled with water, equipped with a torso coil and placed in the isocentre of a 3T MR imaging system (Achieva, Philips Healthcare, Eindhoven, Netherlands) as shown in Figure 2 c. An MR-conditional guidewire (EPflex GmbH, Dettingen, Germany) was manipulated manually and afterwards with robotic assistance during data acquisition.

## 3 Results

This section presents the experimental results of fluoroscopy and MRI studies. Firstly, measured interaction forces between instrumentation and aortic phantom under fluoroscopic guidance are given as mean ± standard deviation for manual and robotic task completion. Results are further summarised in Figure 2 b. Mean forces of LSA cannulation yielded to $(0.37 \pm 0.21)$ N for manual and $(0.18 \pm 0.09)$ N for robotic intervention. For manual and robotic LCCA cannulation, $(0.59 \pm 0.22)$ N and $(0.28 \pm 0.12)$ N were measured, respectively. Finally, manual and robotic RCCA cannulation were determined to $(0.68 \pm 0.33)$ N and $(0.35 \pm 0.18)$ N. Mean differences between groups of manual and robotic intervention were statistically significant for all tasks as indicated by Student's t-test ($p < .05$).

However, results of corresponding cannulation times show a different outcome (mean ± SD). Manual cannulation of the LSA with $(6.7 \pm 4.5)$ s was significantly less time consuming in comparison to the robotic approach with $(50.0 \pm 28.3)$ s. The same applies for cannulation of the LCCA, i.e. $(65.0 \pm 2.8)$ s vs $(32.0 \pm 10.8)$ s . Finally, robotic cannulation of the RCCA with $(43.7 \pm 15.0)$ s was less time consuming in comparison to manual with $(58.3 \pm 24.9)$ s. Fluoroscopy times for robotic manipulation ranged from $(134.0 \pm 48.6)$ s to $(302.0 \pm 56.0)$ s.

Furthermore, the deployment and manipulation of MR conditional guidewires in an abdominal phantom within an MR environment has proven feasibility. The displacement of MR markers on the surface of the guidewire was measured in acquired image data and captured for the entire retraction phase with manual and robotic manipulation. Subtraction images composed of image data from manual and robotic manipulation

did not reveal significant artefacts between both conditions. Beyond that, the signal-to-noise ratio (SNR) of acquired image sequences associated to manual and robotic manipulation did not reveal significant differences between both manipulation conditions (means of approximately 26 dB).

## 4 Discussion

We presented in this contribution a novel versatile robotic master-slave platform for manipulation of standard endovascular instruments in different imaging environments, such as fluoroscopy and MRI. Besides of the pneumatic actuation of all DoF to satisfy MRI safety regulations, the design targets a lightweight, hands-on interventional workflow to support the introduction and acceptance to the surgical theatre. The performance of the proposed platform was evaluated in a pre-clinical phantom study with teleoperated cannulation of three aortic arteries under fluoroscopic guidance. Subsequently, a feasibility study in an MR environment on robotic displacement of MR conditional guidewires demonstrated complementary results towards the aspect of versatility.

All tasks were completed successfully and significantly reduced interaction forces were measured for robotic manipulation. Mean forces were reduced in between 20% to 30% for cannulation of LSA, LCCA, and RCCA in comparison to manual execution. With regard to maximum forces, a reduction of up to 77% has been achieved. In terms of the clinical workflow and safety considerations, findings are very promising to avoid intra-procedural penetration and trauma of vessel walls.

In general, robotic procedure times were increased. However, if a more complex task is performed, such as RCCA cannulation, robotic procedure times were slightly lower than manual. This motivates that the device should be in particular considered in demanding phases of the interventional procedure, where device accuracy for successful deployment and procedure safety are of more importance.

## 5 Conclusion

This contribution presented the design and evaluation of a novel MR safe robotic platform for endovascular surgery. Performance metrics obtained from robotic instrument manipulation under fluoroscopic guidance have proven applicability of the proposed system to successful cannulation of aortic arteries. Preliminary studies in an experimental MRI setting demonstrated that pneumatic actuation can be deployed to drive the instrumentation in close vicinity of the MR bore. In summary, benefits of the device are twofold and contribute to patient outcome and daily routines of clinical users. The results strongly support prospective transitions to MR image guidance in endovascular interventions. In particular, this enables transition of robotic technologies to paediatric endovascular applications that strongly demand for alternative instrument manipulation and imaging, e.g. alternative solutions with absence of ionizing radiation. Moreover, future work is dedicated to four fields of research: 1) integration of MR safe sensing to the robotic mechanism to improve the usability and procedural safety of the system, 2) analysis of advanced pneumatic instrument drivers, 3) automation, and 4) extended pre-clinical user and animal studies under consideration of both imaging modalities.

## References

[1] Joel Kaplan. *Kaplan's Cardiac Anesthesia*. Elsevier, 7th edition, 2017.

[2] G. Dagnino, J. Liu, M. E. M. K. Abdelaziz, W. Chi, C. Riga, and G.-Z. Yang. Haptic feedback and dynamic active constraints for robot-assisted endovascular catheterization. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1770–1775, Oct 2018. doi: 10.1109/IROS.2018.8593628.

[3] H. Rafii-Tari, C. J. Payne, and G.-Z. Yang. Current and emerging robot-assisted endovascular catheterization technologies: A review. *Annals of Biomedical Engineering*, 42(4):697–715, Apr 2014. ISSN 1573-9686.

[4] C. J. Payne, H. Rafii-Tari, and G.-Z. Yang. A force feedback system for endovascular catheterisation. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1298–1304, Oct 2012. doi: 10.1109/IROS.2012.6386149.

[5] V. Groenhuis and S. Stramigioli. Rapid prototyping high-performance mr safe pneumatic stepper motors. *IEEE/ASME Transactions on Mechatronics*, 23(4):1843–1853, Aug 2018. ISSN 1083-4435. doi: 10.1109/TMECH.2018.2840682.

# Refined process model for robotic middle and inner ear access

*F. Müller[1], D. Schneider[1], J. Hermann[1], J. Anso[1], G. Bom Braga[1], M. Matulic[2], S. Weber[1]*

*[1]ARTORG Center for Biomedical Engineering Research, University of Bern*
*[2]CAScination AG, 3008 Bern, Switzerland*

Contact: fabian.mueller@artorg.unibe.ch

***Abstract***

*A recent clinical trial has demonstrated the feasibility of robotic cochlear implantation. From the surface of the mastoid a robotically drilled tunnel to the inner ear allows to insert the electrode array directly through the tunnel into the cochlea. The current robotic workflow can be optimized with regard to preoperative planning, surgical process flow and intraoperative interventions. In order to ensure a reproducible retention of residual hearing in cochlear implantation, a method for consistent planning and atraumatic access to the inner ear is required. The objective of this study was to assess the feasibility of a refined process model for robotic middle and inner ear access. A verification of the model has been conducted by analysing the data from robotically drilled human cadaveric head specimen.*

**Keywords**: Robotic cochlear implantation, inner ear access, minimally invasive

## 1 Problem

Cochlear implantation (CI) is an otologic microsurgical procedure for people with profound sensorineural hearing loss. An electronic device is implanted in the lateral skull base with an electrode array placed into the cochlea. The procedure requires surgeons to perform microsurgical tasks at the limits of their dexterity, made more difficult due to the proximity of important anatomical structures (facial nerve, chorda tympani, ossicles, and external ear canal wall) and the delicate and critical nature of the internal structures of the cochlea. Trauma and disturbances that occur during the opening of the cochlea and insertion of the electrode can lead to hair cell loss, membrane displacement, membrane perforation or new bone formation and fibrosis within the Scala tympani resulting in a loss of residual hearing in up to 50% of patients [1,2]. The investigation of augmenting tools for CI surgery has resulted in the concept of robotic cochlear implantation (RCI) [3].

RCI is a robotic surgical procedure that can be performed without direct visual access to the anatomical structures undergoing surgery [3,4,5]. A tunnel is drilled to the middle ear with a sensor- and image-controlled robot supported by the safety protocols of neuromonitoring and positional estimates using correlations of bone density and drilling force measurements [6]. The procedure relies on precise surgical planning based on preoperative imaging to determine the optimal trajectory to the cochlea [7,8]. On completion of the robotic middle ear access (RMA), the cochlea is manually opened through the transcanal access.

This work has focused on the development of a safe approach for robotic inner ear access (RIA) [9,10]. A sensory guided robotic cochlear access and decision-making process could perform the procedure beyond human sensory and manual limits allowing for a higher consistency of residual hearing preservation. The work presented herein assesses the feasibility of optimizing the current robotic procedure and proposes a reliable and consistent model for robotic access to the inner ear.

## 2 Material and Methods

### 2.2 Refined process model

This work introduces a refined process model for robotic middle and inner ear access that was adapted and extended from the recent RCI procedure [3]. The general model presented here is the basis for a patient-specific, image-based precision access to the middle and inner ear (Fig. 1). This plan includes the geometric definition of the trajectory within the mastoid, task specific parameters for the different machine phases, risk reduction measures and safety features such as intraoperative cone beam computed tomography (CBCT) imaging, neuromonitoring,

bone density based force correlations and intraoperative visualization. In a stepwise process, all relevant stages of inner ear access for RCI were addressed: a precise and image-controlled milling process using a task-specific robotic technology [11]; secondary positional estimates using force correlation and milling depth prediction; and task-specific endoscopic visualization to detect whether the milling process provided a sufficient access to the cochlea while the endosteal membrane remains preserved.

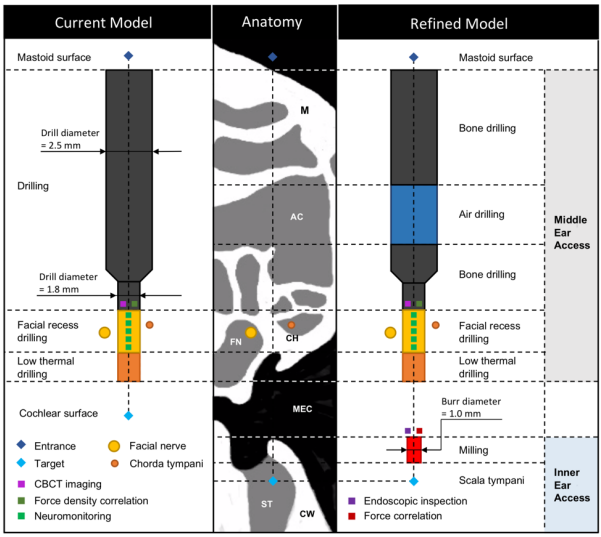



Figure 1: *Left: Current process model for robotic middle ear access; Right: Refined and extended process model for robotic middle and inner ear access; Centre: Anatomy of a human temporal bone shown in the plane of an optimally planned trajectory (M: Mastoid, AC: Air cell, FN: Facial nerve, CH: Chorda tympani, MEC: Middle ear cavity, ST: Scala tympani, CW: Cochlear wall).*

## 2.3 Parameterization of process model

### 2.3.1 Middle ear access parameterization

In the current model of RCI, the above-mentioned elements are identified from the preclinical image data by an otologic planning software (OTOPLAN®, CAScination, Switzerland). For further parameterization of the surgical procedure a task-specific planning algorithm has been developed. An extended optimization of RMA divides the trajectory automatically into anatomically specified process phases by investigating repetitive patterns in the density profile of the preoperative CT scan (Fig. 2) by using machine learning techniques. For middle ear access this includes transitions from high bone density to low density mastoid air cells (AB: Air Drilling → Bone Drilling) and vice versa (BA: Bone Drilling → Air Drilling) as well as the breakthrough into the middle ear cavity (MB). Physical drilling errors are taken into account by varying the position and orientation of the trajectory in an evenly spaced volume around the originally planned trajectory. The specific layout of the alternative trajectories was defined on the basis of the existing accuracy of the robotic system.

### 2.3.2 Inner ear access parameterization

Despite its superior overall accuracy, image guidance is not the only informant for positional accuracy in a robotic surgical procedure such as access to the inner ear. In this work, a secondary localization method based on the correlation of milling force and a positional force prediction for the cochlear wall has been implemented (Fig.3). Similarity measures can be used to determine the best match from a given set of force profile shapes that vary in position and orientation, generated in an evenly spaced volume around the originally planned trajectory. The depth of the lateral and medial wall as well as the optimal milling depth is determined automatically by the process refinement algorithm. In addition, a force profile shape is calculated for all trajectory candidates from the

preoperative CT image taking into account geometrical and dynamical properties of the milling burr. During the robotic procedure, similarity measures are used to determine the best match from the given candidate set to estimate whether the optimal milling depth based on preoperative planning has been reached.

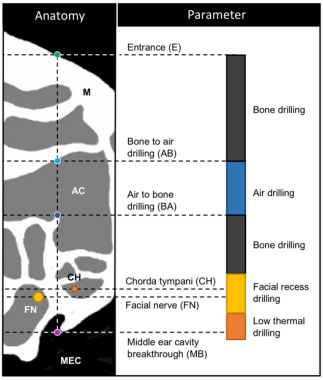



Figure 2: *Left: Relevant anatomy for middle ear access; Right: Transition parameters and corresponding drilling phases for robotic middle ear access.*

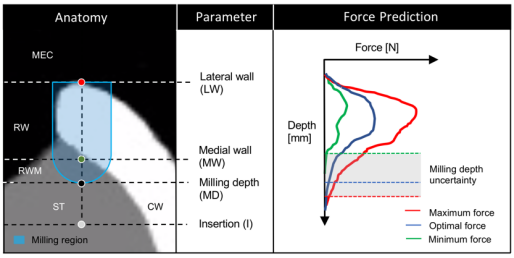



Figure 3: *Left: Anatomy of the round window/cochlear wall showing the planned milling region; Centre: Transition parameters and corresponding milling phase for robotic inner ear access. Right: Force prediction calculated by the algorithm in a range of possible candidate trajectories.*

## 2.4 Ex vivo validation study

For the verification of the new process model, data of n = 10 robotically drilled trajectories to the cochlea in Thiel embalmed human cadaveric head specimen were analysed. The left and right sight of five human cadaveric head specimens were prepared for robotic cochlear implantation by performing a c-shaped retro-auricular incision and implanting four bone anchored fiducial screws (CAScination, Switzerland). For all specimens, cone beam computed tomography (CBCT) images were acquired (XCat, Xoran Technologies, USA) and reconstructed with an isotropic resolution of 0.1 $mm^3$. The otologic planning software (OTOPLAN®, CAScination, Switzerland) was used to assess the temporal bone, facial nerve, chorda tympani, external auditory canal, the ossicles and the cochlear dimensions. An optimal trajectory was planned for each specimen from the mastoid surface to the round window with a diameter of 1.8 mm and a minimum clearance of 0.3 mm to all surrounding anatomical structures. The trajectory was chosen on the basis of the optimal insertion angle that is possible while going through the facial recess. After attaching the patient reference and performing a co-registration of the preoperative image data, middle and inner ear access were performed with the task-specific surgical robotic system (HEARO®, CAScination, Switzerland).

# 3 Results

## 3.1 Validation of middle ear access model

The planning algorithm of the refined process model was successfully applied to the image and acquired sensor data of all n = 10 data sets. In all cases, the transitions of the trajectory segments for the middle ear have been automatically determined and are anatomically correct (Fig. 4). The accuracy of the transition depths for air cells and the middle ear cavity breakthrough determined by a comparison of the transitions BA, AB and MB with the actually measured drilling forces are presented in table 1.

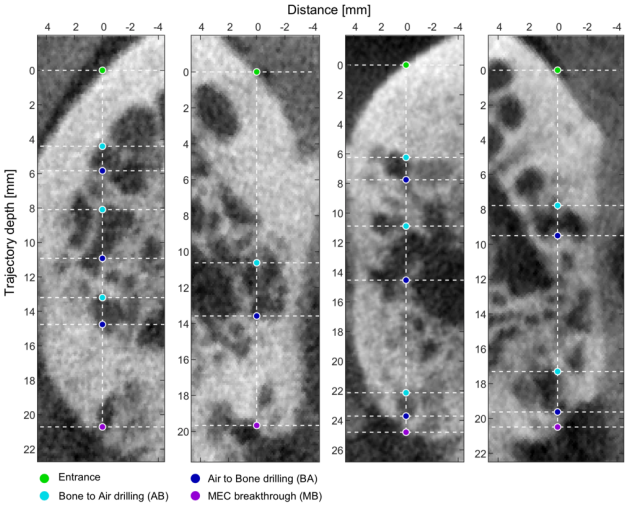



Figure 4: *Pre-operative CT image of four subjects showing the planned trajectory for middle ear access with the corresponding transitions depths determined by the refinement algorithm.*

Table 1*: Average absolute deviation of all determined transition depths.*

| Transition | BA | AB | MB |
|---|---|---|---|
| Avg. deviation [mm] | 0.307 | 0.251 | 0.169 |
| STD [mm] | 0.231 | 0.213 | 0.110 |

## 3.2 Validation of inner ear access model

The lateral and medial boundary of the inner ear (LW, MW) determined by the planning algorithm are as accurate as the planning conducted by an experienced ENT surgeon. For larger deviations (> 0.3 mm), an independent surgeon stated that the algorithm had delivered more accurate results than the initial planner.

From an objective point of view, the predicted and measured force profiles show a good agreement with respect to position, shape and amplitude (Fig 5). A comparison between the manual milling depth as a result of the surgeon's decision and the depth determined by the force correlation based milling depth algorithm shows minor deviations (< 0.31 mm) (Fig. 6). For cases where the manual procedure stopped earlier than the algorithm the last prediction of the algorithm is used as reference for the final milling depth.

In all cases, sufficient access for electrode insertion was created by the manually guided robotic inner ear access procedure. The preservation of the round window (RW) membrane was verified by an ENT surgeon via an endoscope and/or microscope in 9 of 10 cases.

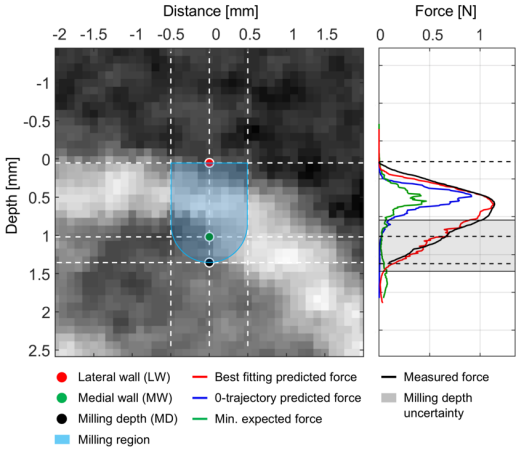



Figure 5: *Left: Pre-operative CT image of one subject showing the planned trajectory and milling region for inner ear access with the corresponding transitions depths determined by the refinement algorithm. Right: Predicted and measured force profile shapes (0-trajectory = planned optimal trajectory).*

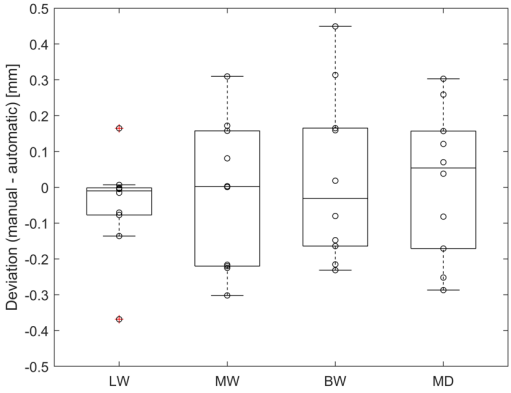



Figure 6: *Deviation between the manual and automatic planning as well as the deviation of the final milling depth. (LW: Lateral wall, MW: Medial wall, BW: Bony overhang width, MD: Milling depth.*

## 4 Discussion

This work introduces a new robotic process model for cochlear implantation that was adapted and extended from the recent robotic procedure with the aim to reduce the human preoperative workload and intraoperative intervention while enabling more efficient CI surgery.

With this study, we were able to show that it is possible to automatically refine the robotic process model during middle ear access and extend the current procedure with the functionality of robotic inner ear access. An automatic approach for anatomical segment selection was described, which represents the basic prerequisite for an automated optimization of the procedure. The introduced methods were validated trough a comparison with the manual planning, drilling and milling performed with the extended RCI procedure on ten human temporal bone specimens. The agreement of the geometric outcome between the manually performed procedure and the retrospectively

analysed automatic approach shows the feasibility of the developed algorithms. A refinement of the process model by adapting and extending the current robotic procedure could eventually enable a more efficient and less traumatic CI surgery. It is generally assumed that the precise milling of an access to the inner ear ensures the integrity of the endosteal membrane which is a basic requirement for maximum residual hearing preservation.

## 5 Conclusion

An extension of the current robotic process model with RIA as well as an automatic refinement of RMA is feasible. The force profile information during inner ear access has proven to be a valuable information for an independent estimation of the position of the burr when accessing the inner ear. Although the presented model for RIA provides sufficient access for electrode insertion, it does not necessarily have to be the optimum in terms of minimizing trauma to the inner ear. Therefore, the presented robotic inner ear access model does not yet allow a final conclusion to be made about the audiological outcome for cochlear implantation.

## Acknowledgments

We would like to thank Marco Matulic, Laetitia Racz-Perroud, Slobodan Koprivica, Daniel Erpenbeck and Dr. Antoine Leouzey from CAScination AG for software and hardware support and Dr. Masoud Zoka Assadi (Med-El Gmbh) for his advice.

## 6 References

[1] Roland, P. S., Gstöttner, W., & Adunka, O. (2005). Method for hearing preservation in cochlear implant surgery. *Operative Techniques in Otolaryngology-Head and Neck Surgery*, *16*(2), 93–100.

[2] James, C., Albegger, K., Battmer, R., Burdo, S., Deggouj, N., Deguine, O., ... Fraysse, B. (2005). Preservation of residual hearing with cochlear implantation: How and why. *Acta Oto- Laryngologica*, *125*(5), 481–491.

[3] Weber S, Gavaghan K, Wimmer W, et al (2017) Instrument flight to the inner ear. Sci Robot 2:eaal4916. https://doi.org/10.1126/scirobotics.aal4916

[4] Caversaccio M, Gavaghan K, Wimmer W, et al (2017) Robotic cochlear implantation: surgical procedure and first clinical experience. Acta Otolaryngol 137:447–454. https://doi.org/10.1080/00016489.2017.1278573

[5] Labadie RF, Balachandran R, Noble JH, et al (2014) Minimally invasive image-guided cochlear implantation surgery: First report of clinical implementation. Laryngoscope 124:1915–1922. https://doi.org/10.1002/lary.24520

[6] Ansó J, Dur C, Gavaghan K, et al. A neuromonitoring approach to facial nerve preservation during image-guided robotic cochlear implantation. Otol Neurotol 2015;37:89-98.

[7] Bell B, Gerber N, Williamson T, et al (2013) In Vitro Accuracy Evaluation of Image-Guided Robot System for Direct Cochlear Access. Otol Neurotol 34:1284–1290. https://doi.org/10.1097/MAO.0b013e31829561b6

[8] Gerber N, Bell B, Gavaghan K, et al. Surgical planning tool for robotically assisted hearing aid implantation. Int J CARS 2014;9:11-20.

[9] Brett, P. N., Taylor, R. P., Proops, D., Coulson, C., Reid, A., & Griffiths, M. V. (2007). A surgical robot for cochleostomy. *Conference Proceedings: International Conference of the IEEE EMBS.*, 1229–32.

[10] Wimmer W, Venail F, Williamson T, et al. Semiautomatic cochleostomy target and insertion trajectory planning for minimally invasive cochlear implantation. BioMed Res Int. 2014;2014:596498.

[11] B. Bell, N. Gerber, T. Williamson, K. Gavaghan, W. Wimmer, M. Caversaccio, S. Weber, In vitro accuracy evaluation of image-guided robot system for direct cochlear access. Otol. Neurotol. 34, 1284–1290 (2013).

# A New Setup for Markerless Motion Compensation in TMS by Relative Head Tracking with a Small-Scale TOF Camera

*M. Gromniak, C. Brendes, A. Schlaefer*

*Technical University of Hamburg, Institute for Medical Technology, Hamburg, Germany*

Contact: martin.gromniak@tuhh.de

***Abstract***

In robotic transcranial magnetic stimulation (TMS) a robot is used to move the magnetic coil so that head movements of the patient are compensated during the treatment. The state of the art is to track the head with the help of reflective marker geometries and an external tracking camera. We propose a new motion compensation setup, where a 3D camera is directly mounted on the robot's arm, rigidly attached to the end effector and the coil. A variant of the iterative closest point algorithm is used to estimate the transformation between the initial pointcloud template of the head and the pointcloud at the current timestep. With an appropriate calibration from camera to robot the robot's goal pose can be calculated so that the transformation between head and coil remains constant. The setup was evaluated with a head phantom and achieved a RMSE of 3.53 mm and 1.22° for head movements in six degree of freedom.

**Keywords**: Motion Compensation, Head Tracking, Transcranial Magnetic Stimulation

## 1 Problem

Robotic motion compensation is a necessary capability for various medical procedures. The general idea is to adapt physical devices so that the relative movement between a target region and the device is canceled out. In robotic radiation therapy, the position of the radiation source is adapted to the movements of the tumor due to e.g. respiration with the help of a robot arm [5][4]. Examples for robotic motion compensation in surgery are laser osteotomy in the presence of breathing motions[3] or drilling stabilization during hip resurfacing prosthesis surgery [12]. Another application of motion compensation is transcranial magnetic stimulation (TMS). In TMS, a magnetic coil is used to stimulate cortical regions of the brain. The medical applications of TMS originate from the treatment of mental illness [7] and have today expanded to treatment of tinnitus [1], diagnosis of the connection between the motor cortex to the peripheral nervous system [13] and pre-surgical mappings of cortex regions [8]. The target area of stimulation is usually very small with a size of $< 1$ cm. Therefore is is necessary that the pose of the coil with respect to the head remains constant during the treatment. Early solutions of this problem consisted of clamping the head of the patients with stereotactic fixtures [14], resulting in a high level of discomfort for the patient. Robotic TMS removes the need for such fixtures, by actively moving the coil with the help of a robot so that the movements of the head are compensated [9] [10]. Thus the problem of estimating the head movement arises. In currently available systems, passive reflective markers are attached to the patients head, e.g. with a headband [9]. Theses reflective marker geometries are tracked by an external tracking camera. With a calibration from tracking camera to robot and eventually robot to magnetic coil the robot can be moved in such a way that a desired relative between head and coil is achieved or maintained respectively. While these systems usually achieve a high motion compensation accuracy they also come with multiple disadvantages. First, optical tracking systems like the fusionTrack by Attracsys or the Polaris by NDI are an expensive part of the overall system. Second, reflective markers cause discomfort for the patient and can be easily displaced, for example by scratching.

In this work, we take a new approach for compensating motions in robotic TMS. Instead of having the tracking system decoupled from an actuator, we mount a 3D camera directly on the robot's end effector, rigidly attached to the coil. From an initial coil-to-head pose, where the camera is pointed at the face of the patient, the system is able to compensate for the relative head motions effectively.

## 2 Material and Methods

### 2.1 Hardware Setup

The hardware setup consists of a Universal Robot 3 (UR3), the TMS coil and the Pico Flexx camera, which is mounted on a cantilever so that it captures the face of the patient frontally. The setup is shown in Figure 1.

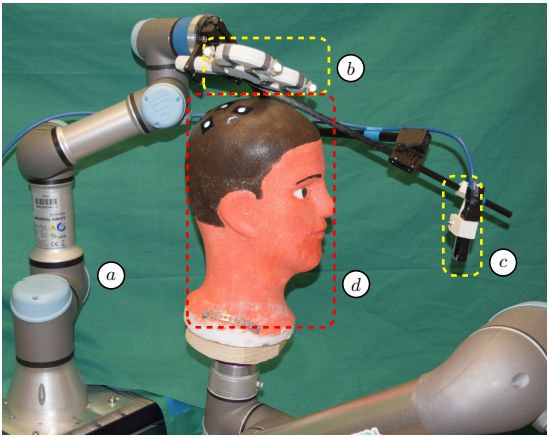



Figure 1: The prototypical setup of the motion compensation system. The UR3 Robot (a) holds a mock-up of the magnetic coil (b). The camera (c) is also rigidly attached to the end effector. The system was evaluated with a styrofoam head phantom (d). For this purpose, reflective markers were attached to coil and head phantom.

#### 2.1.1 Pico Flexx Camera

For the intended setup, the camera must be mounted on the robot and capture images withing a relatively close range of approximately 0.2 m - 0.4 m. The Pico Flexx, a small-scale time-of-flight camera by PMD Technologies is used for this purpose. With a size of 68 mm × 17 mm × 7 mm and a weight of 8 g it is one of the most portable 3D cameras available. It provides depth and infrared images with a resolution of 224 × 171 pixels. Especially the small near plane distance of 0.1 m makes it attractive for the sketched setup. The frame rate was set to 35 FPS, which corresponds to a specified depth sensing accuracy of 1% of the measured distance.

#### 2.1.2 UR3 Robot

For this work, the the Universal Robot 3 (UR3) has been chosen. It is a medium-sized collaborative robot with six degrees of freedom and a reach of 0.5 m. The maximum load at the robot's end effector is 3 kg which is sufficient for a TMS coil, the camera and the cantilever setup. The repeatability of ±0.1 mm qualifies it as a good choice for the TMS setup to maintain proper accuracy throughout the treatment. Another important feature is the internal collision detection as the robot is operating in close proximity to the subject's head. In case an unexpected collision happens the robot performs a emergency stop to prevent further damages to surroundings and itself.

#### 2.1.3 Cantilever Setup

With the cantilever construction, the coil and the camera are attached to the robot's end effector flange. It was designed with the goal to reduce jitter and oscillation of the camera as much as possible, as those directly affect the overall motion compensation performance. Full material 10 mm and 15 mm carbon rods were used for the cantilever, while connector elements were 3D printed from PLA.

### 2.2 Motion Compensation Algorithm

The algebraic motion compensation approach for relative tracking of the head involves the coordinate frames for the robot base (B), the robot's end effector (E), the camera (C) and the head (H). The calculation of the desired pose for the end-effector depending on the estimated displacement of the head is visualized in Figure 2 as a three step process.

1. In the initial configuration, at the start of the motion compensation, the transform from camera to head frame ${}^{C}T_{H}$ is calculated as described in Section 2.2.1.

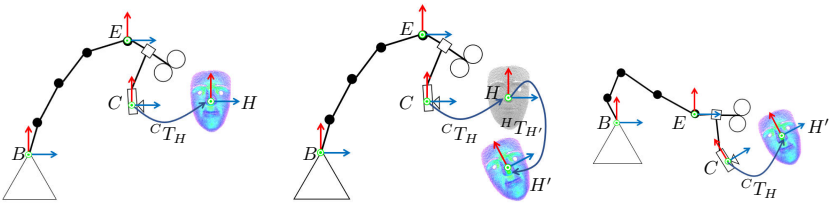



Figure 2: The motion compensation process in three steps. First, the transformation from the camera to the initial head frame is calculated (i). Then, the current pose of the head is calculated (ii). Lastly, the end effector can be moved so that the initial camera to head pose is restored (iii).

2. When a a new pointcloud is received from the camera, a registration is performed which calculates the transform from the the initial head frame to the current head frame ${}^{H}T_{H'}$. The details are described in Section 2.2.2. The current pose of the head in the robot base frame can therefore be derived as:

$$ {}^{B}T_{H'} = {}^{B}T_{E}\,{}^{E}T_{C}\,{}^{C}T_{H}\,{}^{H}T_{H'}. \tag{1} $$

   In this equation, ${}^{E}T_{C}$ is the pose of the camera in the end effector frame. It is retreived with a hand-eye calibration as described in section 2.2.3.

3. Subsequently, the new target pose for the robot can be calculated as

$$ {}^{B}T_{E} = {}^{B}T_{H'}({}^{C}T_{H})^{-1}({}^{E}T_{C})^{-1}. \tag{2} $$

   As a result, the initial transformation between camera and head is restored.

#### 2.2.1 Estimation of the initial head pose ${}^{C}T_{H}$

For only tracking relative motions of the head it is not necessary to track an explicit head coordinate system defined by anatomical landmarks. This omits a potentially error-prone step and simplifies the tracking task. We construct a surrogate head coordinate system, that is is a translated version of the camera coordinate system. Its translational components are the coordinates of the pointcloud's center of mass. Since in the subsequent step, the ICP performs only a relative estimation of the transform to the current pointcloud of the head, the orientation of the initial head frame is arbitrary. Therefore, the rotational part of ${}^{C}T_{H}$ is set to the Identity.

#### 2.2.2 Head Point Cloud registration with the Iterative Closest Point Algorithm

The transformation ${}^{H}T_{H'}$ is the result of the registration between the pointcloud recorded at the initial pose and the pointcloud at the current timestep. The iterative closest point (ICP) algorithm [2] based on the implementation of the PointCloud Library [11] has been used to provide this registration. As a high control frequency is desired in the motion compensation setup and the runtime of the ICP increases with the number of points, the pointclouds were preprocessed before using them in the ICP registration. First, only a subset of most informative points in the pointclouds is used. The points are defined by a rectengular region which covers the eyes, the nose and the mouth as recorded in the initial pose. Second, the pointclouds are downsampled using a voxel grid filter with a size of $0.1\,\text{mm} \times 0.1\,\text{mm} \times 0.1\,\text{mm}$.

#### 2.2.3 Hand-Eye Calibration of Camera and Robot

For the hand-eye calibration of camera and robot, a standard procedure was applied were a stationary checkerboard pattern was placed in the field of view of the camera. The robot moved to 50 random poses, recording the robot pose together with the transform from camera to checkerboard. The QR24 algorithm [6] was used to solve the handeye calibration problem. A mean translational error of 2.947 mm and a rotational error of 2.445° was achieved for the full transformation chain.

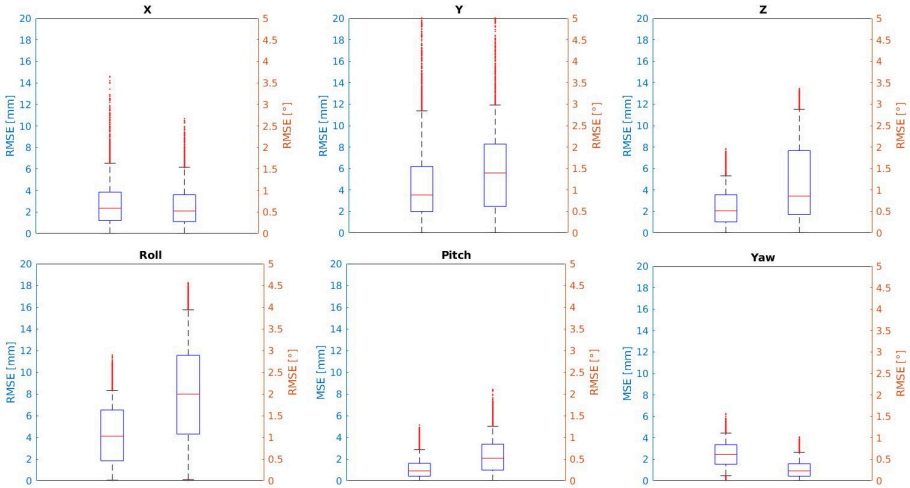



Figure 3: Motion compensation errors for head movements in six degrees of freedom. The left boxplots show the translation error $d_{rmse}$. The right boxplots show the rotation error $r_{rmse}$. The upper and lower whiskers depict the 95 % and 5 % quantile respectively.

### 2.3 Experimental Evaluation

The motion compensation system was evaluated with a styrofoam phantom head which was mounted on a second robot arm. The Atracsys fusionTrack 500 optical tracking camera was used to record the poses of coil and head during the motion compensation. Therefore, reflective markers were mounted on the coil and adhered to the phantom head respectively. The initial positioning of the coil relative to the head was chosen to represent a typical TMS configuration. Head movements in all six degrees of freedom (DOF) were evaluated individually by moving the head repetively only in the specific DOF. The 4D/BTi head coordinate system was used for this purpose. For translational movements in $x$, $y$, and $z$ directions, an amplitude of 100 mm and a speed of $10\,\mathrm{mm\,s^{-1}}$ was applied. For roll, pitch and yaw rotations around the head's center, an amplitude of 10° and and a speed of $1\,^\circ\,\mathrm{s^{-1}}$ was used. All movements were repeated ten times.

Based on the initial transformation from head to coil ${}^{H}T_C^0$ and all subsequent transformations ${}^{H}T_C^i,\ i \in [1, N]$ two error metrics can be constructed. The translational root mean squared error is defined as

$$d_{rmse} = \sqrt{\frac{1}{N}({}^{H}d_C^0 - {}^{H}d_C^i)^2}, \tag{3}$$

where ${}^{H}d_C^*$ denotes the euclidean distance calculated from the translational part of ${}^{H}T_C^*$. The rotational root mean squared error is defined as

$$r_{rmse} = \sqrt{\frac{1}{N}({}^{H}\theta_C^0 - {}^{H}\theta_C^i)^2}, \tag{4}$$

where ${}^{H}\theta_C^*$ is the angle of the the axis angle representation calculated from the rotational part of ${}^{H}T_C^*$.

## 3 Results

The results of the motion compensation in terms of translational and rotational errors are summarized in Figure 3. For all six movement types, the robot was able to successfully follow the head. The RMSE across all movement types are 3.53 mm and 1.22° for translation and rotation respectively. The runtime of the algorithm was on average 23.17 ms and therefore stayed well below the time between two camera frames (28.5 ms).

## 4 Discussion

The results show the feasibility of the presented motion compensation approach. The trajectories used for the phantom head included all DOF and abrupt direction changes. Therefore they resemble a good test environment for real head trajectories during TMS treatments. The results show a relatively large difference in the errors between movement types. For example, errors in y and roll directions are large. One possible explanation is that during those movements the pointclouds look substantially different due to the fixed crop based on the initial pointscloud, and therefore cannot be matched very well. This could be verified by correlating ICP metrics to the tracking errors. Another explanation are differences in the templates, that were each acquired at the beginning of the motion compensation for each movement type. As the pointclouds are inherently noisy, averaging multiple pointclouds could result in a more robust template.

The presented motion compensation setup has multiple advantages, compared to the traditional setup with a tracking camera. First, the price of the Pico Flexx (235€) is substantially lower than the price of conventional tracking cameras (10.000-15.000€), which can be a crucial factor when deciding for or against a robotic TMS system. Additionally, the setup is space efficient, has a quick setup time and the amount of tolerable head movement is only limited by the robot.

## 5 Conclusion

In this work, we presented a new approach for head motion compensation during TMS. A prototype was built and evaluated with movements of a phantom head. The results are promising when considering the accuracy requirements for TMS. Also, the new setup has multiple non-quantitative advantages. For future work, the generalization to different head phantoms and human subjects needs to be shown. Also, the refinement of ICP hyperparameters and preprocessing of the pointclouds in order to achieve a robust head tracking are worthwhile work approaches.

## References

[1] M. Anders, J. Dvorakova, L. Rathova, P. Havrnkov, P. Pelcova, M. Vaneckova, R. Jech, M. Holcat, Z. Seidl, and J. Raboch. Efficacy of repetitive transcranial magnetic stimulation for the treatment of refractory chronic tinnitus: A randomized, placebo controlled study. Neuro endocrinology letters, 31:238–49, 01 2010.

[2] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2):239–256, Feb 1992.

[3] M. Busack, G. Morel, and D. Bellot. Breathing motion compensation for robot assisted laser osteotomy. In 2010 IEEE International Conference on Robotics and Automation, pages 4573–4578, May 2010.

[4] S. Camps, D. Fontanarosa, P. With, F. Verhaegen, and B. Vanneste. The use of ultrasound imaging in the external beam radiotherapy workflow of prostate cancer patients. BioMed Research International, 2018:1–16, 01 2018.

[5] R. Dürichen, T. Wissel, F. Ernst, A. Schlaefer, and A. Schweikard. Multivariate respiratory motion prediction. Physics in medicine and biology, 59:6043–6060, 09 2014.

[6] F. Ernst, L. Richter, L. Matthäus, V. Martens, R. Bruder, A. Schlaefer, and A. Schweikard. Non-orthogonal tool/flange and robot/world calibration. The International Journal of Medical Robotics and Computer Assisted Surgery, 8(4):407–420, 2012.

[7] J. C. Horvath, J. M. Perez, L. Forrow, F. Fregni, and A. Pascual-Leone. Transcranial magnetic stimulation: a historical evaluation and future prognosis of therapeutically relevant ethical concerns. Journal of Medical Ethics, 37(3):137–143, 2011.

[8] S. Krieg, J. Sabih, L. Bulubas, T. Obermller, N. Chiara, I. Janssen, E. Shiban, B. Meyer, and F. Ringel. Preoperative motor mapping by navigated transcranial magnetic brain stimulation improves outcome for motor eloquent lesions. Neuro-oncology, 16, 02 2014.

[9] C. Lebossé, B. Bayle, M. de Mathelin, E. Laroche, and J. Foucher. Robotic image-guided transcrasial magnetic stimulation. International Journal of Computer Assisted Radiology and Surgery, 2006.

[10] L. Matthäus, P. Trillenberg, C. Bodensteiner, A. Giese, and A. Schweikard. Robotized tms for motion compensated navigated brain stimulation. International Journal of Computer Assisted Radiology and Surgery, 1:139–141, 06 2006.

[11] R. B. Rusu and S. Cousins. 3D is here: Point Cloud Library (PCL). In IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, May 9-13 2011.

[12] P. Torres, P. Gonalves, and J. Martins. Robotic system navigation developed for hip resurfacing prosthesis surgery. pages 173–183, 07 2018.

[13] M. Weber and A. Eisen. Magnetic stimulation of the central and peripheral nervous systems. Muscle & nerve, 25:160–75, 02 2002.

[14] C. Weltens, K. Kesteloot, G. Vandevelde, and W. V. den Bogaert. Comparison of plastic and orfit masks for patient head fixation during radiotherapy: Precision and costs. International Journal of Radiation Oncology*Biology*Physics, 33(2):499 – 507, 1995.

# Virtual & Augmented Reality

# A comparison of streaming methods for the Microsoft HoloLens

*V. Kraft[1], J. Strehlow[1], S. Jäckle[2], V. García-Vázquez[3], F. Link[1], F. von Haxthausen[3], A. Schenk[1], C. Schumann[1]*

[1] *Fraunhofer MEVIS - Institute for Digital Medicine, Bremen, Germany*
[2] *Fraunhofer MEVIS - Institute for Digital Medicine, Lübeck, Germany*
[3] *Institute for Robotics and Cognitive Systems, University of Lübeck, Lübeck, Germany*

Kontakt: valentin.kraft@mevis.fraunhofer.de

***Abstract***

The Microsoft HoloLens is an Augmented Reality headset that gained wide adoption in various areas during the last years. Many applications face the challenge of bringing complex contents to the HoloLens, which can be problematic given the limited hardware capabilities. In addition to surface rendering with high polygon counts, this also includes advanced volume rendering, which is often used in medical applications. One solution to enable such challenging visualizations on the HoloLens is to render the contents on another, more capable computer and to stream the final images to the HoloLens. In this work, we implement various methods for streaming custom content to the HoloLens and compare their efficiency when being decoded on the HoloLens. This includes the JPG and the DXT1 image compression formats, as well as the popular H.264 video codec and HoloLens Remoting. Our results indicate that the DXT1 and H.264 formats are fast and well-suited for streaming.

**Keywords**: Augmented Reality, HoloLens, Streaming, Visualization

## 1 Problem

Augmented Reality (AR) has been adopted in many application areas over the last years, including medicine, engineering, and education. In the medical context, AR might be used to transmit radiological data and derived patient specific models into the operating room [1, 2, 3, 4, 5]. That way, the surgeon can explore the data and compare it with the current situation in the situs without the need to look into the direction of a monitor. Furthermore, the interaction with the data might be facilitated with hand gestures. An established device for this kind of AR applications is the Microsoft HoloLens. One aspect that differentiates the HoloLens from many other AR devices, but also from wider adopted Virtual Reality (VR) headsets, is the fact that it has an independent computing unit with its own CPU and GPU. While most other devices require an additional computer (PC, server or smartphone), the HoloLens is able to run applications on its own. A common way to develop such applications is the popular multi-platform game engine Unity[1], which can easily be compiled for the HoloLens. However, as a mobile device, the HoloLens has comparably poor computing power and strong memory limitations [6]. In contrast to that, AR applications and tasks, as for example medical visualization or rendering, tend to become more and more advanced and complex. One way to approach this discrepancy is to reduce the data to be visualized for example by means of mesh simplification [6]. However, for medical applications this approach might be problematic. First, further data manipulation must be carried out extremely careful to not change the content in this context. Second, volume rendering is used quite often, which demands high amounts of GPU memory and can not benefit from mesh reduction. Another common solution to cope with the limited hardware of the HoloLens is to stream the whole content, or parts of it from another computer to the headset, thus moving closer to a setup that resembles lightweight headsets such as the HTC Vive which only serve as a display. This strategy also opens the possibility to reuse existing visualization technology outside of the Unity environment. Streaming, however, can again imply intense computations since it often requires sophisticated compression and decompression algorithms or codecs. At the same time, high framerates and low latencies are crucial for AR in order to not break immersion and allow proper user interaction. Existing or proprietary developed streaming solutions need to fulfill further requirements in order to allow for integration into the overall project pipeline. On the server side, integration into the server must be able from a purely technical standpoint. Common hurdles in this area are mismatching build systems and software stacks which have too many dependencies. On the client side, it might be preferable to still use the Unity engine, which is often used to implement user interface elements and interaction.
Therefore, the question arises, which the best suited codec or framework among the various available codecs

[1] `https://unity.com/`

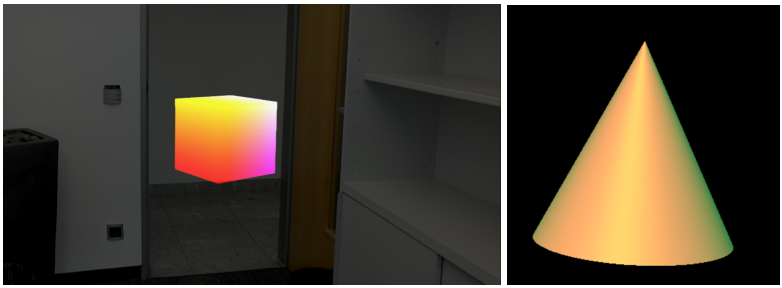


Figure 1: *The content that was used for testing the different streaming approaches. The left image shows the view inside the HoloLens from the user's point of view. The cube was used for testing the 3DStreamingToolkit and HoloLens Remoting. The cone on the right side was used for testing JPG and DXT1. Both objects were animated during the streaming.*

and solutions for streaming to the HoloLens is. We try to answer this question partially by testing a number of streaming solutions for the HoloLens. In this context, we are comparing the JPG and the DXT1[2] image formats, the H.264 video codec [7] and Microsoft's official streaming solution called "HoloLens Remoting".

## 2 Material and Methods

In order to determine the most efficient and best suited streaming solution for the HoloLens, we tested the mentioned codecs and frameworks on the same machine and hardware setup. Since we had a very powerful machine that was capable of encoding most of the codecs in real-time and a simple single-client application in our project context, we mainly focused on the decoding time on the HoloLens. But since mobile devices and clients are mostly not as powerful as the server side, we can assume that the client usually is the bottleneck for single-client applications which makes decoding timings more important than the encoding timings.

To determine the most efficient streaming solution, we measured the timings on the HoloLens for decoding and rendering the received image. We tested two scenarios: single images with 512 by 512 pixels as well as stereo-images with a resolution of 1280 by 720 pixels per eye in the respective codec. All images were rendered as RGB-images with 24Bit color depth. The content of the streamed images that was used for testing is depicted in fig. 1. Since we partially used the examples that were provided by the frameworks to test the approaches, the used image content is not completely identical, but similar enough that it should not make a measurable difference for the timings. The resolutions of the images were equal for each tested codec, however.

Furthermore, we measure the latency between the request of the image and the actual rendering on the HoloLens. Hence, this latency consists of the decoding time, the time for transferring the image data over the network and an assumably constant network respond latency. As stated above, we tried to exclude the encoding from the timings, since we want to focus on the decoding. Only for the H.264 codec, we could not determine an exact timing since it was encoded directly on the GPU but the introduced bias should be negligible according to the official Nvidia measurements[3].

### 2.1 Codec and implementation details

For testing the H.264 codec, we employed the open source 3DStreamingToolkit[4] which includes numerous server and client applications that should be compatible among each other. The communication between client and server applications is handled by a Signaling Server[5], which was running in a Docker on the desktop machine. The whole system is therefore prepared for a multi-client scenario, which we did not use in our project context yet, however. As a protocol, the toolkit creators chose the popular WebRTC[6] framework. The H.264 video

[2] https://en.wikipedia.org/wiki/S3_Texture_Compression
[3] https://developer.nvidia.com/nvidia-video-codec-sdk#NVENCPerf
[4] https://github.com/3DStreamingToolkit/3DStreamingToolkit
[5] https://github.com/3DStreamingToolkit/signal-3dstk
[6] https://webrtc.org/

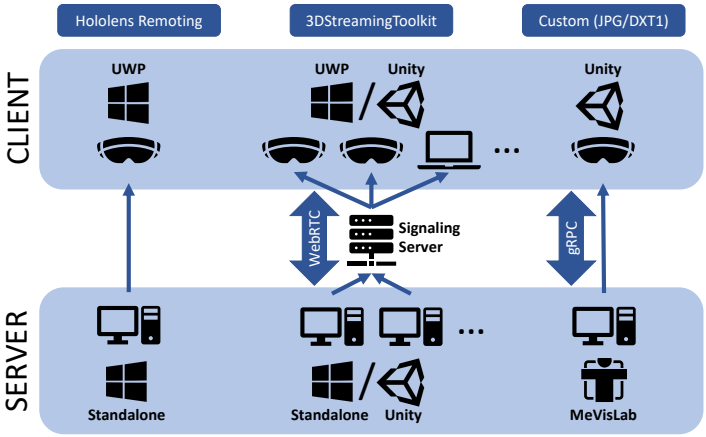



Figure 2: *Schematic overview over the tested streaming solutions.*

codec is efficiently encoded on the GPU using Nvidia's NVenc[7]. Therefore, a Nvidia graphics card is needed for efficient encoding. From the variety of servers and client applications we decided on the native DirectX-based C++ server implementation and the native UWP client application. Unfortunately, the Unity-based server and client applications did not work for us. Generally, the setup of the Toolkit was time-consuming and complex and it is dependent on many different software components, which makes it difficult to integrate in existing software projects and environments.

For the image formats DXT1 and JPG, we used custom implementations for both the server and client side. For the DXT1 format the server side was implemented in C++ in our *MeVisLab* [8] software environment using an open source Real-Time DXT compression library[8]. Since we streamed images without an alpha channel and the DXT1 codec can be encoded very fast, we used the DXT1 compression. The client applications were realised with the Unity Engine in C#. As the protocol for the communication between client and server, we employed Google's gRPC[9] for both codecs. For the JPG encoding, we used a quality factor of 90 percent.

HoloLens Remoting is the official streaming solution from Microsoft but the used codec and protocol are not published. For our tests we used the publicly available "Remoting Host Sample"[10] from Microsoft for the server side and the official "Holographic Remoting Player"[11] for the HoloLens. Since the Holographic Remoting Player is not open source, it is difficult to prove the timings that the application was giving us. According to Microsoft[12], the decoding time is measured as "the average number of rendered frames the remoting player is receiving and rendering per second" and the latency is "the average amount of time it takes for a frame to go from your PC to the HoloLens". To be consistent with the other timings in the table, we converted the framerate to milliseconds and computed the latency to also contain the decoding timing. The various streaming solutions are summarized in fig. 2.

## 3 Results

Table 1 is showing the results of our tests. The used test system consisted of the server side, which is represented by a desktop PC (Intel i7-9800X CPU, 64GB RAM, NVIDIA RTX2080Ti graphics unit, running Windows 10 64Bit) and the Microsoft HoloLens as the client. The two components were connected via WiFi, using a TP-Link Nano Router with 300 Mbps. All the timings were taken over the period of a couple of seconds and then

[7] https://developer.nvidia.com/nvidia-video-codec-sdk
[8] https://github.com/Cyan4973/RygsDXTc
[9] https://grpc.io/
[10] https://github.com/Microsoft/MixedRealityCompanionKit/tree/master/RemotingHostSample
[11] https://www.microsoft.com/de-de/p/holographic-remoting-player/9nblggh4sv40?activetab=pivot:overviewtab
[12] https://docs.microsoft.com/en-us/windows/mixed-reality/holographic-remoting-player

the mean and standard deviation were calculated. To be consistent, we converted the data from framerates to durations where it was necessary.

Table 1: *The timings for the decoding on the HoloLens and the corresponding latency for the various streaming solutions. The given latency is the latency including decoding and transferring over the network, but without encoding. However, the marked entries (*) are measured with encoding, but the encoding time is most likely insignificantly small. All given values are the mean values with standard deviation.*

| **Resolution** | | **JPG** | **DXT1** | **H.264** | **HoloLens Remoting** |
|---|---|---|---|---|---|
| 512x512 | *Decoding & Rendering* $\overline{d}$ | 19 ± 3 ms | 3 ± 1 ms | 18 ± 2 ms | approx. 17 ms |
| | *Latency* $\overline{l}$ | 31 ± 4 ms | 18 ± 1 ms | 94 ± 17 ms* | approx. 32 ms |
| 2*1280x720 | *Decoding & Rendering* $\overline{d}$ | 123 ± 10 ms | 9 ± 3 ms | 20 ± 5 ms | approx. 20 ms |
| | *Latency* $\overline{l}$ | 153 ± 21 ms | 106 ± 16 ms | 104 ± 12 ms* | approx. 42 ms |

# 4 Discussion

Our results as presented in table 1 indicate that the DXT1 codec is by far the fastest solution among the tested variants when it comes to decoding on the HoloLens. This is not surprising since DXT1 was designed in such fashion that it can be efficiently decoded on the graphics hardware. Even for high resolution stereo images, the decoding is still very efficient, even on the comparably weak HoloLens hardware. In comparison to that, the JPG format is significantly slower in decoding and the timings increase faster for higher resolutions as well. While the difference between the decoding times for high resolutions is immense, the overall mean latency $\overline{l}$ does not show such a big difference for the two image formats. This is likely due to the fact that DXT1 is not compressing the data as good as JPG and therefore the data transfer over the network takes longer so that the network transfer becomes the bottleneck at some point. In contrast to that, the H.264 codec shows comparably or almost equal timings for both resolutions. This might be due to the fact that the 3DStreamingToolkit employed a Virtual Sync at 60 Hertz and thus capped the framerate. For H.264, the timings also include the encoding since it was difficult to measure the exact timings on the GPU but according to Nvidia, the encoding timings should approximately be just a few milliseconds on the used graphics card[13] and therefore should not introduce a significant bias. Moreover, H.264 has an efficient GPU-based encoder called NVenc although the 3DStreamingToolkit, which we used for testing, was very difficult to set up and had a lot of other dependencies. HoloLens Remoting also shows reasonable timings for the decoding, quite similar to the results of the H.264 codec. The latencies are even better than the ones from its competitors, but they are so close to the timings for the decoding that it is highly questionable if the given latency really contains all of the required steps. Because the client application is not open source, it is hard to verify the timings and also to integrate HoloLens Remoting into an existing project or software environment.

In our tests, we neglected the influence of the images onto the measured timings. However, depending on the used compression algorithms the image content might have a measurable impact. This should be investigated in future work. Also, we concentrated on transmission and decoding, mainly because we used very potent hardware for the server side. Furthermore, our own implementation for DXT1 encoding is prototypic and only uses the CPU. Once a GPU implementation is available, the encoding should be benchmarked as well, also using less powerful hardware, since our setup cannot be assumed to be always available. We did this whole comparison at a point in time that is questionable as the new HoloLens 2 has just been announced by Microsoft. Still, we needed to find a solution for the HoloLens 1 as it might still take months until the first units of the new version are shipped. Furthermore, the HoloLens 1 is still an expensive and capable piece of hardware, despite the mentioned shortcomings and we plan to support the existing units in the future. However, once the new version is available, an update or extension of our analysis should be carried out.

[13] `https://developer.nvidia.com/nvidia-video-codec-sdk#NVENCPerf`

# 5 Conclusion

In this work we compared several solutions for streaming content to the Microsoft HoloLens. Each of the tested methods has their individual strengths and weaknesses. Our preliminary results indicate that H.264 seems to be a good choice in most of the cases and showing a good trade-off between performance and end-to-end latency, especially when the resolution is getting higher. However, this has to be proven in future work as well. DXT1 is by far the fastest solution when it comes to decoding and should be preferred for lower resolutions and when enough power for encoding is available. In contrast to that, JPG shows both poor decoding performance and latency timings, which makes it not well suited for streaming purposes. HoloLens Remoting shows quite good results but due to its closed source client application it is hard to integrate into existing software environments.

As future work we propose to test more codecs and formats on the Microsoft HoloLens and compare them to the ones mentioned here. In particular, the new H.265 codec[14] seems very promising and we would like to investigate the performance on the HoloLens. In addition to that, we will also extend our analysis to include the encoding performance of the different codecs and formats. Finally, it will be interesting to also test the respective codecs and formats on the new HoloLens 2 once it is available.

# 6 Acknowledgement

This work was funded by the German Federal Ministry of Education and Research (BMBF): project VIVATOP (funding code: 16SV8078) and project Nav EVAR (funding code: 13GW0228).

# 7 References

[1] Khor W.S., Baker B., Amin K., Chan A., Patel K., Wong J., *Augmented and virtual reality in surgery-the digital surgical environment: applications, limitations and legal pitfalls*, Ann. Transl. Med., 4(23), 454 (2016).

[2] Tepper O.M., Rudy H.L., Lefkowitz A., Weimer K.A., Marks S.M., Stern C.S. and Garfein M.D., *Mixed reality with HoloLens: where virtual reality meets augmented reality in the operating room*, Plastic and Reconstructive Surgery, 140(5), 1066-1070 (2017).

[3] Vavra P., Roman J., Zonča P., et al., *Recent Development of Augmented Reality in Surgery: A Review*, Journal of Healthcare Engineering, vol. 2017 (2017).

[4] García-Vázquez V., von Haxthausen F., Jäckle S., Schumann C., Kuhlemann I., Bouchagiar J., Höfer A., Matysiak F., Httmann G., Goltz J. P. et al. *Navigation and visualisation with HoloLens in endovascular aortic repair*. Innovative Surgical Sciences, 3(3), 167-177, De Gruyter (2018).

[5] Reinschlüssel A.V., Münder T., Uslar V.N., Weyhe D., Schenk A., Malaka R. *Tangible Organs: Introducing 3D Printed Organ Models with VR to Interact with Medical 3D Models*. Conference on Human Factors in Computing Systems (CHI), 2019.

[6] Bahirat K., Lai C., Mcmahan R., and Prabhakaran B. *Designing and Evaluating a Mesh Simplification Algorithm for Virtual Reality*. ACM Trans. Multimedia Comput. Commun. Appl., 14(3), Article 63 (2018).

[7] Wiegand T., Sullivan G. J., Bjontegaard G. and Luthra A., *Overview of the H. 264/AVC video coding standard*. IEEE Transactions on circuits and systems for video technology, 13(7), 560-576 (2003).

[8] Ritter F., Boskamp T., Homeyer A., Laue H., Schwier M., Link F., Peitgen H.-O., *Medical Image Analysis: A visual approach.*, IEEE Pulse, 2(6), 60-70 (2011).

[14] `http://hevc.info/`

# Effects of Accuracy-to-Colour Mapping Scales on Needle Navigation Aids visualised by Projective Augmented Reality

*F. Heinrich[1], F. Joeres[1], K. Lawonn[2], C. Hansen[1]*

[1] *University of Magdeburg, Faculty of Computer Science & Research Campus STIMULATE, Germany*
[2] *University of Koblenz-Landau, Faculty of Computer Science, Germany*

Contact: hansen@isg.cs.uni-magdeburg.de

***Abstract***

*Instrument navigation in needle-based interventions can benefit from augmented reality (AR) visualisation. Design aspects of these visualisations have been investigated to a limited degree. This work examined colour-specific parameters for AR instrument navigation, that have not been successfully researched before. Three different mapping methods to encode accuracy information to colour and two colour scales varying different colour channels were evaluated in a user study. Angular and depth accuracy of inserted needles were measured and task difficulty was subjectively rated. Result trends indicate benefits of mapping accuracy to discrete colours based on thresholds and using single hue colour scales that vary in the luminance or saturation channel. Yet, more research is required to validate the exposed indications. This work can constitute a valuable basis for this.*

**Keywords**: projective augmented reality, needle navigation, intra-operative visualisation, colour mapping

## 1 Problem

Surgical navigation systems that guide tracked instruments towards their target positions have beneficial effects on needle-based minimally invasive interventions, such as improved targeting accuracy, decreased procedure time and less required imaging scans [1], [2]. Augmented reality (AR) can be used to visualise this navigation information directly on the patient, thus overcoming issues of commonly used monitor-based systems, e.g., increased mental load, complicated hand-eye coordination and disrupted attention to the patient [3], [4].

Diverse examples of such AR instrument navigation systems can be found in the literature. These system use different display modalities to convey navigation information, like video see-through monitors [5], optical see-trough glasses [6] and projection systems [4]. Most publications in the field primarily focus on the technical realisation and evaluation of navigation systems, while visualisation aspects are only marginally discussed. Yet, a few works explicitly address this topic. Seitel et al. [7] compared four navigation concepts presented on a monitor. Chan and Heng [8] examined different access path visualisations and Mewes et al. [9] investigated effects of two projected visualisation concepts on needle insertions inside an MRI scanner. In previous work [10], we analysed existing visualisation approaches to support needle navigation tasks and compared three projected concepts in terms of accuracy measures and subjectively perceived task difficulty. We also varied between two methods of indicator scaling (i.e. how changes in accuracy translated to changes in the visualisation concepts shape) and three different accuracy-to-colour mapping schemes. However, no clear conclusions could be drawn regarding the latter factor due to a considerably large amount of data that needed to be excluded from the analysis. This topic has also not been discussed elsewhere, to the best of our knowledge. Thus, there is still need for an extensive investigation of different accuracy-to-colour mapping methods for AR needle navigation. This work presents a user study that addressed this topic. Building on our previous results [10], three different colour mapping methods and two colour scales to convey accuracy information for projective AR needle navigation were compared using similar measures.

## 2 Material and Methods

To achieve comparable results to [10], a similar experimental apparatus was implemented. The work carried out a three factor test to evaluate navigation visualisations. These factors were the general navigation visualisation concept, the indicator scaling method and accuracy-to-colour mapping. To focus more on colour-specific visualisation aspects, the concept and indicator scaling factors were held constant for this experiment. Instead, different accuracy-to-colour mapping methods have been investigated. Additionally, the underlying colour scale was varied.

## 2.1 Apparatus

Out of our previously examined concepts, we determined a crosshairs shaped visualisation to be the most effective in terms of accuracy, task completion time and subjective difficulty ratings [10]. Moreover, similar concepts are already implemented by navigation systems in surgical use [1]. Therefore, this concept was used to support needle navigation in this experiment. Figure 1 shows the visualisation's functionality. The needle handle position is projected onto a circular grid in the form of a small crosshairs marker. Angular accuracy is visualised by the distance between the marker and the centre of the grid. For perfect needle orientation, both are aligned. Depth information is visualised by a circular filling of the grid. The filling's diameter represents depth accuracy. The correct insertion depth is reached when the diameter matches with the outer circle of the grid.

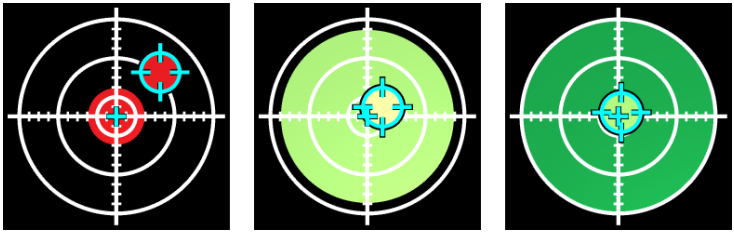


Figure 1: Needle navigation visualisation. A crosshairs grid is projected onto the injection site. A smaller crosshairs marker (blue) represents the needle handle position and the filling of the grid represents insertion depth information. From left to right the insertion angle is steadily improved and the needle is inserted further. Colours encode accuracy information.

The movement of the depth and angle indicators (i.e. filling diameter and marker-to-centre distance) were scaled logarithmically in this experiment. In [10], we compared a linear and and a logarithmic scaling method and concluded that the logarithmic method yielded higher depth accuracy results. Therefore, we implemented a similar method for this experiment, as well. This resulted in more sensitive indicator changes at higher than at lower accuracy levels and thus caused a higher resolution when precision is required.

To get a deeper insight into accuracy-to-colour mapping methods, this factor was adopted from [10]. Hence, a discrete and a continuous colour mapping method were implemented. Figure 2 shows the differences between both alternatives and depicts the threshold levels used for colour changes. Besides a red-yellow-green colour scale using a traffic lights metaphor, a green single hue scale was implemented. The first variant was adopted from [10]. However, the hue channel is best used to differentiate categorical attributes. Accuracy values rather represent quantitative ordered data which is better encoded by the saturation or luminance channel of single hue colour scales [11]. Therefore, the second variant was implemented, as well. Both scales are based on colour maps from *www.ColorBrewer2.org* [12]. In case of colour vision deficiency, scales using blue hues instead of green were provided.


Figure 2: Accuracy-to-colour mapping methods and colour scales used to encode angular and depth accuracy information. Continuous and discrete mapping methods were investigated for each colour scale. Top scale varies by hues and the bottom scale varies by luminance. Thresholds for accuracy levels are reported at the separator lines.

The needle navigation aids have been implemented using the game engine *Unity* (Unity Technologies, USA)

and were displayed by a projector-camera-system. Three Barco F22 WUXGA Digital Light Processing projectors (Barco GmbH, Germany) were mounted above a table and aligned over a common projection area. The projectors were calibrated with the photogrammetric measurement system *ProjectionTools* (domeprojections.com GmbH, Germany). To obtain instrument position data, the IR-based optical tracking system fusionTrack 500 (Atracsys LLC, Switzerland) was installed facing the projection site. Coordinate systems of the projector-camera-system and the tracking camera were registered using an optical IR marker with a known transform in projection space. A mannequin phantom filled with candle gel served as the projection surface. Covered by a sheet of paper, the candle gel created skin-like haptic when inserting a needle. An overview of the setup can be seen in Figure 3.

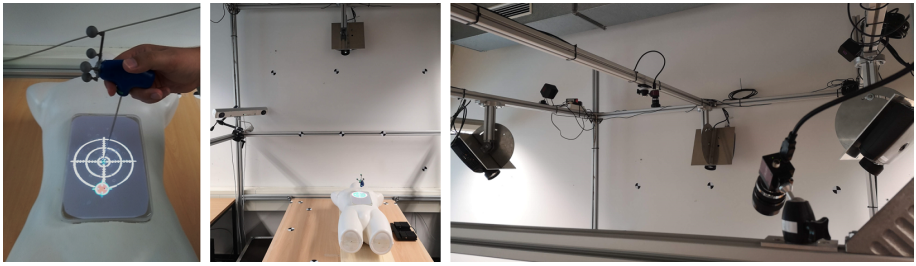


Figure 3: Experimental apparatus. Three projectors augmented a candle gel filled phantom with navigation visualisations. An IR-based optical tracking camera was used to obtain instrument position information.

## 2.2 Evaluation

During the experiment, participants were asked to repeatedly insert a tracked needle into the candle gel filled phantom. Each insertion was guided by the AR navigation visualisation and defined by preset insertion angle and depth parameters, that had to be reached as closely as possible. Participants were asked to keep an accurate insertion angle throughout the insertion process and not to over-insert the needle.

This task did not require specific clinical experience. However, a general understanding of the medical motivations may have been helpful. Therefore, we recruited twelve medical students (10 female, 2 male) for this study. Additionally, three students with a technical background and experience with AR (2 female, 1 male) participated in the experiment. The subjects were between their first and fifth year of university and their age ranged from 21 to 27 years old (median: 24 years). No participant reported colour vision deficiencies.

Two independent variables were regarded for this within-subject design study. First, the discrete and continuous accuracy-to-colour mapping methods were investigated as levels of the factor *Mapping Method*. Additionally, a monochromatic factor level was included. For this variant, angular and depth accuracies were always mapped to the same colour, which was set to the respective colour of the second accuracy level. Secondly, the two colour scales depicted in Figure 2 were considered as levels of the factor *Colour Scale*.

The dependent variables of this experiment were chosen and measured identically to [10]. Throughout each needle insertion, the angle between the current and the planned needle trajectory was calculated. Angular deviation was measured after each $1mm$ of insertion. At the end of each trial, these data points were averaged resulting in a *mean angular deviation*. *Absolute depth deviations* were measured as the absolute value of the difference between the current and the planned insertion depth at the end of each trial. Participants had control of the start and end times of trials by pressing a single-button switch near their standing position. The time in between was measured as the *task completion time*. After each needle insertion, participants rated the subjectively perceived difficulties to find the correct insertion angle and depth on a 6-point Likert scale (i.e. *subjective difficulty rating angle* and *subjective difficulty rating depth*).

Each participant performed a total of twelve needle insertions. Every factor combination between the two independent variables was evaluated twice per user. Before the actual trials began, a training session was conducted, where participants were instructed with the overall needle navigation concept, the colour scales and their tasks during the study. During this session, participants could freely practise the navigated insertion process. Afterwards, the twelve trials began. The order of presented alternatives was partially randomised. First, all six possible factor level combinations were presented in random order. After the completion of these trials, every alternative was presented a second time in the same order as before. By pressing the aforementioned

switch, a new trial began. Participants could then freely select a desired injection site and begin the insertion process. Planned insertion angle and depth were generated randomly for each trial with angles ranging from 0° to 30° around the perpendicular to the injection site and depths between 70$mm$ and 90$mm$. When participants felt confident to have completed the needle insertion, they pressed the switch again and rated the perceived task difficulty dimensions.

## 3 Results

The two factors' effects were analysed with two-way ANOVAs for all dependent variables. Results of the *Mapping Method* factor are summarised in Figure 4 and results of the *Colour Scale* factor are shown in Figure 5. No statistical significance could be shown except for the *Mapping Method* main effects on both subjective difficulty variables. However, the descriptive data reveal some trends worth noticing.

Regarding the *Mapping Method* factor, discrete accuracy-to-colour mapping seems to have yielded the least angular deviation but highest depth deviation. This may be due to the choice of accuracy thresholds and easier noticeable differences between final colours. The most accurate threshold may have been low enough to cause advantageous effects for angular accuracy, but may have been too great for depth accuracy (see Figure 2). Using the monochromatic and continuous methods achieved similar accuracy results. This may indicate that the navigation concept alone was helpful enough to accurately orient and insert needles. Yet, monochromatic accuracy-to-colour mapping was rated to be significantly more difficult compared the other factor levels, which was probably caused by a higher degree of confidence due to additional feedback. The *Colour Scale* factor plots suggest higher achievable accuracies when using the single hue scale. This is consistent with the general classifications of Munzner [11] and may confirm that luminance is a more suitable colour channel to convey accuracy information in projective AR needle navigation than hue.

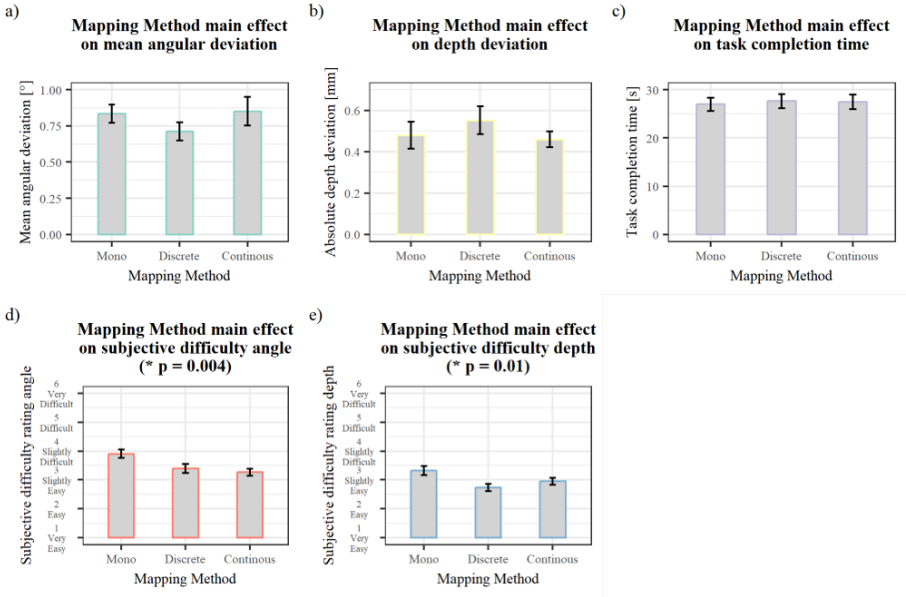



Figure 4: Main effects of the *Mapping Method* factor on: a) angular deviation, b) absolute depth deviation, c) task completion time, d) subjective difficulty angle, and e) subjective difficulty depth. (Error bars represent standard error. * denotes statistical significance.)

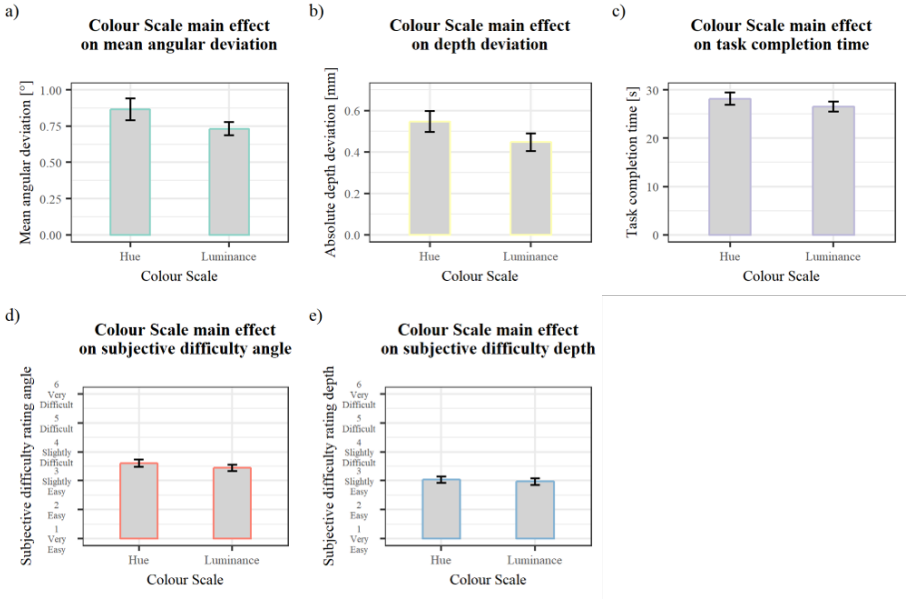



Figure 5: Main effects of the *Colour Scale* factor on: a) angular deviation, b) absolute depth deviation, c) task completion time, d) subjective difficulty angle, and e) subjective difficulty depth. (Error bars represent standard error.)

## 4 Discussion

Some statistically significant effects could be shown by the ANOVAs, although trends could be identified in the descriptive data. Various reasons may have contributed to this. The sample size of the study was rather small. Additionally, the overall navigation visualisation may have already contributed enough information to achieve high accuracy and therefore may have caused a ceiling effect. This is supported by the low angular and depth deviation results for the monochromatic factor level.

Results for the *Mapping Method* factor suggest, that colour is a useful tool to convey additional accuracy feedback and thus increase confidence of correctly performed needle insertions. Moreover, the position of final accuracy level thresholds seemingly influences the insertion precision for discrete colour mapping and may cause higher accuracy than monochromatic or continuous methods. This needs to be further investigated in a future experiment, especially because no statistically significant effects could be shown, yet.

The *Colour Scale* factor yielded results that are consistent with the literature. However, no clear conclusions can be drawn regarding this factor because no statistically significant effects could be shown for this factor either. The trends visible in the descriptive data may also be due to differently perceived navigation support from both monochromatic colour scales, which is indicated by the raw data to a certain degree. The monochromatic colour of the luminance scale (light green) may have been easier to differentiate from crosshairs grid (white) compared to the monochromatic colour of the hue scale (light yellow). Therefore, more research is required to find definitive answers to the question which colour scale yields the best results.

The selected colour scales may have influenced the experiment, as well. Different colours may have been easier to differentiate. For example, more nuances may have been noticeable on a red single hue scale, thus creating benefits for a continuous mapping method. High contrast between colours is especially important for projection systems, that are known to be sensitive to this property. Future research should focus on finding and evaluating suitable colour scales for this display modality.

Moreover, colour anchors in this work were not evenly distributed along the accuracy intervals, which resulted in non-linear colour scales. This was done to facilitate a higher colour resolution at higher accuracy levels, but

may have ultimately influenced the experiment. More research could therefore be conducted to explore effects of linearisation of colour scales. This could also include an investigation of how including more discrete colour steps influences needle insertion accuracy.

## 5 Conclusion

This work investigated colour-specific visualisation parameters for projective AR instrument navigation. These included different methods of accuracy-to-colour mapping and variations of the channel that was used to create a colour scale. The factors were evaluated in a user study with simulated needle insertion tasks. Measured variables were objective performance parameters (angular and depth accuracy ans task completion time) and subjectively perceived task difficulty. Results suggest that discrete accuracy-to-colour mapping has to potential to convey accuracy information the most effectively. This highly depends on the choice of thresholds used for colour mapping. Moreover, for instrument navigation visualisations, colour scales varying the luminance or saturation channel seem to be more expressive than colour scales varying the hue channel. However, extended research is required to find specific answers to the identified open questions regarding instrument navigation visualisation. This work can constitute a solid basis for this.

## 6 Acknowledgements

This project was funded by the DFG (HA 7819/1-2 and LA 3855/1-2).

## 7 References

[1] J. Engstrand, G. Toporek, P. Harbut, *et al.*, "Stereotactic CT-Guided Percutaneous Microwave Ablation of Liver Tumors With the Use of High-Frequency Jet Ventilation: An Accuracy and Procedural Safety Study," *American Journal of Roentgenology*, vol. 208, no. 1, pp. 193–200, 2017.

[2] R. F. Grasso, E. Faiella, G. Luppi, *et al.*, "Percutaneous lung biopsy: Comparison between an augmented reality ct navigation system and standard ct-guided technique," *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, no. 5, pp. 837–848, 2013.

[3] D. Manzey, S. Röttger, J. E. Bahner-Heyne, *et al.*, "Image-guided navigation: The surgeon's perspective on performance consequences and human factors issues," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 5, no. 3, pp. 297–308, 2009.

[4] K. Gavaghan, T. Oliveira-Santos, M. Peterhans, *et al.*, "Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: Phantom studies," *International Journal of Computer Assisted Radiology and Surgery*, vol. 7, no. 4, pp. 547–556, 2012.

[5] A. Seitel, N. Bellemann, M. Hafezi, *et al.*, "Towards markerless navigation for percutaneous needle insertions," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 1, pp. 107–117, 2016.

[6] C. A. Agten, C. Dennler, A. B. Rosskopf, *et al.*, "Augmented Reality-Guided Lumbar Facet Joint Injections," *Investigative Radiology*, vol. 53, no. 8, pp. 495–498, 2018.

[7] A. Seitel, L. Maier-Hein, S. Schawo, *et al.*, "In-vitro evaluation of different visualization approaches for computer assisted targeting in soft tissue," *International Journal of Computer Assisted Radiology and Surgery*, vol. 2, S188–S190, 2007.

[8] W. Y. Chan and P. A. Heng, "Visualization of needle access pathway and a five-DoF evaluation," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 643–653, 2014.

[9] A. Mewes, F. Heinrich, B. Hensen, *et al.*, "Concepts for augmented reality visualisation to support needle guidance inside the MRI," *Healthcare Technology Letters*, vol. 5, no. 5, pp. 172–176, 2018.

[10] F. Heinrich, F. Joeres, K. Lawonn, *et al.*, "Comparison of projective augmented reality concepts to support medical needle insertion," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 6, pp. 2157–2167, 2019.

[11] T. Munzner, *Visualization analysis and design*. AK Peters/CRC Press, 2014.

[12] C. A. Brewer, M. Harrower, and The Pennsylvania State University. (2013). Colorbrewer: Color advice for maps, [Online]. Available: `http://www.colorbrewer2.org` (visited on 06/17/2019).

# Self-Localized Multi-Projector Systems for Surgical Interventions: A User Study

*Shivaraman Ilango, Hardik Dava, Tobias Ortmaier*

*Institute of Mechatronic Systems, Leibniz Universität Hannover, Hannover, Germany*

Contact: shivaraman.ilango@imes.uni-hannover.de

***Abstract***

*Spatial augmented reality (SAR) systems can be used to visualize preoperatively planned surgical information (incision lines, drilling marks etc.) directly onto a patient during surgery. In our prior work, a prototype of a self-localized hand-held SAR system consisting of a pico projector and a stereo camera has been developed for liver surgery applications and evaluated using various user studies and experiments. However, this hand-held system has major disadvantages in terms of illuminance and luminous flux. The visibility of the projected virtual information is very low especially under surgical lights. In this paper, two more self localized video projectors are used along with the hand-held system to increase the visibility of the projected information by superimposing all the three projector systems. The superimposed projectors are also used to provide additional 3D information for localization and orientation (6DoF). Further, view-dependent blurring and dispersion effects caused while projecting onto inhomogeneous surfaces are compensated using these multi-projector systems. In the end, a user study is conducted with 8 test persons to analyze the visibility of the projected information under different surgical light intensities.*

**Keywords**: Spatial augmented reality, self-localization, illuminance, multi-projector, visibility contrast, view-dependent projection

## 1 Problem Statement

Spatial augmented reality systems can replace head mounted displays (HMDs) for most augmented reality (AR) applications [1]. In medical augmented reality, the use of HMDs could restrict the surgeons peripheral vision and motion [2]. On the other hand, SAR systems do not require the use of any special displays and since the display is not associated with each user, it can scale up to group of users, thus allowing collaboration between them. Today, an increasing trend towards projected-based displays for medical augmented reality applications can be noticed. A variety of stationary (e.g., [3, 4, 5], movable (e.g., [6, 7]) and hand-held (e.g., [8, 9]) projectors have been proposed for displaying preoperative surgical information obtained from medical imaging devices directly onto the organs or skin of the patients. These approaches require the use of artificial markers and/or external navigation system for intraoperative registration or tracking. Furthermore, in all the above methods the surgical lights had to dimmed or switched off since the projected information is not visible under these lights.

Another disadvantage of spatial augmented reality in general, is the reduction of dimensionality from 3D to 2D while projecting onto the objects. The points that are not within the illuminated surface cannot be augmented [10]. The methods in [10] describe how a pose of a bone segment could be defined with the help of two projectors.

Finally the SAR systems have issues in terms of color dispersion and blur effects while projecting the virtual information onto complex 3D surfaces with varying depth (seen in Fig.1). This could be solved by using structured light in combination with an integrated CCD sensors. However for projecting onto varying depths, the focus of every single pixel has to be measured and adjusted individually [1]. To overcome this problem, laser projectors are used instead of conventional projectors for always displaying focused images on curved surfaces. However, owing to its cost and safety high illuminance laser projector can only be used in surgery with high safety requirements. Another drawback with using the laser projector is the speckle noise, which is displeasing for a human observer and needs to be reduced to an acceptable level [11]

SAR systems cannot be effectively used for surgical interventions if the following disadvantages are not addressed. Particularly the disadvantage in terms of visibility under surgical light is a major concern. User studies conducted by our group [12] have provided us promising insights and results about preliminary specifications to design a hand-held projector that has an illuminance more than that of the surgical lights. Another approach for solving this issue is to use multi-projector systems that can display seamlessly blended images resulting in brighter superimposed projections under the surgical lights. Further these multi-projector systems can provide additional 3D information for repositioning an organ with respected to the planned virtual information.

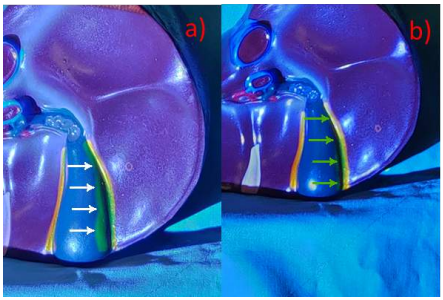



Figure 1: *a) Color dispersion when an incision line is projected on the gall bladder from extreme left view (white arrows) b) Corrected projection using seamless blending from multi-projector systems (green arrows)*

State-of the-art multi-projector approaches are able to display superimposed images onto geometrically and radiometrically complex surfaces [13,14]. However its use in medical interventions is not an area that has been exploited so far. There are few works [10,15,16] that use multi-projectors for medical interventions. However [15,16] deal with only volumetric display using multiple projectors without any emphasis on computer assisted surgical interventions. Our self-localized multi-projection system concept is inspired from [10] where Kahrs et al. used two video projectors to project preplanned surgical information in order to provide 3D information to reposition a bone segment for Cranio-Maxillofacial Surgery. However tracking and registration is realized using artificial markers and optical tracking systems which could be inconvenient as it requires undisturbed free sight from the navigation cameras to the operation area.

In this paper, we develop and study the concept of using self-localized multi-projector systems for surgical interventions. As an evaluation scenario we have chosen an artificial liver phantom and augmented reality incision lines to guide the surgeons during interventions. Further the illuminance of the projectors and visibility contrast of the projected information under the surgical lights is analyzed using different techniques.

## 2 Materials

This section describes the materials used in preforming this study. The study was conducted in an experimental operation room (OR). The experimental setup consists of a the hand-held projection system (AAXA L2 Pico Projector & a stereo camera) and two other stationary video projectors (EPSON EB-W41) to project the planned information (incision lines) on an artificial liver made of silicon under the surgical lights. The combined illuminance of all the three projectors at a distance of 60cm was approximately 18,000 lx. The illuminance of the surgical lights (2 x Polaris 560 DC Dräger) was measured to be approximately 100,000 lx at a distance of one meter. The illuminance was measured using a standard lux meter (Gossen GmbH). The silicon liver is used with red paint pigments to mimic the properties of the human liver in terms of reflection and occlusions. The experimental set up is shown in Figure 2.

## 3 Methods

The algorithms developed for the hand-held projection system [12] is extended for usability with new setup consisting of two additional projectors.

The projectors were calibrated in the same coordinate frame viewing the same calibration object (checker board pattern) using structured light approach [17]. This helps to locate the organ with its orientation in 3D. Before the visualization of the planned information, registration of the pre-operative model (obtained from segmentation of the computer tomography (CT) images) and intra-operative (obtained via structured light reconstruction) scene is performed. This process would help in transferring the planned data to the area of interest by finding the corresponding transformation from the planning system to the patients current position. The intra-operative surface reconstruction is performed by using active optical triangulation, where a sequence of gray code pattern projected is decoded by the camera and using local homogrpahies, the pixels between the camera and the projector are mapped to find the depth. The intra-operative scene model of the liver obtained

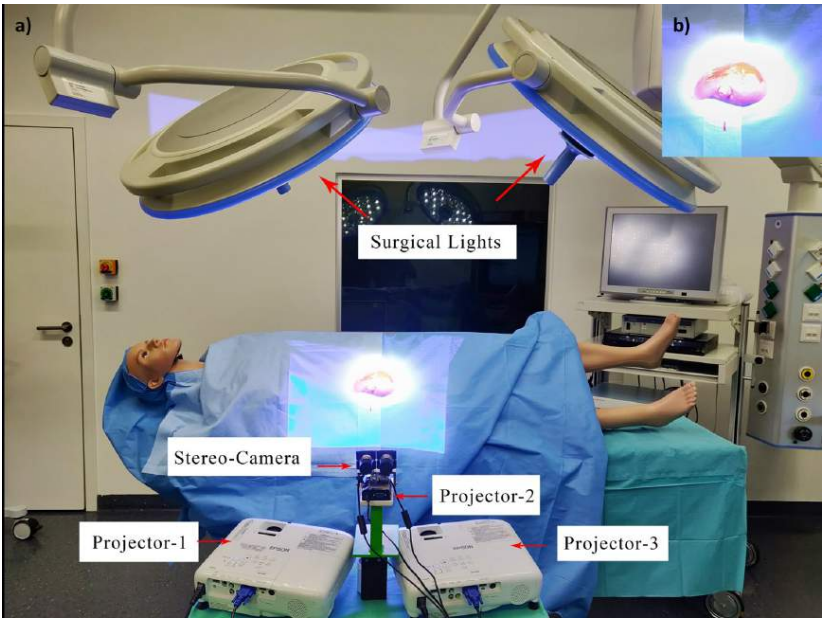



Figure 2: *a) Experimental setup of the study. b) Silicon liver phantom*

from all the three projectors are registered to the pre-operative model via the the Iterative Closet Point (ICP) algorithm in combination with the Fast Point Feature Histograms (FPFH) method for coarse registration [18].

After registration, the projection of planned information along with tracking is initialized. The tracking is based on ORB-SLAM [19]. However to reduce computational cost, latency and facilitate real-time tracking concurrent processing is done using CUDA framework and a GeForce 745 GPU (NVIDIA, Santa Clara, CA, USA). The three projectors and the stereo camera system are placed on a moving platform rigidly fixed with each other. The systems could be moved according to the user's convenience after the tracking has been initialized.

A graphical user interface was developed using Qt libraries to register, track and visualize the planned data on the liver during the user study. Incision lines with different thickness (1mm, 3mm) are projected onto the liver and checked for visibility under the surgical lights. The virtual information projected always has a blue background and yellow foreground as this color combination was considered the best visible under the surgical lights in the user studies performed in [9].

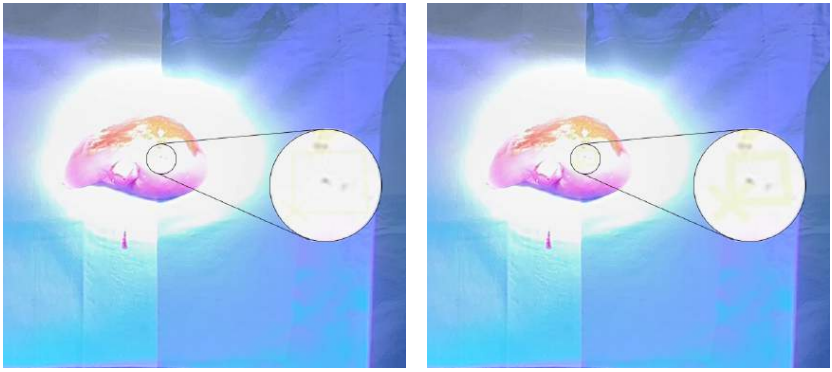
The user study was performed under different surgical light illuminance to help the users assess the visibility of the projected information at different levels. The controlled illuminance of the surgical lights at different brightness levels are shown in Table 1.

| Surgical Light Illuminance (lux) | Brightness Levels (BL) |
|---|---|
| 100,000 | BL5 |
| 85,000 | BL4 |
| 61,000 | BL3 |
| 42,000 | BL2 |
| 19,000 | BL1 |

Table 1: Surgical light illuminance at different brightness levels.

# 4 Results

This section summarizes the results of the user study which is based on the results of eight test persons. The test persons are between the age of 25 and 35. The entire algorithmic pipeline was run on an Intel i5-6500 3.2GHZ, 16GB RAM desktop computer. The planned data after registration is projected on the liver with different thickness as mentioned in Section 3. The redundant projection under the surgical light for 1mm and 3mm projection is shown in Figures 3a & 3b respectively. It can be seen that even with all the projectors switched on the projection is barely visible. However, the visibility is better when compared to only using the hand-held system for projecting the virtual information under the surgical lights. The projection with 3mm thickness is slightly better visible than that with 1mm (almost not visible). Hence the vessel structures/planned information having a thickness below 3mm might not be shown to the surgeons during surgery.

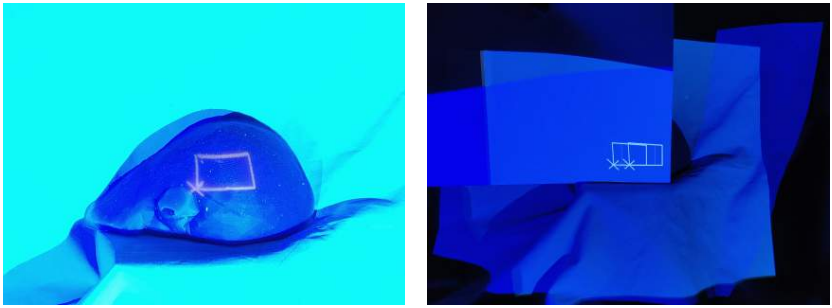


(a) (b)

Figure 3: Redundant projection with (a) 1 mm thickness at at full surgical light intensity (BL5); (b) 3 mm thickness at full surgical light intensity (BL5)

Figures 4a & 4b show the same projections with the surgical light switched off. Figure 4a shows the lines on the liver superimposed and in correct position. Figure 4b shows the lines not overlapping in a correct way. The paper sheet symbolizes the position of the planned incision.

An incision is possible below the correct coincidence of the cross-hairs and the system can be moved intuitively in the direction of the true location. Here instead of using the hand-held projection system, another stationary projector with high illuminance could also be used. Occlusions that occur by the surgeon or other instruments could be avoided with projectors illuminating the scene from different angles.



(a) (b)

Figure 4: (a) Redundant projection with no deviation at BL0; (b) Projected deviation not at the right position (simulated with a paper)

Figures 5 and 6 summarizes the results of the user study done with 8 test persons. It can be seen from the table that at brightness levels 4 and 5, most of the users predicted that the projected incision lines of 3mm thickness was clearly visible under the surgical lights. However, the incision lines of 1mm thickness were barely visible for any of the users until the surgical lights were reduced to brightness level 1


Figure 5: Results of User Study (visibility contrast evaluation of the incision line having 1 mm thickness)

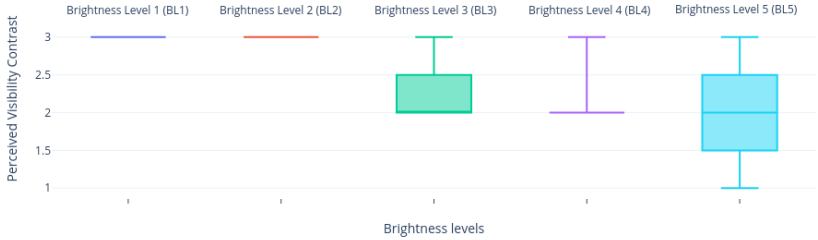



Figure 6: Results of User Study (visibility contrast evaluation of the incision line having 3 mm thickness)

## 5 Conclusion and Outlook

This paper introduces for the first time a study on self-localized multi-projector system for surgical interventions. With this approach it is possible to give additional 3D information to the surgeon. It can also be seen that this approach can be used to increase the visibility of the projected information under the surgical lights. The repositioning and localization of the liver model is also possible with the determination of 6DoF using multiple projectors. The next steps would be to simulate and build a high illuminance projection technology that could be used to provide brighter projections under the surgical lights.

# 6 Referenzen

[1] O. Bimber, G. Wetzstein, A. Emmerling, C. Nitschke, *Enabling view-dependent stereoscopic projection in real environments*, In Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR05), 14-36 (2005)

[2] K. A. Gavaghan, M. Peterhans, T. Oliveira-Santos, M. Reyes, H. Kim, S. Anderegg, S.Weber, *Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: phantom studies*, International journal of computer assisted radiology and surgery, **7(4)** 547-556 (2012)

[3] R. Krempien, H. Hoppe, L. A. Kahrs, S. Daeuber, O. Schorr, G. Eggers, M. Bischof, M. W. Munter, J. Debus, W. Harms, *Projector-based augmented reality for intuitive intraoperative guidance in image-guided 3d interstitial brachytherapy*, International Journal of Radiation Oncology* Biology* Physics, **70(3)** 944-952 (2008)

[4] J. R. Wu, M. L. Wang, K. C. Liu, M. H. Hu, P. Y. Lee, *Real-time advanced spinal surgery via visible patient model and augmented reality system*, Computer methods and programs in biomedicine, **113(3)** 869-888 (2014)

[5] T. Koishi, S. Ushiki, T. Nakaguchi, H. Hayashi, N. Tsumura, Y. Miyake, *A navigation system using projection images of laparoscopic instruments and a surgical target with improved image quality*, In SPIE Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling, **6918** 691810 (2008)

[6] N. Glossop, C. Wedlake, J. Moore, T. Peters, Z. Wang, *Laser projection augmented reality system for computer assisted surgery*, In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 239-246 (2003)

[7] J. P. Tardif, S. Roy, J. Meunier, *Projector-based augmented reality in surgery without calibration*, In Proceedings of the 25th annual international conference of the IEEE engineering in medicine and biology society, 548-551 (2003)

[8] P. Edgcumbe, P. Pratt, G. Z. Yang, C. Nguan, R. Rohling, *Pico lantern: surface reconstruction and augmented reality in laparoscopic surgery using a pick-up laser projector*, Medical image analysis, **25(1)** 548-551 (2003)

[9] K. A. Gavaghan, M. Peterhans, T. Oliveira-Santos, S. Weber, *A portable image overlay projection device for computer-aided open liver surgery*, IEEE transactions on biomedical engineering, **58(6)** 1855-1864 (2011)

[10] L. A. Kahrs, H. Hoppe, G. Eggers, J. Raczkowsky, J. Marmulla, H. Wörn, *Visualization of surgical 3d information with projector-based augmented reality*, Studies in health technology and informatics, **111** 243-246 (2005)

[11] G. Verschaffelt, S. Roelandt, Y. Meuret, W. Van den Broeck, K. Kilpi, B. Lievens, A. Jacobs, P. Janssens, H. Thienpont, *Speckle disturbance limit in laser-based cinema projection systems*, Scientific reports, **5** 14105 (2015)

[12] S. Ilango, M. Knöchelmann, L. A. Kahrs, A. Wolf, T. Ortmaier, R. Lachmayer, *User evaluation study on illumination requirements to design an augmented reality projector for open liver surgery*, In SPIE Illumination Optics V, 10693 (2018)

[13] O. Bimber, A. Emmerling, T. Klemmer, *Embedded entertainment with smart projectors*, Computer, **38(1)** 48-55 (2005)

[14] M. D. Grossberg, H. Peri, S. K. Nayak, P. N. Belhumeur, one object look like another: Controlling appearance using a projector-camera system, In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), **1** I-I (2004)

[15] H. Liao, N. Hata, T. Dohi, *Image-guidance for cardiac surgery using dynamic autostereoscopic display system*, In IEEE International Symposium on Biomedical Imaging:Nano to Macro, 265-268 (2004)

[16] W. Song, Q. Zhu, T. Huang, Y. Liu, Y. Wang, *Volumetric display system using multiple miniprojectors*, SPIE Optical Engineering, **54(1)** 013103 (2015)

[17] D. Moreno, G. Taubin, *Simple, accurate, and robust projector-camera calibration*, In Second International Conference on 3D Imaging, Modeling, Processing, Visualization Transmission, 464-471 (2012)

[18] J. Bergmeier, D.Kundrat, A. Schoob, L. A. Kahrs, T. Ortmaier, *Methods for a fusion of optical coherence tomography and stereo camera image data*, In SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, **9415** 94151C (2015)

[19] R. Mur-Artal, J. D. Tardos, *Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras*, IEEE Transactions on Robotics, **33(5)** 1255-1262 (2017)

# Life-like rendering of inner ear anatomy

*Catherine Tsai[1], Daniel Schneider[2], Jan Hermann[2], Fabian Mueller[2], Niklaus Heeb[1], Alessandro Holler[1], Simon Tschachtli[1], Gabriela O'Toole Bom Braga[2], Stefan Weber[2]*

*[1]Zurich University of the Arts, Knowledge Visualization, Zurich, Switzerland*
*[2]ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*

Contact: daniel.schneider@artorg.unibe.ch

***Abstract***

*Robotic cochlear implantation is a novel surgical technique that potentially improves surgical outcomes of cochlear implantation. The robotic platform requires patient-specific 3D visualizations generated from CT scans to plan and conduct the surgery. Current visualizations use rudimentary colouring and rendering parameters to visualize the surgical plan. Life-like rendering of anatomical and associated structures specific to robotic cochlear implantation was investigated. Microscopic and endoscopic images of ear surgery were analysed to optimize rendering parameters. Through more life-like visualizations, the user satisfaction could be improved.*

**Keywords**: 3D visualization of ear anatomy, surgical planning, robotic cochlear implantation

## 1 Problem

Five percent of the world is deaf or hard of hearing [1]. In severe to profoundly deaf cases, the cochlea - the organ that allows us to hear – is damaged and the cause of hearing deficits. Cochlear implantation (CI) is an otologic microsurgery procedure during which a hearing prothesis is inserted into the cochlea providing patients with hearing sensation [2]. Per year, some 65000 CIs are implanted globally. Because the surgery is at the limits of human visual and tactile skills, 30–50 % of the cases have suboptimal implant performance. Thus, robotic technology has been proposed to overcome human limitations, making the surgery more effective, less invasive, and safer for the patient [3], [4]. The phases of the procedure requiring robotic assistance include surgical planning and robotic drilling through. In the first step, the anatomy of the ear is segmented from a computed tomography (CT) scan and 3D visualizations are generated using a surgical planning software (OTOPLAN®, CAScination, Switzerland). Based on the 3D visualizations, a collision-free drill-trajectory is planned from the surface of the skull to the cochlea, passing the facial and gustatory nerves at distances <0.5 mm. Thus, for safe planning accurate and easy-to-understand 3D visualizations of the ear anatomy are essential. The second step involves robotic drilling of the planned trajectory and inserting the implant into the cochlea through the drill channel. Current surgical planning tools use rudimentary, unnaturally coloured palettes and default lighting and shading parameters to visualize the anatomical models (Figure 1). The aim of this project was to improve the user experience of robotic cochlear implantation surgery planning by manipulating visual properties. Specifically, rendering life-like 3D anatomical and associated structures specific to robotic cochlear implantation was investigated. We believe that with better user experience, user errors can be reduced, and hence patient outcomes improved.

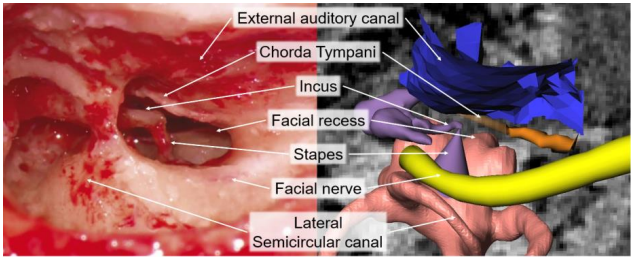



*Figure 1 Real colours and textures of the human ear as exposed during surgery (left) versus rudimentary, unnatural colouring and shading of the ear anatomy as it is currently used during robotic cochlear implantation procedures (right).*

## 2 Materials and Methods

Planning of robotic cochlear implant surgery can be conducted preoperatively and intraoperatively and thus operates most ideally on a tablet or other mobile computer. Due to the dynamic use of the visualizations, the renderings of the surgical plan must run in real-time, in this case on a tablet (Microsoft Surface Pro). Furthermore, during planning the surgeon needs to assess the surgical plan. Thus, the user interface to the 3D visualization must allow for interactive inspection.

A set of anatomical and associated structures specific to robotic cochlear implantation to be considered within this project was defined. The anatomical structures include the temporal bone, facial and gustatory nerves, ossicles, auditory canal, skin and labyrinth (comprising of the semicircular canals and the cochlea). The associated structures include the drill trajectory, implant, a skin-retractor and fiducial-screws which are required for accurate robotic drilling. After segmentation of individual structures from a patient CT scan data, 3D models were generated using a surgical planning software (OTOPLAN®, CAScination, Switzerland) and Amira (ThermoFischer Scientific, USA). According to visual design principles, the form, color, texture, transparency, reflectance and lighting of the structures were manipulated using Cinema4D in an iterative process with feedback on the visualizations from surgeons. Subdivision surface and Phong-shading algorithms were applied to smooth the appearance of the models without distorting measurements required for robotic trajectory planning; large deformations of the shape were avoided to retain safe planning. Microscopic and endoscopic images taken during ear surgeries were used to calibrate the color palette and lighting conditions. Cold-white spotlight aligned with the rendering camera was used to mimic the endoscope or microscope light in the surgery room. Color textures mimicking bone and skin surfaces were generated from the images and applied to the 3D models. Reflectance and glossiness were used to mimic the wet mucosa in the ear. Normal-maps were used to mimic the scattering of the light on the structures' surfaces. Subsurface scattering was applied to mimic the translucency of organic tissue. Furthermore, by color-coding the trajectory to indicate its proximity to the facial and gustatory nerves (green: >1 mm; red: <0.2 mm), the critical part of the surgery is highlighted.

During the design process, a web-based 3D modeling platform (Sketchfab, USA) was used to share selected versions of the scene with study participants to illicit feedback. A real-time capable 3D rendering framework (Qt3D, The Qt Company, Finland) was used to integrate the visualizations into OTOPLAN® to present the work in the context of robotic cochlear implantation. Finally, two surveys were conducted to test the effectiveness of the life-like visualizations compared to the rudimentary original models. In Study 1, a survey was administered to N=10 otolaryngologists. The surgeons were presented with interactive 3D scenes containing the original rudimentary, unnaturally colored visualizations and new life-like visualizations. Based on the visualizations, a questionnaire was administered to assess the user experience regarding the different responses to the two visual styles. In Study 2, N=6 in-person interviews with experienced otolaryngologists were conducted. The surgeons were asked to explore the original and new visualizations in OTOPLAN®. Opinions were obtained regarding words to describe each of the visualizations and overall experience of interacting with the 3D models.

## 3 Results

Lighting, shading, reflectance, glossiness and subsurface scattering parameter values, color textures and normal-maps were elaborated and documented. A rendering of the resulting models after the design process is shown in Figure 2. The 3D models were also integrated into the planning software for robotic cochlear implant surgery (OTOPLAN®). Study 1 revealed that user satisfaction is greatly increased with the life-like visualizations (10 of 10 surgeons completely agreed to the statement "I am more satisfied with the lifelike visualizations compared to the unnatural visualizations.". Study 2 showed that the life-like visualizations were associated with more positive words compared to the unnatural models (Figure 3).

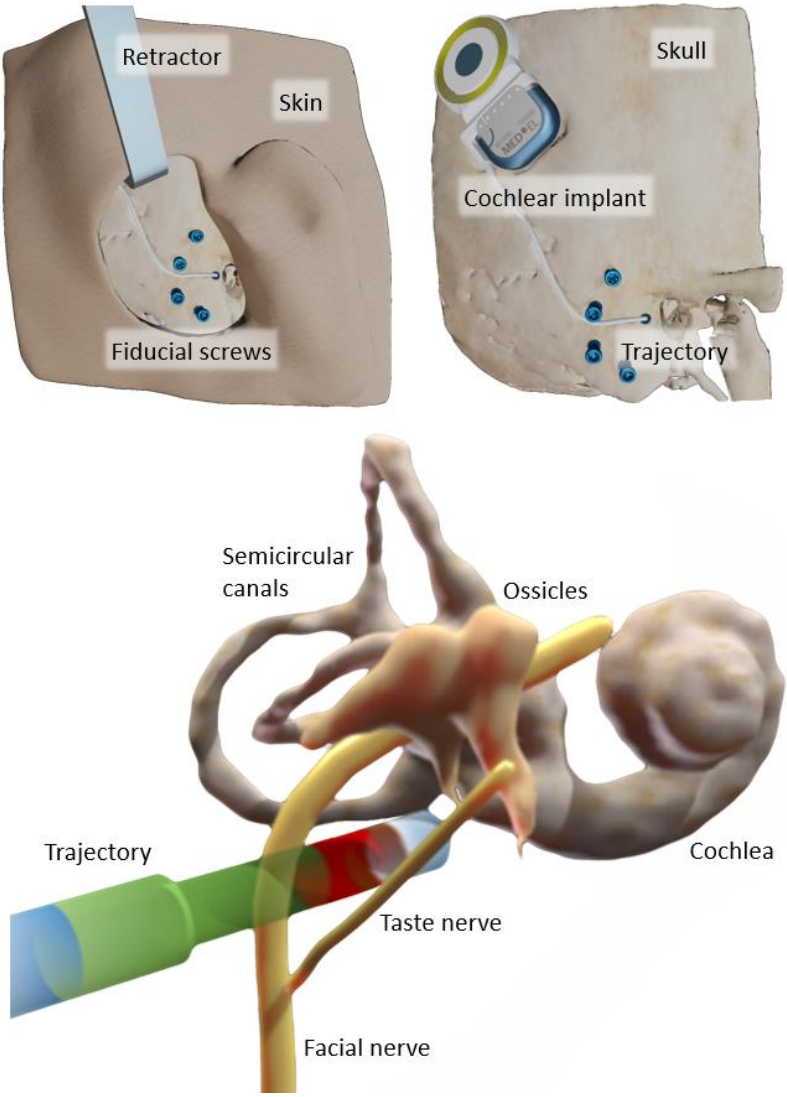



*Figure 2 Lifelike renderings of anatomical and artificial structures.*


*Figure 3 Words used by ear (or head and neck??***) surgeons to describe life-like (left) versus unnaturally coloured (right) visualizations of ear anatomy. Larger words were used more frequently.*

## 4 Discussion

This study highlights the importance of the visual design in positively impacting the user experience of a robotic surgery platform of the ear that must exist in real-time. While the technology behind many novel surgical platforms is highly advanced, some systems like the one used in this study are not fully realized and warrant further development. While unnaturally colored 3D-generated models of anatomy work well in cases in which one must distinguish between different components of the anatomy (for example highlighting a hepatic tumor in green against a naturally colored or grayscale liver), this is not the case in robotic cochlear implantation. In this case, the physical relationship between the structures (needed to program the robotic trajectory) is more important than the need to distinguish between individual anatomical structures. Examples of extremely advanced graphical renderings of anatomy can be found in surgical training/educational platforms like those used in the Touch Surgery app [5] or LevelEx surgical training platforms [6]. Perhaps the same graphical design methods could be applied to platforms for surgical planning and robotic surgery to improve the overall user experience.

## 5 Conclusion

By improving the 3D visualization of anatomical and associated structures specific to robotic cochlear implantation, user satisfaction can be improved. More emphasis should be placed on the visual appearance of the system for the user.

## References

[1] "Deafness and Hearing Loss." [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss. [Accessed: 17-Jun-2019].

[2] "NIH Fact Sheets - Cochlear Implants." [Online]. Available: https://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=83. [Accessed: 17-Jun-2019].

[3] S. Weber *et al.*, "Instrument flight to the inner ear," *Sci. Robot.*, vol. 2, no. 4, p. eaal4916, 2017.

[4] R. F. Labadie *et al.*, "Minimally invasive image-guided cochlear implantation surgery: First report of clinical implementation," *Laryngoscope*, vol. 124, no. 8, pp. 1915–1922, 2014.

[5] "Touch Surgery - Prepare for Surgery." [Online]. Available: https://www.touchsurgery.com/. [Accessed: 20-Jun-2019].

[6] "Level Ex: Video Games for Doctors." [Online]. Available: https://www.level-ex.com/. [Accessed: 20-Jun-2019].

# VR Craniotomy for Optimal Intracranial Aneurysm Surgery Planning

*Belal Neyazi[1], Patrick Saalfeld[2], Philipp Berg[3,5], Martin Skalej[4], Bernhard Preim[2,5], I. Erol Sandalcioglu[1], Sylvia Saalfeld[2,5]*

[1]*Department of Neurosurgery, Otto-von-Guericke University Magdeburg, Germany*
[2]*Department of Simulation and Graphics, Otto-von-Guericke University of Magdeburg, Magdeburg, Germany*
[3]*Department of Fluid Dynamics and Technical Flows, Otto-von-Guericke University of Magdeburg, Magdeburg, Germany*
[4]*Department of Neuroradiology, Otto-von-Guericke University Magdeburg, Germany*
[5]*Forschungscampus STIMULATE, Otto-von-Guericke University of Magdeburg, Magdeburg, Germany*

contact: sylvia.saalfeld@ovgu.de

***Abstract:***
In the last decades, microsurgical clipping of aneurysms is often replaced by endovascular treatment, depending on the aneurysm location and patient-specific anatomy. In consequence, an increase in the complexity of aneurysms to clip represents a challenge for training of neurosurgeons as well as treatment planning in clinical practice. To overcome these limitations, we present a VR environment for the training of craniotomies for microsurgical clipping of intracranial aneurysms. The user can interactively explore the best skull positioning and place an adjustable virtual craniotomy to identify the best view on intracranial aneurysms. We conducted a pilot study for an initial evaluation of the usefulness of our approach yielding a benefit of the VR environment especially for peripheral aneurysms. The usage in clinical practice and treatment planning would require patient-specific data and arbitrary-shaped holes for differentiated bone flaps, extensions we intend to include in the future.

**Keywords: Virtual Reality (VR), Intracranial Aneurysms, Craniotomy**

## 1 Problem

Intracranial aneurysms (IA) are pathologic dilatations of the intracranial vessel wall that bear the risk of rupture with often fatal consequences for the patient. Treatment strategies for IAs have changed substantially during the past two decades, which constitutes a major challenge for the neurosurgical training und preoperative planning. Since the approval of Guglielmi detachable coils in 1995 the clipping first policy paradigm of IA was complimented by endovascular treatment options. In 2005, the International Subarachnoid Aneurysm Trial (ISAT) demonstrated improved mortality rates at one year in patients undergoing endovascular treatment compared to microsurgical clipping [1]. However, long-term results of the ISAT and the later Barrow Ruptured Aneurysm Trial (BRAT) showed an approximation of both treatment modalities [2]. The recent treatment strategy at our Neurovascular Center in the University Hospital Magdeburg takes different pathological aspects of IAs into account [3]. Patients suffering from wide-necked aneurysms of the anterior circulation benefit from microsurgical treatment [4]. Especially, patients suffering from ruptured aneurysms of the middle cerebral artery with accompanying intracerebral hemorrhage show a significant better outcome after microsurgical treatment [5,6].

Nevertheless, a decrease in microsurgical clipping interventions could be noticed [7]. As a result, there was a reduction in the number of clipped aneurysms on the one hand and a considerable increase in the complexity of the aneurysms to be treated by microsurgery, on the other hand. These changes in the treatment paradigm of IAs confront the training in cerebrovascular neurosurgery with considerable challenges. The traditional training, which was based on a long-term assistance and the highest possible number of cases, now has structural limitations.
To overcome these limitations, we developed a virtual reality (VR) based environment to support the training as well as treatment planning for microsurgery of IAs. Although VR simulation programs have been established for several

years for general surgical training, VR procedures are not widely used in neurosurgical clipping and treatment planning [8]. First approaches exist, but they are not tailored to microsurgical clipping [9] or solely allow for an interactive exploration of the vasculature but not for an interaction with the skull [10].

## 2 Material and Methods

### Medical Image Data

For the proposed system, a model of a complete Circle of Willis (CoW) was required, which was extracted from a healthy patient's MRI data set. The dataset was acquired for clinical education using the MAGNETOM Skyra 3T (Siemens Healthineers AG, Erlangen, Germany) with a 20-channel-head/neck coil and a voxel resolution of 0.26mm × 0.26mm × 0.5mm. A 3D surface model was segmented based on the approach described by Glaßer et al. [11].
In order to prepare our VR training environment, we artificially modified the CoW such that different training situations could be simulated. We prepared the following configurations: a healthy CoW, a CoW harboring an aneurysm at the anterior communicating artery, the M2 segment of the middle cerebral artery and two aneurysms at the posterior communicating artery, see Figure 1. The aneurysms were modeled with Sculptris (Pixologic, Inc., United States).

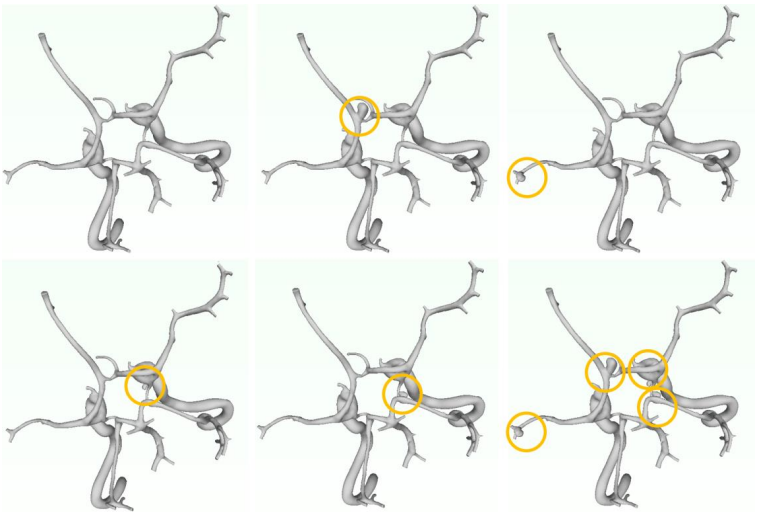


Figure 1: Depiction of the Circle of Willis (CoW) model and the four artificial IAs. From top left to bottom right: Healthy CoW, aneurysm at the anterior communicating artery, aneurysm at the middle cerebral artery, aneurysm at the posterior communicating artery and the internal carotid artery and the posterior artery, respectively, and CoW with all four aneurysms.

### Virtual Environment

Our immersive virtual environment was realized with the low-cost VR headset *HTC Vive* providing a field of view of 110° and a tracking area of 4.5 × 4.5 meters [12]. The interaction is realized with standard HTC Vive controllers providing six degrees of freedom and a sub-millimeter tracking accuracy. Our software prototype was developed with

the *Unity* game engine and the *Virtual Reality Toolkit* for basic VR functionalities such as grabbing as well as the SteamVR platform to control the HTC Vive.
In order to achieve a high degree of immersion, a virtual surgery room as presented by Huber et al. [13] was adapted to our application scenario. Here, the patient was replaced by the skull model including the CoW with the previously described IA configurations. For the virtual craniotomy itself, the user could change the placement and inclination angle of the skull [14]. We implemented two main rotation axes, see the user's menu in Figure 2. Furthermore, the hole representing the craniotomy can be interactively dragged along the skull, revealing the underlying vasculature. Hence, the hole's size can be interactively changed as well. In the software prototype, the hole was realized with a custom spherical cutout-shader that uses two information: (1) a position in world space and (2) a distance. In the shader, the world position of each fragment of the skull is calculated. If this fragment falls into the sphere described by the world position and distance specified in the shader, it is discarded, i.e., not rendered. Additionally, the back faces of the skull are rendered with a monochrome white color. These back faces are revealed on the border of the hole, which effectively gives an impression of the bone thickness. With the provided slider in the user menu, the user can interactively change the sphere's radius.

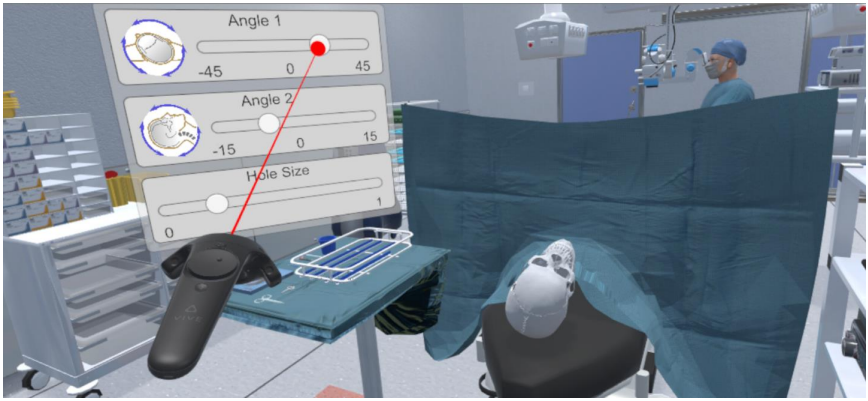



Figure 2: Depiction of the virtual environment with the user menu and the skull including the CoW models representing the patient.

## 3 Results

We conducted a pilot study to test our system and carried out an informal evaluation with an experienced neurosurgeon. We used the *Think-Aloud Method* [15], where the participant is asked to explain his thoughts about the functionalities during the evaluation to achieve qualitative results.
Therefore, we started by demonstrating the environment to the medical expert. Afterwards, the neurosurgeon tested the tool himself and modified the inclination and reclination angle of the head. He was positively surprised by the immersive experience and the quality of the surgery room as well as the 3D models but it must be noted that he was not familiar with such VR systems beforehand.
Required time for getting familiar with the sliders to modify the skull's inclination, to change the hole size and the movement of the hole itself took only very few minutes and we almost immediately could focus on the different aneurysm configurations. The virtual craniotomy is illustrated in Figure 3.

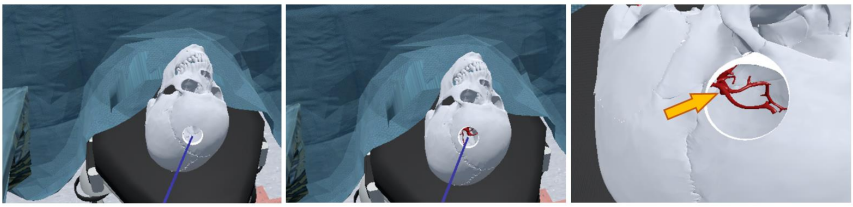


Figure 3: Illustration of the interactive hole definition. After the neurosurgeon is satisfied with the skull's inclination, the hole can be interactively dragged and resized until the trajectory reveals the aneurysm at the anterior communicating artery (arrow) satisfyingly.

When evaluating the different aneurysm cases within the pilot study, the medical expert highlighted the problems of aneurysms at the anterior communicating artery and the corresponding difficulties depending on whether the aneurysm is tilted slightly upwards or downwards. Hence, he praised the possibility of the virtual environment since the modifications can be easily presented. When analyzing the middle cerebral artery aneurysm, he demonstrated the differences in required craniotomy sizes to reveal the aneurysm depending on the skull inclination and the hole placement. This was repeated for the other aneurysms of the anterior part of the CoW. Finally, he used the healthy CoW and showed us challenging IA locations, e.g., near the anterior clinoid process, where the patient-specific anatomy strongly influences the adjustment of the craniotomy. Finally, he pointed out the advantage of exploring such cases in the virtual environment and the benefit due to visualization of the bone thickness with the spherical cutout-shader.

## 4 Discussion

The main advantage of our system was the interactive placement of the craniotomy and the pilot study revealed the strong benefit for determining well suited trajectories of possible interventions. The presented approach indicates that the tailored positioning of the patient's skull and the variable placement of the craniotomy in dependence of the pathology can lead to considerably less extended approaches and therefore less periprocedural trauma for the patient, see Figure 4. Furthermore, the true benefit of so-called skull base approaches could be demonstrated in a vivid and clear fashion for educational purposes. Hence, the medical expert repeatedly expressed enjoyment using the interactive tools and complained about the missing integration in clinical practice for a case discussion.

Although the study by Ekstrand et al. [16] concluded that no significant difference between second-year students using VR and those who did not existed regarding the learning and retention of neuroanatomy, we could easily identify several configurations where a neurosurgeon would highly benefit from a VR planning tool. These cases include peripheral arteries at the middle cerebral artery (as already implemented in the demo cases), the inclination of aneurysms at the anterior communicating arteries (based on the pilot study's demo case) as well as aneurysms near the anterior clinoid process due to the complex skull configuration near this location. The major criticism concerned the round shape of the hole realized with the custom spherical cutout-shader. Future applications should account for arbitrarily shaped holes or rather typical holes created during the clinical interventions. Furthermore, we used an exemplary CoW and skull model. In the next step, patient-specific data including a rendering of the patient's skin should be considered. However, since we do load the models at the beginning, they are easily interchangeable.

In the long picture, we plan to implement a clipping functionality as well. Here, we are also interested in modelling surrounding tissue, e.g. brain matter and specific anatomical landmarks, e.g. the Sylvian Fissure.

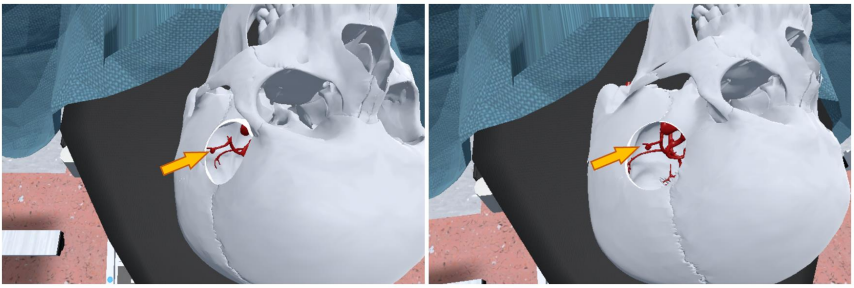


Figure 4: Illustration of the virtual craniotomy for the aneurysm at the middle cerebral artery (arrows); depending on a well-adjusted inclination of the skull and the size of the virtual hole, a smaller hole in the skull would be required (left) or a larger one (right).

## 5 Conclusion

In this work, we presented a VR environment for the training of craniotomies for microsurgical clipping of intracranial aneurysms. Based on a custom spherical cutout-shader to suppress parts of the skull, the user can interactively explore the best skull position and possible trajectory to reveal a well suited view on the IA of interest. A first evaluation indicated the usefulness of our approach, especially considering the challenges regarding neurosurgical training, i.e., less aneurysms are clipped and those who are clipped are often characterized by increased complexity. Future work should include patient-specific data and arbitrary-shaped holes for differentiated bone flaps. We do also intend to include the clipping itself in our environment at a later point in time.

## Acknowledgements

We thank Muttahir Mumtaz for his help building up the VR setup. This work was partly funded by the German Ministry of Education and Research (13GW0095A) within the Forschungscampus *STIMULATE* and the German Research Council (SA-3461/2-1, BE 6230/2-1).

## References

[1] Molyneux A. J., Kerr R. S., Yu L. M., et al., *International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised comparison of effects on survival, dependency, seizures, rebleeding, subgroups, and aneurysm occlusion*. Lancet,366 809–17 (2005)

[2] Spetzler R. F., McDougall C. G., Zabramski J. M., Albuquerque F. C., Hills N. K., Nakaji P., Karis J. P., Wallace R. C., *Ten-year analysis of saccular aneurysms in the Barrow Ruptured Aneurysm Trial*, J Neurosurg, 8:1-6 (2019)

[3] Etminan N, Brown R. D., Beseoglu K., et al., *The unruptured intracranial aneurysm treatment score: a multidisciplinary consensus.* Neurology, 85(10) 881-9 (2015)

[4] Mascitelli J. R., Lawton M. T., Hendricks B. K., et al., *Analysis of Wide-Neck Aneurysms in the Barrow Ruptured Aneurysm Trial*. Neurosurgery, doi: 10.1093/neuros/nyy439 (2018)

[5] Bohnstedt B. N., Nguyen H. S., Kulwin C. G., et al., *Outcomes for clip ligation and hematoma evacuation associated with 102 patients with ruptured middle cerebral artery aneurysms*, World Neurosurg, 80(3-4) 335-341 (2013)

[6] Lee C. S., Park J. U., Kang J. G., Lim Y. C., *The clinical characteristics and treatment outcomes of patients with ruptured middle cerebral artery aneurysms associated with intracerebral hematoma*, J Cerebrovasc Endovasc Neurosurg, 14(3) 181-185 (2012)

[7] Qureshi A. I., Vazquez G., Tariq N., et al., *Impact of International Subarachnoid Aneurysm Trial results on treatment of ruptured intracranial aneurysms in the United States. Clinical article*. J Neurosurg, 114(3) 834-841 (2011)

[8] Kockro R. A., Serra L., Tseng-Tsai, Y., et al., *Planning and simulation of neurosurgery in a virtual reality environment*, Neurosurgery, 46(1) 118-135, (2000)

[9] Antonio B., *Virtual Reality and Simulation in Neurosurgical Training*, World Neurosurgery, 106 1015-1029 (2017)

[10] Ong D., Yesantharao Q., Pakpoor C., et al., *Virtual Reality in Neurointervention*, Journal of Vascular and Interventional Neurology, 10(1) 17-22 (2018).

[11] Glaßer S., Berg P., Neugebauer M., Preim B., *Reconstruction of 3D Surface Meshes for Blood Flow Simulations of Intracranial Aneurysms*, In Proc. der Tagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (CURAC), pp. 163-168 (2015)

[12] Coburn J. Q., Freeman I., Salmon J. L., *A review of the capabilities of current low-cost virtual reality technology and its potential to enhance the design process*, Journal of computing and Information Science in Engineering, 17(3) 031013 (2017)

[13] Huber T., Wunderling T., Paschold M., et al., *Highly immersive virtual reality laparoscopy simulation: Development and future aspects*, International Journal of Computer Assisted Radiology and Surgery, 13(2) 281-290 (2018)

[14] Jägersberg M., Brodard J., Qiu J., et. al, *Quantification of working volumes, exposure, and target-specific maneuverability of the posterional craniotomy and its minimally invasive variants*, World neurosurgery, 101 710-717 (2017)

[15] Jakob Nielsen, *Thinking aloud: The# 1 usability tool*, Nielsen Norman Group (2012)

[16] Ekstrand C., Jamal A., Nguyen R., et al., *Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomized controlled study*, CMAJ Open, 6(1) E103–9 (2018)

# Klinische Anwendungen

# OP-Simulation in der Neurochirurgie durch 3D-gedruckte, vollfarbige Modelle am Beispiel von Schädelbasischondrosarkomen

*Kosterhon M[1], Neufurth M.[2], Schäfer L.[1], Neulen A.[1], Conrad J. [1], Müller W.E.G.[2], Kantelhardt S. R.[1], Ringel F. [1]*
*[1] Klinik und Poliklinik für Neurochirurgie, Universitätsmedizin Mainz, Deutschland*
*[2] Institut für Physiologische Chemie, Universitätsmedizin Mainz, Deutschland*

Kontakt: mikoster@uni-mainz.de

***Abstract***

*Die möglichst komplette Tumorresektion ist für die Behandlung zahlreicher intrakranieller Pathologien unerlässlich. 3D-gedruckte Modelle könnten hier die präoperative Planung vereinfachen, da sie die Simulation von Operationen ermöglichen würden. Gängige einfarbige 3D-Drucktechniken stellen jedoch nur die Oberfläche dar, während darunter liegende Strukturen nicht unterschieden werden können.*
*Es wurden die Möglichkeiten des mehrfarbigen 3D-Drucks für die Operationssimulation mittels binder-jetting-Verfahren (Polymer-Gips-Druck) untersucht. Durch eine Abwandlung des herkömmlichen Aufbereitungsverfahrens medizinischer Bilddaten für den 3D-Druck war es zudem möglich nicht nur die Oberflächen der Modelle, sondern auch innenliegende Strukturen volumetrisch einzufärben. Hierdurch konnten z.B. Strukturen wie ein Gefäß innerhalb eines Tumors in unterschiedlicher Färbung dargestellt werden. Am Beispiel von Patienten mit einem intrakraniellen Chondrosarkom wurde untersucht wie genau der 3D Druck die individuelle Anatomie reproduzieren kann und inwiefern sich das gipsartige Material für die Operationssimulation eignet.*

**Keywords**: 3D-Druck, Operationsplanung, Operationssimulation, Ausbildung

## 1 Problemstellung

Eine vollständige Tumorresektion ist für die Behandlung insbesondere von Schädelbasis-Chondrosarkomen unerlässlich, da Strahlen- und Chemotherapien nur wenig wirksam sind [1, 2]. Die anatomische Komplexität und die Seltenheit der Erkrankung machen die Operation zu einer herausfordernden Aufgabe. Der 3D-Druck eines individuellen Falls könnte die präoperative Planung vereinfachen, da er bei geeigneter Ausführung die Simulation einer Operation ermöglichen würde. Die meisten 3D-Drucktechniken produzieren jedoch nur Objekte mit einer oder wenigen unterschiedlichen Farben. Zudem sind die Objekte oft aus Kunststoff, so dass sie bei einer Bearbeitung mit chirurgischem Werkzeug wie Fräsen oder Stanzen schnell schmelzen oder brechen.

Das sog. binder-jetting-Verfahren, ermöglicht es 3D-Drucke, welche aus schichtweise gehärtetem Gipspulver, erstellt werden, mit nahezu unbegrenzter Farbvielfalt zu erstellen. In der Standardkonfiguration wird jedoch nur die Außenhülle des Objektes eingefärbt, sodass bei einer Bearbeitung mit chirurgischen Instrumenten innenliegende Strukturen nicht mehr unterschieden werden können.

Ziel war es eine Methode zu entwickeln, welche volumetrisch durchgefärbte 3D-Drucke ermöglicht. Zudem sollte untersucht werden, wie genau die 3D-Drucke die Anatomie, die aus den Bildgebungstechniken CT und MRT gewonnen wurden, reproduzieren kann. Zusätzlich wurde untersucht wie gut sich das Material mit herkömmlichen OP-Instrumenten (Fräse, Stanze) bearbeiten ließ.

## 2 Material und Methoden

CT-, CTA- und MRT-Daten von 7 Patienten mit einem intrakraniellen Chondrosarkom wurden in der Software Amira 5.4.5 (Thermo Fischer Scientific, Massachusetts, USA) räumlich miteinander fusioniert, gefolgt von der manuellen Segmentierung relevanter Strukturen wie Knochen, Gefäße, Tumor, Sehnerv und Hypophyse. Die so generierten Drahtgittermodelle wurden dann im stl-Format in die 3D-Druck-Software Netfabb Premium 2019.0 (Autodesk, Kalifornien, USA) importiert. Hier wurden sich überschneidende Modellanteile mittels Bool'scher

Operationen entfernt. Zudem wurden insbesondere für freischwebende Gefäße zusätzliche Abstützungen generiert.

Damit später im 3D-Druck die Objekte auch volumetrisch durchgefärbt sind, wurden ausgewählte Objekte (Tumor und innenliegendes Gefäß) weiter modifiziert. Hierzu wurde die äußere Hülle der Objekte mehrfach dupliziert und gleichzeitig verkleinert, sodass eine sich nach innen fortsetzende, zwiebelschalenartige Anordnung der Hüllen des Objektes resultierte. Anschließend wurden alle Objekte mit entsprechenden Farben versehen und im vrml-Format exportiert.

Die Daten wurden in die proprietäre 3D-Drucker-Software des ZPrinter 450 (3DSystems, South Carolina, USA) geladen und mit einer Schichtdicke von 0,1 mm gedruckt.

Nach dem Druck wurden die Modelle mit einer Härterlösung infiltriert, um sie widerstandsfähig zu machen.

## 3 Ergebnisse

Um die prinzipielle Reproduktionsgenauigkeit der Methode zu untersuchen wurde beispielhaft eines der gedruckten Modelle in einem CT gescannt und mit den anfänglichen 3D-Daten aus der Segmentierung in Nettfabb überlagert. In 50 zufälligen Oberflächenpunkten wurde die Abweichung zwischen den beiden Modellen bestimmt. Es zeigte sich ein mittlerer Abstand von 0,021 mm (SD +/- 0,076 mm) und damit eine hohe Reproduktionsgenauigkeit des 3D-Drucks. Dies ist eine Grundvoraussetzung für die Anwendung von 3D-Drucken zur Operationssimulation.
Zusätzlich wurden an den gedruckten Schädelmodellen eine temporale Kraniotomie und die Entfernung der Tumoren mit einer Fräse erfolgreich durchgeführt. Während der Bearbeitung zeigte sich, dass das Material nicht schmolz oder sich verformte bzw. splitterte.
Bezüglich der volumetrischen Färbung zeigte sich an zahlreichen Testdrucken, die mit einer Bandsäge aufgeschnitten wurden, dass innenliegende Strukturen klar voneinander anhand der Farbe zu unterscheiden waren. Zusätzlich wurde gezeigt, dass es mit einer Fräse möglich war ein Blutgefäß innerhalb eines Tumors vollständig herauszupräparieren, ohne dieses zu beschädigen.

## 4 Diskussion

Die vorgestellte Methode des farbigen 3D-Drucks auf Polymer-Gips-Basis ist ein bereits etabliertes Verfahren, wurde jedoch in der vorliegenden Arbeit um die Möglichkeit erweitert, Objekte volumetrisch durchzufärben, wodurch sie sich besser für die komplexen Anforderungen der Operationssimulation eigneten. Hierbei ist der Farbverbrauch während des Drucks deutlich höher, sodass die Technik nur auf die wesentlichen, anatomisch relevanten Strukturen, die in direkter Nachbarschaft zu den Tumoren lagen angewendet wurde.

Das Material eignet sich insbesondere gut für die Replikation knöcherner Strukturen. Die Simulation weicher Gewebe ist hierdurch jedoch nicht möglich, sodass beispielsweise die Verschieblichkeit und Retrahierbarkeit des Gehirns nicht nachgebildet werden konnte. Hier könnte künftig ein Hybridmodell, das harte und weiche Komponenten, die mit zwei unterschiedlichen Druckverfahren hergestellt werden, weiterhelfen.

## 5 Zusammenfassung

In dieser Fallstudie wurden vollfarbige 3D-Drucke erstellt, die zu visuell ansprechenden und „begreifbaren“ Modellen führten. Die individuelle Anatomie des Patienten konnte hier mit großer Genauigkeit nachempfunden werden. Das volumetrische Färben von Objekten ermöglichte überdies die Simulation auch komplexer Konfigurationen (Gefäß im Tumor). Die Technik könnte somit dazu beitragen patientenindividuelle Fälle realitätsgetreuer mittels 3D-Druck nachzubilden und für die präoperative Operationssimulation oder für Ausbildungs- und Schulungszwecke zu verwenden.

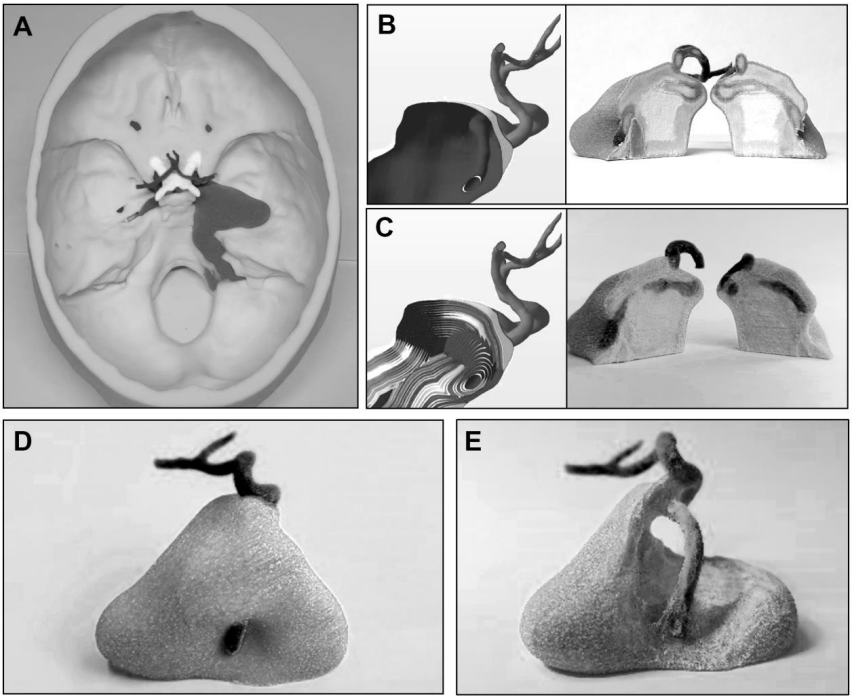



Abbildung 1*: **A:** Beispiel eines farbigen 3D-Drucks eines Patienten mit intrakraniellem Chondrosarkom; **B:** Herkömmlicher Workflow des farbigen 3D-Drucks: das segmentierte 3D-Modell besteht nur aus einer Oberfläche (links) und resultiert in einem 3D-Druck (rechts) der nur an der Oberfläche gefärbt ist (aufgesägtes Modell eines Tumors mit darin verlaufendem Blutgefäß); **C:** Abgewandelter Workflow mit zwiebelschalenartiger Anordnung mehrerer Oberflächen ineinander zur Füllung des Volumens des Objektes (links) und resultierender volumetrisch durchgefärbter 3D-Druck (rechts). **D:** Modell des Tumors mit darin verlaufender Arterie; **E:** Durch teilweise Entfernung des Tumors mit der Fräse kann die innenliegende Arterie herauspräpariert werden.*

## Referenzen

[1] Awad, M., Gogos, A. J. and Kaye, A. H. Skull base chondrosarcoma. *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia*, 24(Feb 2016), 1-5.
[2] Jones, P. S., Aghi, M. K., Muzikansky, A., Shih, H. A., Barker, F. G., 2nd and Curry, W. T., Jr. Outcomes and patterns of care in adult skull base chondrosarcomas from the SEER database. *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia*, 21, 9 (Sep 2014), 1497-1502.

# Konzept eines anatomischen Atlas für medizinische Studien und patientenspezifische 3D-Modelle

*P. Kongtso*[1], *P. Salz*[1], *T. Neumuth*[1], *A.W. Reske*[1,2]

[1] *Universität Leipzig, Innovation Center Computer Assisted Surgery, Leipzig, Deutschland*
[2] *Heinrich-Braun-Klinikum, Klinik für Anästhesiologie, Intensivmedizin, Notfallmedizin und Schmerztherapie, Zwickau, Deutschland*

Kontakt: patrick.kongtso@medizin.uni-leipzig.de

***Abstract***

*Die Zahl untereinander vernetzter medizinischer Geräte wächst, ebenso wie die Menge an medizinischen Patientendaten. Damit steigt auch der Bedarf, diese Daten für gezielte Zwecke zu selektieren, zu strukturieren und leicht abrufbar zu machen. Eine weitere Herausforderung sind fehlende patientenspezifischen Modalitätsdatensätzen im Notfalleinsatz. Mit dem anatomischen Atlas wird eine Plattform geschaffen, die selektierte Patientendaten anwendungsspezifisch aufbereitet und in einer leicht zugänglichen Form verfügbar macht. Das vorgestellte Konzept zeigt eine modular aufgebaute, leicht erweiterbare und skalierbare serviceorientierte Anwendung, die Patientendaten über gängige, domänunabhängige Schnittstellen bereitstellt und fehlende patientenspezifische Daten mit statistischen bzw. durch künstliche Intelligenz unterstützten Methoden errechnet.*

**Keywords**: Patientendatenbank, anatomisches 3D-Modell, künstliche Intelligenz, statistsches Datenmodell, quantitative CT-Analyse, Elektroimpedanztomographie.

## 1 Problemstellung

**Analyse von Patientendaten** Quantitative Analysen von medizinischen Daten [1] machen relevante, aktuelle und einheitliche Datensätze erforderlich. Daten müssen aus verschiedenen Quellen und Systemen, inklusive lokaler Textdokumente und handschriftlicher Notizen, gesammelt und in einer einheitlichen Form abgelegt werden. Überdies ist es notwendig, diese Daten zu erweitern, aktuell zu halten und für eine maschinelle Verarbeitung zu strukturieren.

**Patientenspezifische 3D-Modelle** Die Elektroimpedanztomographie (EIT) ist ein in der Medizin verwendetes bildgebendes Verfahren basierend auf der Messung elektrischer Leitfähigkeiten biologischer Gewebe [2]. Für die Lösung des zugrunde liegenden inversen Problems im Rahmen der Bildrekonstruktion wird ein möglichst präzises 3D-Modell der untersuchten Region benötigt [3]. Bei ambulanten Behandlungen bzw. Notfallsituationen stehen in der Regel keine rezenten Modalitätsaufnahmen des Patienten zur Verfügung. Hinzu kommt das Problem der zeitaufwändigen manuellen Bildsegmentierung und 3D-Modellerstellung. Aus diesen Gründen ist es erforderlich, dass vom EIT-Gerät in Echtzeit auf eine möglichst genaue Approximierung des anatomischen Aufbaus des Brustkorbs des Patienten zugegriffen werden kann.

**Stand der Technik** Aktuell stehen die für medizinische Studien relevante Patientendaten i.d.R. verteilt, inhomogen und schwer zugänglich zur Verfügung oder sind z.T. in digitaler Form gar nicht verfügbar. Moderne EIT-Geräte [4] stellen für die Rekonstruktion lediglich eine geringe Auswahl an Standardmodellen bereit. Die Schnittstellen sind dabei sehr domänspezifisch bzw. nicht standardisiert.

## 2 Material und Methoden

Der anatomische Atlas wird als modulare serviceorientierte Webanwendung realisiert. Dies ermöglicht eine unkomplizierte und standardisierte Maschine-zu-Maschine-Kommunikation auf Basis des Hypertext-Übertragungsprotokolls (HTTP(S)). Damit wird die Möglichkeit eines Datenaustauschs mit einem beliebigen System, das zu der Kommunikation über das HTTP-Protokoll fähig ist, sichergestellt. Weiter ermöglichen der modulare Aufbau und die serviceorientierte Architektur eine einfache Erweiterbarkeit und Skalierbarkeit des Gesamtsystems, unabhängig von der eingesetzten Technologie.

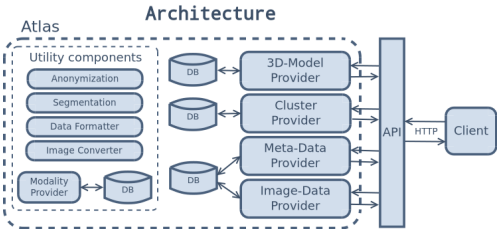



Abbildung 1: *Architektur und Komponenten des anatomischen Atlas.*

Abbildung 1 stellt die Architektur der Plattform dar. Hauptkomponenten bilden die Grundfunktionalität des Systems ab und sind über eine Webschnittstelle (API) erreichbar. Sie verfügen ggf. über eigene Datenbanken (DB). Utility-Komponenten interagieren mit den Hauptkomponenten und übernehmen gezielte Aufgaben.

**Patientendatenbank** Patientendaten werde über eine Webschnittstelle eingelesen und strukturiert abgespeichert. Interne Komponenten bereiten die Daten auf (Anonymisierung, Filterung, Formatierung, Segmentierung, Konvertierung, etc.). Die angefragten Daten werden über dieselbe Schnittstelle ausgegeben.

**Patientenspezifische 3D-Modelle** Patientendaten werden KI-unterstützt in Cluster zusammengefasst. Modalitätsaufnahmen werden weitgehend automatisiert segmentiert. Anschließend werden statistische 3D-Modelle für die jeweiligen Gruppen generiert.[1] Die Datenausgabe erfolgt über die Webschnittstelle.

## 3 Ergebnisse

Der anatomische Atlas bildet eine Datenbank, die für den Allgemeinbedarf in medizinischen Forschungseinrichtungen genutzt oder auch für die Lösung spezialisierter Aufgabenstellungen zugeschnitten werden kann. Der Atlas bedient sich der Methoden aus der KI und der Statistik, um fehlende und patientenspezifische Datensätze zu ergänzen. Die Daten werden über standardisierte und domänunabhängige Schnittstellen ausgetauscht. Eine Prototypimplementierung der Grundstruktur und einiger Komponenten für die Datenanalyse liegt bereits vor.

## 4 Diskussion

Durch den vereinfachten Zugriff auf medizinische Daten ist der Atlas vielfältig einsetzbar. Die erzeugten 3D-Modelle können in Bereichen Verwendung finden, in denen approximierte, an den Patienten angepasste anatomische 3D-Abbildungen benötigt werden, so z.B. für Simulationen von Abläufen, statistische Analysen, Planung von operativen Eingriffen, Erstellung von Phantomen oder Training des medizinischen Personals.

## 5 Zusammenfassung

Das vorgestellte Konzept des anatomischen Atlas für die Analyse medizinischer Daten und Erzeugung patientenspezifischer 3D-Modelle beschreibt ein zugängliches, modulares und leicht erweiterbares System, das in der Lage ist selektierte Daten zu halten, maschinell zu verarbeiten und mithilfe intelligenter Algorithmen zu ergänzen. Die bereitgestellten Daten finden potenziell in einem breiten Spektrum von Gebieten Anwendung.

## 6 Referenzen

[1] A. W. Reske und M. Seiwerts, “Qualitative und quantitative CT-Analysen beim akuten Lungenversagen”, Der Radiologe, Jg. 49, Nr. 8, S. 687–697, Aug. 2009.

[2] I. Frerichs u. a., “Chest electrical impedance tomography examination, data analysis, terminology, clinical use and recommendations: consensus statement of the TRanslational EIT developmeNt stuDy group”, eng, Thorax, Jg. 72, Nr. 1, S. 83–93, 2017.

[3] P. Salz u. a., “Improving Electrical Impedance Tomography Imaging of the Lung with Patient-specific 3D Models”, in Visualization in Medicine and Life Sciences, L. Linsen u. a., Hrsg., The Eurographics Association, 2013.

[4] I. Frerichs u. a., “Chest electrical impedance tomography examination, data analysis, terminology, clinical use and recommendations: consensus statement of the TRanslational EIT developmeNt stuDy group - Online supplement 1: Execution of EIT chest measurements”, eng, Thorax, Jg. 72, Nr. 1, S. 83–93, 2017.

[1] Die zu benutzenden KI-Algorithmen und 3D-Modelgenerierungsmethoden sind Gegenstand dieses Forschungsprojekts.

# Evaluation of different bladder phantoms for panoramic cystoscopy

*Ralf Hackner[1], Karl-Ernst Grund[2], Daniela Franz[1], Philipp-Fabian Müller[3], Niels Lemke[4], Thomas Wittenberg[1,5]*

*[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany*
*[2]Tübingen University Hospital, Tübingen, Germany*
*[3]Freiburg University Hospital, Freiburg, Germany*
*[4]Schölly Fiberoptic GmbH, Denzlingen, Germany*
*[5]Chair of Computer Graphics, University Erlangen, Germany*

Kontakt: ralf.hackner@iis.fraunhofer.de

***Abstract***

***Scope:*** *Panorama endoscopy is a new technique to enlarge a surgeon's field of view during an endoscopic "key-hole" procedure. Within cystoscopy it can be used during the examination of the bladder wall in order to enhance the documentation. For testing and evaluation of the panorama image stitching system with real endoscopy equipment, phantoms of the bladder are a good choice.* ***Materials:*** *In this work we assess four bladder phantoms with different geometry, surface textures, stiffness and different materials with cystoscopes with 0 and 30 degrees regarding the suitability for image stitching. For the generation of the image panoramas we apply an already established real-time stitching method that has successfully proven with image sequences from real human tissue.* ***Results:*** *Panoramas of all combinations of phantoms and endoscopes were computed. Using landmarks in the phantoms the maximum extension of the panoramas were assessed. Deformable phantoms perform better than static models. 30-degree cystoscopes yield larger panoramas.*

**Keywords**: Panoramic imaging, cystoscopy, bladder phantoms

## 1 Introduction and Objective

In the past years, we developed a software system to create panorama images of the bladder (and other body hollow organs) during endoscopic examinations in real time [1, 2]. Such endoscopic image panoramas enlarge the field of view for the surgeons, provide visual feedback if there are yet unobserved areas on the hollow wall and improve the quality of documentation. Figure 1 gives an example of a small image panorama of the bladder wall, computed retrospective from a cystoscopy sequence obtained from a routine bladder wall examination in the Freiburg University clinics. The software system interacts with the user in various ways: If due to a fast movement or blurry images the generation of the panorama cannot be continued, the user is requested to move the endoscope back to a position that is already integrated in the panorama in order to continue the stitching process with better images. Testing such interactions with retrospective image data as e.g. prerecorded videos from real examinations is hardly possible. Also live demonstrations of the stitching can't be performed based on prerecorded videos. Hence, as an alternative, bladder phantoms can be applied, made from silicone, rubber and other materials. An evaluation of various such phantoms will help us to improve the stitching procedures as well as improve the phantoms to be more realistic.

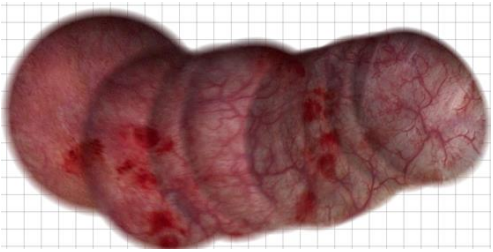


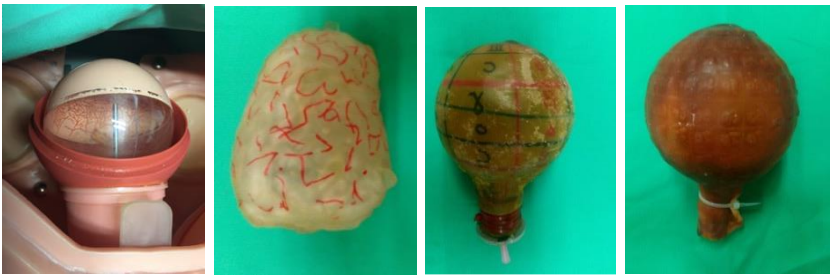
Figure 1: *Example of a panoramic image of the bladder, based on a cystoscopy examination from the Freiburg University Clinic*

Various bladder phantoms are currently available and applied in the field of urology. Nevertheless, most of these phantoms serve dedicated purposes, such as education in anatomy, ultrasound imaging assessment or endoscopy. One necessity for generation of cystoscopic panorama images is that availability of a tissue-like texture or somehow texturized surface inside the phantom. As standard phantoms for cystoscopy (as e.g. the VirtaMed *UroSim* System [3], the CLA bladder phantom [4], or the Simbionix *UroMentor* [5]) do not provide prominent and sufficient texture of the bladder wall, several alternative phantoms have been designed, built and tested with respect to panoramic image stitching of cystoscopic video data. Hence, for our research we apply various self-constructed phantoms [1] and phantoms from the Tübingen University Hospital (see section 2.1). A phantom similar to our self-constructed one (Fig. 2.a) was also used by [6] for experiments to reconstruct the 3D-geometry of a bladder from endoscopic images. Lurie et al [7] suggested a 3D-printed bladder phantom, with various layers, which inspired us to construct a less complicated phantom (*phantom B*, see below)

# 2 Materials and Methods

## 2.1 Phantoms

For our experiments, four bladder phantoms have been used. As one preposition a somehow textured surface of the bladder walls is required in order to provide sufficient information for the stitching algorithm. ***Phantom A*** is based on an acrylic-glass sphere, with a circular opening of 1 cm diameter for the insertion of the cystoscope. The distal hemisphere is colored in pink, and a vascular scheme has been added manually, cf. Figure 2.a. Only one hemisphere is covered with paint. ***Phantom B*** is a silicon-based 3D-print (see Figure 2.b). The 3D data used for the printing was obtained from an interactive segmented bladder based on an MRI image volume [8]. After the segmentation, the volumetric data is converted to a 3D-mesh, and "holes" have been closed. The location of the "entrance" to the lumen is anatomical correct at the urethra. For the vascular texture, the outside of the phantom, has been augmented manually. ***Phantom C*** (from the Tübingen University Hospital) has been made of silicone and is based on a handmade plaster-negative. For orientation purposes as well as texture, a grid with eight vertical and eight horizontal lines and a set of 64 different symbols have been added manually, cf. Figure 2.c. Finally, ***phantom D*** (also from Tübingen University Hospital) is a variation of phantom C, see Figure 2.d. Here the "negative" cast is a 3D-print, where elevated letters and numbers have been added to the surface. In the positive silicon-phantom these letters and numbers are depicted in the inside of the phantom and hence provide a texture and structures to support the stitching process.
While *phantom A* is static, *phantoms B* to *D* are partially deformable due to the use of soft silicon material. *Phantom B* is the only one with correct anatomical form and extension, while the other three phantoms approximate the bladder volume by a sphere.



(a) (b) (c) (d)

Figure 2: *Images of the bladder phantoms used for the experiments: (a) hand-colored acrylic glass sphere, (b) 3D-printed silicon phantom, (c) hand-colored silicon bladder, (d) silicon bladder phantom with imprinted structures.*

## 2.2 Endoscopic Imaging

For the cystoscopic imaging of the above described bladder phantoms, a state-of-the-art cystoscopy system was applied. This system consists of a light source (Richard Wolf, Knittlingen; ENDOLIGHT Flex LED), a 0-degree 4 mm cystoscope (Richard Wolf, SN 5000321214), a 30-degree 4 mm cystoscope (LUT HDScope, Schölly, Denzlingen; SN552687), an endoscopic video camera (Richard Wolf, ENDOCAM) and an endoscope video platform (Richard Wolf, ENDOCAM Logic HD).

## 2.3 Panoramic Image Processing

In the past years, we developed a software system allowing to compute and display large-scale panoramic images of the bladder in real-time during a diagnostic cystoscopy [1,2] in real time. The software has been optimized for the usage with the bladder [1,2], but also works with other spherical organs [9,10]. The system is connected to the endoscope controller (see section 2.3) by a frame-grabber to acquire the live video stream from the cystoscopy system during the examination of the phantom in real time with approximately 25 frames per second. For the image stitching process (cf. Figure 3), in a first step, a color-shading correction is performed [11] in order to equalize the inhomogeneous endoscopic illumination. This is followed by a features extraction process.
We use SURF features and key points for the tracking process and the registration. In the next step a *local* a frame-to-frame registration is performed along the determined track of the video-sequence. In a final step, a *global registration* process based on bundle adjustment corrects possible misalignments and provides a *loop-closing* [1].

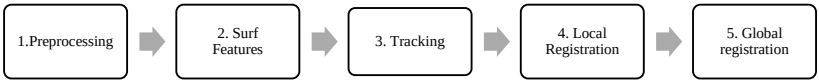



Figure 3: *Overview of panorama generation*

## 2.4 Quality Criteria

An important quality criterion of an endoscopic panorama image is its ***completeness***, thus every region of the bladder that is observable for the used cystoscope (0 or 30 degrees), should also be depicted in the corresponding panorama. Closely related to the completeness (no holes visible in the panorama) is the ***extension***. As our bladder phantoms are mainly spherical objects, we defined an imaginary fixpoint in the center of the observed hemisphere (the “south pole”) directly opposite to the insertion opening of the bladder-model at the urethra (“north pole”), see Figure 4. In an assumed spherical phantom this extension can be assessed by the *degree of latitude* between the fix point (being the “south pole”) and the maximum observable landmark on the border of the panorama. It can be assumed, that with a standard rigid endoscope, where the tip is located in the center of the sphere and moved using standard tilting, rotating, pushing and pulling maneuvers, the complete distal hemisphere form the “south-pole” to the “equator” at 90 degrees latitude can be captured (cf. Figure 4a. and b.) Hence it is of special interest how far the panorama can be extended *“beyond the equator”*. Thus, for each panorama the *maximum extension* is approximated by the distance of a landmark furthest away (in degrees of latitude) from the “south pole” (see Figure 4.b and Table 1). This angle limits the observed area. It is strongly dependent on the viewing angle of the endoscope used (0 or 30 degree), as well as the geometry and diameter of the entrance point (urethra).

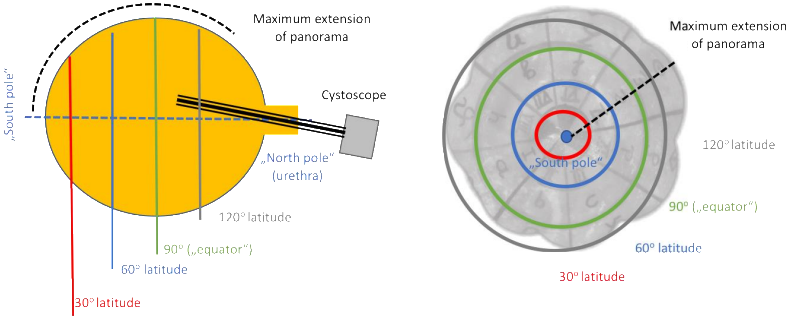


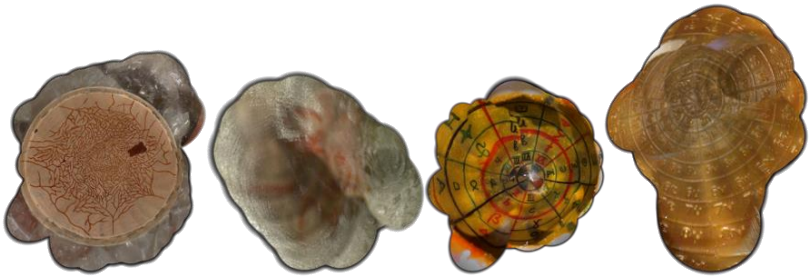

Figure 4: *Left Side: Lateral sketch pf a spherical phantom with denoted fixpoints ("south pole”) and urethra (“north pole”) and degrees of latitude. The dotted line refers to a possible maximum extension of the generated panorama. Right side: related virtual 2D panorama of a spherical phantom, with "south pole” and degrees of latitude marked. The dotted line refers to the maximum extension of the generated panorama.*

One further important criterion for the evaluation of panoramic images is the ***quality of the stitching process*** itself. If the algorithm fails to extract a sufficient amount of unique features, single tiles might be placed at the wrong position or are sheared or stretched in a wrong manner. While extreme shearing can be detected and the algorithm

drops affected tiles, misplaced frames are hard to detect automatically without a ground truth. Therefore, we inspected the resulting panorama images manually for stitching errors.

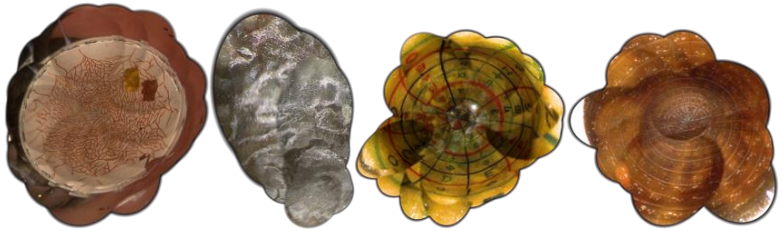
## 3 Results

In our tests were able to obtain panorama images for all four phantoms with a 0-degree as well with a 30-degree cystoscope. Representative panorama images are depicted in Figure 5 (0-degree) and Figure 6 (30-degree). The cystoscope tip was moved from the center of the hemisphere ("south-pole") to the border towards and beyond the "equator" (see Figure 4) and backwards again. This motion was repeated with an angular shift resulting in star-shaped tracks and panoramas. The stitching algorithm is applied in an identical manner for all tests, except of the parameter for the diameter of the region of interest. We adapted this parameter to the different optical projections of cystoscopes onto the camera. The used set of parameters has been previously proven to work well with the texture of real human bladders [2].



*(a)* *(b)* *(c)* *(d)*

Figure 5: *Panoramic Images recorded with a **0-degree cystoscope** from the bladder phantoms in Figure 2: (a) hand-colored acrylic-glass sphere phantom with a gummy bear as "lesion" at two 'o clock, (b) 3D-printed silicon phantom, (c) hand-colored and hand-labeled silicon phantom, (d) silicon bladder phantom with imprinted structures.*



*(a)* *(b)* *(c)* *(d)*

Figure 6: *Panoramic Images recorded with a **30-degree cystoscope** from the bladder phantoms in Figure 2: (a) hand-colored acrylic-glass sphere, (b) 3D-printed silicon phantom, (c) hand-colored silicon phantom, (d) silicon bladder phantom with imprinted structures.*

Approximating the maximum observation angles in case of *phantom B* was only possible with low precision, since the landmarks applied on the outside of the 3D-printed phantom were not clearly visible from the inside. Hence, we used a laser pointer on the outside to identify the corresponding structures inside the phantom. *Phantom A* incorporates only one painted hemisphere, thus we were not able to identify landmarks in the panorama image with an angle of more than 90 degree. To obtain the outermost reachable angle we added some markers from

outside. See Tables 1 and 2 below for results. Table 1 provides the *maximum observed angles* between the central landmark "south-pole" and the landmarks depicted in the periphery of the panorama (borders of the panorama).. Table 2 shows the number of *visible stitching errors* in relation to the *total number of frames* in the panorama image.

Table 1*: maximum angle between the central landmark ("south pole") and outermost landmark that can be observed with the endoscope.*

| ***Cystoscope*** | ***Phantom A*** | ***Phantom B*** | ***Phantom C*** | ***Phantom D*** |
|---|---|---|---|---|
| 0 degree | >90 º | 117 º | 130 º | 135 º |
| 30 degree | >90 º | 131 º | 155 º | 144 º |

Table 2*: Visible stitching errors in relation to the total number of frames in the panorama image*

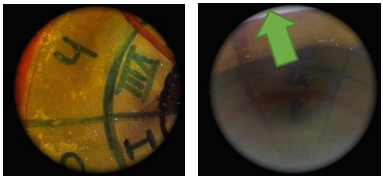
| ***Cystoscope*** | **Phantom A** | **Phantom B** | **Phantom C** | **Phantom D** |
|---|---|---|---|---|
| 0 degree | 2 /87 | 30/92 | 5/95 | 5/128 |
| 30 degree | 8 /112 | 25/87 | 9/83 | 6/180 |

# 4 Discussion

Despite of the vascular pattern added on the outside of the 3D printed *phantom B* (being the only one with correct anatomical shape) we were not able to obtain a panorama image with sufficient quality. Even the entire outer surface is covered with a vascular pattern, it is not visible from the inside at all locations, as can be seen in Figure 5.b. The reason is the reduced transparency of the material due to the ringing effect, which is typical for 3D printers and the ensured differences in the thickness of the shell. Compared to the work of Lurie at al [7] with a similar 3D printed phantom, the external texture lacks the needed structure. Nevertheless the complex 3-layer approach of [7] providing an *internal vascular texture* is currently hard to reproduce with limited resources. Hopefully the phantom of [7] can be assessed or borrowed for further experiments.

For the other evaluated phantoms (A, C, D), we achieved very satisfying results. The computed panoramas depict wide parts of the observed phantom surface. Unsurprisingly, the captured panorama images cover a larger part of the surface with the *30-degree* side-view endoscope than with the 0-degree frontal view cystoscope. The hand-drawn lines on *phantom A* and *C* provide a sufficient number of features for the stitching process, as well as the riffle structure of *phantom D*. For *phantoms C* and *D*, the obtained angles in the panorama are close to the observable maximum and are beyond the "equator". *Phantom A* is a special case, since only one hemisphere has been covered with usable textures. Stitching worked also for a fringe outside the painted part (most likely due to scratches in the shell and structures outside), but we were not able to identify unique landmarks or measure the angle here. The rate of stitching errors is low for all phantoms.

As it can be seen in Figures 7, phantom *C* contains a number of dark frames in the fringe area. The material of the phantom is reflecting light well. When the endoscope is moved close to the shaft to record the area around it, it can happen that small parts of the shaft come very close to the endoscope. These small areas become over-exposed, due to reflection, and the camera controller tries to compensate that, by reducing the exposition. The luminance correction used in the algorithm [11] was not able to compensate this completely.



*(a)* *(b)*

Figure 7: *An image frame with regular exposure (a) and a partially over exposed frame (b) from phantom C.*

# 5 Conclusion

We compared four different bladder phantoms with respect to the question, whether their surface structure and geometry is suitable for the generation of panorama images. We were able to obtain panorama images for all phantoms in our experiments, but with strongly varying quality. The ringing structure typical for 3D prints does not provide sufficient texture and reflective surfaces can cause problems. Engraved and painted patterns (*phantoms A, C* and *D*) on the surface seem to work equally well. The two deformable phantoms (*C* and *D*) yield larger panoramas than the static phantoms (*A* and *C*). On all phantoms side-view 30-degre cystoscopes provide panoramas with a larger extensions than the frontal-view 0-degree endoscopes.
A still open challenge is the design and construction of a phantom that resamples the real anatomical geometry of the real bladder as the 3D printed version one and also having a suitable real vascular texture in the inside. Here, some ideas of Lurie et al [7] might be considered for future experiments. Since the riffle structures from *phantom D* also worked, it could either be an attempt to add artificial riffles to the next version of the 3D print in the inside, so that it can be used as phantom directly, or as a negative cast for a silicon phantom like *phantom D*.

# Acknowledgement

This work has partially been supported by the German Ministry for Research and Education (BMBF) under the contract 03ZZ0444E (subproject **"**Interactive visualization of 2D/3D panorama endoscopy data of the bladder") within the Project "3D-Sensation – Uro-MDD".

# References

[1] T Bergen*: Real-time Endoscopic Image Stitching for Cystoscopy*. PhD Thesis, Univ. Koblenz (2017).

[2] M Kriegmair, T Bergen, M Ritter, P Mandel, MS Michel, T Wittenberg, C Bolenz: *Digital Mapping of the Urinary Bladder: Potential for Standardized Cystoscopy Reports*. Urology. pii:S0090-4295(17) 30151-6. 2017

[3] *VirtaMed - UroS™ for TURP, TURB and Laser BPH training*. Retrieved on 2019-07-01, https://www.virtamed.com/en/medical-training-simulators/uros/

[4] *CLA-Urological Examination Phantom for Endoscopic Operations*. Retrieved on 2019-07-01, https://www.somso.de/en/anatomie/medizinischephantome/urology-catheterisation/ts-64/

[5] *Simbionix URO Mentor.* Retrieved on 2019-07-01, https://www.3dsystems.com/medical-simulators/simbionix-uro-mentor

[6] Q Péntek, S Hein, A Miernik, A Reiterer*: Image-based 3D surface approximation of the bladder using structure-from-motion for enhanced cystoscopy based on phantom data*. Biomed. Eng. 63(4), 461-466 (2018)

[7] KL Lurie, GT Smith, SA Khan, JC Liao, AK Ellerbee. *Three-dimensional, distendable bladder phantom foroptical coherence tomography and white lightcystoscopy*. J Biomedical Optics 19(3), 036009 (2014). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3951584/pdf/JBO-019-036009.pdf

[8] M Pfeifle, S Born, J Fischer, F Duffner, M Hoffmann, D Bartz, *VolV - Eine OpenSource-Plattform für die Medizinische Visualisierung*, CURAC, Karlsruhe (2007)

[9] R Hackner, A Hann, A Meining, M Raithel, T Wittenberg, *Panoramic Endoscopy of the Stomach: First results from Phantom and Patient Data*. CURAC, Leipzig (2018)

[10] T Bergen, P Hastreiter, T Xu, C Münzenmayer, M Buchfelder, S Schlaffer, T Wittenberg*: Real-time panorama imaging of sphenoid sinuses from monocular endoscopic views: an update based on clinical experiments.* Proc's CURAC, pp. 32-35. 2014

[11] C Münzenmayer, F Naujokat, S Mühldorfer, T Wittenberg, *Enhancing Texture Analysis by Color Shading Correction*, 9. Workshop Farbbildverarbeitung (2003)

# Introducing Virtual & 3D-Printed Models for Improved Collaboration in Surgery

*A. Reinschluessel[1], R. Fischer[2], C. Schumann[3], V. Uslar[4], T. Muender[1], U. Katzky[5], H. Kißner[5], V. Kraft[3], M. Lampe[5], T. Lück[6], K. Bock-Müller[5], H. Nopper[6], S. Pelzl[7], D. Wenig[1], A. Schenk[3], D. Weyhe[4], G. Zachmann[1], R. Malaka[1]*

[1] *Universität Bremen, Digital Media Lab, TZI, Bremen, Germany*
[2] *Universität Bremen, CGVR Lab, Bremen, Germany*
[3] *Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany*
[4] *University Hospital for Visceral Surgery, Medical Campus University of Oldenburg, Oldenburg, Germany*
[5] *szenaris, Bremen, Germany*
[6] *cirp GmbH, Heimsheim, Germany*
[7] *apoQlar, Hamburg, Germany*

Kontakt: malaka@tzi.de

***Abstract***

*Computer-assisted surgery and the use of virtual environments in surgery are getting popular lately, as they provide numerous benefits, especially for visualisation of data. Yet, these tools lack features for direct and interactive discussion with remote experts and intuitive means of control for 3D data. Therefore, we present a concept to create an immersive multi-user system, by using virtual reality, augmented reality and 3D-printed organ models, which enables a collaborative workflow to assist surgeries. The 3D models will be an interaction medium to provide haptic feedback as well as teaching material. Additionally, multiple depth cameras will be used to provide remote users in the virtual environment with a realistic live representation of the operating room. Our system can be used in the planning stage, intraoperatively as well as for training. First prototypes were rated as highly useful by visceral surgeons in a focus group.*

**Keywords**: computer-assisted surgery, 3D-printing, virtual reality, augmented reality

## 1 Problem

Digital tools, such as visualization software, can be powerful instruments for surgeons to plan complex surgeries computer-assisted [7], [18], [21], [27]. While the planning data from these tools is available for the local surgeon, only few approaches exist to transfer the data into the intervention room or to allow for discussions and interactions with remote personnel and experts. Virtual and augmented reality (VR and AR) might provide help in various ways in the context of surgery, such as preoperative planning [7], [18], [21], [27]. Creating a *multi-user* immersive VR environment (or AR) for preoperative planning, intraoperative support, and training, comes with challenges, such as: (1) transmitting big amounts of data with high update rates and low latency, (2) creating a sufficient immersion and (3) an intuitive interaction, as the user experience is always a crucial factor. First examples for surgical applications using either VR or AR have been proposed [13], [16], but in general they restrict themselves to single parts of the process such as the visualization of CT (computer tomography) and MRI (magnetic resonance imaging) data [9], planning based on this data [18], or aiding the surgical intervention using this data [22], [25], [26].

In contrast, our research aims at creating a system which supports a broader spectrum of activities of the surgeon in the following three phases:

(1) *preoperative:* discussing radiological images and derived data, and planning the operation steps together with (remote) colleagues

(2) *intraoperative:* performing the surgery while having access to the planning data and if necessary being able to call in a colleague (via telepresence) to assist

(3) *training:* using the case data for teaching, training, in demonstrations, or patient education (beforehand)

To reach this goal, we will tackle the earlier mentioned challenges regarding data transfer, immersion, and (multi-user) interaction. In addition, most existing approaches are limited to displaying images and quantitative data on screens. Live discussions with remote experts based on the real organ or an accurate patient model do not exist to the best of our knowledge. Interacting with medical images, whether remotely or not, happens mostly on a 2D screen with a mouse. In the case of VR and AR environments or 3D displays, abstract gestures

or handles are used for interaction [9], [17], [18]. But surgeons and physicians heavily rely on their tactile sensations and their visual thinking. Hence, one of their essential abilities is to use their anatomical knowledge to interpret the spatial relations of the case at hand based on the available radiological image data and on what they see and feel in the situs. Therefore, an obvious requirement for a surgical VR/AR system is to support this ability. To be best prepared for the real case, surgeons typically plan the intervention beforehand, which is also a critical stage in decision making [14]. 3D-printed organ models are already used for different purposes in medicine [12], e.g., prints of liver (parts) for planning [28]. Nevertheless, a review by Martelli et al. [12] found just 158 cases scientifically reported in a time span of 10 years (2005-2015). This leads to the conclusion, that the technique of creating a model from computed tomography (CT) (or MRI) data [23] is not common yet and requires further research. To summarize the research results: VR and AR miss the "realistic" sensation of what the users see, but offers a variety of options to show the important information from and in the image data. The 3D-prints just show a selected view of the images and do not provide any further displaying options. To get the "best of both worlds", we aim at combining both modalities to match the haptic perception and the rich visualization possibilities of VR and AR to increase the knowledge of the surgeon. As the AR technology develops quickly, AR will be treated as equal to VR while pursuing this aim.

In the field of telemedicine, most research focused on remote-controlled minimal-invasive operations [24] and rarely on systems supporting the actual procedure [15]. There are two commercial telemonitoring systems available [2], [20] and both use video streams to the remotely working medical staff. Even systems with AR support use tablets and video streams [3]. Research on the effect of such systems shows, that there is no difference if there is a remote or a local mentor [19] and the system use leads to better results but takes longer [4]. The reason might be the technology, as most proposed systems rely on using depth cameras for skeleton tracking, which are mapped onto avatars [5] in the application. Avatars are important, as their quality has an influence on behaviour and team performance [8], [10] . However, the need for extensive pre-processing and the big data volumes make the usage in real-time VR and AR applications difficult [1], [6].

As pointed out in the previous paragraphs, to reach our aim of a multi-user VR/AR environment to support surgeons along with different phases a lot of basic research needs to be done, and the presented paper will not solve these problems yet. But in the following sections, we will present our approach and initial steps towards a solution. To get started, we ran a focus group with five visceral surgeons. We got valuable feedback, especially about the interaction part and which aspects are important for the surgeon's work.

## 2 Material and Methods

Our general goal is to assist surgeries, from the planning stage to the actual intervention, and also during education and training of becoming a surgeon, using and combining VR and AR as well as 3D-printed organs as tangible user interfaces. In the following two sections we will first present details on our technical concept of our idea and then shortly present the procedure of a focus group we ran to evaluate our concept idea.

### 2.1 Technical Concept

To reach our goal of supporting surgeons to work collaboratively and effectively on the same set of data, sharing and visualizing information in real time and over distant locations is crucial. The creation of the data that forms the basis for the virtual environments and the 3D-prints involves the following steps:

1. medical image acquisition (CT, MRI) at the clinical site
2. medical image data analysis by medical-technical radiology assistants, including delineation of relevant structures (organs, vessels, tumors) and planning of resection planes
3. deriving tissue qualities including disease state and softness
4. conversion into 3D models (polygonization) and application of textures to visualize the disease state
5. creation of 3D-printed models under consideration of the derived textures and tissue softness values

These steps are carried out as a separate process before the data is actually used and may take days including the 3D-printing. As a long term goal, this process should be automatized as far as possible. Furthermore, high standards in data security including anonymization and secure data transfer need to be established.

As possible hardware to view the data in all three phases mentioned in section 1, recent VR and AR devices like HTC VIVE[1] and Microsoft HoloLens[2] are of interest. The HoloLens will be used by the surgeons

[1] https://www.vive.com
[2] https://www.microsoft.com/de-de/hololens

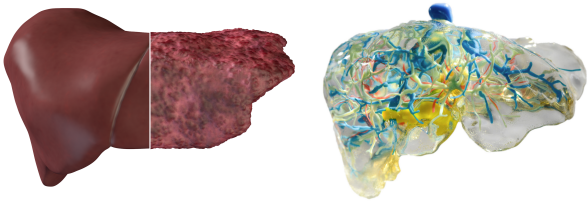


Figure 1: *Left: Exemplary textured 3D model showing varying liver tissue states (left: healthy, right: cirrhotic) Right: a transparent 3D-print of a liver*

in the operating room (OR) to explore the data and to interact with it. The HTC VIVE, in turn, will be applied by the remote personnel or experts in a way that they can discuss and interact within the multi-user virtual environment. Based on virtual representations of patient specific organs, all users will then be able to collaboratively examine and annotate the model. Also, VR and AR components of our architecture are coupled, which enables a seamless communication between them, e.g., displaying planning results and annotations made in VR on the AR-glasses during a surgical intervention. For the realization of such a shared immersive AR/VR environment we need to tackle the following technical challenges:

- realistic rendering of the anatomical models including textures to illustrate the tissue state (Fig. 1 left)
- creation of a realistic 3D environment for the telepresence user in the intraoperative phase

For the rendering of anatomical structures we will explore both, surface rendering and volume rendering, to determine the best possible visualization of the relevant anatomical structures and the resection proposals for the various application scenarios and display technologies. Common to both approaches is the need to focus on the relevant information. Hence, we will carry out studies to determine for each scenario which structures should be visualized in which way. Furthermore, we will investigate technologies for realistic rendering to raise the level of immersion. In the case of AR, this will also necessitate to explore methods for streaming as the hardware of the HoloLens is not suited for advanced volume rendering. To provide a realistic looking live 3D environment for the telepresence user in the intraoperative phase, we will use multiple depth cameras which record and stream the surgical intervention into the virtual environment. Here, a three-dimensional representation of the fused data, which includes the color as well as the depth information, is shown in form of a point cloud. Since the streaming of the data of multiple depth cameras over network needs a considerable amount of bandwidth, we will make heavy use of compression algorithms, although it is important to keep the computational complexity and therefore the latency low. We plan to enhance the current algorithms and develop specialised ones in the future to further reduce the needed bandwidth and latency. In this regard, we will also explore different transmission techniques and formats as well as filtering algorithms. Depending on the use case, the 3D-printed organ models used as tangible user interfaces need to be for instance deformable, and/or transparent. Since for some use cases, tactile feedback should be as realistic as possible, we need to conduct studies to measure organ softness depending on the respective disease. Furthermore, as the 3D-printed organ models need to be integrated into the VR, tracking using reflective markers will also be employed. Hardware solutions are for example OptiTrack[3] or Brainlab[4].

## 2.2 Focus Group

In order to investigate the potential of our approach, we performed a focus group with five physicians from the University hospital for Visceral Surgery in Oldenburg. The experts were a head physician, two chief residents and two residents of the visceral surgery department. The system was introduced to the physicians through a verbal description and three accompanying prototypes. One was a 3D-printed liver model (see Fig. 1 right) and material probes with different properties in terms of softness to introduce the haptic component of our idea. The second was a VR prototype with a virtual model of the same liver. The liver model was attached to a VR controller and could be inspected by turning the controller. With the second controller participants could annotate the virtual model. Additionally, the physicians could explore an application on the HoloLens, showing a human skull. The discussion of the focus group was recorded and analyzed descriptively afterwards.

[3] https://optitrack.com
[4] https://www.brainlab.com/de/

# 3 Results

The discussion in the focus group revealed that the presentation of the organ is a central aspect. The surgeons requested functionalities like marking, scaling, and changing the views, for instance by hiding vessels, or showing surrounding tissue. As for the 3D-printed organ model, they did not care about the colour when the model is used for controlling in the VR. When talking about training, they favoured two versions: an abstract visualization and a realistic visualization. The abstract one should be used mainly for the surgery planning in order to learn about the spatial arrangement of structures in the organ. The realistic visualization aims at forcing users to use their haptic senses to explore the model. The surgeons additionally suggested to overlay the relevant structures with AR.
The 3D-printed model was accepted very quickly and the surgeons expressed a huge benefit of soft models with embedded hard structures as tumors for education purposes. Landmarks, such as the ligament of the liver, should also be visible and palpable in a 3D-print. The participants were divided over the benefit for patient individual models, also in a soft print fashion, as they discussed the costs/benefit ratio and the environmental aspect. Additionally, the size was discussed intensively, as a real sized model might be beneficial for educational aims, but as an interaction device it was judged as too big and heavy.
The 3D-printed model also led to valuable observations regarding the interaction, as surgeons turned the model and pointed with their finger while showing their fellow colleagues the details of the presented model. One participant's suggestion was to record the planning process to inform the staff in the operation room beforehand of the procedure. Asked about planning with remote colleagues the participants were enthusiastic about the idea as current solutions involve screen sharing and Skype[5] or similar services, which is not a "safe solution according to good clinical practice" and brings with it high concerns with regards to patient data confidentiality.

# 4 Discussion

The results from our focus group show that the concept idea of our system is received very well by the participating physicians. They all agreed that our system would improve the current workflow and the combination of AR/VR with haptic models opens up new possibilities. Consistent with literature findings [14], the surgeons highlighted the benefit of better spatial perception by using a 3D model. Based on their interaction with the model and how they discussed the model at hand, it was obvious that this way of visualization and interaction encourages exchange between the experts.

The results of the focus group can be used to differentiate the use of technology between the three scenarios or phases of section 1: preoperative planning phase, intraoperative phase and the training scenario.

In the *preoperative planning phase* the surgeon can review patient's data and plan the surgery, either in VR or in AR. Especially in VR the 3D-printed model, either a general or case-specific model, will aid as an interaction device to control the virtual model. But also in AR the 3D-printed model can be used for the same purpose and might be overlaid with additional information, as the physicians positively discussed. Finding the optimal rendering solution for VR as well as AR, i.e. the HoloLens, will be an important aspect for the success of the system. As our goal is to create a multi-user application, the interaction using a 3D-printed model at one location and the purely virtual model at another location simultaneously will be a strong research focus. Research on how to present several users adequately and with low latency accompanies this research aspect. Also the size of the organ model will be of research interest, to ensure an interaction that is not fatigueing, which was a concern of the participants.

The information from the planning phase will be available in the *intraoperative setting*. In the OR AR technology will be used to ensure the surgical staff has a clear view of the situs. At the same time the surgeon can view the planning data whenever needed and interact with it with hand gestures. As the surgeons stressed the benefits of being able to bring in remote experts who get a realistic impression of the current situation in the OR, an important research topic will be how to transfer point cloud data between different locations with low latency. This will enable remote experts to view the live situation of the OR in 3D and they will be able to get all spatial information necessary to support the surgeon. This technology can give surgeons the possibility to consult specialists who can help with complicated or unusual situations and give valuable advise as they are aware of the current state of the intervention.

As the *education* of future physicians and the *training* of surgeons is an important aspect for our future, the third application area of our concept focuses on these topics. Thereby all technology developed for the scenarios preoperative planning and intraoperative support can be used to train on real cases. Furthermore, the 3D-prints can be used for a variety of use cases: First, transparent models in real size can be used to teach and

[5] `https://www.skype.com/de/`

train the spatial relations of internal structures of the organs as the participants highlighted the importance of this ability. Second, opaque haptically realistic models with varying softness can be used to train visual and tactile diagnostic skills. These models can either be general examples or case-specific models, which will be reused from a real case. Printing haptically realistic models matching a liver with cirrhosis or tumors inside is challenging. Research has shown, that current 3D-printing material is not soft enough to mimic human tissue and just workarounds like air pokes and vents in the 3D-print can get (nearly) satisfactorily results [11]. Biological materials like collagen are not suitable, as our models are supposed to be long lasting for repetitive used in lectures. Therefore, creating a realistic and long lasting 3D-print will be of research interest together with proper didactic integration.

## 5 Conclusion

Computer-assisted surgery becomes more common and is required by surgeons to help them plan complex surgeries, but the respective tools are lacking collaborative features as well as haptic feedback. In this work we presented a concept idea of a system, that can be used to support planning and execution of the surgery as well as training and education. We will combine a multi-user virtual and augmented reality environment with 3D-printed organ models as tangible user interfaces. Additionally, depth cameras will be used for a live reconstruction of the surgical intervention, so surgeons and remote personnel will be able to collaboratively view and manipulate detailed 3D data interactively in an immersive environment. In addition to describing the concept, we presented results of a focus group with five visceral surgeons, who tested first partial prototypes. The surgeons rate the approach as highly useful and highlighted the advantage of easier grasping the spatial relations and discussing with remote colleagues, which would greatly improve the planning phase of surgery and further steps.

## 6 Acknowledgment

This project has received funding from the German Federal Ministry of Education and Research (BMBF) in the grant program “Gesundes Leben” (healthy living).

## 7 References

[1] J. Achenbach, T. Waltemate, M. E. Latoschik, and M. Botsch, “Fast generation of realistic virtual humans,” in Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, ACM, 2017, p. 12.

[2] R. Agarwal, A. W. Levinson, M. Allaf, D. V. Makarov, A. Nason, and L.-M. Su, “The roboconsultant: Telementoring and remote presence in the operating room during minimally invasive urologic surgeries using a novel mobile robotic interface,” Urology, vol. 70, no. 5, pp. 970–974, 2007.

[3] D. Andersen, V. Popescu, M. E. Cabrera, A. Shanghavi, B. Mullis, S. Marley, G. Gomez, and J. P. Wachs, “An augmented reality-based approach for surgical telementoring in austere environments,” Military medicine, vol. 182, no. suppl_1, pp. 310–315, 2017.

[4] D. Andersen, V. Popescu, M. E. Cabrera, A. Shanghavi, G. Gomez, S. Marley, B. Mullis, and J. P. Wachs, “Medical telementoring using an augmented reality transparent display,” Surgery, vol. 159, no. 6, pp. 1646–1653, 2016.

[5] A. Baak, M. Müller, G. Bharaj, H.-P. Seidel, and C. Theobalt, “A data-driven approach for real-time full body pose reconstruction from a depth camera,” in Consumer Depth Cameras for Computer Vision, Springer, 2013, pp. 71–98.

[6] M. Dou, S. Khamis, Y. Degtyarev, P. Davidson, S. R. Fanello, A. Kowdle, S. O. Escolano, C. Rhemann, D. Kim, J. Taylor, et al., “Fusion4d: Real-time performance capture of challenging scenes,” ACM Transactions on Graphics (TOG), vol. 35, no. 4, p. 114, 2016.

[7] I. Endo, R. Matsuyama, R. Mori, K. Taniguchi, T. Kumamoto, K. Takeda, K. Tanaka, A. Köhn, and A. Schenk, “Imaging and surgical planning for perihilar cholangiocarcinoma,” Journal of Hepato-Biliary-Pancreatic Sciences, vol. 21, no. 8, pp. 525–532, 2014.

[8] B. S. Hasler, B. Spanlang, and M. Slater, “Virtual race transformation reverses racial in-group bias,” PloS one, vol. 12, no. 4, e0174965, 2017.

[9] F. King, J. Jayender, S. K. Bhagavatula, P. B. Shyn, S. Pieper, T. Kapur, A. Lasso, and G. Fichtinger, "An immersive virtual reality environment for diagnostic imaging," Journal of Medical Robotics Research, vol. 1, no. 01, p. 1 640 003, 2016.

[10] S. F. van der Land, A. P. Schouten, F. Feldberg, M. Huysman, and B. van den Hooff, "Does avatar appearance matter? how team visual similarity and member–avatar similarity influence virtual team performance," Human Communication Research, vol. 41, no. 1, pp. 128–153, 2015.

[11] J. Maier, M. Weiherer, M. Huber, and C. Palm, "Imitating human soft tissue on basis of a dual-material 3d print using a support-filled metamaterial to provide bimanual haptic for a hand surgery training system," Quantitative imaging in medicine and surgery, vol. 9, no. 1, p. 30, 2019.

[12] N. Martelli, C. Serrano, H. van den Brink, J. Pineau, P. Prognon, I. Borget, and S. E. Batti, "Advantages and disadvantages of 3-dimensional printing in surgery: Asystematic review," Surgery, vol. 159, no. 6, pp. 1485–1500, 2016, ISSN: 0039-6060.

[13] T. Mazur, T. R. Mansour, L. Mugge, and A. Medhkour, "Virtual reality–based simulators for cranial tumor surgery: A systematic review," World neurosurgery, vol. 110, pp. 414–422, 2018.

[14] T. Morineau, X. Morandi, N. L. Moëllic, S. Diabira, L. Riffaud, C. Haegelen, P.-L. Hénaux, and P. Jannin, "Decision making during preoperative surgical planning," Human Factors, vol. 51, no. 1, pp. 67–77, 2009.

[15] T. A. Ponsky, M. Schwachter, J. Parry, S. Rothenberg, and K. M. Augestad, "Telementoring: The surgical tool of the future," European Journal of Pediatric Surgery, vol. 24, no. 04, pp. 287–294, 2014.

[16] Y. Pulijala, M. Ma, M. Pears, D. Peebles, and A. Ayoub, "Effectiveness of immersive virtual reality in surgical training—a randomized control trial," Journal of Oral and Maxillofacial Surgery, vol. 76, no. 5, pp. 1065–1072, 2018.

[17] A. V. Reinschluessel, J. Teuber, M. Herrlich, J. Bissel, M. van Eikeren, J. Ganser, F. Koeller, F. Kollasch, T. Mildner, L. Raimondo, et al., "Virtual reality for user-centered design and evaluation of touch-free interaction techniques for navigating medical images in the operating room," in Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM, 2017, pp. 2001–2009.

[18] B. Reitinger, A. Bornik, R. Beichel, and D. Schmalstieg, "Liver surgery planning using virtual reality," IEEE Computer Graphics and Applications, vol. 26, no. 6, pp. 36–47, Nov. 2006, ISSN: 0272-1716.

[19] J. Rosser, M. Wood, J. Payne, T. Fullum, G. Lisehora, L. Rosser, P. Barcia, and R. Savalgi, "Telementoring," Surgical endoscopy, vol. 11, no. 8, pp. 852–855, 1997.

[20] B. El-Sabawi and W. Magee III, "The evolution of surgical telementoring: Current applications and future directions," Annals of translational medicine, vol. 4, no. 20, 2016.

[21] A. Schenk, D. Haemmerich, and T. Preusser, "Planning of image-guided interventions in the liver," IEEE pulse, vol. 2, no. 5, pp. 48–55, 2011.

[22] J. H. Shuhaiber, "Augmented reality in surgery," Archives of surgery, vol. 139, no. 2, pp. 170–174, 2004.

[23] R. Sodian, D. Schmauss, C. Schmitz, A. Bigdeli, S. Haeberle, M. Schmoeckel, M. Markert, T. Lueth, F. Freudenthal, B. Reichart, et al., "3-dimensional printing of models to create custom-made devices for coil embolization of an anastomotic leak after aortic arch replacement," The Annals of thoracic surgery, vol. 88, no. 3, pp. 974–978, 2009.

[24] R. H. Taylor, A. Menciassi, G. Fichtinger, P. Fiorini, and P. Dario, "Medical robotics and computer-integrated surgery," in Springer handbook of robotics, Springer, 2016, pp. 1657–1684.

[25] J. Wang, H. Suenaga, K. Hoshi, L. Yang, E. Kobayashi, I. Sakuma, and H. Liao, "Augmented reality navigation with automatic marker-free image registration using 3-d image overlay for dental surgery," IEEE transactions on biomedical engineering, vol. 61, no. 4, pp. 1295–1304, 2014.

[26] E. Watanabe, M. Satoh, T. Konno, M. Hirai, and T. Yamaguchi, "The trans-visible navigator: A see-through neuronavigation system using augmented reality," World neurosurgery, vol. 87, pp. 399–405, 2016.

[27] J. H. Yoon, J.-I. Choi, Y. Y. Jeong, A. Schenk, L. Chen, H. Laue, S. Y. Kim, and J. M. Lee, "Pre-treatment estimation of future remnant liver function using gadoxetic acid mri in patients with hcc," Journal of hepatology, vol. 65, no. 6, pp. 1155–1162, 2016.

[28] N. N. Zein, I. A. Hanouneh, P. D. Bishop, M. Samaan, B. Eghtesad, C. Quintini, C. Miller, L. Yerian, and R. Klatte, "Three-dimensional print of a liver for preoperative planning in living donor liver transplantation," Liver Transplantation, vol. 19, no. 12, pp. 1304–1310, 2013.

# Interaktive, browserbasierte 3D-Darstellung von medizinischen Planungsdaten

*A. Leuchtenberg[1,2], O. Burgert[1], M. Schöller[1]*

*[1]Hochschule Reutlingen, Informatik Fakultät, Reutlingen, Deutschland*
*[2]Technische Hochschule Mittelhessen, Fachbereich Gesundheit, Gießen, Deutschland*

Kontakt: adina.leuchtenberg@ges.thm.de

***Abstract***

*In der Orthopädie werden Robotersysteme bereits seit mehreren Jahren erfolgreich unterstützend eingesetzt. Dieser Ansatz erfordert die vorgelagerte Erstellung eines digitalen Modells auf Basis von medizinischen Bilddatensätzen. Die Erstellung und Überprüfung der Modelle soll in einer browserbasierten Client-Server-Anwendung erfolgen. Hierfür ist die Darstellung von zweidimensionalen und dreidimensionalen Datensätzen erforderlich. Basis dieses Papers ist die Entwicklung eines Ansatzes zur interaktiven, browserbasierten dreidimensionalen Darstellung medizinischer Planungsdaten. Die Anwendung stellt ein Proof of Concept dar, ob die bestehenden Desktopanwendungen zur Darstellung von Planungsdaten, ersetzt werden können. Mit Hilfe des Frameworks AMI.js wurde die Anwendung umgesetzt. Sie erfüllt alle definierten Anforderungen und kann somit die aktuellen Desktopanwendungen ersetzen.*

**Keywords**: AMI.js, DICOM, Webanwendung

## 1 Problemstellung

Das vorliegende Paper stellt die Ergebnisse einer Evaluation verschiedener Frameworks, für die Anwendung eines konkreten Unternehmens dar. Der Anwendungsfall für das Unternehmen fokussiert dabei auf Hüft- und Knieendoprothetik. Dabei können Roboter unterstützend eingesetzt werden. Mit Hilfe eines Roboterarms werden Eingriffe präziser und sicherer.

Um jedoch eine Operation mit Unterstützung durchführen zu können, muss diese Operation vorher geplant werden. Dafür existiert spezielle Planungssoftware. In diesen Desktopanwendungen werden medizinische Daten als dreidimensionale Modelle dargestellt.

Die Installation und Wartung der Planungssoftware erfolgten auf entsprechenden Computersystemen. Ziel der Arbeit war es, zu analysieren und auszuwerten, ob diese Anwendungen durch browserbasierte Anwendungen ersetzt werden können. Ist dies mit den aktuellen Webtechnologien möglich, entfallen die Installation und Wartung der Anwendungen auf den Planungsrechnern. Die Anwendung kann an zentraler Stelle, wie zum Beispiel der Cloud oder einem separat bereitgestellten Server installiert, gewartet und aktualisiert werden. Für die Ausführung der Anwendung kann zusätzlich nahezu jeder beliebige Standardrechner verwendet werden.

Nachdem die Anforderungen für die zu entwickelnde Anwendung definiert wurden, wurden verschiedene Frameworks für die browserbasierte Darstellung von medizinischen Daten auf ihre Eignung analysiert. Für die Auswertung der Ersetzbarkeit der Desktopanwendung wurde dann eine prototypische Anwendung entwickelt. Anhand dieser ist analysiert worden, ob eine browserbasierte Anwendung die Desktopanwendung ersetzen kann. Auf Datensicherheitsaspekte wurde in der Arbeit nicht eingegangen, da diese nur eine Evaluation verschiedener Frameworks darstellen sollte.

## 2 Material und Methoden

Die Anwendung wurde entwickelt, um die aktuellen Anwendungen im Planungsprozess einer Operation zu ersetzen. Dafür wurden über Use Cases, Funktionale (F) und Nicht-Funktionale (NF) Anforderungen definiert. Anhand der Anforderungen wurden Kriterien zum Vergleich von bereits existierenden Frameworks ausgearbeitet.

Nach einer Recherche über Frameworks zur Darstellung von medizinischen Daten wurden im Anschluss die drei Frameworks XTK.js, AMI.js und X3DOM ausgewählt und in den verschiedenen Kriterien verglichen um zu entscheiden, welches Framework das geeignetste für die Anwendung ist. Daraufhin wurde die Anwendung prototypisch mit dem ausgewählten Framework entwickelt.

## 2.1 Anforderungen an die Anwendung

Die Anforderungen wurden in Muss-, Soll- und Kann-Anforderungen aufgeteilt. Die folgende Tabelle 1 zeigt jedoch nur die Funktionalen Muss-Anforderungen und die Nicht-Funktionalen Anforderungen.

Tabelle 1: *Funktionale Muss und Nicht-Funktionale Anforderungen*

| | Anforderung |
|---|---|
| F010 | Das System muss dem Anwender die Möglichkeit bieten, einen Patienten anhand einer ID auszuwählen. |
| F020 | Das System muss dem Anwender die Möglichkeit bieten, eine Studie von einem bestimmten Patienten über eine ID auszuwählen. |
| F030 | Das System muss dem Anwender die Möglichkeit bieten, eine Serie von einer Studie anhand einer ID auszuwählen. |
| F040 | Das System muss dem Anwender die Möglichkeit bieten, ein Implantat über den Typ und die Größe auszuwählen. |
| F050 | Das System muss fähig sein, einen DICOM-Datensatz von dem DICOM Provider zu laden. |
| F060 | Das System muss fähig sein, ein Implantat von dem Implantat Provider zu laden. |
| F070 | Das System muss fähig sein, einen DICOM-Datensatz und ein Implantat als dreidimensionales Modell zu rendern. |
| F080 | Das System muss dem Anwender die Möglichkeit bieten, die dreidimensionale Darstellung in den Freiheitsgraden zu drehen. |
| F090 | Das System muss dem Anwender die Möglichkeit bieten, das Implantat relativ zu dem DICOM-Datensatz in den Freiheitsgraden zu bewegen. |
| | |
| NF010 | Die Ladezeit der Bilddaten bei einer Datensatzgröße von 50MB muss unter 10 Sekunden sein. |
| NF020 | Die Framerate darf bei folgender Hardware, Prozessor: Intel(R) Core(TM) i7-6820HQ, CPU: 2.70GHz, RAM 8.00GB, GPU: NVIDIA Quadro M2000M, 8GB bei einer Datensatzgröße von 50MB, ohne Interaktion des Anwenders nicht unter 30 FPS fallen. |
| NF030 | Die Anwendung muss allein durch den Browser Chrome, Version 66.0.3359.117 nutzbar sein, ohne dass für das Laden und Rendern der Daten, zusätzliche Software oder Plug-ins installiert werden müssen. |

## 2.2 Auswertung der Frameworks

Für die Entwicklung der webbasierten Anwendung wurden, für die Darstellung der zwei- und dreidimensionalen Inhalte, Frameworks welche auf WebGL basieren, verwendet. Giokas et al. listen in ihrem Paper über Webanwendungen zur Darstellung von medizinischen Volumendaten drei Frameworks, welche für die grafische

Darstellung von medizinischen Daten in Webanwendungen genutzt werden können. Dies sind XTK, VJS und X3DOM. [1] Mittlerweile wird das Framework VJS laut des Github Repositories A Medical Imaging JavaScript Toolkit (AMI) genannt. Aufgrund der Vielzahl an WebGL Frameworks wird sich auf diese drei als mögliche Frameworks für die Entwicklung der Anwendung, beschränkt.

X3DOM wurde von Behr et al. entwickelt, um dreidimensionale Daten auf Basis von X3D, auch ohne Plug-in, im Browser darstellen zu können. Dabei verfolgt die Darstellung durch das X3DOM Framework einen deklarativen Ansatz, setzt jedoch auf WebGL auf. [2] Birr et al. entwickelten eine Anwendung für Medizinstudenten, welche mit der Anwendung die Anatomie der Leber lernen können. Das Ziel der Entwickler war es, eine Anwendung zu entwickeln, welche patientenspezifische Daten als interaktive dreidimensionale Modelle innerhalb einer Webanwendung darstellt. [3]

Das X Toolkit, auch XTK genannt, ist ein JavaScript Framework um medizinische Bilddaten interaktiv über WebGL darzustellen. Es wurde von Haehn et al. am Kinderkrankenhaus in Boston zur Darstellung von Bildgebungen des zentralen Nervensystems entwickelt. [4]

A Medical Imaging JavaScript Toolkit (AMI) ist ein Framework zur Darstellung von medizinischen Bilddaten. Es ist der Nachfolger des Frameworks XTK und wird ebenfalls am Bostoner Kinderkrankenhaus entwickelt. Es ist genau wie XTK Open Source und wurde entwickelt, um verschiedenste medizinische Bilddaten zwei- und dreidimensional darzustellen. Dabei setzt es nicht direkt auf WebGL auf, sondern verwendet das JavaScript Framework Three.js. [5] Three.js ist ein Framework für die Darstellung von dreidimensionalen Inhalten über WebGL. Es verbirgt die genauere WebGL- Struktur und -Programmierung und vereinfacht somit die Erstellung von dreidimensionalen Szenen erheblich. [6]

Die drei Frameworks wurden in verschiedenen Kriterien, welche aus den Anforderungen entstanden sind, miteinander verglichen. Kriterien waren beispielsweise, das Datum des letzten Releases, dass die Technologie Open Source ist oder welche Modelltypen bereits geladen werden können. Beim Vergleich der Frameworks, mit Hilfe von kleinen Testanwendungen, kam heraus, dass das Framework AMI.js für die Umsetzung der Anforderungen am geeignetsten ist. Dafür sprechen vor allem die Aktualität des Frameworks und dass komprimierte und unkomprimierte DICOM-Daten geladen werden können. Im Gegensatz hierzu kann XTK nur unkomprimierte DICOM-Daten laden. Zusätzlich hat AMI.js eine schnellere Ladezeit wie XTK.js und alle Funktionalitäten von Three.js können ebenso genutzt werden. Das Framework X3DOM schied aus, da es trotz Komponente, welche eigens für die Darstellung von medizinischen Daten entwickelt wurde, laut Literatur nicht an die Performanz und Qualität einer Desktopanwendung herankommt. [7]

# 3 Ergebnisse

Die Anwendung wurde mit Hilfe des Frameworks AMI.js entwickelt. Bei der Entwicklung wurde dabei ein angepasstes Modell View Controller Pattern umgesetzt. Die Anbindung an die jeweiligen Server, zum Laden der Daten, erfolgte über REST-Schnittstellen.

Abbildung 1 zeigt die komplette Nutzeroberfläche der Anwendung. Die linke Seite zeigt die Multiplanaransicht bereits mit einem DICOM-Datensatz geladen. Es lässt sich in der linken oberen Ecke das dreidimensionale Modell und in den anderen Teilen jeweils die zweidimensionalen Ansichten erkennen. Die rechte Seite der Oberfläche der Anwendung besteht aus den Interaktionsmöglichkeiten für den Nutzer. Dieser kann DICOM-Datensätze, Implantate und Segmentierungsdaten auswählen. Die Modelle werden jeweils von einem Server, welcher die Daten bereitstellt, geladen. Außerdem kann der Anwender das Implantat in den Freiheitsgraden drehen und bewegen.

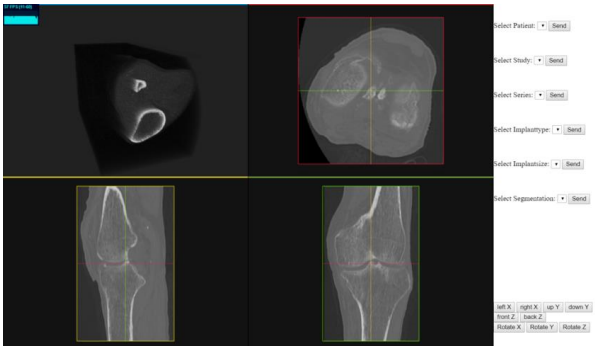



Abbildung 1: *Benutzeroberfläche der Anwendung*

Abbildung 2 zeigt die Multiplanaransicht mit einem Segmentierungs-Datensatz geladen, dargestellt in gelb. Außerdem ist ein Implantat in rot dargestellt. Beides sind STL-Dateien.

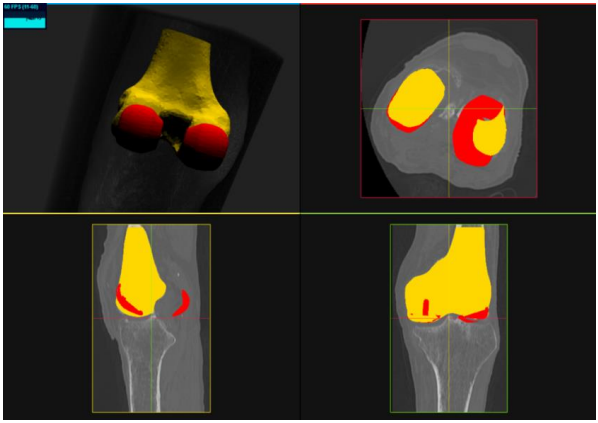


Abbildung 2: *Implantat und Segmentierung innerhalb der Anwendung*

## 4 Diskussion

Das Ziel der Arbeit war es, zu analysieren ob bestehende Desktopanwendungen für die Darstellung von medizinischen Daten, insbesondere Planungsdaten von Hüft- und Knieoperationen, durch browserbasierte Anwendungen ersetzt werden können. Um dies zu überprüfen wurde eine interaktive Anwendung entwickelt, welche DICOM-Datensätze und Implantats- bzw. Segmentierungsdaten als zweidimensionale und als dreidimensionale Modelle rendern kann. Der Anwender kann zusätzlich mit der dreidimensionalen Szene interagieren, indem er heranzoomen, sie bewegen oder drehen kann. Auch ist es möglich, durch die einzelnen Schichten der zweidimensionalen Daten zu navigieren. Ebenso ist die Anwendung an Drittsysteme wie ein DICOM-Server angebunden.

Das Framework AMI.js war als Basis für die Entwicklung der Anwendung sehr gut geeignet. So bringt es Funktionalitäten wie das Laden und Rendern der Daten mit. Die zahlreichen Beispiele, wie die Multiplanaransicht, welche in dessen Github Repository bereitgestellt werden, konnten als Einstieg für die Implementierung genutzt und erweitert werden.

Ein Problem bei der Entwicklung der Anwendung war das Laden des DICOM- Datensatzes über den DICOM-Server. Der erste Versuch war es, den gesamten Datensatz als Zip-Archiv herunterzuladen, es dann zu entpacken und die entpackten Daten an AMI zu übergeben. Jedoch hat AMI die entpackten Daten nun nicht mehr als DICOM-Daten erkannt. Auch die Daten als Zip-Datei zu speichern und lokal zu laden um sie dann erst zu entpacken führte zu demselben Fehler. Der nächste Versuch war die entpackten Daten lokal auf dem Rechner zu speichern um sie im Anschluss zu laden. Auch dies hat nicht die gewünschte Lösung gebracht. Der Browser Chrome hat nur die letzte Datei gespeichert. Als Alternative wurde der Browser Firefox getestet, doch hier musste jedes Mal auf den Button Speichern geklickt werden. Dies wird bei einem Datensatz mit über 200 Bildern zu umständlich. Gelöst werden konnte das Problem nur, indem jedes Bild einzeln vom Server geladen wird und nicht alle zusammen als Zip-Daten. Somit kann AMI den Link der einzelnen Bilder direkt interpretieren.

Ein weiteres Problem war die Datensatzgröße. Beim Versuch einen DICOM- Datensatz, ein Implantat und eine Segmentierung mit einer gesamten Größe von 400MB zu laden brach WebGL und somit die gesamte Anwendung zusammen. WebGL bekommt nur 2GB des GPU-Speichers zugewiesen. Diese sind fest und können nicht verändert werden. Somit konnte der Datensatz nicht geladen werden und es musste sich auf einen kleineren Datensatz beschränkt werden. Im Verlauf der Entwicklung wurden somit immer nur Datensätze von bis zu 50MB geladen. Diese Laden problemlos.

Das Rendern des dreidimensionalen Modells erreichte Anfangs nur maximal 10 Frames pro Sekunde. Daraufhin wurde es nur noch nach Interaktion des Nutzers gerendert um bei einem Datensatz mit einer Größe von 50MB mindestens 30 Frames pro Sekunde zu erzielen. Ohne eine Interaktion bei dem dreidimensionalen Modell liegen die Frames pro Sekunde zwischen 50 und 60.

## 5 Zusammenfassung

Die entwickelte Anwendung zeigt, dass alle zuvor definierten Funktionalen- und Nicht-Funktionalen Anforderungen erfüllt werden konnten. Die Anfangs definierte Bedingung, die Desktopanwendungen können ersetzt werden, wenn die browserbasierte Anwendung alle Muss-Anforderungen erfüllt, ist somit erfüllt. Neben den Muss-Anforderungen sind auch die Kann- und Soll-Anforderungen erfolgreich umgesetzt worden.

## Referenzen

[1] Giokas K, Bokor L, Hopfgartner F, Arbelaiz A, Moreno A, Kabongo L, et al., editors, *Volume Visualization Tools for Medical Applications in Ubiquitous Platforms*, EHealth 360. Springer International Publishing, (2017).

[2] Behr J, Eschler P, Jung Y, Michael Z, *X3DOM: A DOM-based HTML5/X3D integration model*, Proceedings of the 14th International Conference on 3D Web Technology, ACM, Darmstadt, Germany, S. 127–135 (2009).

[3] Birr S, et al., *The LiverAnatomyExplorer: A WebGL-Based Surgical Teaching Tool*, IEEE Computer Graphics and Applications, 33.5, S. 48–58 (2013).

[4] Haehn D, Rannou N, Ahtam B, Grant E, Pienaar R, *Neuroimaging in the Browser using the X Toolkit*, Frontiers in Neuroinformatics, 101 (2014).

[5] Rannou N, Bernal-Rusiel JL, Haehn D, Grant PE, Pienaar R, *Medical imaging in the browser with the A* Medical Imaging (AMI) tool- kit*, ESMRMB Annual Scientific Meeting 2017, (2017).

[6] Cabello R, *Three.js*, (2018). Available from: https://github.com/mrdoob/three.js/.

[7] Arbelaiz A, Moreno A, Kabongo L, Garc´ıa-Alonso A, *X3DOM volume rendering component for web content developers*, Multimedia Tools and Applications, 76(11) (2017).

# VR Multi-user Conference Room for Surgery Planning

*Oleksii Bashkanov*[1], *Patrick Saalfeld*[1], *Hariharasudhan Gunasekaran*[1], *Mathews Jabaraj*[1], *Bernhard Preim*[1], *Tobias Huber*[2], *Florentine Hüttl*[2], *Werner Kneist*[2], *Christian Hansen*[1]

[1] Faculty of Computer Science & Research Campus *STIMULATE*, Otto-von-Guericke University Magdeburg, Germany

[2] Department of General, Visceral and Transplant Surgery, University Medicine of the Johannes Gutenberg-University Mainz, Germany

Contact: saalfeld@isg.cs.uni-magdeburg.de

***Abstract***

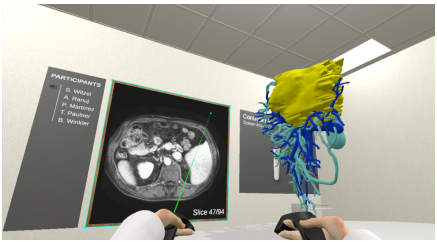
*Preoperative planning is a fundamental precondition for the success of the surgery. In the course of planning, the appropriate decision making must take into account the individual anatomical characteristics of the organs and the patients physical condition. Virtual reality (VR) based systems enable interaction with 3D organ models, which allows surgeons to mentally reconstruct the patient-specific organ structure more easily. Furthermore, the importance of proper team interaction and collaboration among surgeons must not be underestimated. In this work, we present the prototype for a multi-user conference room for surgery planning inside VR, where users can benefit from interaction with 3D organ models as well as 2D gray-value images. This system also enables the discussion of the surgical problems over distance. We chose liver surgery planning for evaluation purposes, but this prototype is also functional for planning other surgical procedures. A pilot study showed that surgeons found this tool helpful in preoperative planning routines. They suggest enhancements relating to avatar appearance and advance 3D model interaction.*

**Keywords**: Virtual reality, human-computer-interaction, surgical decision support, surgical planning

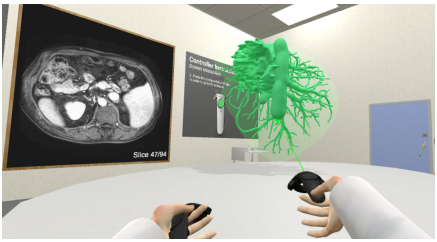
## 1 Introduction

Successful surgery can be ensured only if surgeons can mentally build spatial relationships between anatomical and pathological structures. Additionally, preoperative planning of surgery procedures highly depends on computer assistance [1]. To support surgeons in preoperative planning, various planning software solutions were introduced [2]. As a basis, these solutions mostly visualize 3D medical image data as 2D slices. Surgeons use these to mentally reconstruct the internal representation of the liver and building spatial relationships of the particular patient. This task can be challenging even for well-trained surgeons [3, 4]. A 3D organ visualization can provide substantial support for such actions. Several possibilities allow obtaining a 3D model with detailed inner structures of the liver surfaces, its vascular structures, and tumors from the tomographic data. Exploring the 3D liver model on a standard 2D display dismiss the benefit of 3D display techniques since they cannot convey depth cues such as binocular disparity and motion parallax. In contrast, a VR system mimics how we perceive the physical world. With head-mounted displays (HMD), the user gets the impression of seeing real 3D objects [5]. Therefore, VR has proved itself to be an effective tool for numerous surgical simulations, including the training of fundamental surgical skills used in laparoscopic surgery [6, 7].

The surgical team should share a mental model [8] based on information about the liver characteristics to anticipate surgical errors and ensure patients safety [9]. Thus, collaborative VR interaction techniques are crucial to perform surgery planning. Exchange of experience and knowledge among physicians is mostly performed via face to face interaction, team meetings, video-/phone call or even social media [10]. All these variants have limitations. Face to face interaction lowers possibilities for miscommunication; however, it is not possible over different locations. Video/-phone calls raise the chance for miscommunication because the participants only can communicate with each other in a limited way. With all these solutions, the data is still viewed on a flat screen. Our paper proposes solutions, where such a form of communication takes place in a shared virtual environment, thus preventing misinterpretation of difficult concepts. Such multi-user virtual interaction also enables the exchange of competence across distances. There are scenarios where physicians need to discuss a complex surgical case while being located across the globe. Furthermore, through raising the accessibility of VR technology and its integrity, the overall communication process among surgeons can be improved.




(a) *By pointing to the DICOM viewer with the virtual ray and pressing the touchpad on different positions, the user is able to scroll through the DICOM slices.*




(b) *The user interacts with the 3D liver planning model by pointing with a virtual ray on it. Now, she is able to move and rotate it.*

Figure 1: *Interaction with medical 3D and 2D data.*

## 2 Materials and Methods

To realize the immersive environment, we chose the low-cost and accessible VR headset *HTC Vive*. It provides a wide field of view of 110 ° and offers a large tracking area of roughly 4.5 × 4.5 meters, allowing the user to experience spatial immersion in a room-scale virtual environment [5]. For interaction, we use standard HTC Vive controllers that provide six degrees of freedom and sub-millimeter tracking accuracy.

We use the game engine *Unity* as a development environment since it natively supports the SteamVR platform, which provides a single interface that operates with all major VR headsets, including the HTC Vive. Additionally, the *Virtual Reality Toolkit* (VRTK) is used as it provides common VR solutions such as grabbing and locomotion. Voice transmission is included to enable voice communication if users are at different locations. The users are represented with humanoid avatars. These components are described in more detail below.

### 2.1 Virtual Conference Room

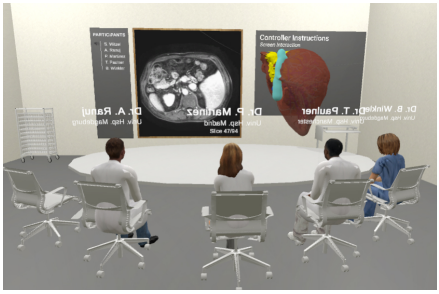
Room-scale in VR plays a significant role in ensuring a user-friendly experience. Ideally, virtual DICOM screen, 3D organ model, and other users avatars should be displayed together in sight to minimize mental effort of switching attention between them. After trying several setup options, we decided on room arrangement with an oval table, with all users sitting at one side, and DICOM screen, on the opposite side of the table (see Fig. 2a). The 3D organ model is floating above the table. The seat arrangement is organized in such fashion that seats are taken from edges to the center as users connect to the application since the seats on the edges allow unrestricted viewing of the 2D image and 3D liver model.

### 2.2 Virtual Object Interaction

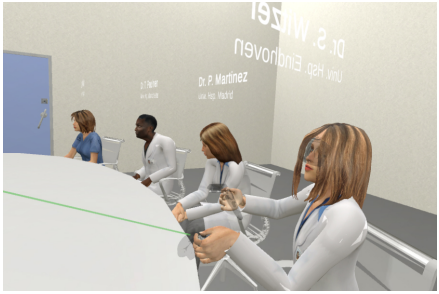
We employ a detailed 3D liver planning model with meta-information containing the liver surface, segmented tumor(s), and vascular system that was bought over a MeVis distant service. Additionally, we load the corresponding DICOM data set containing the patients liver into the virtual environment. A direct grabbing technique [11], where the user touches the object with the controller to enable the interaction with it, is not appropriate, because not all users can reach virtual object or virtual DICOM screen. Instead, a virtual ray is used for interaction. To change the current slice of the DICOM viewer, the user points the virtual ray to the virtual screen and use the touchpad to scroll through the slices (see Fig. 1a).

The interaction with the 3D liver planning model is realized through a direct and indirect technique, respectively. The direct technique allows the user to grab the model with the virtual ray (see Fig. 1b). Once the trigger button is pressed, the model is *pinned* to the ray and can be moved and rotated simultaneously. Although this technique is natural and easy to learn, it makes it difficult to rotate the model in a specific orientation without moving it. Therefore, a rotation mode is used, allowing to fix the object position while rotating it.

Each structure of the 3D liver model can be shown or hidden. The user can control the appearance of these structures with a combination of pressed buttons and the touchpad on the HTC Vive controller. During virtual surgical planning, the aforementioned components can provide surgeons with essential visual support.




(a) *The virtual environment is modelled as a conference room. Users sit around a table, facing a DICOM viewer at a wall and the 3D model is floating over the table.*




(b) *Female and male avatars. A user can be identified by the name-tags shown above each avatars headset.*

Figure 2: *Virtual environment setup and humanoid avatars.*

## 2.3 Humanoid Avatar

Social presence is an essential factor in ensuring the high quality of communication in a computer-mediated context [12]. Humanoid avatars could enable a higher degree of social presence inside the virtual environment. The systematic conducted by Oh et al. [13] has shown that humanoid avatars, controlled by an actual human, generally could enable a higher degree of social presence in VR than agents, controlled by animation script. It is possible to represent them (1) with complete body and predefined animations, (2) with complete body and a one-to-one mapping of the users movement to the avatars movement and (3) avatar body that consists only of the user's head and hands with a one-to-one mapping of the user. According to the experiment conducted by Heidickers et al. [14], the second option generates the highest co-presence and behavioral interdependence.

The *Autodesk Character Generator* was used to create the avatars. This tool allows exporting FBX files, including materials that can be used directly in Unity. The imported avatar consists of a hierarchical structure of objects representing each body part. We created three female avatars and two male avatars wearing doctor's white coats (see Figures 2a and 2b). Though the avatars do not resemble the original user, they can be identified using the individual name-tag shown above each avatar headset.

Since the HTC Vive provides only three tracking positions (head, two hands via controller), the remaining posture of the avatar has to be calculated. Here, the *Final IK* library was used, which provides an inverse kinematics solver to calculate the position of the non-tracked body parts. With inverse kinematics, e.g., the hand of the avatar can follow one HTC Vive controller. The wrist, elbow, and shoulder joints adjust automatically to maintain their proper orientation toward the hand.

## 2.4 Voice Transmission

Communication is important to convey the information and share knowledge. Real-time communication is vital in a multi-user application where many activities can be done in parallel. During surgery planning, it is crucial for physicians to discuss the complex interior structure of the liver. For example, the translation and rotation of the liver or interaction with DICOM viewer by more than one user needs real-time communication to avoid confusion and improve efficiency. Additionally, voice communication can enhance the sense of immersion [15].

We use the *Photon Voice* library, which uses *Opus codec*, an efficient and high-quality audio format for voice transmission. We used *Photon* as this uses a cloud with dedicated servers to reduce latency. A left screen lists all involved users in the scene (see Figure 3). We allow the viewing of active users, muted users, and currently speaking users. The icon would appear only when the user is active. We use the voice detection feature of the *Photon voice* library to highlight which user is speaking right now by turning the icon green. The icon becomes red when the user is muted. The mute function helps to reduce the echo when the two users are in the same room.

## 2.5 Network Management

Unity provides its own directly integrated multiplayer network functionality. It has a high-level scripting API (HLAPI) that grants access to basic commands that cover most of the standard requirements for a multi-user

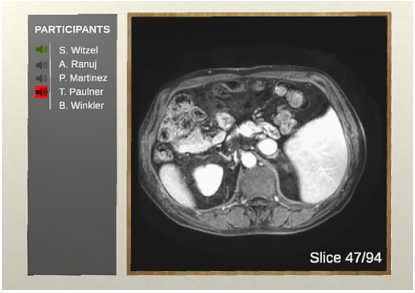



Figure 3: *On the left, a participant list is shown. Speaking users are highlighted with a green speaker icon. It is possible to mute participants per user, which is helpful if two participants are in the same room. On the right, a DICOM viewer displays the transverse plane of a medical data set. All participants can scroll through the slices via the controllers touch-pad.*

application. With the HLAPI the same application runs on every personal computer (PC). One PC takes the role of the server, but since the application on each PC is the same, this role can be taken by any user. The following logic is implemented in the prototype to hide this decision and network management from the physician. On application startup, the prototype connects to a Unity server and searches for an existing running server. If one is found, it connects to it and signs in as a client. If no server is identified, the application takes this role. Therefore, the physicians only have to start the application on every PC in any order.

### 2.6 Pilot Study

We showed our prototype to two liver surgeons in an informal pilot study. They assessed the prototype as helpful for liver surgery planning, already by using it as a single-user application. The spatial impression of the liver supports the definition of resection areas for tumor resection very well. The multi-user aspect enhances this impression further, as the two physicians could point at areas of the liver and discuss a strategy together. However, they criticized the appearance of the avatars. Reasoned by the inverse kinematics approach, the physicians real body posture and virtual posture were not corresponding from time to time. This reduces the feeling of body ownership and, thus, the perceived immersion.

## 3 Conclusion

We presented a virtual conference room for surgery planning on the liver, as an example, where multiple physicians from different locations can connect in a shared environment. They can interact simultaneously with a 3D liver model similarly as one could interact with real 3D objects. This gives numerous possibilities to improve how physicians derive a surgery plan. For example, the efficiency of communication is improved, and ambiguity is reduced by combining voice communication and a visual realistic humanoid representation of each user. Additionally, this representation of users leads to a higher sense of immersion, which is essential for a virtual reality application. Different representations of patient data as 3D planning models and 2D planar images gives the users different perspectives to prepare the surgery plan. However, VR headsets displays are not yet suitable for clinical decision making based on virtual DICOM screen due to display quality limitations. Once VR headsets meet regulatory requirements and obtain standardized display and grayscale resolution, they could be used as a primary visual interface with virtual DICOM screen.

For future work, several aspects can be improved and included to use our prototype as a virtual surgery planning system. One important aspect is to integrate a virtual resection [16] simulation to enable the liver cutting that can be performed by all physicians together. Tools for resection and corresponding simulation methods should be included that could enable multiple users to interact with the liver model in such a way that resection strategies could be devised in real time. The resection plane could add value showing the underlying medical data right on the 3D liver. For the avatar representation, synchronization of the lip movements with voice transmission would make the avatars more realistic. Furthermore, user interaction can be supported by using alternative input devices, such as VR gloves or IR camera systems.

## Acknowledgements

This work was funded by the Federal Ministry of Education and Research of Germany under grant number 16SV8054.

## References

[1] Julie Hallet, Brice Gayet, Allan Tsung, Go Wakabayashi, Patrick Pessaux, and 2nd International Consensus Conference on Laparoscopic Liver Resection Group, *Systematic review of the use of pre-operative simulation and navigation for hepatectomy: current status and future perspectives*, Journal of Hepato-Biliary-Pancreatic Sciences, **22**(**5**), pp. 353–362 (2015)

[2] Christian Hansen, Stephan Zidowitz, Bernhard Preim, Gregor Stravrou, Karl Oldhafer, and Horst Hahn, *Impact of Model-based Risk Analysis for Liver Surgery Planning*, International Journal of Computer Assisted Radiology and Surgery, **9**(**2**), pp. 473–480 (2014)

[3] Vimalraj Velayutham, David Fuks, Takeo Nomi, Yoshikuni Kawaguchi, and Brice Gayet, *3D visualization reduces operating time when compared to high-definition 2D in laparoscopic liver resection: a case-matched study*, Surgical endoscopy, **30**(**1**), pp. 147–153 (2016)

[4] Caitlin T Yeo, Andrew MacDonald, Tamas Ungi, Andras Lasso, Diederick Jalink, Boris Zevin, Gabor Fichtinger, and Sulaiman Nanji, *Utility of 3D reconstruction of 2D liver computed tomography/magnetic resonance images as a surgical planning tool for residents in liver resection surgery*, Journal of surgical education, **75**(**3**), pp. 792–797 (2018)

[5] Joshua Q Coburn, Ian Freeman, and John L Salmon, *A review of the capabilities of current low-cost virtual reality technology and its potential to enhance the design process*, Journal of computing and Information Science in Engineering, **17**(**3**), p. 031013 (2017)

[6] Tobias Huber, Markus Paschold, Christian Hansen, Tom Wunderling, Hauke Lang, and Werner Kneist, *New dimensions in surgical training: immersive virtual reality laparoscopic simulation exhilarates surgical staff*, Surgical endoscopy, **31**(**11**), pp. 4472–4477 (2017)

[7] Tobias Huber, Tom Wunderling, Markus Paschold, Hauke Lang, Werner Kneist, and Christian Hansen, *Highly Immersive Virtual Reality Laparoscopy Simulation: Development and Future Aspects*, International Journal of Computer Assisted Radiology and Surgery, **13**(**2**), pp. 281–290 (2018)

[8] Michael J Burtscher and Tanja Manser, *Team mental models and their potential to improve teamwork and safety: a review and implications for future research in healthcare*, Safety Science, **50**(**5**), pp. 1344–1354 (2012)

[9] Yael Einav, Daniel Gopher, Itzik Kara, Orna Ben-Yosef, Margaret Lawn, Neri Laufer, Meir Liebergall, and Yoel Donchin, *Preoperative briefing in the operating room: shared cognition, teamwork, and patient safety*, Chest, **137**(**2**), pp. 443–449 (2010)

[10] Justin P Wagner, Amalia L Cochran, Christian Jones, Niraj J Gusani, Thomas K Varghese Jr, and Deanna J Attai, *Professional use of social media among surgeons: results of a multi-institutional study*, Journal of surgical education, **75**(**3**), pp. 804–810 (2018)

[11] Florian Heinrich, Sebastian Rohde, Tobias Huber, Markus Paschold, Werner Kneist, Hauke Lang, Bernhard Preim, and Christian Hansen, *VR-basierte Interaktion mit 3D-Organmodellen zur Planung und Simulation laparoskopischer Eingriffe*, In: Proceedings of the Annual Meeting of the German Society of Computer- and Robot-Assisted Surgery (CURAC), pp. 57–62 (2018)

[12] Sara Kiesler, Jane Siegel, and Timothy W McGuire, *Social psychological aspects of computer-mediated communication.* American psychologist, **39**(**10**), p. 1123 (1984)

[13] Catherine S Oh, Jeremy N Bailenson, and Gregory F Welch, *A systematic review of social presence: definition, antecedents, and implications*, Front. Robot. AI 5: 114. doi: 10.3389/frobt (2018)

[14] Paul Heidicker, Eike Langbehn, and Frank Steinicke, *Influence of Avatar Appearance on Presence in Social VR*, In: Proceedings of IEEE Symposium on 3D User Interfaces (3DUI) (Poster), pp. 233–234 (2017)

[15] J-C Bolot and Sacha Fosse-Parisis, *Adding voice to distributed games on the internet*, In: Proceedings of IEEE Conference on Computer Communications, pp. 480–487 (1998)

[16] Olaf Konrad-Verse, Arne Littmann, and Bernhard Preim, *Virtual resection with a deformable cutting plane.* In: SimVis, pp. 203–214 (2004)

# Comparing the cutting characteristics of a clinically relevant CO2 laser to a diode pumped Er:YAG laser

*E. Reins[1], H. Wurm[2], K. Stock[2], T.K. Hoffmann[1] , P.J. Schuler[1]*

*[1] Department of Otorhinolaryngology, Head and Neck Surgery*

*Ulm University Medical Center, Frauensteige 12, 89075 Ulm, Germany*

*[2]Institute for Laser Technology in Medicine and Measurement Technique (ILM)*

*Ulm University, Helmholtzstraße 12, 89081 Ulm, Germany*

Contact: elisabeth.reins@uniklinik-ulm.de

***Abstract***

*The $CO_2$ laser causes a comparatively large coagulation zone. The erbium:YAG (Er:YAG) laser achieves less coagulation damage through high tissue absorption. We compared the diode pumped Er:YAG laser of the Institute of Lasertechnology Ulm with the standard $CO_2$ laser. 3 sections (1.5cm) were performed on the mucosa of porcine tongues, respectively. The lasers were used with 7.7W at a speed of 2, 5 and 10mm/s. Azan-stained histological specimens (n = 18) were examined microscopically for depth of cut and width of the coagulation zone. The Er:YAG laser cuts twice as deep as the $CO_2$ laser and causes a lower coagulation width than the $CO_2$. In vivo experiments to assess improved hemostasis and reduced scarring are being prepared.*

*Keywords: laser surgery, CO2, Er:YAG*

## 1 Background

The $CO_2$ laser is frequently used in head and neck surgery, and it is applied for standard procedures such as the stapedotomy in otosclerosis. But during the last years it has gained more importance in the excision of tumors especially of the oropharynx and larynx. Because of its relative high absorption rate, the penetration depth is limited and deeper tissues are not affected (1). Ideally a surgical laser should precisely excise the target while causing local hemostasis without injuring surrounding structures. Through high surrounding temperatures, the $CO_2$ laser causes a large coagulation zone with a possible deterioration of surrounding structures. The erbium-doped yttrium aluminium garnet (Er:YAG) laser has a high absorption coefficient in water ($\lambda$ 2.94 µm) inducing a higher tissue absorption and a smaller coagulation damage as compared to the $CO_2$ laser. Until today, the flashlamp pumped Er:YAG laser has been used in the ablation of soft and hard tissue (2-4). In this study we, therefore, compared our diode pumped Er:YAG laser with the $CO_2$ laser used in standard medical procedures in our ENT department. The new system allows higher repetition up to 2kHz. The diode-pumped laser is basically maintenance-free and guarantees reliable operation for several thousand hours (5).

## 2 Material and Methods

The diameter of focus size for both lasers was 500µm. The size and the focus plane was analyzed by moving an alignment paper through the focus plane area with the translation stage at a speed of 30 mm/s. Therefore the alignment paper was moved in x and z direction at the same time. After determination of the position all cuts were performed with the translation stage in the focal plane. The $CO_2$ laser of the company Lumenis® was set on continuous pulse. The diode pumped Er:YAG laser of the ILM Ulm (DPM40, Pantec Engineering AG) was set on a current of 300A, at repetition rate of 200Hz and pulse duration of 154µs. Both lasers were used with a total power of 7.7W. Multiple laser sections were performed on mucosa of porcine tongues. To this end, equivalent tissue samples (thickness 1cm) were cut from the lateral part of the tongue. The mucosa samples were then placed on a sample holder. The sample holder was moved by a translation stage with constant velocity (**figure 1**). The idea was to vary the energy per position by varying the velocity. In addition, a higher velocity, for example performed by a scanner unit, would decrease procedure duration. We chose 2 mm/s to imitate hand held cutting, 10 mm/s for scanner speeds and a third velocity in between at 5 mm/s. Therefore, three cuts of 1.5cm length were performed at a velocity of 2, 5 and 10mm/s on three samples, respectively.

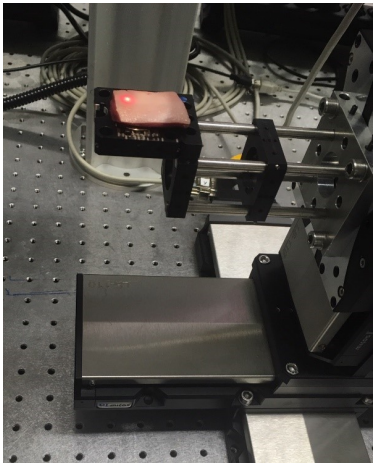
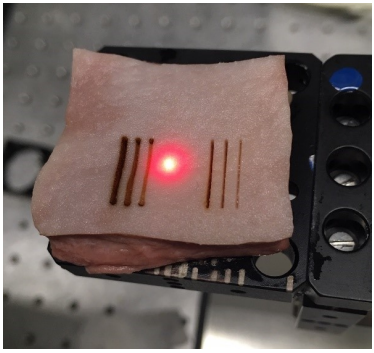
a)  b) 

Figure 1: Sample holder with mucosa samples before a) and after b) sections were performed with the CO2 laser (left side) and Er:YAG (right side) at 2, 5 and 10mm/s, respectively.

After the laser treatment, the samples were stored in formaldehyde for at least 3 days.

Paraffin embedded sections (n=18) were stained with Azan (). The cutting depth and the width of the coagulation zone were analyzed via microscope (ZEISS, Axiophot) on two different histological slides at three points, respectively (**figure 5**).

## 3 Results

At a velocity of 2mm/s (depth of the cuts: 2240 vs. 1298µm), 5mm/s (1124 vs. 551µm) and 10mm/s (710 vs. 380µm) the Er:YAG laser cuts twice as deep as the $CO_2$ laser (**figure 2**). At the same time the Er:YAG laser causes a lower coagulation width than the $CO_2$ laser at a velocity of 2mm/s (96 vs. 234µm), 5mm/s (67 vs. 123µm) and 10mm/s (31 vs. 91µm; **figure 3**).

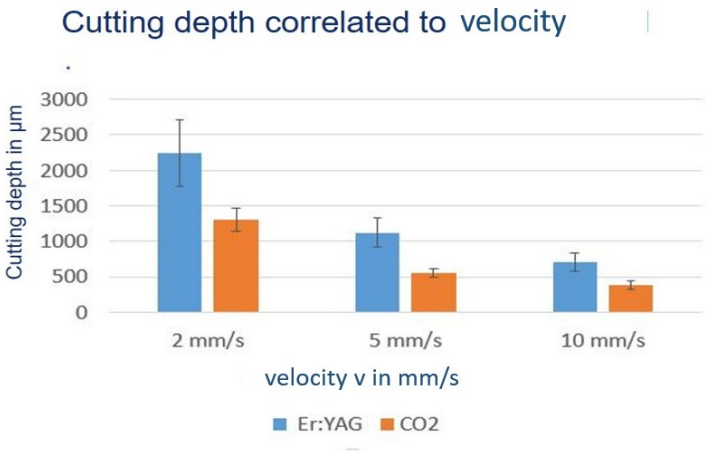



Figure 2: Bar graph showing the cutting depth correlated to velocity at 2, 5 and 10mm/s. Whiskers represent the standard deviation.

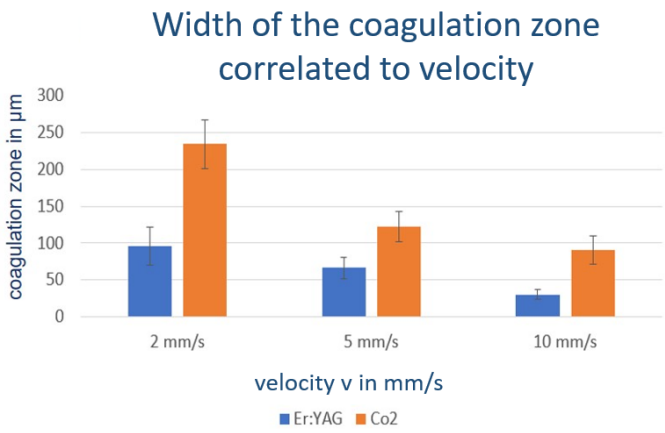



Figure 3: Bar graph showing the width of the coagulation zone correlated to cutting speed at 2, 5 and 10mm/s. Whiskers represent the standard deviation.

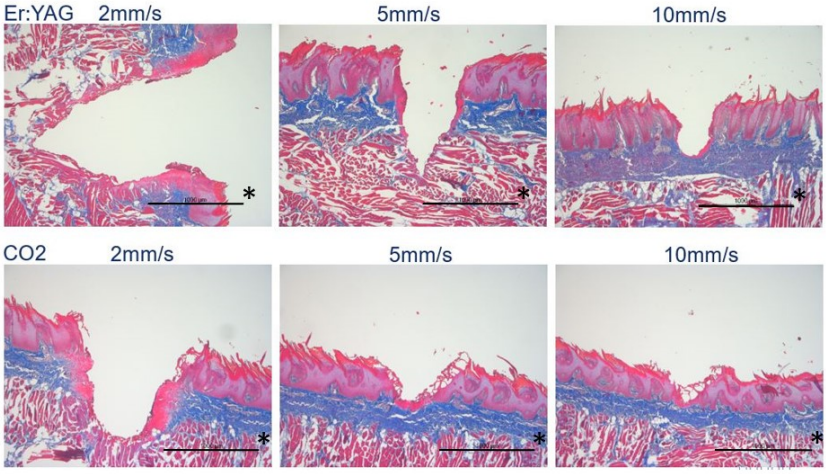



Figure 4: Representative examples of Azan-stained sections. * Scale at 1000µm.

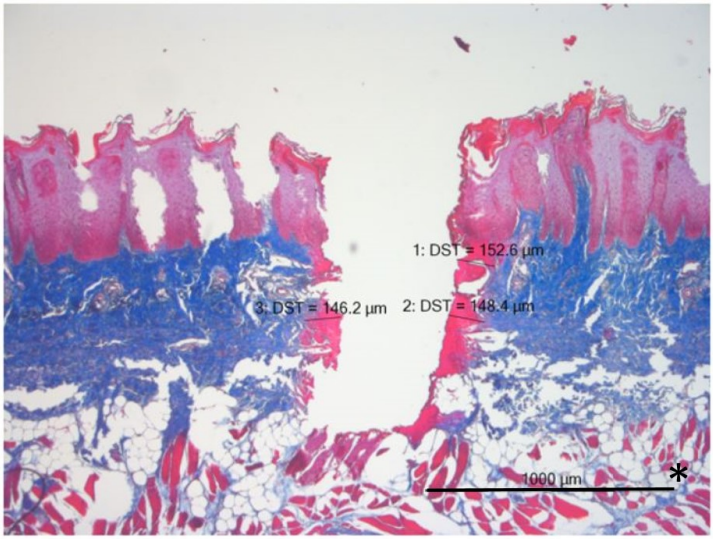



Figure 5: Representative image of Azan-stained section (CO2 Laser at velocity 2mm/s). Width of the coagulation zone was analyzed via microscope on two different histological slides at three points, respectively.

## 4 Conclusion

In the *ex vivo* animal experiments, the Er:YAG laser shows better cutting properties than the $CO_2$ laser. *In vivo* experiments to assess improved hemostasis and reduced scarring are being prepared.

## 5 References

1. Marchese MR, Scorpecci A, Cianfrone F, Paludetti G (2011) "One-shot" CO2 versus Er:YAG laser stapedotomy: is the outcome the same? Eur Arch Otorhinolaryngol. 268(3): p. 351-6.

2. Nagel, D., "The Er:YAG laser in ear surgery: First clinical results," Lasers Surg.Med., 21, 79-87 (1997)

3. Hale GM, Querry M R, (1973) Optical constants of water in the 200-nm to 200-µm wavelength region. Appl. Opt., 12, 555-563 (1973).

4. Hibst, R (1992) Mechanical effects of erbium: YAG laser bone ablation. Lasers Surg Med, 12.2: 125-130

5. Stock K, Meitinger D, Hausladen F, Stegmayer T, Wurm H (2019) Primary investigations on defined thermal effects on soft tissue using a diode pumped Er:YAG laser system. BiOS 2019 Conference paper.

# Computer-assistance in minimally invasive endopancreatic surgery

*Philip C. Müller[1], Caroline Haslebacher[2], Benjamin Eigl[2,3]*

*1Department of Surgery, University Hospital Zürich, Zürich, Switzerland*
*[2]CAScination, Bern, Switzerland*
*[3]ARTORG Center, Bern, Switzerland*

Kontakt: Benjamin.eigl@cascination.com

***Abstract***

*Endopancreatic surgery stands for an experimental minimally invasive technique, where a pancreatic resection is performed with a rigid endoscope from inside the pancreas. Yet, the missing haptic feedback and the decreased field of view pose major challenges for the surgeon. Using computer-assistance pre- and intraoperative image data can be combined to provide additional visualization support. The aim of this proof of concept study was to combine the novel surgical technique (endopancreatic surgery) with computer assistance for the resection of invisible pancreatic tumours. Computer assisted endopancreatic surgery was feasible in a 3D pancreas model and allowed safe resection of the invisible pancreatic tumours. Pure intraparenchymal landmark registration showed the best accuracy with a fiducial registration error of 2.24mm (1.40-2.85). The preliminary results showed an accurate registration using internal landmarks, which is the cornerstone of a reliable image-guided navigation.*

**Keywords**: Pancreatic surgery, pancreas, computer-assistance

## 1 Introduction

Minimally invasive pancreatic surgery is a young surgical field, the first laparoscopic distal pancreatectomy was performed in 1994 and soon after, the first robotic pancreatic resection followed.[1], [2] Since then, minimally invasive resections have become the standard of care for pancreatic left resections in most high volume centers.

As open and minimally invasive pancreatic surgeries are still associated with a high morbidity, current research is looking for further reduction of the surgical trauma to improve the perioperative outcome. A minimization of the surgical trauma may be possible when natural orifices are chosen to access abdominal organs.[3] In natural orifice transluminal endoscopic surgery (NOTES), access to the abdominal cavity is not limited to incisions into the abdominal wall. Rather, NOTES uses natural orifices, such as the vagina, the anus, and the mouth. While pure NOTES, which uses only one natural orifice, is not widely used in surgical practice, hybrid NOTES, which uses small incisions in the abdominal wall plus access through a larger natural orifice, has become a valuable surgical technique.

The concept of hybrid NOTES lead to the idea of approaching the pancreas with a rigid endoscope over its natural connection from the duodenum to the pancreatic duct.[4] Similar to the transurethral access for resection of the prostate, laparoscopic access to the pancreatic duct is established via the duodenum and the papilla. A rigid endoscope is used to enable a stable access and the use of rigid endoscopic instruments. The concept was developed as an alternative to duodenum-preserving pancreatic head resections in benign diseases such as chronic pancreatitis or as an alternative to enucleation for benign cystic lesions. Compared to conventional therapies the so called endopancreatic surgery (EPS) has several advantages: First, there is no need for a pancreaticojejunal anastomosis with the associated risk of insufficiency; second, lower morbidity and enhanced recovery are assumed through the minimally invasive access; and third, the procedure can be repeated as often as required.

We have previously developed EPS as an experimental technique for resection of pancreatic tissue from inside the pancreatic duct in ex-vivo bovine and in-vivo porcine experiments.[4]–[6] In the experimental setting, the resections were made according to the visual feedback from the endoscope; however, the study demonstrated the risk of organ perforation with this technique. Furthermore, the orientation inside the pancreas was challenging and structures not visible by the endoscope are impossible to target.

The aim of this project is to implement computer-assisted image-guidance to assist EPS to achieve a more precise and safer resection. In the long-term, EPS has the potential to fill the gap between endoscopic and surgical treatments for chronic pancreatitis and benign cystic lesions.

## 2 Material and methods

The experimental setup was conducted using a pancreatic phantom, a surgical navigation system with dedicated software for endoscopic targeting of lesions (Figure 1).

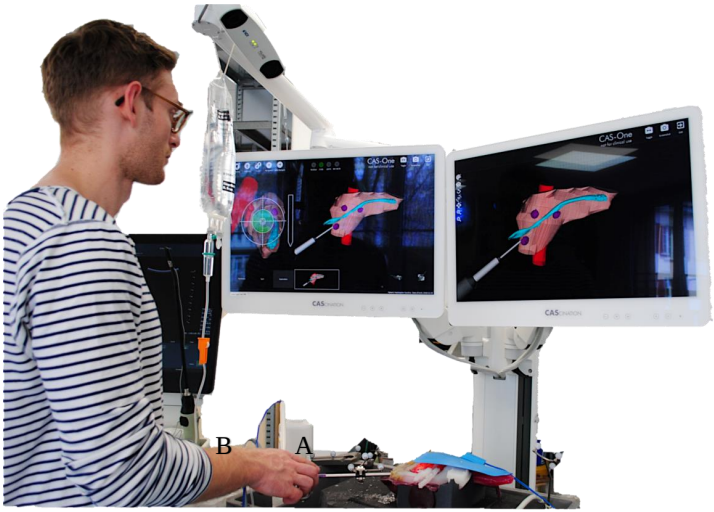



Figure 1: *Experimental setup consisting of the navigation platform, ultrasound device, pancreas phantom (**A**) and endoscopic laparoscope (**B**).*

### 2.1 Endopancreatic approach

The principle of EPS is shown in Figure 2. For the experiment, the 3-D pancreatic model is placed in a box trainer. Two 12-mm and one 5mm trocar are placed. Opposite to the duodenal papilla a 3cm duodenotomy is performed (1A). Another 5-mm trocar is placed on the right lateral abdominal wall close to the duodenum and a rigid endoscope with a diameter of 2.7 mm (Karl Storz GmbH&KG, Tuttlingen, Germany) inserted through the duodenotomy and over the papilla into the pancreatic duct (1C). Purisole solution (Fresenius Kabi, Bad Homburg, Germany) is used for irrigation. Under visual control the endoscope is advanced in the pancreatic duct and resections are performed with a monopolar loop (1D).

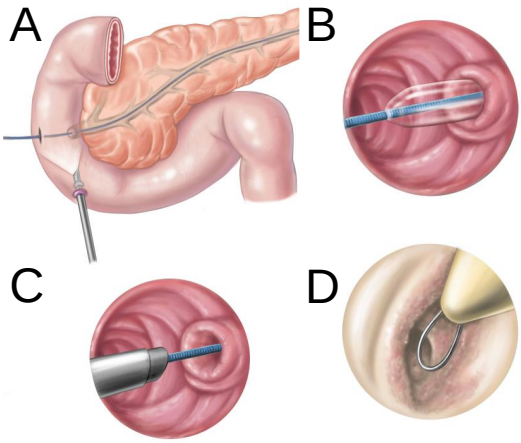



Figure 2: ***A****. A duodenotomy is made opposite to the papilla.* ***B.*** *If necessary, the papilla is dilated with a balloon.* ***C.*** *The endoscope is advanced over the papilla into the pancreatic duct.* ***D.*** *A resection of pancreatic parenchyma is performed from inside the duct.*

## 2.2 Pancreas phantom

In a first step a 3-D pancreatic model, which is compatible with current intraoperative imaging modalities such as ultrasound and computed tomography, was developed. The model includes three intrapancreatic "tumors" and the surrounding structures at risk such as aorta, portal vein, duodenum and the pancreatic papilla. We followed the protocol from Chmarra et al. [7] to achieve appropriate soft tissue characteristics.

## 2.3 Computer-assistance

Computed tomography images of the 3-D printed model are obtained. The images will be segmented with a medical grade computer software to highlight the pancreatic duct, organ borders, the three intrapancreatic tumors and surrounding vessels. Then the original CT-images and the segmentations are transferred to the CAS-One platform (Cascination AG, Bern, Switzerland). The system consists of a processing tower, two wide-screen touch monitors, and an infrared tracking camera (*Figure 1*). The laparoscopic instruments, the ultrasound and the endoscope for the endopancreatic resection are attached with retroreflective markers for spatial tracking using an optical approach. Geometric instrument calibration is carried out to localize the instrument tips as well as to compute the ultrasound parameters. To achieve the registration between the pre-operative data with the intraoperative scene, landmark points are acquired using either a rigid pointing tool or the ultrasound and matched with their 3D counterparts. The CAS-One solution provides the surgeon with the following information: 3D visualization of the tool positions relative to the 3D scene and virtual guidance to navigate towards a defined target.

## 2.4 Study design

The purpose of the study was to evaluate three techniques to achieve best possible local registration. In the first study group only surface landmarks were chosen (Papilla, superior pancreatic border, V. mesenterica superior). In the second group two landmarks were registered by ultrasound (two tumors) and one surface landmark was chosen (Papilla). While in the third group, landmarks were all acquired by ultrasound (three tumors). The measures taken were the Fiducial Registration Error (FRE) as well as the time for the registration attempt. After the registration computer assisted endopancreatic resection of three pancreatic tumours that were endoscopically invisible but

visible on the preoperative CT scan, was attempted. The study was conducted by a trained medical professional and each registration attempt was conducted 5 times.

## 3 Results

Pure intraparenchymal landmark registration (Figure 3) showed the best accuracy with a fiducial registration error of 2.24mm (1.40-2.85), however the registration was more time consuming compared to the surface landmark registration (Table 1). After the registration the segmented CT image was superimposed on the ultrasound image, which enabled the surgeon to find the tumours visible on the CT scan but not on the ultrasound image or the endoscopic image (Figure 4). Computer assisted endopancreatic surgery was feasible and allowed safe resection of the three invisible pancreatic tumours.

Table 1: *Different registration approaches with corresponding accuracy and registration time*.

| Study group | Registration approach | Fiducial Registration Error (mm) | Registration Time (min) |
|---|---|---|---|
| **1** | Surface landmarks | 3.46 (2.25-4.85) | 01:51 (01:33-02:05) |
| **2** | Surface landmarks and intraparenchymal landmarks | 2.46 (1.60-3.35) | 02:12 (01:54-02:33) |
| **3** | Intraparenchymal landmarks | 2.24 (1.40-2.85) | 02:58 (01:42-04:04) |

Median (interquartile range)

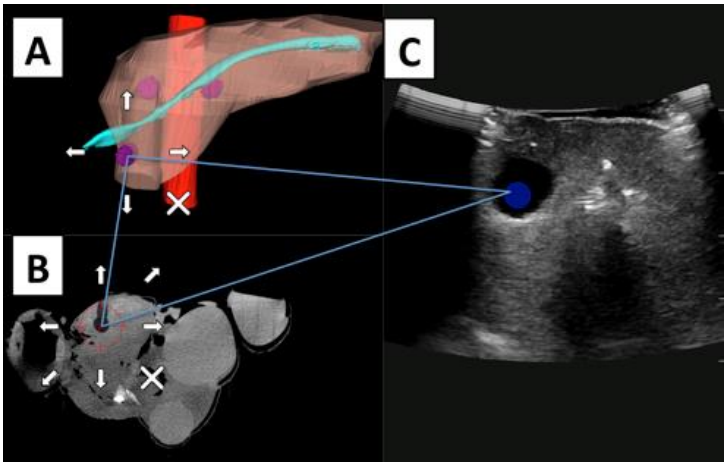



Figure 3: *Ultrasound based registration with an intraparenchymal landmark, in this case a pancreatic tumor.* ***A.*** *Depicts the segmented CT image, where the tumor in the processus uncinatus is selected. After identifying the same tumor on the ultrasound image (**C**), the corresponding structure is identified on the CT image (**B**).*

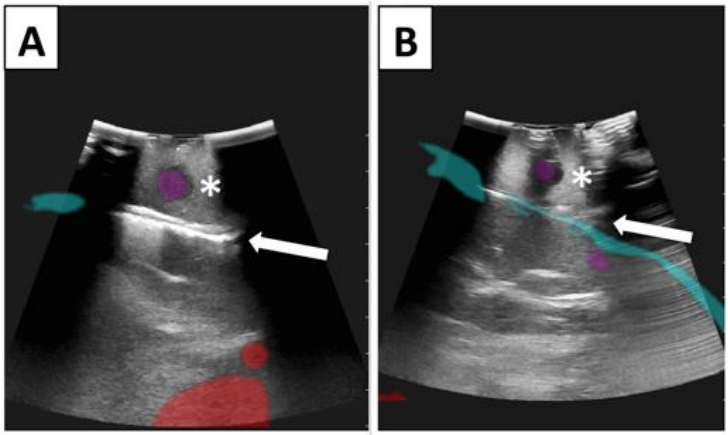



Figure 4: ***A&B.*** *After the registration the software overlays the segmented structures such as the tumor in purple (asterisk) and the pancreatic duct in light blue (arrow).*

## 4 Discussion

This proof of principle study showed that the novel minimally invasive endopancreatic surgery can be combined with computer assistance to display structures visible on the CT scan but not on the ultrasound or endoscopic image. With the help of computer assistance the surgeon is able to navigate inside the pancreas and perform a safe resection from inside the organ.

The registration using internal landmarks is essential for appropriate image to patient registration. To provide reliable navigation, the landmarks for registration purposes need to be in the proximity of the targeted lesions. Therefore, the landmark acquisition with ultrasound is recommended over rigid pointing tools. A shortcoming of the optical tracking approach is the inability to track the flexible laparoscopic ultrasound. Therefore, we aim to integrate electromagnetic tracking into this workflow as it does not suffer from the line of sight problem and allows to track flexible instruments.

## 5 Summary

The combination of computer assistance and the novel minimally invasive endopancreatic surgery allowed a safe pancreatic tumour resection from inside the organ. Computer assistance is of paramount importance for the orientation and the safety of endoscopic pancreatic surgery. If endopancreatic surgery can be performed safe and precise, this novel technique has the potential to replace open pancreatic surgical treatments that are known for major morbidity and mortality.

## References

[1] M. Gagner, A. Pomp, and M. F. Herrera, “Early experience with laparoscopic resections of islet cell tumors.,” *Surgery*, vol. 120, no. 6, pp. 1051–4, Dec. 1996.

[2] M. Kornaropoulos *et al.*, “Total robotic pancreaticoduodenectomy: a systematic review of the literature,” *Surg. Endosc.*, vol. 31, no. 11, pp. 4382–4392, Nov. 2017.

[3] D. C. Steinemann *et al.*, “Meta-analysis of hybrid natural-orifice transluminal endoscopic surgery *versus* laparoscopic surgery,” *Br. J. Surg.*, vol. 104, no. 8, pp. 977–989, Jul. 2017.

[4] P. C. Müller *et al.*, “Transduodenal-transpapillary endopancreatic surgery with a rigid resectoscope: experiments on ex vivo, in vivo animal models and human cadavers,” *Surg. Endosc.*, Mar. 2017.

[5] P. C. Müller, D. C. Steinemann, P. Sauer, K. Z’graggen, G. R. Linke, and B. P. Müller-Stich, “Balloon

Dilatation of the Minor Duodenal Papilla Up to 4 mm is Safe in a Porcine Model," *Surg. Laparosc. Endosc. Percutan. Tech.*, May 2017.

[6] P. C. Müller *et al.*, "Transpapillary endopancreatic surgery: decompression of duct system and comparison of greenlight laser with monopolar electrosurgical device in ex vivo and in vivo animal models," *Surg. Endosc.*, vol. 32, no. 7, pp. 3393–3400, Jul. 2018.

[7] M. K. Chmarra, R. Hansen, R. Mårvik, and T. Langø, "Multimodal phantom of liver tissue.," *PLoS One*, vol. 8, no. 5, p. e64180, 2013.

# Analysis of insertion angles of lateral wall cochlear implant electrode arrays in computed tomography images

P. Aebischer[1,2], S. Meyer[1,2], M. Caversaccio[1,2], W. Wimmer[1,2]

[1]*ARTORG Center for Biomedical Engineering Research, University of Bern, Switzerland*
[2]*Department of ENT, Head and Neck Surgery, Inselspital, University of Bern, Switzerland*

Contact: philipp.aebischer@artorg.unibe.ch

***Abstract***

*Insertion angles of lateral wall cochlear implant (CI) electrode arrays into the cochlea were computed from pre- and post-implantation computed-tomography recordings. Due to the nature of straight arrays, these angles give a lower bound on the anatomically achievable insertion angles in CI implantation.*

**Keywords**: Cochlear Implant, scala tympani, hearing preservation, insertion angle

## 1 Problem

The cochlear implant is a neural prosthesis that works by electrically stimulating the auditory nerve cells, thereby bypassing the acoustic part of the human ear. It consists of an electrode array inserted into the cochlea, an implanted receiving coil and an external transmission coil connected to an audio processor worn behind the ear, converting input from a microphone to signals sent to the electrodes.

The electrode array is preferably inserted through the round window into the scala tympani. Lateral wall electrodes are designed to slide along the outer wall of the scala tympani and are ideally inserted parallel to the cochlear centerline [1]. However, this is usually not possible due to the location of the facial nerve and chorda tympani [2]. The insertion of the array may cause trauma to the modiolus, basilar membrane and osseous spiral lamina. Minimizing intracochlear trauma is especially important for the preservation of residual hearing but maintaining intracochlear structures can reduce formation of fibrosis tissues and neo-ossification which otherwise can lead to increased impedance [3] and is therefore also desirable in regular implantation [4].

## 2 Material and Methods

Coregistered pre- and post-implantation computed-tomography recordings (slice thickness 0.2 mm) were available for 39 patients implanted with a MED-EL Flex$^{28}$ electrode array. The cochlea was manually segmented in the preop images and the electrode in postop images. The centerline of the basal turn of the cochlea and the electrode array was computed using a 3-D medial axis thinning algorithm. A local coordinate system was defined as proposed by Verbist et al. [5] by manual placement of markers at the round window, center of the modiolus and helicotrema.

## 3 Results

The angle of the electrode array at the entry point with respect to the ideal insertion was 16.5±8.8°, with the corresponding component in the basal plane $\varepsilon$=11.8±9.2° and the out-of-plane component $\partial$=10.9±7.2°. The first point of contact between the electrode tip and the intracochlear wall (assuming an insertion along the tangent of the final orientation of the inserted electrode) was at an angular insertion depth of 25.7±23.0°.

## 4 Discussion

The obtained values give a lower bound for anatomically achievable insertion angles, as the flexible straight arrays naturally try to minimize curvature at the insertion site. Insertions with an angle of approach above 15° were

classified as inaccurate and showed significantly increased intracochlear trauma by Torres et al. [3]. The entry angle of half of the arrays analyzed here (51%) deviate more than 15° from the scala tympani. This leads to an early and steep contact of the array with the intracochlear structures, which may negatively affect surgical outcome.

## 5 Conclusion

Pre- and post-implantation computed-tomography imaging of cochlear implants allows to estimate anatomically achievable insertion angles of the electrode array. Insertion angles below 15° correlated with less traumatic insertions [3] were not attainable for half of the examined patients.

## References

[1] X. Meshik, T. A. Holden, R. A. Chole, and T. E. Hullar, *Optimal Cochlear Implant Insertion Vectors*, Otol. Neurotol., vol. 31, no. 1, pp. 58–63, (2010)

[2] H. A. Breinbauer and M. Praetorius, *Variability of an Ideal Insertion Vector for Cochlear Implantation*, Otol. Neurotol., vol. 36, no. 4, pp. 610–617, (2015)

[3] R. Torres et al., *Cochlear Implant Insertion Axis Into the Basal Turn: A Critical Factor in Electrode Array Translocation*, Otol. Neurotol., vol. 39, no. 2, pp. 168–176, (2018)

[4] M. L. Carlson et al., *Implications of Minimizing Trauma During Conventional Cochlear Implantation*, Otol. Neurotol., vol. 32, no. 6, pp. 962–968, (2011)

[5] B. M. Verbist et al., *Consensus Panel on a Cochlear Coordinate System Applicable in Histologic, Physiologic, and Radiologic Studies of the Human Cochlea*, Otol. Neurotol., vol. 31, no. 5, pp. 722–730, (2010)

[6] R. Torres et al., *Variability of the mental representation of the cochlear anatomy during cochlear implantation*, Eur. Arch. Otorhinolaryngol., vol. 273, no. 8, pp. 2009–2018, (2016)

[7] W. Wimmer et al., *Semiautomatic Cochleostomy Target and Insertion Trajectory Planning for Minimally Invasive Cochlear Implantation*, BioMed Res. Int., vol. 2014, (2014)

# Convolutional Neural Networks im laparoskopischen Trainingssetting

*P.Beyersdorffer[1], K. Jansen[2], J. Miller[2], Peter Wilhelm[2], A. Kirschniak[2], J. Rolinger[2]*

*[1]Fakultät für Informatik, Hochschule Reutlingen, Deutschland*
*[2]Klinik für Allgemeine-, Viszeral- und Transplantationschirurgie, Universitätsklinikum Tübingen, Deutschland*

Kontakt: andreas.kirschniak@med.uni-tuebingen.de

***Abstract***

*In der Laparoskopie befinden sich der Operationssitus sowie die verwendeten Instrumente nicht im direkten Sichtfeld des Operationsteams. Unkontrollierte Bewegungen des Laparoskops oder der Instrumente können Verletzungen von umliegenden Gewebestrukturen verursachen. Es soll daher eine Anwendung konzipiert werden, welche diese risikobehafteten Situationen in Echtzeit erkennt und ein audiovisuelles Feedback an das Operationsteam generiert. Dies ermöglicht die Sensibilisierung für solche Situationen im Kontext einer laparoskopischen Trainingsumgebung und birgt das Potential die Wahrscheinlichkeit für deren Auftreten zu reduzieren. Die vorliegende Arbeit untersucht diesbezüglich die Machbarkeit einer binären Klassifizierung von laparoskopischen Bilddaten unter Verwendung eines Convolutional Neural Network als erster Schritt eines entsprechenden Entwicklungsprozesses. Die Akquise der benötigten Bilddaten von insgesamt sechs simulierten Cholezystektomien erfolgte im Rahmen eines laparoskopischen Trainingskurses. Auf der Basis einer effizienten Vorverarbeitung der generierten Trainingsdaten mittels „selective data augmentation" konnte ein entsprechendes künstliches neuronales Netzwerk konfiguriert werden. Die Verifizierung auf unbekannten Testdaten erbrachte in 88 % der Fälle eine korrekte Zuordnung. Durch eine Steigerung der Datenmenge erscheint eine Verbesserung der Fähigkeiten der neuronalen Netzstruktur und darauf aufbauend die Realisierung eines Echtzeit-Feedbacksystems möglich.*

**Keywords**: Convolutional Neural Networks, Visuelles Instrumententracking, Chirurgisches Training

## 1 Problemstellung

Die Arbeitsweise in einem Operationssaal wurde im Zuge der Etablierung minimalinvasiver Techniken konzeptionell verändert. Besonders die visuelle Informationsübertragung mittels laparoskopischem Kamerabild bringt eine Komplexitätssteigerung mit sich. Der Operationssitus sowie die verwendeten chirurgischen Instrumente befinden sich nicht mehr im unmittelbaren Sichtfeld des Operationsteams. Dennoch müssen die laparoskopischen Instrumente sicher über den dargestellten digitalen Bildausschnitt im dreidimensionalen Raum des Operationssitus positioniert werden. Ungeeignete Kameraeinstellungen sowie unkontrollierte Bewegungen der Instrumente können zu unbeabsichtigten Verletzungen benachbarter Strukturen führen [1,2]. Diese bleiben während des eigentlichen Eingriffs oft unbemerkt und demaskieren sich erst im späteren klinischen Verlauf [3].
Seit Etablierung der minimalinvasiven Chirurgie wurden vielfach Versuche unternommen, die Patientensicherheit zu erhöhen und unerwünschte Operationsergebnisse zu vermeiden [4,5]. Dabei spielt das chirurgische Training an geeigneten Modellen – wie zum Beispiel dem Boxtrainer – eine essenzielle Rolle [6]. In Bezug auf Langzeiterfahrung, weltweite Verbreitung und hohe Fallzahlen kann die laparoskopische Cholezystektomie als viszeralchirurgische Indexoperation bezeichnet werden [7]. Der Eingriff hat in diesem Zusammenhang einen erheblichen Stellenwert im Rahmen der chirurgischen Ausbildung sowie des laparoskopischen Trainings [8,9].
Trotz zunehmender Wahrnehmung des Sicherheitsaspekts durch die chirurgischen Fachgesellschaften und der Weiterentwicklung der verwendeten technologischen Verfahren, hat das Thema Patientensicherheit in der Laparoskopie nach wie vor eine hohe und aktuelle Relevanz [10].
In der vorliegenden Arbeit soll die Machbarkeit einer binären Klassifizierung von laparoskopischen Bilddaten unter Verwendung eines Convolutional Neural Network (CNN) evaluiert werden. Dies stellt den ersten Schritt in der graduellen Entwicklung eines automatisierten, kontextsensitiven Systems dar, welches ein audiovisuelles Echtzeit-Feedback zur Unterstützung des Operationsteams zur Verfügung stellt. Hierdurch soll in einer Trainingsumgebung die Aufmerksamkeit für solche Situationen sensibilisiert und damit die Wahrscheinlichkeit für deren Auftreten reduziert werden.

## 2 Material und Methoden

Die realitätsnahe Simulation der laparoskopischen Cholezystektomie am Boxtrainer bietet die Möglichkeit für diesen Eingriff ein spezifisches Convolutional Neural Network zu (CNN) konfigurieren. Im Anbetracht des einführend geschilderten Sicherheitsgedankens resultiert folgende Klassifizierungsaufgabe: Es soll mittels CNN entschieden werden, ob die Position des aktuell verwendeten laparoskopischen Arbeitsinstruments in Relation zum präsentierten Kamerabild ungesehene Gewebeverletzungen zulässt. Dies ist dann der Fall, wenn sich der schneidende bzw. hochfrequenzbetriebene Effektor im Situs befindet, aber nicht durch das Laparoskop erfasst wird. Das entsprechende Instrument wird also zu diesem Zeitpunkt nicht durch das Operationsteam visualisiert und stellt somit eine potentielle Gefahr für iatrogene Verletzungen dar.
Ein effizienter Workflow für die Vorverarbeitung der zugrundeliegenden Trainingsdaten ermöglicht die Abbildung dieser Klassifizierungsaufgabe durch die neuronale Netzstruktur. Nach der Aufzeichnung von sechs simulierten Cholezystektomien am Boxtrainer, standen 124.000 unsortierte Bilddaten zur Verfügung (Abbildung 1). Initial wurde eine Teilmenge von 25.000 dieser Bilder von zwei unabhängigen Experten gelabelt. Die Sichtbarkeit eines schneidenden bzw. hochfrequenzbetriebenen Instruments bestimmt dabei die Zuordnung zu einer der beiden Klassen. Diese Vorsortierung bildet die Grundlage, um einer ersten CNN-Instanz die Klassifizierung der restlichen Datenmenge mittels „supervised learning“ zu ermöglichen [11]. Für die Kontrolle der resultierenden Aufteilung konnte wie in [12] vorgestellt, die Wahrscheinlichkeit hinzugezogen werden, mit welcher das Bild von dem CNN der jeweiligen Klasse zugeordnet wurde. Diese Vorstufe des finalen CNN ordnete 72,86 % der unsortierten Daten ihre korrekte Klasse zu. Durch Anwendung von „oversampling“ [13] in Kombination mit „selective data augmentation“ konnte eine Gleichverteilung auf die Klassen des Datensatzes mit insgesamt 190.000 laparoskopischen Bildern erstellt werden. Der balancierte Datensatz ermöglicht den Transfer einer auf dem ImageNet Datensatz vortrainierten ResNet50 Architektur auf die zu erlernende Klassifizierungsaufgabe [14, 15].

## 3 Ergebnisse

Die Fähigkeit des CNN eine potenziell gefährliche Situation zu erkennen wurde über 50 Epochen trainiert und auf separierten Testdaten (10% der Gesamtdaten = 18.960 Bilder, zufällig selektiert) bewertet. In 88 % der Fälle konnte eine korrekte Zuordnung getroffen werden (Abbildung 2). Die übergeordnete Bewertung eines möglichen Fehlverhaltens des künstlichen neuronalen Netzes über den F1 Score, beläuft sich über die Kombination von Precision (0,9261) und Recall (0,8432) auf 0,8826.

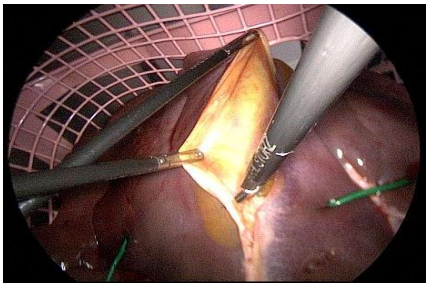


*Abbildung 1- Laparoskopische Sicht während der simulierten Cholezystektomie (Beispielaufnahme mit einem schneidenden Instrument)*

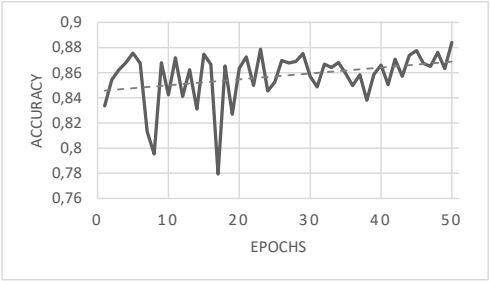



*Abbildung 2 – Erreichte Accuracy mittels Adam Optimizer, Kreuz-Entropie als Fehlerfunktion, Lernrate von $10^{-4}$ mit einer Reduzierung von $10^{-4}$/50 pro Epoche*

## 4 Diskussion und Zusammenfassung

Das entwickelte CNN bearbeitet die vorliegende Klassifizierungsaufgabe anhand von Bilddaten eines komplexen und realitätsnahen Laparoskopiesettings am Boxtrainer. Dabei wurden keine Manipulationen an den laparoskopischen Instrumenten oder dem Boxtrainer vorgenommen, um die Erkennungsrate zu erhöhen. Bisher publizierte Herangehensweisen für die automatisierte Analyse von Instrumentenbewegungen basieren in der Regel

auf einer Modifikation der Trainingsumgebung. Dies kann unter anderem durch magnetische Sensorik [16,17], farbliche Markierungen [18], Verwendung kinematischer Daten [19] oder die Reduzierung der Trainingsumgebung auf chirurgische Basisfertigkeiten [16-20] erreicht werden. Vergleichbare, intelligente Systeme für die klinische Anwendung verfolgen jedoch andere Ziele, wie beispielsweise die intraoperativen Prozess- und Workflowoptimierung durch Detektion der jeweiligen OP-Phase [21,22].
Die Umsetzung dieses Projekts belegt die Machbarkeit der binären Klassifizierung laparoskopischer Bilddaten mittels CNN. Das neuronale Netzwerk bildet die Grundlage für die Entwicklung einer Anwendung, welche das laparoskopische Videosignal in Echtzeit analysiert, um sicherheitskritische Positionierungen der Instrumente zu erkennen und damit perspektivisch eine intraoperative Risikobewertung der chirurgischen Maßnahme zu ermöglichen. Durch ein audiovisuelles Feedback soll die Aufmerksamkeit des Operationsteams für solche Situationen sensibilisiert und damit die Wahrscheinlichkeit für deren Auftreten reduziert werden. Ausblickend kann formuliert werden, dass durch eine Steigerung der Datenmenge eine weitere Verbesserung der Klassifizierungsleistung möglich erscheint. Die Erhöhung des Umfangs der Trainingsdaten von 25.000 auf 190.000 mit der korrespondierenden Verbesserung der berechneten Accuracy von 0.7286 zu 0.8822 unterstützt diese These. Zur Erweiterung der Heterogenität der Trainingsdaten könnte die risikosensible, binäre Klassifizierung auf dem öffentlich zugänglichen Cholec80 Datensatz [21] angewendet werden.
Neben dem beschriebenen audiovisuellen Feedback ergeben sich weitere Anwendungsmöglichkeiten im Rahmen der chirurgischen Trainingsumgebung. Denkbar ist zum Beispiel die Anwendung eines solchen Systems in Prozessen zur automatisierten Evaluation von chirurgischen Fertigkeiten oder in die gezielte Steuerung von hochfrequenzbetriebenen Effektoren zur Vermeidung akzidenteller, iatrogener Verletzungen von Nachbarstrukturen.

## Referenzen


[1] J. T. Bishoff, M. E. Allaf, W. Kirkels, R. G. Moore, L. R. Kavoussi, F. Schroder, *Laparoscopic bowel injury: incidence and clinical presentation*, The Journal of Urology, 161(3) 887-890 (1999)

[2] N. O. Machado, *Duodenal injury post laparoscopic cholecystectomy: Incidence, mechanism, management and outcome*, World Journal of Gastrointestinal Surgery, 8(4) 335-344 (2016)

[3] S. Cassaro, *Delayed manifestations of laparoscopic bowel injury*, The American Surgeon 81(5) 478-482 (2015)

[4] S. M. Strasberg, M. Hertl, N. J. Soper, *An analysis of the problem of biliary injury during laparoscopic cholecystectomy*, Journal of the American College of Surgeons, 180(1) 101-125 (1995)

[5] U. C. Niwa, S. Axt, C. Falch, S. Muller, J. A. Kreuzer, P. Nedela, A. Kirschniak, *Laparoscopic cholecystectomy as standardized teaching operation to treat symptomatic cholecystolithiasis*, Zentralblatt für Chirurgie, 138(2) 141-142 (2013)

[6] M. Waseda, N. Inaki, L. Mailaender, G. F. Buess, *An innovative trainer for surgical procedures using animal organs*, Minimally Invasive Therapy and Allied Technologies 14(4-5) 262-266 (2005)

[7] R. Aggarwal, O. T. Mytton, M. Derbrew, D. Hananel, M. Heydenburg, B. Issenberg, C. MacAulay, M. E. Mancini, T. Morimoto, N. Soper, A. Ziv, R. Reznick, *Training and simulation for patient safety*, Quality and Safety in Health Care 19 Suppl 2 i34-43 (2010)

[8] R. Aggarwal, T. Grantcharov, K. Moorthy, T. Milland, P. Papasavas, A. Dosis, F. Bello, A. Darzi, *An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room*, Annals of Surgery, 245(6) 992-999 (2007)

[9] K. E. Roberts, R. L. Bell, A. J. Duffy, *Evolution of surgical skills training*, World Journal of Gastroenterology 12(20) 3219-3224 (2006)

[10] P. H. Pucher, L. M. Brunt, R. D. Fanelli, H. J. Asbun, R. Aggarwal, *SAGES expert Delphi consensus: critical factors for safe surgical practice in laparoscopic cholecystectomy*, Surgical Endoscopy, 29(11) 3074-3085 (2015)

[11] F.Y. Osisanwo, J. E. T. Akinsola, *Supervised Machine Learning Algorithms: Classification and Comparison*, International Journal of Computer Trends and Technology, 48(3) 128-138 (2017)

[12] C. Zhang, W. Tavanapong, *Real Data Augmentation for Medical Image Classification*, Lecture Notes in Computer Science, 10552 67-76 (2017)

[13] M. Buda, A. Maki, *A systematic study of the class imbalance problem in convolutional neural networks*, Neural Networks 106 249-259 (2018)

[14] A. Krizhevsky, I. Sutskever, *ImageNet Classification with Deep Convolutional Neural Networks*, 25th International Conference on Neural Information Processing Systems, Nevada (2012)

[15] K. He, X. Zhang, *Deep Residual Learning for Image Recognition*, 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas (2016)

[16] M. Uemura, M. Tomikawa, F*easibility of an AI-Based Measure of the Hand Motions of Expert and Novice Surgeons*, Computational and Mathematical Methods in Medicine, 1-6 (2018)

[17] Y. A. Oquendo, E. W. Riddle, *Automatically rating trainee skill at a pediatric laparoscopic suturing task*, Surg Endosc, 32 1840-1857 (2018)

[18] G. A. Alonso-Silverio, F. Pérez-Escamirosa, *Development of a Laparoscopic Box Trainer Based on Open Source Hardware and Artificial Intelligence for Objective Assessment of Surgical Psychomotor Skills*, Surgical Education: Training for the Future, 25(4) 380-388 (2018)

[19] Z. Wang, A. Majewicz Fey, *Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery*, International Journal of Computer Assisted Radiology and Surgery 13(12) 1959-1970 (2018)

[20] I. Funke, S. T. Mees, J. Weitz, S. Speidel, *Video-based surgical skill assessment using 3D convolutional neural networks*, International Journal of Computer Assisted Radiology and Surgery 14(7) 1217-1225 (2019)

[21] A. P. Twinanda, S. Shehata, *EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos*, IEEE Transactions on Medical Imaging, 36(1) 86-97 (2017)

[22] S. Bodenstedt, M. Wagner, *Unsupervised temporal context learning using convolutional neural networks vor laparoscopic workflow analysis*, arXiv: 1702.03684 (2017)

# Towards Automatic Visual Inspection in a Laparoscopy Box Trainer using an Instance Segmentation Deep Learning Architecture

*Harold Jay Bolingot[1], Tomohiro Shibata[1]*

*[1]Human and Social Intelligence Systems Laboratory, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan*

Contact: bolingot.harold681@mail.kyutech.jp

***Abstract***

*In general, data-driven evaluation of laparoscopic surgical skills currently focus on motion analysis of tools, wherein the movement of laparoscopic tools is recorded. Several studies have validated the quality of tool motion as a standard to represent the quality of a surgical action and therefore, the subject's surgical competency. However, skill assessment methods relying on tool motion usually ignore the surgical objects and how they are manipulated in training, which makes this skill assessment framework limited with respect to the globally accepted standards and criteria for evaluating fundamental laparoscopy skills such as OSATS and GOALS. This research aims to extend the application of computer vision in laparoscopy training by tracking tool motion, to include surgical object detection, tracking, and instance segmentation. Specifically, an instance segmentation deep learning algorithm called YOLACT is tested on a modified FLS peg transfer exercise in a box trainer. Experiments show a mAP score of 61.09 for bounding boxes and 39.71 on masks, on a limited training dataset, which is comparable with standard tool tracking detectors based on deep learning algorithms. In terms of speed performance, the YOLACT algorithm delivered impressive performance at almost 20 fps playback, which is near-real time at 52-ms latency. This research clearly demonstrates the potential of using instance segmentation algorithms for the purpose of understanding the outcomes achieved on the surgical field view, which can reinforce tool motion analysis methods in automatically evaluating the performance of surgeons during training.*

**Keywords**: laparoscopic surgery, YOLACT, deep learning, object tracking, instance segmentation

## 1 Problem

This research aims to reinforce the data-driven skill assessment framework for laparoscopic surgery training that is mostly limited to tool motion analysis [1, 2]. By proposing a system for assessing the state of surgical objects in a box trainer using a computer vision architecture based on instance segmentation algorithms, the process of skill assessment can be improved with the introduction of a system component that analyzes the state of the surgical objects being manipulated by the tools. the unique attributes of laparoscopy images, which include the presence of specular reflections, complicated geometry, color variability, occlusion, and existence of various endoscopic artefacts such as motion blur and smoke. Instance segmentation algorithms can focus on regions-of-interest in laparoscopy images which include specific tissues and anatomical structures. In the context of box trainers, instance segmentation algorithms can be used to focus specifically on the object states, such as deformability, or if they retain the same shape and structure throughout the training. The deformability of surgical objects can imply whether too much force is exerted on the object, which can be used to judge the ability of handling the tissues with care and utmost dexterity, an aspect that cannot rely on tool motion tracking alone. Instance segmentation algorithms are capable of providing not just masks to specify regions-of-interests, but also bounding boxes and class confidence score, which are standard in ordinary object detection algorithms.

## 2 Material and Methods

Images of the surgical view through the RGB endoscopic camera were utilized for training a deep learning object detector and tracker. A popular deep learning algorithm known as YOLACT [3] was used. YOLACT [3] (You Only Look At the CoefficienTs) is a more optimized version for instance segmentation, which has secured a good reputation for its speed and accuracy trade-offs. YOLACT uses ResNet-101 with FPN (Feature Pyramid Networks)

that helps in creating pyramids of feature maps of high-resolution images rather than the Conventional Pyramid of Images approach, therefore reducing the time and requirements of computational capabilities. YOLACT [3] was tested in this experiment to track two kinds of objects within the box trainer's surgical view: pegs and graspers. Pegs are the objects being picked up and carried around by the laparoscopic graspers. The grasper's end-effectors are being detected and tracked. YOLACT was pre-trained on COCO during development. The weights generated from the COCO dataset were then incorporated in a model that was trained on our dataset of labels of the pegs and end-effectors.

The setup of the box trainer and the camera view are depicted on Fig. 1. The RGB camera used is the Intel RealSense Depth Camera D415. It is chosen due to its appropriate size, weight, and image resolution (1920x1080) for the box trainer. The laparoscopic graspers used are training prototypes, and the pegs and gridded board were all 3D printed in the laboratory. Fig. 2 shows examples of ground truth mask labels on the training dataset. Table 2 describes the dataset generated using the setup in Fig. 1.

Figure 1*: Left: setup of the box trainer. Right: camera view.*

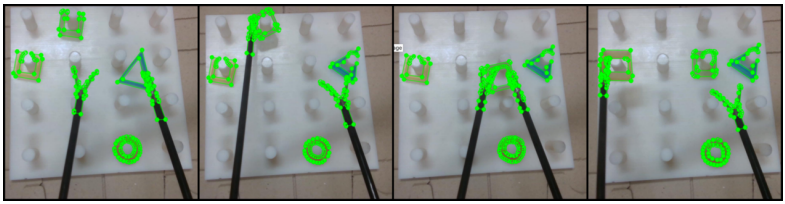


Figure 2*: Examples of the ground truth labels on the training dataset.*

Table 1*:* Dataset characteristics.

| **Images (1920 by 1080 pixels)** | | **Labels: polygon masks** | |
|---|---|---|---|
| Training | Validation | Objects | Tools |
| 155 images | 47 images | red_peg | left_ef |
| | | yellow_peg | |
| | | green_peg | right_ef |
| | | blue_peg | |

## 3 Results

Fig. 3 shows the training loss attained from training the YOLACT algorithm on the dataset described in Table 1. The losses plotted are class confidence, mask loss, bounding box loss, and segmentation loss. Fig. 4 shows the graph of mean average precision for all the classes' bounding box and masks through iterations up to 30k iterations. By the 80,000$^{th}$ iteration, the mask loss, bounding box loss, and segmentation have remained consistent while the confidence loss has experienced a slight uptick. However, the mAP scores have slightly experienced overfitting within the 80k to 100k iterations. Fig. 5 shows examples of the resulting masks of some test images. The class

confidence scores have been consistently 1.00 on all cases of test data: offline video, images, and webcam real-time view. The masks are accurate but not perfect. There are some deformities in the mask; the entirety of the object shapes are not covered fully by the mask, but the masks properly encapsulate the shapes outside the poles. In addition, slight variations in the camera angle do not impact the quality of the masks, as seen on the examples in Fig. 5. Fig. 6 shows average precision for all classes at certain IoU thresholds.

Table 2*:* Accuracy, speed performance, and latency of the algorithm.

| | | |
|---|---|---|
| **Final mAP** | Bounding box | **61.09** |
| | Mask | **39.71** |
| **Speed performance** | Processing | **18 fps** |
| | Video playback | **19 fps** |
| **Latency** | **~52 ms** | |

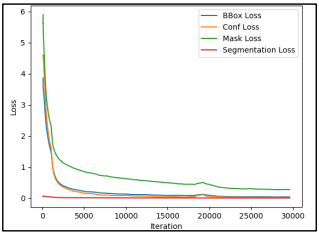



Figure 3*: Training loss curve (up to 30k iterations).*

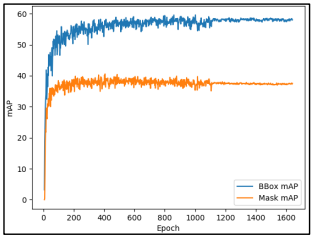



Figure 4*: Mean average precision for all classes combined (up to 30k iterations).*

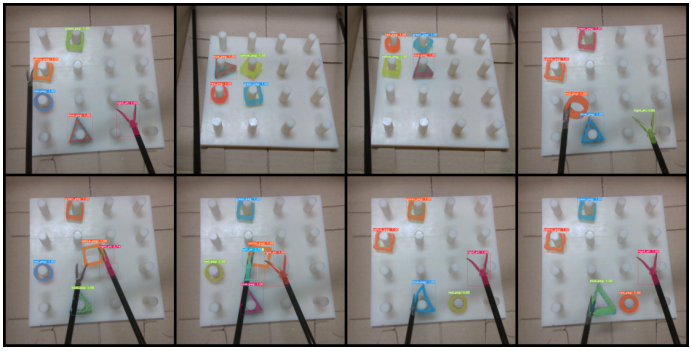


Figure 5*: Examples of masks output.*

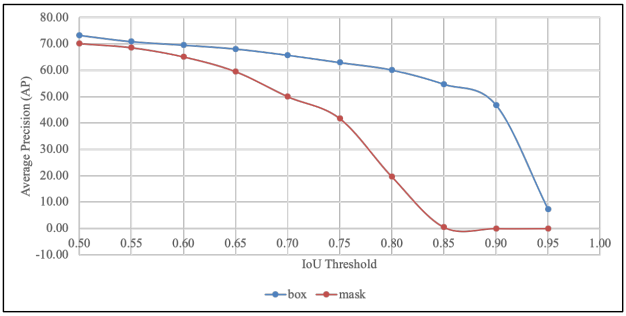



Figure 6*: Average precision for all classes at certain IoU thresholds.*

# 4 Discussion

Table 3 shows the two criteria specified under two standard assessment guidelines used in laparoscopy surgery training: the Objective Structured Assessment of Technical Skills (OSATS) [7] and the Global Operative Assessment of Laparoscopic Skills (GOALS) [5], and whether tool motion analysis and automatic visual inspection are effective and useful in each criterion. As it is shown, tool motion analysis mostly considers the psychomotor skills, while automatic visual inspection mostly considers the effect of tool manipulation on the surgical targets, such as tissues, which is equally important in skills analysis but always mostly ignored if methods only rely on tool motion analysis.

Table 4 presents the comparison of the developed prototype with other works on using object detection and classification for improving skill assessment and training in laparoscopy (bold indicates best performance). Jin, et al. (2018) used Faster R-CNN, a region-based convolutional neural network, on a dataset of 923 instances of 7 classes of end-effectors of 7 tools, while Kyungmin, et al. (2018) object detection system YOLO9000 on a dataset of 8,108 images on the same dataset. As can be inferred from a table, the direct relationship between the size of training dataset and mAP score is not clearly established, although the availability of over 8,000 images on the algorithm the second work must have contributed to the extremely high mAP and processing speed. YOLO9000 is a recently upgraded algorithm of its predecessor YOLO. Despite our experiment having the smallest dataset, it is still relatively comparable with the first work that leveraged on a larger dataset compared to ours. The biggest difference between our work and on the two works is the focus away from tool motion analysis to automatic visual inspection for integration in skill analysis and training, with training conducted on surgical object representations as pegs on the peg transfer exercise. This has not been taken into consideration in the two works featured as well as largely in the research literature on automated skill analysis. However, it is understood that several works have attempted various computer vision techniques on surgical tissue segmentation on endoscopy and laparoscopy images; however, those implementations are targeted towards integration in image-guided surgery, and not for use in automatic skill analysis and training.

Insofar as instance segmentation is concerned, YOLACT shows potential in being an effective algorithm for use in automatic visual inspection of surgical objects that can be used to improve automatic skill training and analysis assessment methods that are limited on tool motion analysis. The goal of automatic visual inspection is not to replace tool motion analysis, but to work alongside it to complete the full picture of laparoscopy skill as envisioned in the OSATS and GOALS guidelines.

Table 3: Use of Tool Motion Analysis and Automatic Visual Inspection (potential use) in OSATS and GOALS Assessment Methods.

| Criteria | | **Tool Motion Analysis** | **Automatic Visual Inspection** |
|---|---|---|---|
| **OSATS [7]** | Respect for Tissue | × | ✓ |
| | Time and Motion | ✓ | ✓ |
| | Instrument Handling | ✓ | × |
| | Knowledge of Instrument | ✓ | ✓ |
| | Use of Assistants | (−) | (−) |
| | Flow of Operation and Forward Planning | ✓ | ✓ |
| | Knowledge of Surgical Procedure | × | ✓ |
| **GOALS [5]** | Depth Perception | ✓ | × |
| | Bimanual Dexterity | ✓ | × |
| | Efficiency | ✓ | ✓ |
| | Tissue Handling | × | ✓ |

Table 4: Comparative analysis of our test of YOLACT versus other works.

| | Jin, et al. (2018) [4] | Kyungmin, et al. (2019) [6] | Ours |
|---|---|---|---|
| Objects | **7 classes of end-effectors** | **7 classes of end-effectors** | Grasper, pegs (2 categories, 6 classes) |
| Algorithm | Faster-RCNN | YOLO9000 | YOLACT |
| Training dataset size | 923 instances | 8,108 images | 404 instances |
| Surgical environment | Real | Real | **Box trainer** |
| Bounding box mAP | 63.1 | **84.7** | 61.09 |
| Mask mAP | N/A | N/A | **39.71** |
| Processing speed | N/A | **38 fps** | 18 fps |
| Video playback speed | 5 fps | None | **19 fps** |
| Latency | 200 ms | N/A | **~52 ms** |
| Real-timeness | No (they claim yes) | Not tested | **Near real-time** |
| Skill Analysis and Training | Tool motion analysis, tested with subjective correlation with peer review | To be used for tool motion analysis (but not explicitly tested) | **To be used for automatic visual inspection of surgical objects** |

## 5 Conclusion

The skill assessment and training process for laparoscopy on a box trainer is intended to be improved through the development and integration of instance segmentation technique on laparoscopy images based on a popular deep learning architecture. The proposed system aims to contribute a layer of visual inspector that verifies the state of the surgical objects used in training, specifically to focus on the deformability of surgical objects and detection of various shapes and geometry in the surgical field view. The implementation of YOLACT achieves relatively

successful performance on a limited dataset, achieving about 61.09 mAP on bounding boxes and 39.71 mAP on masks, while being near-realtime at 20 fps video playback. It is tested on the peg transfer exercise which allows the implementation to extend beyond tools and include surgical objects used in training. The implementation and this demonstration of YOLACT instance segmentation paves the way for further progress and development of automatic visual inspection that can work alongside tool motion analysis to complete the fuller picture of skill assessment based on global standards such as OSATS and GOALS.

## References

[1] Chmarra, M. K., Klein, S., de Winter, J. C. F., Jansen, F.-W., & Dankelman, J. (2009). Objective classification of residents based on their psychomotor laparoscopic skills. Surgical Endoscopy, 24(5), 1031–1039. doi:10.1007/s00464-009-0721-y.

[2] Fard, MJ, Ameri, S, Darin Ellis, R, Chinnam, RB, Pandya, AK, Klein, MD. Automated robot-assisted surgical skill evaluation: Predictive analytics approach. Int J Med Robotics Comput Assist Surg. 2018; 14:e1850. https://doi.org/10.1002/rcs.1850.

[3] Bolya, Daniel, et al. “YOLACT: Real-Time Instance Segmentation.” ArXiv:1904.02689 [Cs], Apr. 2019. arXiv.org, http://arxiv.org/abs/1904.02689.

[4] Jin, Amy, et al. “Tool Detection and Operative Skill Assessment in Surgical Videos Using Region-Based Convolutional Neural Networks.” ArXiv:1802.08774 [Cs], Feb. 2018. arXiv.org, http://arxiv.org/abs/1802.08774.

[5] “GOALS.” C-SATS, http://www.csats.com/goals. Accessed 27 May 2019.

[6] Jo, Kyungmin, et al. “Robust Real-Time Detection of Laparoscopic Instruments in Robot Surgery Using Convolutional Neural Networks with Motion Vector Prediction.” Applied Sciences, vol. 9, no. 14, July 2019, p. 2865. DOI.org (Crossref), doi:10.3390/app9142865.

[7] Martin, J. A., et al. “Objective Structured Assessment of Technical Skill (OSATS) for Surgical Residents.” British Journal of Surgery, vol. 84, no. 2, Feb. 1997, pp. 273–78. DOI.org (Crossref), doi:10.1046/j.1365-2168.1997.02502.x.

# Maps, Colors, and SUVs for Standardized Clinical Reports

*Nico Merten*[1,2], *Philipp Genseke*[3], *Bernhard Preim*[1,2], *Michael C. Kreissl*[1,3], *Sylvia Saalfeld*[1,2]

[1] Research Campus *STIMULATE*
[2] Department of Simulation and Graphics, Otto-von-Guericke University
[3] Department of Radiology and Nuclear Medicine, University Hospital Magdeburg
Magdeburg, Germany

Contact: nmerten@isg.cs.uni-magdeburg.de

***Abstract***

*We present a map-based visualization approach of anatomical landmarks and lymph node stations in the thorax. Moreover, a method is presented to create color-codings for such maps.These color-codings are combinations of two categorical color scales for said map elements, and an alternating, sequential color scale for multiple Region of Interest (ROI) segmentations. In clinical practice, nuclear medicine physicians use ROI segmentations to compute Standardized Uptake Values (SUVs), which are important clinical parameters to quantify and compare suspicious findings in PET images, e. g., to assess treatment responses. Therefore, we additionally present a guide to compute reproducible SUVs from PET images and DICOM tag values, since the majority of related work offers theoretical information only.*

**Keywords**: Map-Based Visualizations, Color-Coding, SUV Computation, Clinical Report Generation

## 1 Problem

In recent years, the (semi-)automatic generation of clinical reports, which can be assessed on screen and paper, has gained increased importance [1, 13]. Two important items in such reports are patient-specific diagnoses and therapy decisions. These are usually discussed and documented during interdisciplinary consultations of clinical experts, such as radiologists, oncologists, surgeons, and nuclear medicine physicians. Such consultations are intrinsically visual, since the aforementioned decisions are derived from assessing medical image data, e. g., PET/CT images. Moreover, using facilities that create documents with a standardized structure and layout benefits the process of mental information retrieval.

However, such reports are usually documents that solely consist of text and tables. On the one hand, this results in documents that reflect cases very objectively, e. g., by listing accurate measurements or by using quasi-standardized phrasing to describe cancerous infiltration of tissue. On the other hand, visualizations can enhance such documents. First, text-only documents do neither reflect the visual nature of the aforementioned decision-making processes, nor are the capabilities of our visual apparatus adequately used, i. e., to conceive images in seconds [15]. Secondly, since reporting is a combined effort of various authors, for each part, the degrees of quality and detail can vary, which in turn may hinder reproducible interpretations of findings or decisions [11]. Finally, although clinical experts are well-trained to extract crucial pieces of information out of large texts, this process is prone to errors, because decision-altering details can remain occult to readers.

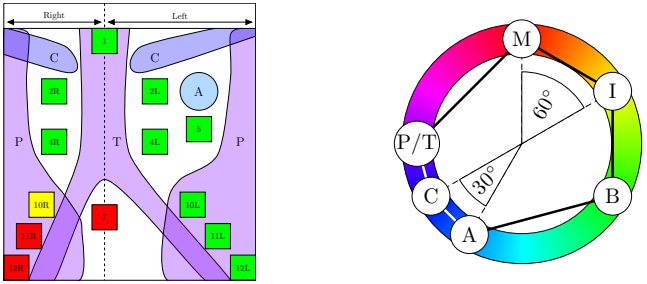



*Fig. 1: The proposed map visualizations of LNS staging diagnoses and anatomical landmarks (left) and their default color-coding (right). The color wheel represents the HSV color model. The colors are plotted in clockwise direction, with **M**align at red (0/360° for H), **I**nconclusive at yellow (60°), and **B**enign at green (120°). The asymmetrical map layout represents the unequal primary bronchi lengths between the **T**rachea and **P**arenchyma, and to avoid visual clutter due to the **A**orta and LNS 5. At the top, the **C**laviculae separate LNS 1 from 2R/L.*

A map-based visualization approach is presented to enhance reports by stylized depictions, which is inspired by other map-based visualizations in clinical workflows, e. g., Bull's Eye Plots to diagnose myocardial perfusion defects [6, 7]. Here, maps depict anatomical landmarks and *Lymph Node Stations* (LNSs) in the thorax. These landmarks are simplified frontal cross-sections of the aorta, trachea, clavicle bones and lung parenchyma, while the LNSs are anatomical regions that group multiple nearby lymph nodes [9]. The two main contributions of this work are two color-coding methods and a guide to compute SUVs from PET image values, e. g., from *Region Of Interest* (ROI) segmentations. Selecting *appropriate* color-codings is an important aspect for map and medical image visualizations [2]. Thus, methods are presented that generate color-codings for the aforementioned map elements and ROIs using the HSV color model. Moreover, a guide is presented to compute reproducible *Standardized Uptake Values* (SUVs). There exists related work that describes the computation of SUVs, but in most cases this process is described in theory only [3]. If more practical descriptions are provided, the lack of specific input and output values makes it challenging to reproduce results.

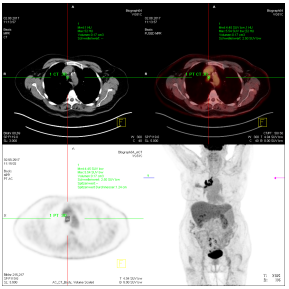
## 2 Material and Methods

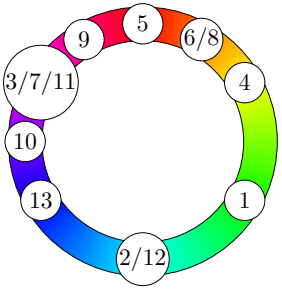
In this section, the color-coding methods and SUV computation guide are presented. Both contributions are extensions of a larger, semi-automatic processing pipeline for clinical reports (cf. Fig. 3), which is implemented in MeVisLab 2.8.2 [8, 12].

### 2.1 Color-Coding – Lymph Node Stations and Anatomical Landmarks

The proposed approach requires a color-coding for three LNS staging diagnoses, namely benign, malign and inconclusive, and four anatomical landmarks, namely the aorta, lung parenchyma, clavicle bones, and trachea. These landmarks provide spatial references for the LNS [9]. For simplification, the same hue is assigned to the trachea and parenchyma landmarks, since they belong to the same organ.

This results in six map elements that can be regarded as nominal data. Nominal data is usually color-coded via categorical color scales, which introduce large hue differences to generate color palettes [2]. In Figure 1, the default color-coding is depicted in the HSV color wheel, with the hue axis plotted in clockwise direction. For example, **M**align marks red (0/360°) and **B**enign marks green (120°). Since it is important to visually distinguish the individual LNS staging diagnoses, they are placed in 60°steps from each other. This guarantees that they are placed in different, major color zones. In the default case, the three diagnostic categories are placed on red, yellow and green, which results in the easy-to-understand *traffic light color-coding*. In contrast, landmark hues are placed in 30°steps from each other with the pivotal landmark hue (here: **C**laviculae) being defined as the complementary color of the pivotal LNS staging hue (here: **I**nconclusive). On the one hand, this makes resulting landmark colors more similar, but still easy to distinguish. On the other hand, for the default version, this results in landmark colors having blue-like tones, which draw minimal visual attention. This hue placement is an adaption of the harmonic *Y-Type* template by Itten [4], which creates two non-overlapping, complementary color groups.






| # | Name | H | S | V |
|---|---|---|---|---|
| 1 | Green | 120 | 1 | 1 |
| 2 | Cyan | 180 | 1 | 1 |
| 3 | Magenta | 300 | 1 | 1 |
| 4 | Yellow | 60 | 1 | 1 |
| 5 | Dark Red | 0 | 1 | 0.5 |
| 6 | Brown | 30 | 1 | 0.5 |
| 7 | Dark Magenta | 300 | 1 | 0.5 |
| 8 | Orange | 30 | 1 | 1 |
| 9 | Pink | 330 | 1 | 1 |
| 10 | Purple | 270 | 1 | 1 |
| 11 | Light Magenta | 300 | 0.5 | 1 |
| 12 | Teal | 180 | 0.5 | 0.5 |
| 13 | Gray | 240 | 0.2 | 0.75 |

*Fig. 2:* ***Left****: An exported screenshot of the VG51C toolbox with a ROI around LNS 10R.* ***Middle/Right****: The first thirteen ROI colors in the toolbox depicted in a HSV color wheel and listed in a table. For simplification purposes, only the hues are depicted. While many colors are defined in some sections, e. g., around magenta (300° for H), some areas are barely used, e. g., around green (120° for H). Given the irregular color placement, it appears that the used color palette is hard-coded.*

Table 1: *Solutions for Eq. 1 and resulting hues for $N = 9$ with respect to the default color hexagon from Figure 1.*

| | | Hue for $i$-th Color via $(ROI\boldsymbol{C}_i)_H$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $B_H$ | $R_H$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 120 | 240 | 144 | 216 | 288 | 168 | 240 | 312 | 192 | 264 | 336 |

Assuming that the saturation and value components of all colors equal 1.0, from a geometrical point of view, this results in an irregular, convex hexagon. Color changes can be geometrically interpreted as rotations (hue), scaling (saturation), or vertical translations (value) of the hexagon inside the HSV color model cylinder. When using the aforementioned report generation software, users can adapt the color-coding interactively by changing the hue, saturation, or value parameters of one color and, consequently, all colors are updated.

## 2.2 Color-Coding – Adding Region of Interest Colors

In addition to color-coding LNSs and landmarks, ROI segmentation masks can also be color-coded. Existing clinical software toolboxes, e. g., *Siemens Syngo VG51C* (Siemens Healthcare, Erlangen, Germany; VG51C), offer ROI segmentation, SUV computation, and image-based export functionalities that can be combined and used for reports. In clinical practice, ROI segmentation tools are used to mask suspicious image regions and after ROI definition, medical images are superimposed by color-coded outlines or areas. For example, in $^{18}$F-FDG PET scans, suspicious hotspots are encompassed and assessed, since they can point to cancerous tissue.

The first thirteen ROI colors of the VG51C toolbox are shown in Figure 2. While some hue sections are barely used, e. g., around green, some areas are cluttered, e. g., around orange and magenta. The latter results in a shortcoming, since the depicted color palette is used for ROIs that superimpose color-coded PET images. In the VG51C software, a black-red-orange-yellow-white color scale is used for PET images, which, in combination with the ROI color-coding, can result in ambiguous color information on screen and paper. Consequently, this can hinder reproducibility of ROI segmentations via reports. To overcome this, the VG51C software offers additional gray scale views of PET and CT slices that have a high contrast to ROI boundary colors (cf. Fig. 2).

In the following, a color-coding method for ROI segmentations is presented that overcomes the aforementioned shortcoming. For simplification purposes, the default color-coding from Figure 1 will be assumed, e. g., with the saturation and value of all ROI colors equal 1.0. However, when LNS and landmark colors are adapted, ROIs color-codings are affected similarly, i. e., moving them through the HSV model cylinder. This is a useful feature, since the aforementioned PET image color-coding is manufacturer-dependent, which makes the proposed methods feasible for various clinical software toolboxes. The proposed method is expressed by two equations, which require the number of ROIs ($N$), the *type* of $N$, $t(N)$, and the hue range $R_H$ in which ROIs are defined. $R_H$ is defined by $R_H := 360 - min(\overline{B_H M_H})$, which is the minimum distance between the **M**align and **B**enign hues (cf. Fig. 1). The first equation is given by

$$(ROI\boldsymbol{C}_i)_H = B_H + \left( \frac{R_H}{(N+1)} \times step_{t(N)}(\boldsymbol{C}_i) \right), \ \forall i \in [1, N] \tag{1}$$

which defines a hue for the $i$-th ROI for $N$-many colors. The second equation, namely $step(C_i)$, depends on $t(N)$ and assigns different hues to subsequently defined ROIs, e. g., between the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ ROI color. $N$ can be one of four types and, consequently, a different step equation is used: $t(N)$ can be even ($step_e(C_i)$), odd and not prime ($step_o(C_i)$), or prime with an even (e. g., 5, $step_{pe}(C_i)$) or odd (e. g., 7, $step_{po}(C_i)$) number of elements to the left and right of the pivotal element, e. g., 4 for $N = 7$. In the following, the four step equations are listed:

$$step_e(C_i) = \left( 1 + \sum_{j=1}^{i-1} \begin{cases} 4 & \text{, if } j = \dfrac{N}{2} \\ 2 & \text{, otherwise} \end{cases} \right) \%(N+1) \qquad step_o(C_i) = \left( 1 + \sum_{j=1}^{i-1} \begin{cases} maxPF(N)+2 & j\%\left(\dfrac{N}{maxPF(N)}\right) = 0 \\ maxPF(N) & \text{, otherwise} \end{cases} \right) \%(N+1)$$

$$step_{pe}(C_i) = \left( 1 + \sum_{j=1}^{i-1} \begin{cases} \dfrac{N+1}{2}+2 & \text{, if } isEven(j) \ \wedge \ j < \dfrac{N-1}{2} \\ \dfrac{N-1}{4}+1 & \text{, if } j = \dfrac{N-1}{2} \\ \dfrac{N-1}{4}+2 & \text{, if } j = \dfrac{N+1}{2} \\ \dfrac{N+1}{2}+2 & \text{, if } isOdd(j) \ \wedge \ j > \dfrac{N+1}{2} \\ \dfrac{N-1}{2}+2 & \text{, otherwise} \end{cases} \right) \%(N+1) \quad step_{po}(C_i) = \left( 1 + \sum_{j=1}^{i-1} \begin{cases} \dfrac{N+1}{2}+2 & \text{, if } isEven(j) \ \wedge \ j < \dfrac{N-1}{2} \\ \dfrac{3N-1}{4} & \text{, if } j = \dfrac{N-1}{2} \\ \dfrac{3(N+1)}{4} & \text{, if } j = \dfrac{N+1}{2} \\ \dfrac{N+1}{2}+2 & \text{, if } isOdd(j) \ \wedge \ j > \dfrac{N+1}{2} \\ \dfrac{N-1}{2}+2 & \text{, otherwise} \end{cases} \right) \%(N+1)$$

For each color $C_i$, the equations return a factor that donates the hue of $ROI\boldsymbol{C}_i$ relative to $B_H$. The function $maxPF(N)$ in $step_o(C_i)$ returns the maximum of the prime factorization of $N$, e. g., 5 for $N = 15$. In Table 1, results using $step_o(C_i)$ and $N = 9$ are listed, while the resulting color palette is depicted in Figure 3.

### 2.3 SUV Computation

SUVs are computed by combining PET image and DICOM tag values, i. e., patient data and information about the administered radiopharmaceutical and, subsequently, they are used for diagnoses and therapy decisions. Generally, they describe how much of an administered radiopharmaceutical is absorbed by the patient's anatomy. This provides an indication of how metabolically active anatomical regions are, while suspiciously high uptakes can point to cancerous or metastatic tissue. There exists related work that describes the computation of SUVs, but in most cases the computation process is only described from a theoretical point of view [3]. Moreover, if more descriptions are presented, no further information about specific input and output values is provided, which makes it challenging to reproduce results [14]. Therefore, the method in Algorithm 1 below describes how PET image values are converted into $\text{SUV}_{BW}$ values, which are SUVs normalized with the patient's bodyweight. In clinical toolboxes, SUVs are usually computed for the single *hottest* voxel value ($\text{SUV}_{max}$) or the average of all ROI-masked PET image values ($\text{SUV}_{\mu}$). However, Algorithm 1 can be applied to any PET image value. There exist various types of SUVs, but this work is focused on the computation of $\text{SUV}_{BW}$ values, since they are used in the VG51C software and allow for an evaluation regarding computation accuracy.

## 3 Results

The results are depicted in Figure 3. On the left, an exemplary report is presented, which is divided into four parts. At the top, information about the hospital and responsible physicians are compiled. In the second part, patient-specific data, an N-staging suggestion, and a therapy decision is listed. Furthermore, a color-coding legend is shown that can be used to interpret the 2D map on the right with LNS diagnoses (cf. Fig. 1). The suggested N-staging is derived from the LNS diagnoses and represents the N-component of the *Tumor, Node, and Metastasis* (TNM) staging system that describes suspicious findings in nearby lymph nodes [10]. In the third part, a frontal PET *Maximum Intensity Projection* (MIP) image with color-coded ROIs is shown and the table next to it lists additional information for each ROI, namely a user-defined name, its color, the corresponding PET slice, and its volume. Moreover, the table also lists PET quantification data for all ROIs, namely their respective $\text{Count}_{max}$, which is the highest PET image value, the resulting $\text{SUV}_{max;BW}$ value, which was obtained with Algorithm 1, and a $\text{SUV}_{\mu;BW}$ value, which is the average $\text{SUV}_{BW}$ inside each ROI. The color wheels on the right depict ROI color-codings for varying $N$ using the respective step equation. Since nine ROIs were defined for the depicted report, the upmost color wheel was used for color-coding. In the last part, acquisition- and SUV computation-related data is listed, which is intended to create a strong connection between the underlying image data, the ROI segmentations and SUV computations, and the diagnosis and therapy decisions derived thereof.

## 4 Discussion

When comparing the baseline color-coding from the VG51C toolbox with the proposed method (cf. Figs. 2 and 3), the proposed method does not define colors in the hue range of color-coded PET images, which minimizes the risk of misinterpretations. However, this only applies if the default traffic light color-coding is used. Moreover,

```
# Load PET Data Set
PETdataSet = dicomRead(fileOnHardDrive)
# Get PET Image Value
PETimageValue = PETdataSet.getVoxelValue(x, y, z)
# Get DICOM Tags (expected units are shown in square brackets, e.g. [s] for seconds)
radioStartTime = PETdataSet.getTagValue([0x0018, 0x1072]) # [hhmmss]
radioTotalDose = PETdataSet.getTagValue([0x0018, 0x1074]) # [Bq]
radioHalfLife = PETdataSet.getTagValue([0x0018, 0x1075]) # [s]
acquisitionDate = PETdataSet.getTagValue([0x0008, 0x0022]) # [yyyymmdd]
acquisitionTime = PETdataSet.getTagValue([0x0008, 0x0032]) # [hhmmss]
seriesDate = PETdataSet.getTagValue([0x0008, 0x0021]) # [yyyymmdd]
seriesTime = PETdataSet.getTagValue([0x0008, 0x0031]) # [hhmmss]
patientWeight = PETdataSet.getTagValue([0x0010, 0x1030]) # [Kg]
rescaleIntercept = PETdataSet.getTagValue([0x0028, 0x1052]) # [scalar]
rescaleSlope = PETdataSet.getTagValue([0x0028, 0x1053]) # [scalar]
# Compute SUV
administrationDateTime = seriesDate + "." + radioStartTime # [yyyymmdd.hhmmss]
seriesDateTime = seriesDate + "." + seriesTime # [yyyymmdd.hhmmss]
radioDecayTime = seriesDateTime - administrationDateTime # [s]
radioDecayDose = radioTotalDose * exp(-radioDecayTime * (log(2)/radioHalfLife)) # [Bq]
BWscaleFactor = (patientWeight * 1000) / decayedDose # [g/Bq]
SUVbw = (PETimageValue + rescaleIntercept) * rescaleSlope * BWscaleFactor
```

*Alg. 1: $\text{SUV}_{BW}$ Computation via DICOM tag values.*

the proposed palettes are not hard-coded and depend on the number of defined ROIs. If implemented, the current palette should not be updated with each new ROI, i. e., updating the color-coding repeatedly. Practically, a fixed color palette size should be defined by default and adapted during report generation and export.

The proposed color-coding methods use the HSV color space. It can be argued that this color space represents a fair tradeoff between usability for non-visualization domain experts and perceived color differences between ROI colors. On the one hand, for subsequently defined ROIs, the step equations guarantee a large hue difference, which makes the colors easy to distinguish, even if they are defined in the same image. Since the proposed equations define ROI colors with respect to $B_H$, the first ROI colors are defined near $B_H$ and the color palette will slowly converge towards $M_H$. When using software tools to define ROIs, clinical experts encompass suspicious image regions, i. e., regions that will most likely be diagnosed to be inconclusive or malign. This lowers the risk of mentally connecting ROI color-codings with image regions diagnosed as malign, which could alter diagnoses and therapy decisions. On the other hand, equidistantly defined colors in HSV color space do not result in equidistantly perceived color differences. Using perceptually uniform color spaces, e. g., CIELAB, could result in better results, but using them as GUI widgets, e. g., the CIELAB *shoe sole*, would be less beneficial for clinical toolkits than a simple color wheel. Furthermore, in contrast to the VG51C color palette, the proposed methods only alter hue values and keep the saturation and value components constant (cf. Fig. 1 and 2). This limits the proposed methods, since varying saturations and values could also offer better results, but it is challenging to perceive ROIs colors with low saturation and value components due to grayscale MIPs of PET images. Practically, the definition of color palettes should generally be done in the background and only brought to the users' attention if specifically queried, e. g., to define special color palettes in the case of color vision deficiencies.

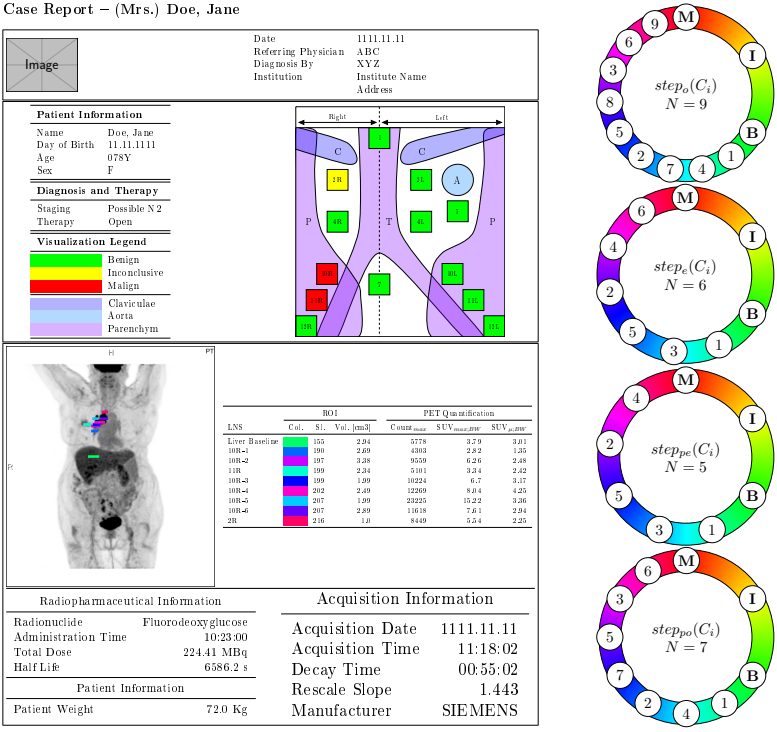



*Fig. 3: On the left, an exemplary report with a color-coded 2D Map from Figure 1, color-coded ROIs, and $SUV_{BW}$ values is depicted. The four color wheels on the right show ROI color-codings for varying Ns. For the report, the upmost color wheel with $N = 9$ was employed and the respective values are listed in Table 1.*

To evaluate the accuracy of Algorithm 1, resulting $SUV_{BW}$ values were compared to SUVs that were obtained using the VG51C software (cf. Fig. 2). This was done by manually reconstructing ROI segmentations from exported screenshots. Regarding $SUV_{max;BW}$ values, for each ROI, there were no differences and thus, the presented algorithm produces reliable and reproducible results. However, this was not the case for $SUV_{\mu;BW}$ values, since even small differences between ROIs can result in significantly different SUVs, which can influence diagnosis and therapy decisions [5].

Currently, the proposed map-based approach only depicts diagnoses with respect to the N-staging component for thoracic lymph nodes [10]. However, the basic idea behind the presented approach can be extended to cover the tumor and metastasis components of the TNM staging system, or to cover other organs, e. g., the liver or prostate. One shortcoming of the visualization-supported reports is that color printers are required, which have higher running costs compared to monochrome printers, but the enhanced reports have the potential to be better understood by a broader clinical audience.

## 5 Conclusion

We presented how clinical reports can be supported via standardized map-based visualizations. Two color-coding methods were presented. The first method can be used for map-based depictions of LNSs and anatomical landmarks, and the second method can be applied to color-code an arbitrary number of ROI segmentations in PET image data. The second method defines colors in a hue range that can be adjusted to not visually overlap with PET image color-codings to decrease the risk of misinterpretations. Finally, a guide was presented to compute reproducible $SUV_{BW}$ values. Where applicable, methods were compared to a clinical toolbox.

**Acknowledgements** This work is partly funded by the Federal Ministry of Education and Research within the Forschungscampus *STIMULATE* (Grant Number 13GW0095A).

**Supplementary Material** The colored version can be found at `http://www.vismd.de`

## References

[1] Birr, S., Dicken, V., et al., *3D-PDF: Ein interaktives Tool für das onkologische Reporting und die Operationsplanung von Lungentumoren*, In: Proc. of German Society of Computer- and Robot-Assisted Surgery (CURAC), Magdeburg, Germany, Sept. (2011)

[2] Brewer, C. A., *Color Use Guidelines for Mapping and Visualization*, Vis Mod Cartogr, **2**, pp. 123–148 (1994)

[3] Foster, B., Bagci, U., et al., *A Review on Segmentation of Positron Emission Tomography Images*, Comput Biol Med, **50**, pp. 76–96 (2014)

[4] Itten, J., *The Art of Color*, ISBN: 0471289280, New York, USA (1973)

[5] Joo Hyun, O, Lodge, M. A., et al., *Practical PERCIST: A Simplified Guide to PET Response Criteria in Solid Tumors 1.0*, Radiology, **280**(**2**), 576—584 (2016)

[6] Klein, J. L., Garcia, E. V., et al., *Reversibility Bull's-Eye*, J Nucl Med, **31**(**7**), pp. 1240–1246 (1990)

[7] Kreiser, J., Meuschke, M., et al., *A Survey of Flattening-Based Medical Visualization Techniques*, **37**(**3**), pp. 597–624 (2018)

[8] Merten, N., Genseke, P., et al., *Towards Automated Reporting and Visualization of Lymph Node Metastases of Lung Cancer*, In: Proc. of Workshop Bildverarbeitung für die Medizin (BVM), Lübeck, Germany, Mar. (2019)

[9] Mountain, C. F., Dresler, C. M., *Regional Lymph Node Classification for Lung Cancer Staging*, Chest, **111**(**6**), pp. 1718–1723 (1997)

[10] Rami-Porta, R., Crowley, J. J., et al., *The Revised TNM Staging System for Lung Cancer*, Ann Thorac Cardiovas, **15**(**1**), pp. 4–9 (2009)

[11] Reiner, B., Siegel, E., *Radiology Reporting: Returning to Our Image-Centric Roots*, Am J Roentgenol, **187**(**5**), pp. 1151–1155 (2006)

[12] Ritter, F., Boskamp, T., et al., *Medical Image Analysis*, IEEE Pulse, **2**(**6**), pp. 60–70 (2011)

[13] Rössling, I., Dornheim, J., et al., *The Tumor Therapy Manager*, In: Proc. of Information Processing in Computer-Assisted Interventions (IPCAI), Berlin, Germany, June (2011)

[14] Soongsathitanon, S., Masa-Ah, P., et al., *A new Standard Uptake Values (SUV) Calculation Based on Pixel Intensity Values*, Int J Math Comput Simul, **50**(**1**), pp. 26–33 (2012)

[15] Tufte, E. R., *The Visual Display of Quantitative Information*, ISBN: 0961392142, Cheshire, CT, USA (2001)

# Accuracy of electrode position in sphenopalatine ganglion stimulation in correlation with clinical efficacy

*G. Santos Piedade, T. Klenzner, J. Vesper, P. J. Slotty*

*University Hospital Düsseldorf, Department of Neurosurgery, Düsseldorf, Deutschland*

Sphenopalatine ganglion (SPG) stimulation is a well-established treatment for chronic cluster headache. The target for the ATI SPG microstimulator in the pterygopalatine fossa (PPF) corresponds to the classical anatomic description of SPG location: posterior to the middle nasal turbinate, between vidian canal (VC) and foramen rotundum (FR). Although it is recommended to insert at least two contacts in this position, the correct placement is due to particularities of the anatomical region in many cases not possible. It is not known whether a suboptimal electrode placement interferes with the postoperative outcomes. SPG stimulation was performed in 13 patients between 2015 and 2018 in the Department of Neurosurgery at the University Hospital in Düsseldorf, intraoperative CT documented lead placement in relationship with osseous structures and the data were correlated to the preoperatively planned electrode position, as well as to clinical data regarding characteristics of cluster attacks and to stimulation parameters. Patients with a reduction of 50% or more in pain intensity after stimulation or in pain frequency were considered responsive.
A total of eleven patients (84.6%) responded adequately to SPG stimulation, eight being frequency responders (61.5%), one acute responder (7.7%) and two frequency and acute responders (15.4%). In seven cases, there were less than two electrodes between VC and FR, there was no significant correlation with negative stimulation results ($p = 0.91$). The mean distance between pre- and postoperative images was 4.85 mm (SD 2.41 mm), no significant correlation with postoperative outcomes was found ($p = 0.84$) and mean distance was even greater in responders (4.91 mm vs. 4.53 mm). In all cases, the closest contact to VC was inside the stimulation area. The most frequent side effect were transitory sensory disturbances that resolved in 3 months after surgery (53.8%). Concluding, there is no significant correlation between the suboptimal lead position and worse postoperative outcomes. Preoperative planning may be helpful, but not determinant to patient's responsiveness to SPG stimulation.

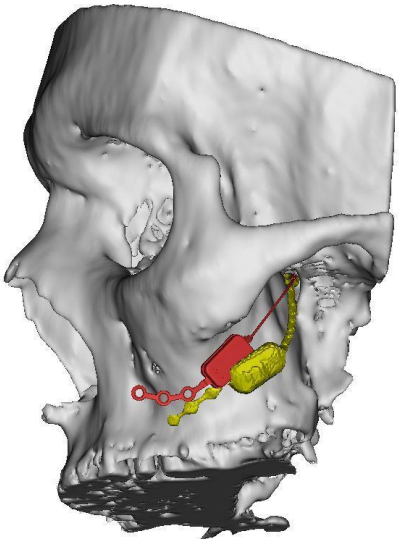


Fig. 1*: The implanted microstimulator system is depicted in yellow, red indicates the preoperative planning.*

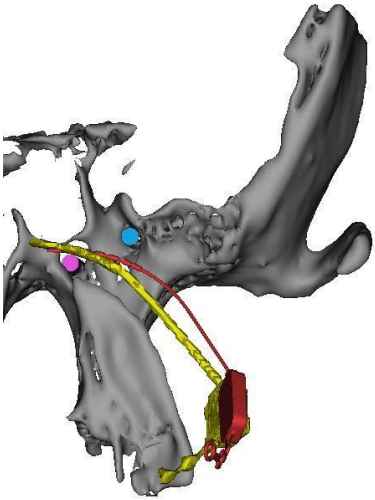


Fig. 2*: Three-dimensional reconstruction of the sphenoid bone. Implanted lead is depicted in yellow, the preoperatively planned ideal position is shown in red. Vidian canal is indicated with a pink point, foramen rotundum in blue.*

# Reproducibility Evaluation of Palpation of Anatomical Landmarks for Estimation of the Patient Location

*Nazario Carlos Aguilar Hidalgo[1,2], Lukas Brand[1], Oliver Burgert[1], Eckhart Fröhlich[3]*

*[1]Hochschule Reutlingen, Fakultät Informatik, Forschungsgruppe Computer Assisted Medicine (CaMed), Reutlingen, Deutschland*
*[2] Universidad de Sevilla, ETSII, Sevilla, España*
*[3] Department of Internal Medicine I (Gastroenterology, Hepatology, Infectious Diseases), University Hospital Tübingen, Tübingen, Germany*

Kontakt: oliver.burgert@reutlingen-university.de

***Abstract***

*This study is about estimating the reproducibility of finding palpation points of three different anatomical landmarks in the human body (Xiphoid Process and the 2 Hip Crests) to support a navigated ultrasound application. On 6 test subjects with different body mass index the three palpation point were located five times by two examiners. The deviation from the target position was calculated and correlated to the fat thickness above each palpation point. The reproducibility of the measurements had a mean error of ≈13.5 mm +- 4 mm, which seems to be sufficient for the desired application field.*

**Keywords**: Surgical Navigation, Palpation, Anatomical Landmarks

## 1 Problem

In many clinical applications of surgical navigation, the determination of the patient position and its tracking is a crucial step. In most surgical applications, a reference body is attached to the patient in a reproducible way, e.g. by using bone screws, head clamp, jaw splint, or headbands. Some of those techniques work for different sessions over time (e.g. the jaw splint or bone screws) whilst others like headbands are not guaranteed to be positioned at the same position in the next session. In such cases it is necessary, to calibrate the reference body to the anatomy of the patient for each examination in a reproducible way.

In our use case, we want to track the position of an ultrasound (US) probe in relation to the patient body in order to be able to produce reproducible ultrasound images in different examination sessions [1]. Since we are focusing on abdominal US in a general internal medicine setting, we cannot use invasive techniques like bone screws or expensive techniques like individually produced jaw splints. Furthermore, there are little anatomical landmarks on the upper body, which are clearly visible on one hand, and invariant against body posture on the other hand (e.g. the belly moves depending on the patient posture). Therefore, we decided to use palpation points as anatomical landmarks, in our case the Xiphoid Process and the two Hip Crests. To evaluate whether this approach is feasible, we evaluated the reproducibility of the palpation of the three anatomical landmarks in a first evaluation study.

## 2 Material and Methods

We recruited six test subjects with different body mass index (BMI) [2], categorized in low BMI (<= 18.5), neutral BMI (18.5 - 24.9) and high BMI = from 25), one male and one female test subject in each category. For this study, all test subjects agreed on the participation in the study and the usage of the gathered data. For each test subject, each palpation point was palpated five times by two different examiners. The examiners were computer science students, trained by a clinical expert in ultrasound examinations prior to the study. The test subjects did not move during the test and the position of the palpation points was measured using an optical tracking system (NDI Polaris Vicra) and the software NDI Track for reading the position values. For each palpation point the fat thickness was measured using an US device (esaote MyLab Sat) which was also used to confirm the correct palpation reference point.

# 3 Results

The reproducibility of the measurements had a mean error of ≈13.5 mm +- 4 mm (14.8 mm for examiner one and 12.4 mm for examiner two) calculated over all palpation points. Figure 1 (left) shows the distribution of errors. Figure 1 (right) shows the correlation between measured fat thickness and measurement accuracy.

Mean time for the measurements including ultrasound verification and fat measurement per test subject with BMI < 19 is 433.8s, for BMI 19-25 436.2 and for BMI >25 535.8s.

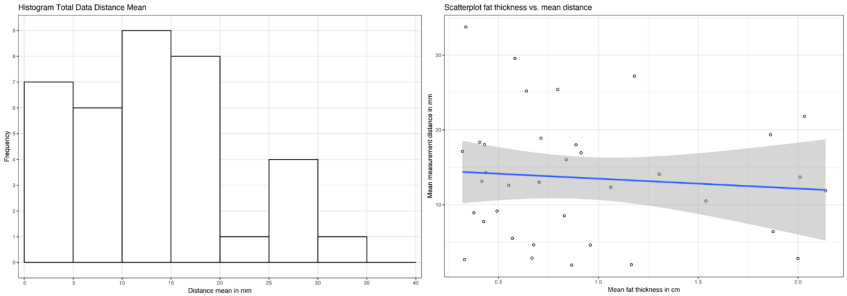



*Figure 1: Histogram of the mean distances in mm to the reference points of the total dataset (left) and Scatterplot comparing fat thickness and mean distance. Blue line shows a linear model fitted to the data (using geom_smooth method lm in R). Gray shows the confidence region. (right)*

# 4 Discussion

The BMI and fat thickness above the palpation point does not seem to influence the measurement quality significantly, but they influenced the measurement time. We assume that this effect is less prominent for clinical experts with larger palpation experience. The mean error of 13.5 mm +- 4 mm for each palpation point might be sufficient for the navigation of the US probe in our setting.

# 5 Conclusion

The reproducibility study showed that palpation is error prone and for our application a careful palpation and location of the anatomical landmarks is crucial. The mean errors are small enough to further evaluate this approach and in a next step we will evaluate the resulting target registration error and the influence on the recorded ultrasound images.

# Acknowledgements

We would like to thank our test subjects for the participation in this study.

# References

[1] Grupp P., *Untersuchung der Anforderungen an ein System zur Unterstützung der Reproduzierbarkeit von Ultraschalluntersuchungen*. connect (IT) (2018). ISBN: 9 -83000-586453

[2] Rothman, K. J., *BMI-related errors in the measurement of obesity*, International Journal of Obesity 32, S3, S56.

# Anrufmanagementsystem: ermöglicht der „kognitive OP Saal“ die Reduktion unnötiger Stressfaktoren während eine Operation?

**N. Samm[1 2], N. Kohn[1], D. Ostler[1], T. Vogel[1 2], N. Marahrens[1], M. Kranzfelder[1 2], D. Wilhelm[1 2], H. Friess[2], H. Feussner[1 2]**

[1] Forschungsgruppe MITI, Klinikum Rechts der Isar, Technische Universität München, Trogerstr. 26, 81675 München
[2] Klinik und Poliklinik für Chirurgie, Klinikum Rechts der Isar, Technische Universität München, Ismaninger Str. 22, 81675 München

Kontakt: Nicole.Samm@tum.de

**Einleitung:**

Die telefonische Erreichbarkeit eines Chirurgen als Entscheidungsträger bei der Versorgung seiner Patienten ist von großer Bedeutung. Andererseits haben ständige Unterbrechungen des chirurgischen Workflows im OP durch eingehende Anrufe einen Einfluss auf die Konzentration des Arztes und können dadurch zu einem negativen Einfluss auf die Patientensicherheit und das operative Ergebnis führen.

**Ziele:**

Ziel dieser Studie war es, bei unterschiedlichem medizinischen Fachpersonal zu evaluieren, ob ein kontextsensitives Anrufmanagementsystem, welches den Anrufer über das aktuelle Stresslevel im OP und über die verbleibende Rest-OP Zeit informiert, die telefonische Belastung des OP-Teams reduzieren kann.

**Methodik:**

46 Probanden aus unterschiedlichen Funktionsbereichen des Klinikum rechts der Isar bekamen 3 Szenarien mit unterschiedlicher Dringlichkeitsstufe vorgelegt. Szenarien der Klasse III waren von geringer Wichtigkeit. Die der Klasse II hatten eine höhere Bedeutung für die Patientenversorgung und Szenarien der Klasse I die höchste Dringlichkeitsstufe. Eine zweite Dimension war die Angabe der ungefähren restlichen Operationszeit (60, 30 oder 15 Min.). Die 3. Dimension war eine Angabe des Systems zur aktuellen Stresssituation im OP (normal, angespannt, kritisch). Die Probanden sollten anhand der Informationen im Rückschluss auf den Grund Ihres Anrufes entscheiden ob sie durchgestellt werden möchten oder nicht.

**Ergebnisse:**

Es zeigt sich ein klarer Zusammenhang zwischen der Bedeutung der Anfrage und der Anzahl der durchgestellten Anrufe. Bei Anfragen von geringer Bedeutung (Klasse III) verzichten 99% aller Probanden auf das Durchstellen ihres Anrufs (4 von 414 Anrufen). Bei Szenarien der Klasse II (mittlere Dringlichkeitsstufe) wurden 7% aller Anrufe an den Chirurgen weitergeleitet (29 von 414 Anrufen). In der höchsten Dringlichkeitsstufe (Klasse I) entschieden sich 78% aller Probanden für eine Verbindung zum Chirurgen (324 von 414 Anrufen). Die Angabe der verbleibenden Operationszeit durch das System zeigt nur einen Einfluss auf die Szenarien der Klasse I (höchste Dringlichkeitsstufe). Bei einer verbleibenden restlichen Operationszeit von 15 Minuten (63%, 57 von 90 gesperrten Anrufen) entschieden sich hier die meisten Probanden gegen das Durchstellen.

**Schlussfolgerung:**

Die Definition der drei Szenarien mit unterschiedlicher Dringlichkeitsstufe in Kombination mit der Abschätzung und Angabe der verbleibenden restlichen Operationszeit durch das System geben dem Anrufer einen klaren Eindruck von der aktuellen Situation im OP und wirken somit einer unnötigen Störung des Operationsablaufs entgegen.

# Autorenverzeichnis

**Y**



**Z**

Die 18. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie in Reutlingen ist das deutschsprachige Forum für den Austausch aktueller Methoden, Ergebnisse und Erkenntnisse in der computer- und roboterassistierten Chirurgie.

Dieser Band umfasst die wissenschaftlichen Beiträge, die im Rahmen dieser Tagung präsentiert wurden.