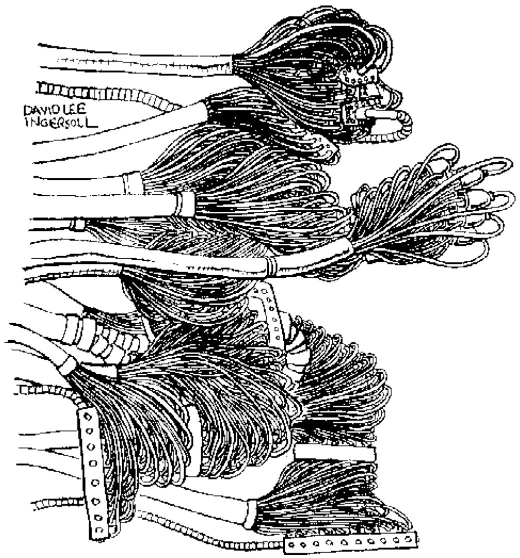
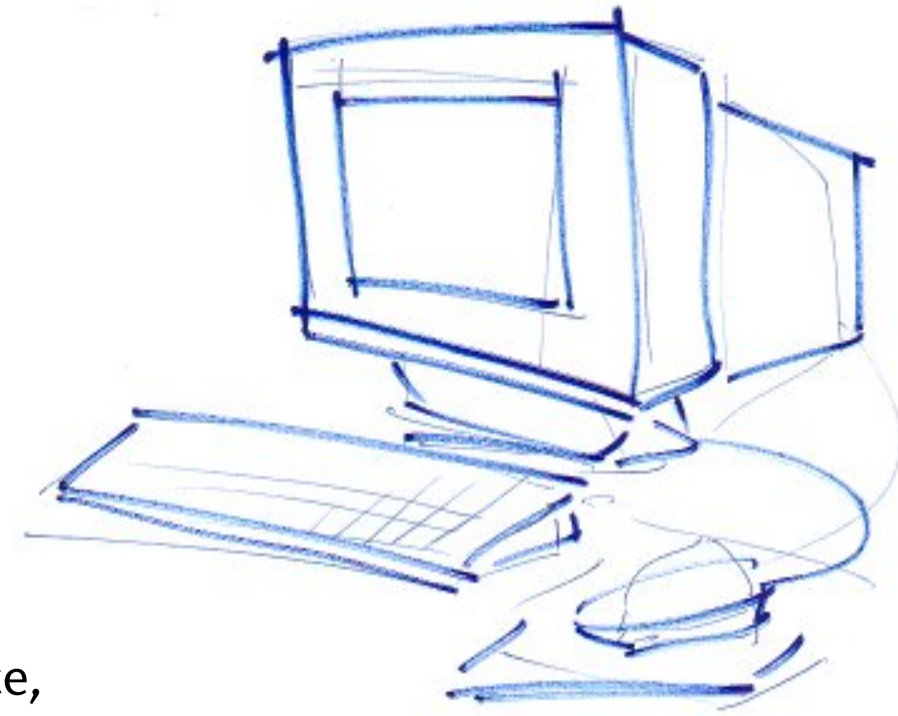
Virtual Reality & Physically-Based Simulation
Principles of Input Devices



G. Zachmann
University of Bremen, Germany
cgvr.cs.uni-bremen.de

The "Bill Buxton Test"

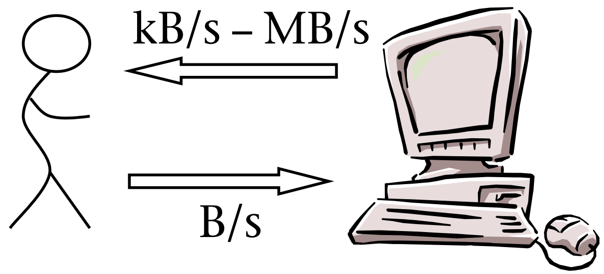
- Draw a computer within 15(!) seconds
- Ca. 80% of all people draw something like this:
  - Monitor
  - Keyboard
  - Mouse
- Remarkable:
  - No "computer" in the drawing!
  - Message: users don't see
    the system as a computer,
    they just see a device on its surface,
    and they just perceive some kind of I/O behavior
    ⟶ the interaction with the device is critical for success/failure



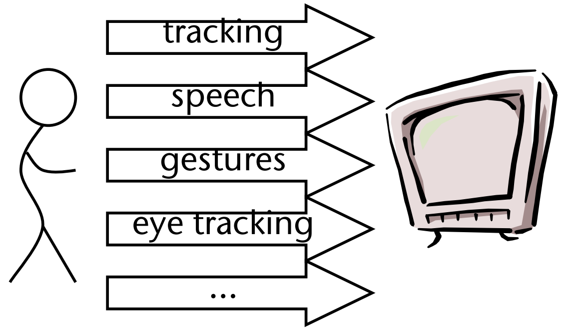
Extreme Examples of "*Intrusive*" I/O Devices

The Promise of Virtual Reality

- Problem of conventional input devices: bandwidth



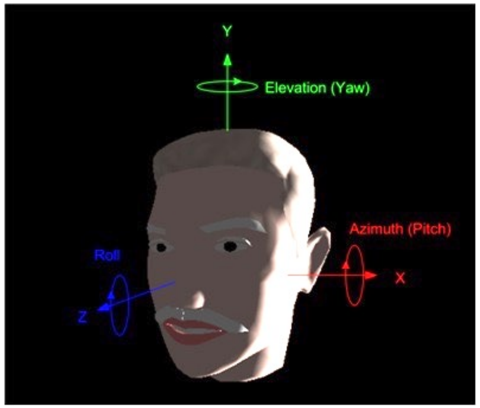
- Multimodal input = input using different modalities, e.g., tracking and voice
  - Post-WIMP interfaces ("WIMP" = windows, icons, menus, pointers)

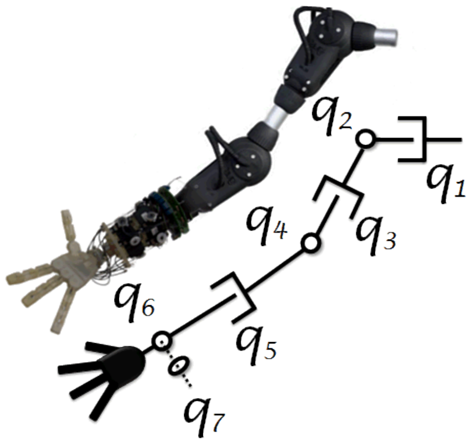


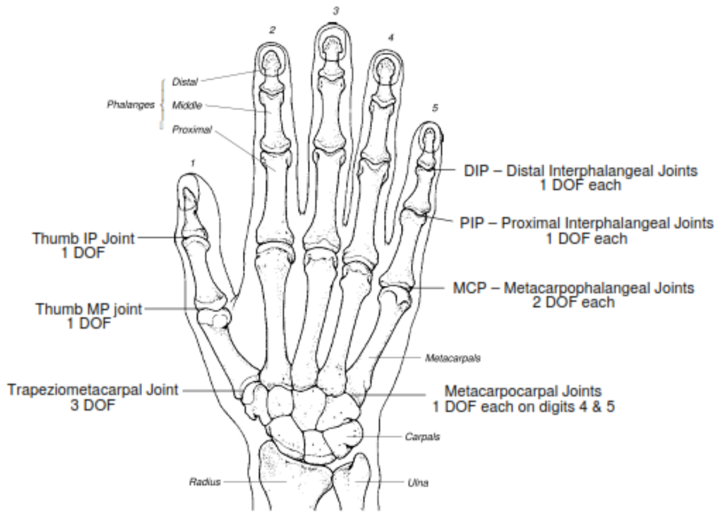
- The "right" interaction paradigms will be paramount

Degrees of Freedom

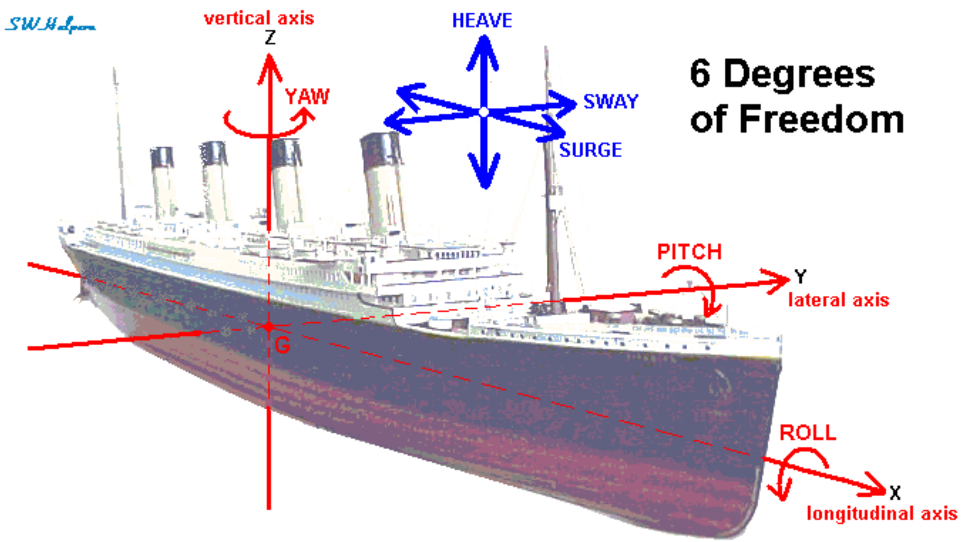
- Definition Degrees of Freedom (DOFs) :=
  number of free variables describing the state of a system
- Quiz about DOFs:
  - How many DOFs does our wrist joint have?
  - The head?
  - One human arm?
  - Our hand?



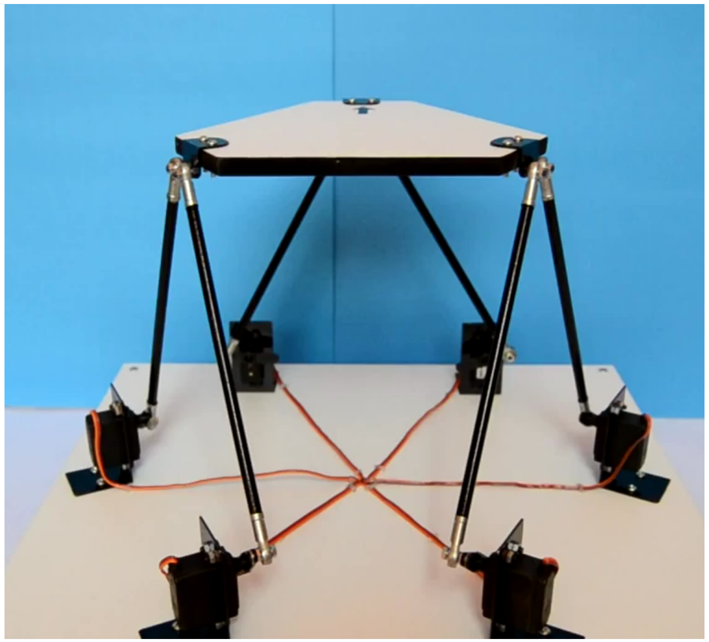




- A ship's pose

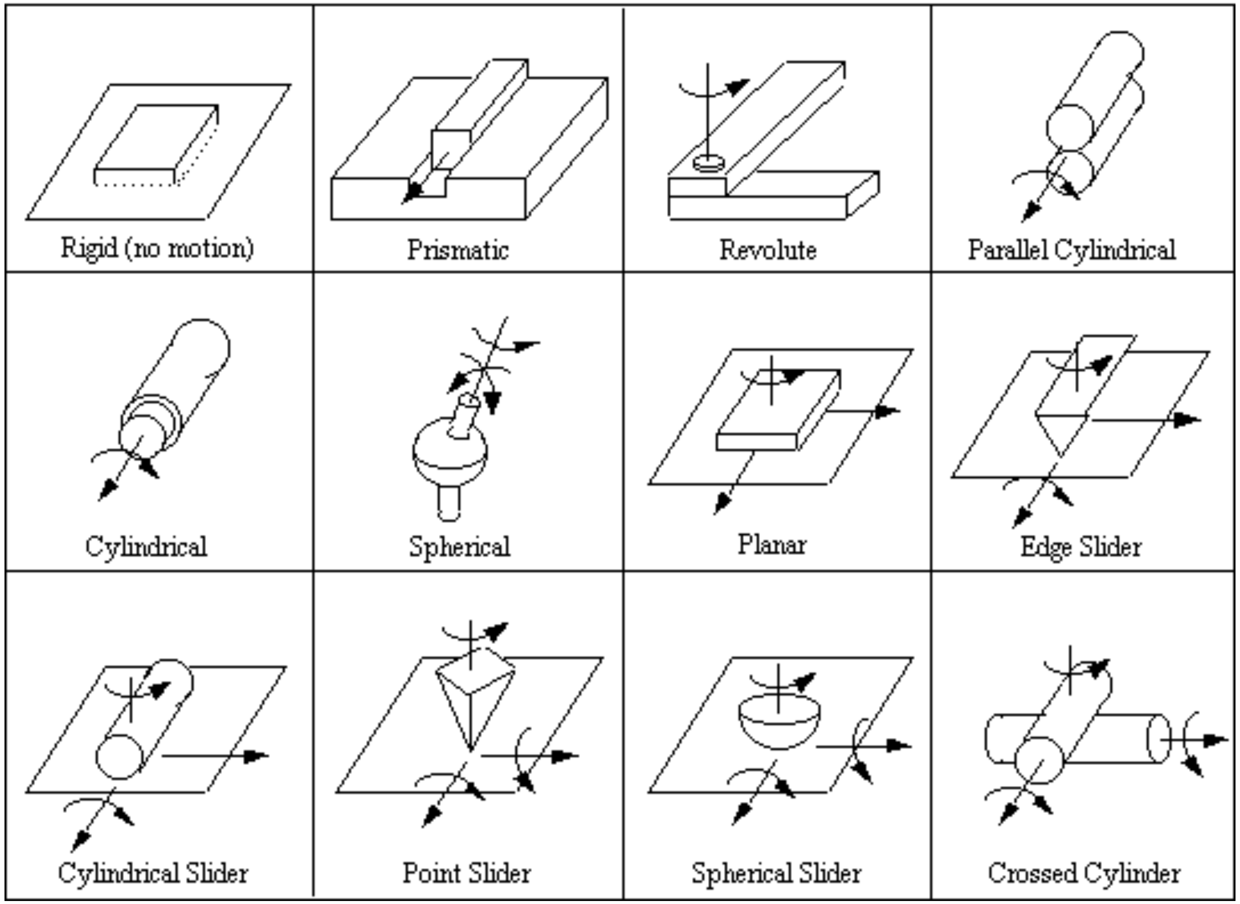


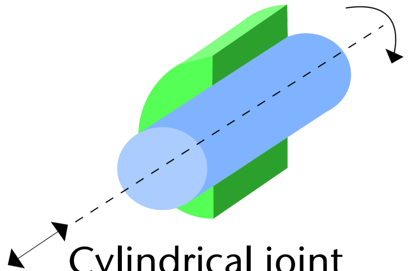
- The Stewart motion platform
  - How many independent DOFs?
  - How many dependent DOFs?



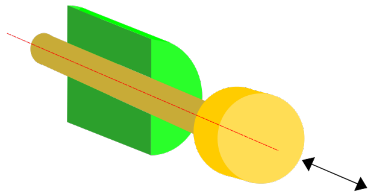
Video

- The main kinematic joints/pairs

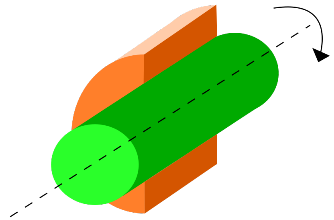




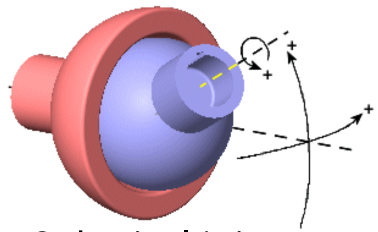
Cylindrical joint



Prismatic joint (slider)

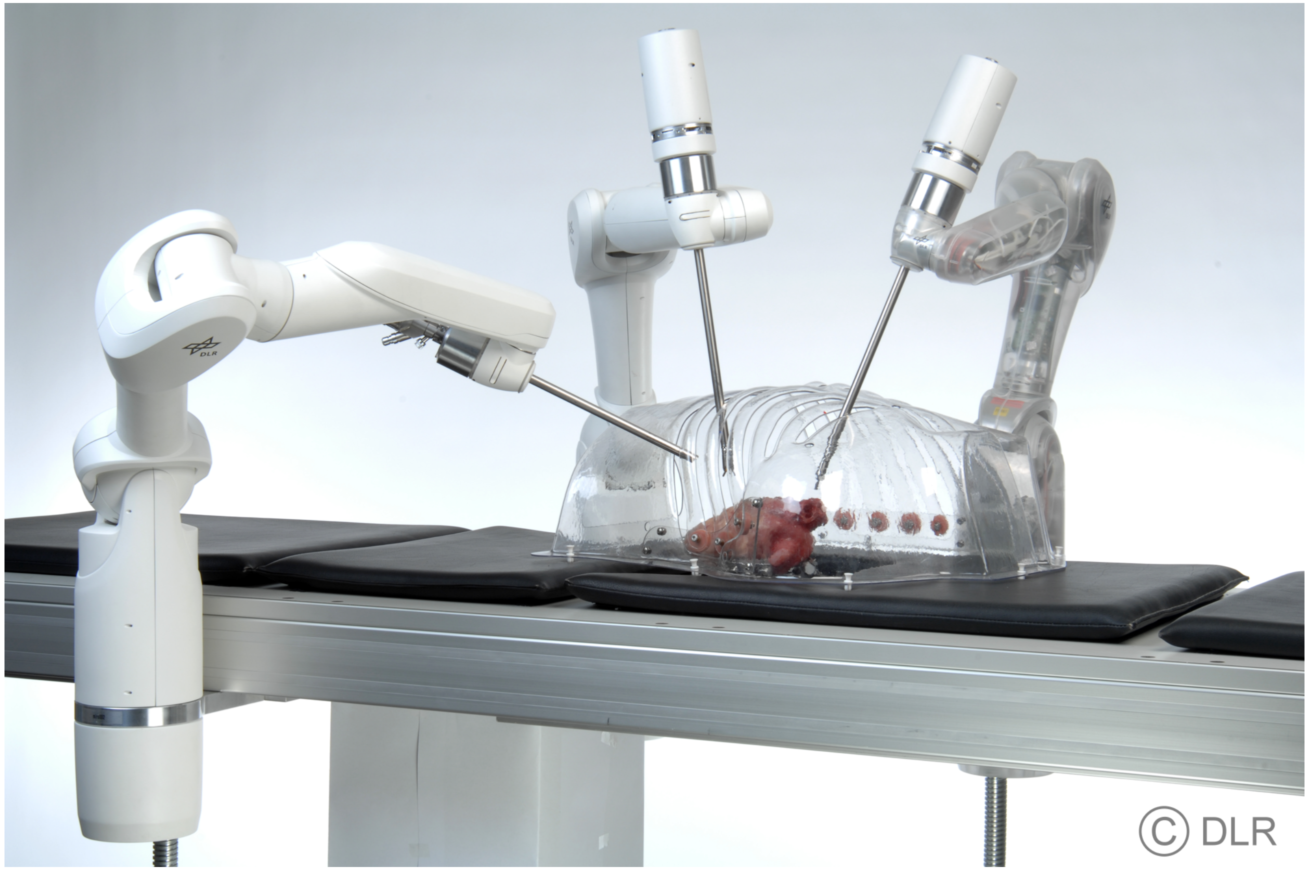


Revolute joint

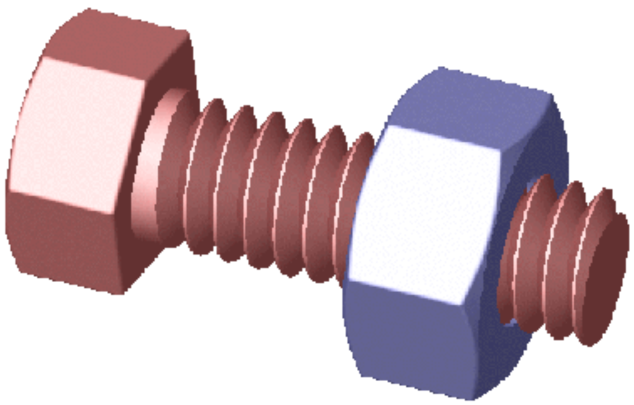


Spherical joint

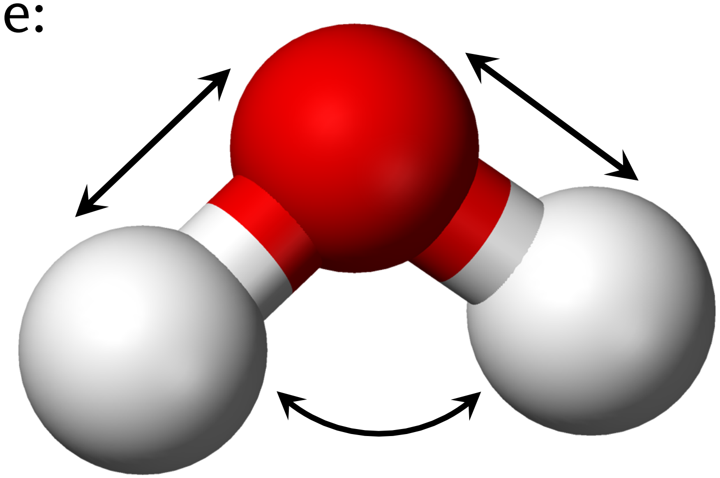
- How many independent DOFs in one robot arm of this surgery robot?



- The screw joint:
  - Joint with coupled rotational and translational degrees of freedom
  - One independent DOF, two(!) dependent DOFs



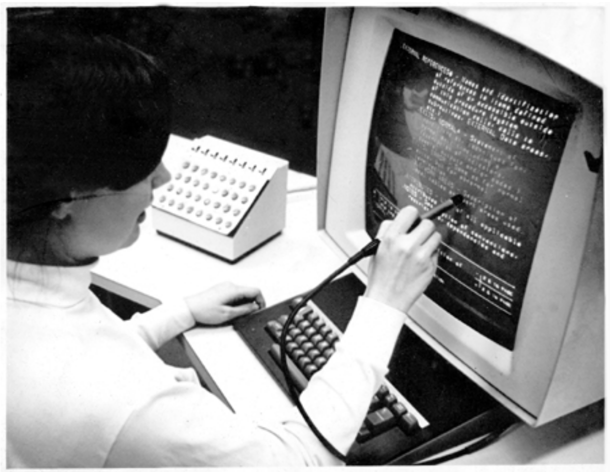
- The internal DOF's of a water molecule:




A fish about to have no degrees of freedom
!!

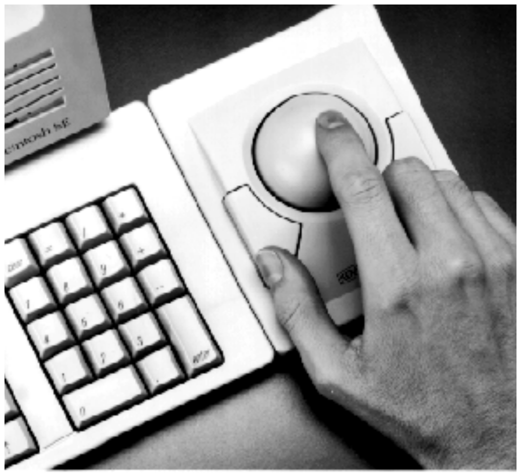
Classical Input Devices

- Mouse:
  - Precise, inexpensive
  - Only 2D, input of orientations is cumbersome
- Drawing tablet:
  - Precise, very well suited for ... drawing
  - 2D, input of orientations is virtually impossible
- Light pen (early version of touch/tablet screen)



The Virtual Trackball

- Interaction task: rotate an object around an arbitrary axis
- Interaction device: classic 2D mouse
  - Would be trivial with real trackball
- Problem: how to enter orientations with a mouse?



The Interactions Metaphor

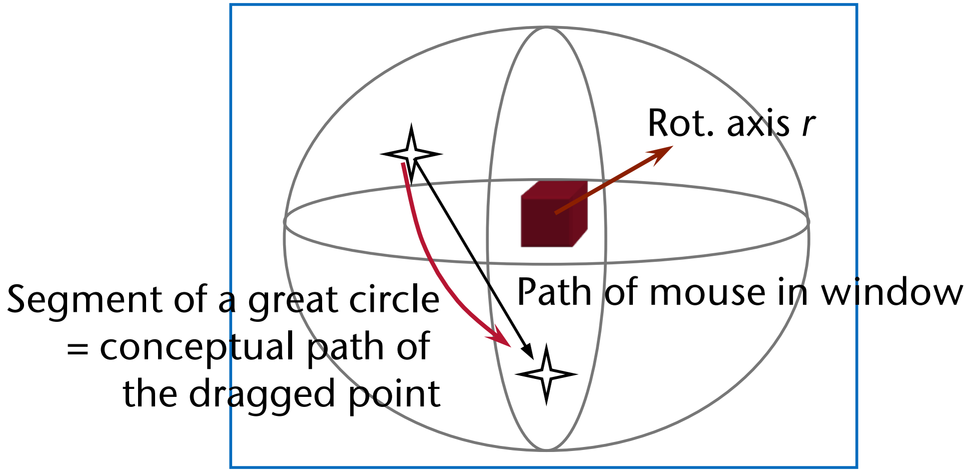
- Idea:
  - Conceptually, put a (virtual) sphere around the object
  - The sphere can rotate only abouts ist center
  - With the mouse, you drag points on the surface of the sphere
- Given: 2D points start = (x_1,y_1), end = (x_2,y_2)
- Wanted: rotation axis **r**
- Computation:
  1. Derive 3D points

$$\mathbf{p}_i = (x_i, y_i, z_i)$$

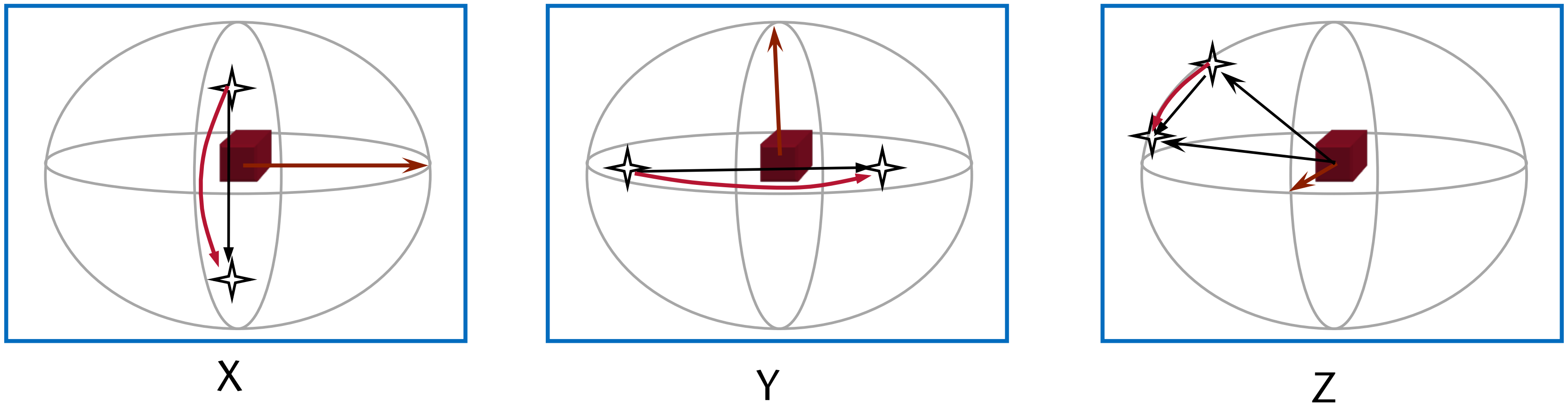
$$z_i = \sqrt{1-(x_i^2 + y_i^2)}$$

  2. Rotation axis

$$\mathbf{r} = \mathbf{p}_1 \times \mathbf{p}_2$$



- If $\mathbf{p}_1$ = first mouse click, $\mathbf{p}_2$ = current mouse pos. → not intuitive
- If $\mathbf{p}_1$ = mouse pos. as of last frame, $\mathbf{p}_2$ = current mouse pos. → intuitive, but rotation exactly about z-axis impossible



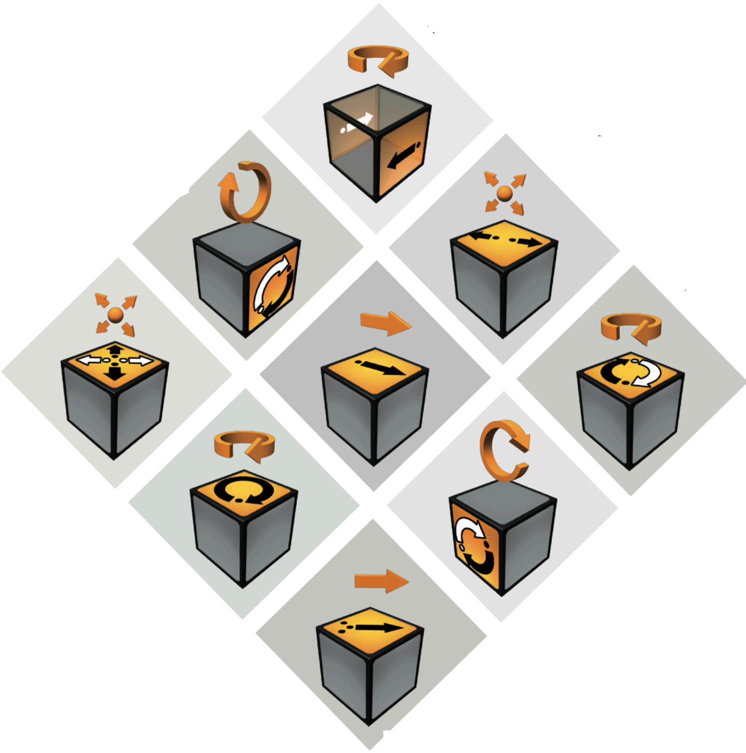
- Improvements / variants:
  - "Spinning trackball": "re-grabbing" the sphere is less often necessary
  - "Snapping": allows precise rotation around world/object coord. axes
  - In case $\mathbf{p}_2$ leaves the ellipse → could use different 3D surface that can be attached continuously to sphere (e.g., hyperboloid)

Remarks

- Rotation axis **r** is given in the camera coordinate frame!
  - You need it in the world frame or object frame
    - Depending on whether the rotation is to be applied to the object before or after all other transformations
- Warning: with variant 2 ("incremental trackball"), a lot of small rotations need to be accumulated! (one per frame) ⟶ consider numerical robustness and drift

Cubtile

- 5 multi-touch surfaces arranged in a cube
- Bonus: very neat illumination ☺
- How many DOFs?



Isotonic vs Isometric Sensing

- Definition isotonic sensing device:
  The user can move the device (or just that DoF) all the way without changing muscle tone.
  - Isotonic = "same muscle tone (tension) during contraction"

- Definition isometric sensing device:
  The device (or just that DoF) does not move when the user pushes/pulls the device (in theory, at least)
  - Isometric = "same muscle metric (length) during contraction"

Example for Isometric Device: Spacemouse

Example for Isotonic Device: Control Action Table

Rotations: controlled by an isotonic sensing mode (cyclic)
Translations: controlled by an isometric sensing mode (infinite)

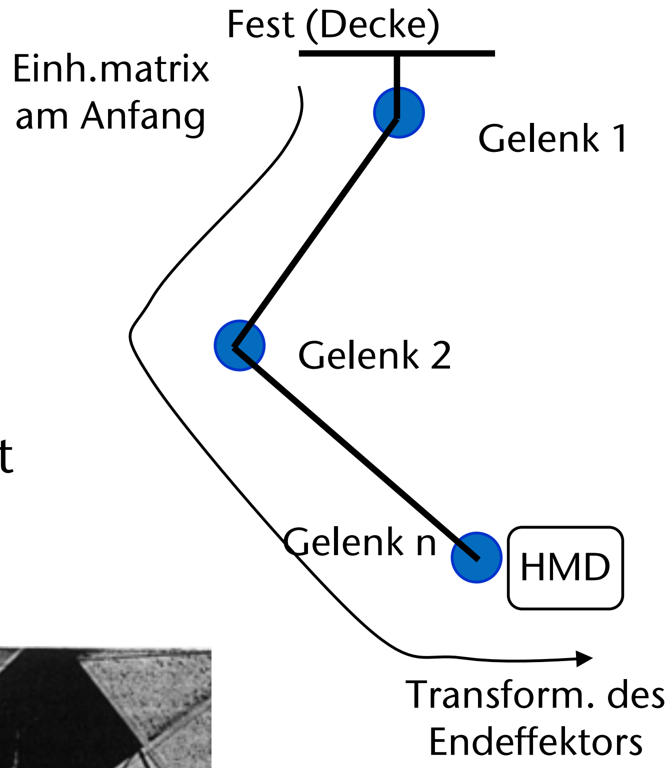
Tracking the User

- Task: determine "where is X of the users?"
  - X = head, hand, eyes, feet, whole body, …
- Requirements:
  - *Non-intrusive*
  - High precision (1 mm)
  - Low latency (1 msec)
  - High *update rate* (100 Hz)
  - Works in all environments and conditions
  - Large working volume
- Doesn't exist (yet?)!

- Technologies for tracking:
  - Mechanical
  - Electro-magnetic
  - Acustic (ultra sound)
  - Optical
  - Computer vision-based
  - Inertia sensors
  - Laser
  - GPS
  - Hybrids

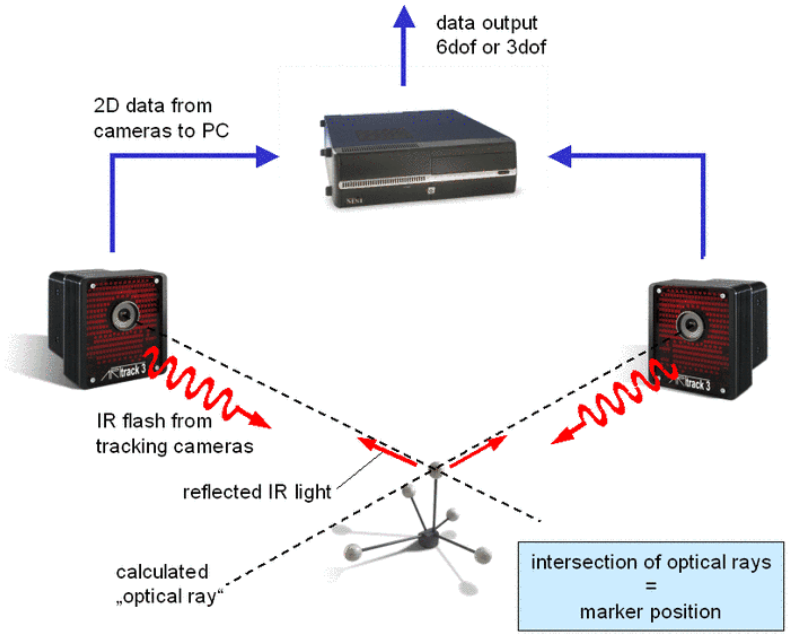
Mechanical

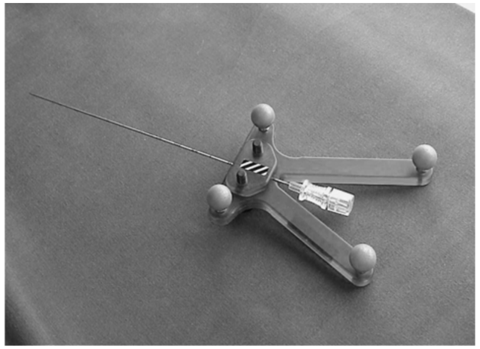
- Advantages:
  - Precision
  - Low Latency
  - No distortion by metal in environment
- Disadvantages:
  - Uncomfortable
  - Working volume
  - "Dead" zones
  - Intrusion
  - Calibration
  - Inertia b/c of mass

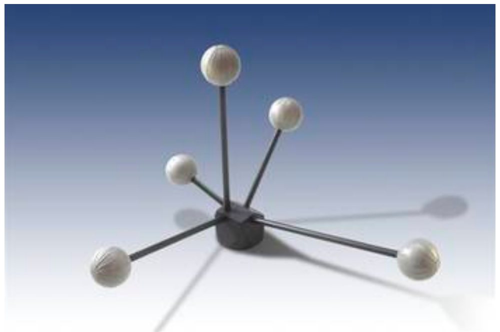


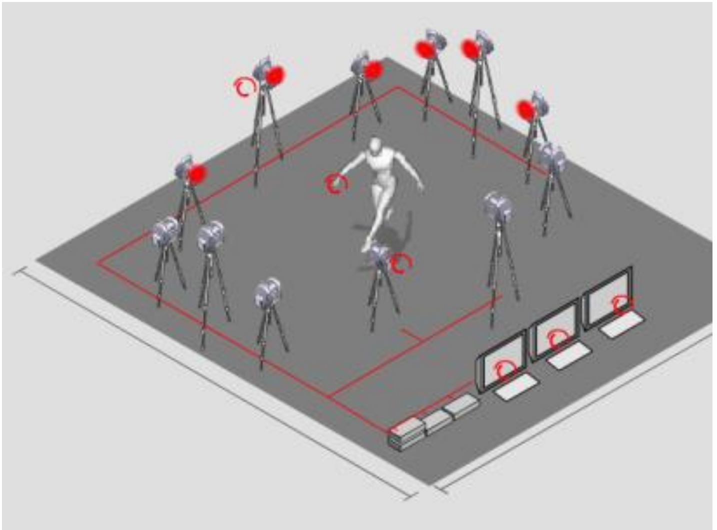
Optical Tracking

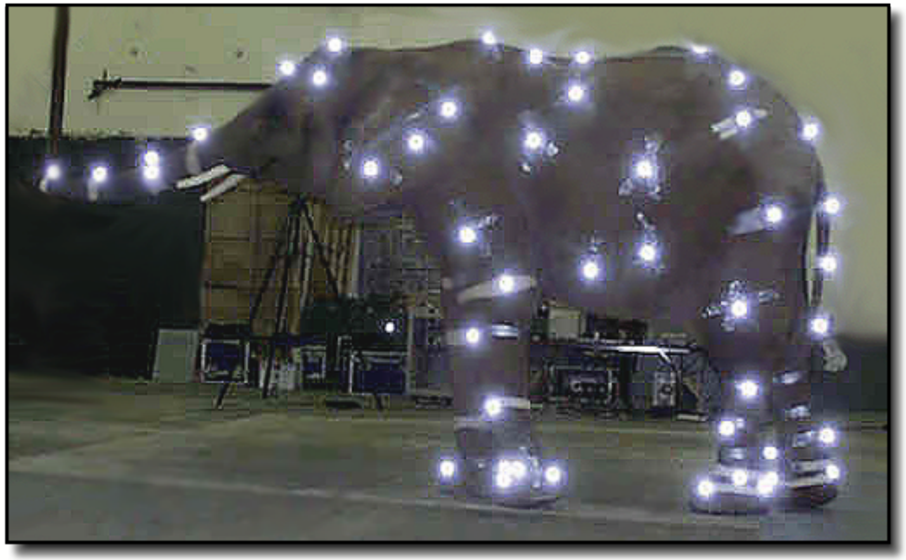
- Idea: track highly reflective markers using IR cameras
- 1 marker → position
  - By way of triangulation
- ≥3 markers (a "*rigid body*") → position and orientation
- Standard technology for body tracking in animation studios and for game dvelopment
  - Motion capturing (MoCap)





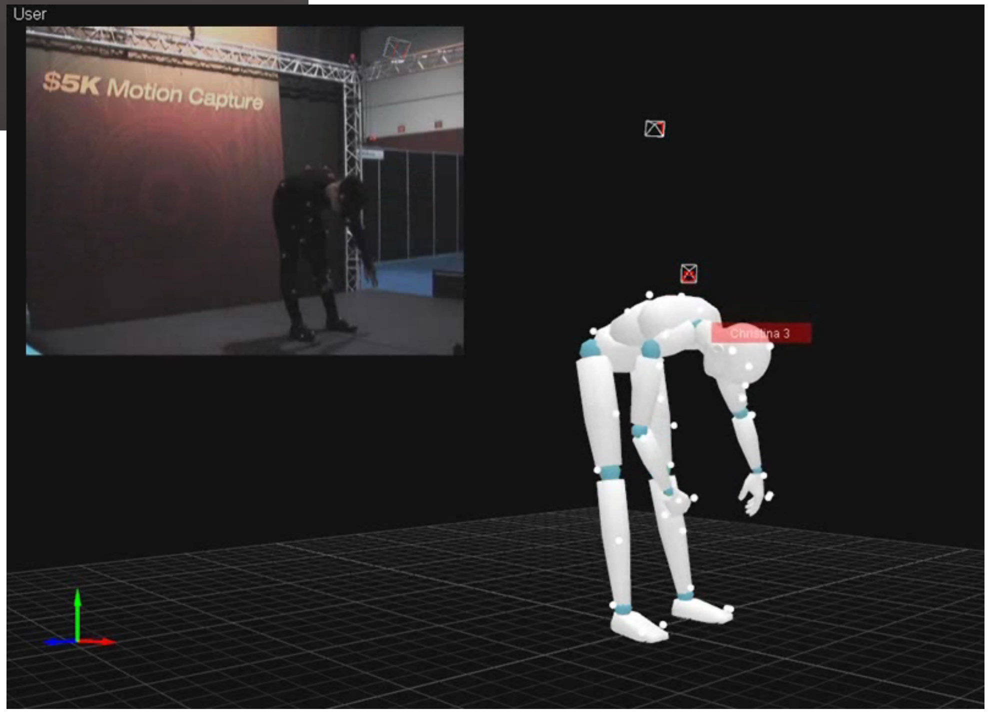




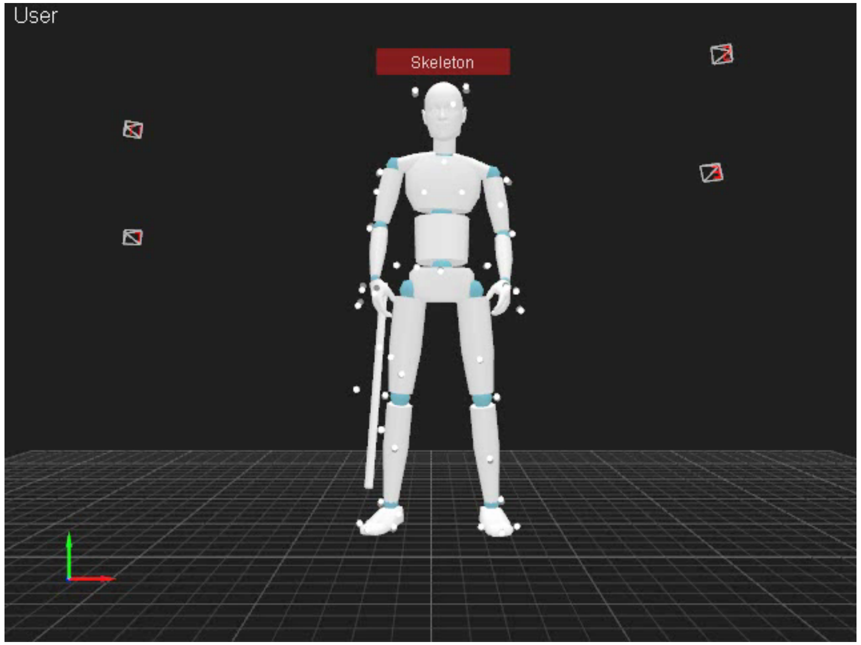


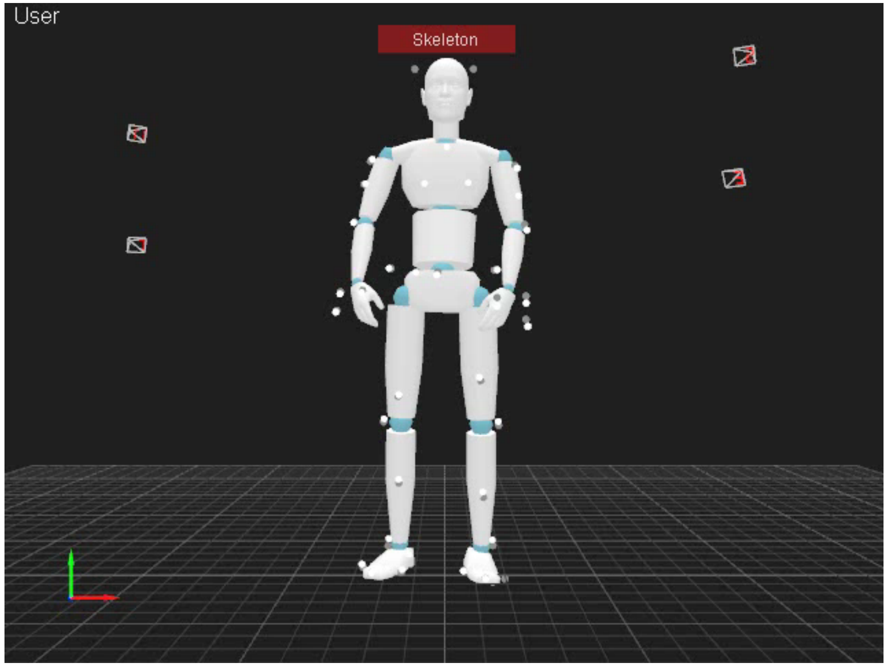
- Advantages:
  - Free movement for users / actors
  - Large working volume
  - High *sampling rate* (typically 120-250 Hz)
  - *Facial animation* is possible, too
- Disadvantages:
  - *Line-of-sight* needed (mitigation: lots of cameras)
  - Price ($40,000 – $140,000)
    - New systems cost only about $6,000

Fluid Images



NaturalPoint (OptiTrack)


User
Skeleton


User
Skeleton

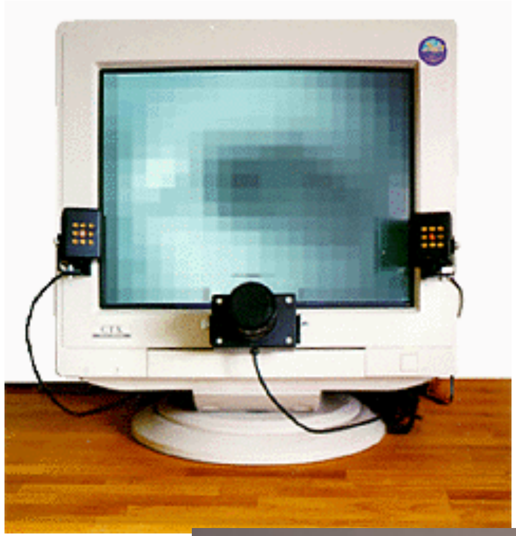
Optical Tracking "*Inside Out*"

- Camera on the head of the user, array of markers (e.g., pulsed LEDs) on the ceiling
- Advantages:
  - Only 1 camera needed
  - Fast (up to 1 msec, 1000 Hz)
  - Precise (1/10 mm)
- Disadvantages:
  - How to track a user's hand?
  - Time- and hardware-consuming installation
- Example: UNC's "HiBall"

  http://www.cs.unc.edu/~tracker/

Eye Tracking

- Where are the user's eyes?
  → eye tracking
- In which direction does the user look?
  → eye gaze tracking
- Applications:
  - Head tracking
  - Controlling LODs
  - Autostereo monitors
- Problems:
  - Precision
  - Sometimes additional hardware is needed



Acoustic Tracking

- Similar to sonar:
  - 1 ultra sound source
  - 3 receivers (for 3 DOFs)
  - Travel time → position
- Advantages:
  - Very inexpensive
- Disadvantages:
  - Echos
  - *Line-of-sight* prerequisite
  - 3 transmitters needed for 6 DOFs
  - Small range
  - Precision: speed of sound depends on air temperature, humidity, etc.

Inertia Sensors

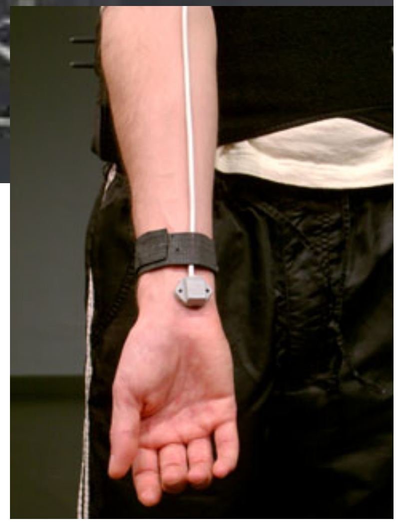
- Measures acceleration in one direction
- Advantages:
  - No transmitter necessary
  - Very small sensors
- Disavantages:
  - Drift
- Often combined with other tracking technologies to compensate for drift, e.g., ultra sound

Laser Tracker

- Measures just distance / position
- So far being used only in manufacturing industries (CNC machines)

Electromagnetic Tracking

- Transmitter =
  - 3 orthogonal coils (using 3 different frequencies)
  - Emit 3 orthogonal electromagnetic fields
- Sensor = receiver =
  - 3 orthogonal coils, too
  - Receive 9 signals in total
- Phase shifts between transmitted and receive signal → distance
- Strength of the 9 different signals → orientation
- Advantages:
  - Small sensors; Working volume = 3 m (or more)
- Disadvantages:
  - Tethering (cables)
  - Metal in environment has severe impact in field distortions
  - Noise



Characteristics of Tracking Systems in Gerenal

1. # DOFs
2. Precision, drift, replicability
3. Update rate, latency
4. Noise
5. Additional buttons
6. E*ase-of-use*, *tethering* (=cables) – *unintrusiveness*!
7. Working volume
8. Price

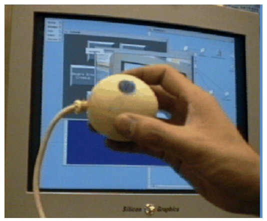
3D Pointers / Stylus / "Controllers"

- Analogue to 2D mouse
- Hardware = tracker with buttons
  - Sometimes with additional joystick, etc.
- Names: *flying mouse*, *flying joystick*, *wand* (= Stab), *bone*, *fly-stick*, etc…
- Advantage: physical object induces a strong feeling of presence while grasping a virtual object

Zhai and Milgram's Directness Continuum for Input Devices

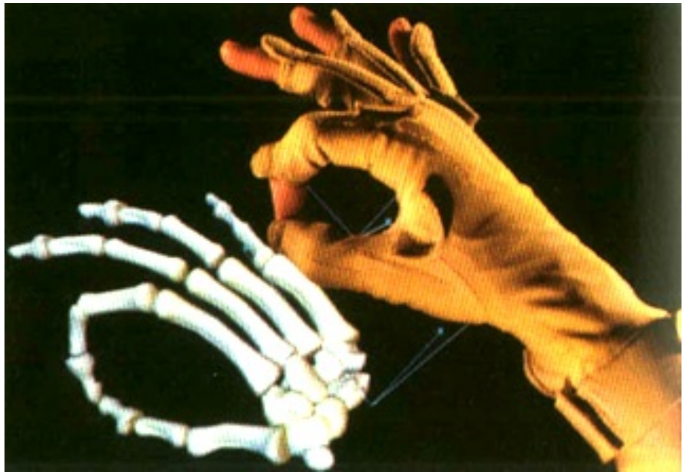
Isomorphic input ← **Directness / Indirectness** → **"Magic tools"**

	Position control			Velocity control	Higher order control (e.g., accel.)
	Absolute		Clutch ("Nachfassen")		
	1:1	scaled			
2-DOF example	*Touch screen*	*Drawing tablet*	*Mouse*	*Joystick*	?
6-DOF example	*Tracking the user*	*Finger-ball*	*Cyber-glove*	*Spacemouse*	?



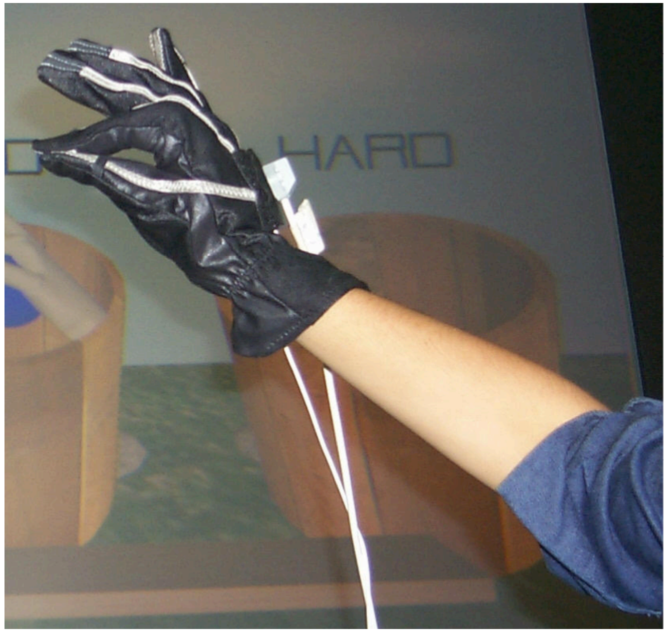
The Dataglove

- "Tracks" fingers of human hand = measures angles of joints
- One of the very early VR devices
- Different numbers of sensors:
  - 12 = 4 (thumb) + 4x2 (2 sensors per finger)
  - 22 = 4 (thumb) + 4x3 (3 sensors per finger) + 3 sensors between fingers + 1 sensor on back of hand (Handrücken)
- Sensor technologies:
  - Glass fibers (not very robust)
  - Bimetallic strips
- Disadvantages:
  - Low precision
  - Glove by and itself (not really accepted)

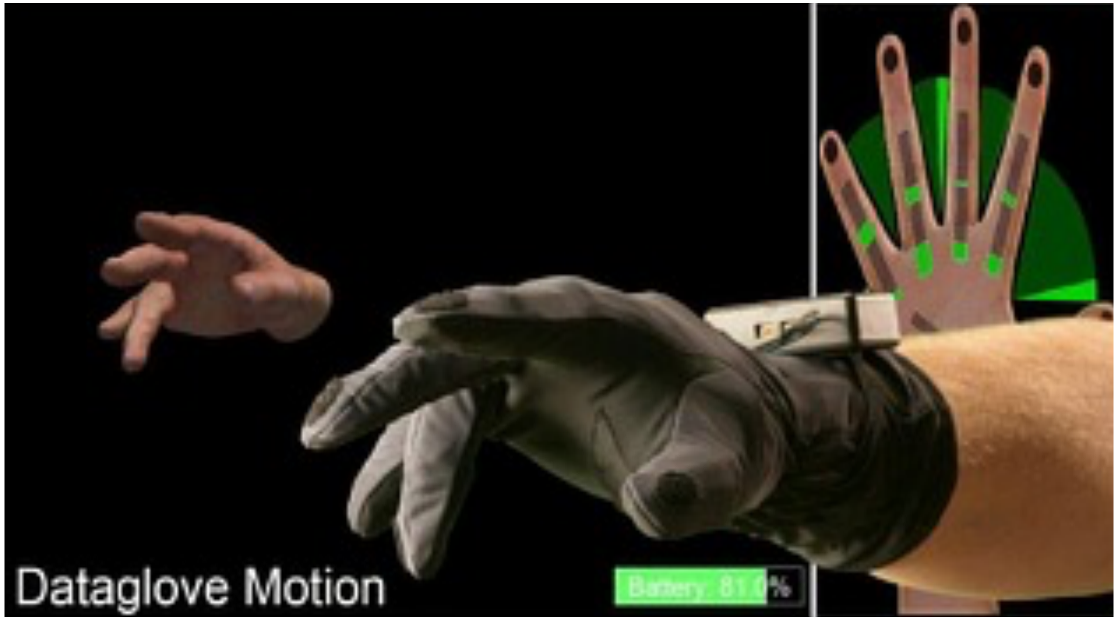


Variants

- Pinch glove:
  - No tracking, just detects contact between finger tips → each finger acts like a button
- Usefuly only using 2 tracked pinch gloves; with, though, pretty clever navigation and manipulations can be performed:
  - Grasping and moving
  - Scaling (using *handles* à la Inventor)
  - Will be presented later …
- Disadvantage: a virtual hand cannot be rendered

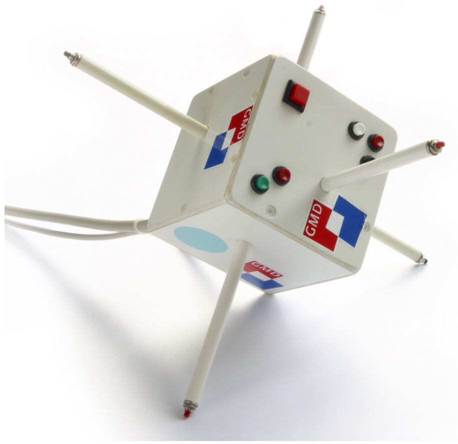


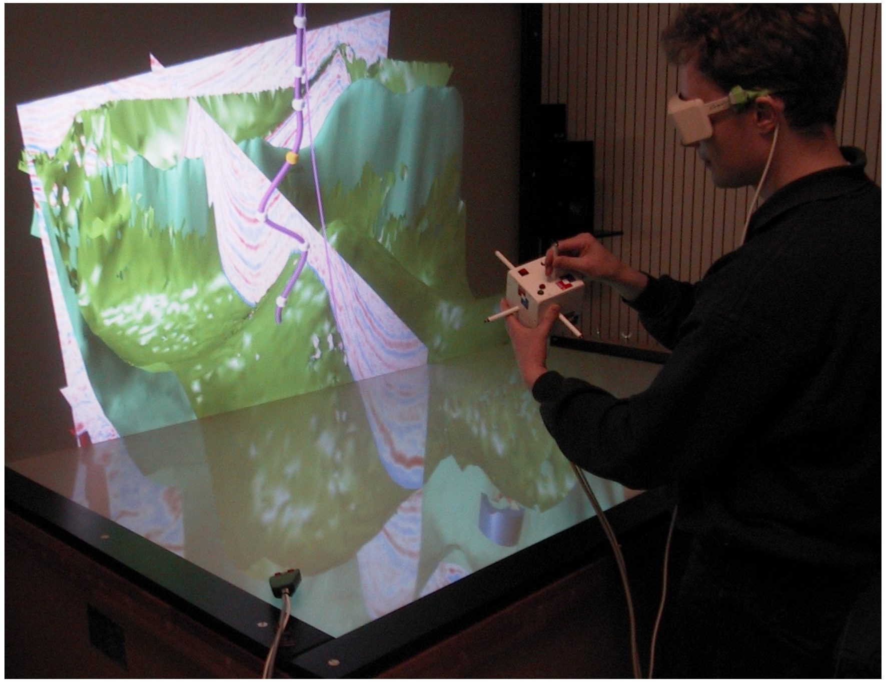
- The P5 from Virtual Realities (www.vrealities.com):

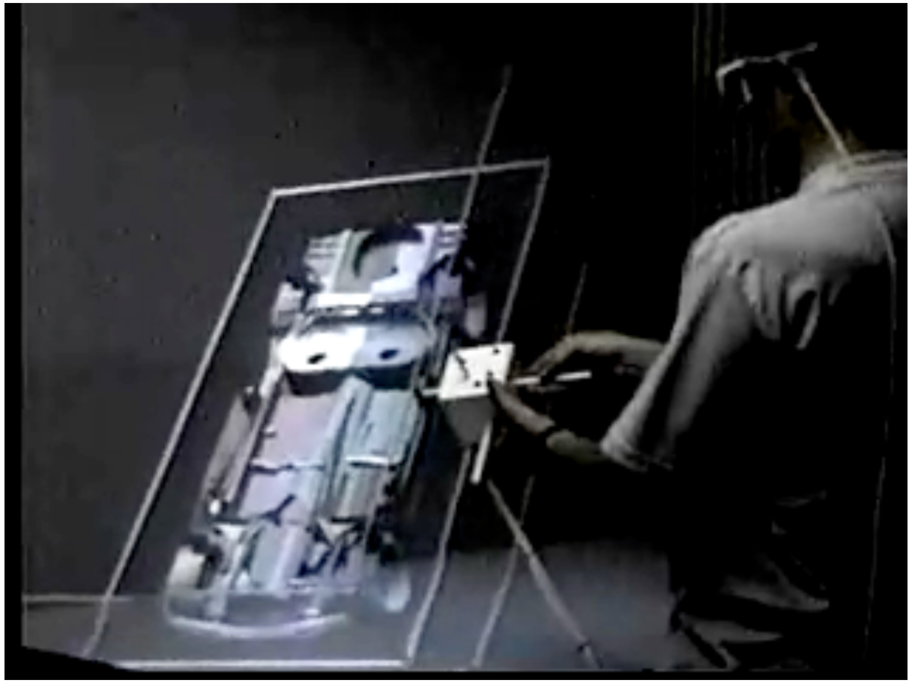


Other High-Dimensional Input Devices

- Cubic Mouse:
  - Number of DOFs = 9



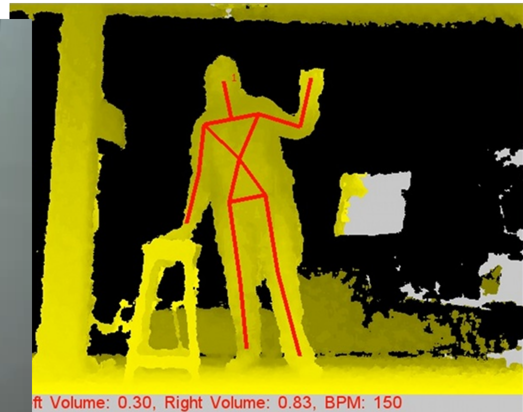


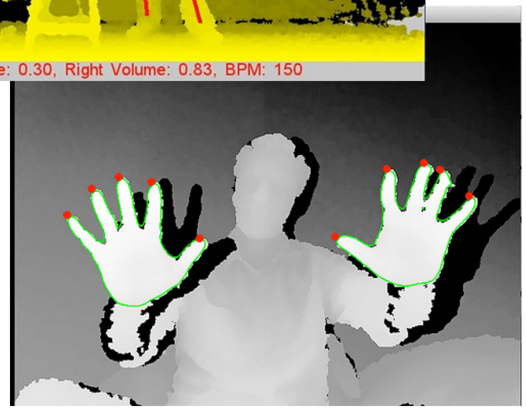


3D Range Sensors

- First consumer device: Microsoft Kinect
- Deliver depth image (range image)
- Lead to so-called natural user interaction (NUI)

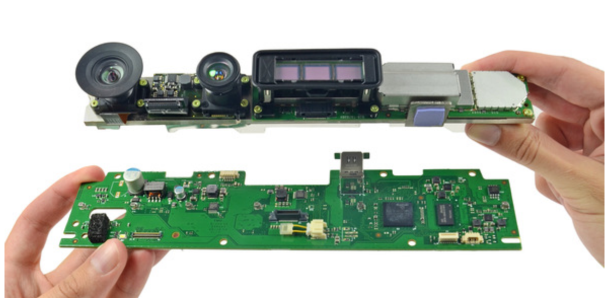
Photo (CC-BY-SA-NC) Kyle McDonald; used by permission

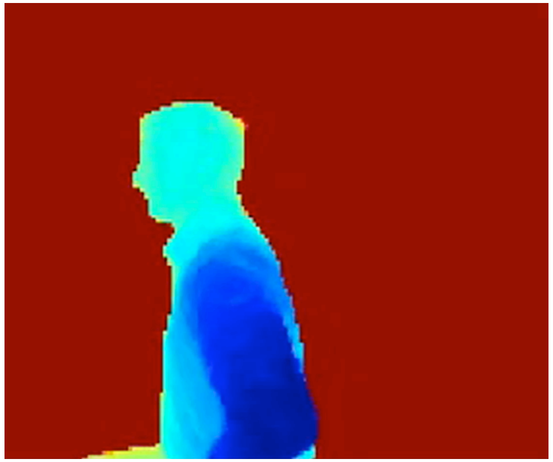


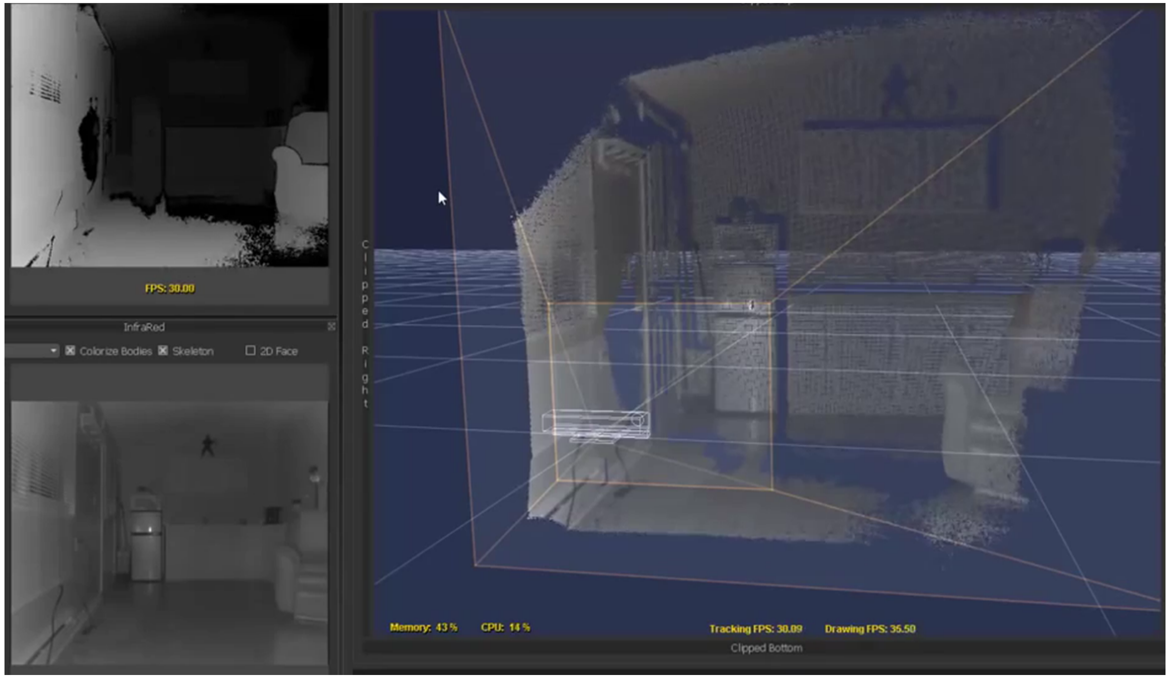


Depth Sensors / Range Image Sensors

- Kinect is the most prominent sensor



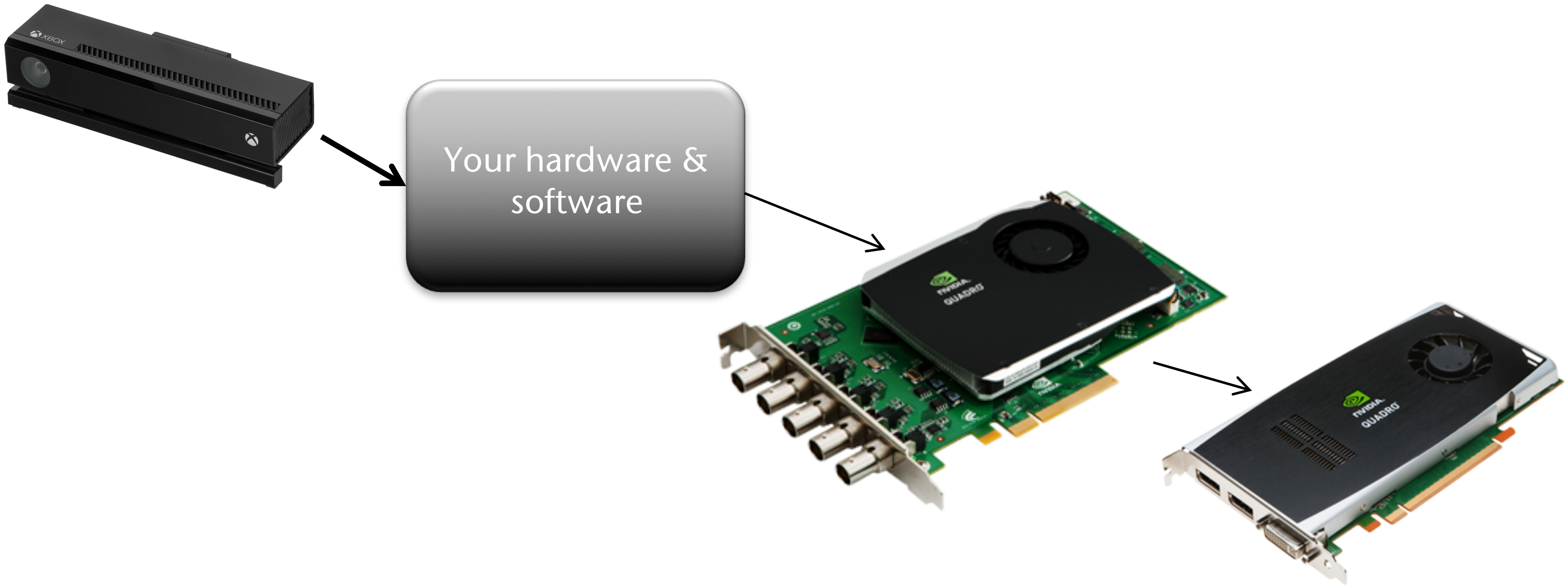




- Lead to new category of user interfaces: NUI = Natural User Interface
  - But this vision existed from the beginning of VR

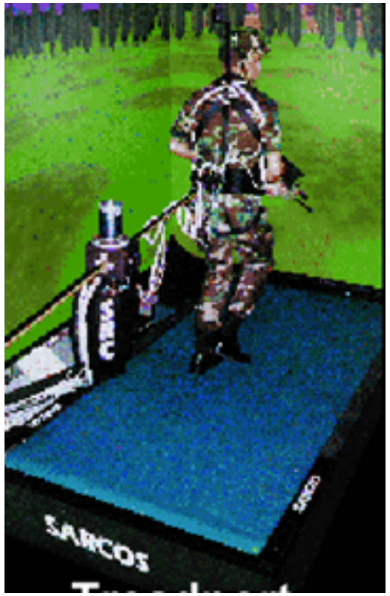
Bachelor / Master Thesis

- Build hardware in order to connect a Kinect directly to a GPU
- There is a card that can capture of SD-, HD-, 3G-SDI video footage, in real time, directly to the GPU memory
- So, "only" need to turn the Kinect output into an SD/HD stream

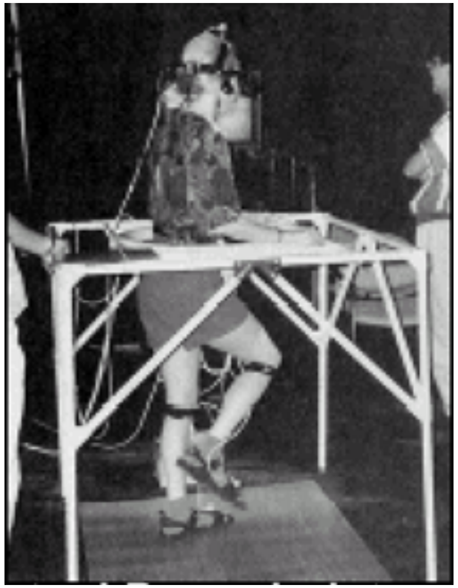


Locomotion Devices

Sarcos, Utah



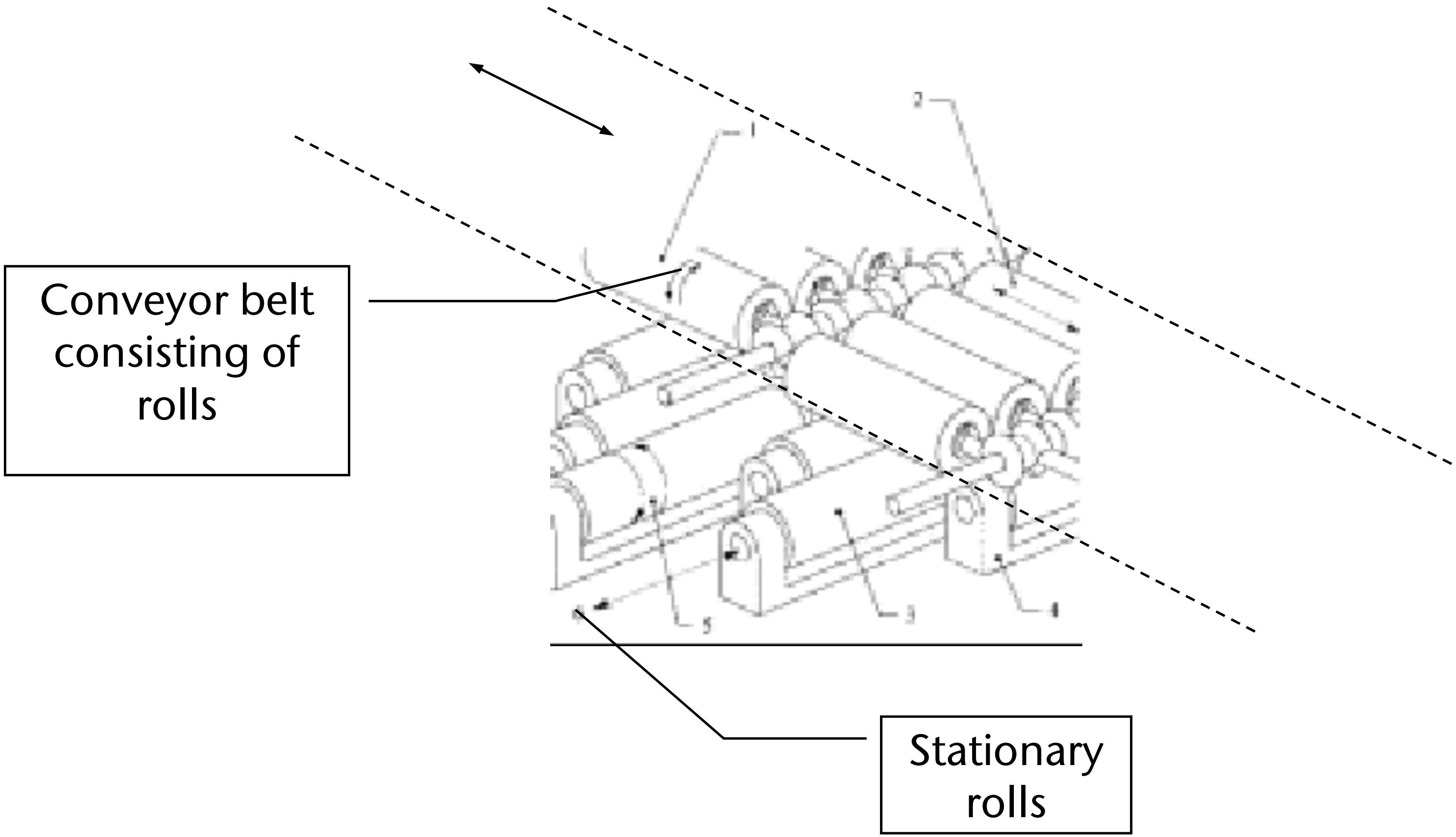
Sarcos



Uni Tsukuba, Japan

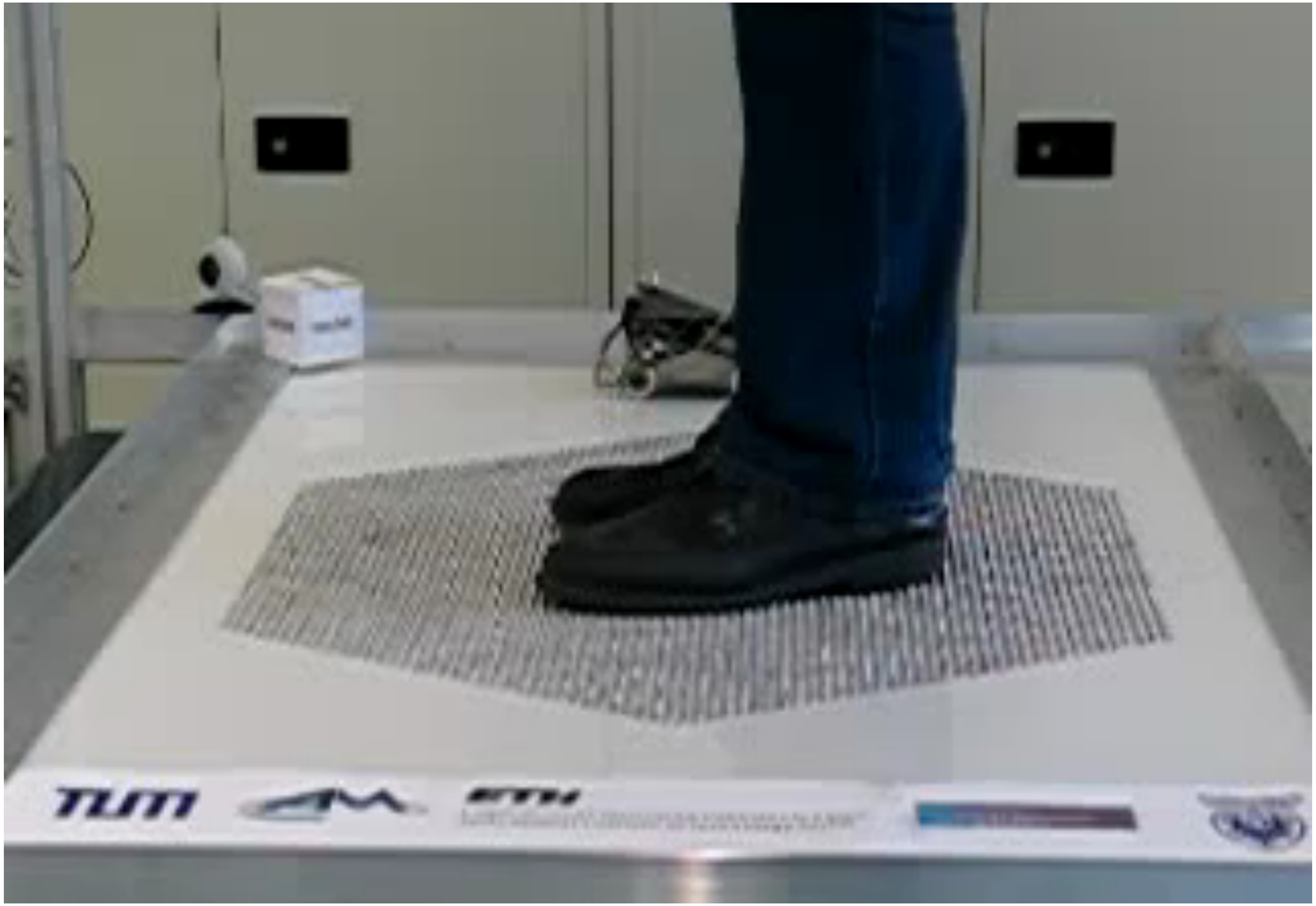
- *Omni-directional treadmill* (omni-direktionale Tretmühle)

Virtual Space Devices, Inc.

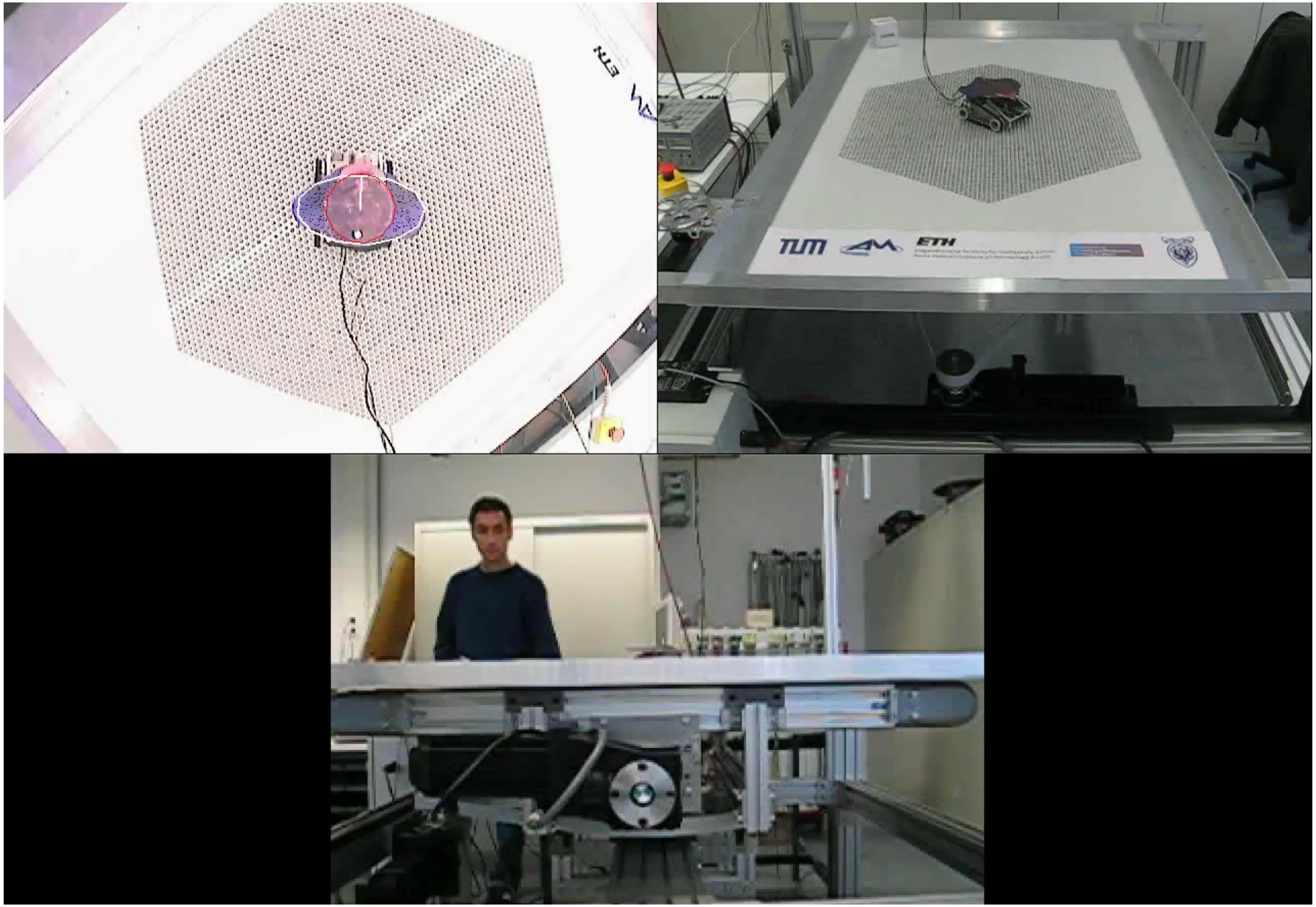

Conveyor belt consisting of rolls
Stationary rolls

Cyberwalk omni-directional treadmill, 2005-2008
TU München

Mode of operation of the Cyberwalk omni-directional treadmill



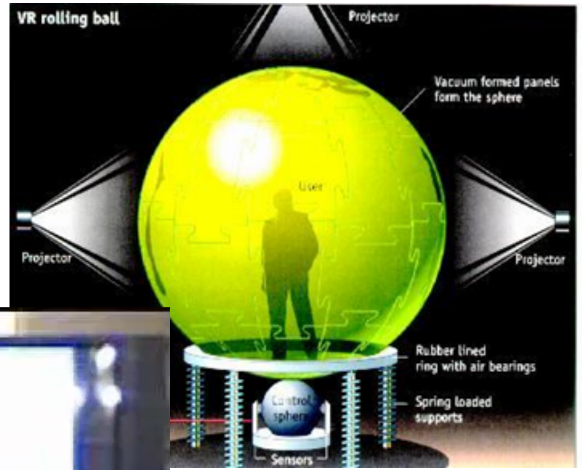
CyberCarpet
Martin Schwaiger, Dr. Thomas Thümmel, TU München



CyberCarpet's mode of operation

- Possible applications:
  - Research on behavior & cognition, brain research
  - Sports medicine
  - Training of soldiers and security staff
  - Fun parks (?)
  - Architecture:
    - Visualization and realistic exploration of architectural designs
    - Test of escape routes
  - Tests on humanoid robots

VirtuSphere



Frühe Studie

www.virtusphere.com

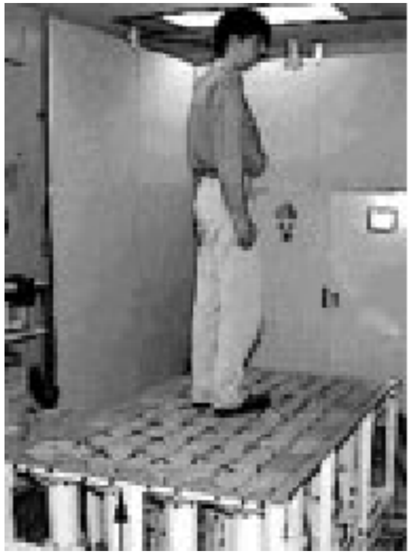

SPIEGEL ONLINE

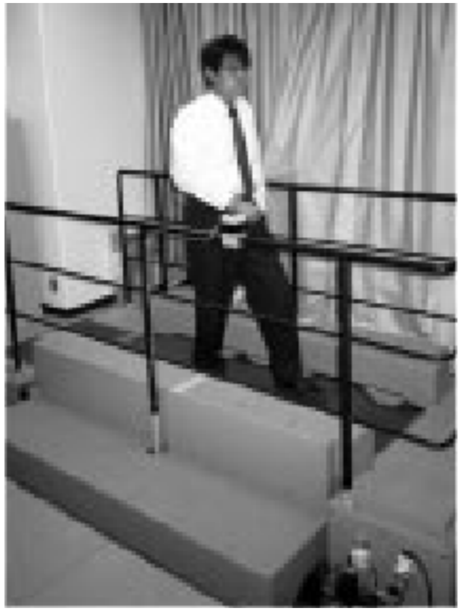
Virtuix: Omnidirectional treadmill for the home [2013]

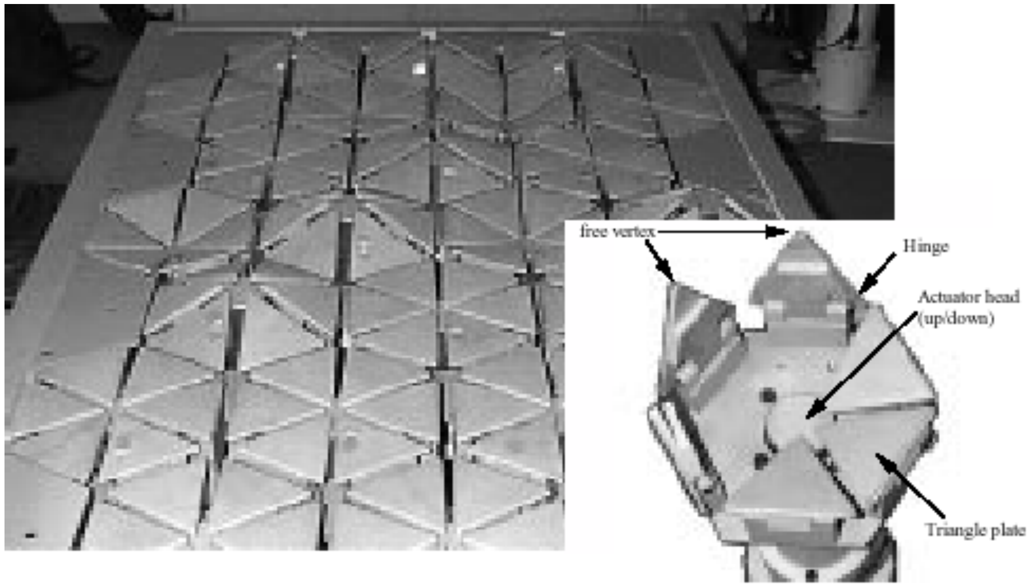
Cyberith Virtualizer

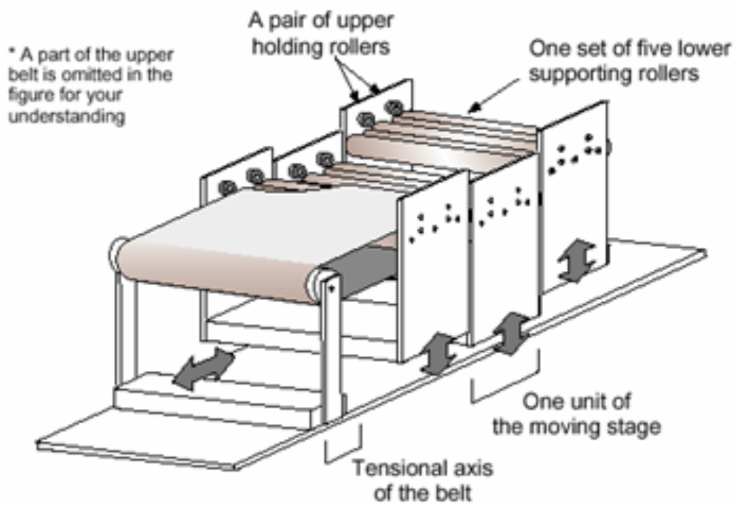
CirculaFloor, 2006

Simulation of Ground for Real Walking



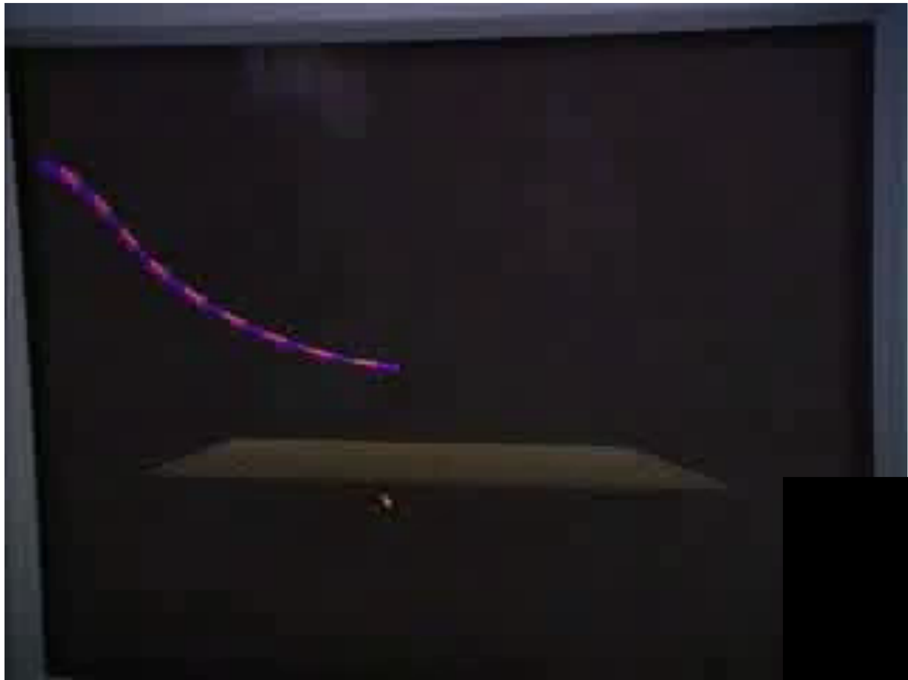




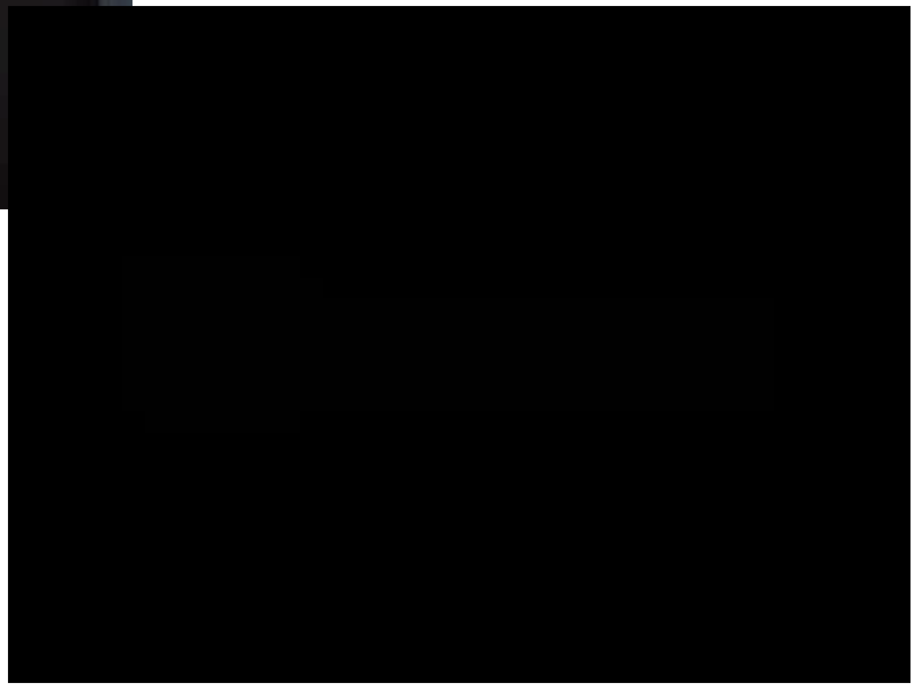


Other Locomotion Devices

Unconventional Input Devices



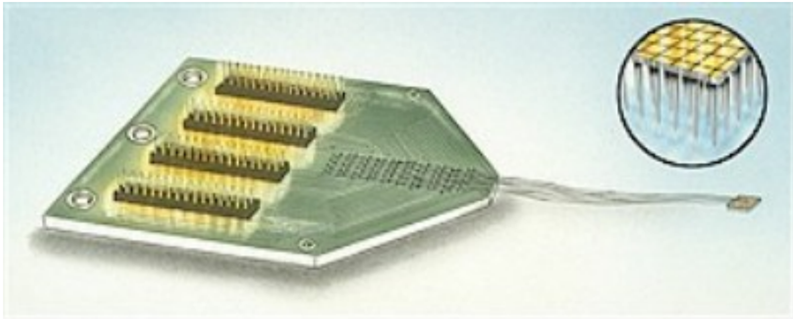
The Shape tape



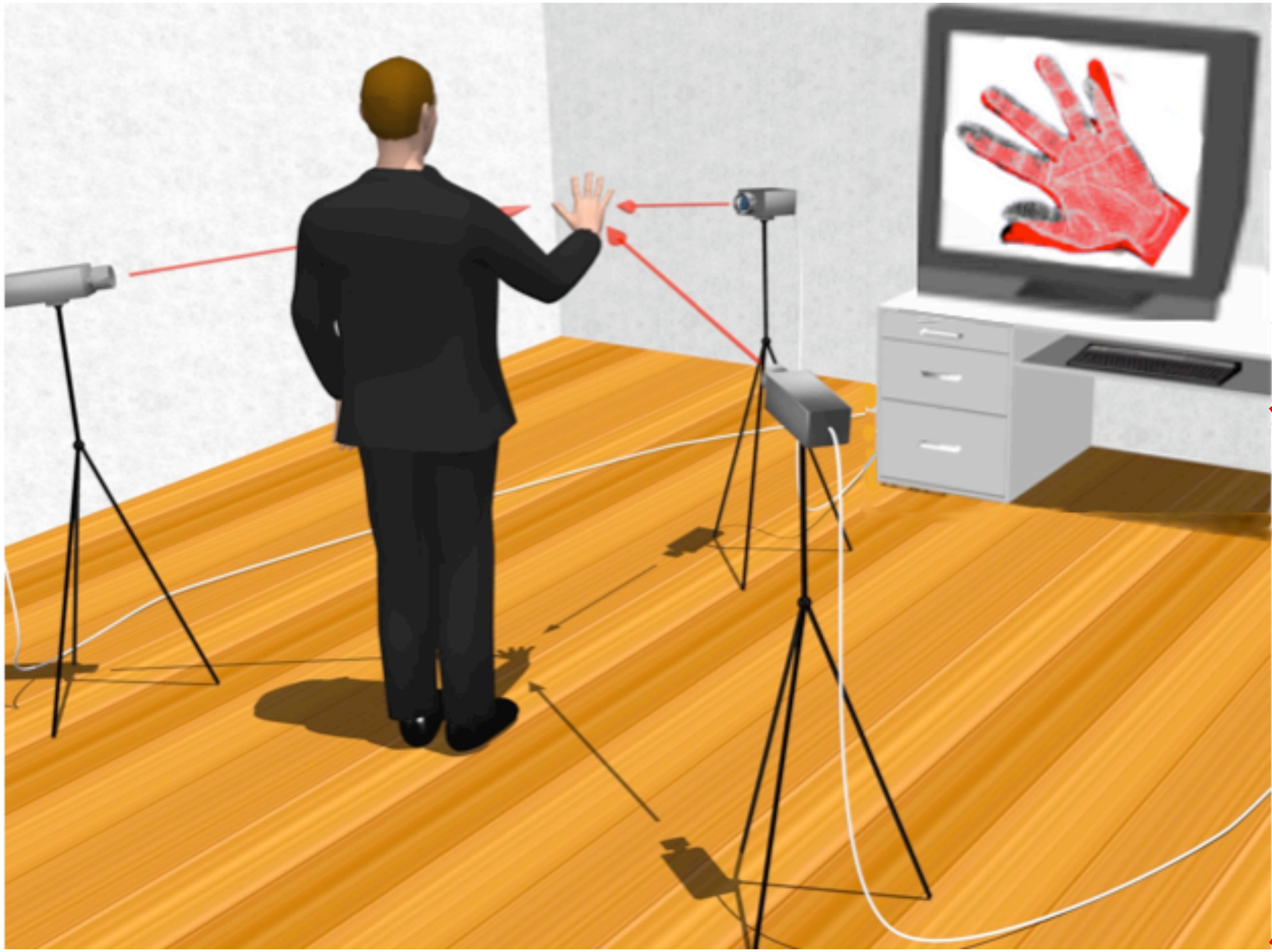
Virtual Keyboard

Brain Computer Interfaces

- Idea: control the machine by your brain only (no intermediary devices)
  - So far: EEG
  - SciFi: implant



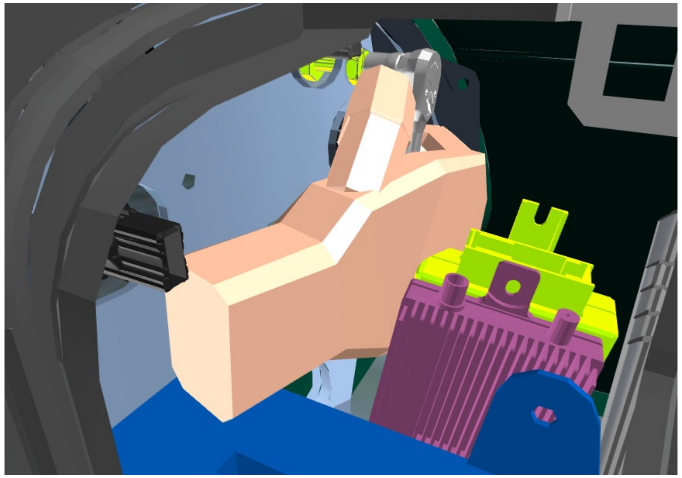
Ultimate Goal: *No* Input Devices

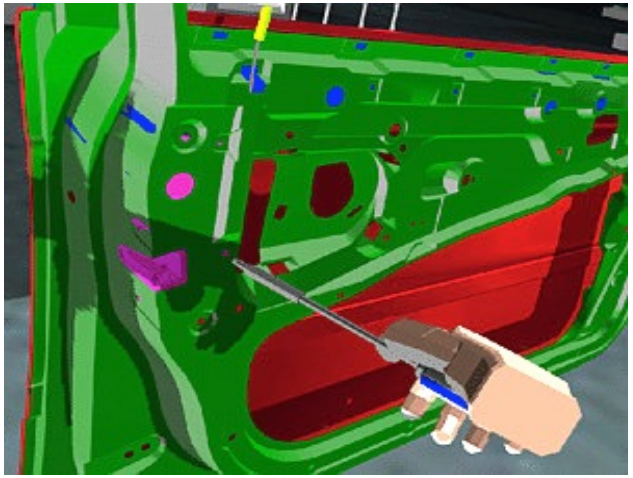


Goal: track human hand in real-time *without* any markers, gloves, etc.

Possible Applications

Virtual Assembly Simulation



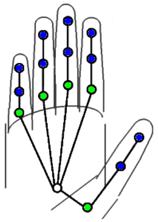


Intra-operative Interaction with Devices

Challenges

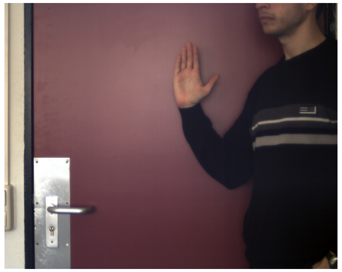
- 26 degrees of freedom

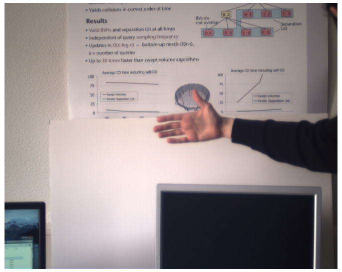
global



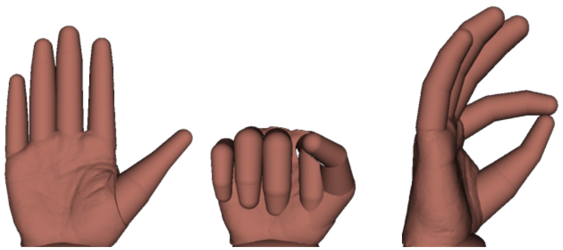
local

- Unconstrained background

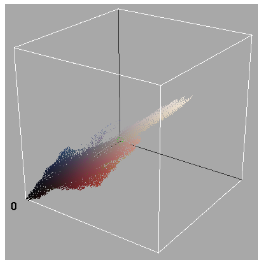




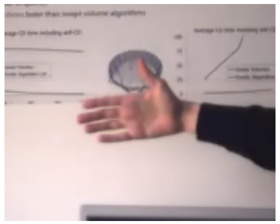
- Apparent shape variability and many self-occlusions



- Camera limitations and difficult lighting conditions



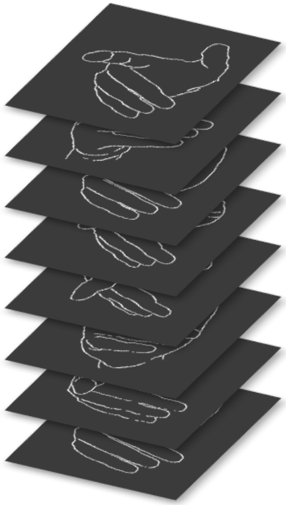
- Real-time tracking under all conditions



Model-Based Tracking by Detection Using Templates

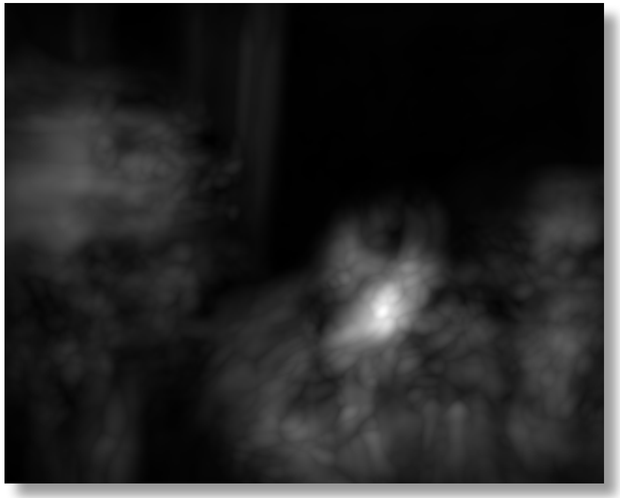
Query image

⊕



Set of template images

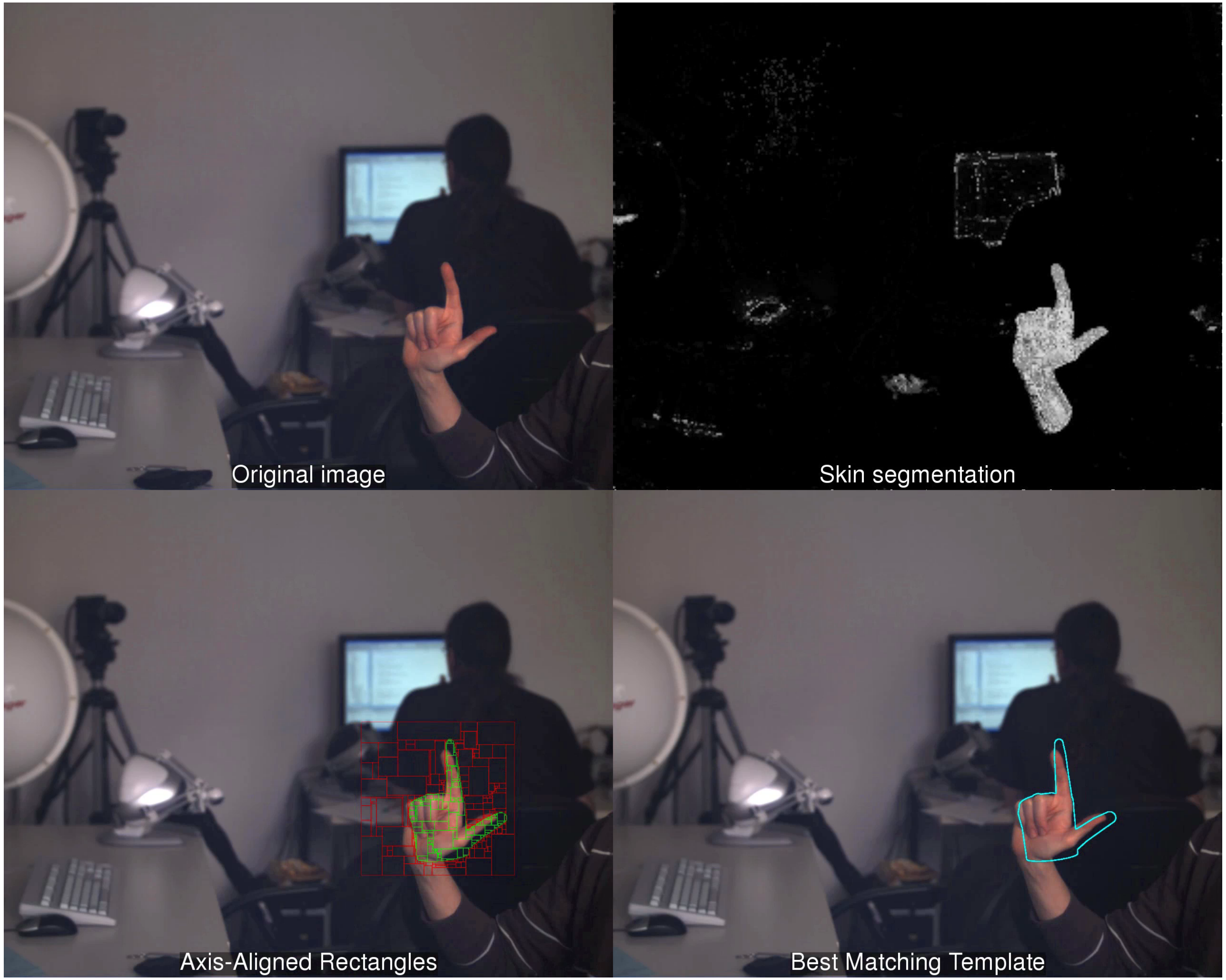
=



Confidence map = largest similarity value over all templates

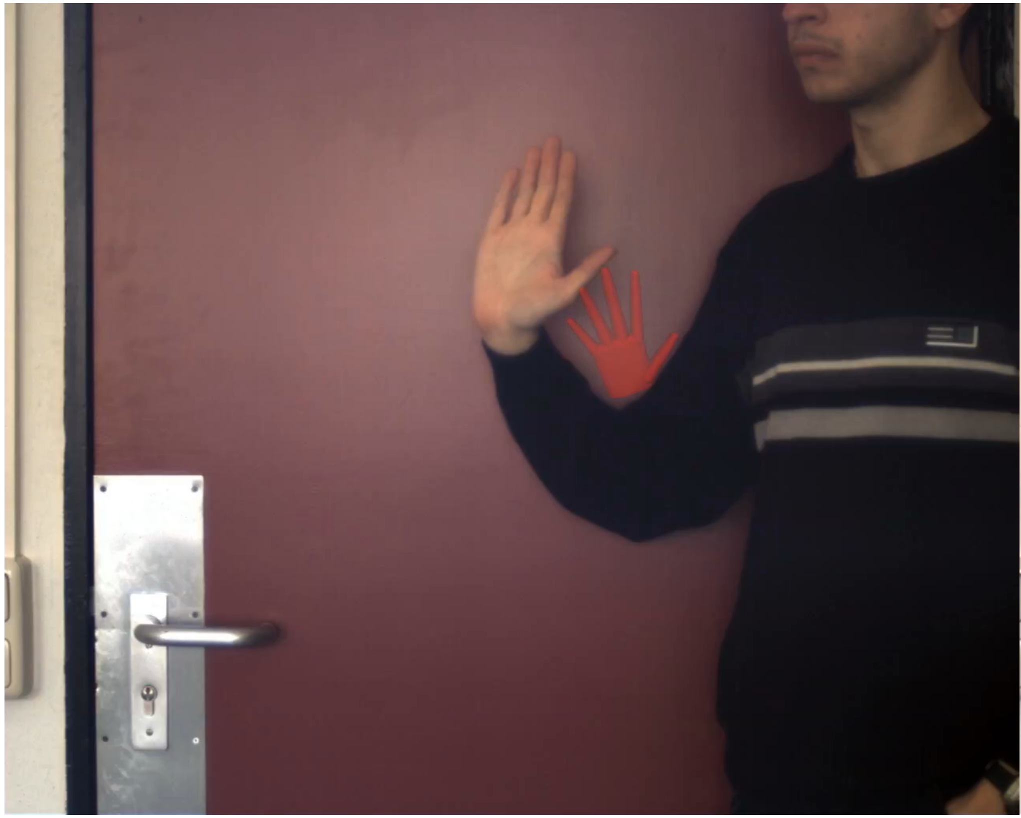
for each position in the query image:
  for each template:
    compute similarity value

Video


Original image
Skin segmentation
Axis-Aligned Rectangles
Best Matching Template

Results for Segmentation-Free Tracking

With segmentation, joint probability

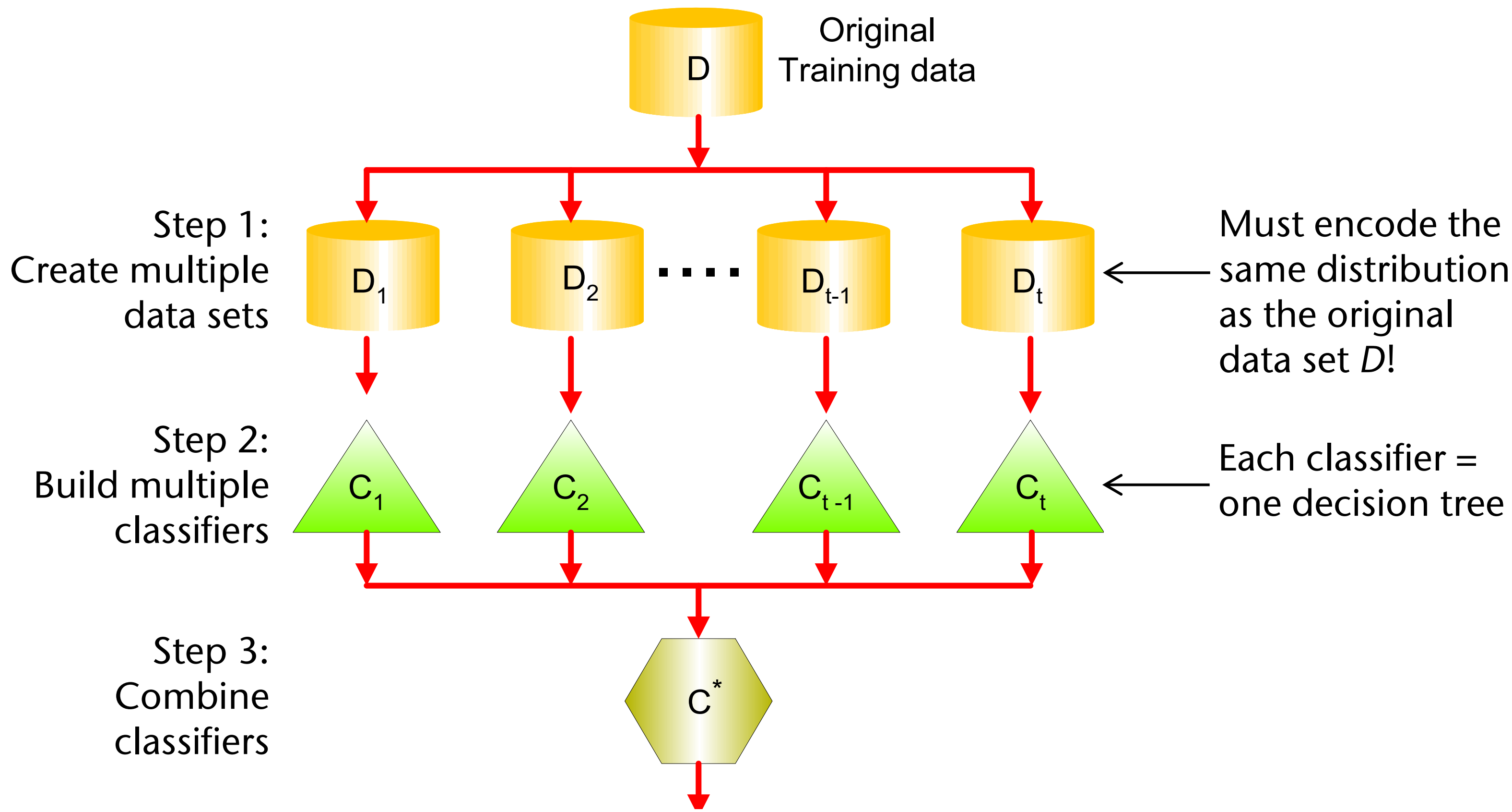


Without segmentation, color distribution

1200 templates, on a Geforce GTX 480 GPU → 20+ FPS

Work in Progress: Using Random Forests

- Ensemble of decision trees as weak classifiers
- Bagging of training data and random subspaces during training



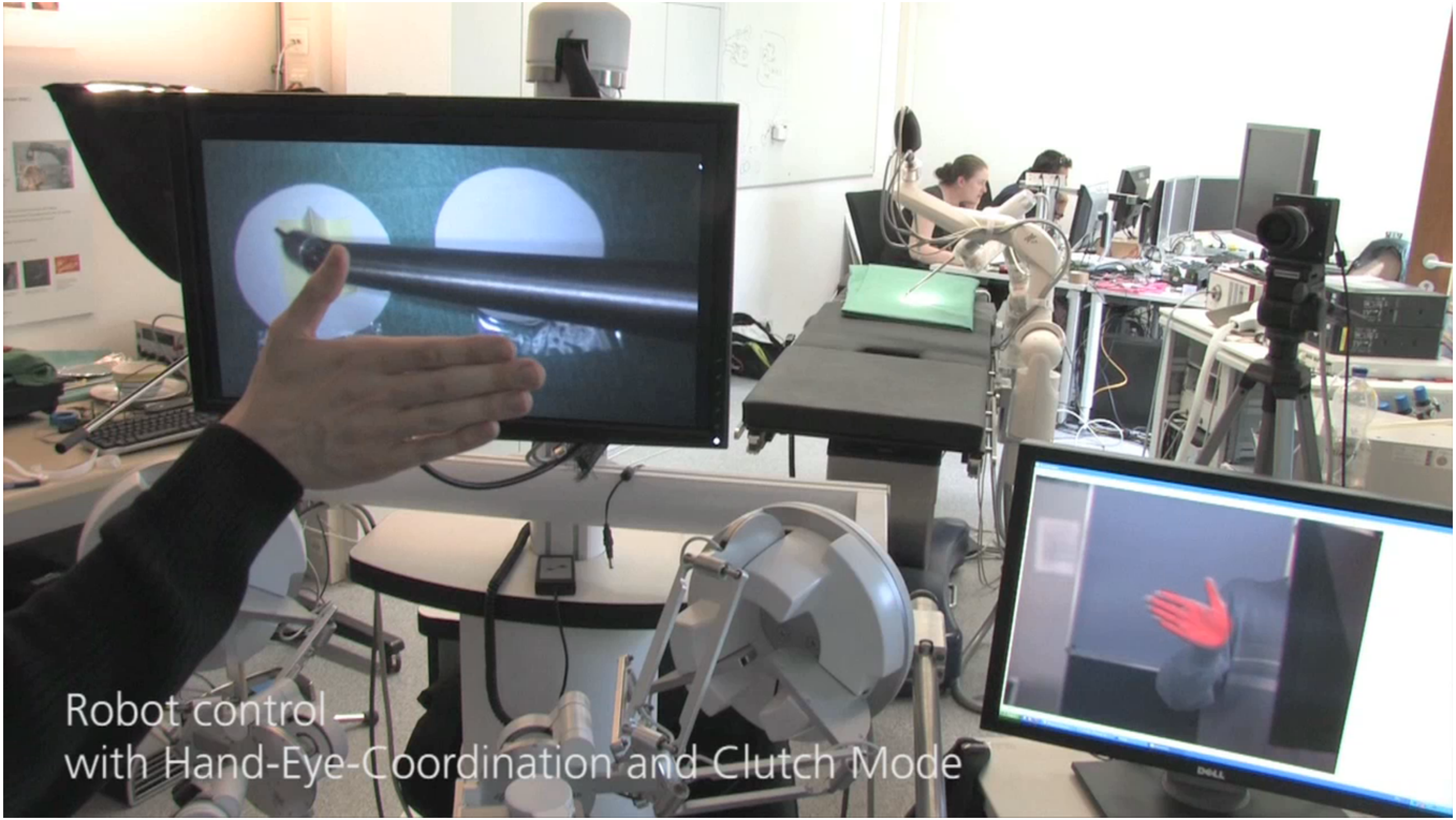
Preliminary Results

- 2.4ms per frame (CPU)
- Number of trees = 200
- Features considered at each split = 256
- No limit on tree depth
- $r \in [2, 8]\,\mathrm{mm}$ rectangle size
- Regression forest: ~5x longer training, no better match quality

Original depth image

Hand pose detected by random forest

A Possible Application: Control of Micro-Surgery Robots



Together with DLR, Institute for Mechatronics, Oberpfaffenhofen

	Proprioception	Consistent	Useable in lap or the side	Haptics capable	Unencumbered	Physical buttons	Hands free to interact with real world	General Purpose
Hand								
World-Grounded Devices	✓	✓		✓	✓	✓	✓	
Non-Tracked Hand-Held Controllers		✓	✓	✓		✓		
Bare Hands	✓				✓		✓	✓
Tracked Hand-Held Controllers	✓	✓	✓	✓		✓		✓
Hand Worn	✓	✓	✓	✓		✓	✓	✓
Non Hand								
Head Tracking	✓	✓					✓	✓
Eye Tracking							✓	
Microphone			✓		✓		✓	✓
Full-Body Tracking	✓	✓	✓	✓			✓	✓
Treadmills	✓	✓			✓		✓	