Interaction Metaphors for Collaborative 3D Environments

G. Zachmann

University of Bremen, Germany

zach@informatik.uni-bremen.de

Learntec, Karlsruhe, Germany, 2014

Virtual/3D Environments (VEs) at Home

- In the old days:

15,000 – 60,000

~500,000

~20,000

- Today:

~500

~5,000

~200

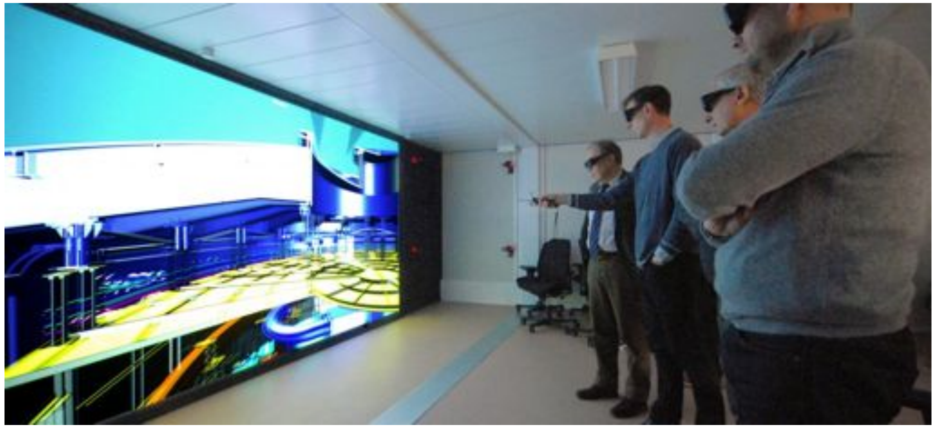
Collaborative Virtual Environments

- Definitions:
  - CVE = *shared* virtual environment that contains virtual representations of real objects/abstract data *and* users (avatars)
  - CVE = VE + CSCW
- Classification by kind of participants: same vs. different domain of expertise

Massively multiplayer online game

Moonbase Alpha, NASA

Simultaneous engineering teams (SET)



Walk-through, ITER

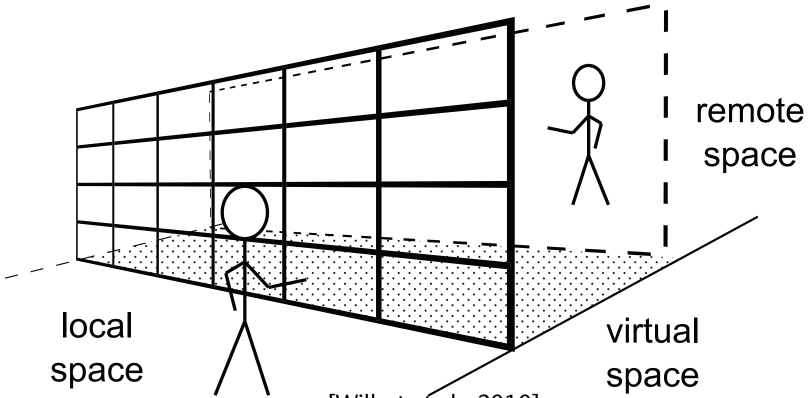
Classification by Place and/or Time

	Synchronous (same time)	*Asynchronous (different times)*
Co-located (same place)	**Face-to-face** ◦ SETs ◦ Shared wall displays (powerwall, workbench, ...) ◦ One set of input devices for the "driver"	**Continuous task** ◦ No collaborative VEs yet ◦ Conventional "war rooms", post-it communication ◦ Large public displays(?) ◦ Touchless input(?)
Remote (different places)	**Remote collaboration** ◦ Video conference ◦ Simultaneous interaction with shared virtual objects ◦ Second life et al., MMOGs	**Communication + Coord.** ◦ Wiki's (Wikipedia) ◦ Email ◦ Version control (software, ...) ◦ (Second life et al.)

Tele-Immersion for Remote Collaboration

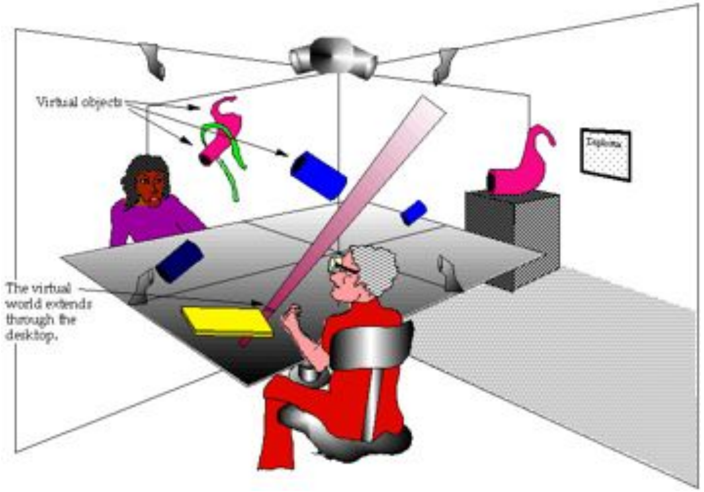
- Lots of commercial products for "telepresence":
  - But are they *immersive*?
  - Do they create the feeling of *presence*?
- Goal: a truly shared space
- Metaphor: *Extended Window*
  - Display: large video wall
  - Head-tracked users ⟶ center of projection for remote environment
  - Creates illusion of looking through a "window" into the collaborator's physical space

©IVCE



[Willert et al., 2010]

- Benefits:
  - Natural scale
  - The virtual space "between" the two collaborators can be populated with virtual objects or information visualizations
  - Natural & intuitive navigation
  - Motion parallax ⟶ increased presence
  - Gaze awareness: each user sees where other user is looking at; users can establish eye contact

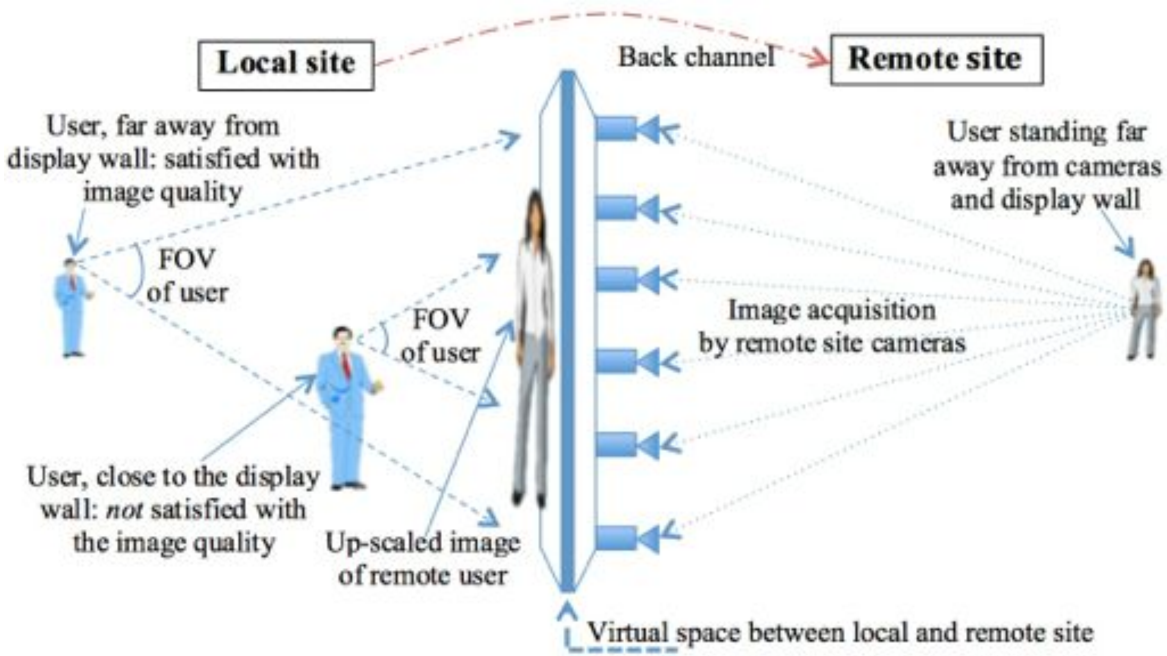


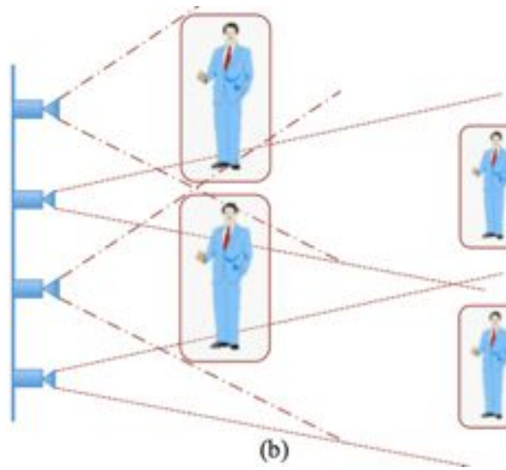
[Kurillo et al., CVPRW, 2010]

- Problem: need a camera image of *remote* environment/user from viewpoint of *local* user
- Solution: micro-lens camera array embedded in video wall

[Willert, Ohl, Lehmann, Staadt, 2010]

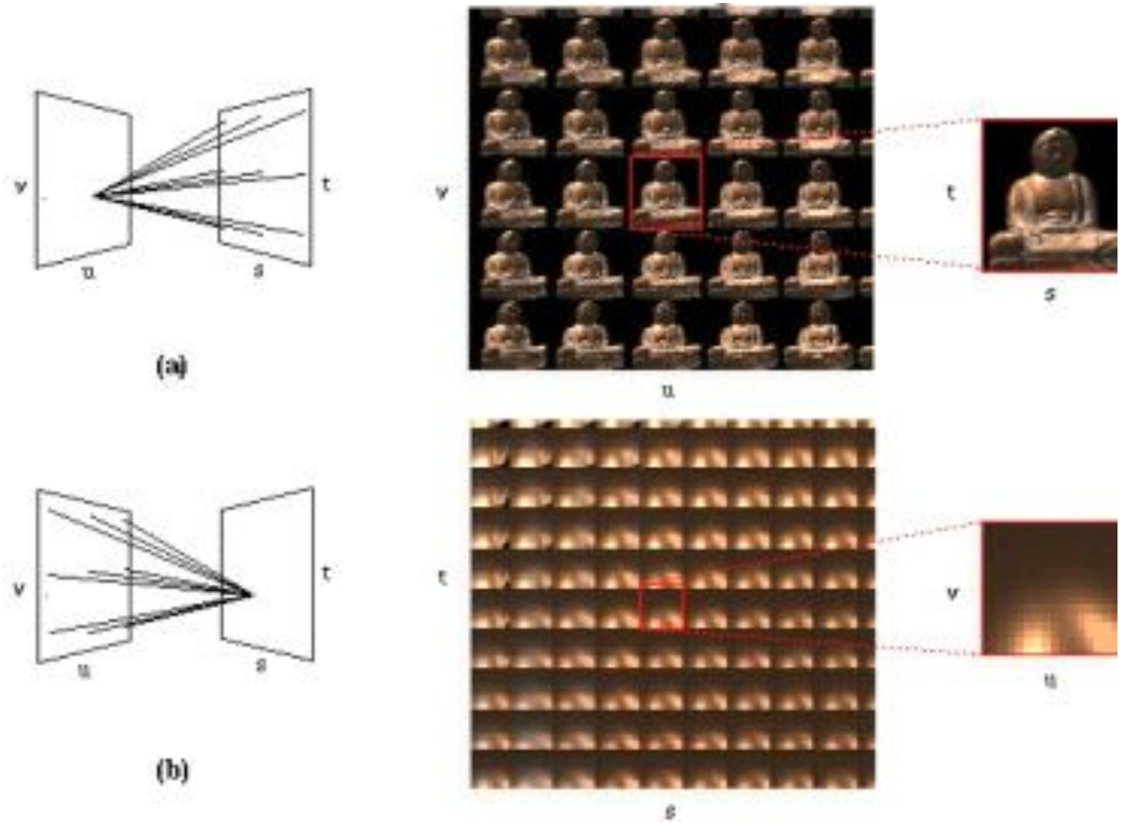
- Problem: insufficient resolution, if local user approaches local display
- Solution: super-resolution images by weighted camera fusion & cameras with different field-of-views





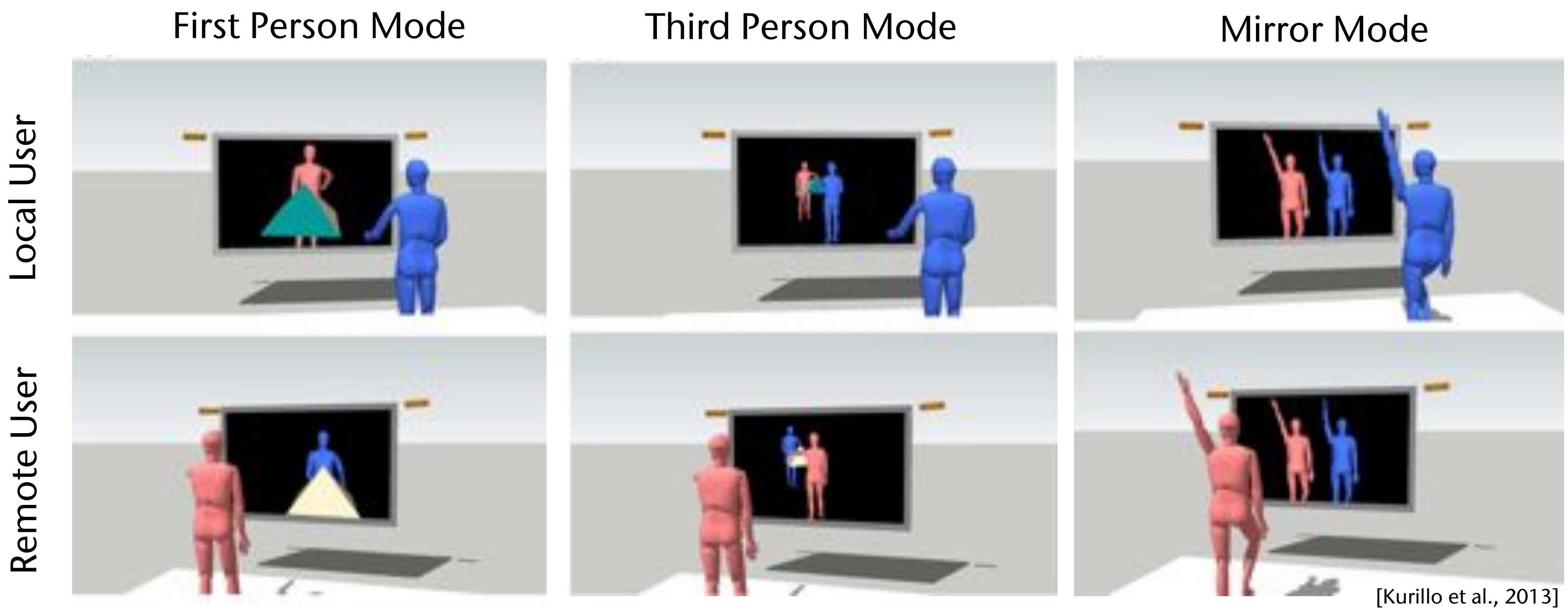
[Islam et al., 2013]

- Problem: camera array outputs essentially a light field ⟶ huge amount of data
- Solutions:
  - a) Transmit local user's viewpoint to the remote site ⟶ extract parts of remote camera images needed to assemble image for local user



  - b) Compress light field (neighboring camera images differ only slightly)

Other Interaction Modes for 3D Tele-Immersion



[Kurillo et al., 2013]

Physically correct *Extended Window* metaphor; each user sees the other and the virtual objects at the physically correct position; virtual viewpoint is always coincident with real viewpoint

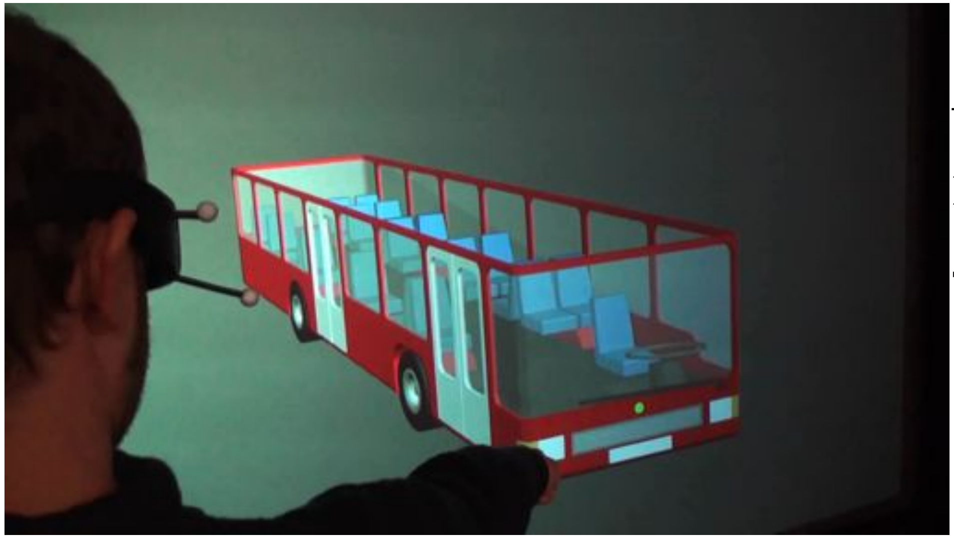
Each user looks over their virtual avatar's shoulder; virtual viewpoint is usually fixed, or can be controlled using some input device; can be useful if display is mono-scopic

Camera image from self is superimposed in a mirrored fashion on remote video stream; could be useful for physical instruction; problems: correct handling of mutual occlusion

A Problem with Co-Located CVEs

- Assume this situation: one stereo display wall, several users in front of it
- Problem with a single-tracked projection (stereo or mono): only the viewpoint of the *tracked* user is correct, only she will see a correct image!
- Example: communication via pointing fails

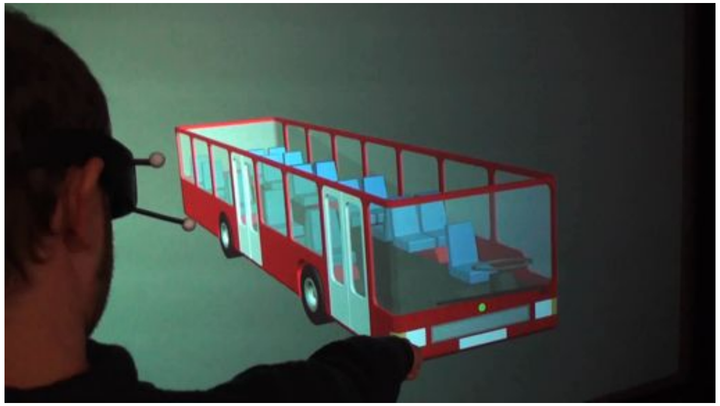
Image's perspective is correct for the user



Image's perspective is correct for the (real) camera

[Kulik et al., ACM Trans. Graph. 30, 6, 2011]

- With perspectively *correct* projections for *all co-located* users, the shared 3D space will become coherent for all users
- Consequence: direct communication (including *pointing!*) in co-located CVEs is possible

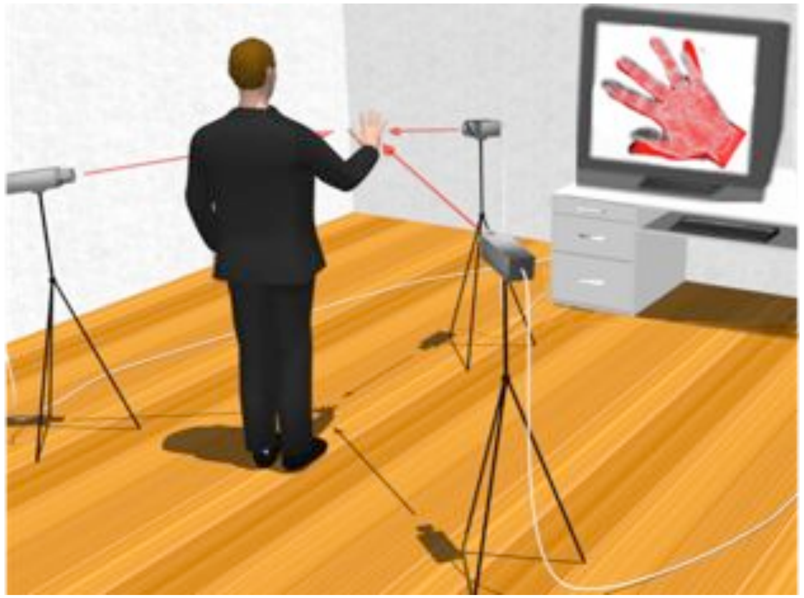


[Kulik et al., ACM Trans. Graph. 30, 6, 2011]

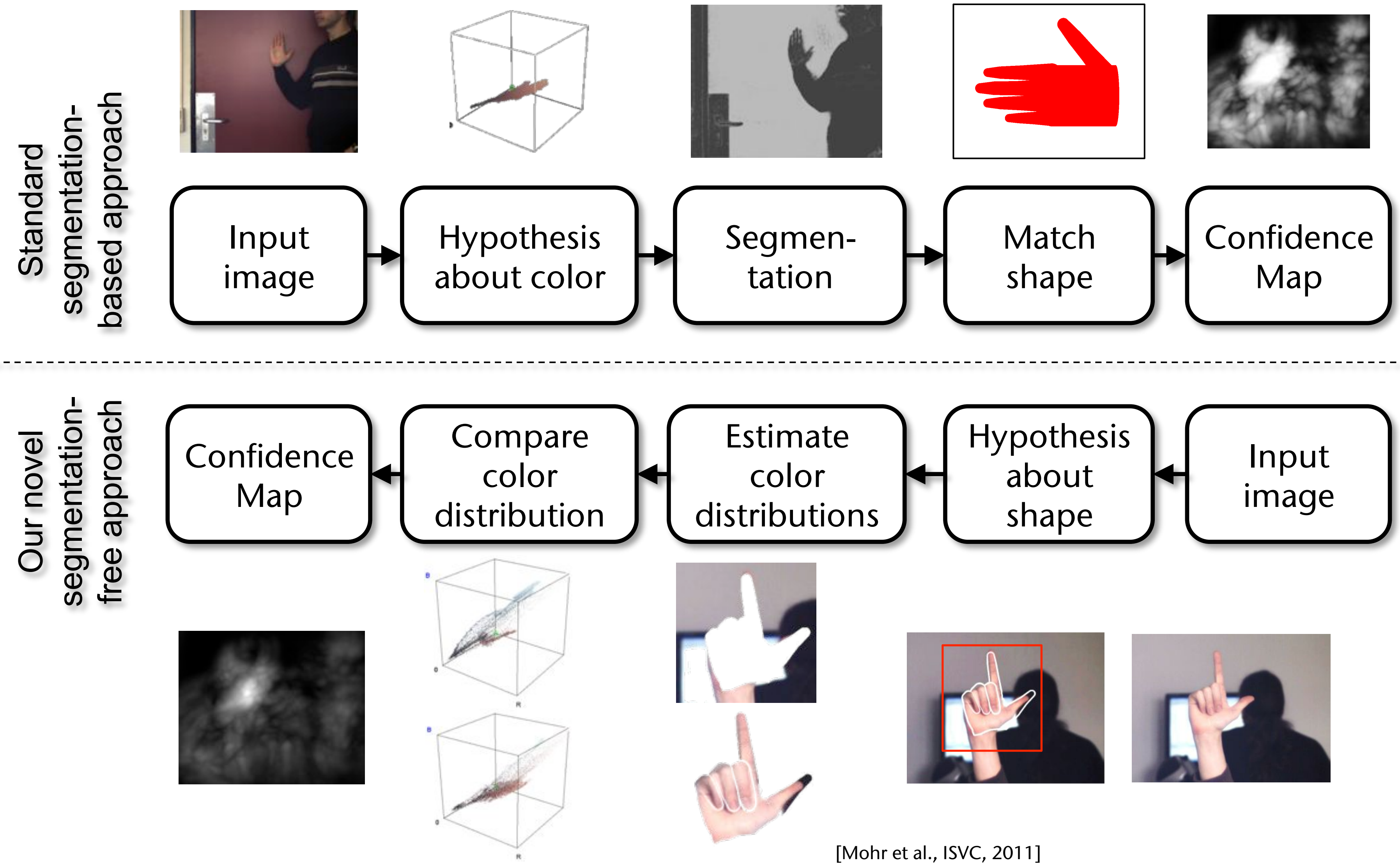
Kitamura et al. 2001

Camera-Based Articulated Object Tracking

- Goal: "Kinect for the human hand"
  - Markerless tracking of human hand with cameras
  - Challenges: high-dimensional configuration space (20+6 DOFs), Real-time, large working volume, lots of self-occlusions



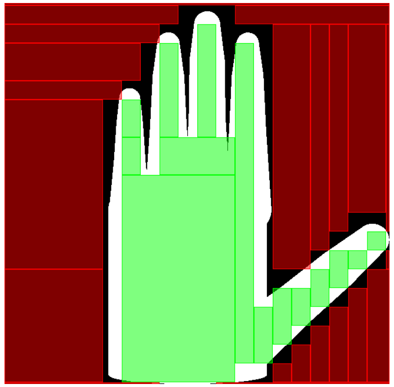
A Segmentation-Free Approach

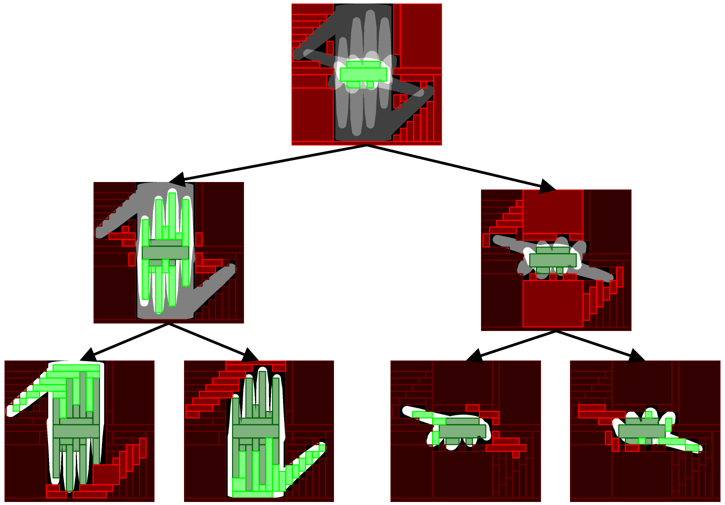


[Mohr et al., ISVC, 2011]

Fast Area-Based Template Matching

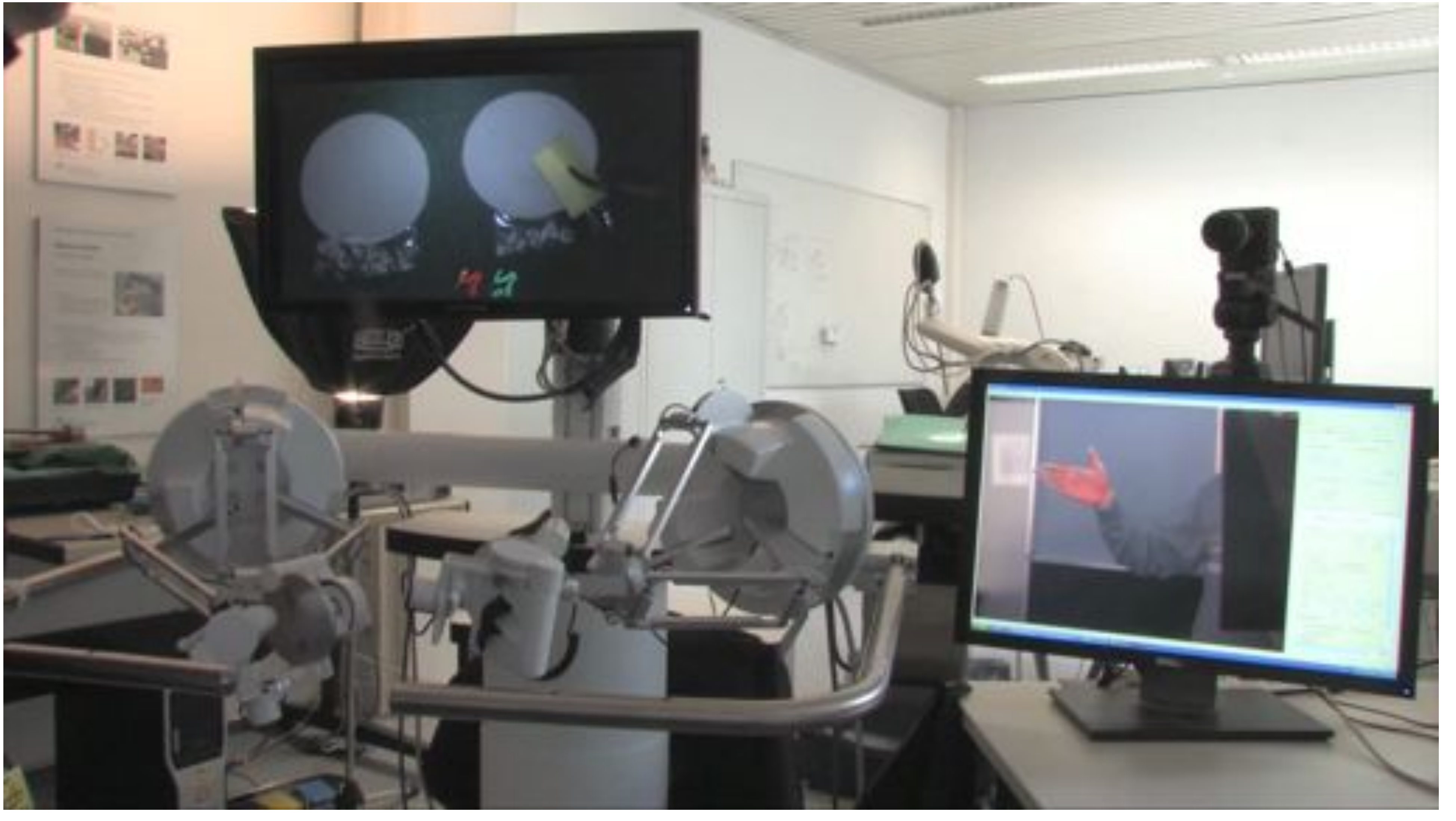
- Novel representation for templates: rectangle coverings
- Advantages:
  - Matching time no longer depends on image or template resolution
  - Speedup = 10-25 x
  - Easy to turn into hierarchical matching algorithm → complexity = $O(\log n)$ for n templates!





[Mohr et al., BMVC, 2010]

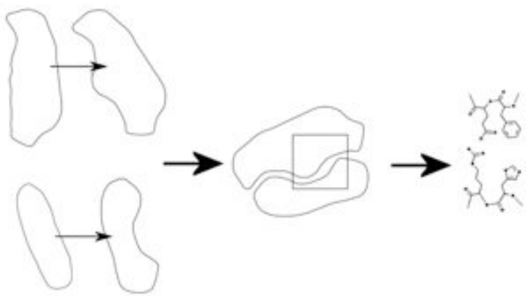
One Possible Application: Touch-less Control of Robots

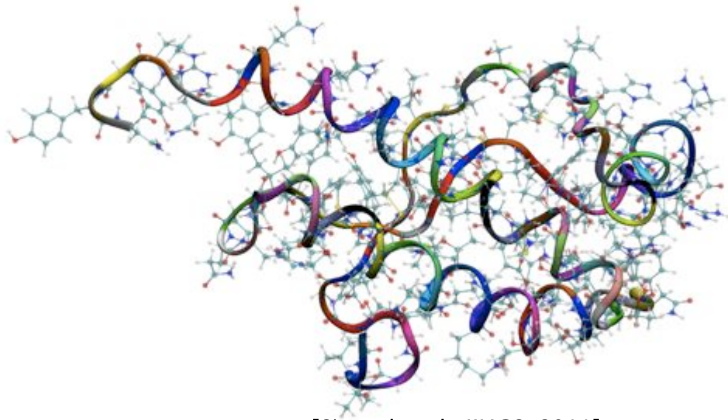


With DLR, Oberpfaffenhofen: touch-less hand-based control of the surgery robot MiroSurge

Collaborative 3D Search and Selection

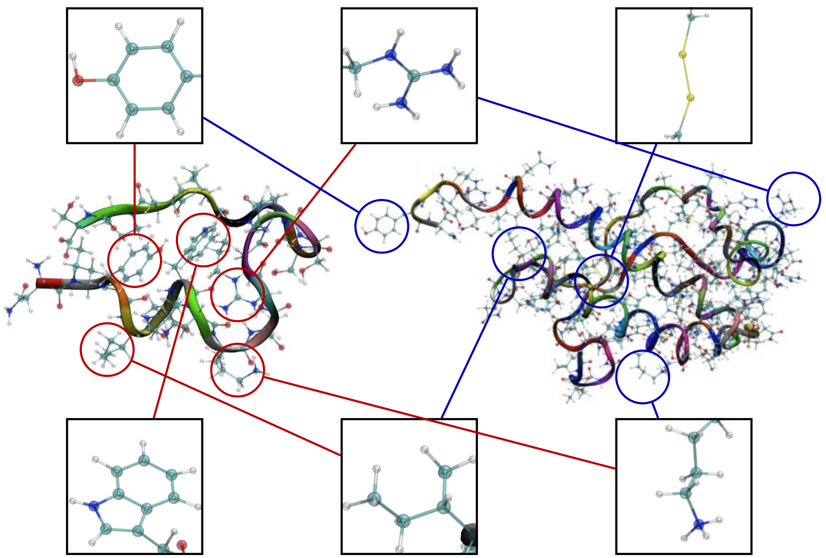
- Studied in a particularly complex VE: molecule analysis & design
- Molecular docking is done in 3 stages:





[Simard et al., IJHCS, 2011]

- Very frequent task in all 3 stages:
  - *Finding* a target (structure or residue)
  - *Grabbing* the target (using a virtual handle)



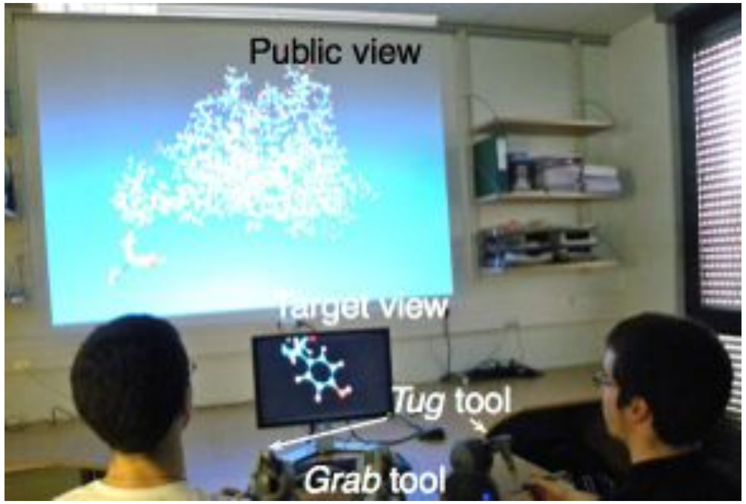
Conditions of the Experiment

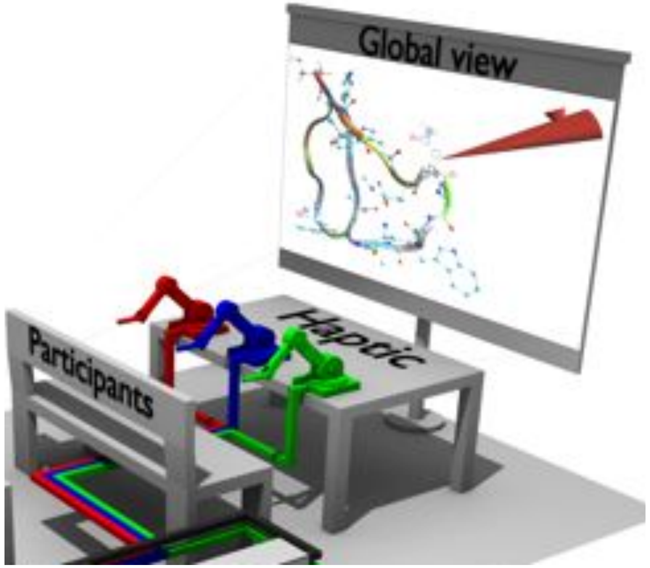
A. Two co-located, synchronously collaborating subjects:

- Left hand of one subject controls orientation of molecule
- Right hands of both users can point and tug at molecule parts
- Requires good mutual understanding of partner's workspace & actions

B. One subject:

- Left hand controls orientation of molecule ("scene in hand" metaphor)
- Right hand moves occluding parts of molecule away





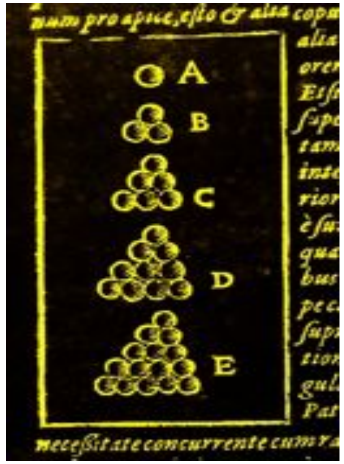
Results

- Tasks with low complexity do not require collaboration
  - Collaboration does not speed up task completion time
- Collaboration (2 subjects) can speed up task completion time by up to a factor 2
  - Reason? (Social facilitation [Triplett, 1898] and/or synergy)?
- *Average affinity* (e.g., student-supervisor) is better than high affinity
- The best strategy here: both should work on neighboring regions
  - "Best" in the sense of 3 criteria: completion time, effectiveness of coordination, amount of verbal communication

Collision Detection as Enabling Technology

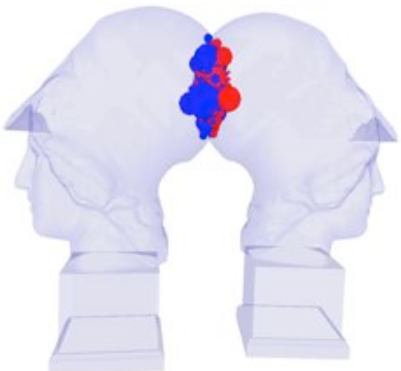
Sphere Packings

- Have a long history …

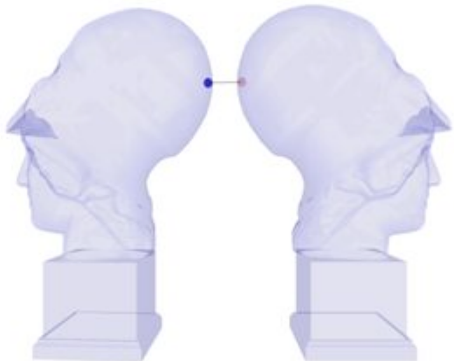


Johannes Kepler
(1571 – 1630)

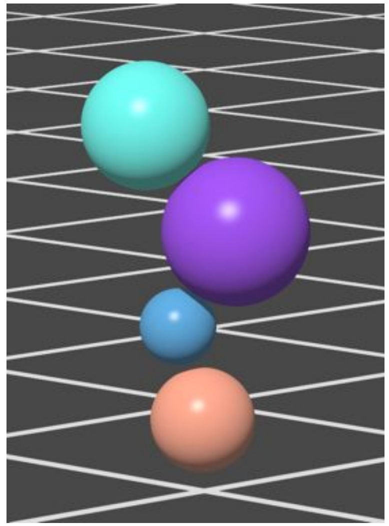
- Collision detection
  based on sphere packings:



~10 microseconds



1 microsecond



[Weller et al., Siggraph Asia, 2010]

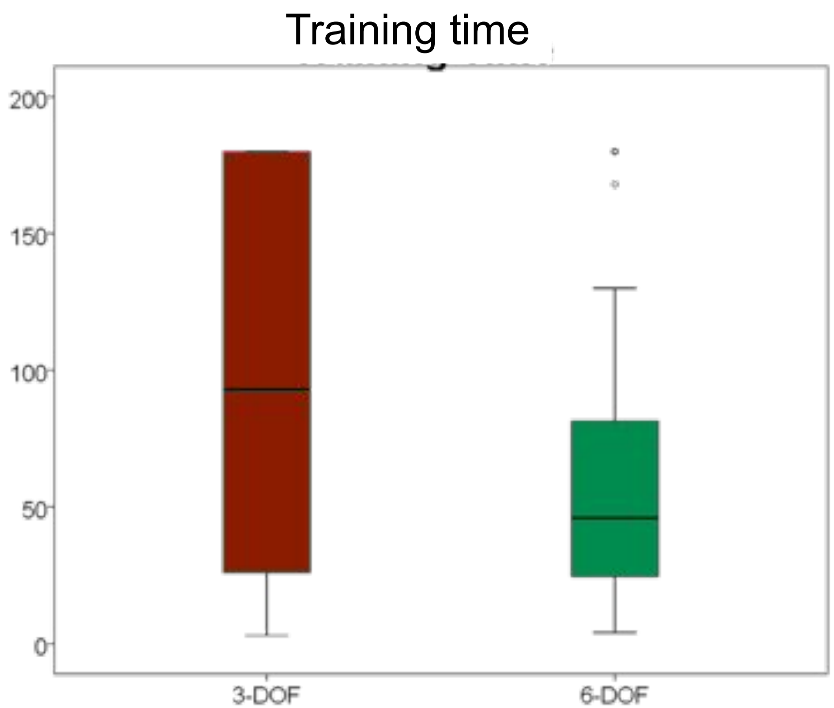
Application: Collaborative Haptic Workspace

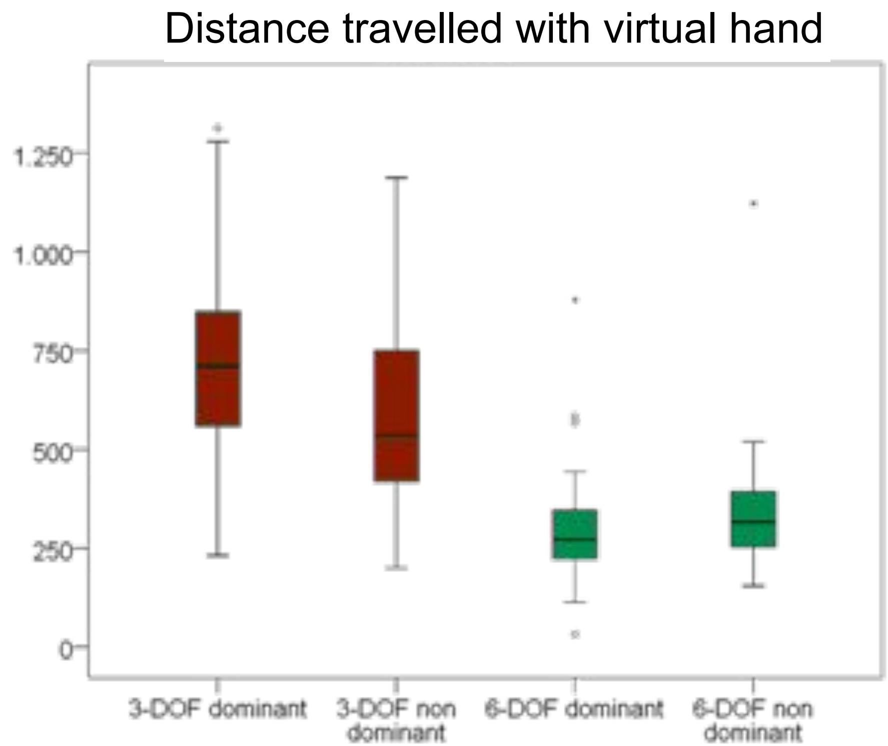
12 moving objects ; 3.5M triangles ; 1 kHz simulation rate ; intersection volume ≈ 1-3 msec

User Study: 3 DOFs vs 6 DOFs

- Conventional wisdom in VR: restrict number of DOFs for precise manipulation
  - E.g. [Veit, Capobianco, Bechmann, VRST 2010]
- Research questions:
  - Is that true when force-feedback is given?
  - If not, is the benefit worth the extra dollars?

Results: User Performance

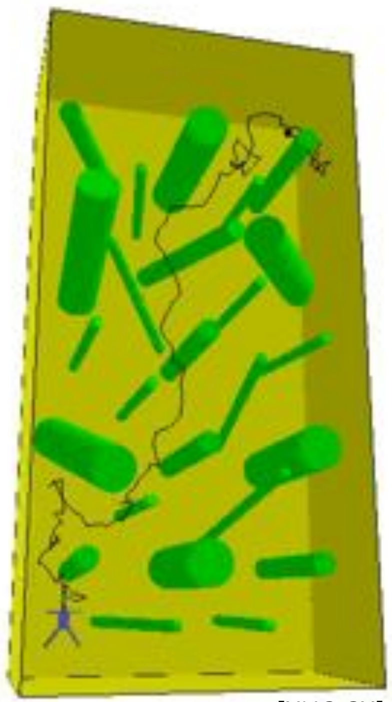




[Weller et al., ICCE, 2011]

The Piano Movers' Problem

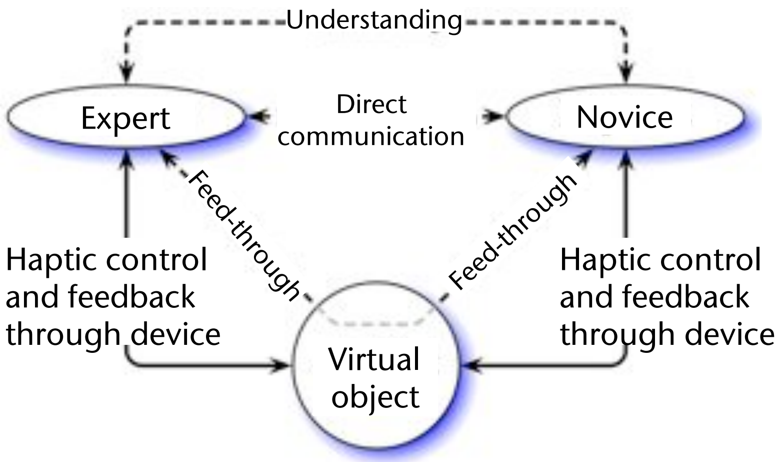
- Path planning problem: find a path (including rotations) for moving an object (piano) from A to B (without moving/hitting anything else)
- Application: assembly simulation (and many others)
- Question: does collaboration in a virtual environment help?



[UNC-CH]

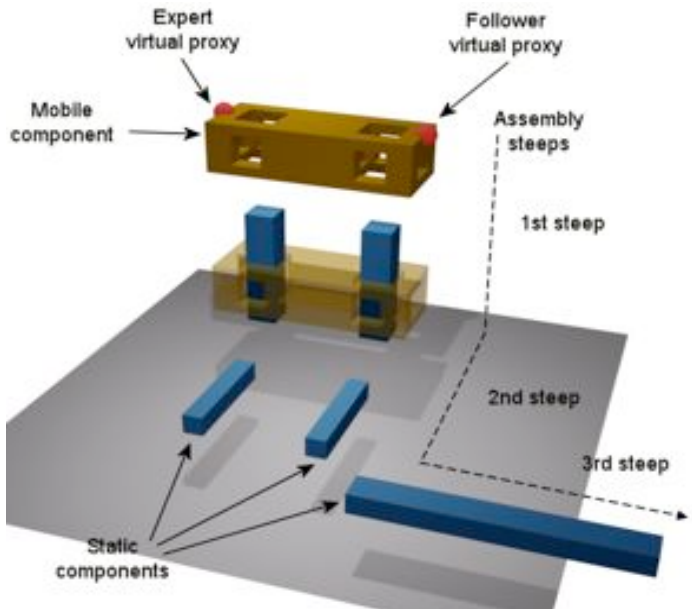
Haptic Communication in Collaborations

- The task: a collaborative assembly task
- Experiment setup:
  - Two users, one expert, one novice
  - Each with one 3-DOF haptic device (Phantom)
  - Only oral (direct) and haptic (indirect, feed-through) communication



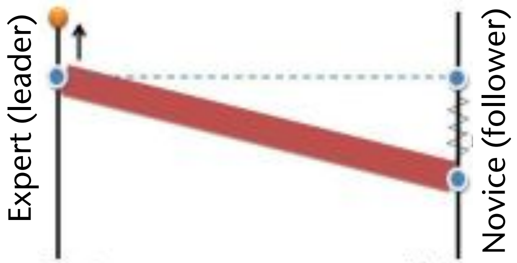
[Simard et al., Virtual Reality, Springer, 2011

- Assembly task:
  - 3 distinct legs of assembly path
  - Expert knows exact movements
  - Novice is guided by haptic feed-through communication metaphors from expert



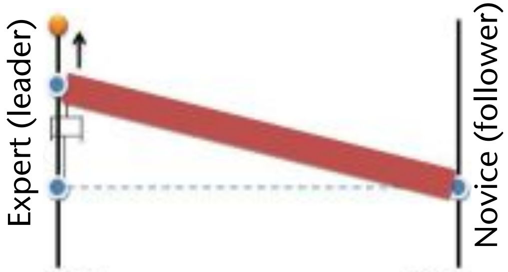
- Haptic *feed-through metaphors*:

Spring metaphor



Force on follower ~ distance |expert – follower|. Force pulls novice in the right direction

Viscosity metaphor



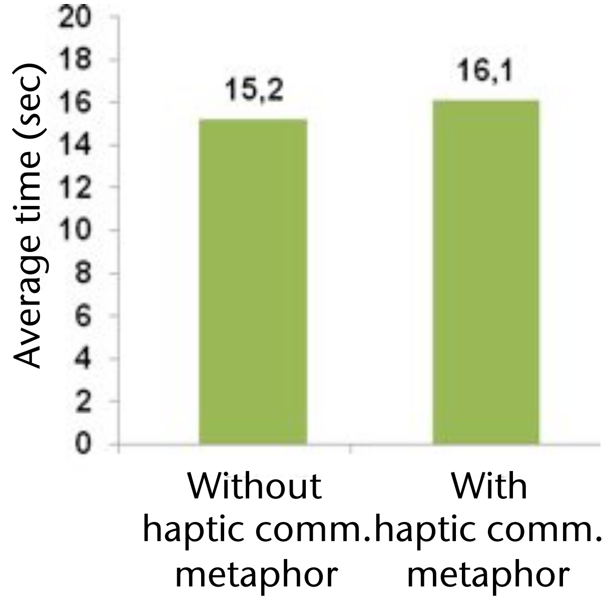
Force on expert ~ negative relative velocity. Force drags expert back, if follower lags behind

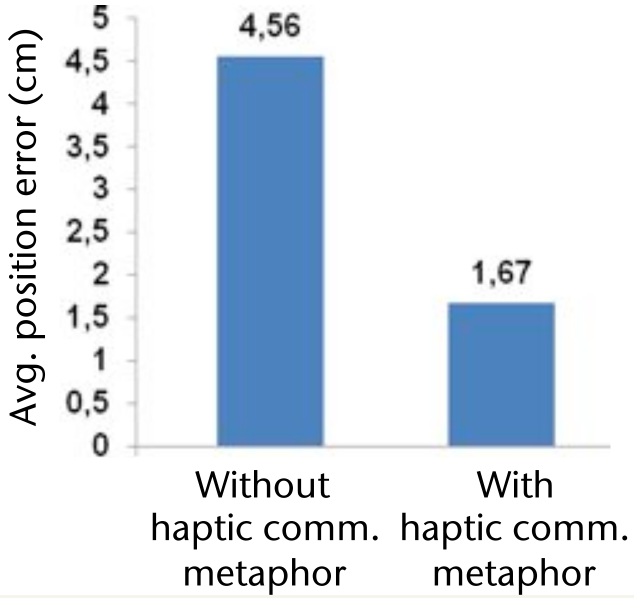
Vibration metaphor

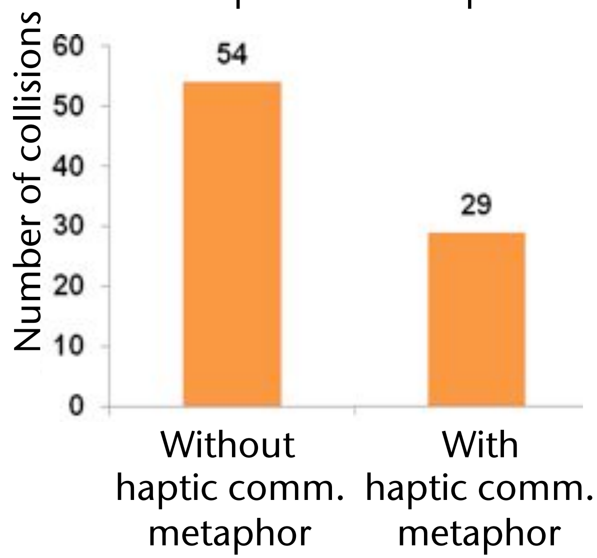
No force, followers gets vibration signal whenever leader changes direction

Results

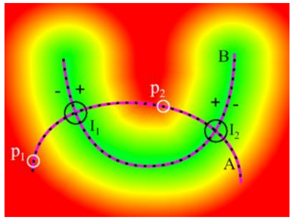
- Bad news: task completion time does not change significantly
- Good news: collaboration improves manual precision

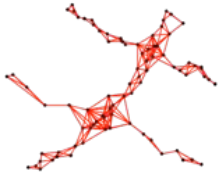
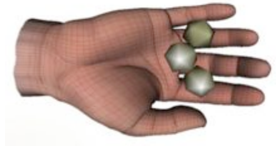


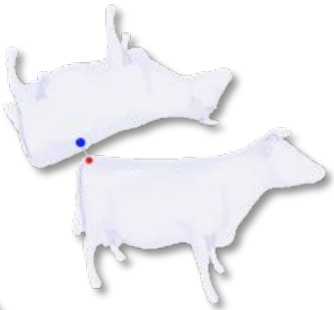




Thank You!

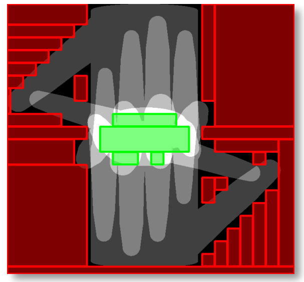






G. Zachmann, zach@cs.uni-bremen.de , http://cgvr.cs.uni-bremen.de/