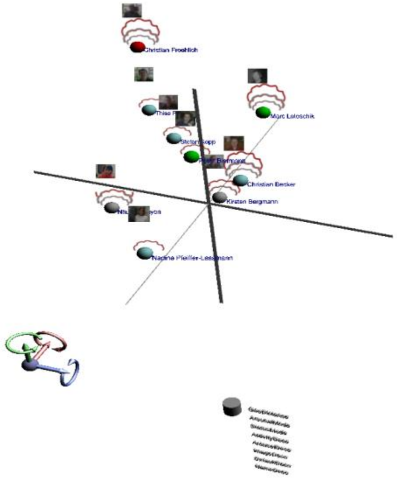
3D FIXATIONS IN REAL AND VIRTUAL SCENARIOS

Thies Pfeiffer, A.I. & V.R. Group, Bielefeld University

Background

- multimodal human computer interaction
- situated natural communication (gaze, gesture, speech)
- natural interaction with dense information displays

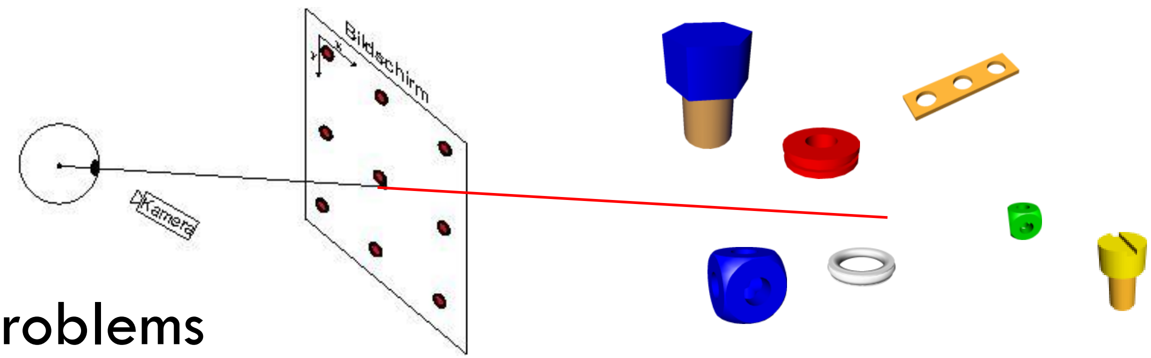


Motivation

- Why should we be interested in automatic reconstructions of the fixated area within 3D space?
  - gaze is essential in natural communication
    - turn-taking (negotiating who‘s up to speak next)
    - focus of attention (resolving references, deictic gaze)
  - basic research
    - visual world paradigm in 3D (e.g. spatial relations regarding the distance from the observer)
  - application
    - virtual agents (Duchowski et al. 2004)
    - optimized rendering in virtual reality (Lübke et al. 2000)
    - selecting / picking objects
      (Tanriverdi und Jacob 2000; Duchowski et al. 2002; Barabas et al. 2004)

State of the Art

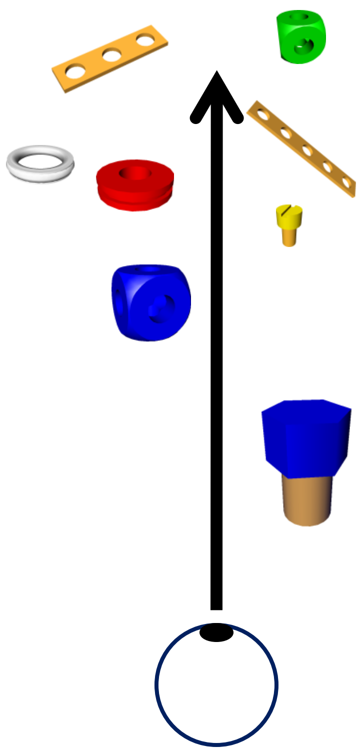
- monocular fixations extended to 3D
  1. calculate 2D fixations on a display
  2. extrapolate by casting a ray from the eye through the fixation into the scene



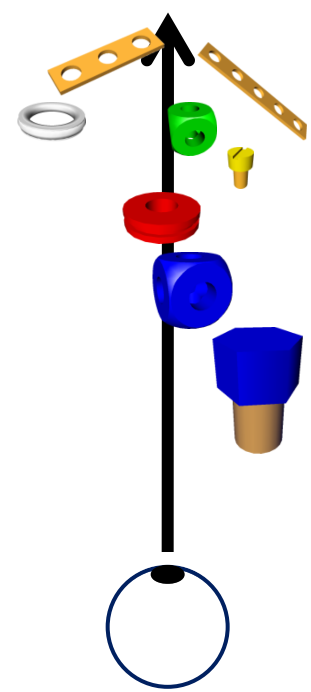
- problems
  - naive 3D fixations only possible when the ray hits an object
  - foreground vs. background problematic
  - ambiguities

Ambiguities

- Underspecification

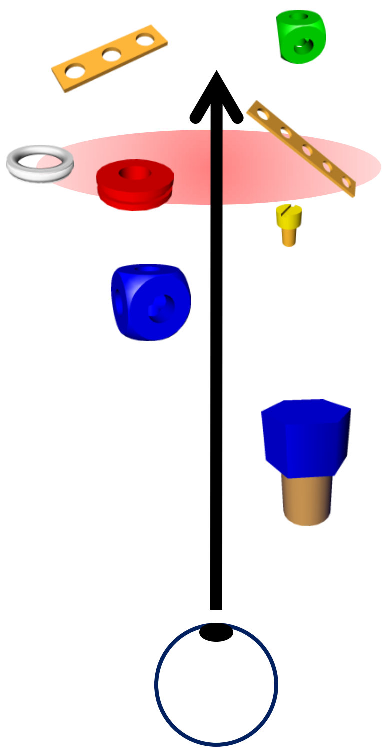


- Overspecification

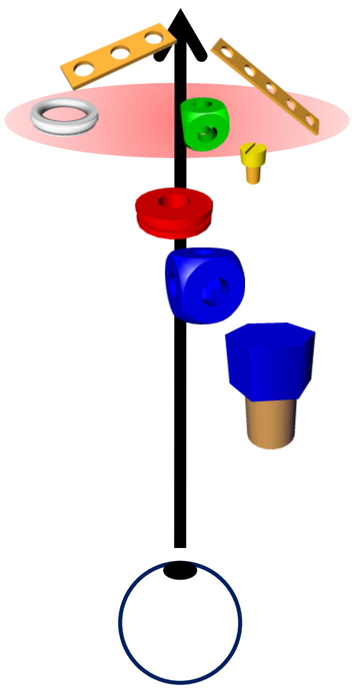


Ambiguities

- Underspecification

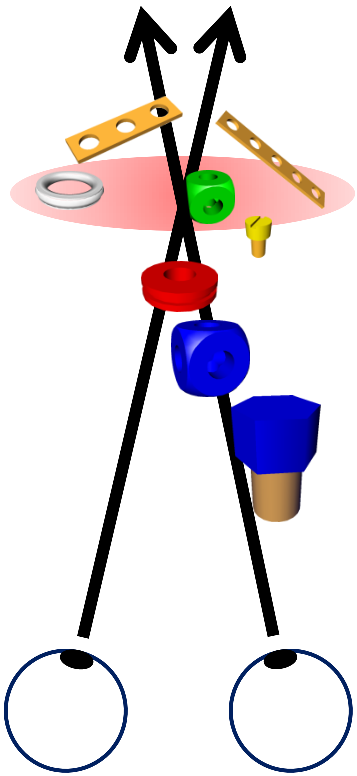


- Overspecification



Idea: determine the depth of the fixation

Ambiguities

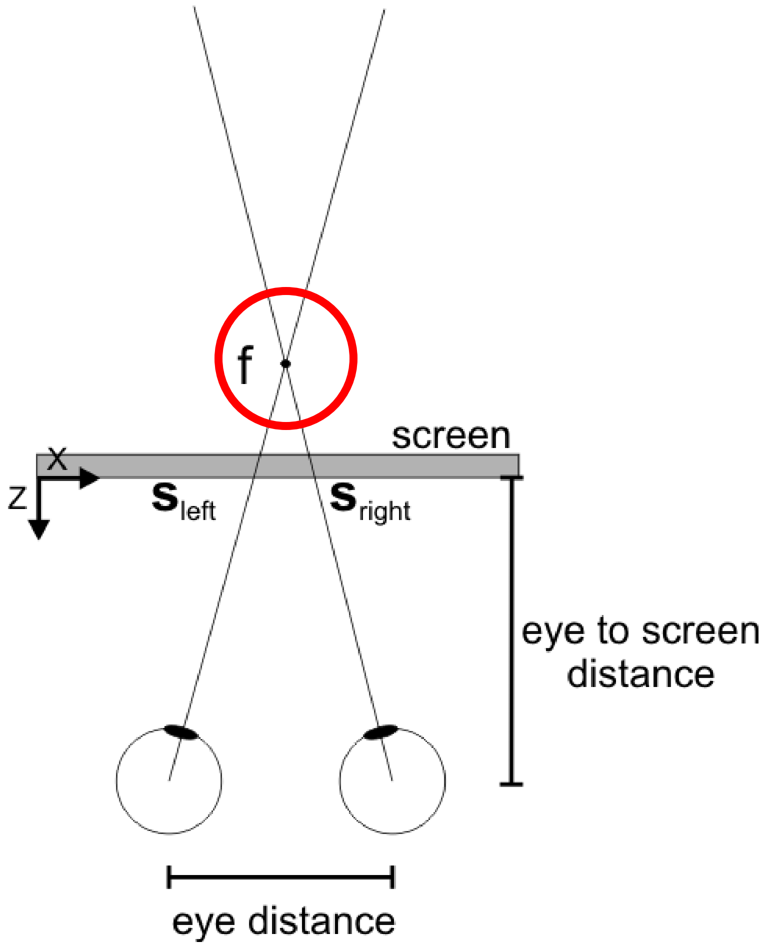


Idea: determine the depth of the fixation

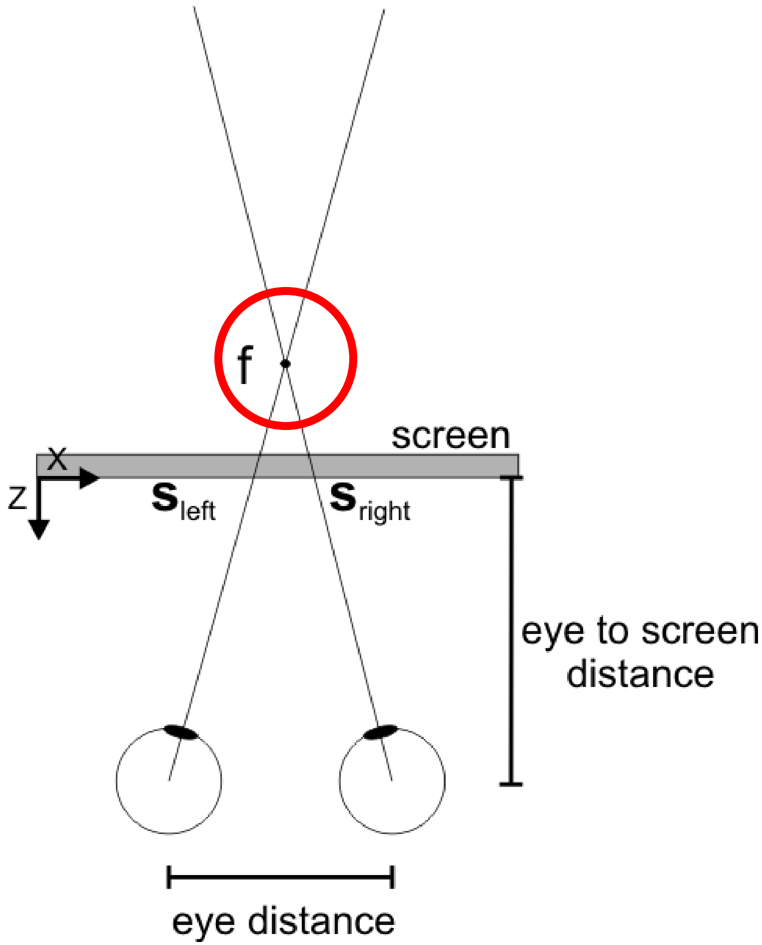
Open Questions

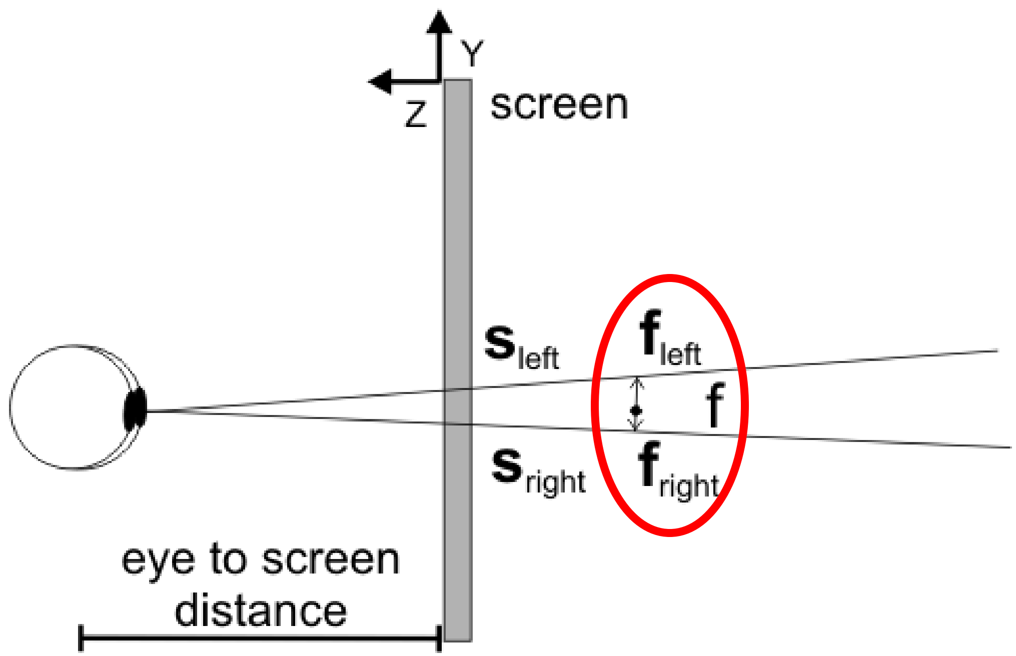
- What features can be used to reconstruct (in parts) the fixated area in 3D space?
  - accomodation
  - vergence
- What algorithms can be used?
  - geometric
  - adaptive (PSOM)
- How accurate does the eyetracker need to be?
  - low-res vs. high-res

Geometric Approach



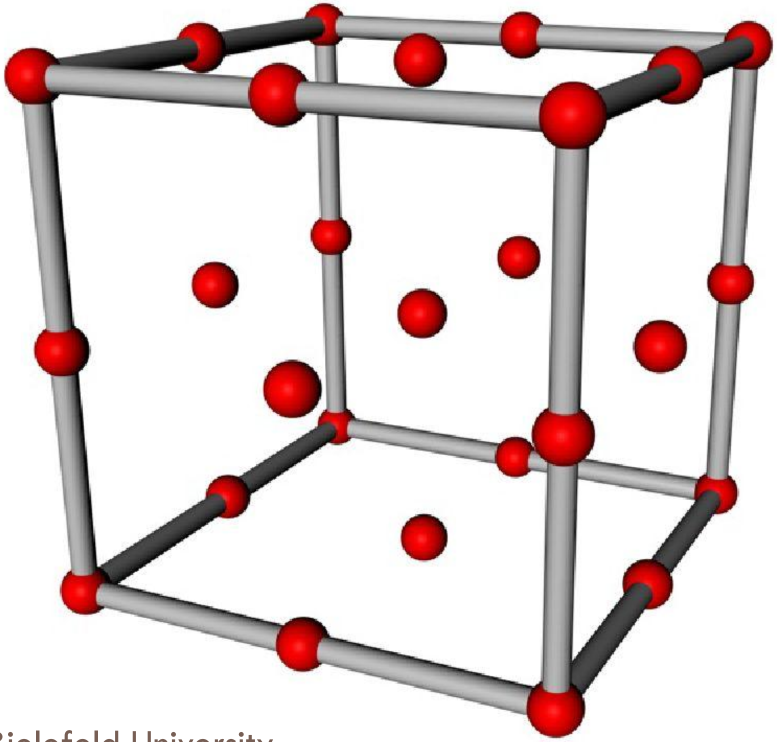
Geometric Approach



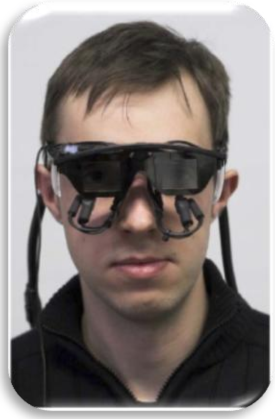


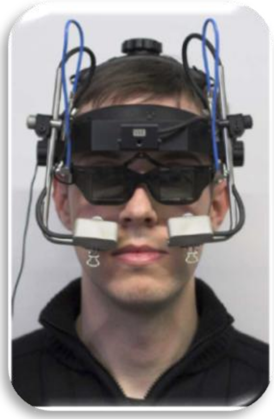
Parameterized Self-Organizing Map

- developed by Ritter in 1993
- applied to anaglyphic stereo images by Essig et al. in 2006
- PSOM
  - input
    $(x_l, y_l), (x_r, y_r), x_r - x_l$
  - output
    (x, y, z)



Eyetrackers – Technical Details



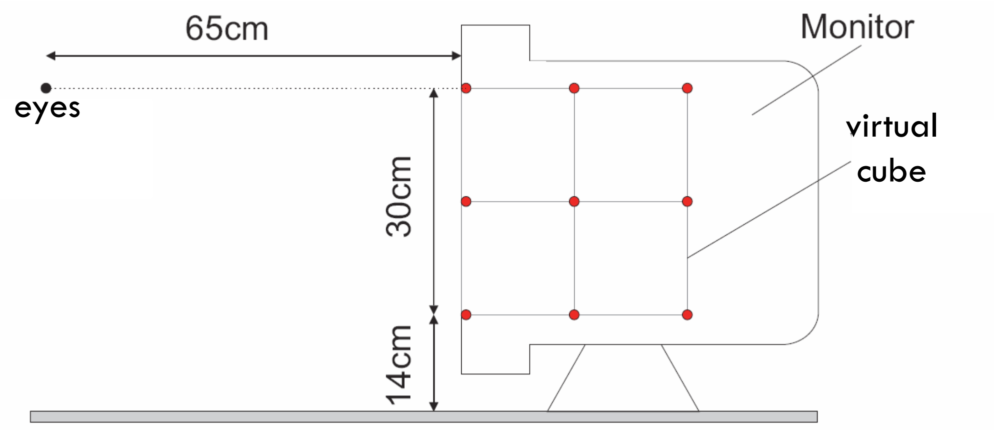


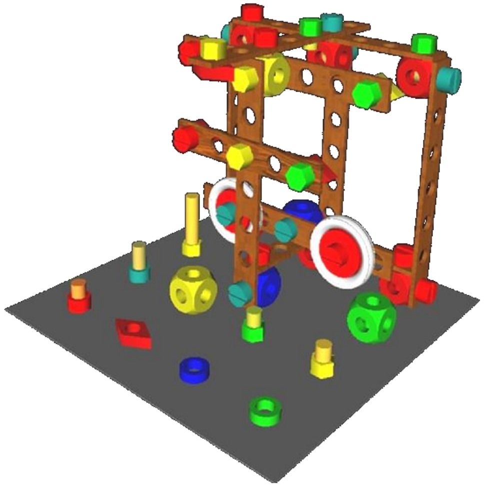
	Arrington PC60	SMI EyeLink I
temporal resolution	30 Hz / 60 Hz	250 Hz
optical resolution	640x480 / 320x240	not specified
mean error	0.25° - 1.0°	< 1.0°
Accuracy	0.15°	0.01°
compensation of head movement	not included	± 30° horiz. ± 20° vert.

Study

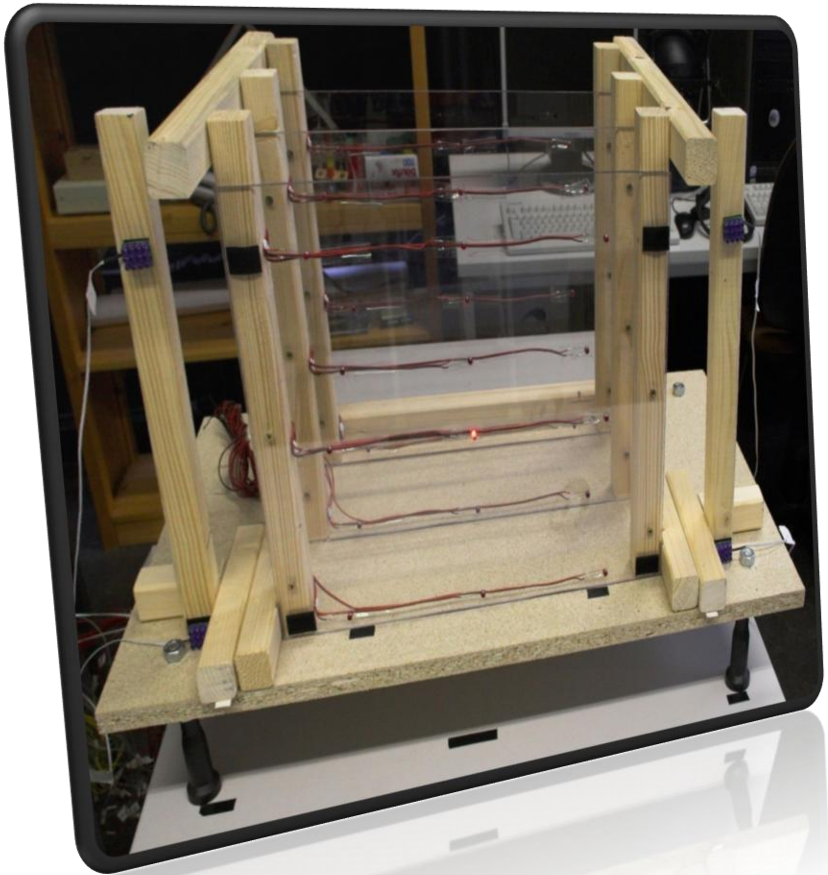
- 10 students tested
- **Hypotheses**
  - (a) **PSOM is better:**
    The PSOM is more accurate than the geometric solution.
  - (b) **EyeLink is better:**
    The SMI EyeLink I will deliver more accurate results than Arrington Research's PC60.
  - (c) **Real is better:**
    In the real scenario we will be able to get more accurate results than in the virtual scenario.

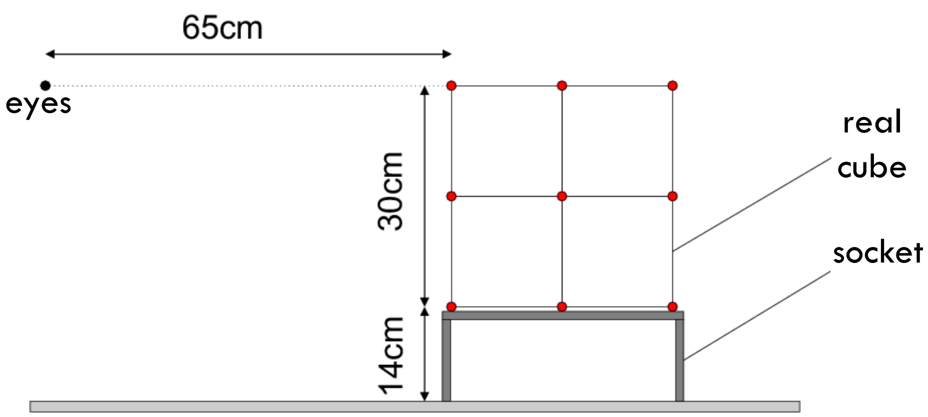
Scenario – Virtual Reality





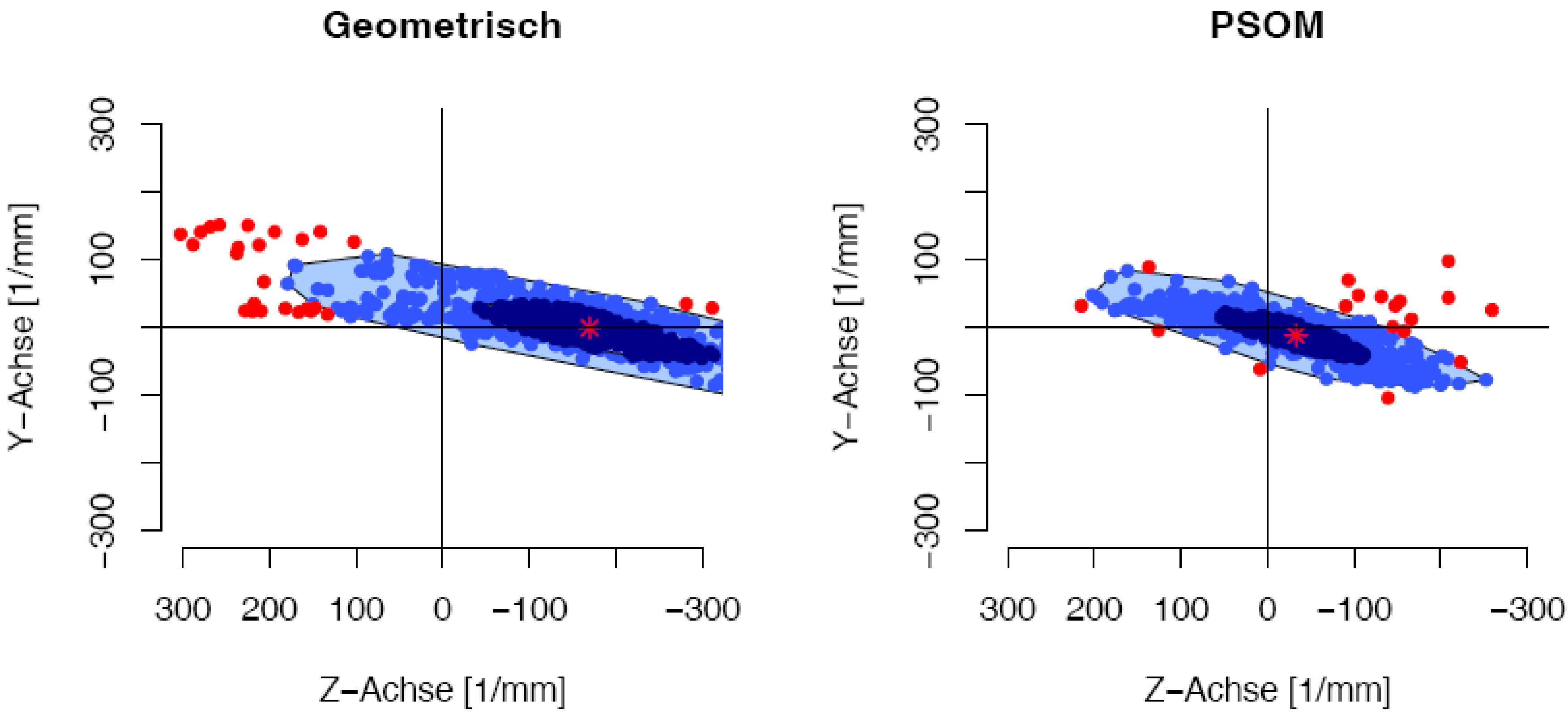
Scenario - Reality



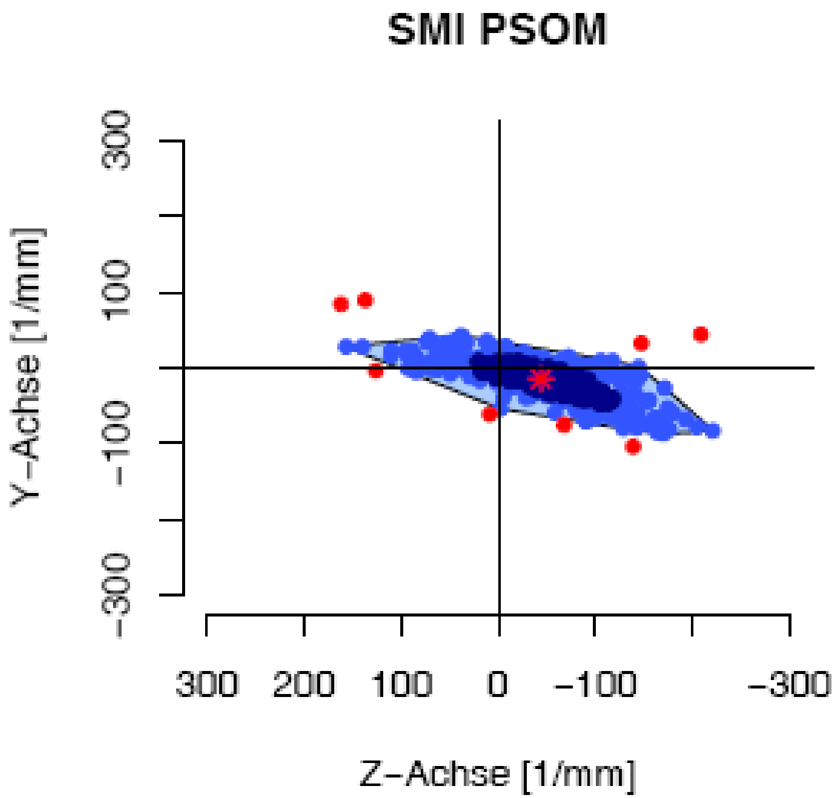


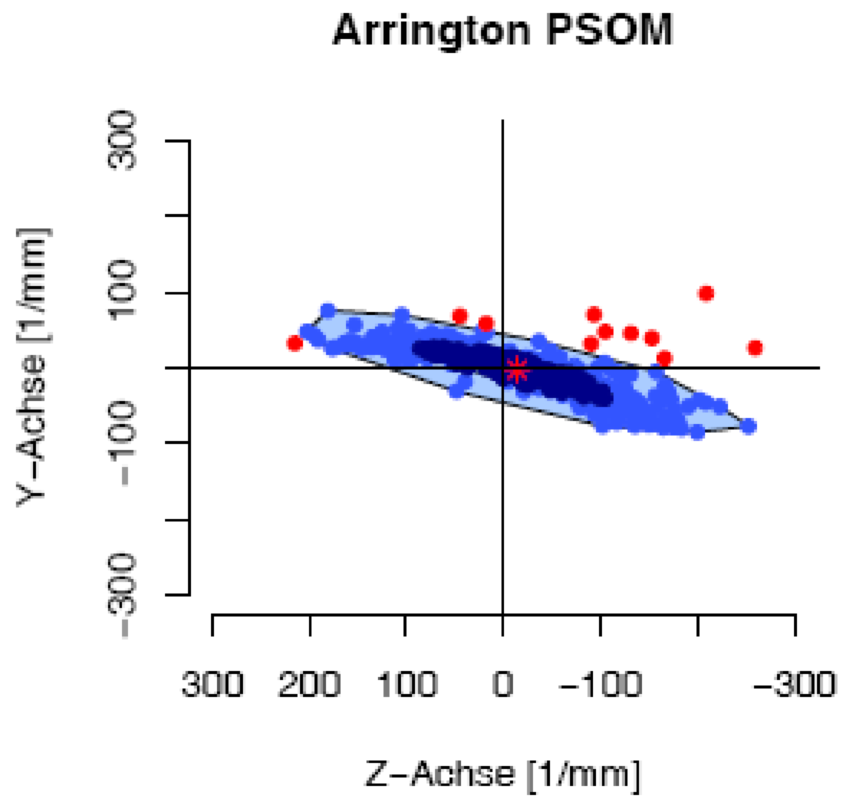
Results

Results: Geom. vs. PSOM



Results: SMI vs. Arrington



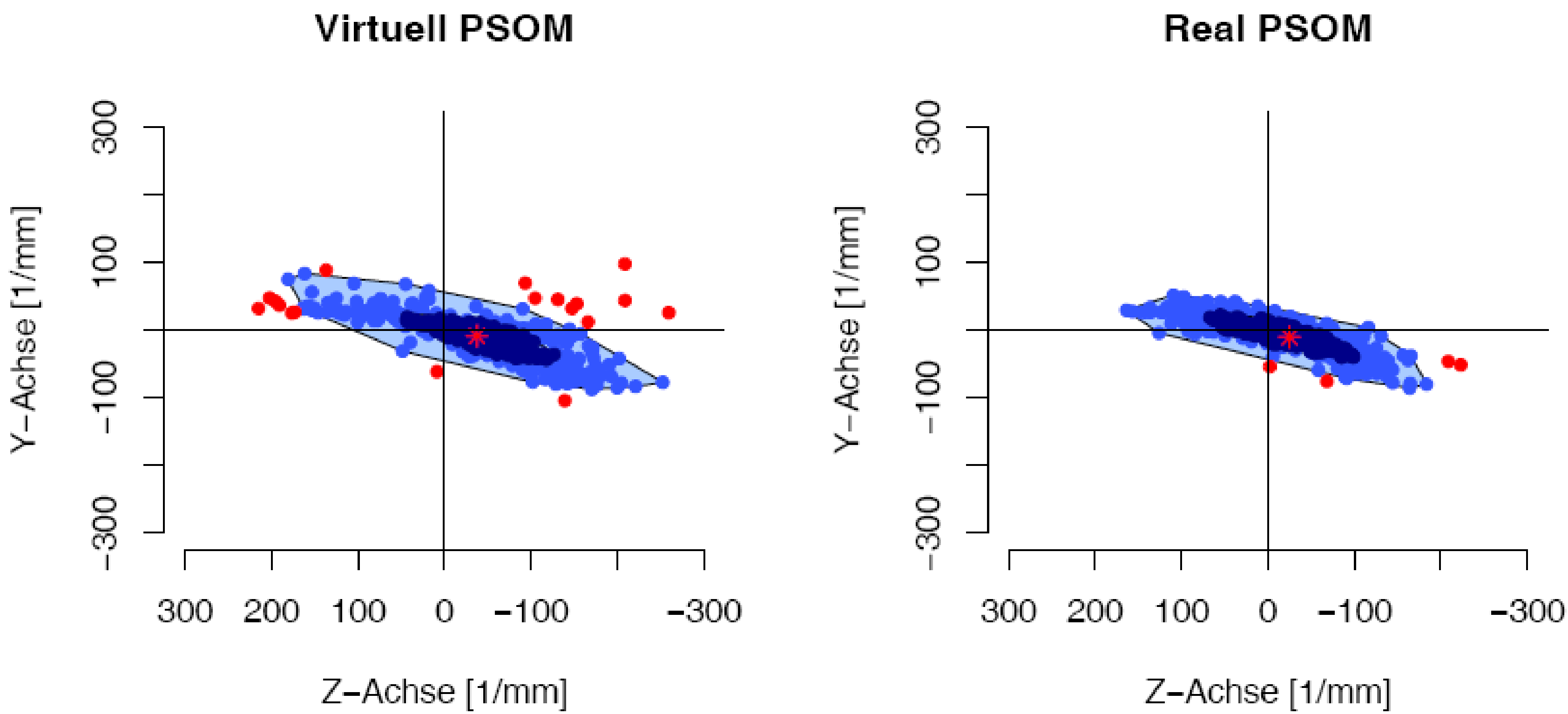


Results

device	algorithm	normally distributed	mean	difference btw. algorithms	nominal error	standard deviation
Arr.	*geom.*	no, $p < 0,001$	-195,77 mm	sig. $p < 0,001$	sig. $p < 0,001$	526,69 mm
	PSOM	yes, $p = 0,943$	-18,75 mm		sig. $p = 0,005$	96,92 mm
SMI	*geom.*	no, $p = 0,038$	-248,55 mm	sig. $p < 0,001$	sig. $p < 0,001$	149,3 mm
	PSOM	yes, $p = 0,661$	-70,57 mm		sig. $p < 0,001$	60,06 mm

- **a) is true: PSOM is more accurate and more precise**
  - significant lower nominal error
  - lower standard deviation
- **b) is twofold:**
  - **Arrington is more accurate**
  - **SMI is more precise**

Results: Virtual vs. Real



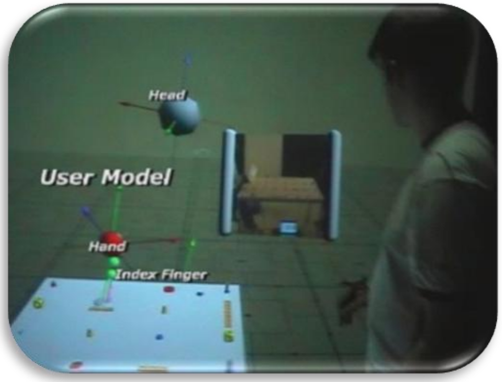
Results: Virtual vs. Real

Value	Virtual	Real
normally distributed	Yes, p = 0.074	Yes, p=0.511
mean	-44.66 mm	-17.24 mm
std. deviation	84.61 mm	69.37 mm

- **c) is true: Real is better**

Discussion

- 3D fixations can be reconstructed measuring the vergence angle and applying a PSOM algorithm
- accuracy is good, precision is less then expected from literature (Essig et al. 2006)
  - but “real world” objects have been used (not dots)
- current advice for basic research
  - distribute critical objects at least 30cm apart when working with near objects
- next study will involve a larger scenario in VR (3m x 3m x 3m)



… in collaboration with

Universität Bielefeld

Matthias Donner
Dipl. Informatiker

Dr. Marc E. Latoschik

Prof. Dr. Ipke Wachsmuth

PASION

Psychologically Augmented Social Interaction Over Networks

European Union program under IST, January 2006 - December 2009