Natural Space as Interface

Interaction Research at UCSB's Four Eyes Laboratory

Tobias Höllerer

holl@cs.ucsb.edu

"Four Eyes" Lab
Imaging, Interaction, and Innovative Interfaces

Computer Science Department
University of California, Santa Barbara

Outline

- **Lab Overview**
- **Selected Research Projects**
  - Interactive Tools for Augmented Reality "X-ray vision"
  - HandVu: Hand gestures for mobile computing
  - Photorealistic Real-Time AR in Unprepared Environments
  - Multi-Flash Imaging
  - Facial Expression Analysis
  - Constraint-Based Interaction with RNA Molecules
- **Conclusions**

The Four Eyes Lab

M. Turk

- Research in the Four "I"s of Imaging, Interaction, and Innovative Interfaces
- 2000/01 Matthew Turk starts the "i-lab"
- 2002/03 Tobias Höllerer joins, new lab space, new name
- 2004 First PhD Student Graduates (Mathias Kölsch)
- 2004/05 Visiting Professors/Researchers
  June-Ho Yi,
  Hyoung Gon Kim,
  Ismo Rakkolainen

The Four Eyes Lab

- **Current Size:**
  - 2 Permanent Faculty
  - 2 Visiting Faculty
  - 1 Post-doctoral Researcher
  - 7 Ph.D. Students
  - 4 Masters Students
  - 3 Undergraduate Researchers

UCSB

The Four Eyes Lab

~20 Research Projects in

- Computer Vision Methods & Applications
- Perceptual & Multimodal Interfaces
- 3D Imaging and Interaction
- Augmented Reality
- Mobile HCI

Four Eyes Lab: General Motivation

Provide better, more compelling HCI technologies in a variety of important computing environments

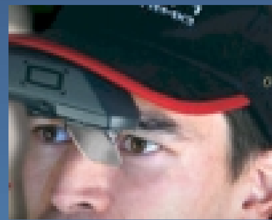
- Desktop, immersive, VR
- Mobile, ubiquitous, AR

…and for a variety of application areas

- Sciences, entertainment, digital art, visualization, …

Investigate fundamental issues in developing robust, real-time, working technologies for interactive systems

Augmented Reality



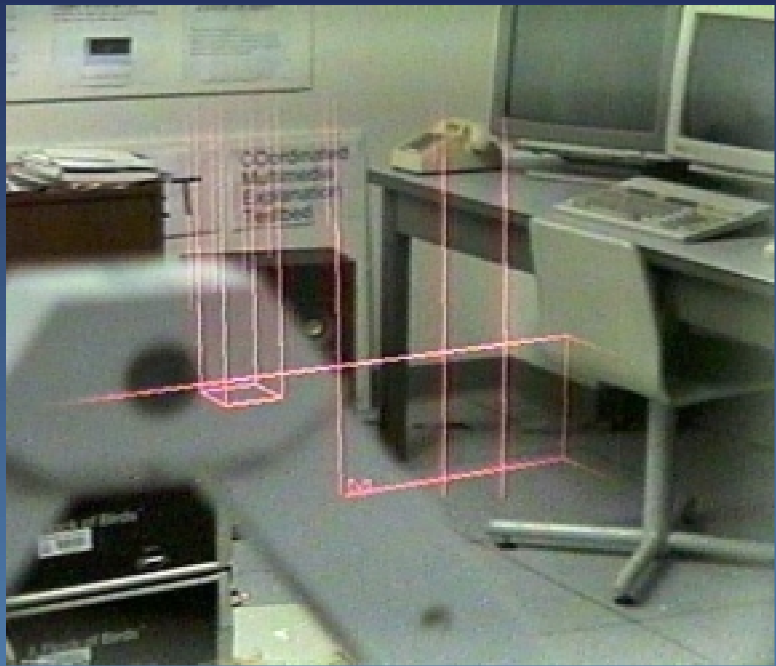
1) Blends real and virtual, in real environment
2) Real-time interactive
3) Registered in 3-D

Mobile Augmented Reality

- **Overlay information anywhere, anytime**
- **Many application areas**
  - Tourism, journalism, architecture, construction, maintenance, repair, military, training, medicine, entertainment, …
- **General purpose situated UI for wearable computers**
  - Navigational aids, communication aids, personal situated information DB, general UI for appliances

UCSB

X-Ray Vision in AR



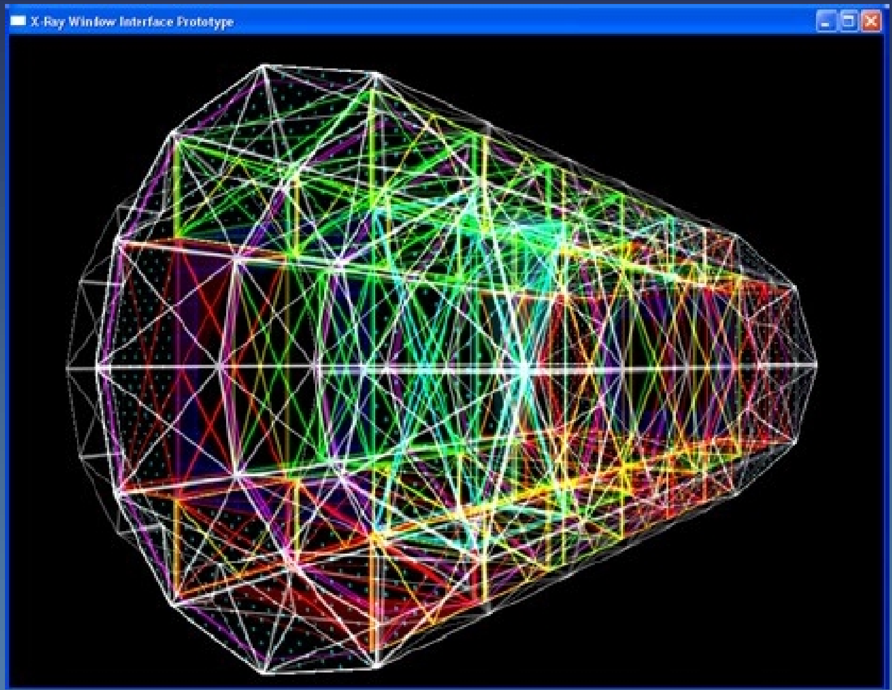
Architectural Anatomy

Feiner, Webster, Krueger, MacIntyre, & Keller, 1995

Augmented Reality Visualization for Laparoscopic Surgery

Fuchs, Livingston, Raskar, Colucci, Keller, State, Crawford, Rademacher

More Motivation



W. White (S. Illinois U., NASA Johnson Space Center)

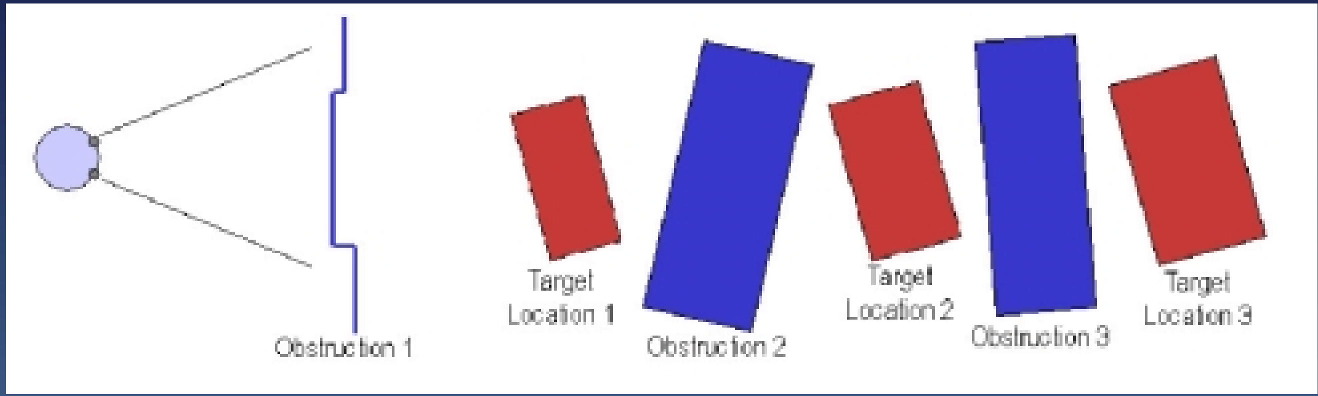
X-Ray Window: Portable Visualization on the International Space Station,

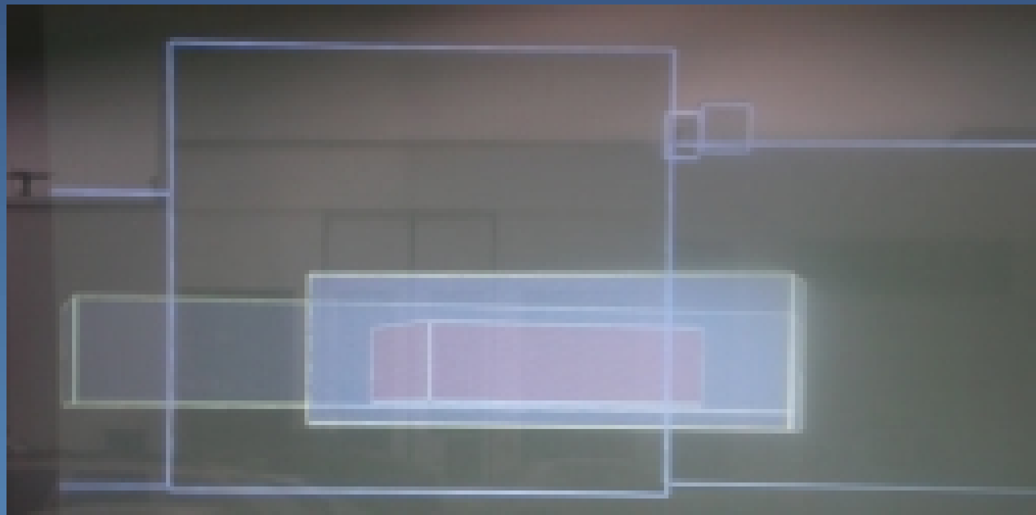
Technical Sketch, SIGGRAPH '04

X-Ray Window Rendering of ISS *Destiny* Module, with Color-Coded Stowage and Equipment Components

UCSB

Direct Display Methods





Resolving Multiple Occluded Layers in Augmented Reality
Livingston, Swan, Gabbard, Höllerer, Hix, Julier, Baillot, Brown,
ISMAR ‘03

Tangible Interaction with Occluded Infrastructure

Provide *interactive tools* to explore occluded infrastructure from a mostly static vantage point.

Slice through the environment in front of you.

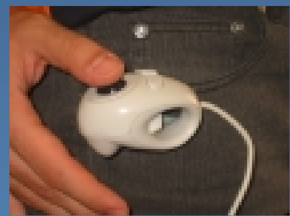
Integrate birds-eye overviews

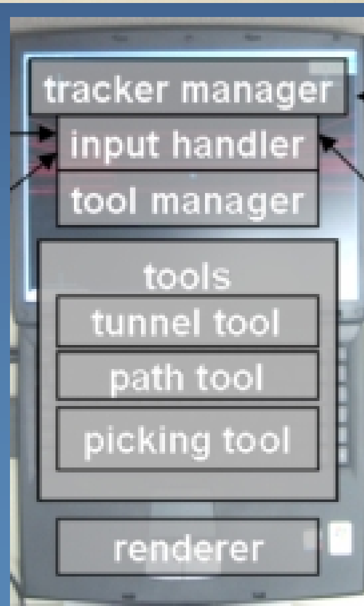
Implementing X-Ray Vision

Hardware Overview

Video See-Through AR
(Tradeoff: FOV / immediacy)

Focus on far-field AR,
Interaction from a relatively static viewpoint.





Running off a 2.2GHz laptop computer with NViDIA Quadro4 Go graphics.

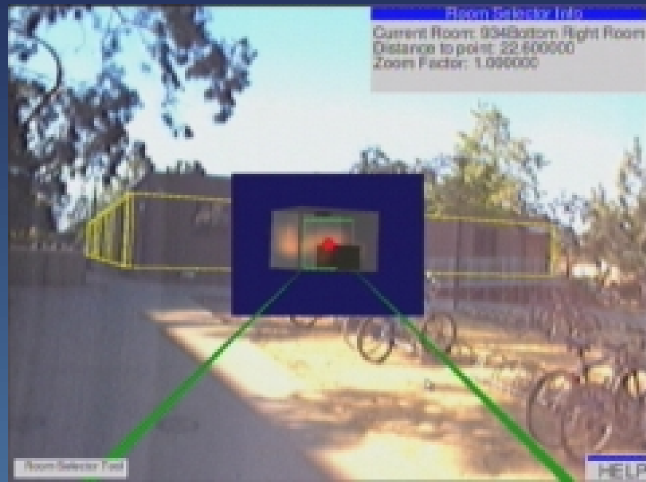
Implementing X-Ray Vision

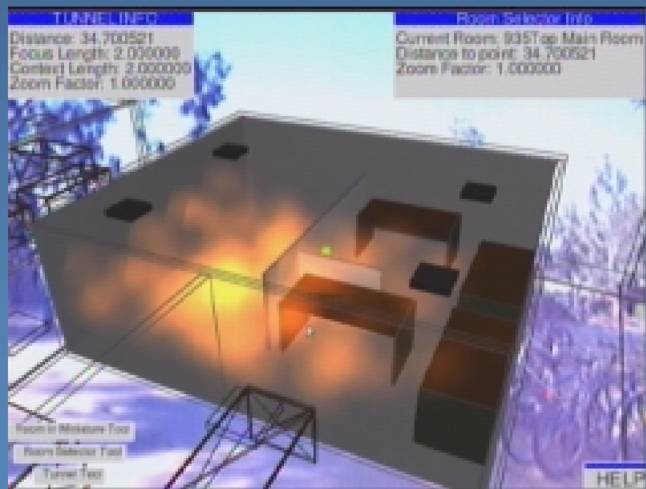
Software / UI Overview

Implements interactive techniques in the form of a small toolset.

First person tools use virtual lenses and three dimensional sliders

Third person tools allow users to control a virtual camera to view objects





The Toolset

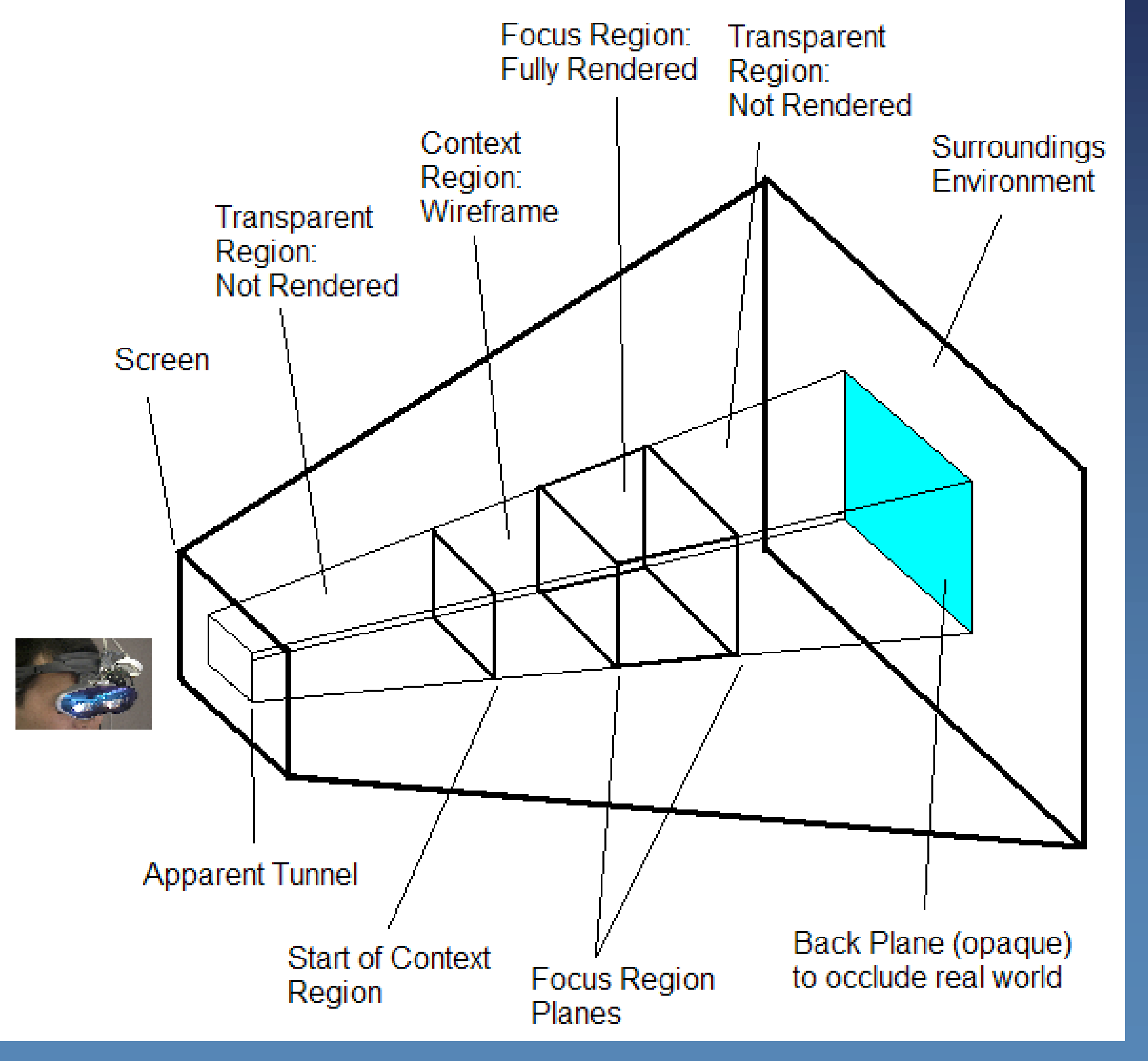
The Toolset examines two main approaches: Volume based and Room based

- Volume tools work on general geometry
- Room tools assume that structures are modeled as collections of rooms, and allow the user to examine buildings on a room by room basis.

General idea: assume limited semantic information. Shifts burden to modeling.

UCSB

An “X-Ray Lens”: The Tunnel Tool



The Tunnel Tool

The Tunnel Tool

- Good for viewing data clouds (temperature, wireless connectivity, …)
- Good for viewing buildings whose main walls are aligned with, or orthogonal to, viewing direction
- Can be confusing when walls are cut at angles.

UCSB

The Room Selector Tool

- The Room Selector Tool is a room-based, first person tool
- Assumes rooms to be represented in the environment model
- Intuitive – This is what many people imagine "x-ray vision" looks like

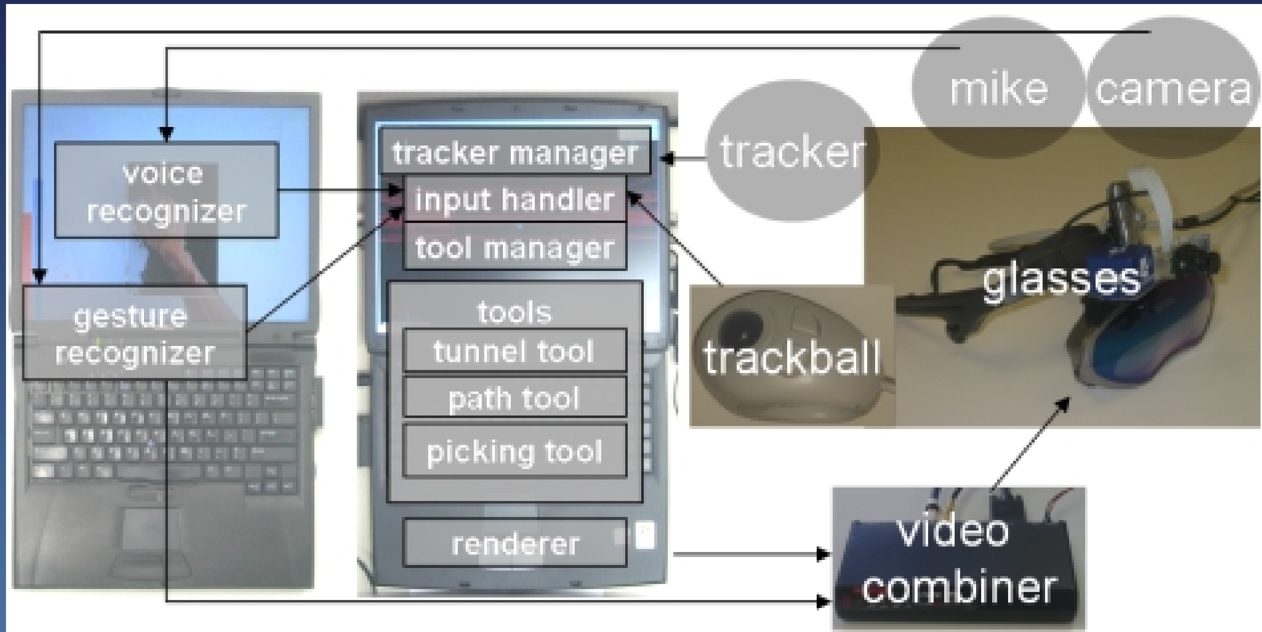
The Room Selector Tool

- **But what happens when objects in the room occlude other objects?**

Putting it all together

Video

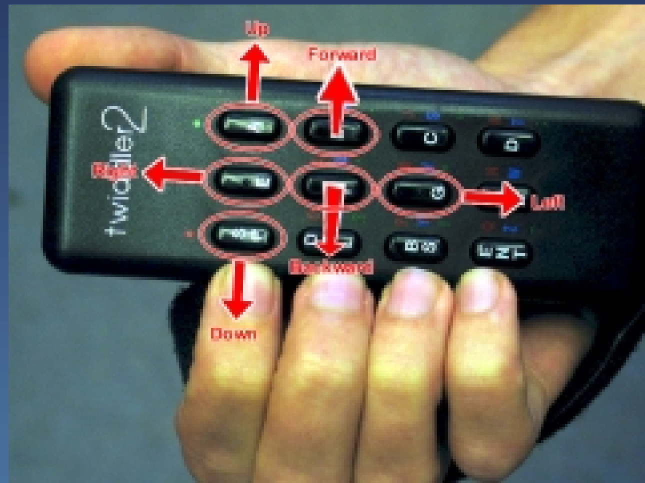
Architecture



Evaluation of Techniques for Interaction (3D Cursor Control) at a Distance

Four techniques tested

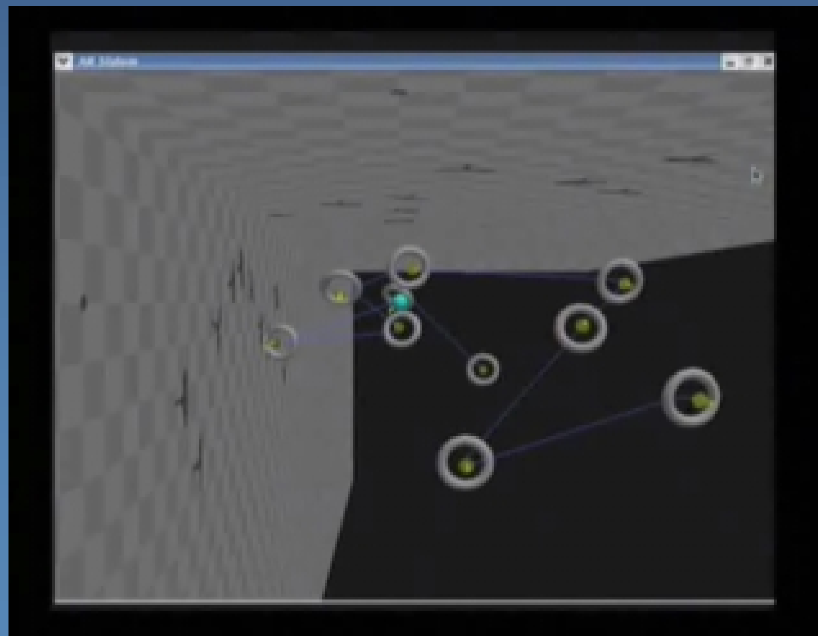
- T1 uses a Twiddler2 keyboard
- T2 uses a Rocket Mouse finger trackball
- T3 uses head orientation and two buttons
- T4 uses head orientation alone (plus a button to switch between modes)



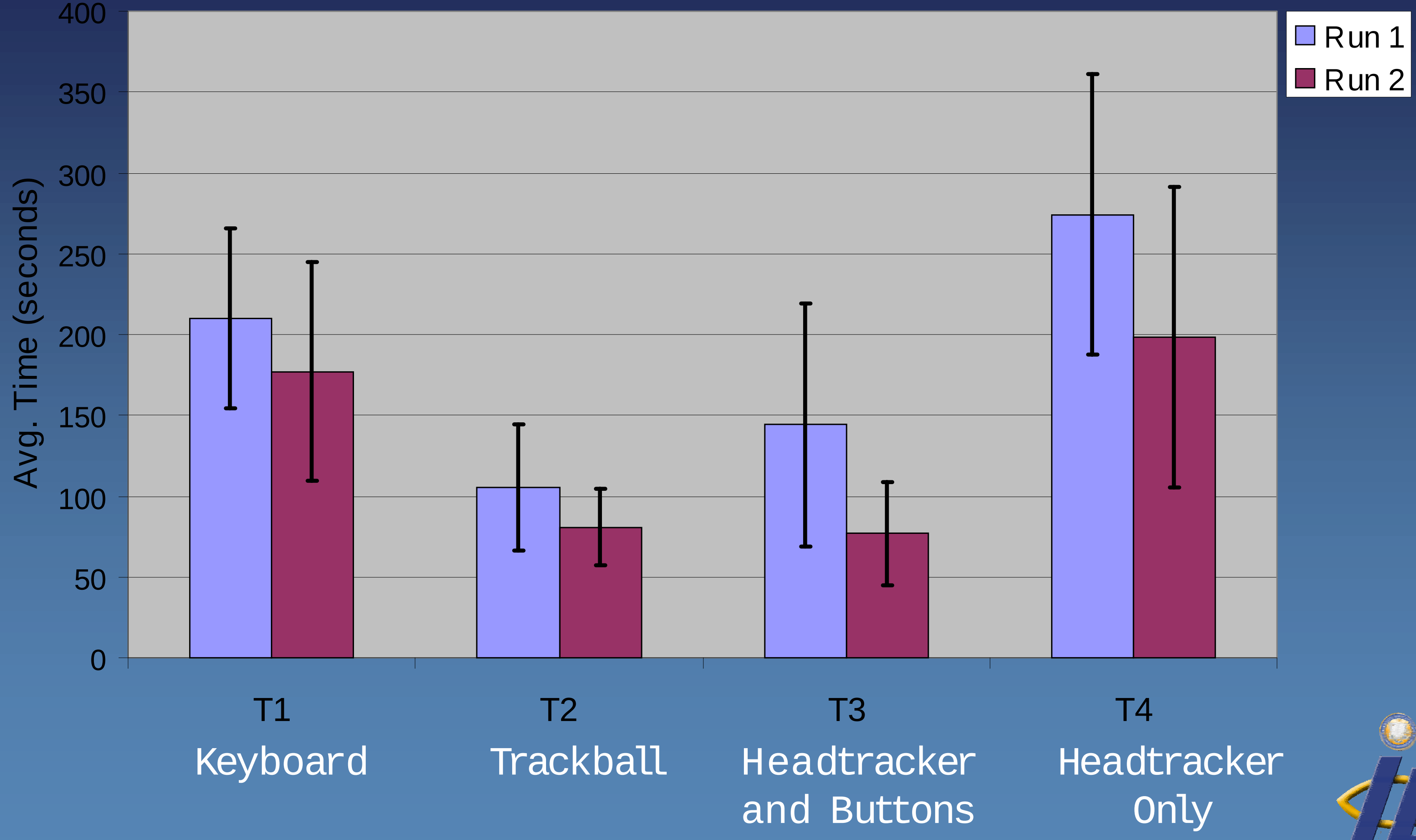
Study details

- Study in VR with users sitting
- 13 users – 9 male / 4 female
- 10 gate course

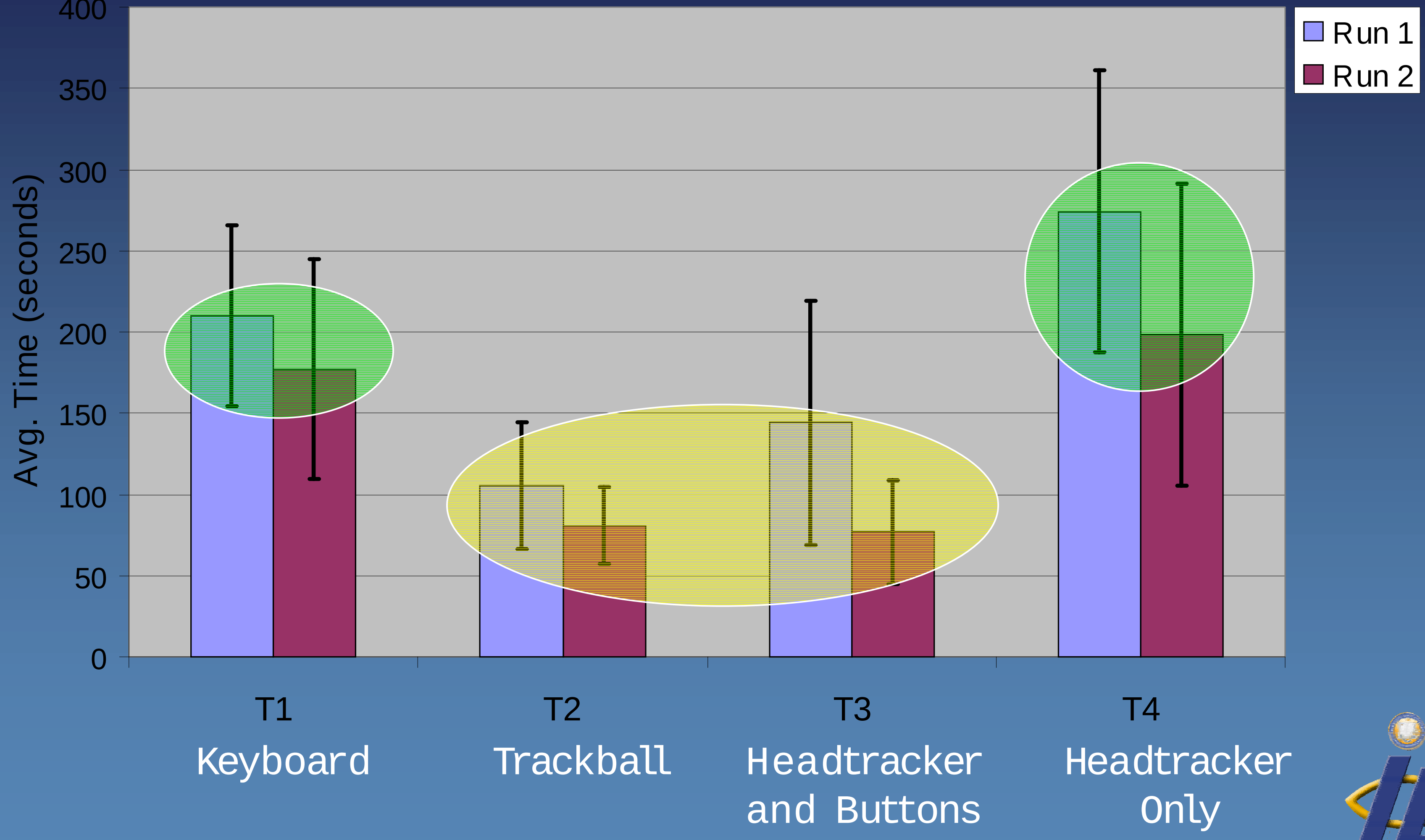
- Training session for each technique
- Two test-runs per technique
- Order of techniques was permuted between users



Study Results (Times)



Study Results (Times)



Summary of Results

- T2 (Trackball) and T3 (Headtracker and Buttons) performed nearly equally well
- Users found T2 (Trackball) more favorable than it objectively performed
- Times and error rates were not directly correlated

Interaction Design Recommendations

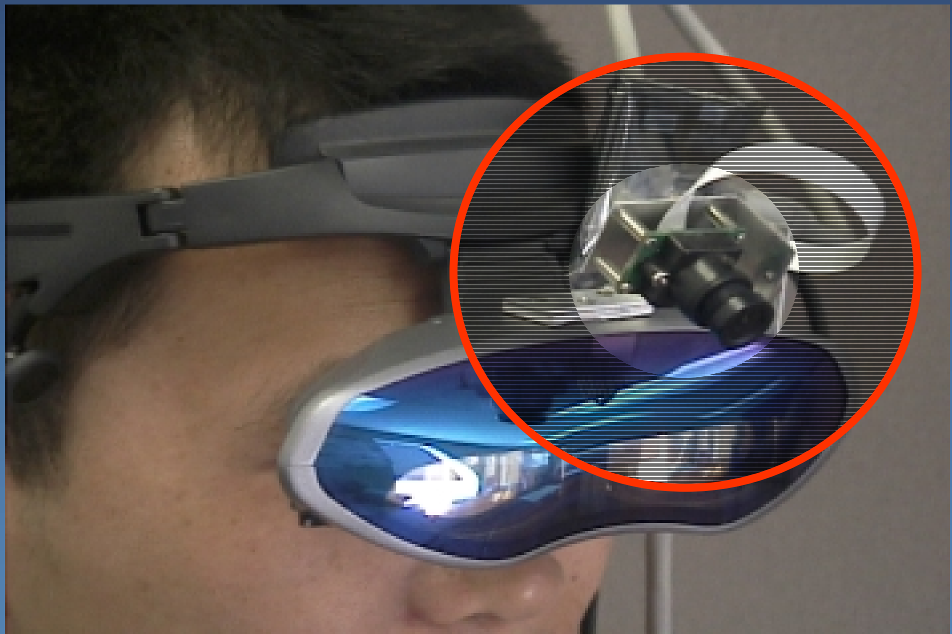
Trackball well suited for 3D cursor control from a distance

Head tracking can be a useful input technique in certain cases, particularly panning

Head tracking is commonplace in AR

Keyboard has low error rate, but slow times

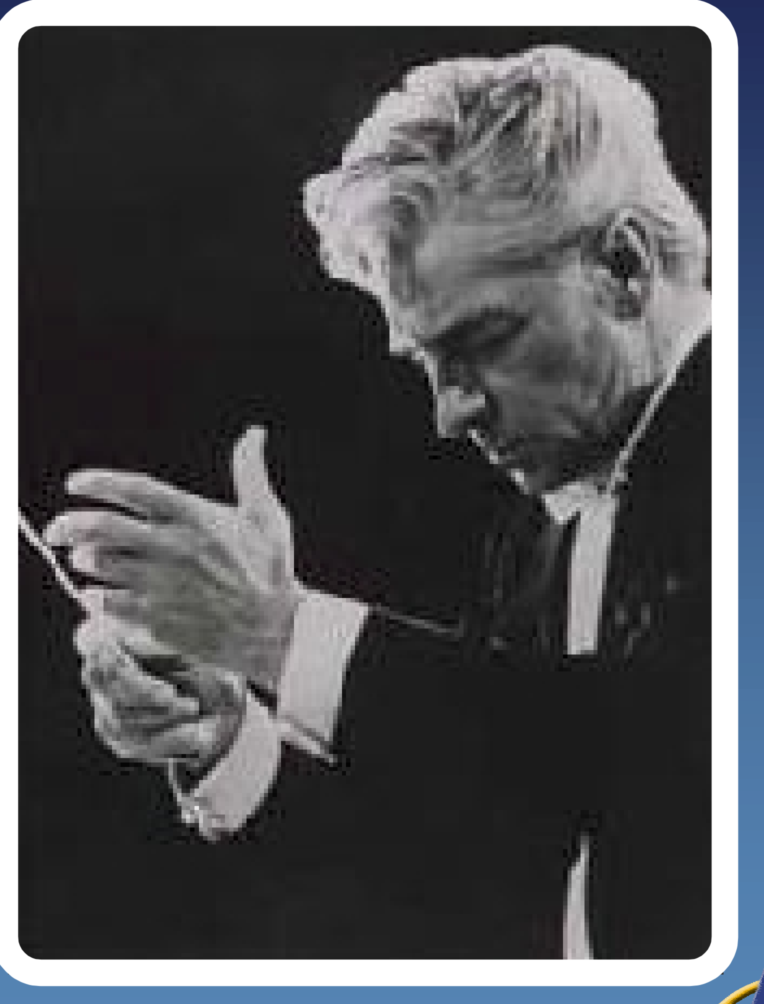
Vision-based interfaces for mobility



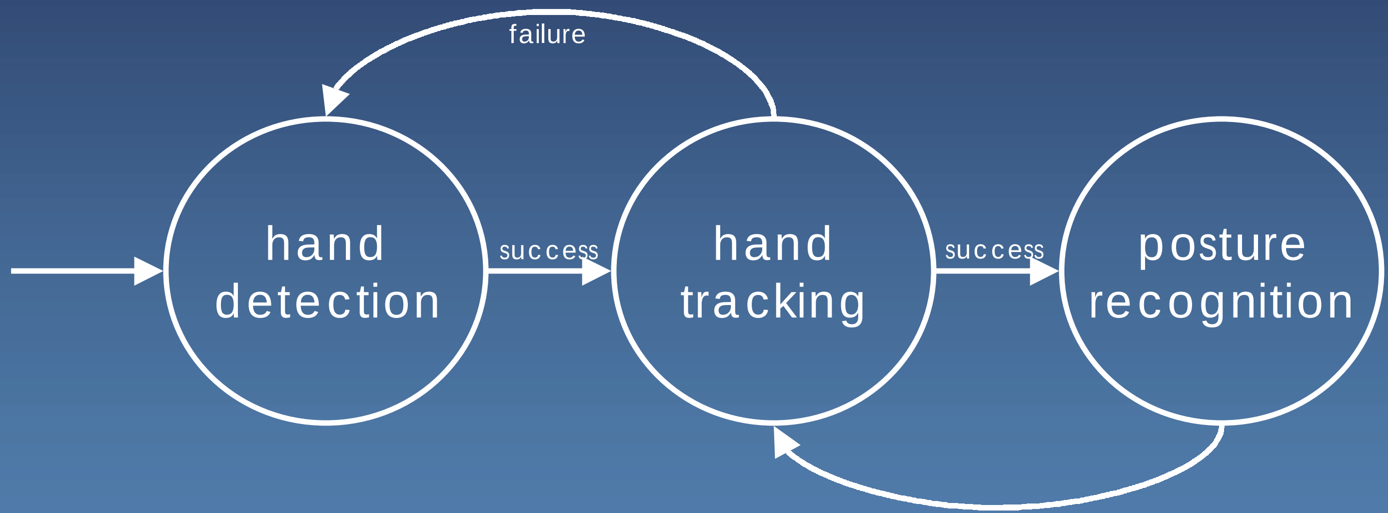
UCSB

Hand gestures

- unobtrusive – no gloves or devices
- lightweight, mobile, wireless
- good for adverse environments
  - sand, moisture, silence, noise
- camera's versatile applications
- hardware advances and price drops

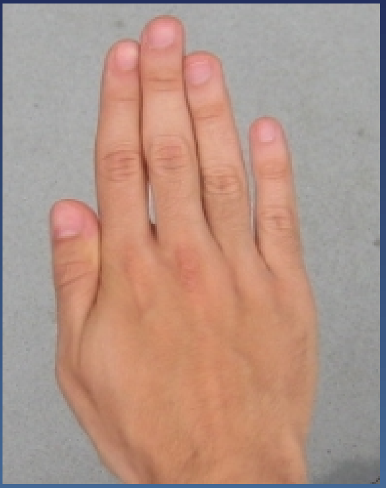


HandVu Vision-Based Interface Toolkit (M. Kölsch)



Robust hand detection

[Kölsch & Turk: Face and Gesture Recognition 2004]



extension to [Viola & Jones 2001]

detection rate: 92 %

false positive rate:

$1.01*10^{-8}$, or

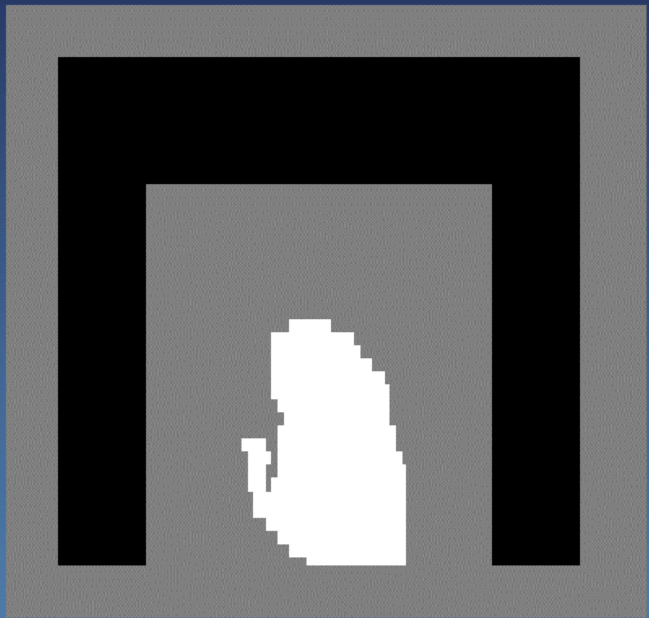
one in 279 VGA-sized image frames,

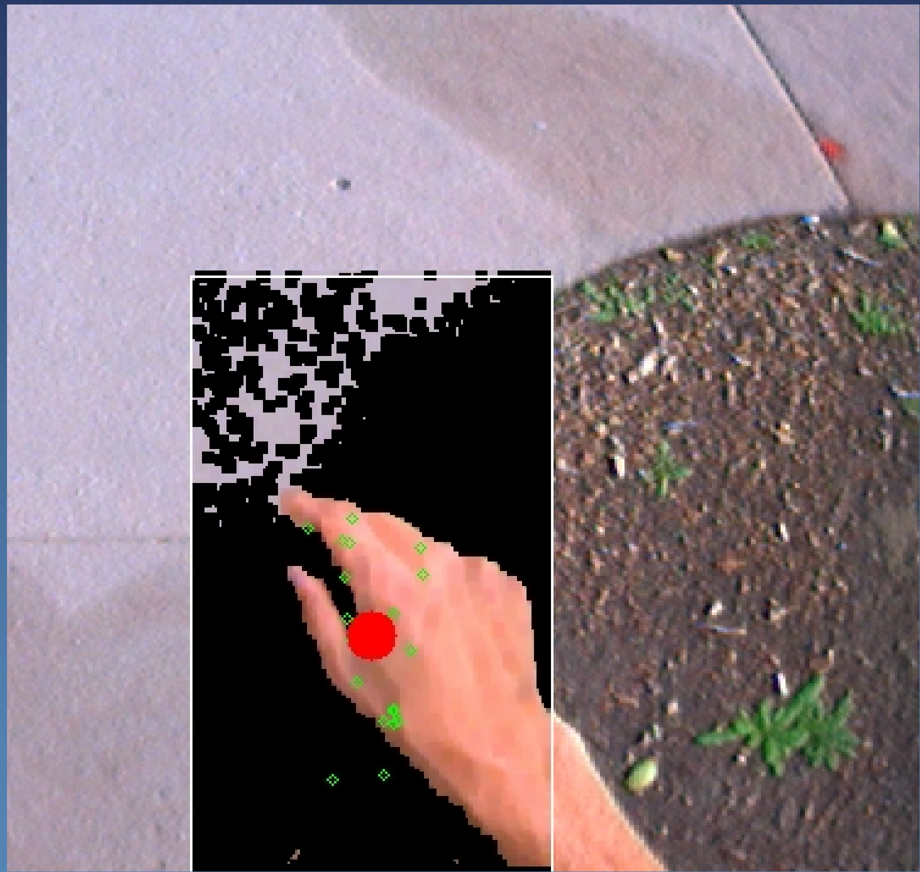
one in ~1500 in our 218x308-size area

with color verification:

few false positives per hour live video

Online learned hand color





Flock of Features: Fast 2D hand tracking

tracking of articulated objects

- location in 2D image plane
- no finger configuration

fast: 5-18 ms on 3GHz

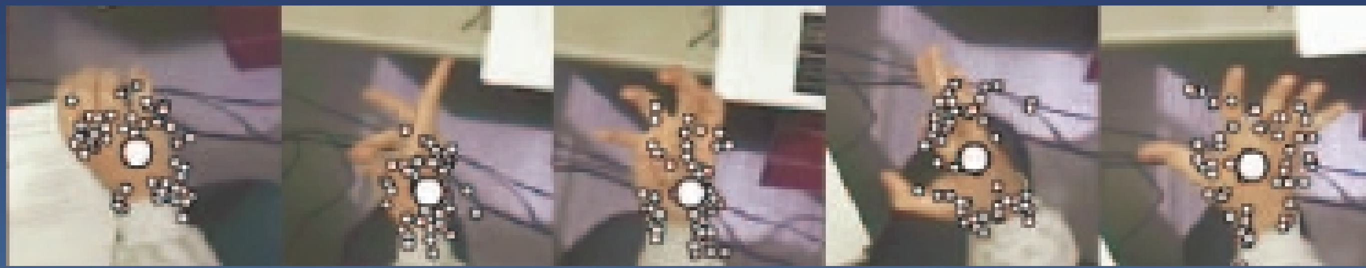
robust:

- despite arbitrary backgrounds
- despite changing lighting conditions

UCSB

Flock of Features tracking

individual KLT features (grey level)



loose, global "flocking" constraints:

- not too close to another feature
- not too far from flock

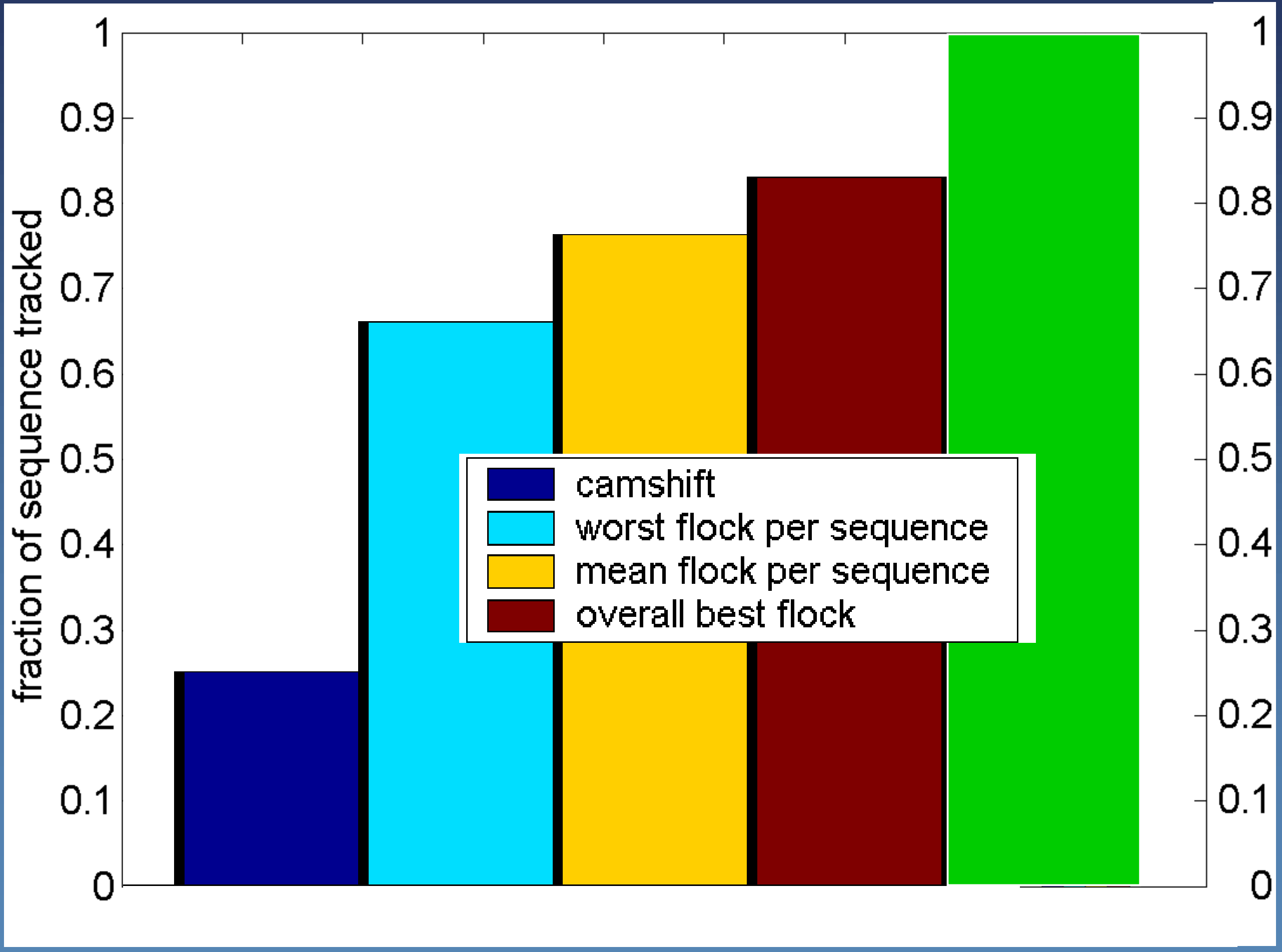
color probability for backup consultation

→ multi-cue integration

→ overcomes single-cue failure modes

Flock of Features tracking

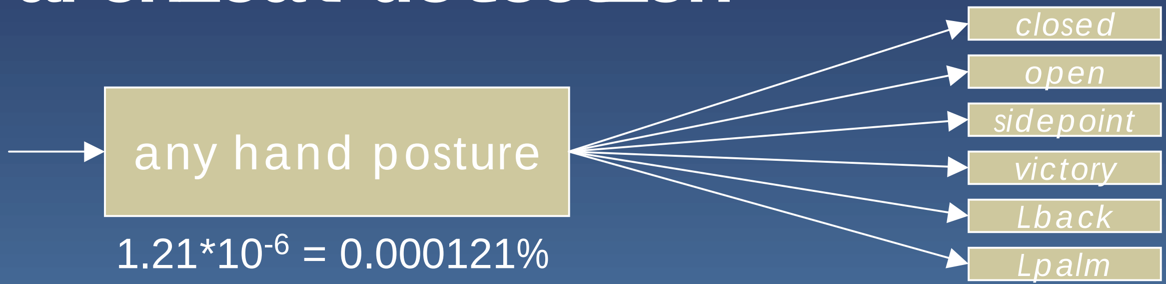
Flock of Features evaluation



[Kölsch & Turk: RTV4HCI (at CVPR) 2004]

Hand posture recognition

hierarchical detection



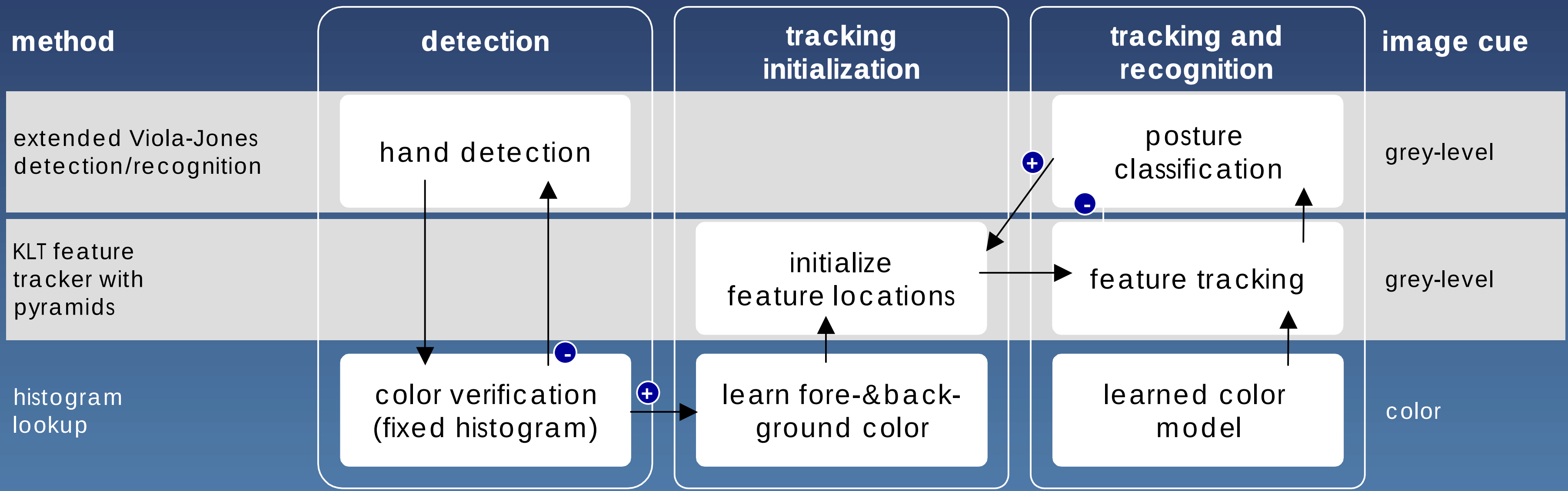
$1.21*10^{-6} = 0.000121\%$

distinguishes six postures

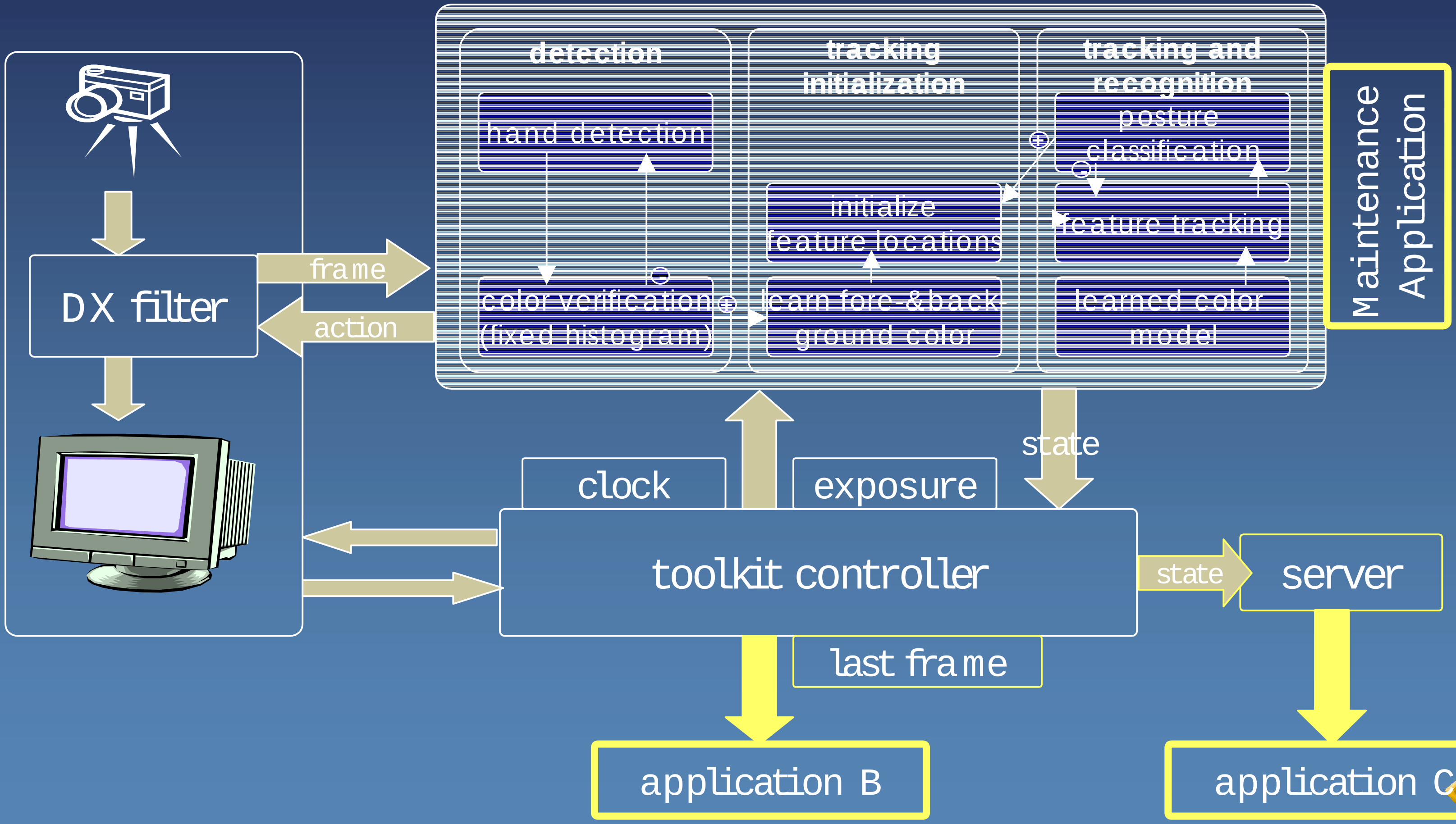
re-initialization of tracking

- learn color model
- set feature locations

HandVu's vision methods



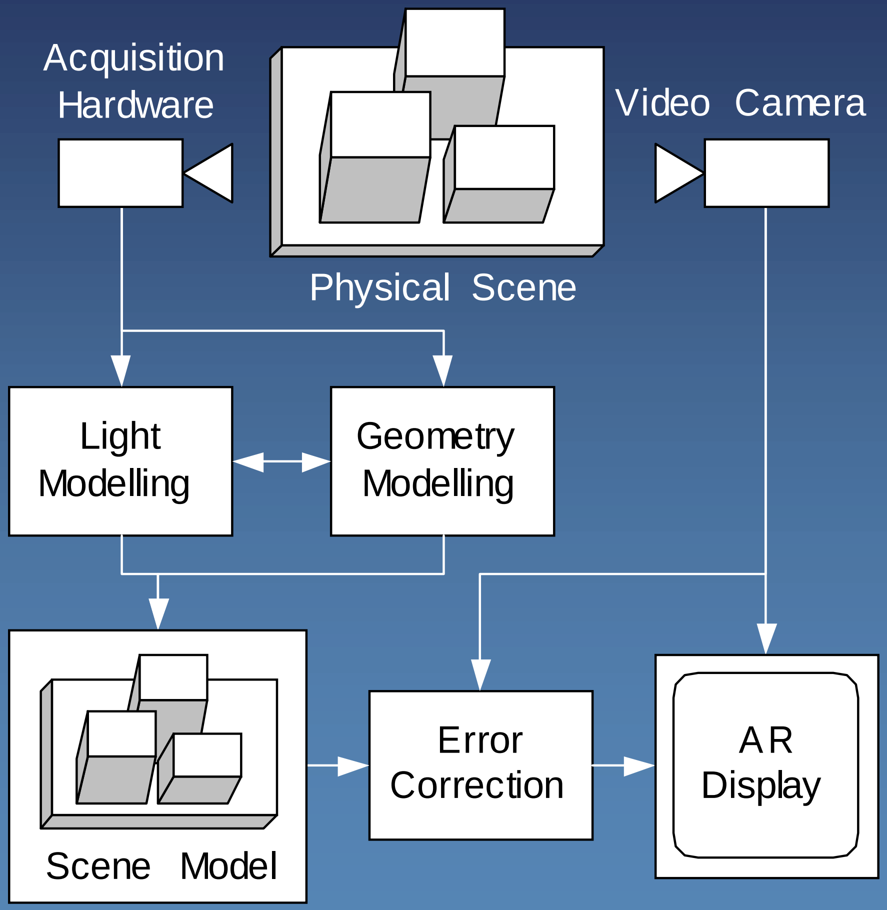
HandVu toolkit



Photorealistic Real-Time AR in Unprepared Environments (S. DiVerdi, T. Höllerer)

- Seamless integration of virtual and physical worlds is a goal of AR.
- Good integration requires accurate geometric registration, as well as consistent illumination.
- Existing realistic mixed reality systems require extensive start up costs - calibration, scene modeling, instrumentation, etc.

Photorealistic Real-Time AR in Unprepared Environments



- System will use a variety of data acquisition hardware.
- Aggregated light and geometry information must be stored in unified scene data structure with low incremental update cost.
- Video input will be used for error correction and video see-through display.

Photorealistic Real-Time AR in Unprepared Environments

http://www.middlebury.edu/stereo/

Approximate Geometry Acquisition

- Physical geometry necessary for lighting and shadows.
- Dynamic response to geometry changes requires (semi-) automatic acquisition techniques.
- Stereo vision algorithms for automatic updates.
- Interactive (marker sweep, line markup) intervention for semi-automatic updates.

Photorealistic Real-Time AR in Unprepared Environments

© P. Debevec. Rendering Synthetic Objects Into Real Scenes, SIGGRAPH 98. http://www.debevec.org/Probes

Lighting and Shadows

- Dynamic response to lighting environment changes requires automatic acquisition techniques.
- Tracked light probe provides partial information.
- Currently, shadows from IBL are only done with offline rendering techniques → turn interactive
- Process light probe to find small number of bright area light sources for shadow generation.
- Real-time environment map

UCSB

Photorealistic Real-Time AR in Unprepared Environments

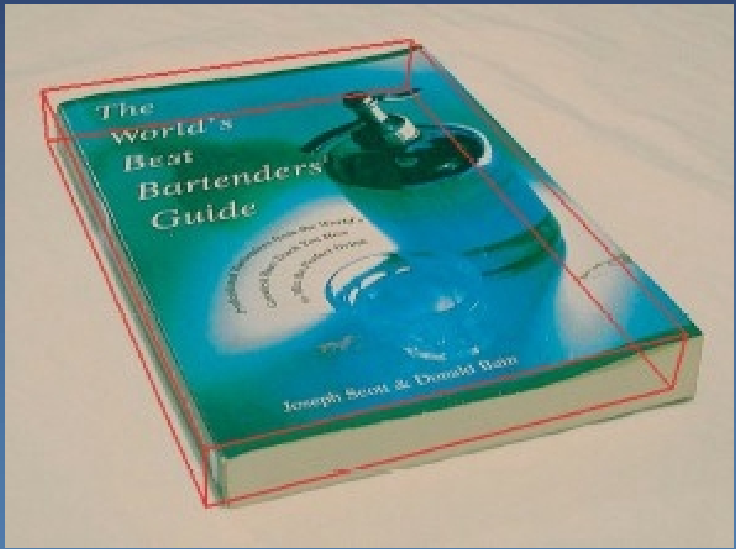


Image-Based Error Correction

- Applying virtual lighting and shadows requires pixel-perfect registration. Errors in registration are easily detectable near edges.
- Registration errors are difficult to eliminate, especially with automatic geometry acquisition techniques.
- Video of scene can be used to correct errors by making rendered lighting edges correspond with depth discontinuities

Multi-flash imaging for depth discontinuities

SIGGRAPH 2004 and CVPR RTV4HCI 2004

Joint work with MERL, MIT

General idea:

- Take *N* images with *N* differently located flashes
- Reason about the cast shadows to determine depth discontinuities

Applications:

- Non-photorealistic rendering
- Finger spelling (for sign language recognition)

Depth Discontinuity Detection

(R. Feris w. R. Raskar (Mitsubishi Electric Research Labs) and M. Turk)

Source object, no shadows

Multi-Flash Camera

Multi-Flash Imaging

(R. Feris w. R. Raskar (Mitsubishi Electric Research Labs) and M. Turk)

Direct Multi-Flash Result

Canny Edge Detector

Multi-Flash Application: Finger spelling recognition

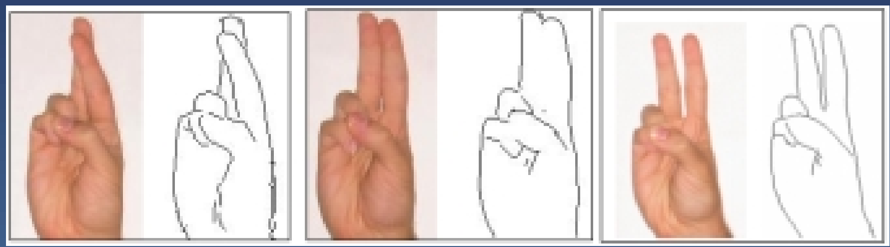
Vision-based finger spelling recognition using depth discontinuities

Part of sign language

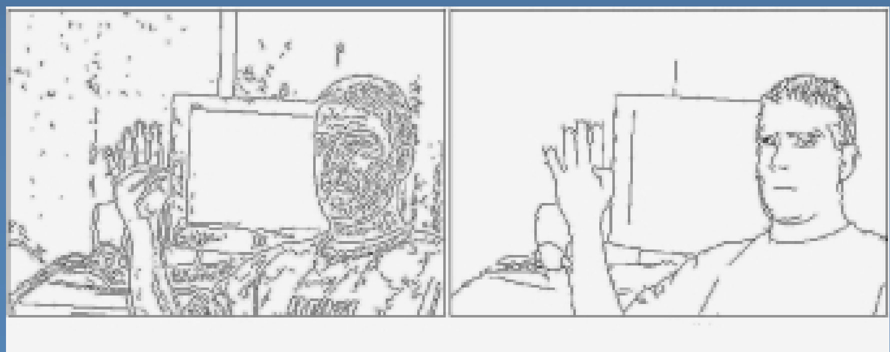
- Lots of occlusion!

Recognition

Worst cases reported by a glove-based system

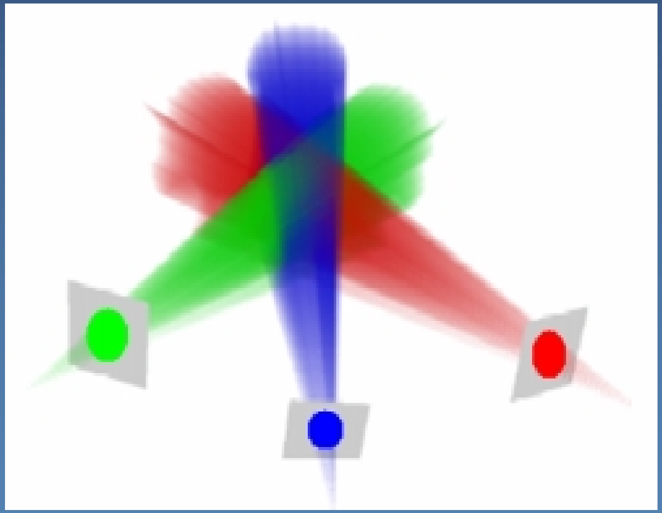


Clutter Removal



96% of correct matches, compared with 88% with Canny edges

Ongoing work: Exploit variable wavelength



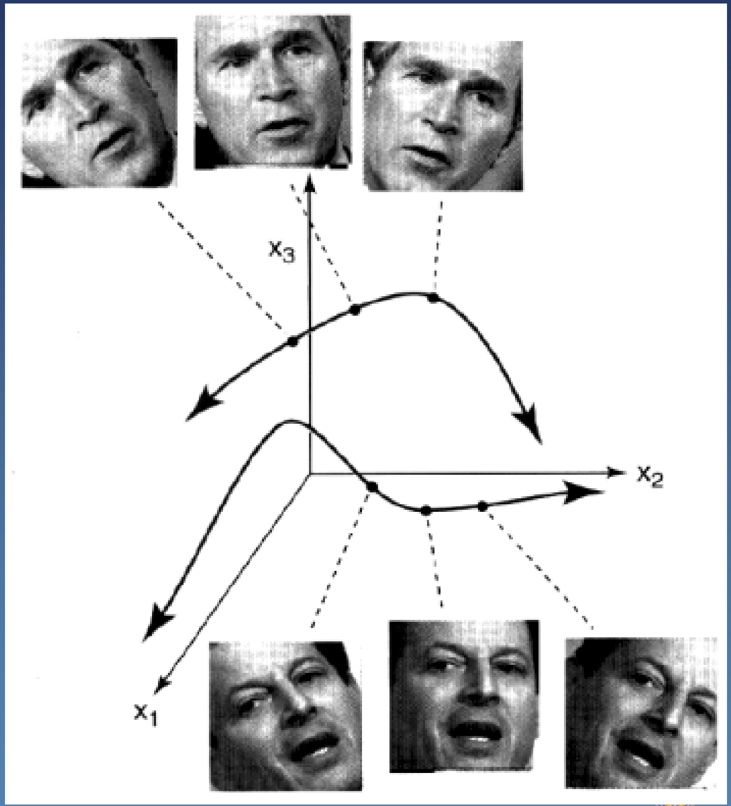
Facial expression analysis
(Y. Chang, M. Turk)

Facial expression <u>representation</u> and <u>visualization</u>

Use non-linear manifolds to represent dynamic facial expressions

Intuition:

- The images of all facial expressions by a person makes a smooth manifold in (high-dimensional) image space, with the "neutral" face as the central reference point

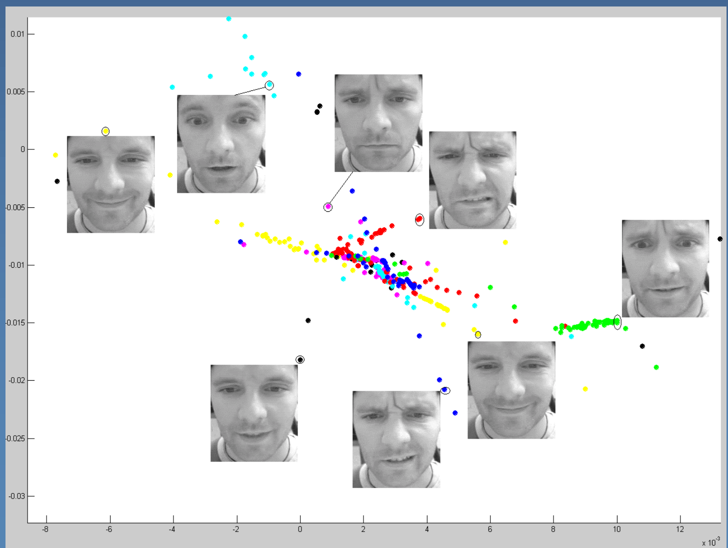


Expression analysis: Non-linear manifold

The manifold is embedded in the high dimensional image space

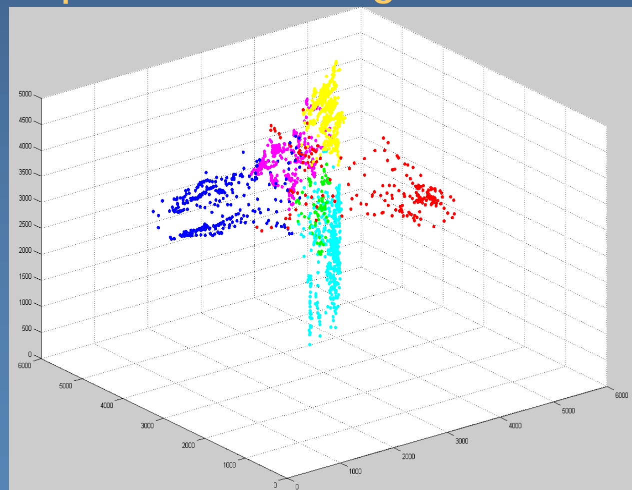
Two techniques:

- Locally linear embedding, Lipschitz embedding



First two dimensions of LLE

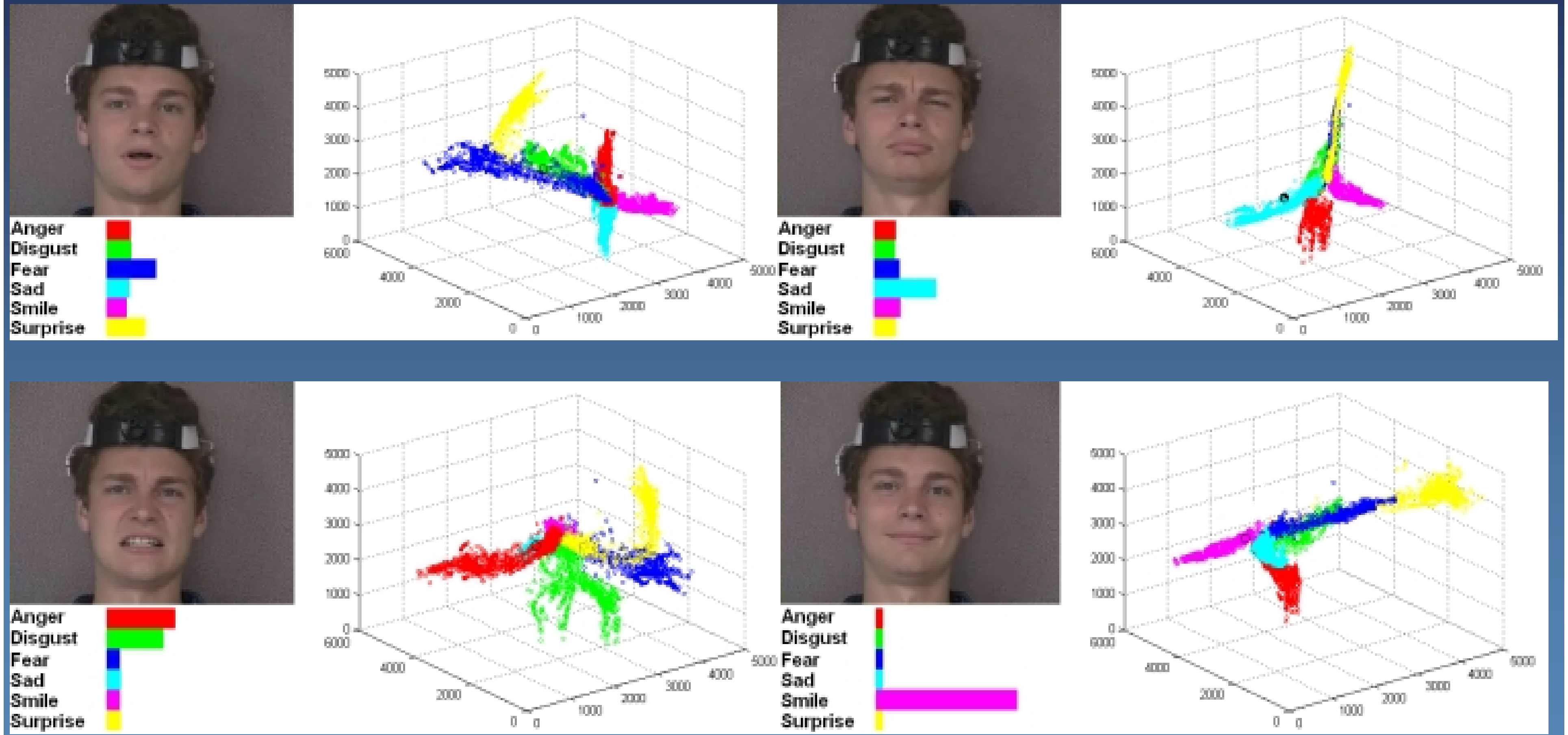
Expression = angle of manifold



First three dimensions of Lipschitz embedding: Good clustering for similar facial expressions

UCSB

Manifold visualization of expression



Example

Constraint-based Interaction with RNA Molecules (Helly Kwee, T. Höllerer, L. Jaeger)

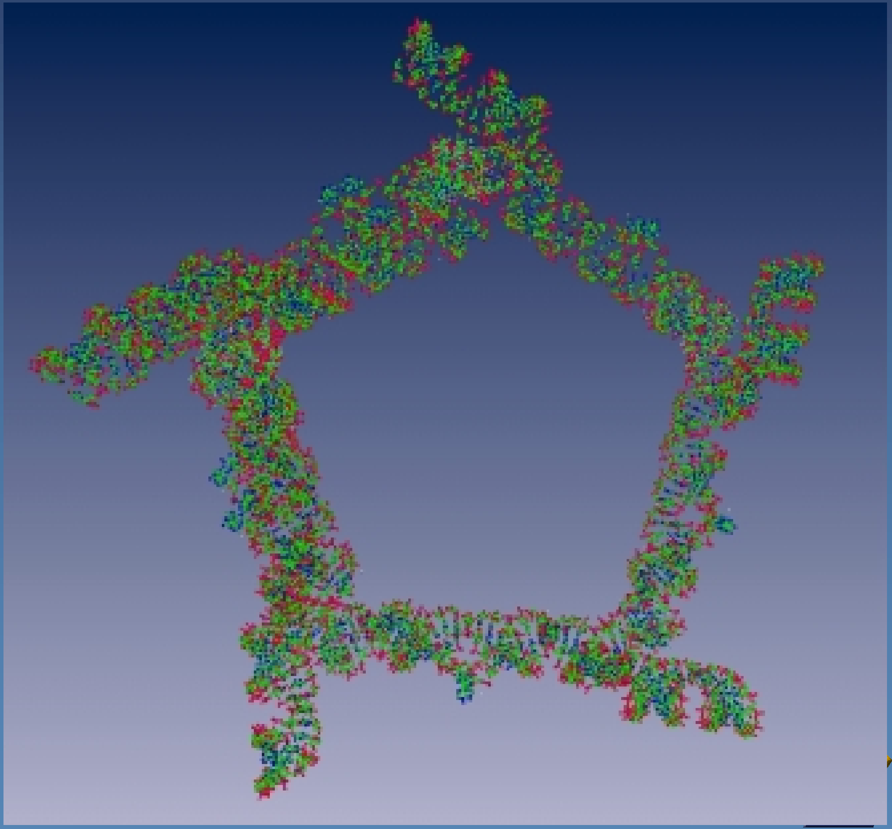
Long-Term Goal:

Assemble new shapes from RNA motifs

"Playing Lego with Biomolecules"

Interaction Research:

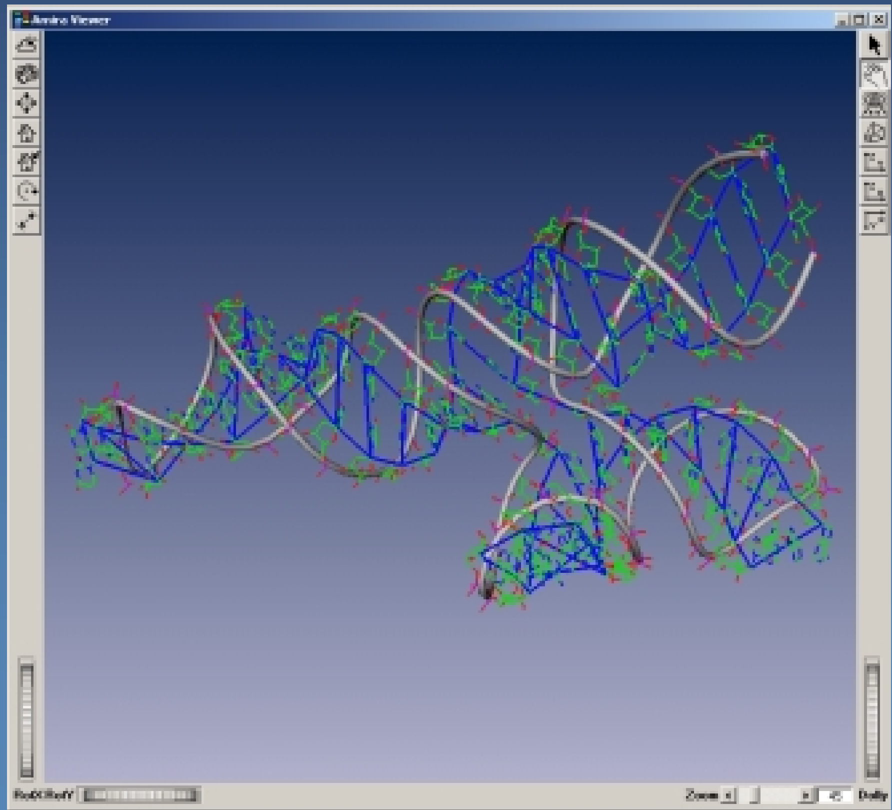
- Simulation Environment for RNA molecule assembly
- How to communicate biochemical constraints in interaction?



UCSB

Constraint-based Interaction with RNA Molecules (Helly Kwee, T. Höllerer, L. Jaeger)

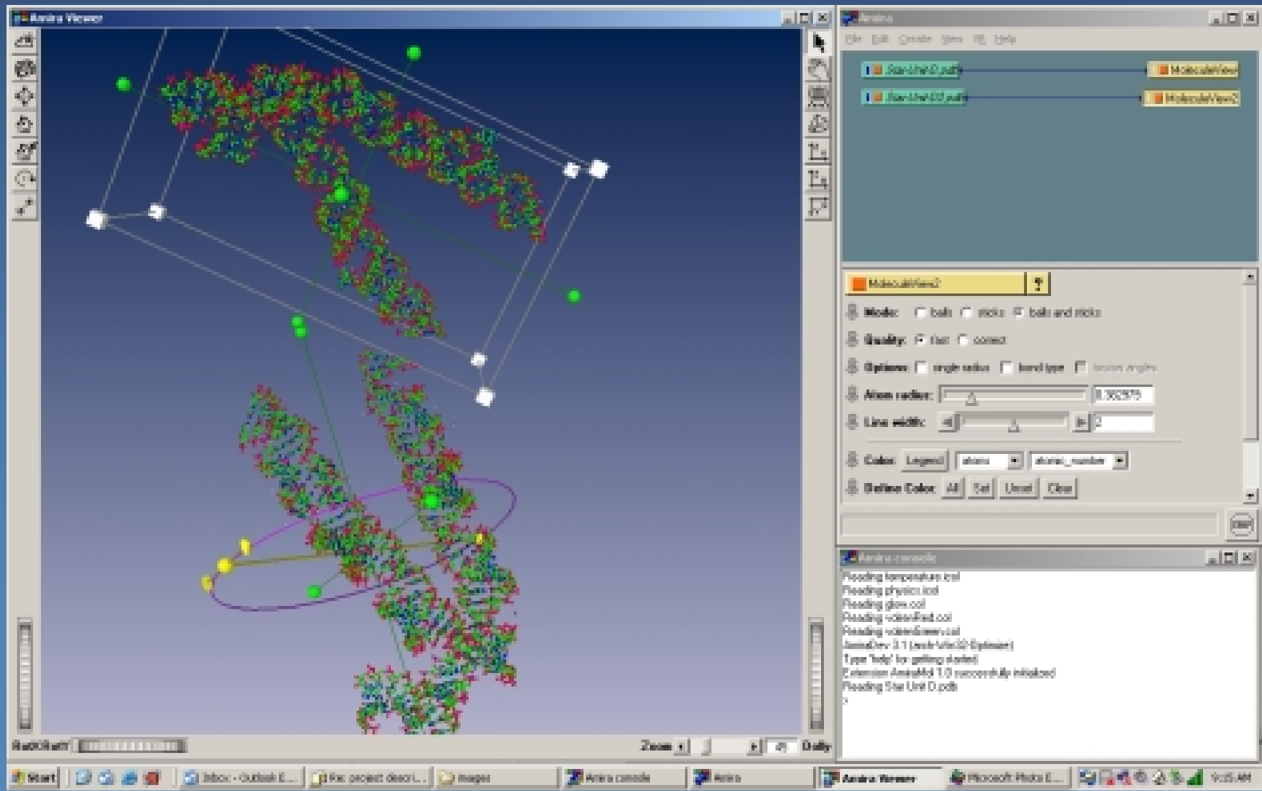
- Emulating Forces as Visualization Constraints
- Compare to Haptic Feedback (Sensable Omni, Delta)
- Interactive Drag'n'Drop 3D Databank of Motifs
- Novel Visualizations



Constraint-based Interaction with RNA Molecules (Helly Kwee, T. Höllerer, L. Jaeger)

International Collaboration with ZIB Berlin, Germany

Integrating new interaction, simulation, and visualization techniques into TGS/Mercury Amira



Outline

- **Lab Overview**
- **Selected Research Projects**
  - Interactive Tools for AR “X-ray vision”
  - HandVu: Hand gestures for mobile computing
  - Photorealistic Real-Time AR in Unprepared Environments
  - Multi-Flash Imaging
  - Facial Expression Analysis
  - Constraint-Based Interaction with RNA Molecules
- **Conclusions**

Conclusions

- Overview of just 6 projects of currently about 20
- Tangible Space very much at the center of our research
- The computer has to be tought how to enhance rather than dominate a space.
- Take into account people's way of interacting with each other

Four Eyes Lab – Summary

General motivation

- Provide better, more compelling HCI technology in many computing environments

Technologies and research

- Fundamental issues in developing robust, real-time, working computer vision technologies for interactive systems
  - Multidisciplinary approach
- Multimodal integration
  - Speech, sound, haptics, user modeling, gestures, visualization
- Main application areas
  - General HCI, entertainment, digital art, visualization,

Ubiquitous Computing, Our Version

… Next comes ubiquitous computing, or the age of *calm technology*, when technology recedes into the background of our lives … (M. Weiser, 1991)

Making UIs ubiquitous:

- Electronically enhance the physical environment
- Electronically enhance the user's perception of the environment
- Enhance the computer's perception of the user