Vision Based Interaction

Matthew Turk

Computer Science Department and
Media Arts and Technology Program

University of California, Santa Barbara

http://www.cs.ucsb.edu/~mturk

Schedule

- Vision based interaction – background and motivation
- VBI-related projects in the Four Eyes Lab
- The Allosphere
- Late afternoon group project

CVPR4HB Mission Statement

A widely accepted prediction is that computing will move to the background, weaving itself into the fabric of our everyday living spaces and projecting the human user into the foreground. To realize this prediction, next-generation computing will need to develop anticipatory user interfaces that are human-centered, built for humans, and based on naturally occurring multimodal human communication. Emerging interfaces will need to include the capacity to understand and emulate human communicative intentions as expressed through behavioral cues such as affective and social signals.

My background

1982	BS, Virginia Tech
1984	MS, Carnegie Mellon University
1984-87	Martin Marietta Aerospace
1991	PhD, MIT Media Lab
1992	Postdoc, LIFIA (Grenoble, France)
1993-94	Teleos Research
1994-2000	Microsoft Research
2000-pres	UC Santa Barbara

My background

1982	BS, Virginia Tech
1984	MS, Carnegie Mellon University
1984-87	Martin Marietta Aerospace
1991	PhD, MIT Media Lab
1992	Postdoc, LIFIA (Grenoble, France)
1993-94	Teleos Research
1994-2000	Microsoft Research
2000-pres	UC Santa Barbara

Robotics and vision

Face recognition

Vision-based interaction, multmodal interfaces

Computer vision, multimodal interfaces, digital media, …

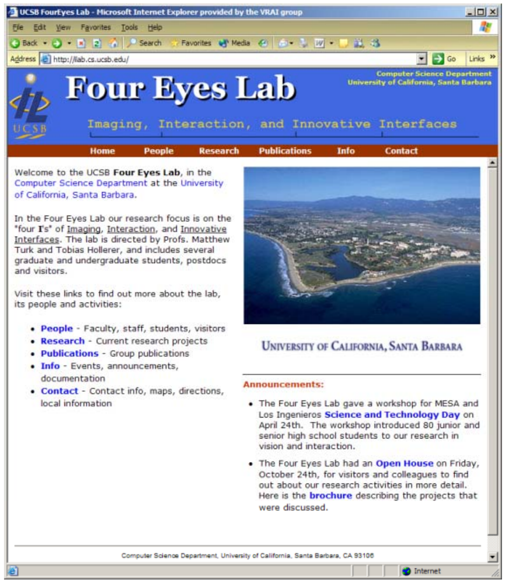
UCSB Four Eyes Lab

4 I’s: **I**maging, **I**nteraction, and **I**nnovative **I**nterfaces

Co-directors: Matthew Turk and Tobias Höllerer

Research in computer vision and human-computer interaction

– Vision based and multimodal interfaces
– Augmented reality and virtual environments
– Mobile human-computer interaction
– Multimodal biometrics
– Novel 3D displays and interaction
– Activity recognition and surveillance
– ….



http://ilab.cs.ucsb.edu

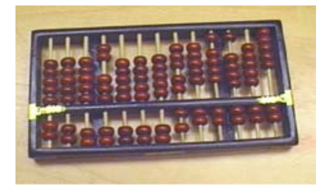
The history of computing

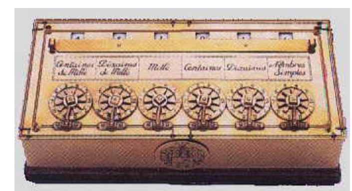
Purposes:

- Counting, manipulating numbers
- Assessing taxes, determining projectiles
- Creating tables of numbers
- Simulation (predicting the weather, the economy, material processes)
- Word processing and spreadsheets
- Email
- Audio + video display
- Mobile, multimedia communication

Form factors:

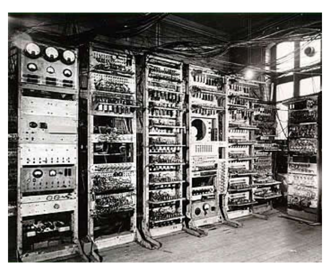
- Mainframes
- Lab computers
- Desktop
- Handheld
- Cell phone
- Immersive
- Wearable





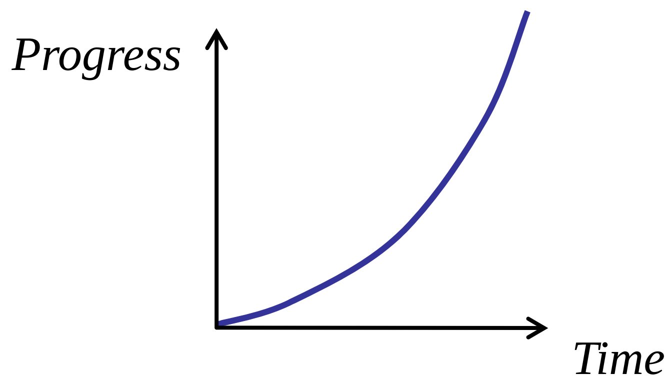
Environments:

- Building
- Laboratory
- Desk
- Coffee shop
- Airport
- Everywhere



Computing has changed...

- Form, function, and context have all changed dramatically
- The *central data element* of computing has evolved:
  - Numbers
  - Text
  - Image
  - Audio+video
  - 3D
  - ...
  - All data underlying communication

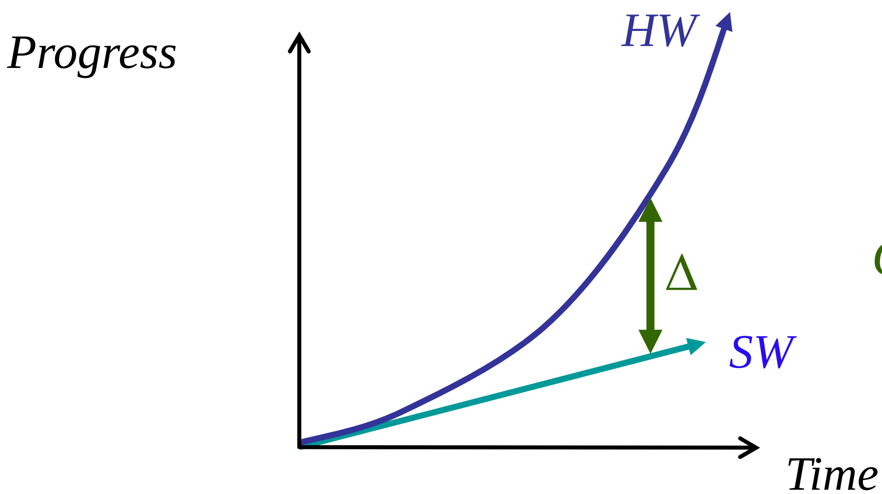


Moore's Law

- What has driven all this?

But there has been no Moore's Law progress for human-computer interaction!

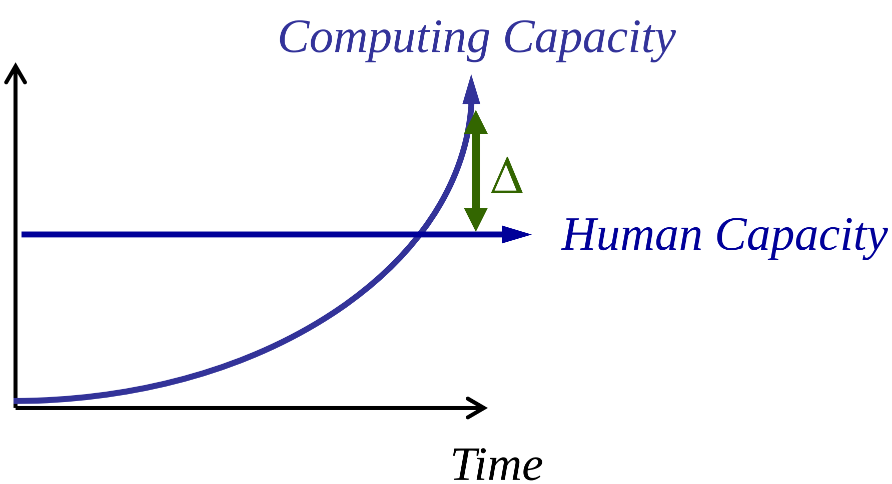
The curse of the delta



Curse of the delta!

Another view:

There’s no Moore’s Law for people!



The result

Video

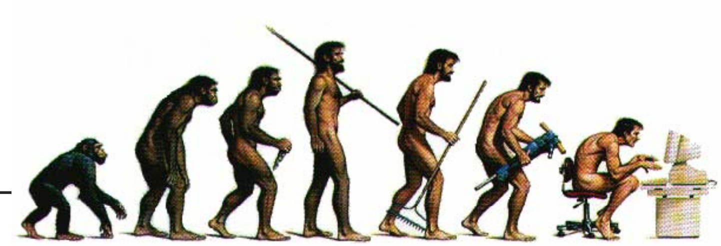
What to do?

- Maybe we need to rethink the way we interact with computers

- Question: What’s the ultimate user interface?
  a) A well-designed machine/instrument
  b) An assistant or butler
  c) None! UIs are a necessary evil
  d) All of the above

- UI Goals:
  – Transparency
  – Minimal cognitive load
  – Task-oriented, not technology-oriented
  – Ease of learning, ease of use (adaptive)

Evolution of user interfaces



When	Implementation	Paradigm
1950s	Switches, punched cards	None
1970s	Command-line interface	Typewriter
1980s	Graphical UI (GUI)	Desktop
2000s	**Perceptual UI (PUI)**	**Natural interaction**

Perceptual Interfaces

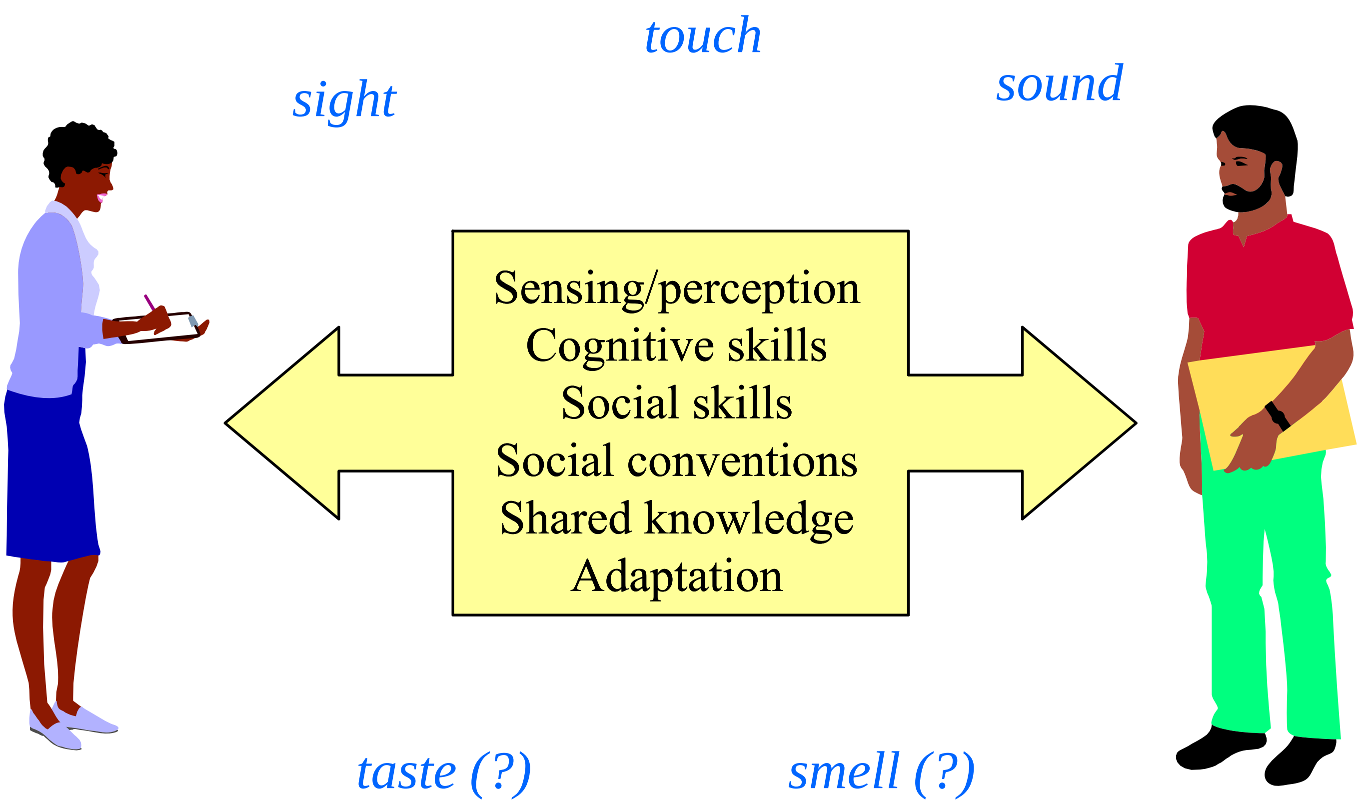
Highly interactive, multimodal interfaces modeled after natural human-to-human interaction

- Goal: For people to be able to interact with computers in a similar fashion to how they interact with each other and with the physical world

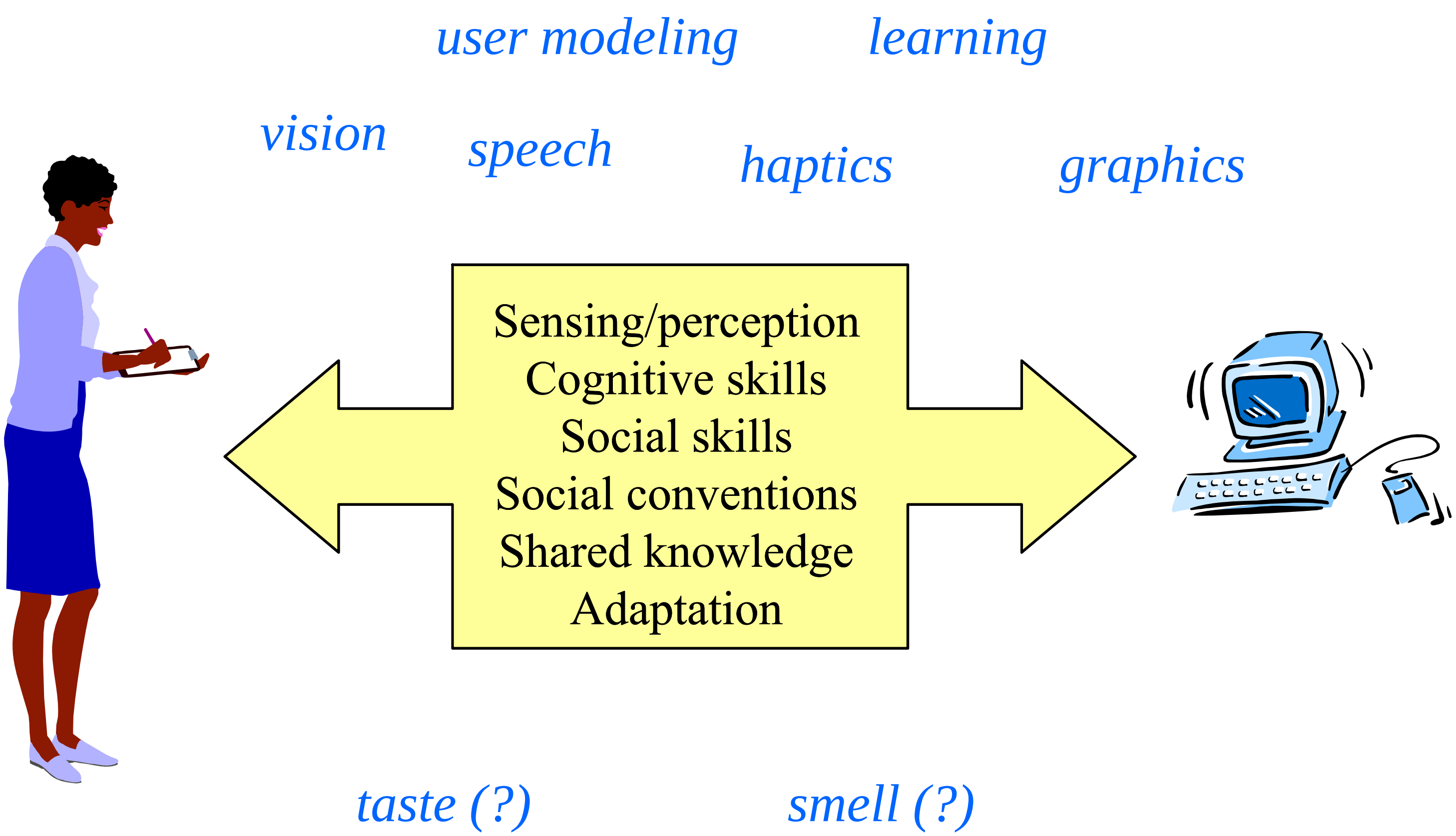
Not just passive

Multiple modalities, not just mouse, keyboard, monitor

Natural human interaction

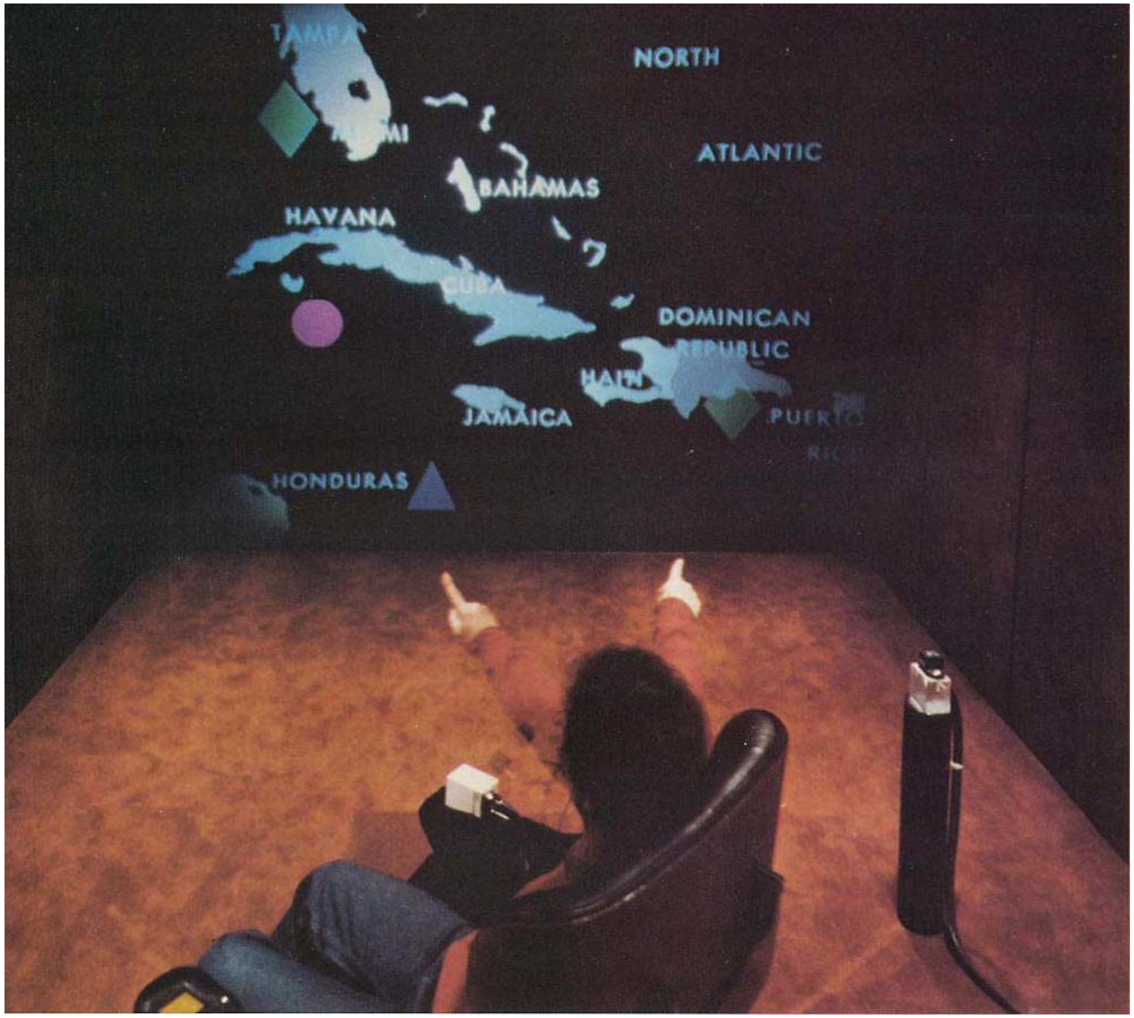


Perceptual and multimodal interaction



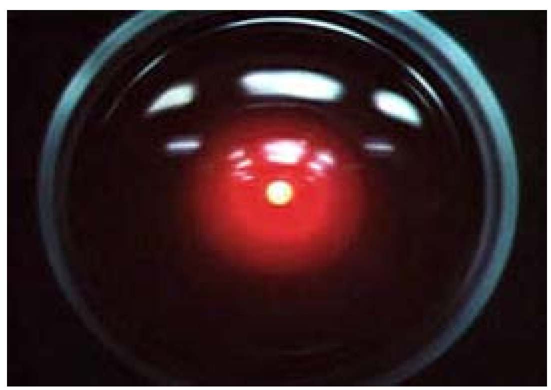
Early example

“Put That There” (Bolt 1980)



Video

Other examples...



Control vs. awareness/context

- Almost all current UI requires explicit (foreground) interaction
  - Intentional control or communication w/ computer
  - Often high physical and cognitive engagement

- Very few examples of system awareness
  - Touching or releasing an input device
  - User presence, location, attention, mood, arousal
  - Back channels of communication (e.g., nodding, “hmm”)

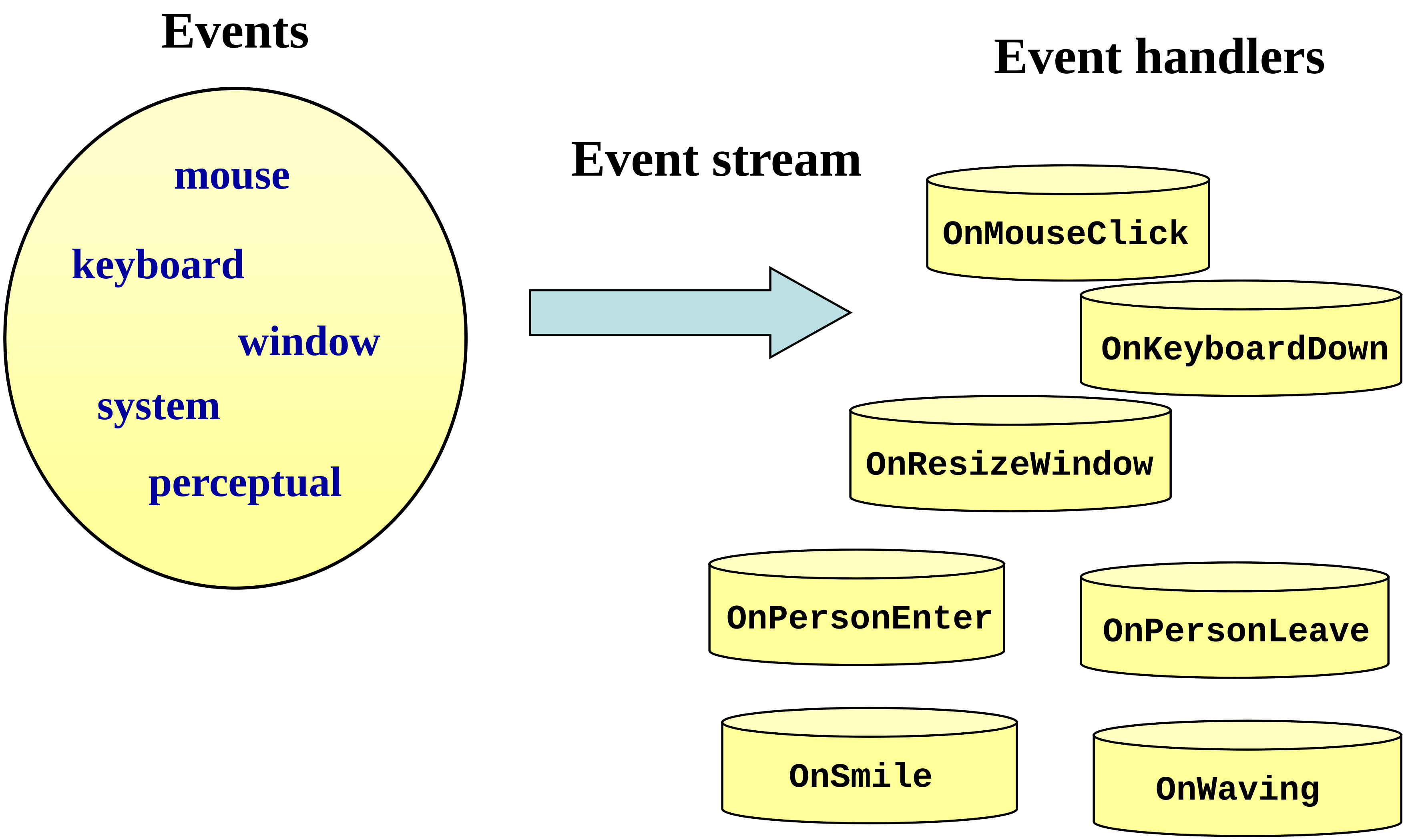
How can achieve the goals of PUI?

- To develop powerful, adaptive, compelling **multimodal interfaces** that reach well beyond the GUI, researchers need to develop and integrate various relevant sensing, display, and interaction technologies, such as:

Speech recognition
Speech synthesis
Natural language processing
Vision (recognition and tracking)
Graphics, animation, visualization

Haptic I/O
Affective computing
Tangible interfaces
Sound recognition
Sound generation
User modeling
Conversational interfaces

A strawman PUI architecture



Strawman PUI

- Superset of the GUI
- Adds perceptual <u>events</u>
- Presents a common, unified approach to PUI-based application development
- Platform opens the door to thousands of developers

Some issues

- Is the event-based model appropriate?
- What defines a perceptual event?
- Is there a useful, reliable subset of perceptual events?
- Non-deterministic events
- Future progress (expanding the event set)
- Input/output modalities? (vision, speech, haptic, *taste, smell?)*
- Allocation of resources
- Multiple goal management
- Training, calibration
- Quality and control of sensors
- Privacy

Direct Manipulation objection

- Shneiderman (and others): HCI should be characterized by
  - direct manipulation
  - predictable interactions
  - giving responsibility to the users
  - giving users a sense of accomplishment
- Argument against intelligent, adaptive, agent-based, and anthropomorphic interfaces – and PUI
- ... But is it really **either/or**? Perhaps not.

PUI/multimodal interface research status

- Young field
- Growing interest
- Resonates with researchers with a wide range of interests (not just HCI researchers, or vision researchers, or …)
- Mixing up the “gene pool “
- Many existing projects and research efforts
- But … still asking basic questions
- Still narrow participation (but growing)

PUI, MLMI, ICMI

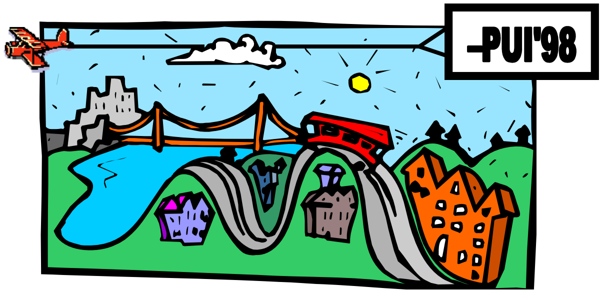
ICMI (1996, 1999, 2000, 2002-2010)
PUI Workshop (1997, 1998, 2001)
MLMI (2004-2008)

http://www.acm.org/icmi

2001 PUI

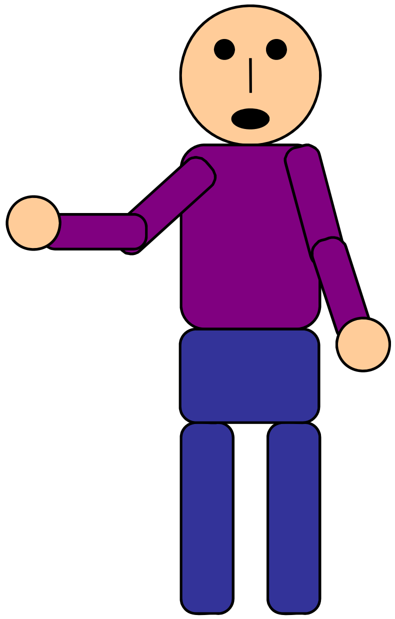
ICMI
SUPPORTING HUMAN PERFORMANCE AND EXPRESSION

ICMI 2003



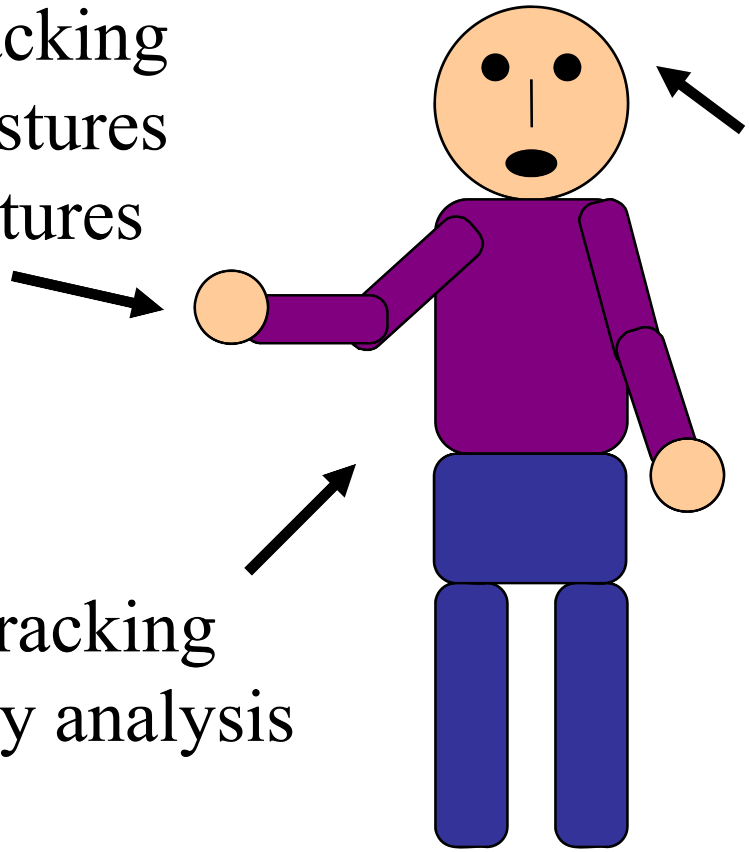
Vision Based Interfaces (VBI)

- Visual cues are important in communication!
- Useful visual cues
  - Presence
  - Location
  - Identity (and age, sex, nationality, etc.)
  - Facial expression
  - Body language
  - Attention (gaze direction)
  - Gestures for control and communication
  - Lip movement
  - Activity



VBI – using computer vision to perceive these cues

Elements of VBI



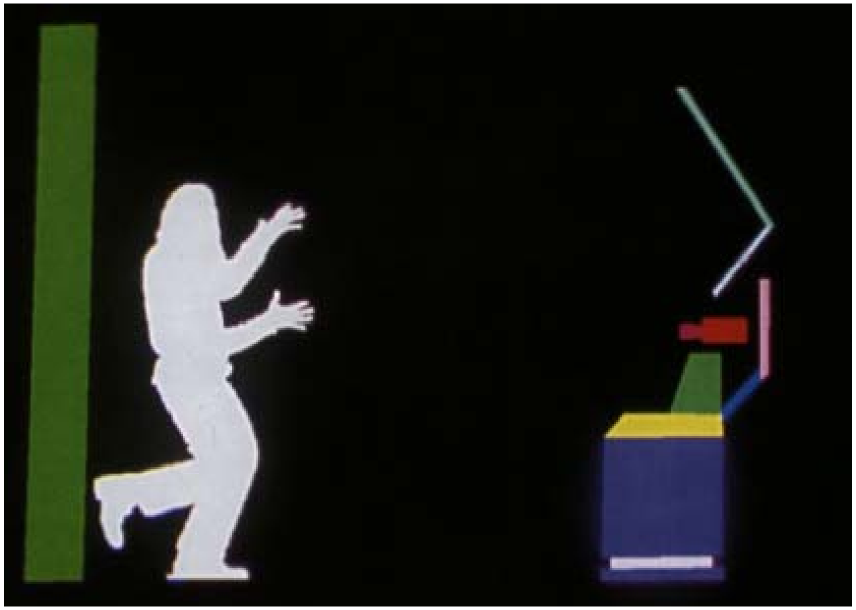
Some VBI application areas

- Accessibility, hands-free computing
- Entertainment and gaming
- Interactive art
- Social interfaces/agents
- Teleconferencing
- Improved speech recognition (speechreading)
- User-aware applications
- Intelligent environments
- Biometrics
- Movement analysis (medicine, sports)
- Visualization environments

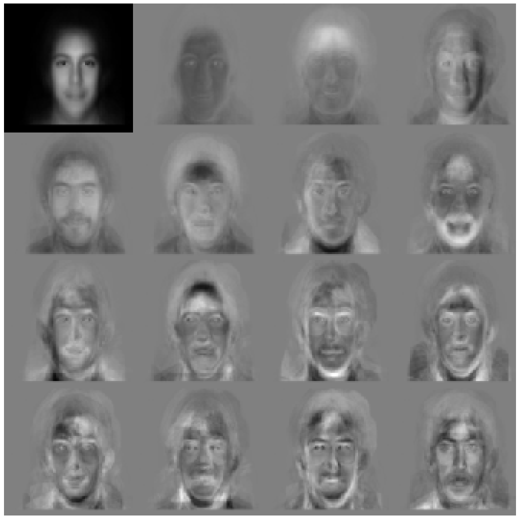
What makes VBI difficult?

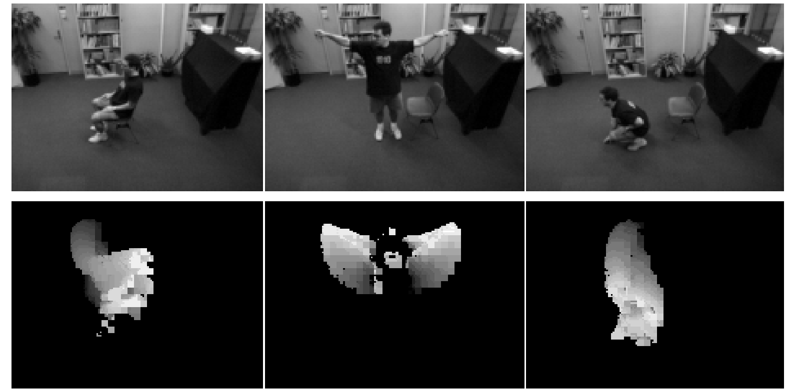
- User appearance
  - size, sex, race, hair, skin, make-up, fatigue, clothing color & fit, facial hair, eyeglasses, aging….
- Environment
  - lighting, background, movement, camera
- Multiple people and occlusion
- Intentionality of actions (ambiguity)
- Speed and latency
- Calibration, FOV, camera control, image quality

Some VBI examples

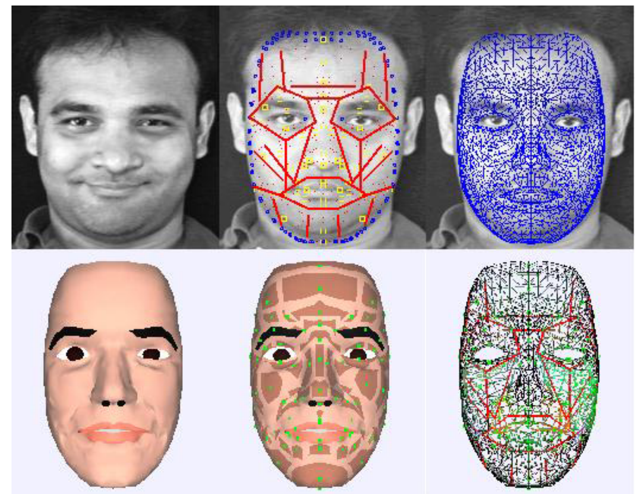


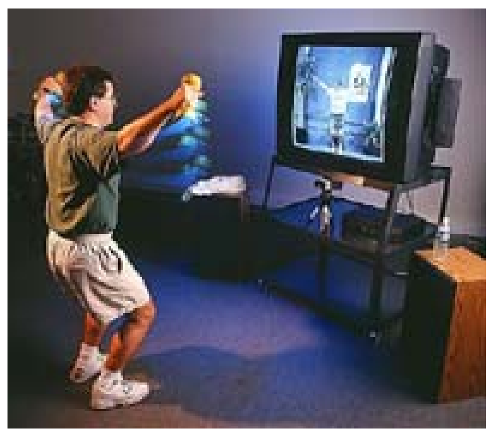
Myron Krueger
1980s

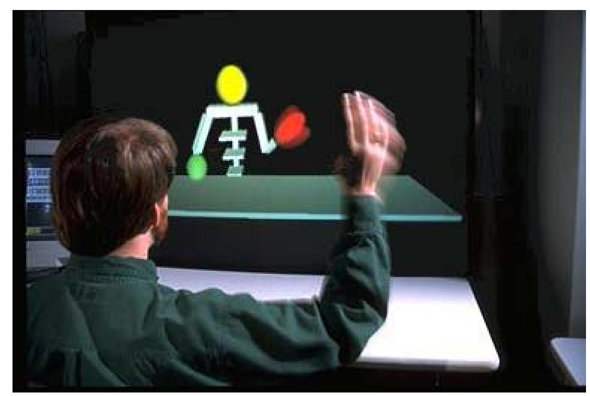




MIT Media Lab
1990s







HMM based ASL recognition

Video

The KidsRoom

Video

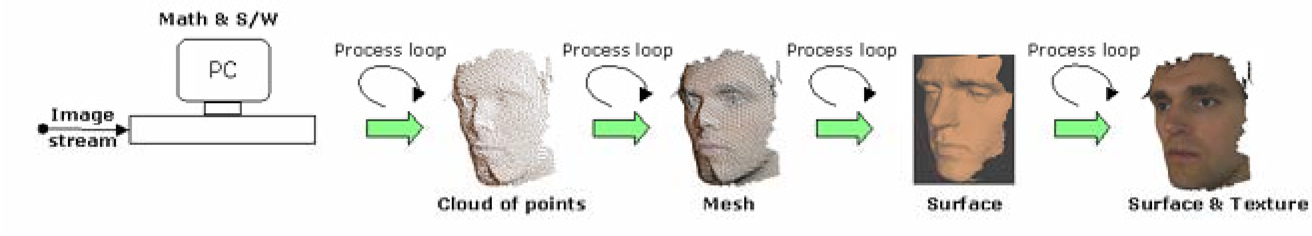
Interaction using hand tracking

Video

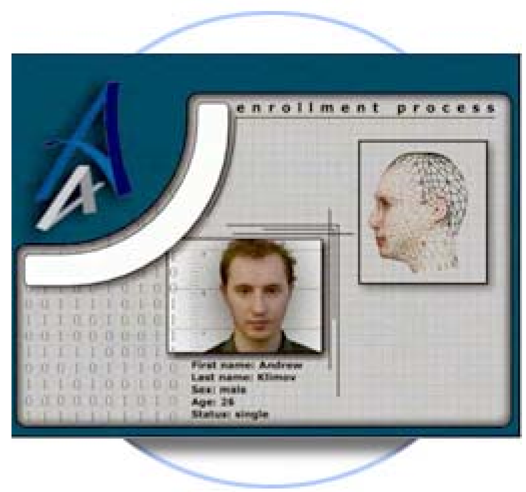
Gesture recognition

Video

Video



Commercial systems 2000s



Sony EyeToy

Video

Reactrix

Video

Microsoft Kinect (Project Natal)

- RGB camera, depth sensor, and microphone array in one package
  - Xbox add-on
  - RGB: 640x480, 30Hz
  - Depth: 320x240, 16-bit precision, 1.2-3.5m
- Capabilities
  - Full-body 3D motion capture and gesture recognition
    - Two people, 20 joints per person (??)
    - Track up to six people
  - Face recognition
  - Voice recognition, acoustic source localization

Video

Where we are today

- Perceptual interfaces
  - Progress in component technologies (speech, vision, haptics, …)
  - Some multimodal integration
  - Growing area, but still a small part of HCI
- Vision based interfaces
  - Solid progress towards robust real-time visual tracking, modeling, and recognition of humans and their activities
  - Some first generation commercial systems available
  - Still too brittle
- Big challenges
  - Serious approaches to modeling user and context
  - Interaction among modalities (except AVSP)
  - Compelling applications

Moore's Law progress

Year 1975

0.001 CPU cycles/pixel of video stream

Year 2000

57 cycles/pixel

Year 2025

3.7M cycles/pixel
(64k x speedup)

Killer app?

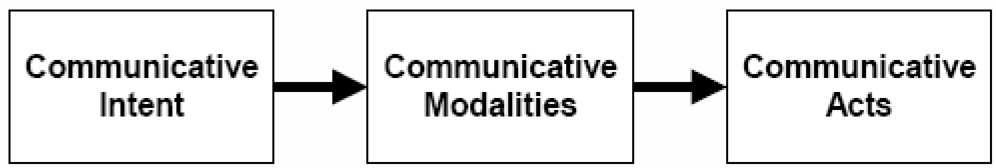
- Is there a "killer app" for vision-based interaction?
  - An application that will economically drive and justify extensive research and development in automatic gesture analysis
  - Fills a critical void or creates a need for a new technology
- Maybe not, but there are, however, <u>many</u> practical uses
  - Many that combine modalities, not vision-only
- This is good!!
  - It gives us the opportunity to do the **right thing**
    - **The science of interaction**
      - Fundamentally multimodal
      - Understanding people, not just computers
      - Involves CS, human factors, human perception, ….

Some relevant questions about gesture

- What is a gesture?
  - Blinking? Scratching your chin? Jumping up and down? Smiling? Skipping?
- What is the purpose of gesture?
  - Communication? Getting rid of an itch? Expressing feelings?
- What does it mean to do gesture recognition?
  - Just classification? ("Gesture #32 just occurred")
  - Semantic interpretation? ("He is waving goodbye")
- What is the context of gesture?
  - A conversation? Signaling? General feedback? Control?
  - How does context affect the recognition process?

Gestures

- A gesture is the act of expressing communicative intent via one or more modalities



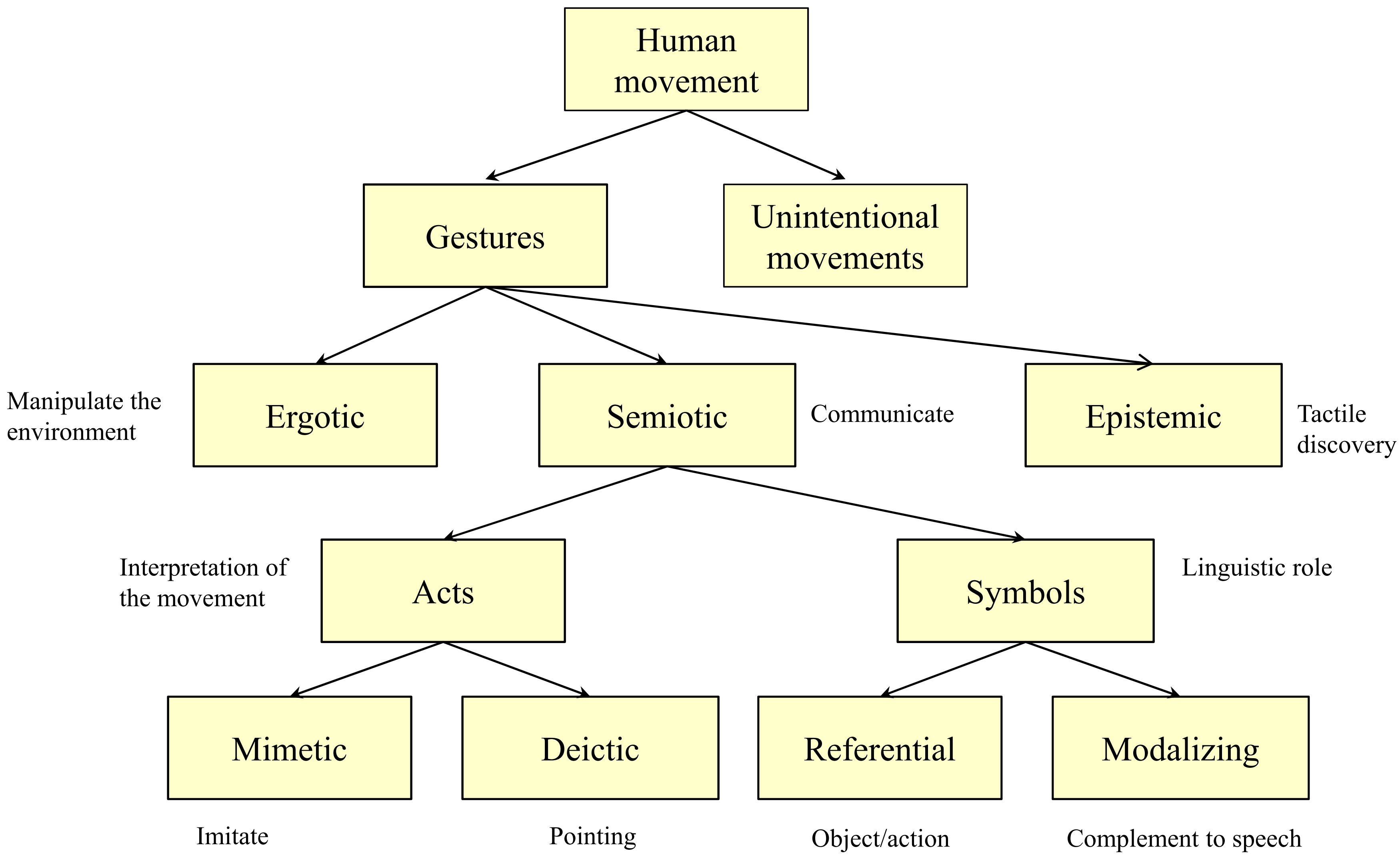
- Hand and arm gestures
  - Hand poses, signs, trajectories…
- Head and face gestures
  - Head nodding or shaking, gaze direction, winking, facial expressions
- Body gestures: involvement of full body motion
  - One or more people

Gestures (cont.)

- Aspects of a gesture which may be important to its meaning:
  - Spatial information: where it occurs
  - Trajectory information: the path it takes
  - Symbolic information: the sign it makes
  - Affective information: its emotional quality

- Some tools for gesture recognition
  - HMMs
  - State estimation via particle filtering
  - Finite state machines
  - Neural networks
  - Manifold embedding
  - Appearance-based vs. (2D/3D) model-based

A gesture taxonomy



Kendon's gesture continuum

- Gesticulation
  - Spontaneous movements of the hands and arms that accompany speech
- Language-like gestures
  - Gesticulation that is integrated into a spoken utterance, replacing a particular spoken word or phrase
- Pantomimes
  - Gestures that depict objects or actions, with or without accompanying speech
- Emblems
  - Familiar gestures such as "V for victory", "thumbs up", and assorted rude gestures (these are often culturally specific)
- Sign languages
  - Well-defined linguistic systems, such as ASL

McNeill’s gesture types

- Within the first category – spontaneous, speech-associated gesture – McNeill defined four gesture types:
  - **Iconic** – representational gestures depicting some feature of the object, action or event being described
  - **Metaphoric** – gestures that represent a common metaphor, rather than the object or event directly
  - **Beat** – small, formless gestures, often associated with word emphasis
  - **Deictic** – pointing gestures that refer to people, objects, or events in space or time.

Gesture and context

- *Context* underlies the relationship between gesture and meaning
- Except in limited special cases, we can't understand gesture (derive meaning) apart from its context
- We need to understand both gesture **production** and gesture **recognition** together (not individually)
- That is, "gesture recognition" research <u>by itself</u> is, in the long run, a dead end
  - It will lead to mostly impractical toy systems!

So… the bottom line

- Gesture recognition is not just a technical problem in Computer Science
- A multidisciplinary approach is vital to truly "solve" gesture recognition – to understand it deeply
  - "Thinkers" and "builders" need to work together
- Still, there is low-hanging fruit to be had, where specific gesture-based technologies can be useful before all the Big Problems are solved
  - (Good…!)

Guidelines for gestural interface design

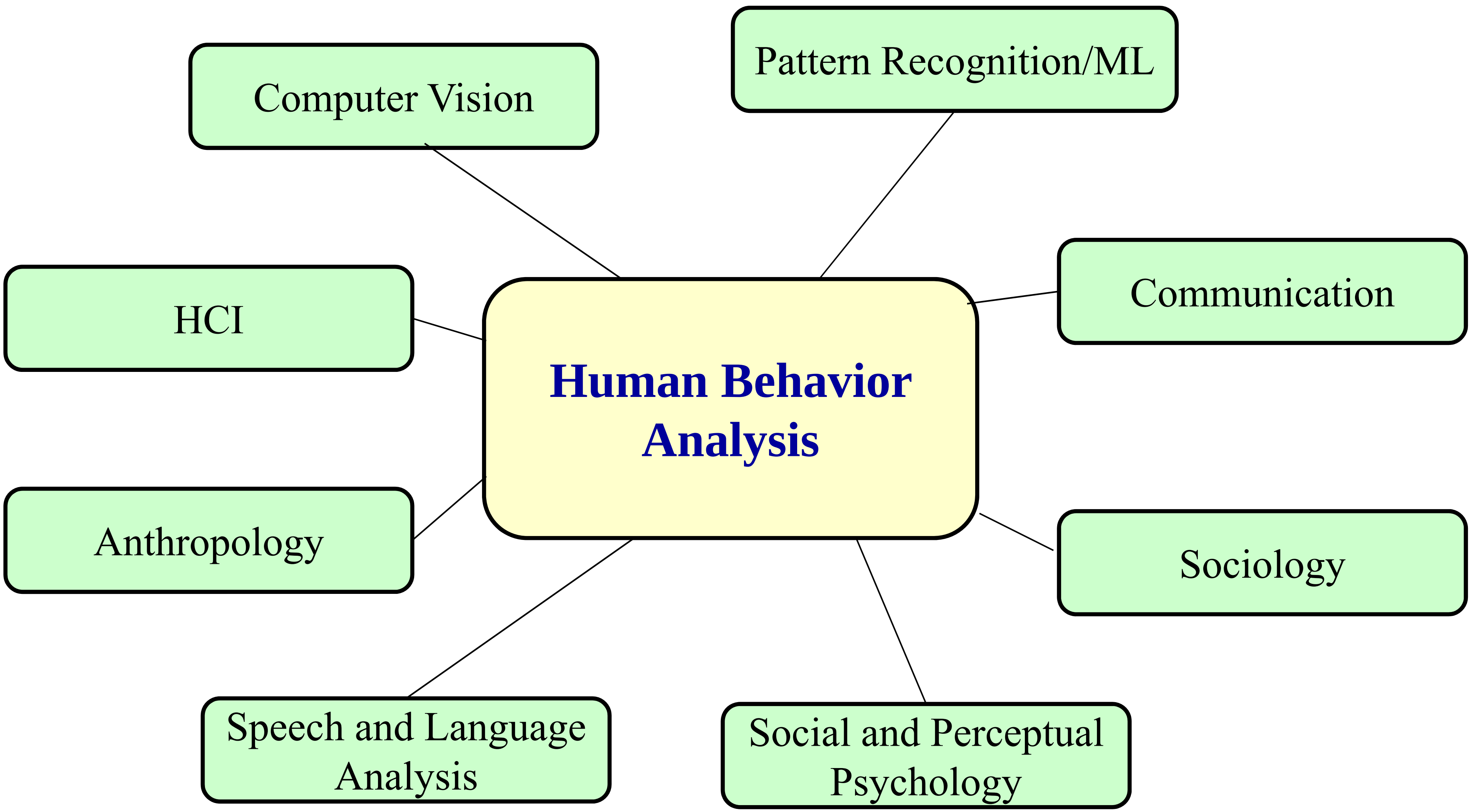
- **Inform the user.** People use different kinds of gestures for many purposes, from spontaneous gesticulation associated with speech to structured sign languages. Similarly, gesture may play a number of different roles in a virtual environment. To make compelling use of gesture, the types of gestures allowed and what they effect must be clear to the user.
- **Give the user feedback.** Feedback is essential to let the user know when a gesture has been recognized. This could be inferred from the action taken by the system, when that action is obvious, or by more subtle visual or audible confirmation methods.
- **Take advantage of the uniqueness of gesture.** Gesture is not just a substitute for a mouse or keyboard.
- **Understand the benefits and limits of the particular technology.** For example, precise finger positions are better suited to data gloves than vision-based techniques. Tethers from gloves or body suits may constrain the user’s movement.

Guidelines for gestural interface design (cont.)

- **Do usability testing on the system.** Don't just rely on the designer's intuition.
- **Avoid temporal segmentation if feasible.** At least with the current state of the art, segmentation of gestures can be quite difficult.
- **Don't tire the user.** Gesture is seldom the primary mode of communication. When a user is forced to make frequent, awkward, or precise gestures, the user can become fatigued quickly. For example, holding one's arm in the air to make repeated hand gestures becomes tiring very quickly.
- **Don't make the gestures to be recognized too similar.** For ease of classification and to help the user.
- **Don't use gesture as a gimmick.** If something is better done with a mouse, keyboard, speech, or some other device or mode, use it – extraneous use of gesture should be avoided.

Guidelines for gestural interface design (cont.)

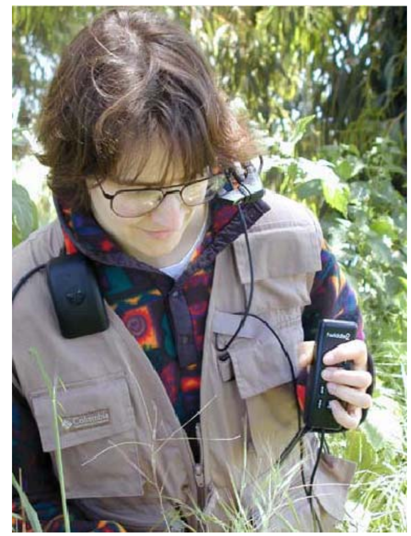
- **Don’t increase the user’s cognitive load.** Having to remember the whats, wheres, and hows of a gestural interface can make it oppressive to the user. The system’s gestures should be as intuitive and simple as possible. The learning curve for a gestural interface is more difficult than for a mouse and menu interface, since it requires *recall* rather than just *recognition among a list.*
- **Don’t require precise motion.** Especially when motioning in space with no tactile feedback, it is difficult to make highly accurate or repeatable gestures.
- **Don’t create new, unnatural gestural languages.** If it is necessary to devise a new gesture language, make it as intuitive as possible.


Human Behavior Analysis
Computer Vision
Pattern Recognition/ML
HCI
Communication
Anthropology
Sociology
Speech and Language Analysis
Social and Perceptual Psychology

Some VBI-related research at the UCSB Four Eyes Lab

HandVu: Gestural interface for mobile systems

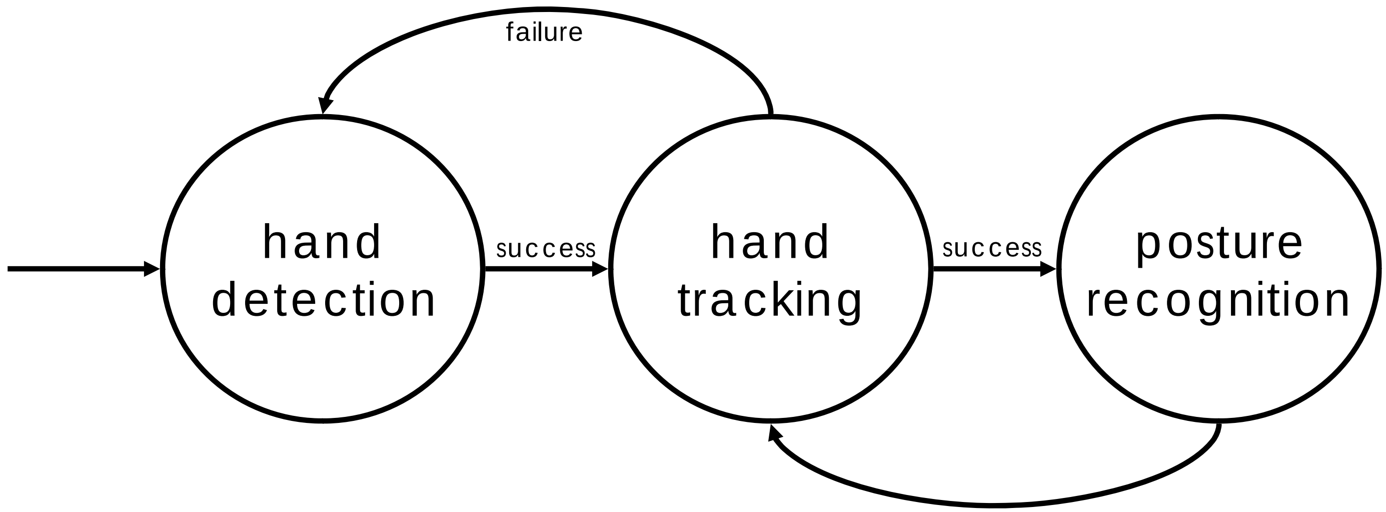
- Goal: To build highly robust CV methods that allow out-of-the-box use of hand gestures as an interface modality for mobile computing environments



System components

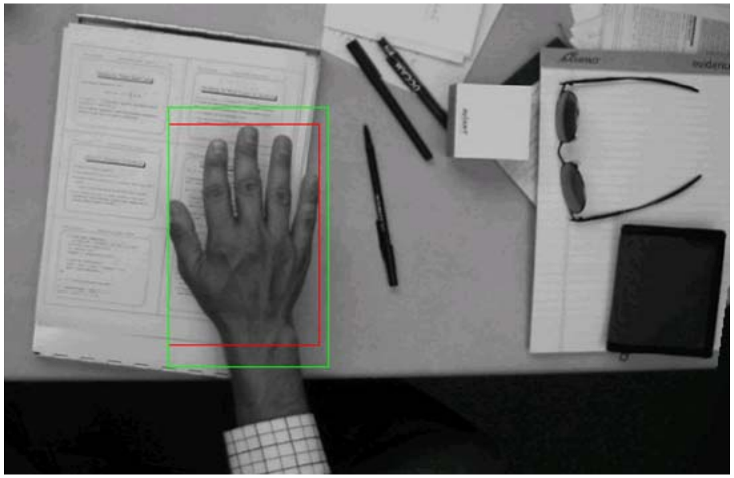
- Detection
  - Detect the presence of a hand in the expected configuration and image position
- Tracking
  - Robustly track the hand, even when there are significant changes in posture, lighting, background, etc.
- Posture/gesture recognition
  - Recognize a small number of postures/gestures to indicate commands or parameters
- Interface
  - Integrate the system into a useful user experience

HandVu


failure
hand detection
success
hand tracking
success
posture recognition

Robust hand detection

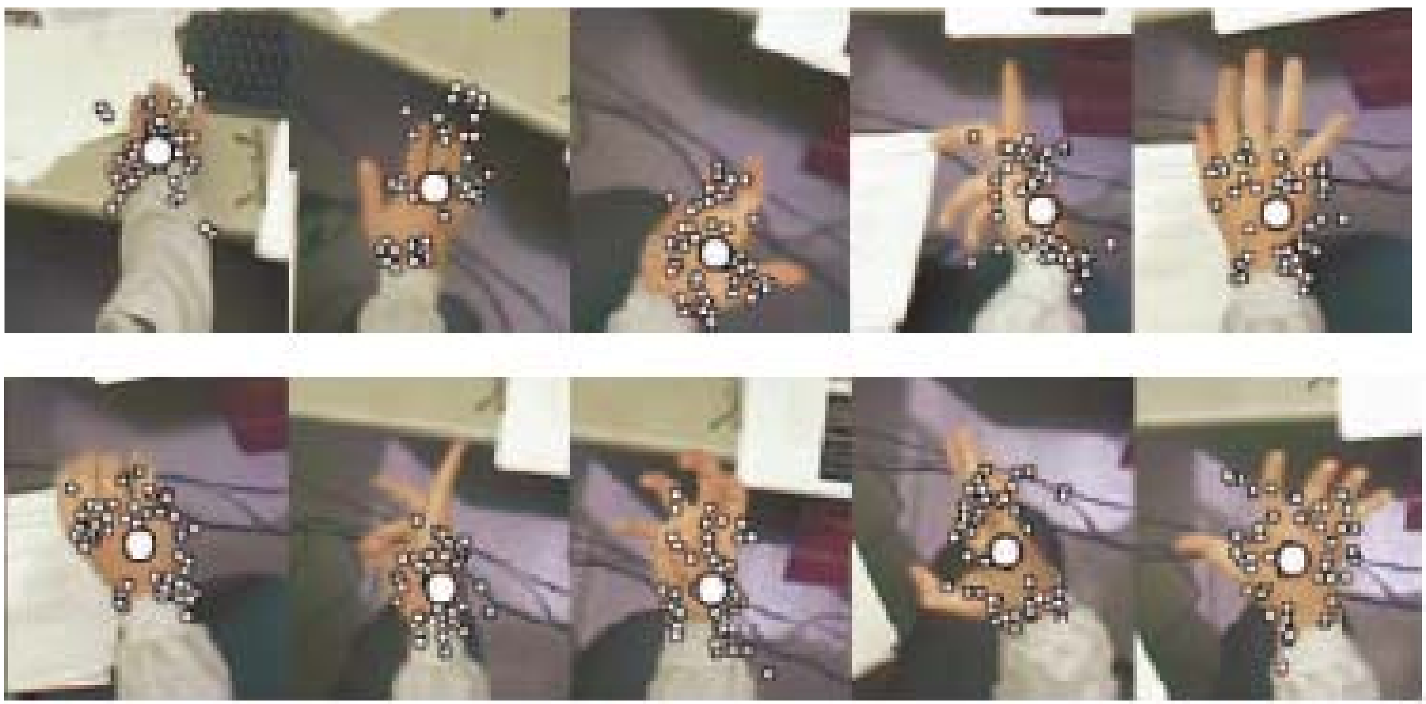
- Detection using a modified version of the Jones-Viola face detector, based on boosted learning



- Performance:
  - Detection rate: 92%
  - False positive (fp) rate:
    - 1.01×10^{-8}
    - One false positive in 279 VGA-sized image frames
  - With color verification: few false positives per hour of live video!

Hand tracking

- “Flocks of Features”
  - Fast 2D tracking method for non-rigid and highly articulated objects such as hands
  - KLT features + foreground color model



Tracking

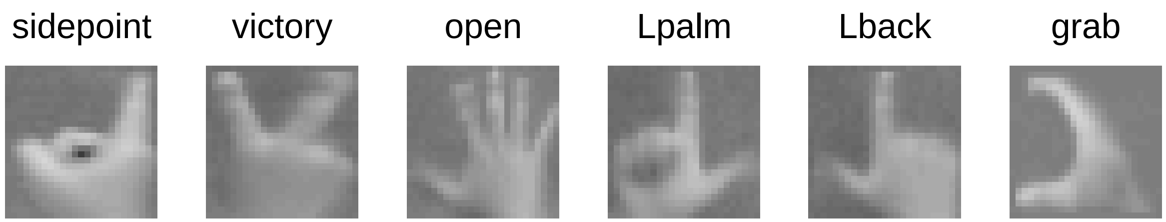
Video

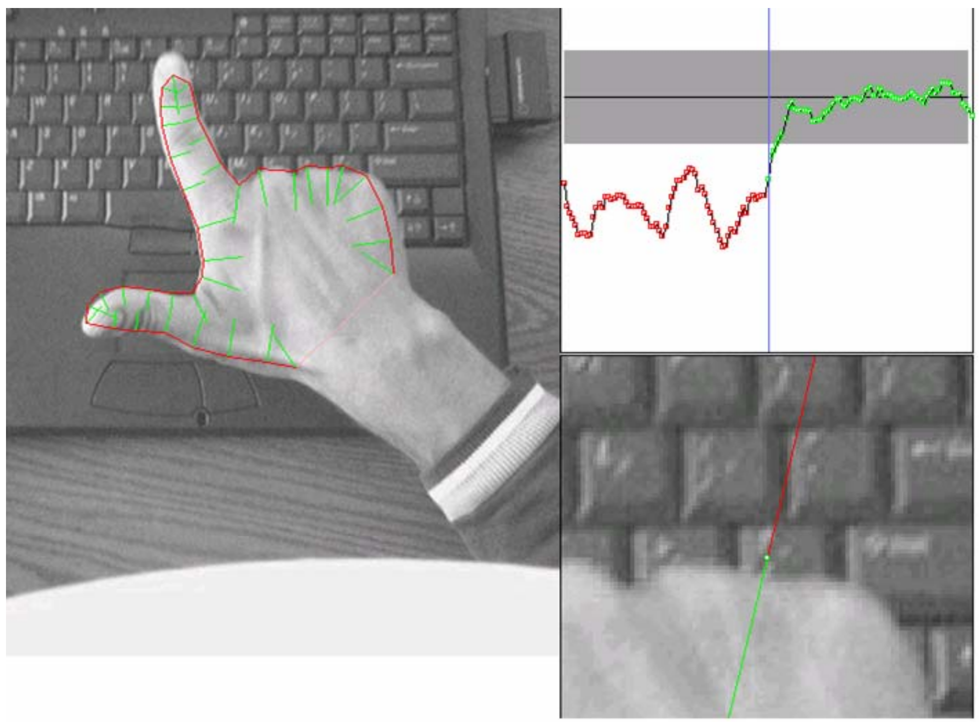
HandVu application

Video

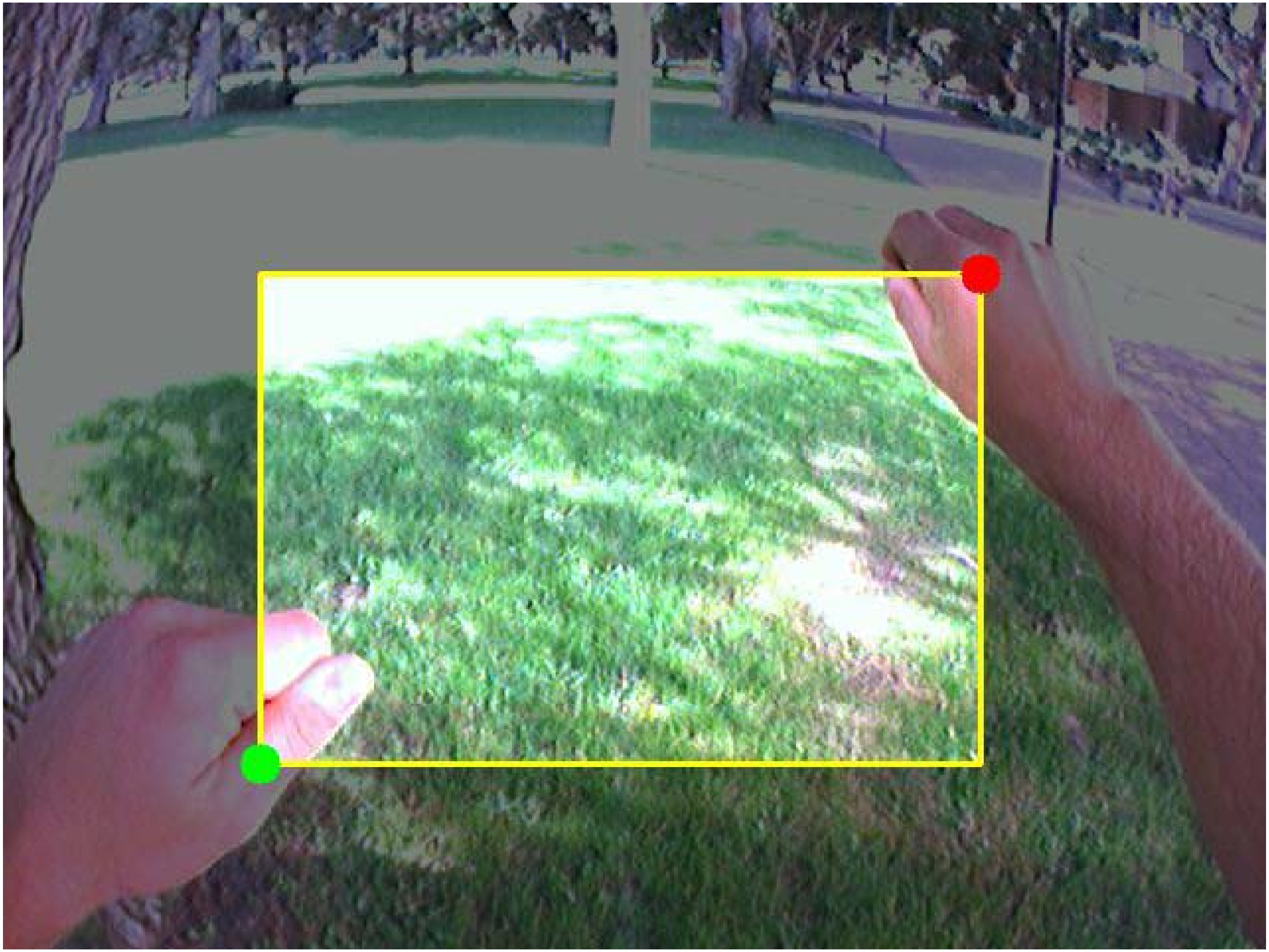
Gesture recognition

- Really view-dependent **posture** recognition
  - Recognizes six hand postures





Driving a user interface



An AR application

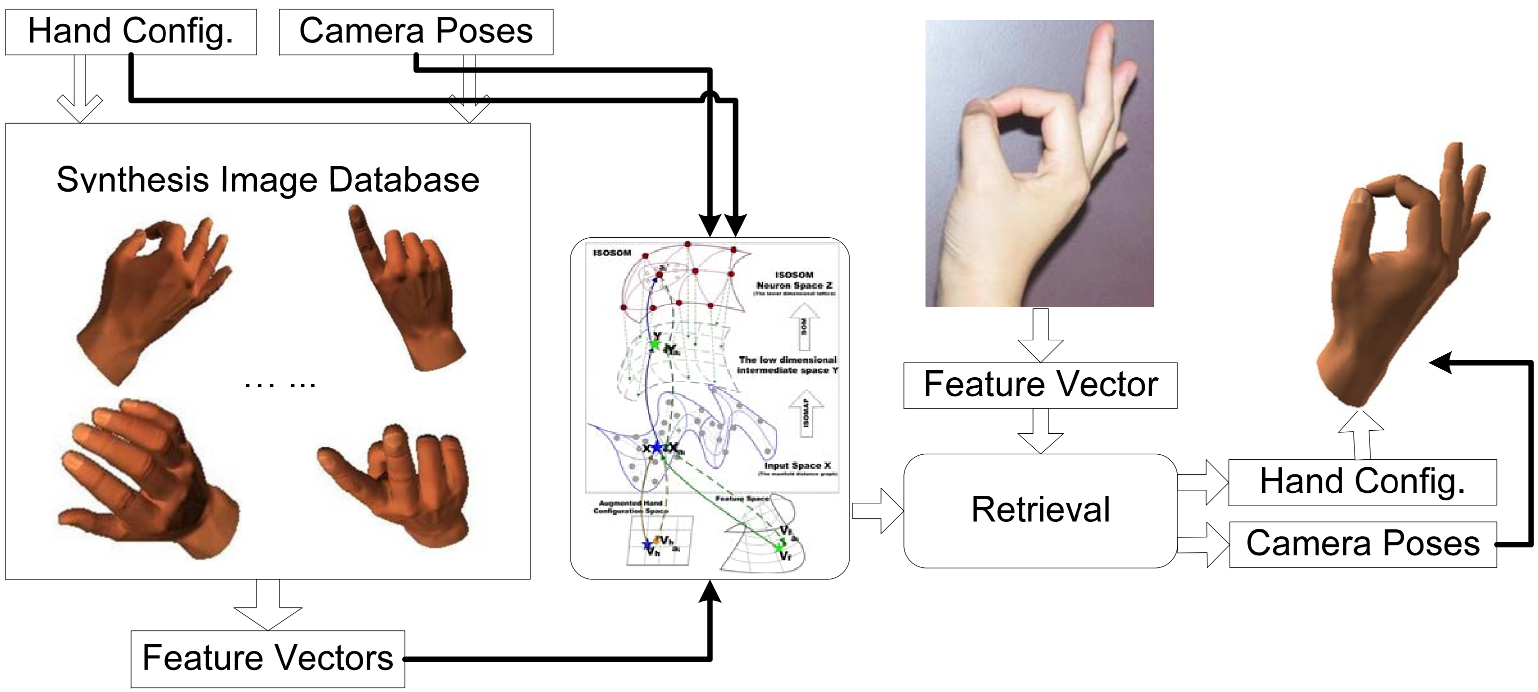
HandVu software

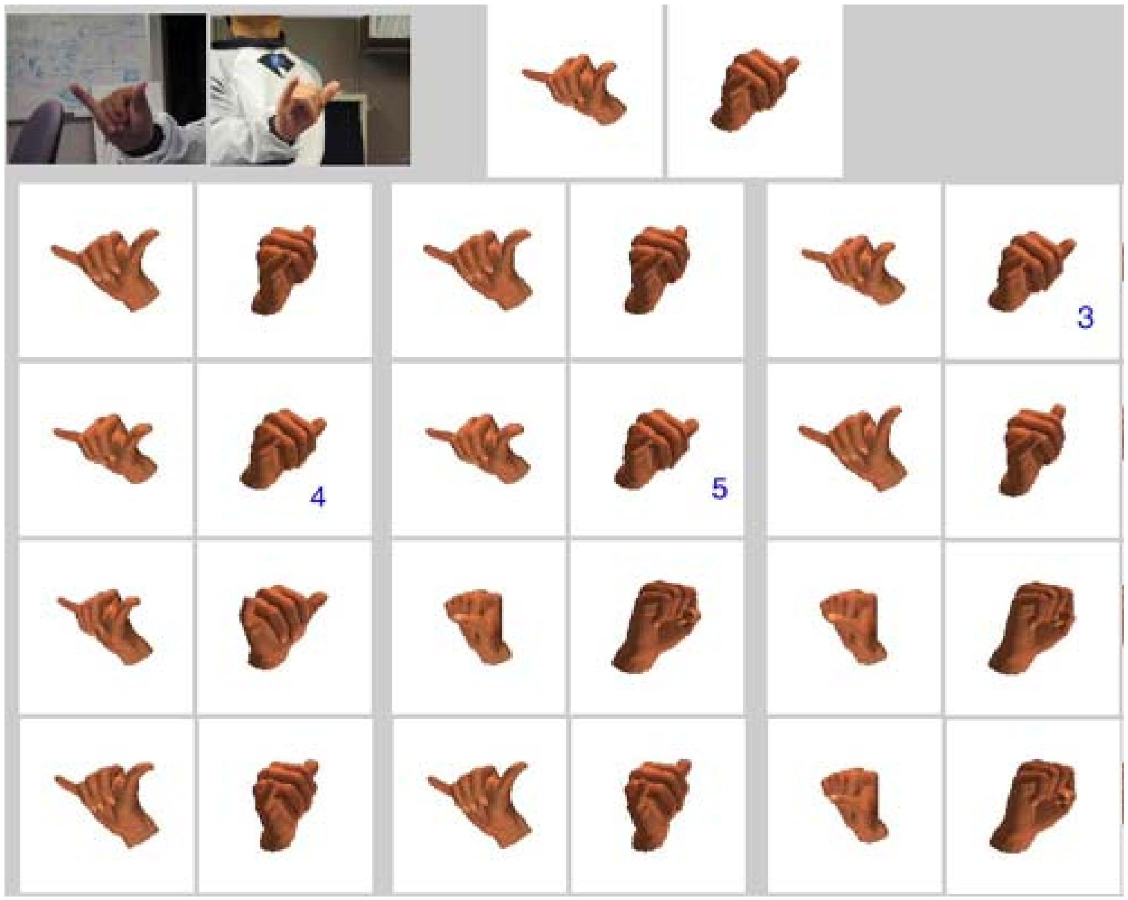
Google: “HandVu”

- A library for hand gesture recognition
  - A toolkit for out-of-the-box interface deployment
- Features:
  - User independent
  - Works with any camera
  - Handles cluttered background
  - Adjusts to lighting changes
  - Scalable with image quality and processing power
  - Fast: 5-150ms per 640x480 frame (on 3GHz)
- Source/binary available, built on OpenCV

Multiview 3D hand pose estimation

- Appearance based approach to hand pose estimation
  - Based on ISOSOM (ISOMAP + SOM) nonlinear mapping
- A MAP framework is used to fuse view information and bypass 3D hand reconstruction

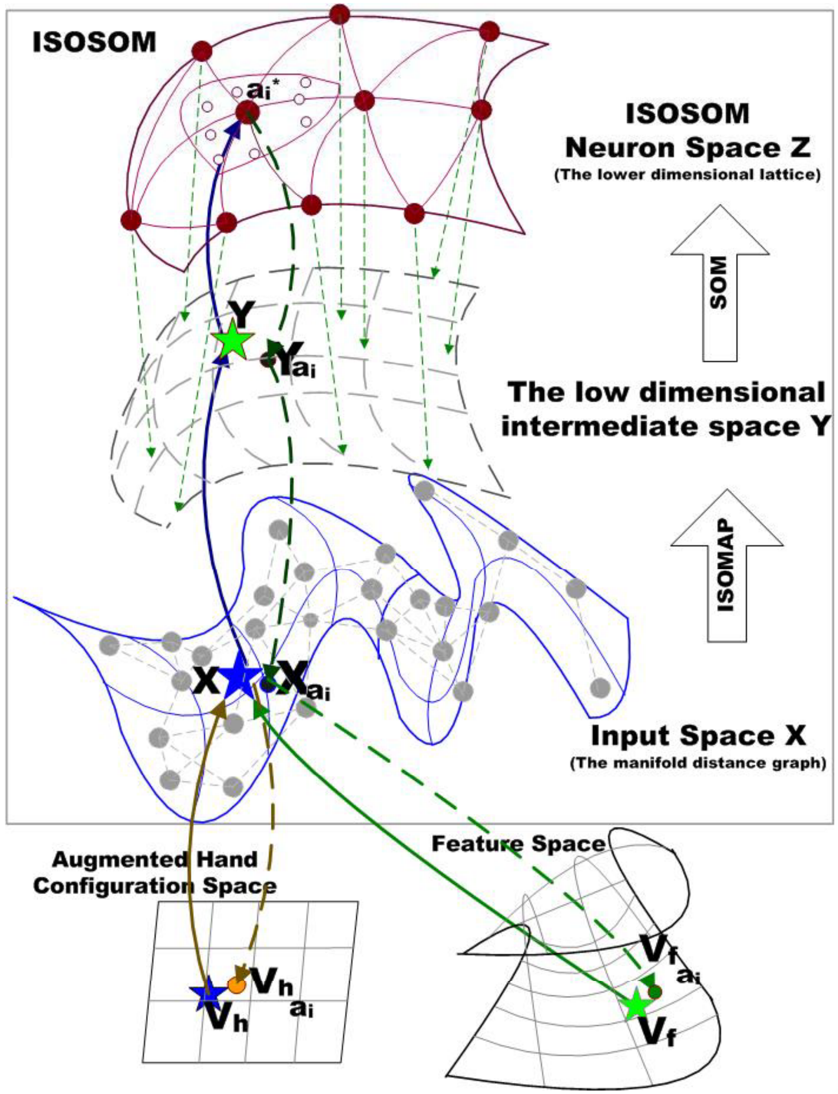




The retrieval results of the MAP framework with two-view images

Isometric self-organizing map (ISOSOM)

- A novel organized structure
  - Kohonen's Self-organizing Map
  - Tenenbaum's ISOMAP
  - To reduce information redundancy and avoid exhaustive search by nonlinear clustering techniques
- Multi-flash camera for the depth edges
  - Less background clutters
  - Internal finger edges

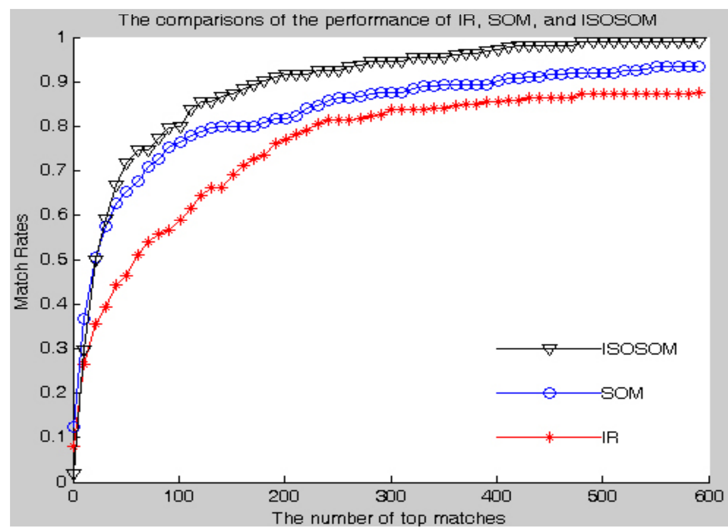


Experimental Results

Pose retrieval results

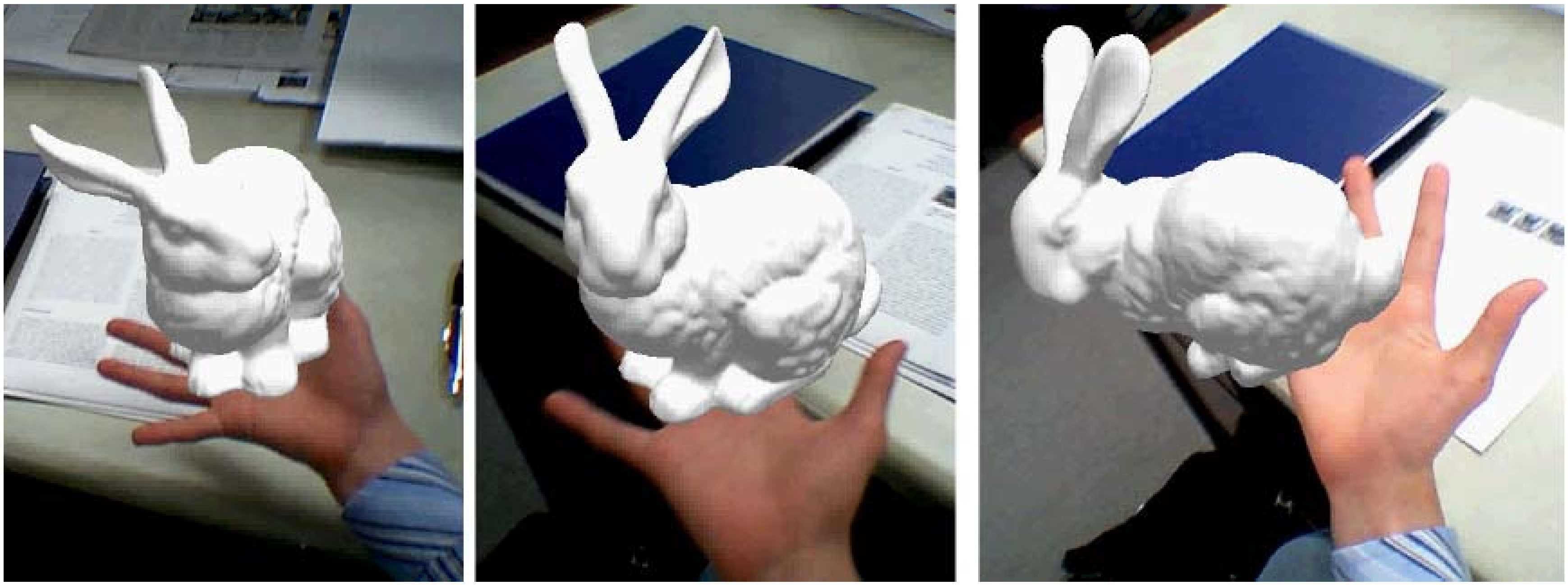
Number	IR	SOM	ISOSOM
Top 40	44.25%	62.39%	65.93%
Top 80	55.75%	72.12%	77.43%
Top 120	64.60%	78.76%	85.40%
Top 160	70.80%	80.09%	88.50%
Top 200	76.99%	81.86%	91.59%
Top 240	81.42%	85.84%	92.48%
Top 280	82.30%	87.17%	94.69%

The correct retrieval rates



The performance comparisons

HandyAR: Inspection of objects in AR

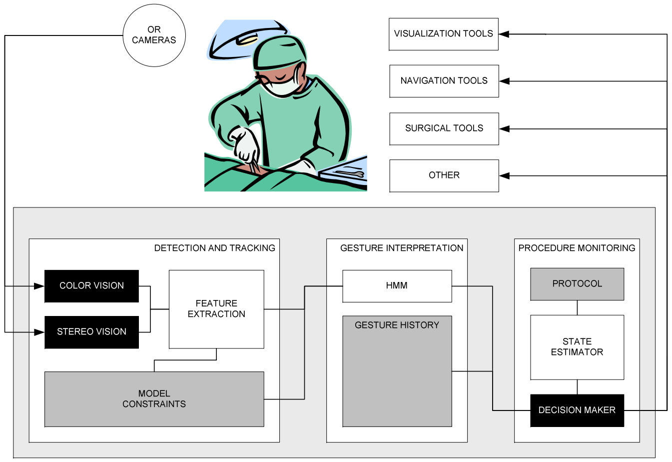


HandyAR

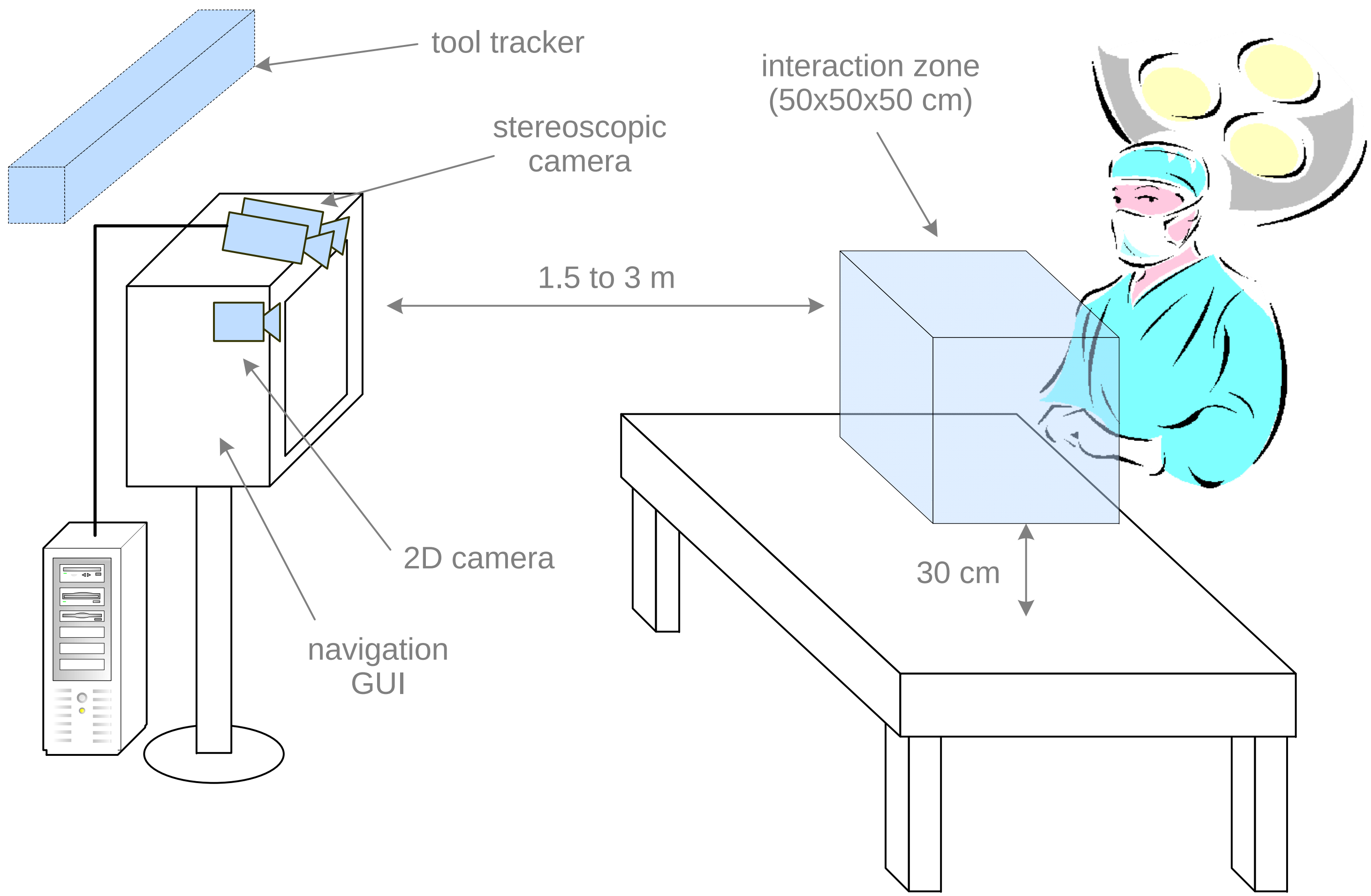
Video

Surgeon-computer interface

S. Grange, EPFL



Uses depth data (stereo camera) and video


tool tracker
stereoscopic camera
interaction zone (50x50x50 cm)
1.5 to 3 m
2D camera
navigation GUI
30 cm

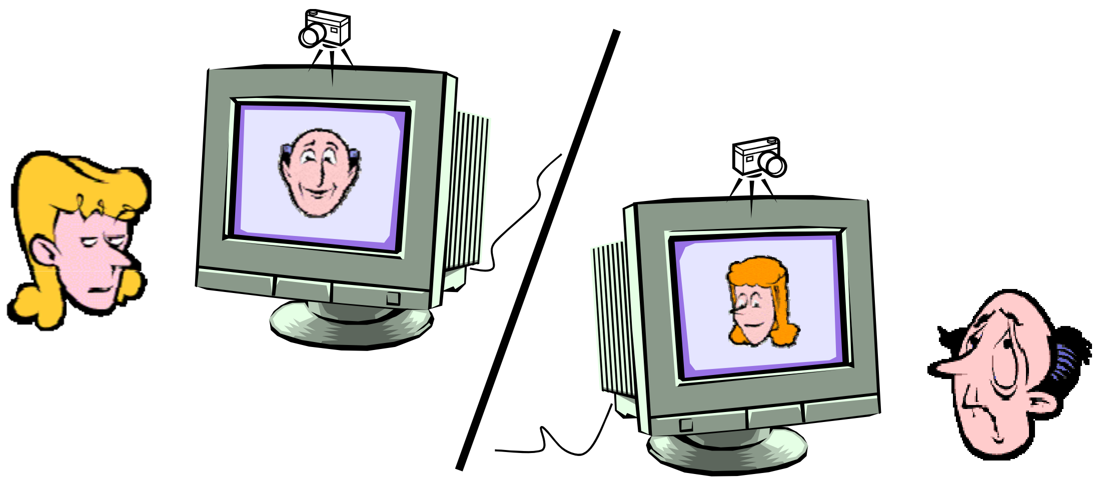
Video

Video

Video

Transformed Social Interaction

Studying nonverbal communication by manipulating reality in collaborative virtual environments

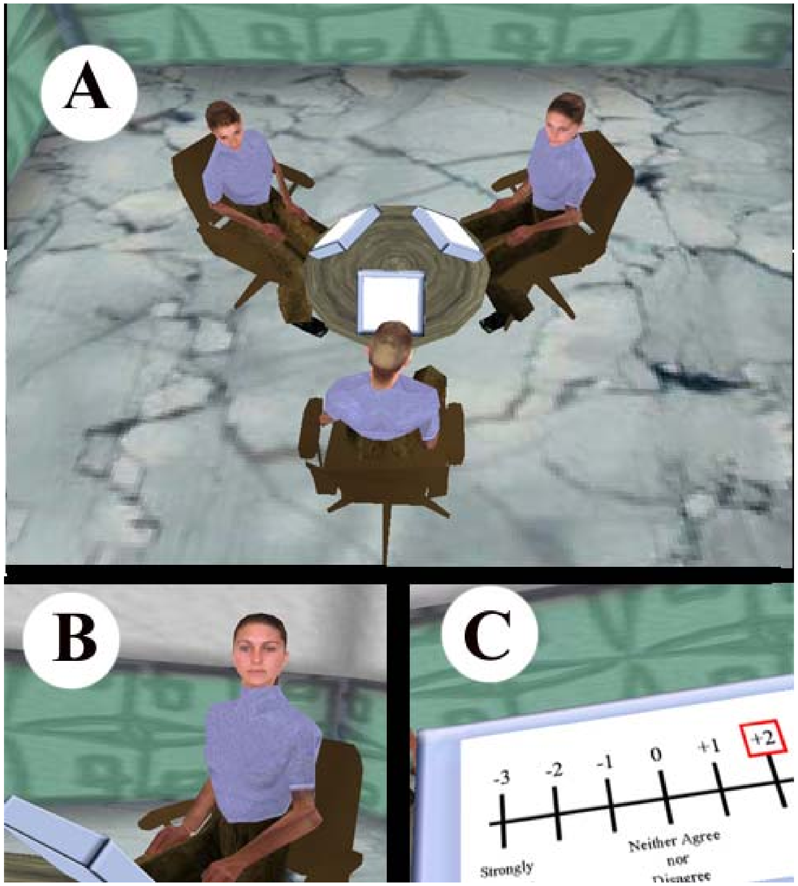


Manipulating appearance and behavior

- Visual nonverbal communication is an important aspect of human interaction
- Since behavior is decoupled from its rendering in CVEs, the opportunity arises for new interaction strategies based on manipulating the visual appearance and behavior of the avatars.
- For example:
  - Change identity, gender, age, other physical appearance
  - Selectively filter, amplify, delete, or transform nonverbal behaviors of the interactant
  - Culturally sensitive gestures, edit yawns, redirect eye gaze, …
  - Could be rendered differently to every other interactant

Transformed Social Interaction (TSI)

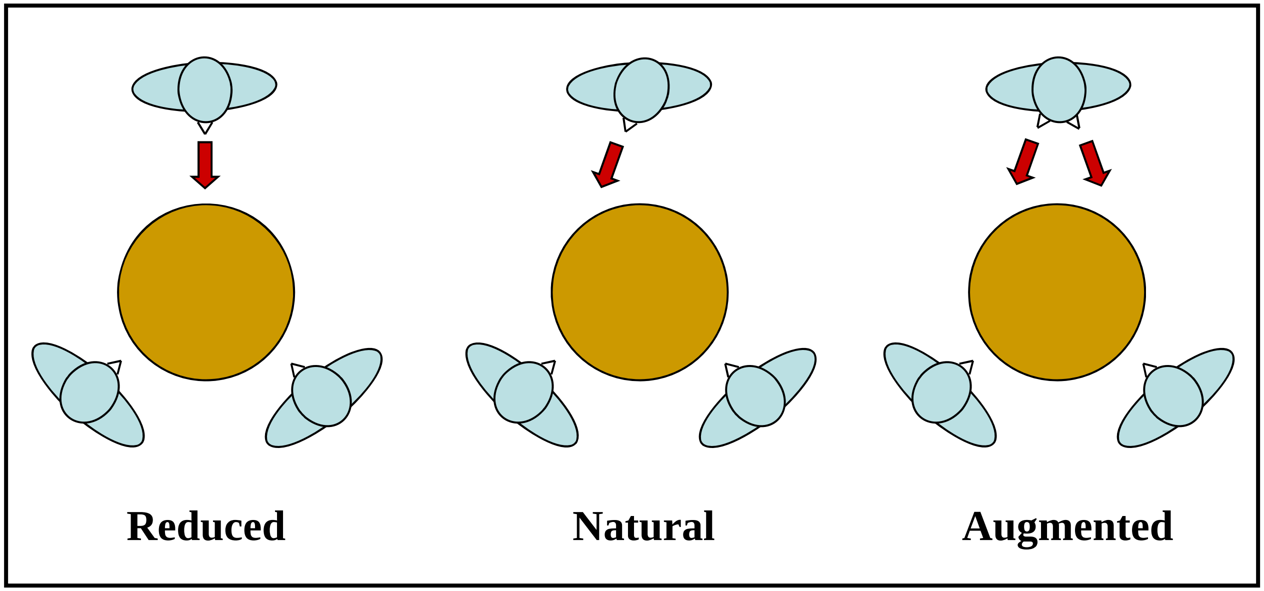
- TSI: Strategic filtering of communicative behaviors in order to change the nature of social interaction



A TSI experiment: Non-zero-sum gaze

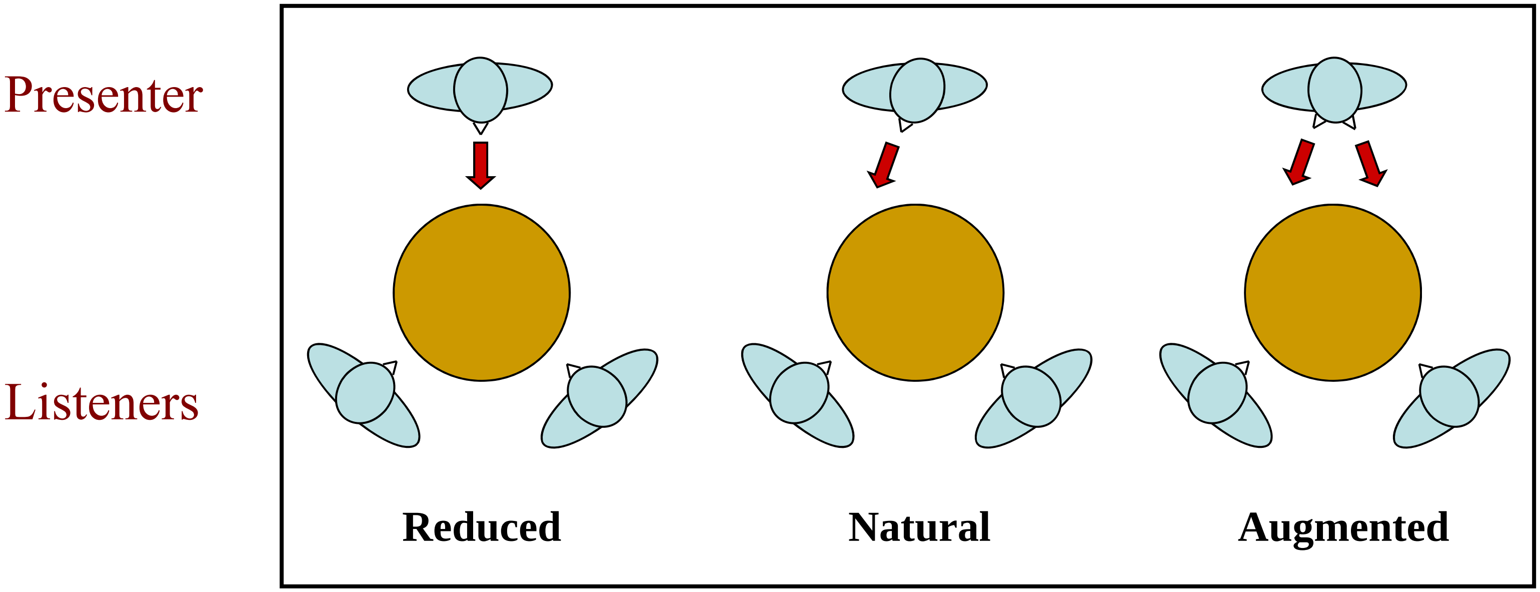
Presenter

Listeners



- Is it possible to increase one’s power of persuasion by “augmented non-zero-sum (NZS) gaze”?
  - Presenter gives each participant > 50% of attention
- Experiment: A presenter tries to persuade two listeners by reading passages of text. Gaze direction is manipulated.

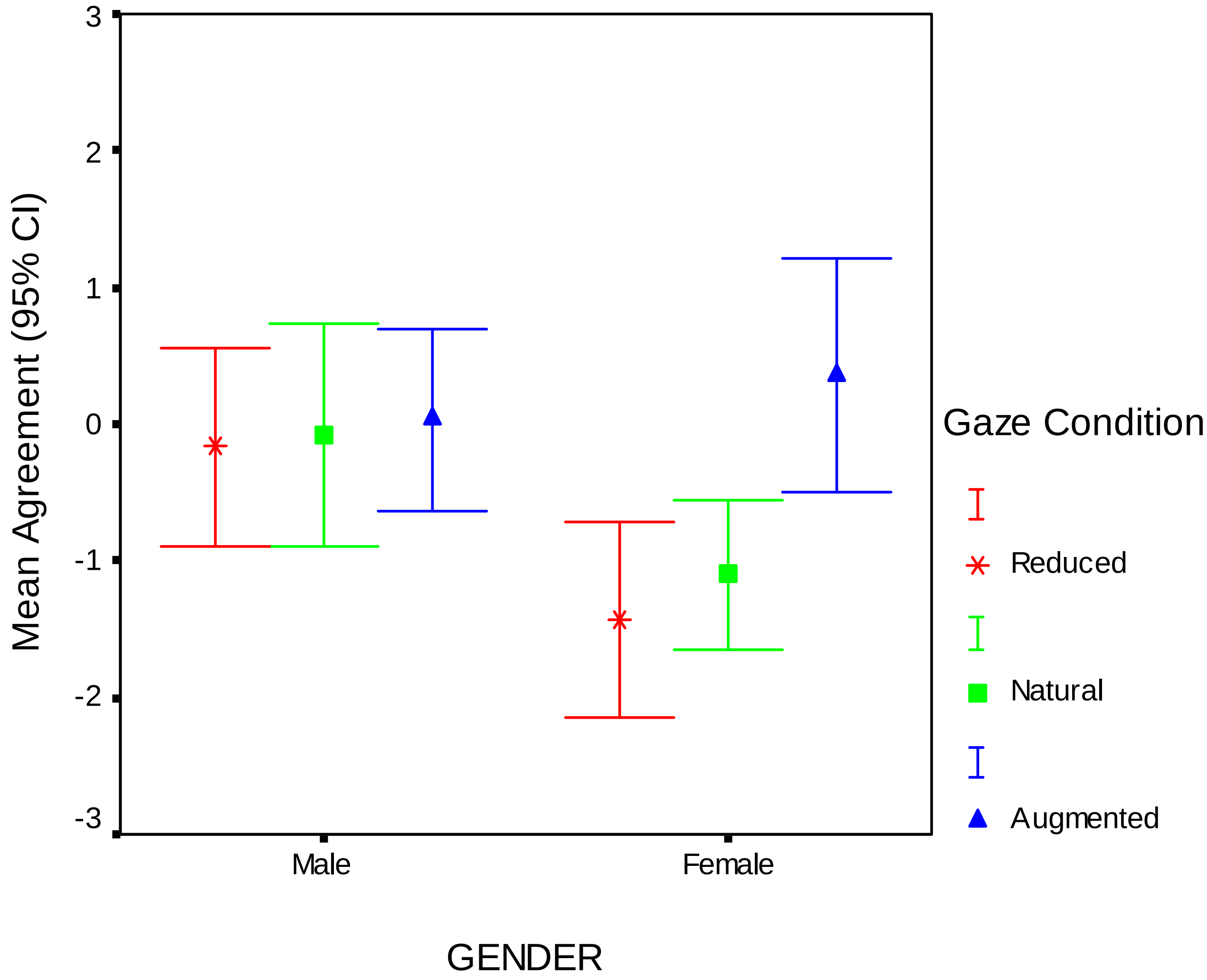
Non-zero-sum gaze



- Three levels of gaze of the presenter:
  - **Reduced**: no eye contact
  - **Natural**: unaltered, natural eye contact
  - **Augmented**: 100% eye contact

NZSG conditions

Initial results



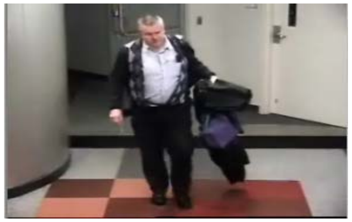
TSI conclusions

- TSI is an effective paradigm for the study of human-human interaction
- TSI should inform the study and development of multimodal interfaces
- TSI may help overcome deficiencies of remote collaboration and potentially offer advantages over even face-to-face communication
- This is just one study, somewhat preliminary – others are in the works….

PeopleSearch: Finding Suspects

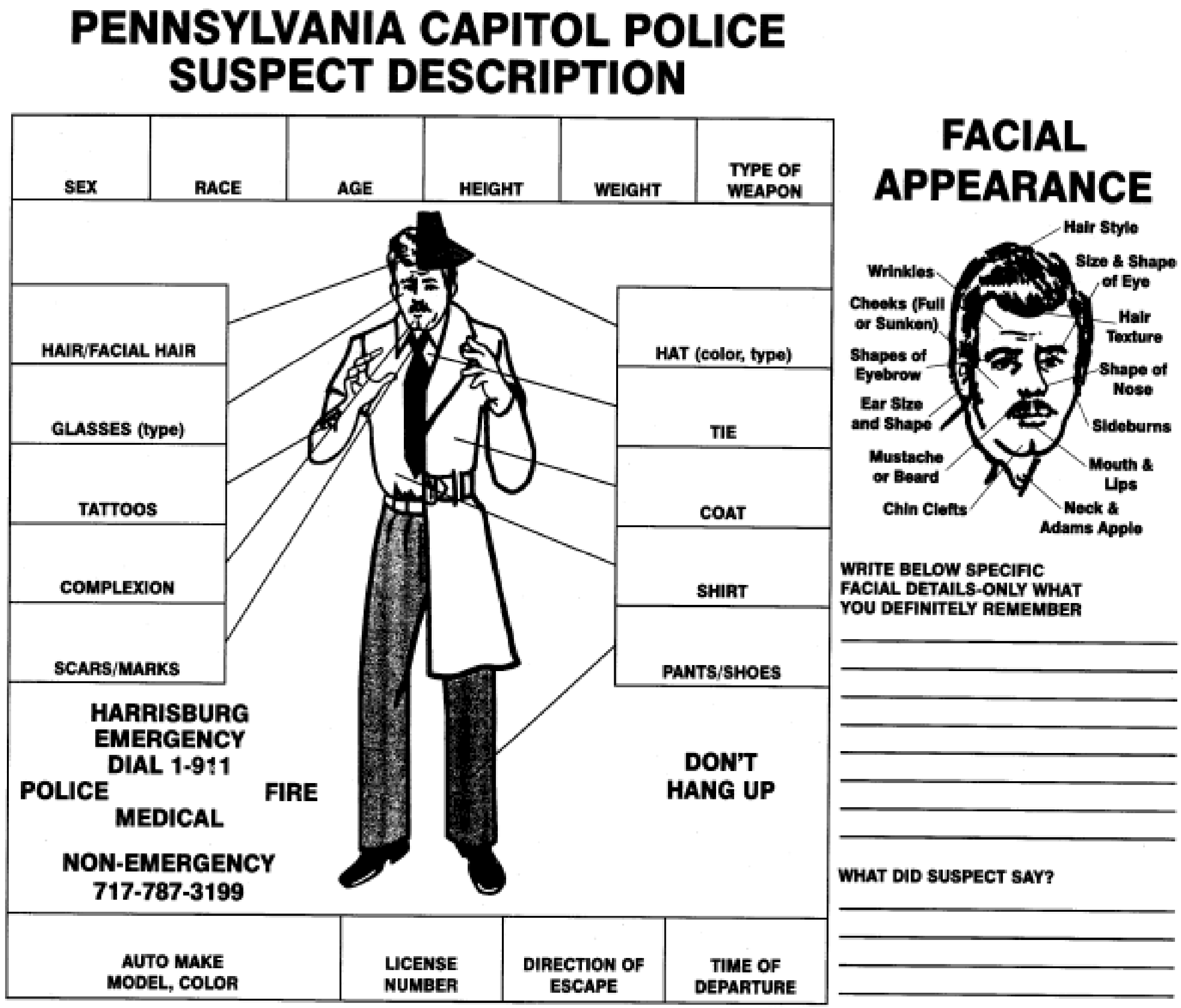
PeopleSearch

- Video Security Cameras
  - Airports
  - Train Stations
  - Retail Stores
  - Etc.



- For
  - Eyewitness descriptions
  - Missing people
  - Tracking across cameras
- Large amounts of video data
  - How to effectively search through these archives?

Suspect Description Form



PENNSYLVANIA CAPITOL POLICE
SUSPECT DESCRIPTION

SEX	RACE	AGE	HEIGHT	WEIGHT	TYPE OF WEAPON

HAIR/FACIAL HAIR

GLASSES (type)

TATTOOS

COMPLEXION

SCARS/MARKS

HAT (color, type)

TIE

COAT

SHIRT

PANTS/SHOES

HARRISBURG
EMERGENCY
DIAL 1-911
POLICE
FIRE
MEDICAL

NON-EMERGENCY
717-787-3199

DON'T
HANG UP

AUTO MAKE MODEL, COLOR	LICENSE NUMBER	DIRECTION OF ESCAPE	TIME OF DEPARTURE

FACIAL
APPEARANCE

WRITE BELOW SPECIFIC FACIAL DETAILS-ONLY WHAT YOU DEFINITELY REMEMBER

WHAT DID SUSPECT SAY?

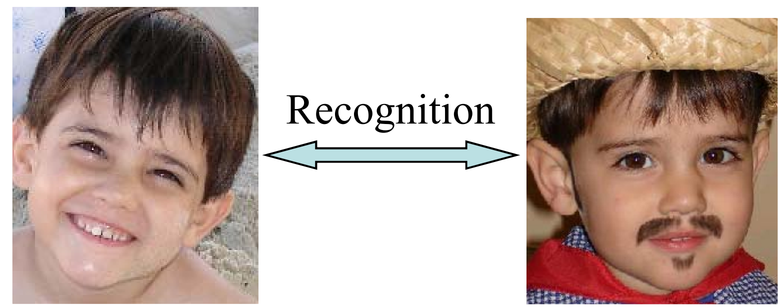
Problem definition

- Given a Suspect Description Form, build a system to automatically search for potential suspects that match the specified physical attributes in surveillance video

- Query Example: "Show me all bearded people entering IBM last month, wearing sunglasses, a red jacket and blue pants."

Face Recognition

- Long-term recognition (need to be robust to makeup, clothing, etc.)
- Return the identity of the person
- **Not reliable under pose and lighting changes**

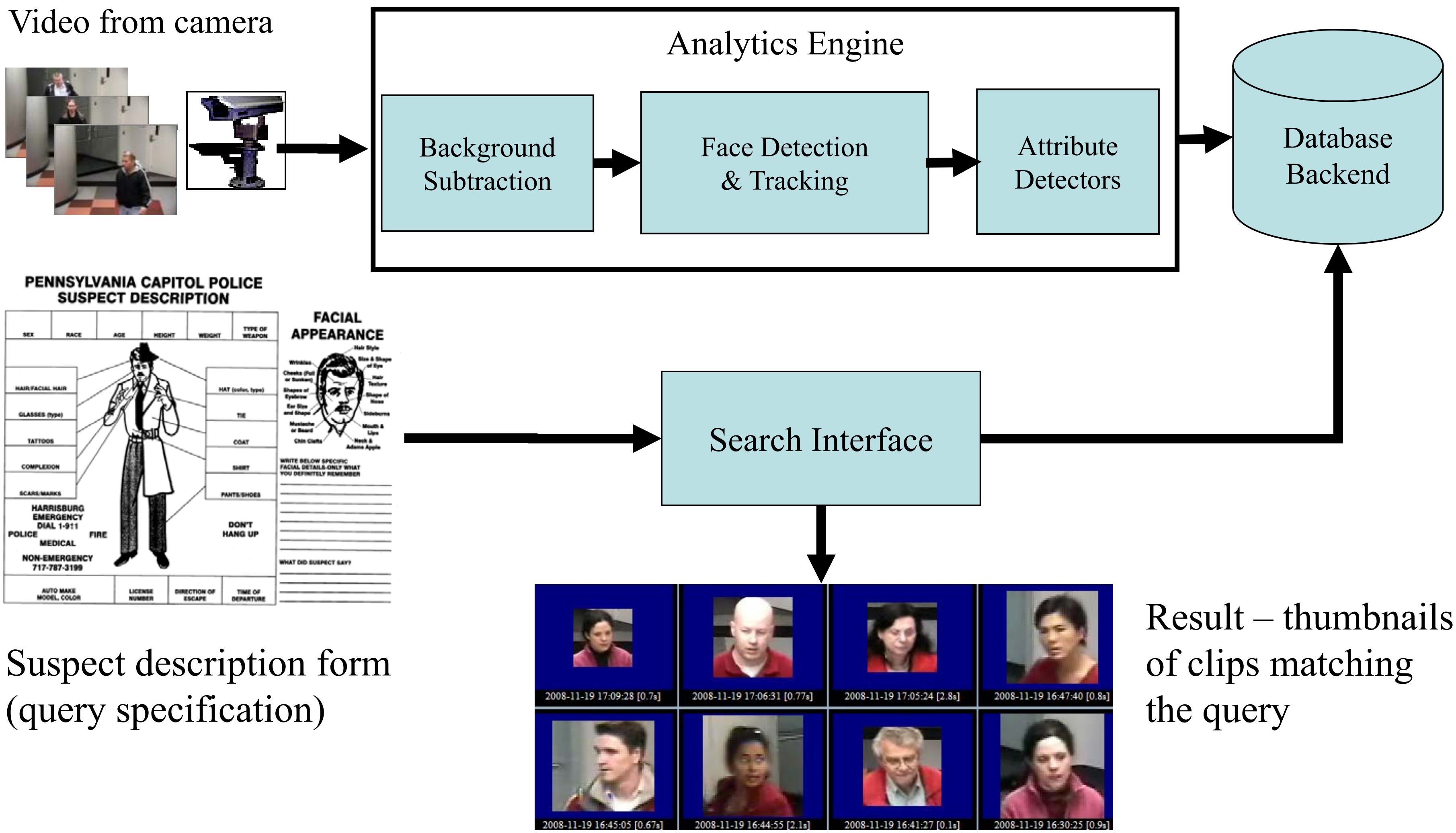


Our Approach: People Search by Attributes

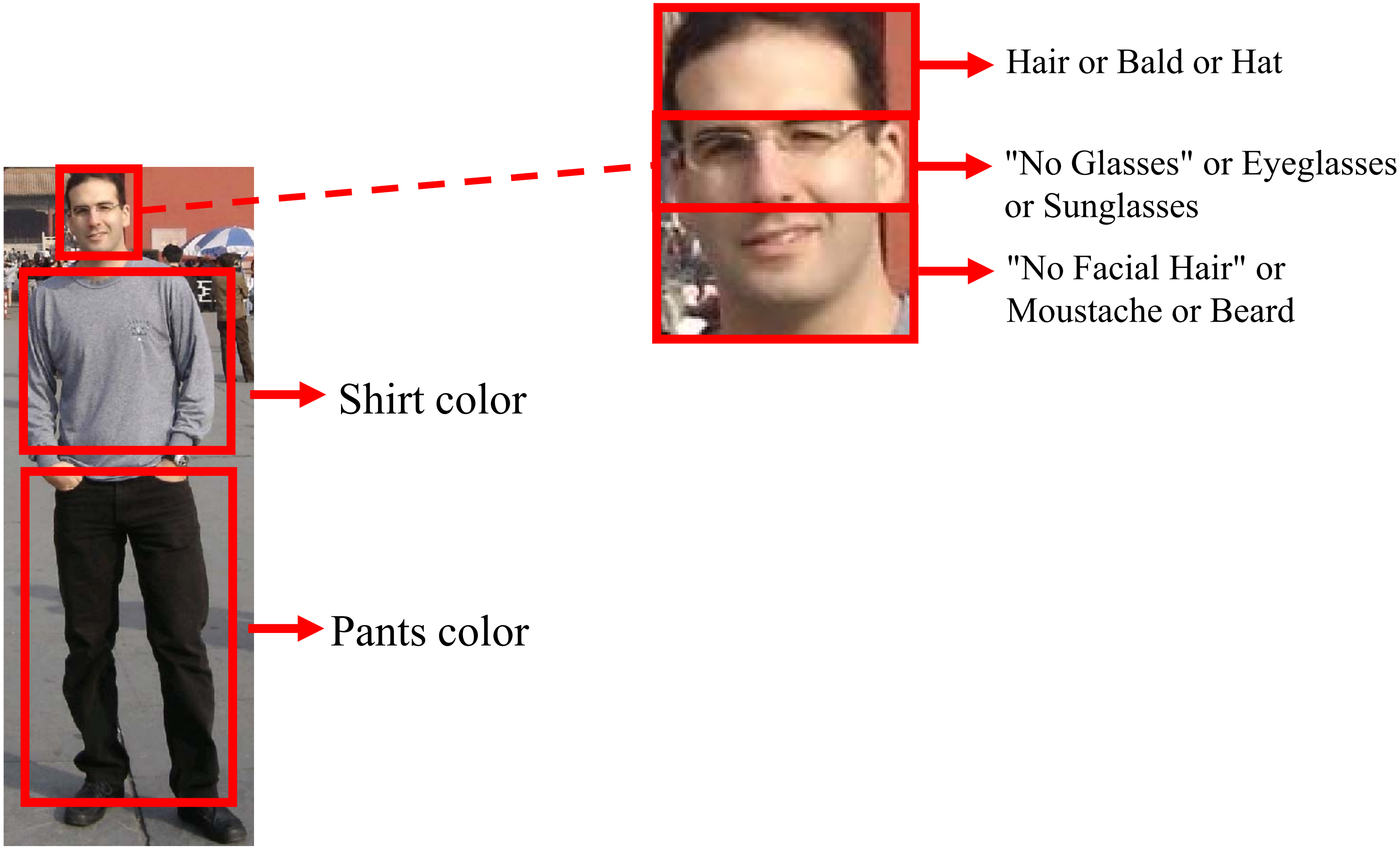
- Short-term recognition (take advantage of makeup, clothing, etc.)
- Return a set of images that match the search attributes
- **Based on reliable object detection technology**

Query: Show me all people with moustache and hat

System overview



Human body analysis

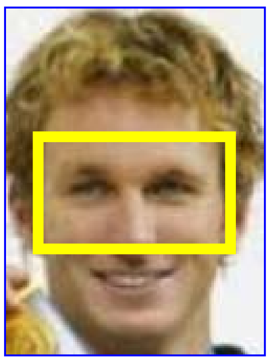


Bald

Hair

Hat

No Glasses



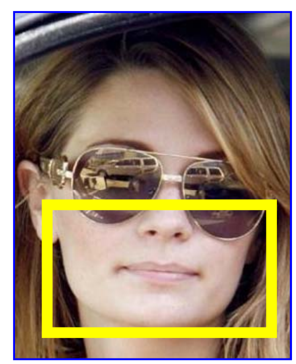
Sunglasses

Eyeglasses

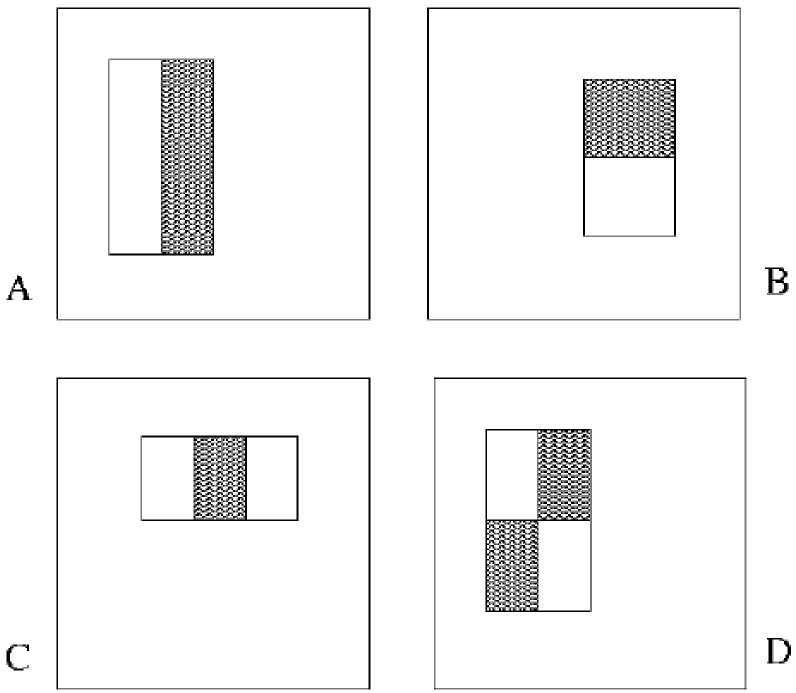
Beard

Moustache

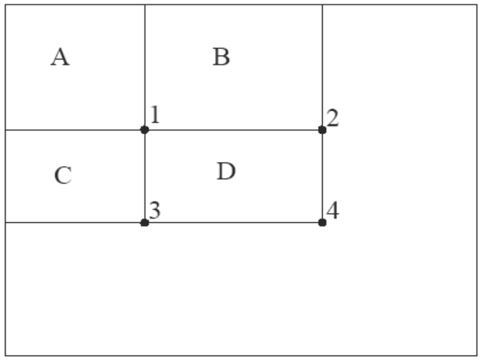
No Facial Hair



Adaboost learning w/Haar features



Integral Image

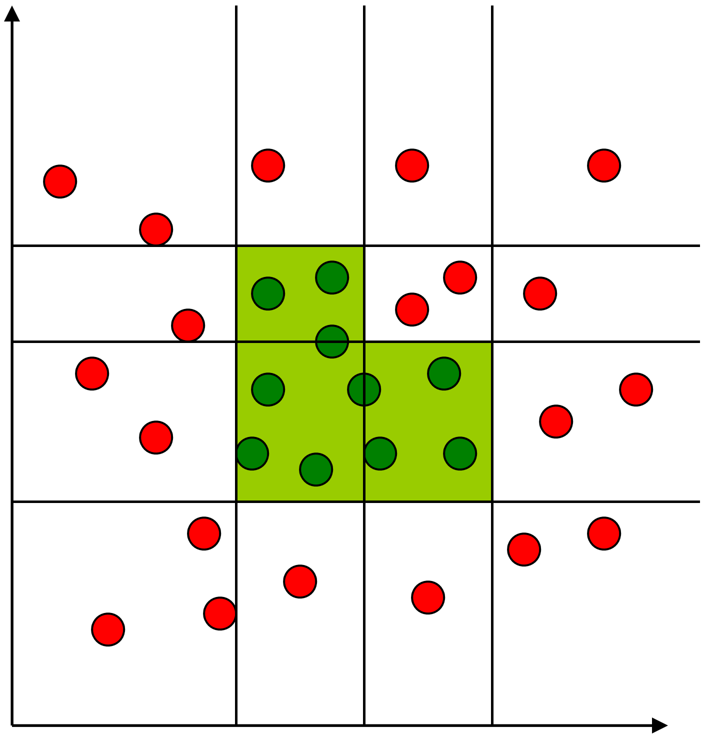


$$D = ii(4) + ii(1) - ii(2) - ii(3)$$

$$= (A+B+C+D)+(A)-(A+B)-(A+C)$$

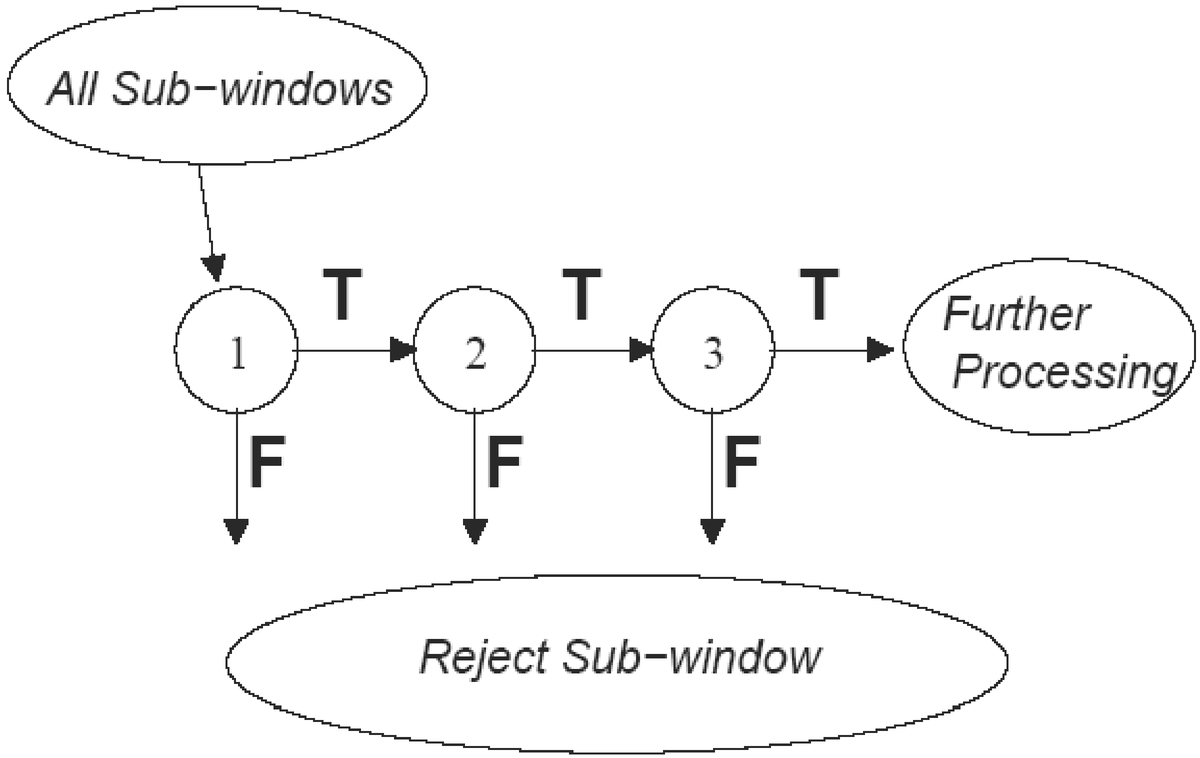
Adaboost learning

- Adaboost creates a single strong classifier from many weak classifiers



- Initialize sample weights
- For each cycle:
  - Find a classifier that performs well on the weighted sample
  - Increase weights of misclassified examples
- Return a weighted combination of classifiers

Cascade of Adaboost classifiers

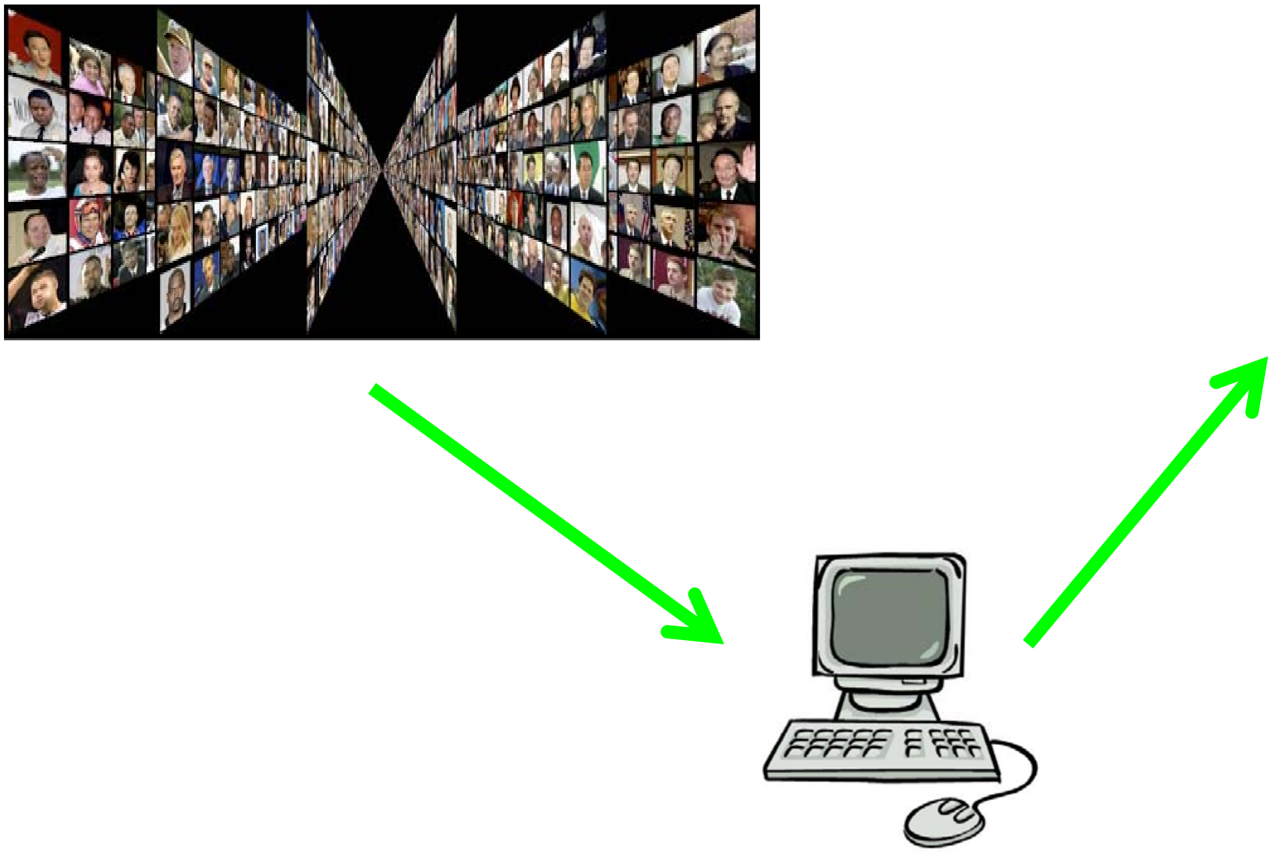


Applying the detector

Search over all possible window positions and scales

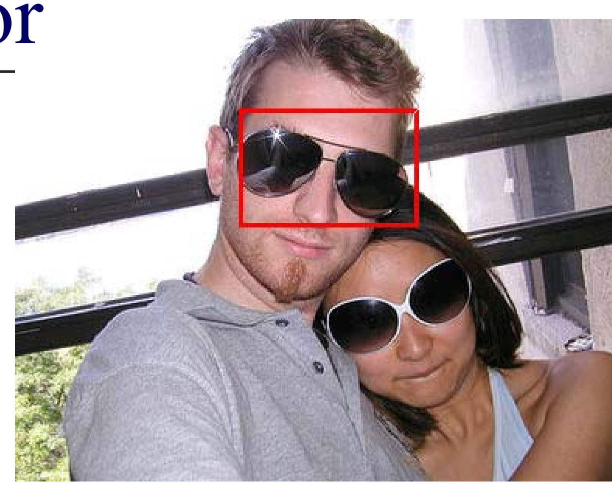
Apply the learned Adaboost classifier using the cascade scheme of Viola & Jones for each window position/scale

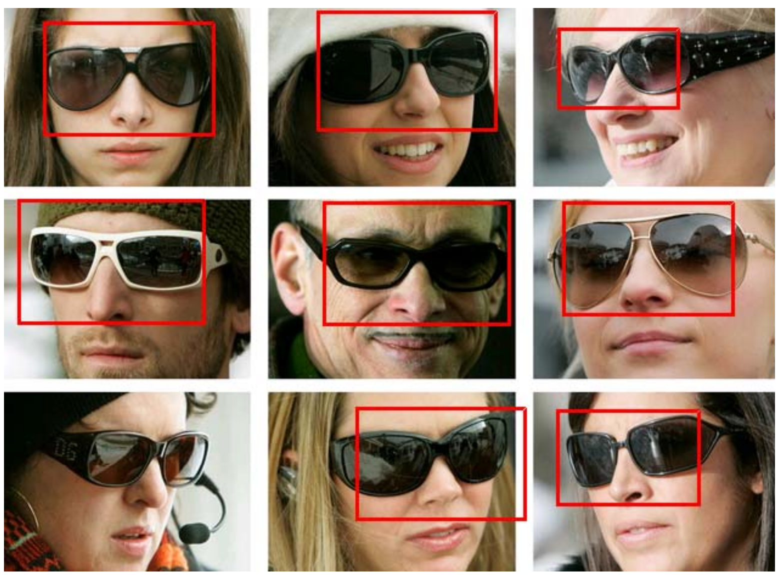
Multiple detector learning

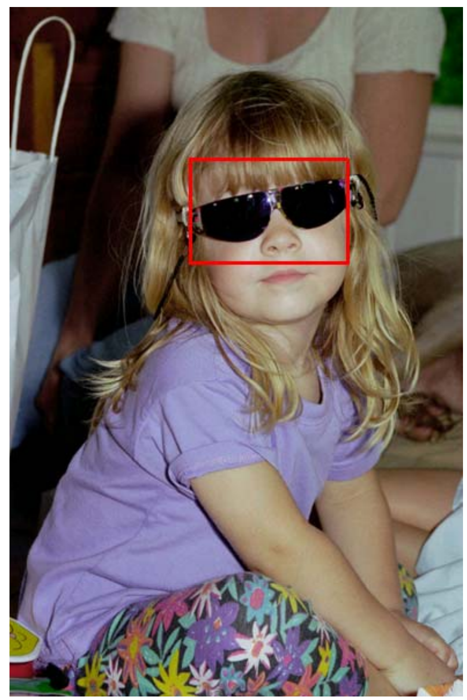


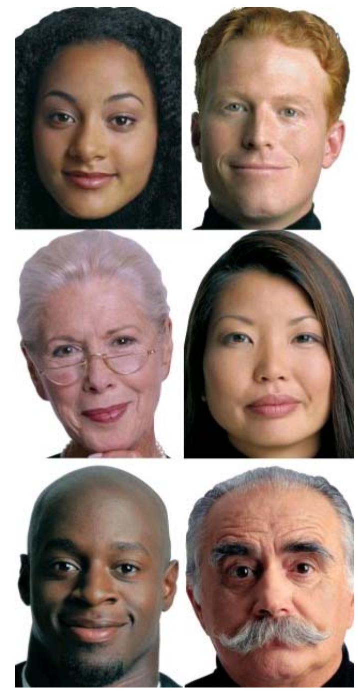
Beard Detector
Moustache Detector
"No Facial Hair" Detector
Sunglasses Detector
Eyeglasses Detector
"No Glasses" Detector
Bald Detector
Hair Detector
Hat Detector

Results: Sunglasses Detector

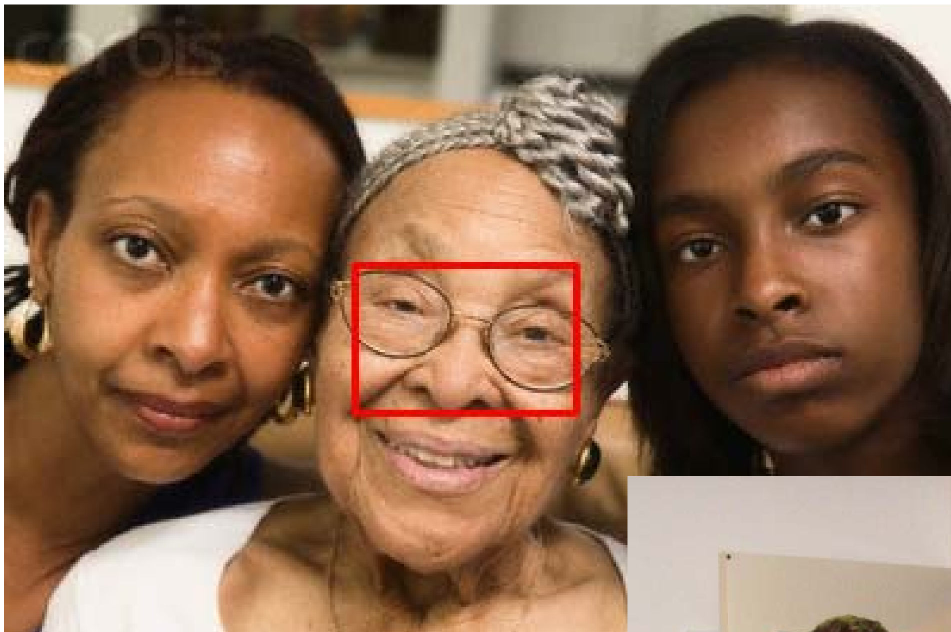


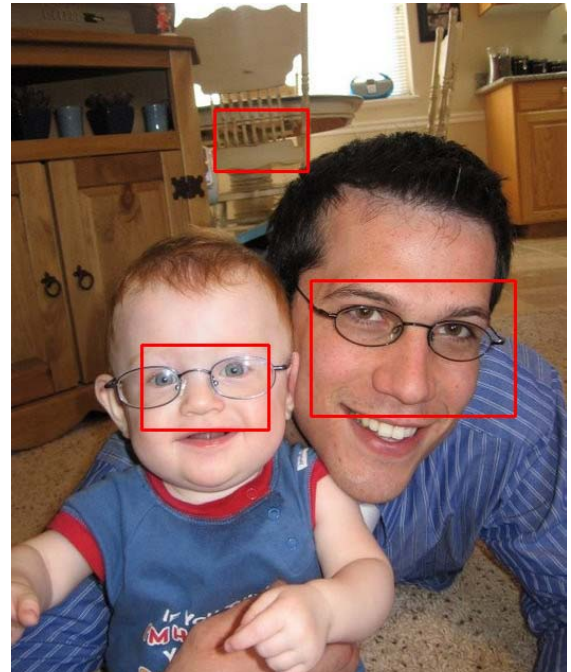


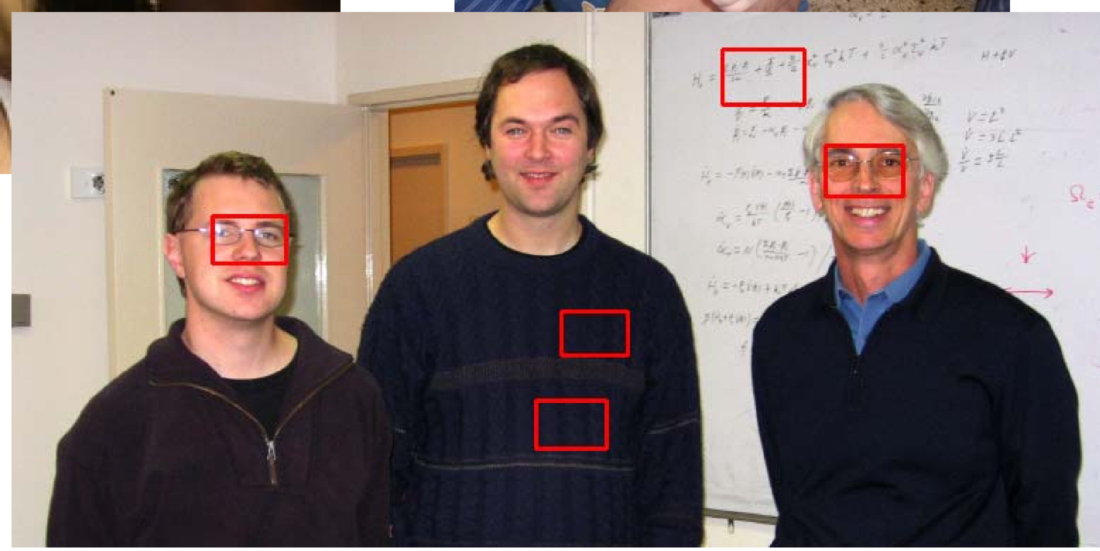




Results: Eyeglasses Detector



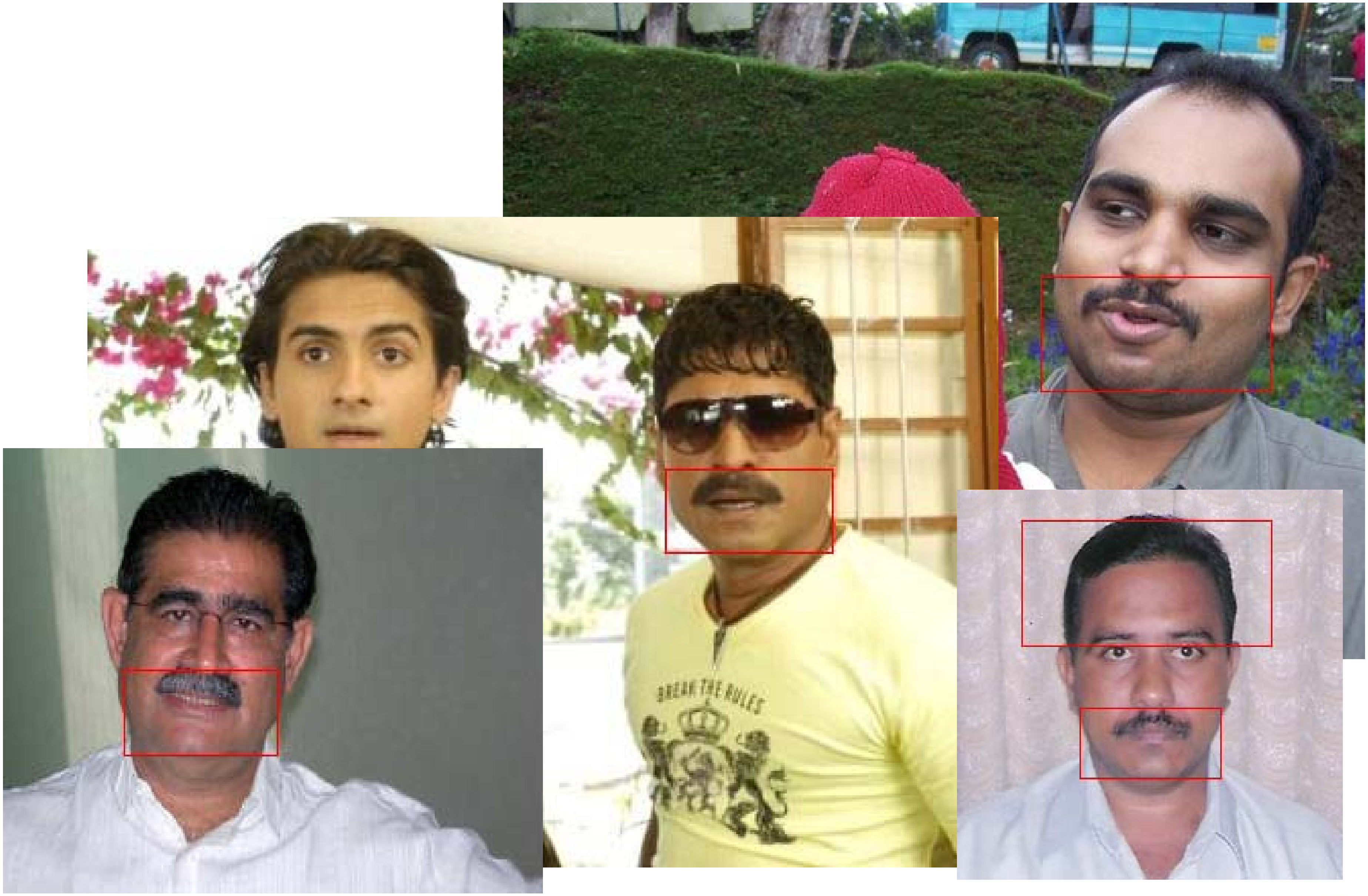




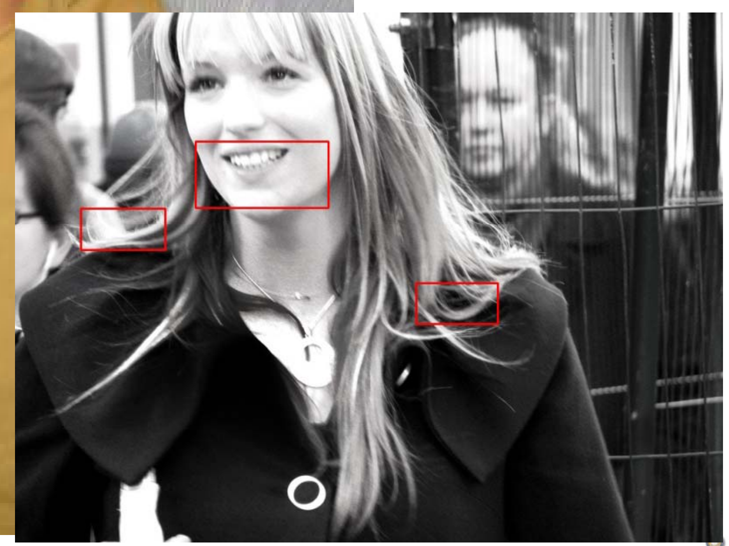
Results: "No Glasses" Detector

Results: Beard Detector

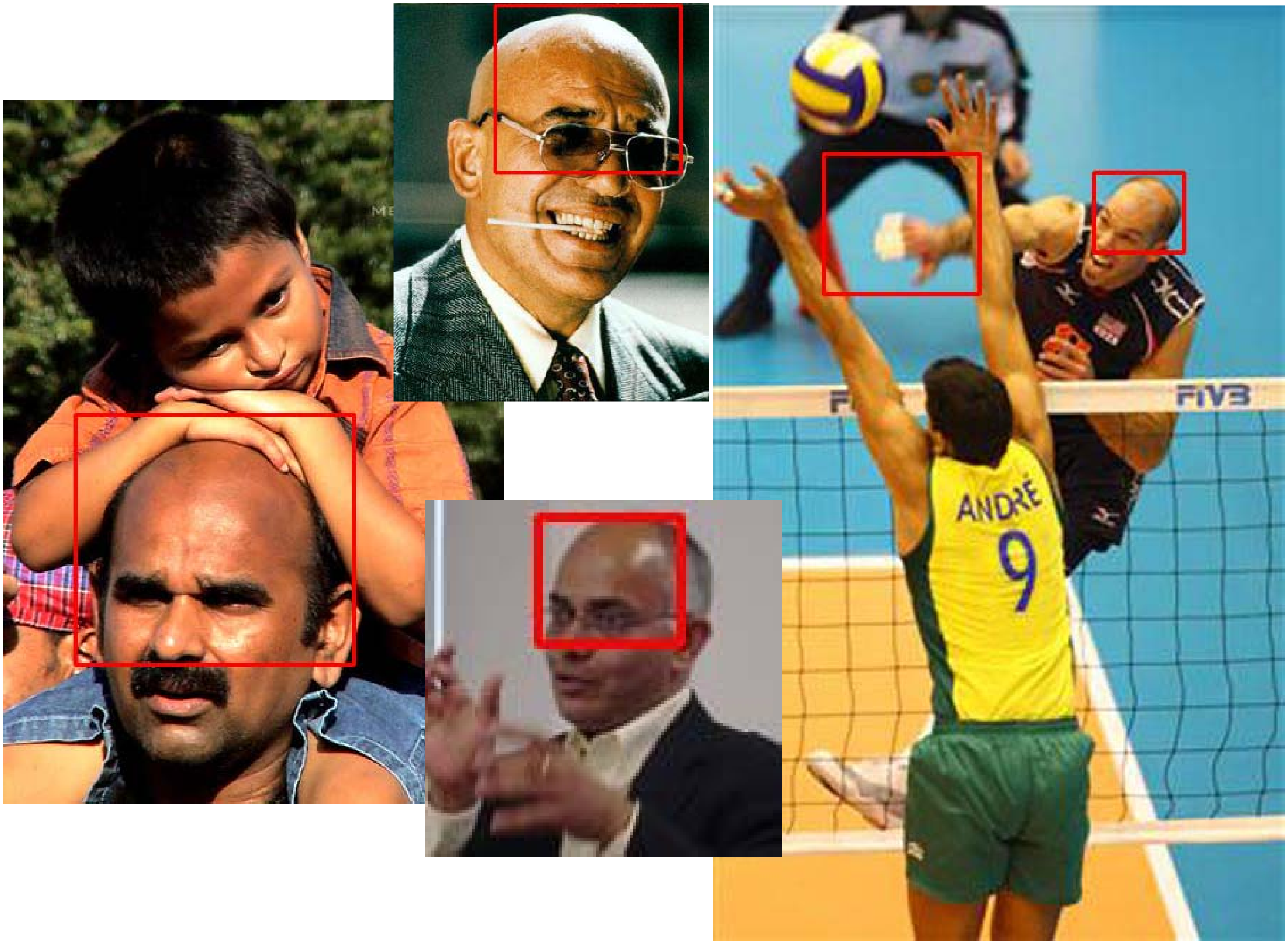
Results: Moustache Detector



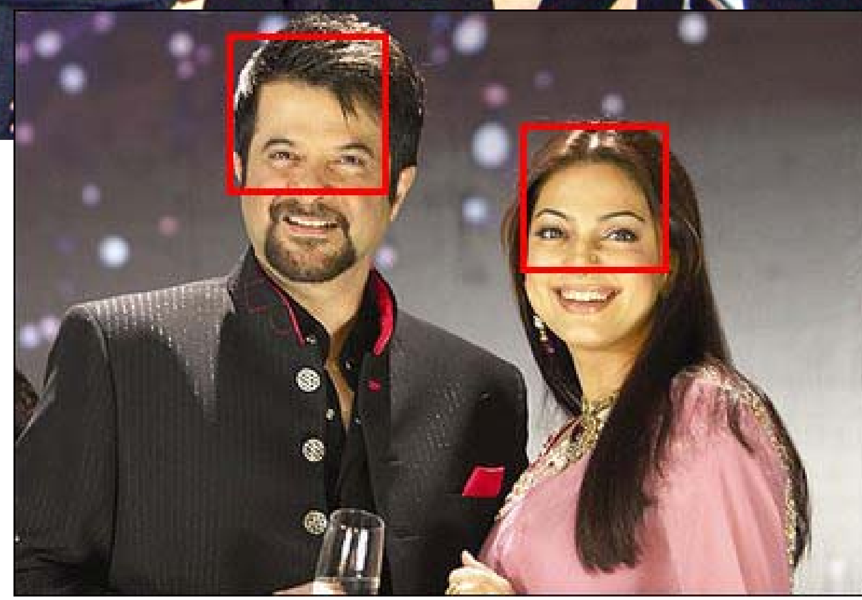
Results: "No Facial Hair" Detector



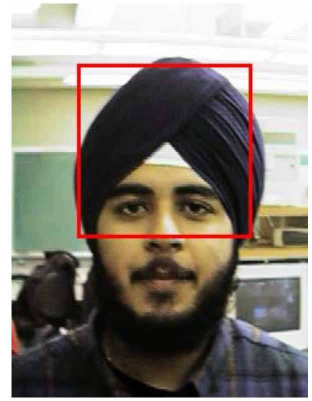
Results: Bald Detector

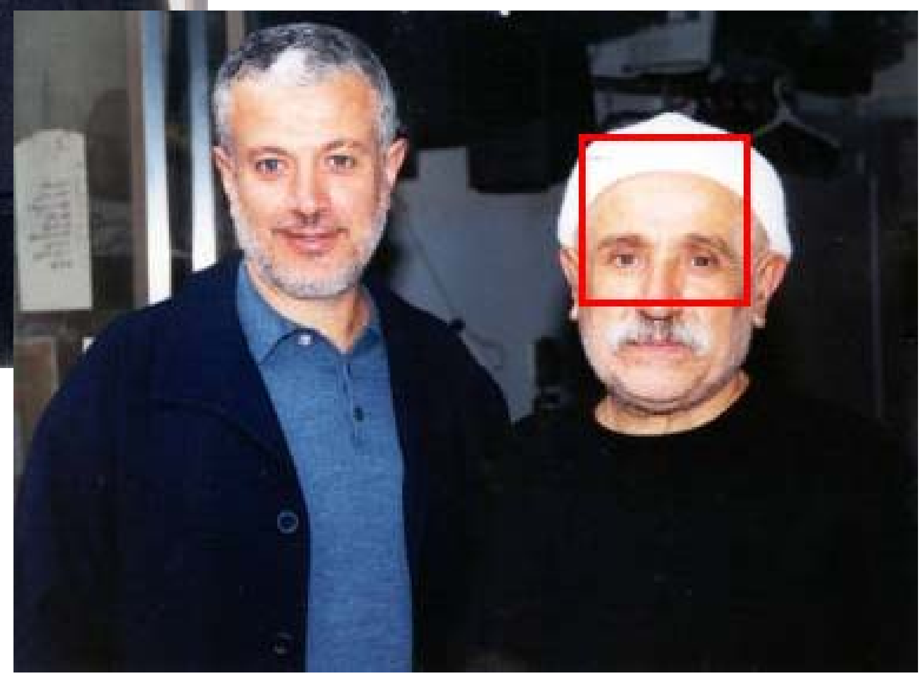


Results: Hair Detector

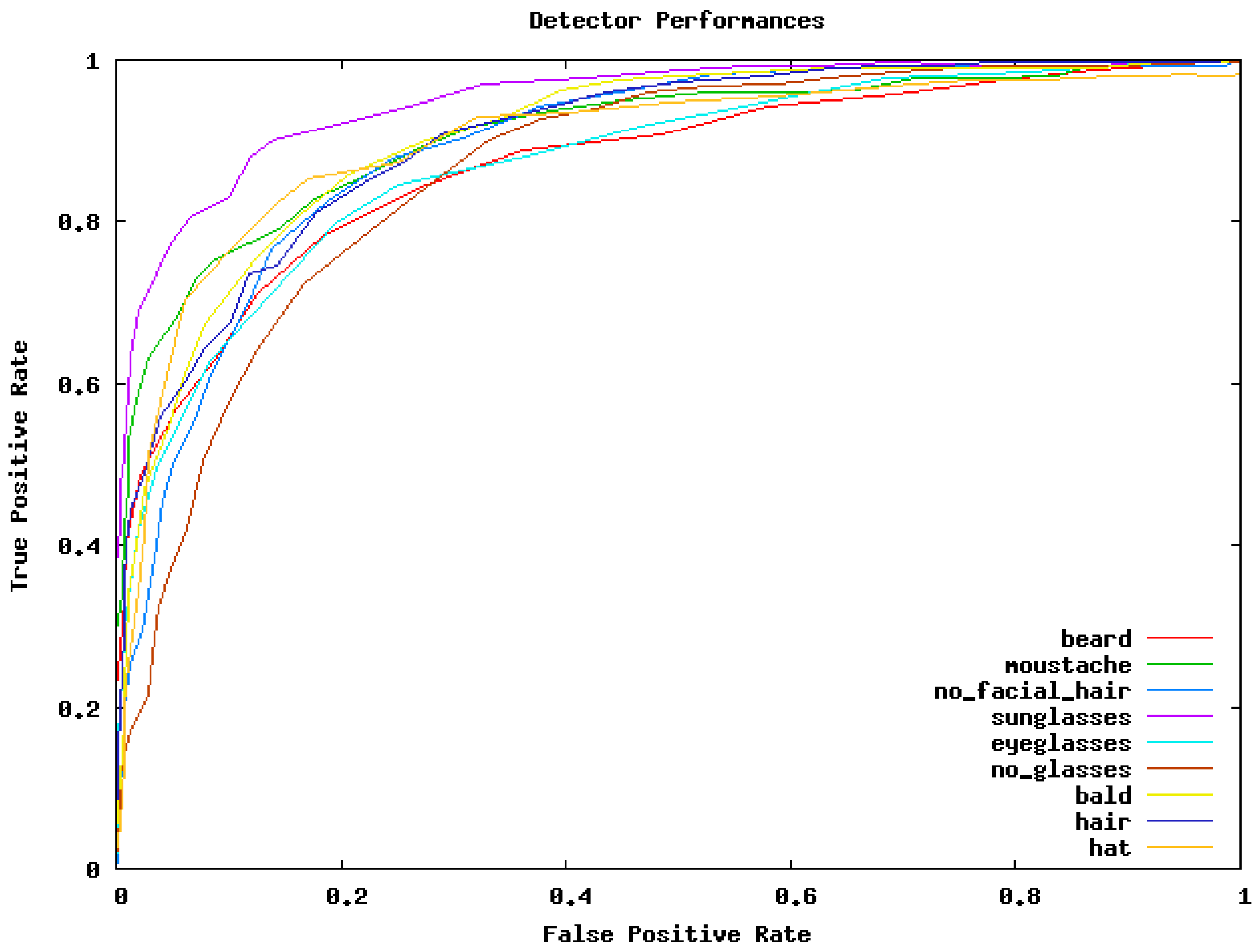


Results: Hat Detector



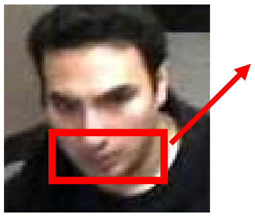


Performance evaluation



Examples of failure cases

(a) Lower Face Part

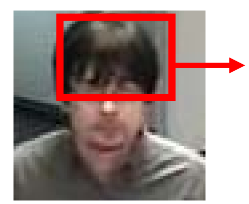


Shadow looks like beard

(b) Middle Face Part

Shadow looks like sunglasses

(c) Upper Face Part



Fringe confused with hat

Multispectral/IR

Attribute detection in multispectral images

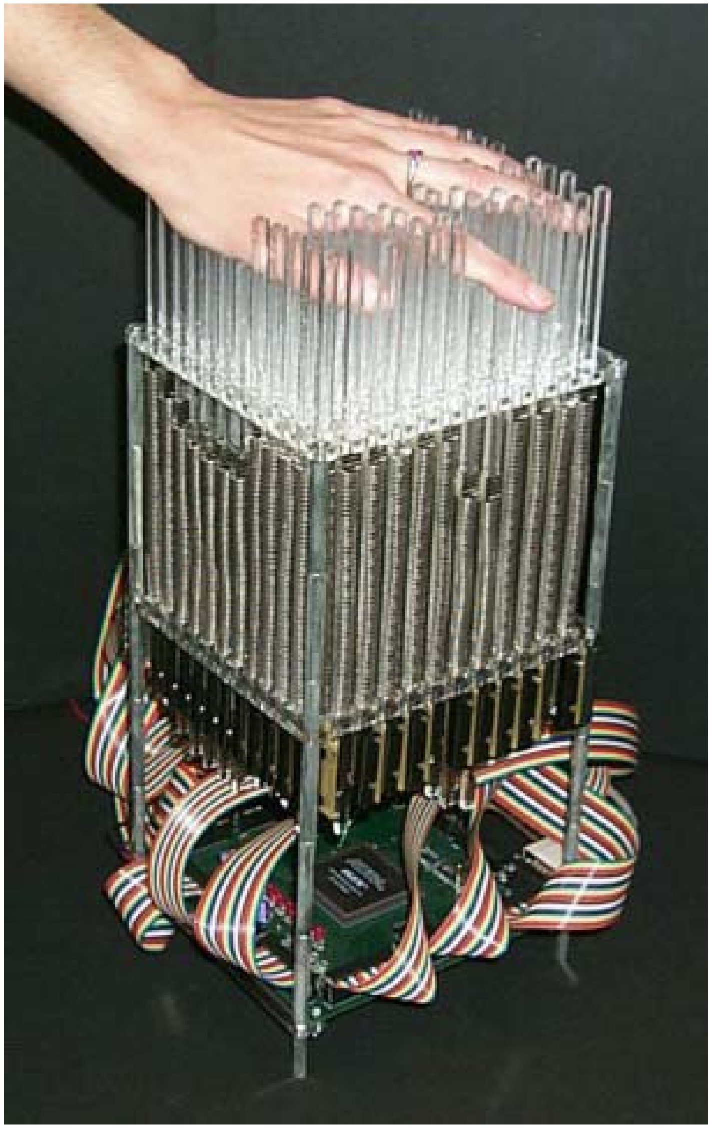
Media Arts and Technology (MAT)

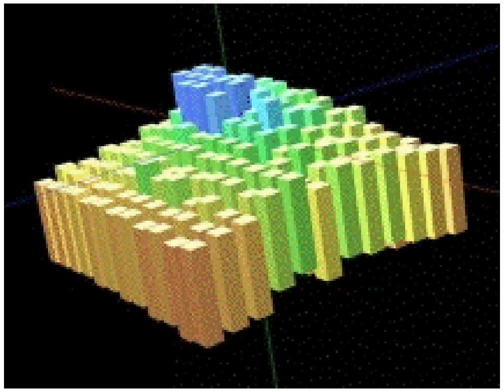
- ***Media Arts and Technology*** is an transdisciplinary graduate program at UCSB, founded to pursue emerging opportunities for education and research at the intersection of Art, Science, and Engineering.

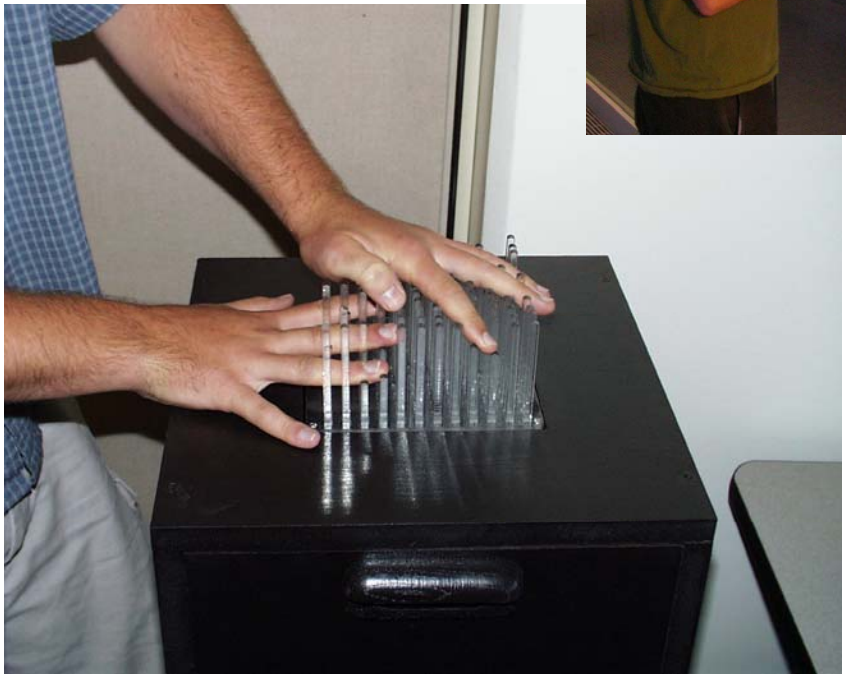
Media Arts and Technology
Graduate Program

Media Arts and Technology
Graduate Program

Devices for interactivity



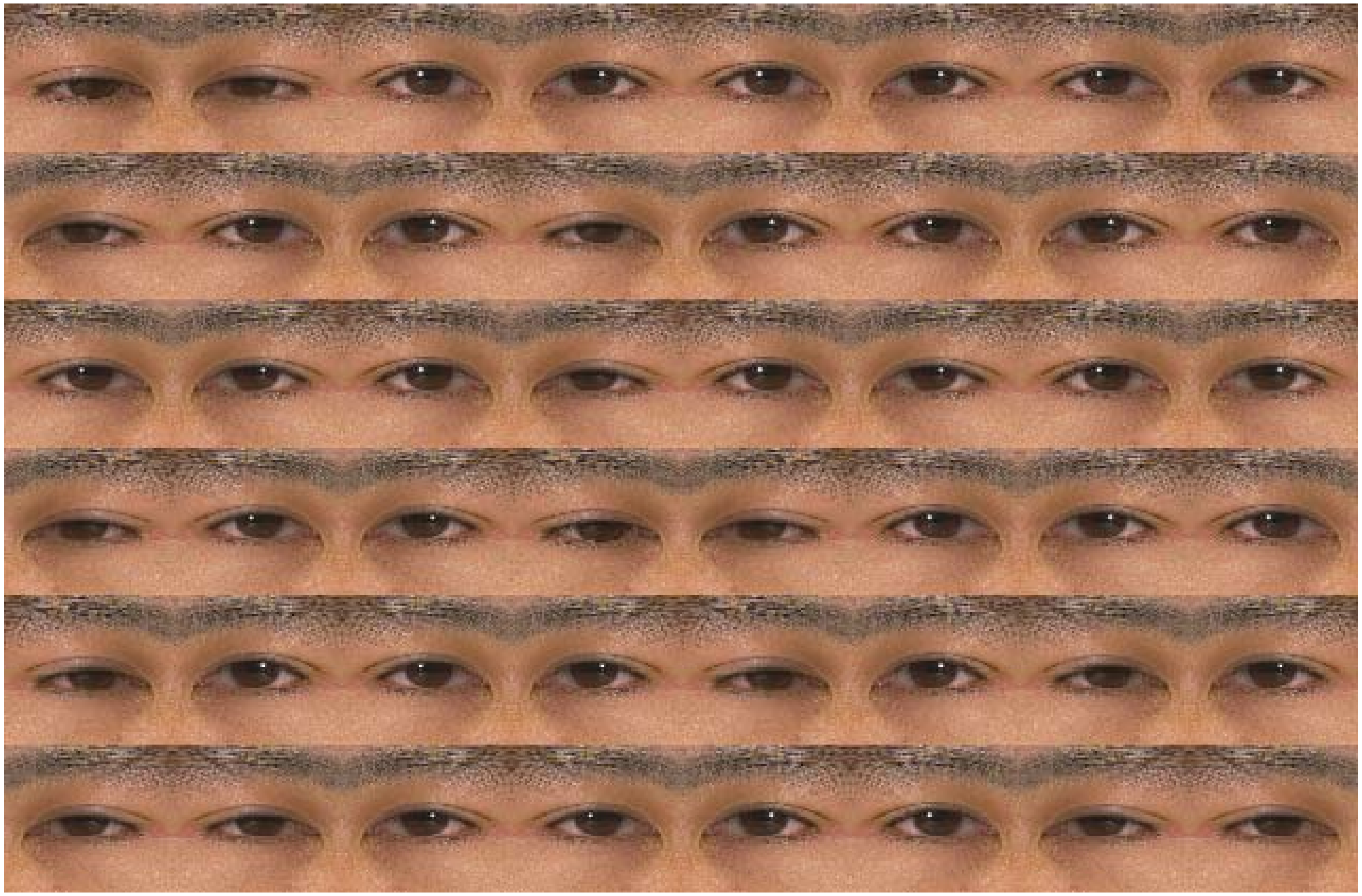




Interactive art

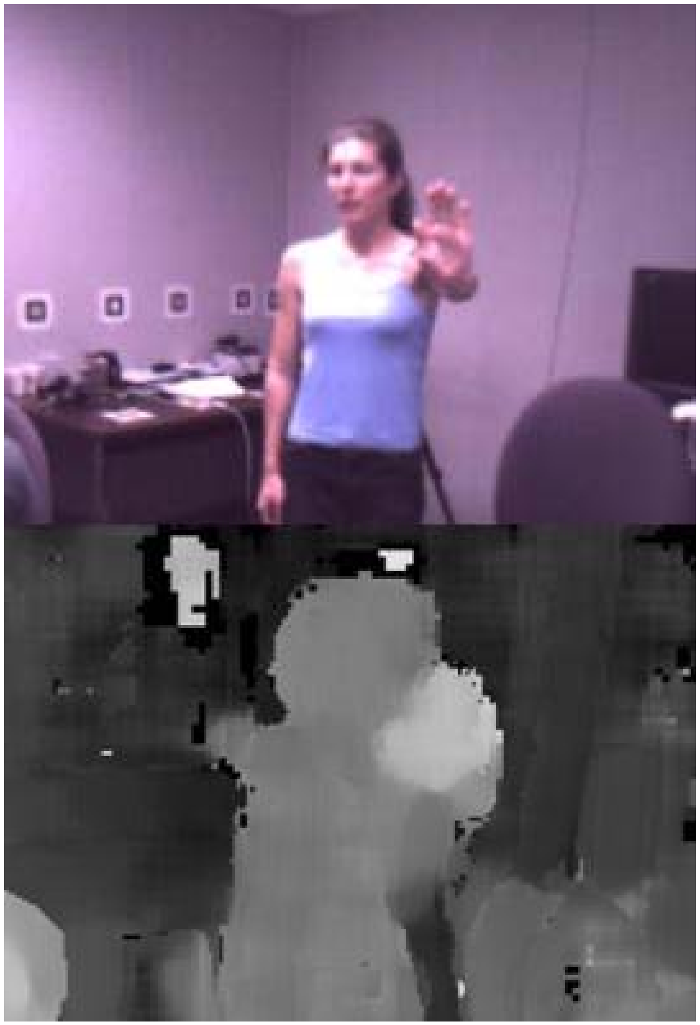
Sensing/Speaking Space @ SFMOMA

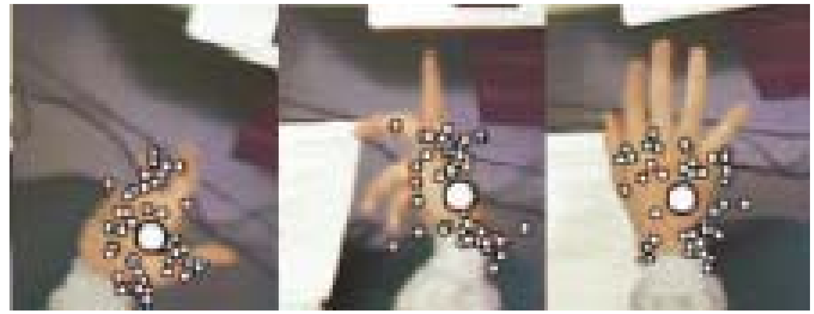
Algorithmic art

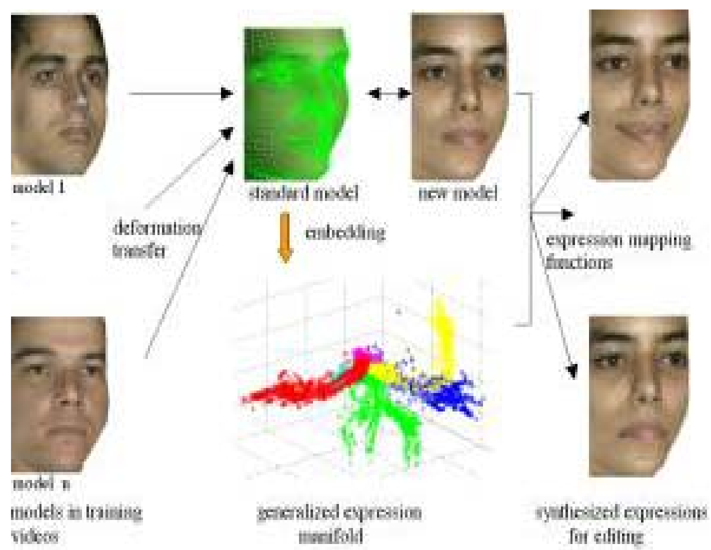


Blink @ SBMA

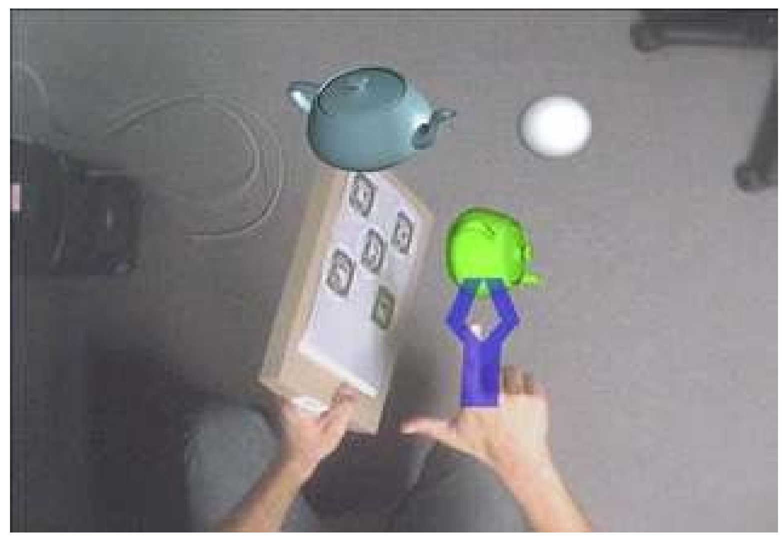
Tracking and recognition





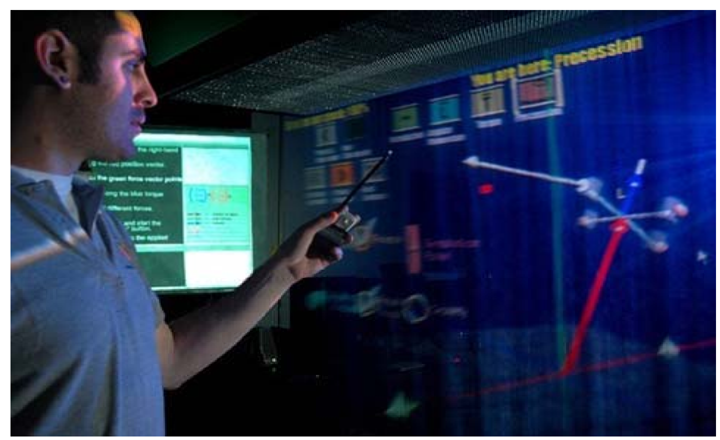


Augmented environments

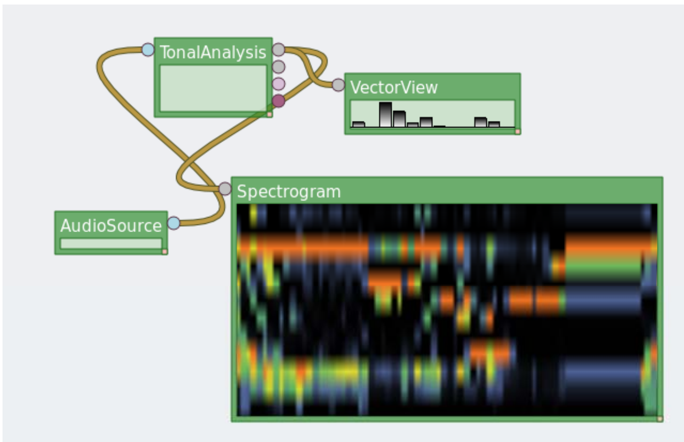


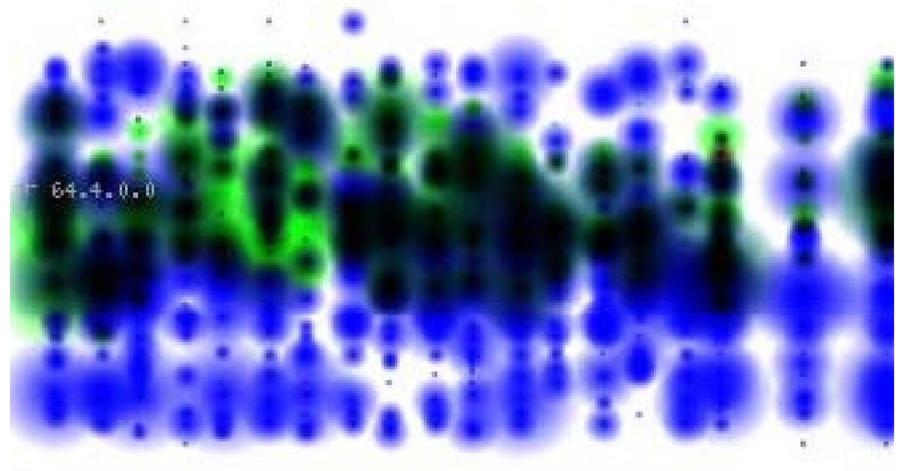
MAT Media Arts and Technology
Graduate Program

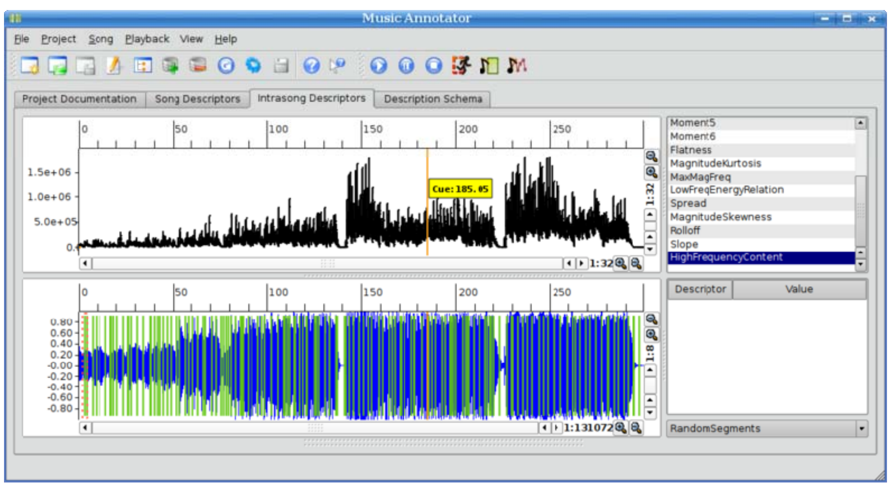
Interactive displays

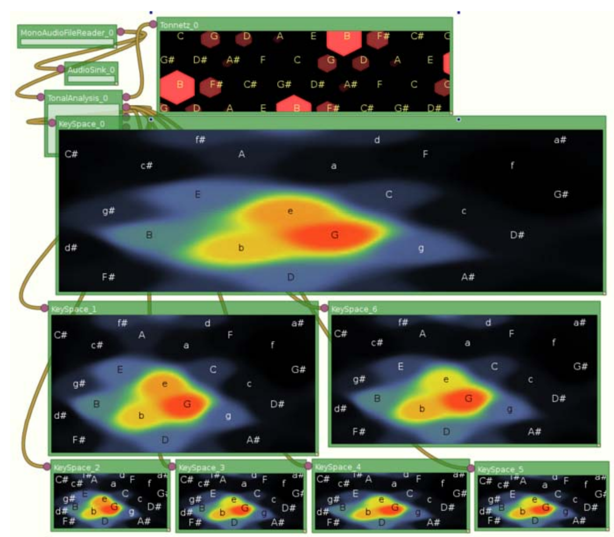


Sound synthesis



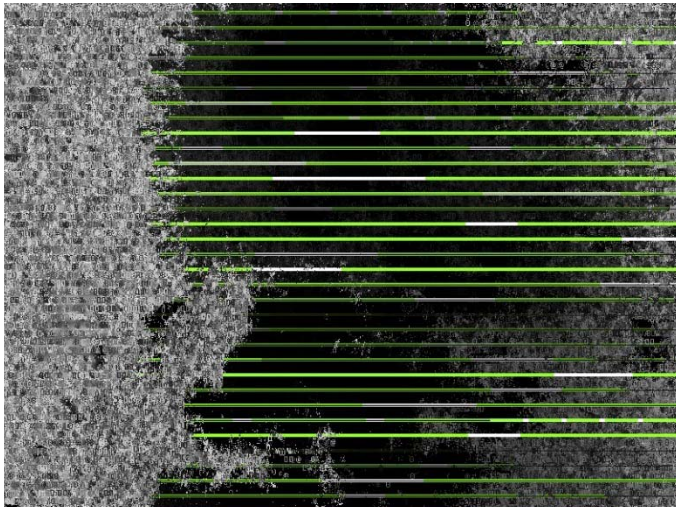


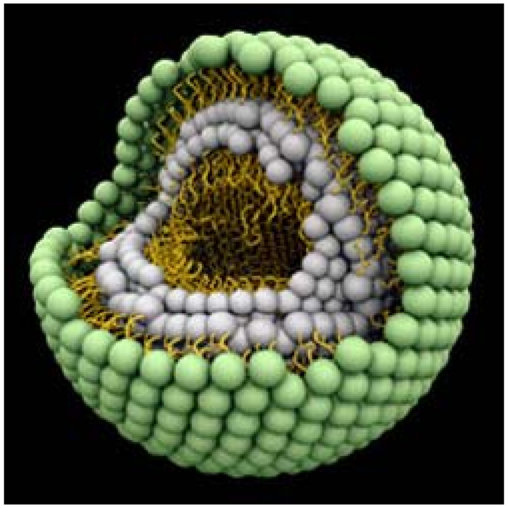


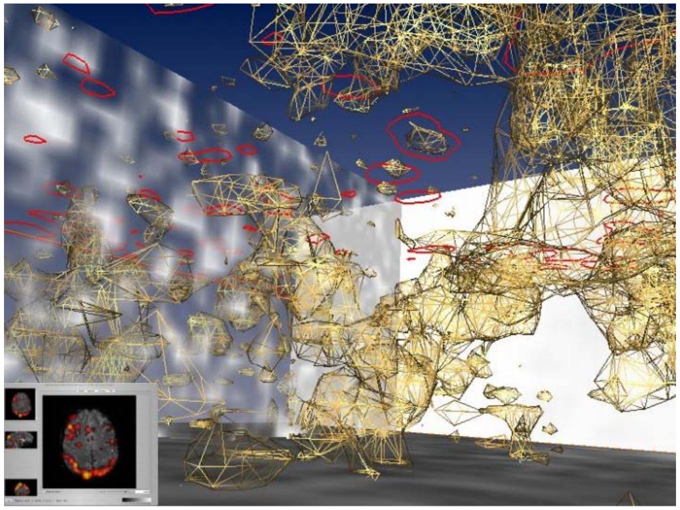


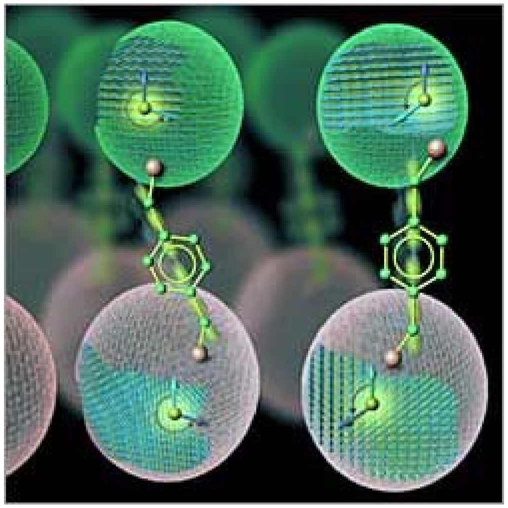
Media Arts and Technology
Graduate Program

Scientific visualization and auralization

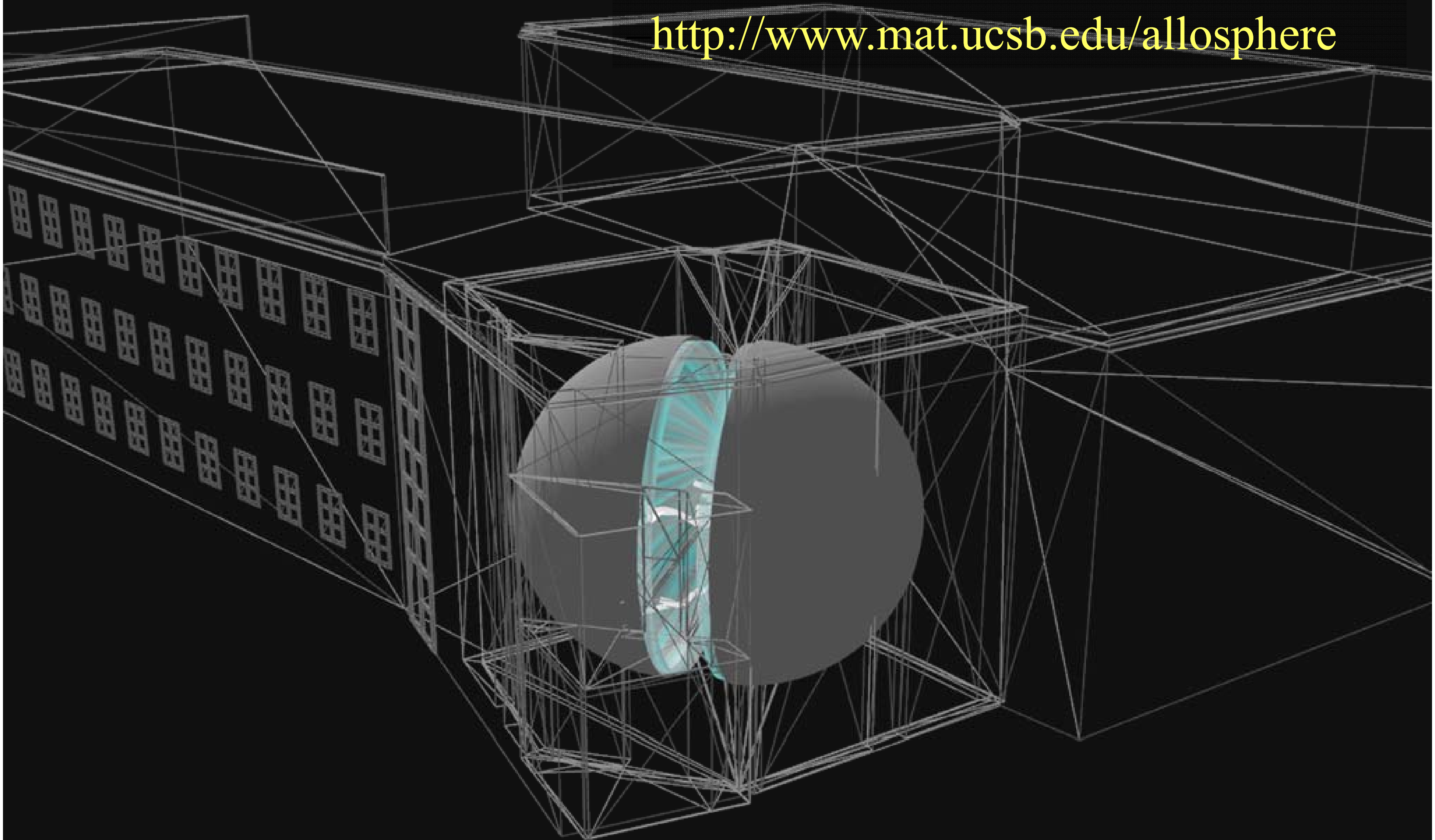








http://www.mat.ucsb.edu/allosphere



The Allosphere

What is the Allosphere?

- A three-story anechoic space containing a built-in spherical screen, 10m in diameter, and a walkway through the center
- A large-scale immersive surround-view instrument
- A digital media center in the California Nanosystems Institute
- A cross-disciplinary community around the UCSB Media Arts and Technology Program

- An advanced instrument for scientific research
  - The manipulation, exploration and analysis of large-scale data sets
- ... and for artistic exploration

Acknowledgements

- Tobias Höllerer, Mathias Kolsch, Rogerio Feris, Ya Chang, Haiying Guan, Changbo Hu, Longbin Chen, Sebastien Grange, Charles Baur, Taehee Lee, Ismo Rakkalainen, Ramesh Raskar, Andy Beall, Jim Blascovich, Jeremy Bailenson, Daniel Vaquero, JoAnn Kuchera-Morin, Allosphere group
- MERL, IBM, Nokia
- NSF

Computer Science Department and

Media Arts and Technology Program

University of California, Santa Barbara

http://www.cs.ucsb.edu/~mturk