The Camera Mouse and Other Interfaces for People with Severe Motion Impairments

CS 640 AI, Fall 2023

Margrit Betke

Department of Computer Science

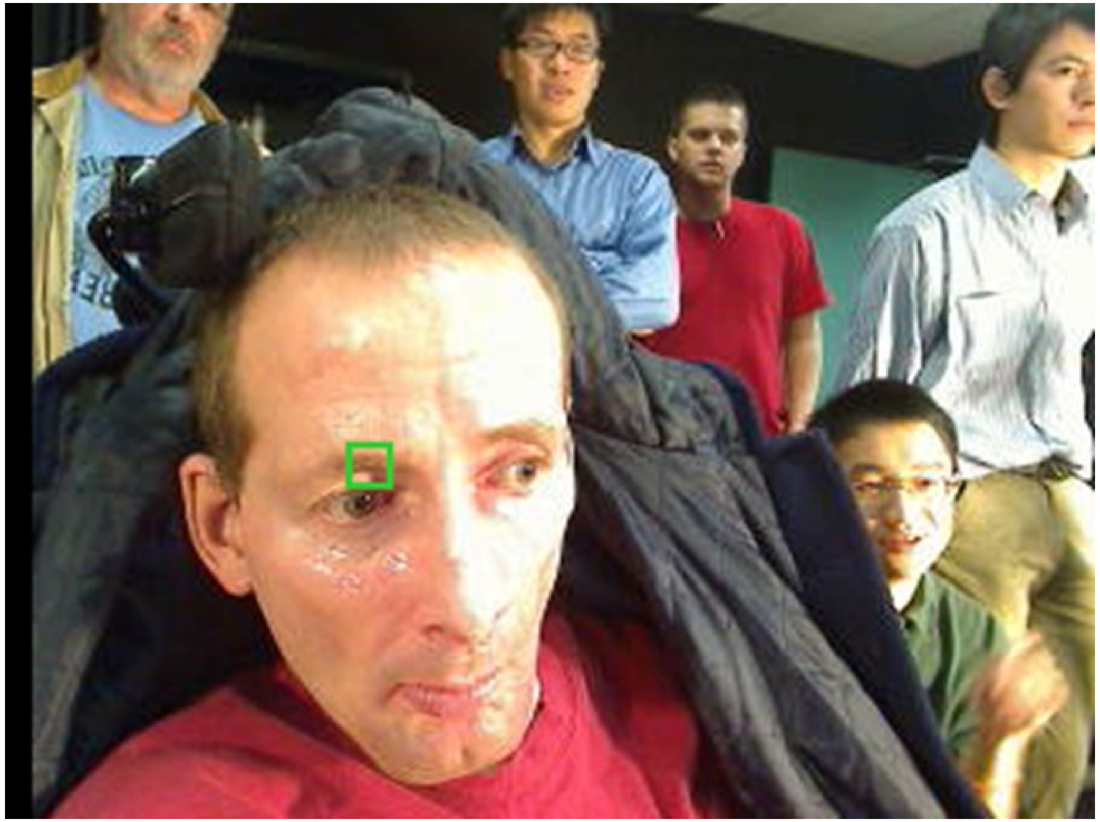
Boston University

Learning Outcomes of this Lecture

- Understand communication needs of users with severe physical disabilities
- Can explain how traditional and neural-network-based computer vision techniques can be used to detect and track facial features
- In particular, can explain correlation-based template tracking
- Understand mapping from camera coordinate system to screen coordinate system for converting facial feature movements to mouse pointer movements
- Can describe assistive software for people with motion disabilities
- Know about various input mechanisms: facial features, gaze direction, eyebrow raises, blinking, etc.
- Understand difficulties in conducting experiments with users with severe disabilities
- Can explain Fitts' law

Rick Hoyt, our Interface Tester

- Boston Marathon icon Rick Hoyt dies at 61 – YouTube, CBS



Users in Need

Millions of people with

- Severe cerebral palsy (CP)
- Traumatic brain injury
- Stroke
- Multiple sclerosis (MLS), muscular dystrophy
- ALS (Lou Gehrig's disease)

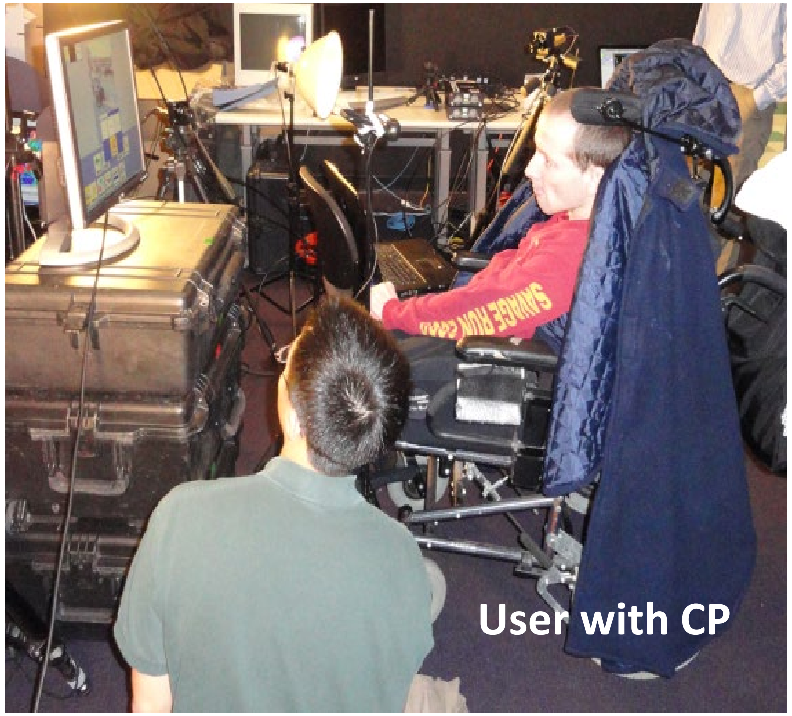
cannot communicate with traditional means:

- Often nonverbal
- Limited voluntary motion

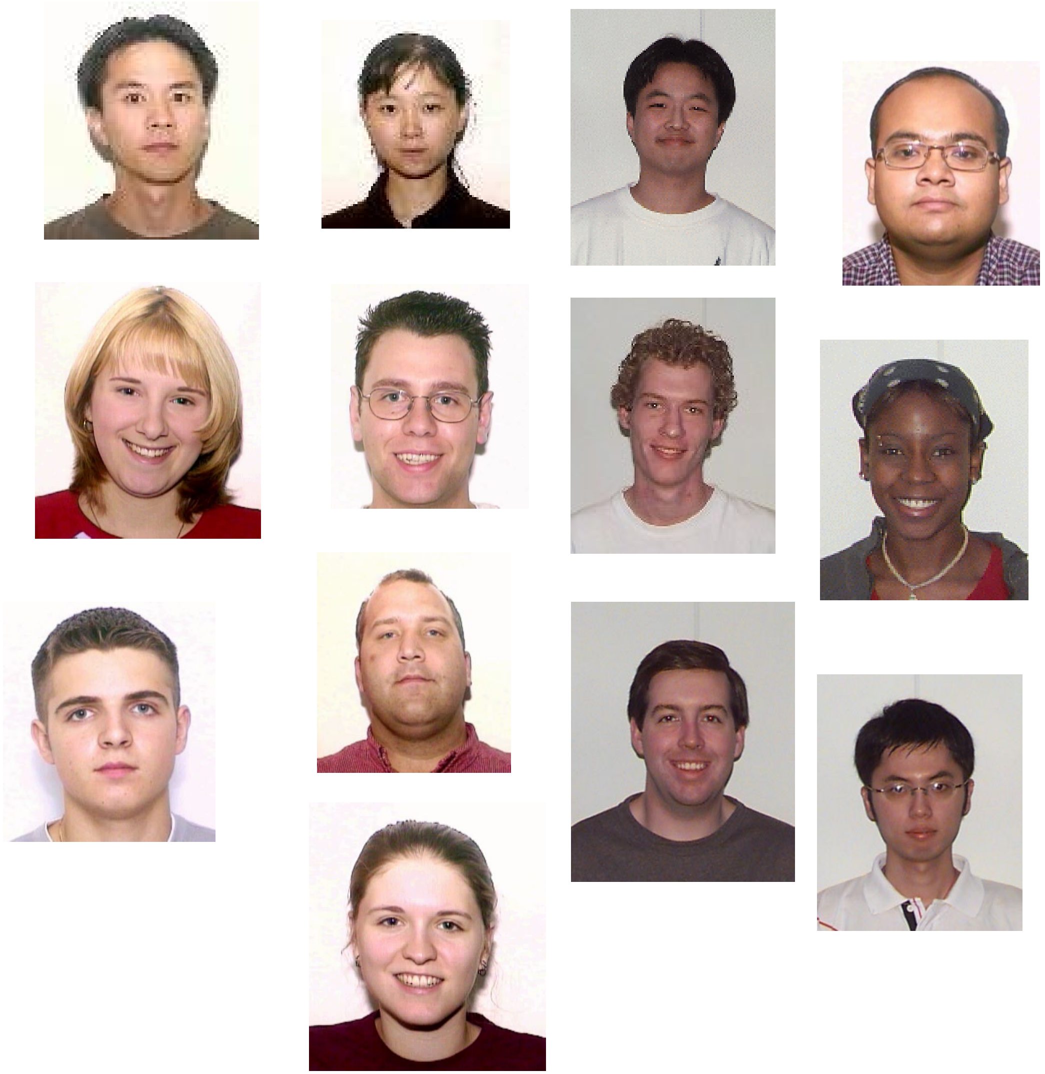
Lack of communication ability ≠ lack of active minds!

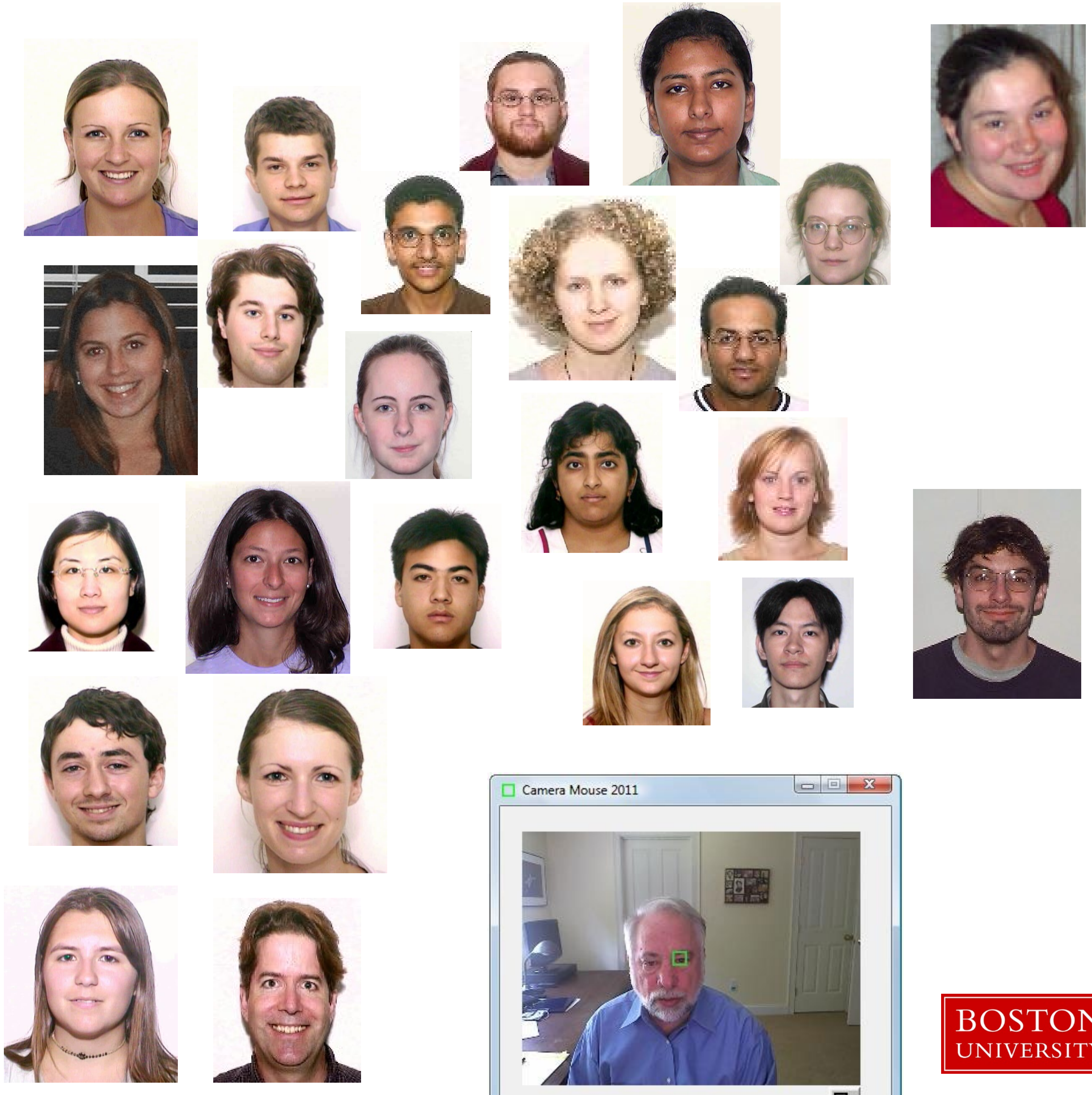
Communication technology often not available or too expensive, inefficient, difficult, etc.

Intelligent assistive interfaces can greatly improve the lives of people with severe paralysis



My Team (+ many)





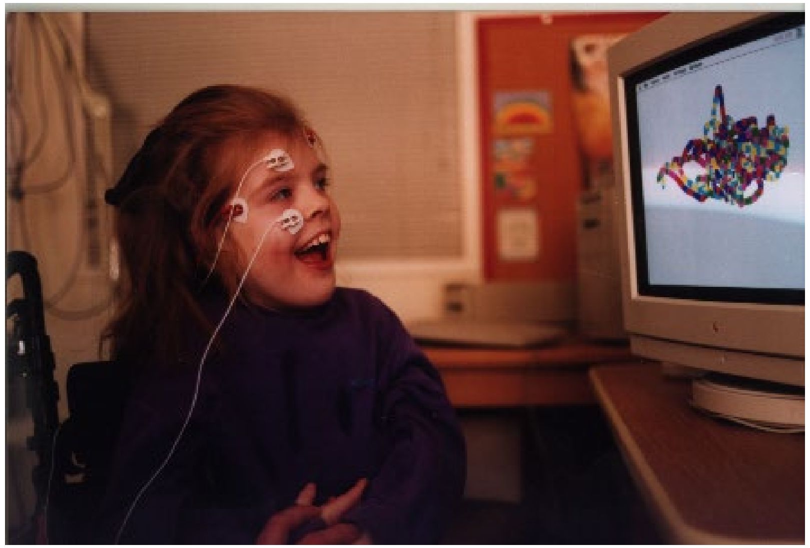
Traditional Approaches

Binary Switch
(blue button)

- Touch Switches
  - Hit plate
  - Wobble stick
  - Grip handle
  - Pinch
  - Pull string
- Photocell Switches
- Sip or Puff Switches
- Voice activated Switches

Traditional Approaches

EagleEyes by J. Gips at Boston College



Severe paralysis may leave the eyes as the only muscles that a person can control

Gaze direction is detected through electro-oculography

Traditional Approaches

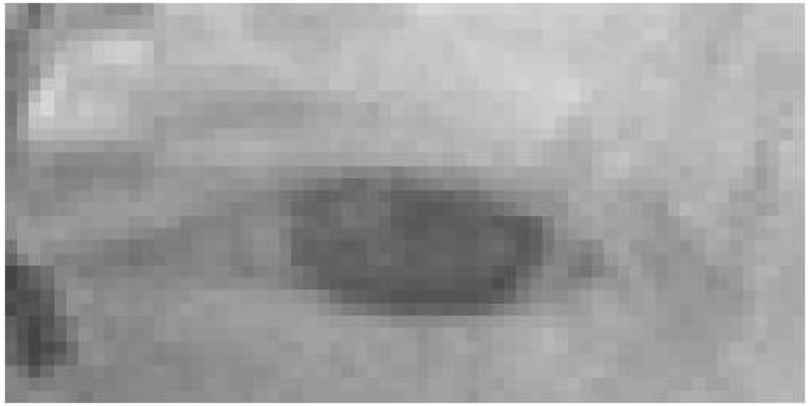
Active Infrared Lighting for Gaze Detection

Non-commercial custom-made hardware

(e.g., IBM's Blue Eyes)

Relatively expensive commercial hardware

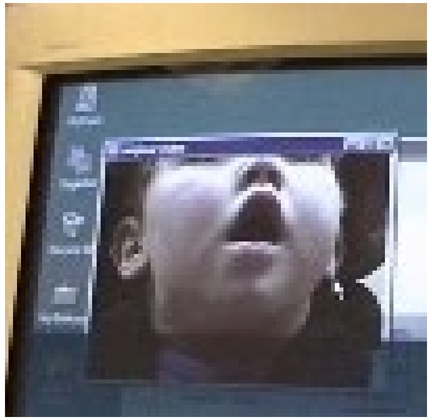
(e.g., Applied Science Laboratories)



- Calibration procedure?
- Long term effect of infrared light on eyes?
- Costs?

Gestures of MS/CP Patients

The Camera Mouse



The Camera Mouse

- Camera-based tracking of body to enable control of a mouse pointer
- Has been commercialized and sold to individuals, schools, and hospitals in the US and Europe

User with ALS

Camera Mouse Users (2001)

Age	Gen-der	Condition	Continuing to Use?	
2	M	CP	Y	Obtaining a system for home.
3	F	CP	Y	First user with home system.
6	F	CP	Y	Spelled name. Obtaining a home system.
8	M	CP	Y	Spells naughty words and laughs.
11	M	CP	Y	Obtaining a home system.
14	M	CP	Y	Spells words. Obtaining a home system.
15	M	CP	N	Close, but could not control reliably.
19	M	CP	N	Does not have sufficient muscle control.
23	M	TBI	N	Does not have sufficient muscle control.
31	M	TBI	Y	Spelled "TAKE OFF DAD"
37	M	CP	Y	Spelled "MERRY CHRISTMAS"
58	M	CP	Y	Spells, explores internet on home system.

Impact

- Numerous camera mouse users with cerebral palsy, multiple sclerosis, ALS, traumatic brain injuries
- Camera Mouse currently used

  In Schools
  In Hospitals
  In Nursing Homes
  At Home

  in Australia, England, Indonesia, Ireland, Turkey, USA, Uzbekistan

- 26 schools in Northern Ireland obtained the Camera Mouse in 2003
- Free download from www.cameramouse.org since April 2007
- Several millions of downloads

Thank-you email from a user

From: Mesutulbe [mailto:mesutulbe@gmail.com]
Sent: Thursday, July 27, 2006 1:46 AM
Subject: I'm grateful to you

I'm so grateful to you. Because I'm a MS(Multpl Sclerosis) patient since 20 years. I can't move my finger. However I'm in internet fr 10 hour at this time. Thank you very much for everything.

Sincerly

Dr.Mesut

Thank-you email from a user

Sent: Thursday, 4 September 2008 16:53:58 +1000

G'day,

[..] a great big thank you from all the way over here in Australia. I have found your fantastic program while looking for an alternative to the expensive trakir and other type head tracking software. [..] I already have a degree of lack of movement in my right arm, shoulder, through to my elbow which also incorporates my ability to grip tightly with my right hand. I am right handed so this is a real pain [..] as well as my condition giving me great pain in movement. I do not take analgesics, I try and keep surfing and swimming and try to keep moving the parts as much as possible. What I am finding is that the program works very well, I am using my daughter's ASUS lap top with built in camera, I can control the cursur OK and feel with more practice will become very proficient. I thank you and your team for your outstanding work in this area.

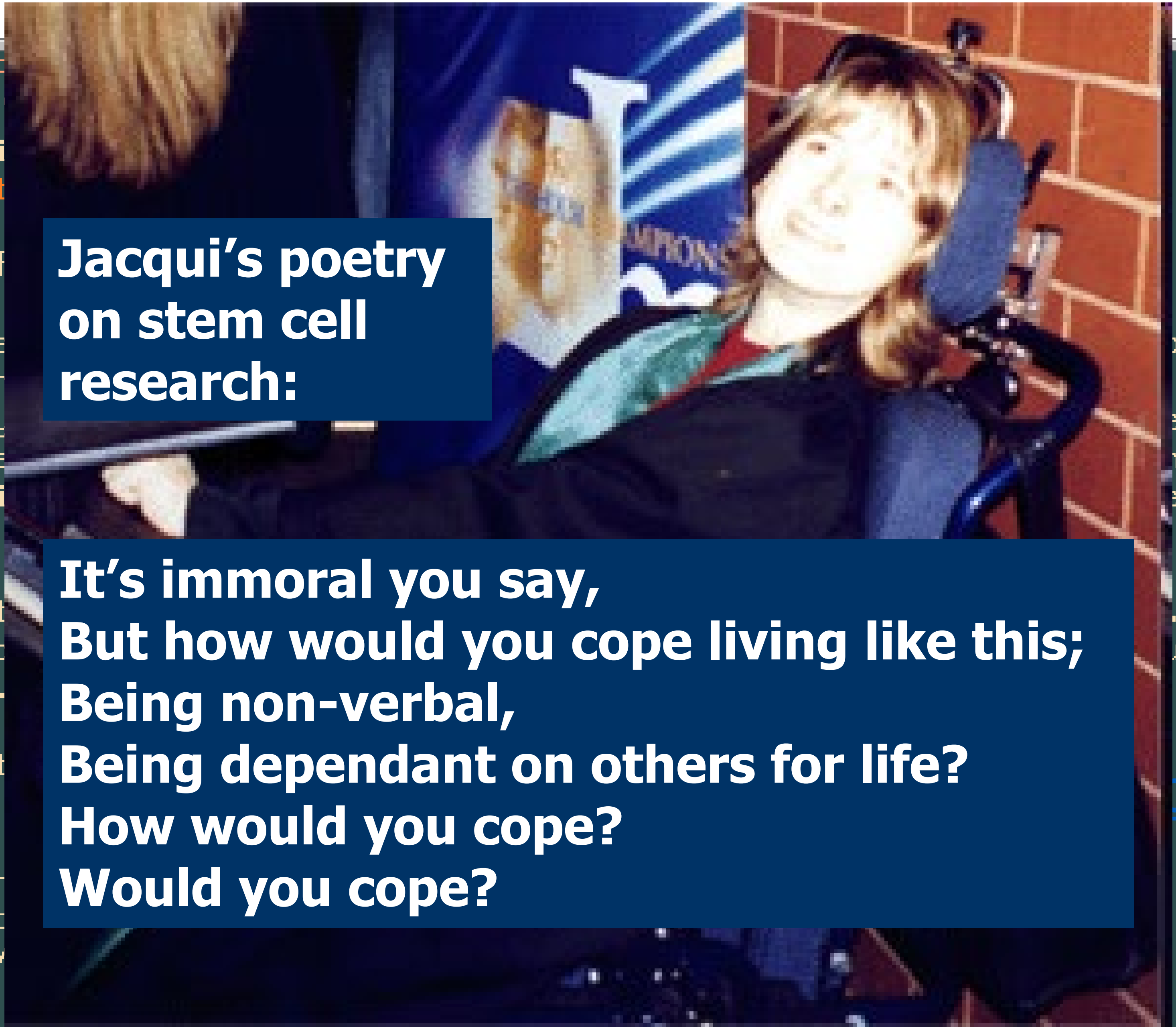
Best Regards to you all Over there.

L. P.

Thank-you email from a user

```
Date: Fri, 9 Jan 2009 14:29:42 +1100 (AUS Eastern Daylight Time)
From: "Jacqui Rogers" <jacquelinerogers@optusnet.com.au>
To: <gips@bc.edu>
Subject: thank you

Hi Professor,

My name is Jacqui and I have severe Athetiod Cerebral Palsy.  For about
sixteen years my only form of computer access was with the scanning program
E Z Keys with a headswitch. Now thanks to you I'm using Camera Mouse with
the onscreen keyboard program Grid Keys. Having an alternative to the
headswitch has really changed my life as I don't suffer from fatigue due to
typing as much now.

I can't express how grateful I am to you for devolping Camera Mouse.
Thankyou so much!!! I hope you don't mind but I've put a link to your site
on www.jacquirogers.id.au

Kindest Regards,
Jacqui Rogers.


----------------------------
www.jacquirogers.id.au ; www.gigsnreviews.com (music blog) ;
http://thinkingpoet.blogspot.com/ (poetry blog)
If you have time, please click on ads on my blogs. Thank you!
```

Thank-you email from a user



Ongoing Camera Mouse Work

From: <hXXX@brown.edu>
Sent: Tuesday, August 18, 2020 8:00 AM
To: Betke, Margrit <betke@bu.edu>
Subject: Camera Mouse Development

Hi Dr. Betke,

I am reaching out because you are the only other listed collaborator on the Camera Mouse website. It appears that Dr. Gips has unfortunately passed away in 2018, and there aren't any updates on the Camera Mouse website to indicate if further development of Camera Mouse is in progress.

It looks like that while Camera Mouse is free to download and use, it is not open source -- is that right? I am XXX at Brown University, and I was interested in improving Camera Mouse for personal use (my father has advanced Multiple Sclerosis and can no longer use his hands or arms). While there are some commercial head trackers available, the cost is quite prohibitive.

Do you have any information about what languages/software were used to create Camera Mouse, or if the backend code is available for non-commercial tinkering?

Thank you in advance for your help!

How does the Camera Mouse work?

How does the Camera Mouse work?

Up to 2018: http://www.cameramouse.org

Traditional computer vision techniques:

Tracking with normalized correlation and optical flow

In 2023:

Modern NN-based computer vision techniques:

Face and facial feature detection

2023 Camera Mouse Demo on Eagle Aliens

How does the Camera Mouse work?

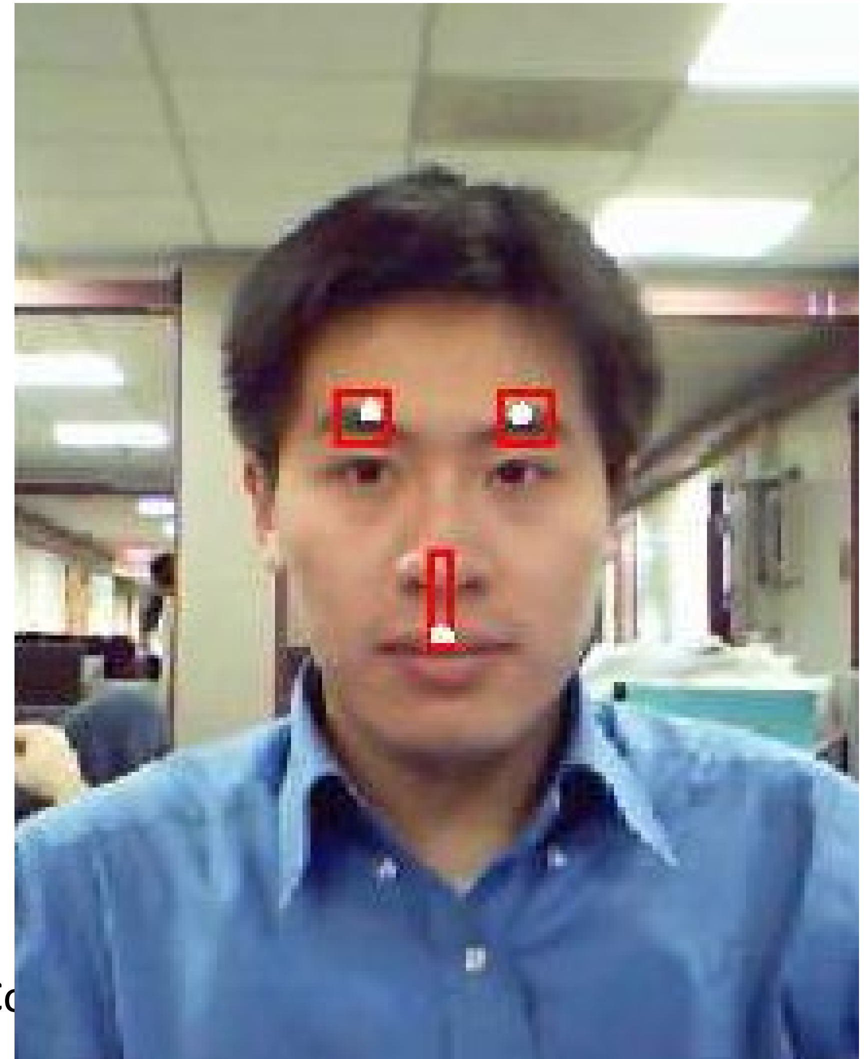
Up to 2018: http://www.cameramouse.org

Traditional computer vision techniques:

Tracking with normalized correlation and optical flow

Initialization of Tracker

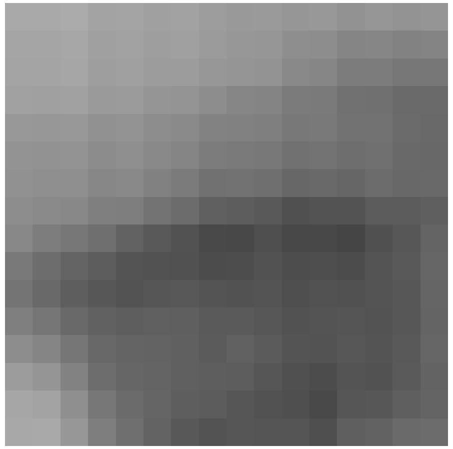
Manual or automatic initialization of feature to track

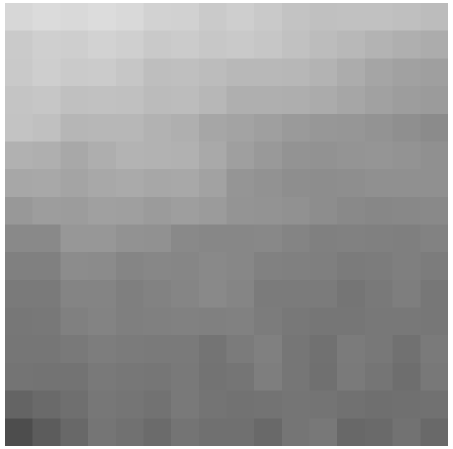
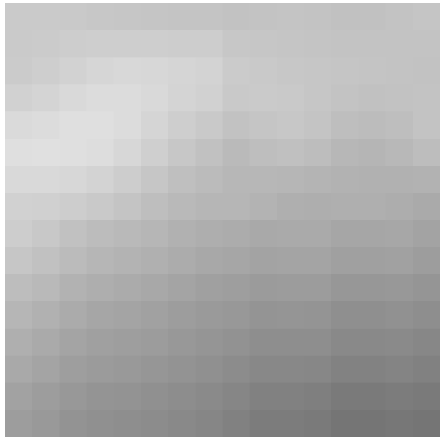


3 Tracking Points

Correlation-based Tracking

Templates = grayscale sub-images




Method:
Search for Best Matching Template

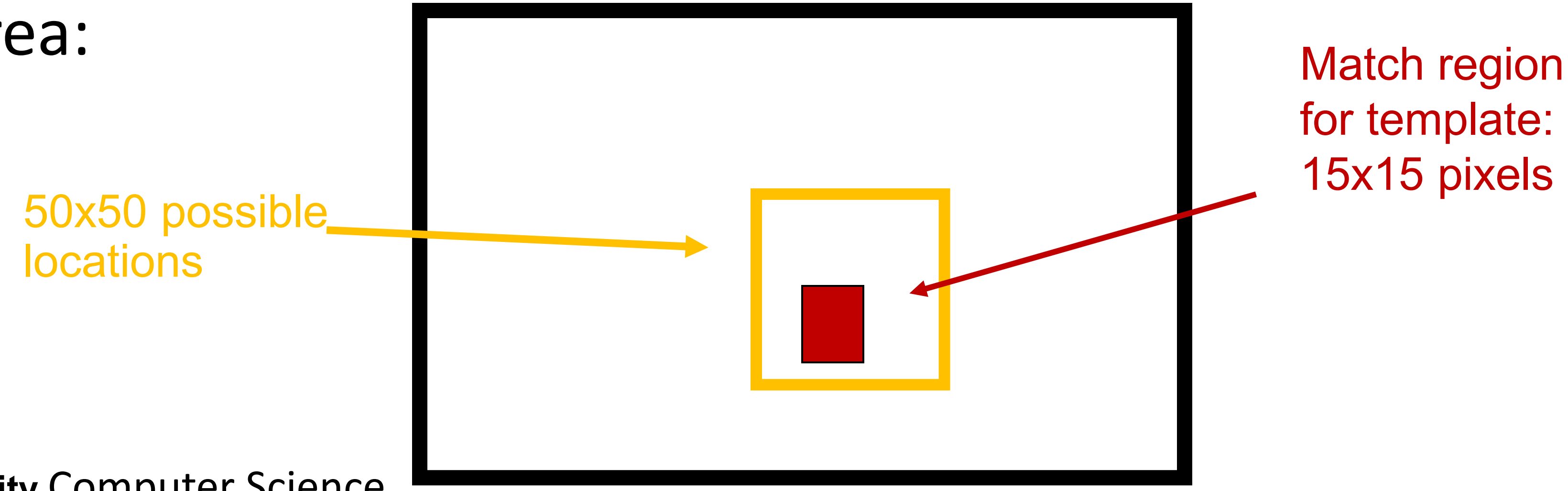
Normalized Correlation Coefficient (NCC):

$$r(s,t) = \frac{A\sum s(x,y)t(x,y) - \sum s(x,y)\sum t(x,y)}{\sigma_s \sigma_t}$$

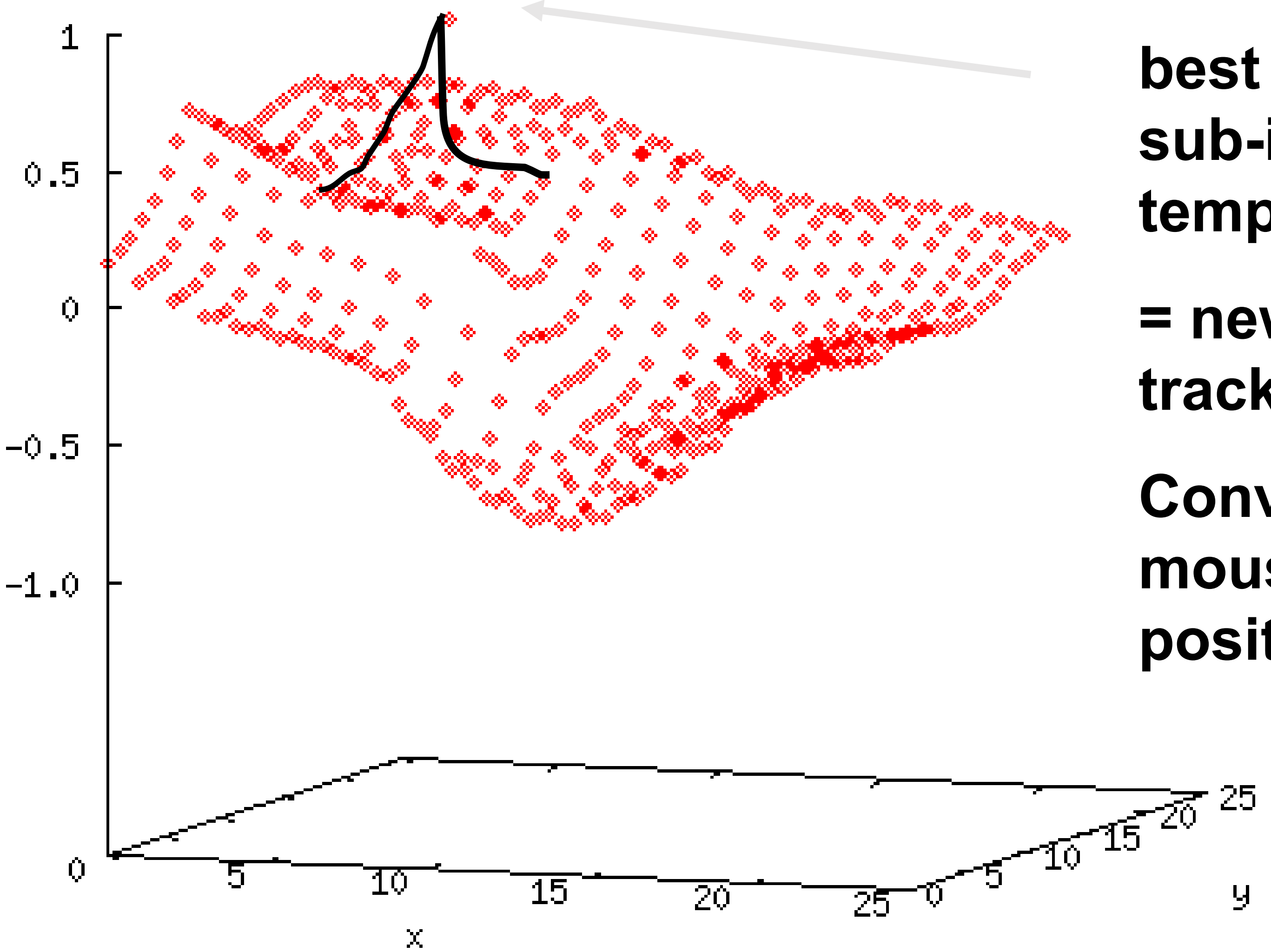
$A =$ number of pixels in template

$$\sigma_s = \sqrt{A\sum s(x,y)^2 - (\sum s(x,y))^2} \qquad \sigma_t = \sqrt{A\sum t(x,y)^2 - (\sum t(x,y))^2}$$

Search area:



Correlation Surface



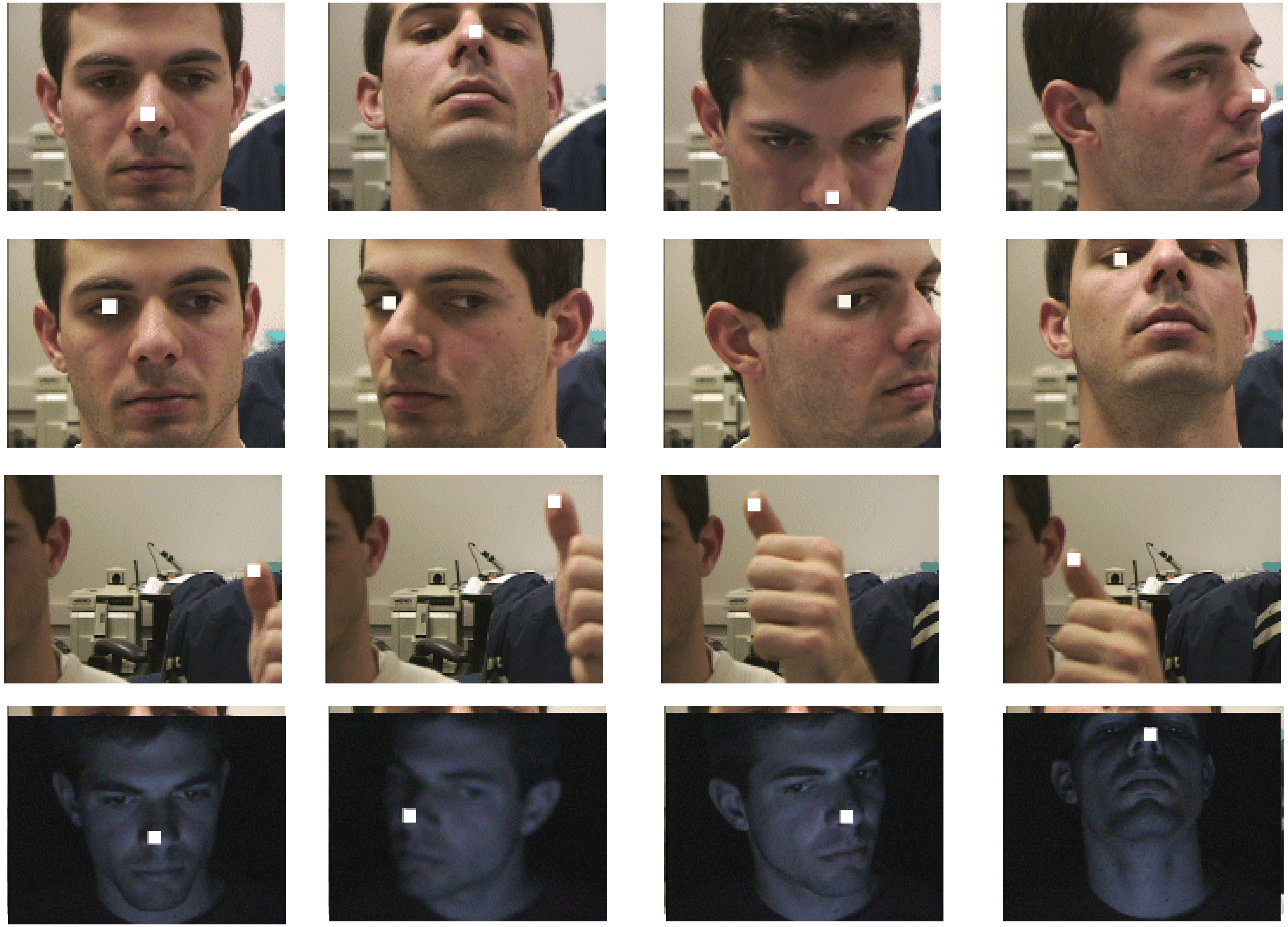
Peak =

best match of sub-image with template

= new location of tracked feature

Convert to new mouse pointer position

Tracked Features



How does the Camera Mouse work?

In 2023:

Modern NN-based computer vision techniques:

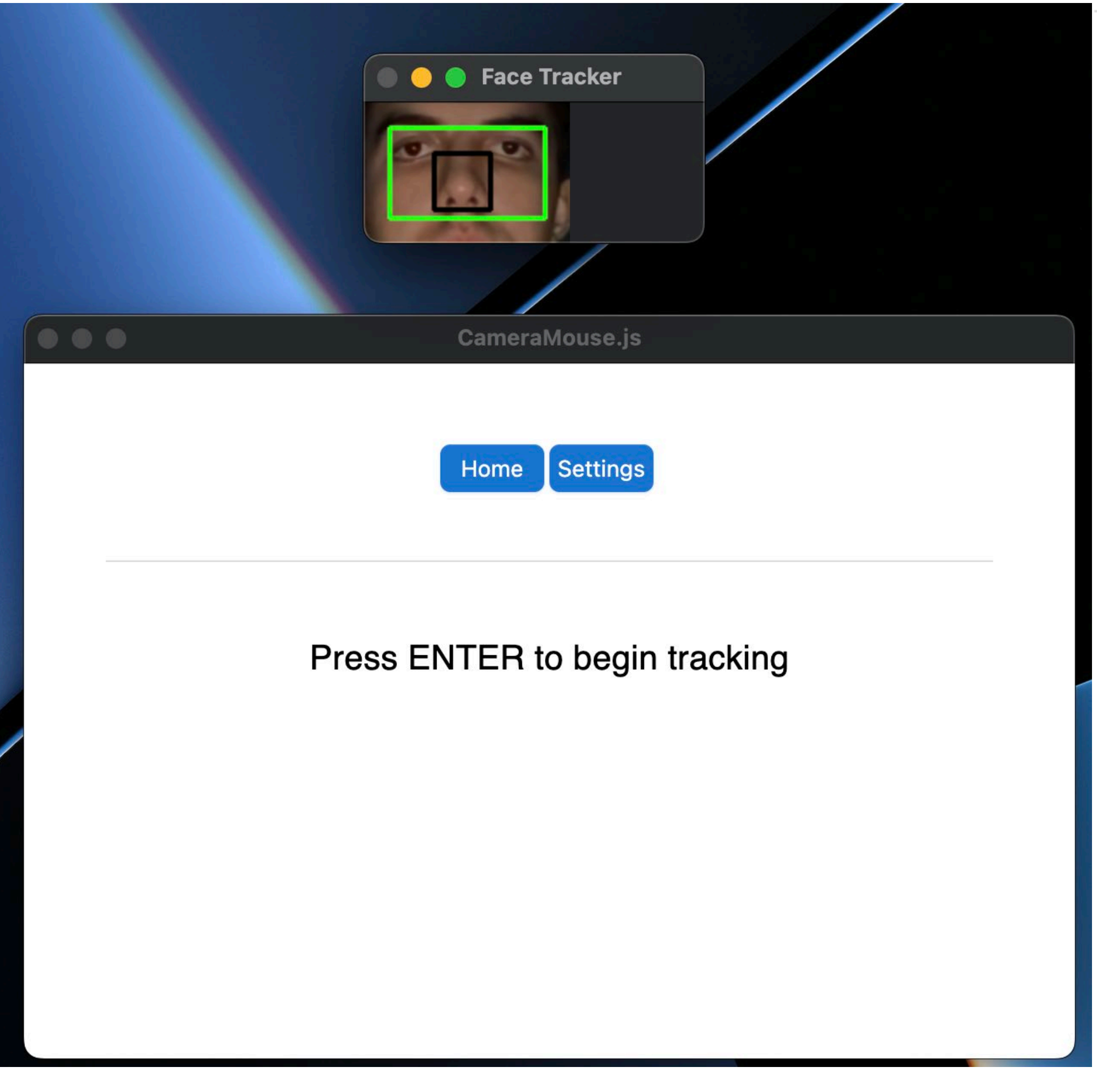
Face and facial feature detection

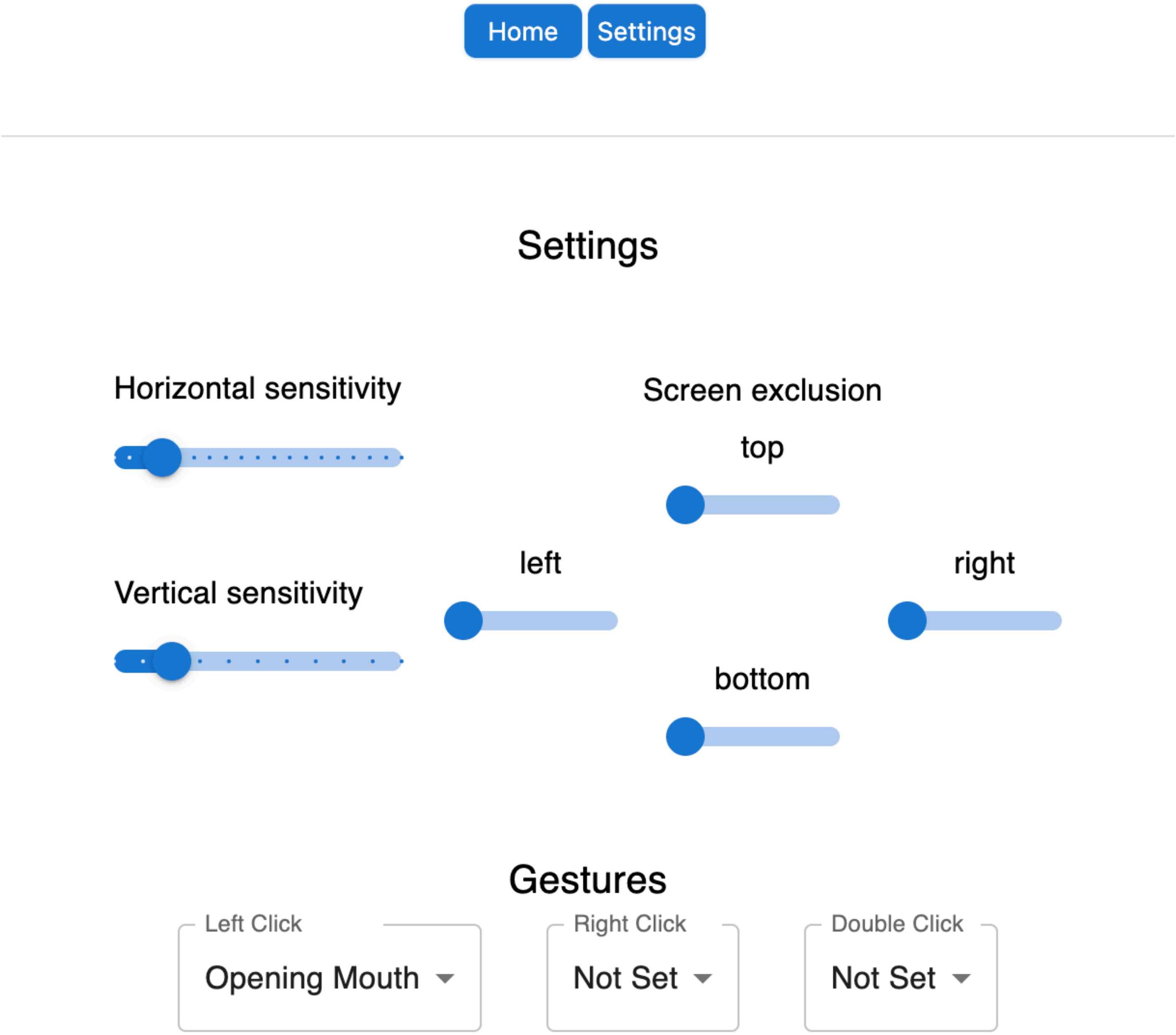
How does the 2023 Camera Mouse work?

2 components: Interface Core and Video Processing

- The Interface Core component provides the graphical user interface (GUI), handles mouse pointer movement, command issuing, and configuration changes.
- The Video Processing component handles facial feature detection and tracking and the mapping from camera to computer screen coordinates.
- The Interface Core is coded using the Electron open-source desktop application development framework (Node JS version 18).
- The Video Processing component is coded in Python 3.8 and uses the default video camera of the user’s computer (desktop or laptop).
- The two components communicate over HTTP.

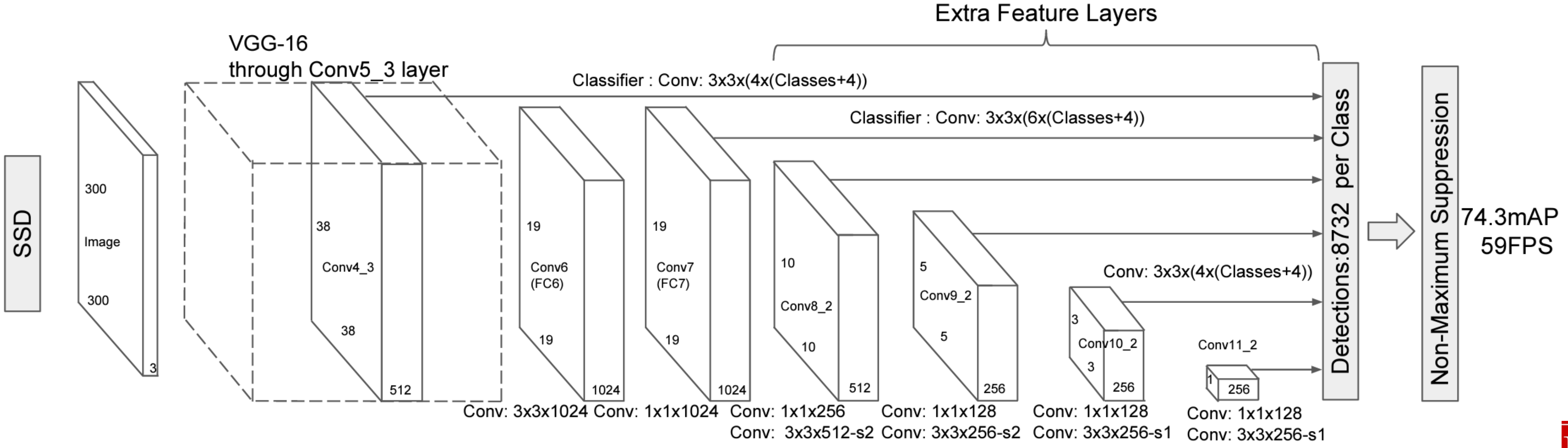
2023 Camera Mouse



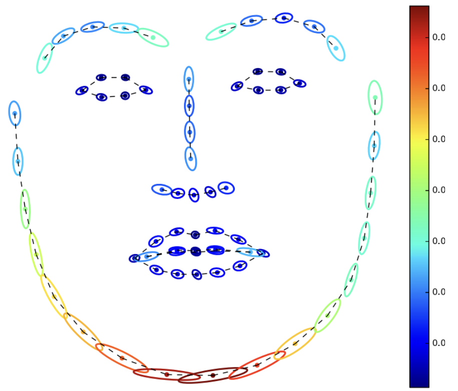


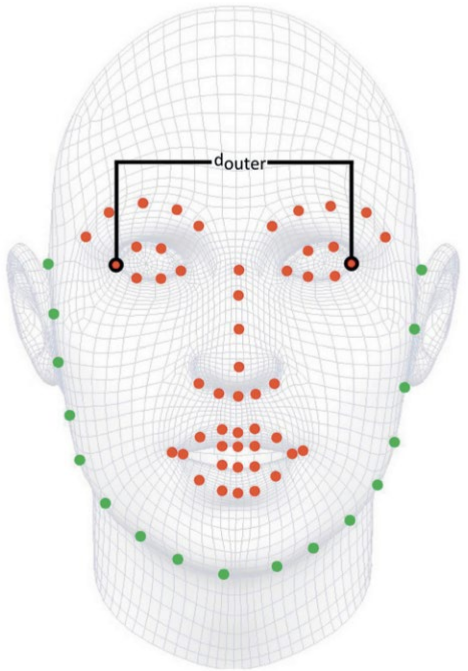
Face Detector

- CM 2023 uses the OpenCV implementation of the Single-Shot Multi Box Detector (SSD) Model by [Liu et al.]
- Useful for real-time face detection from built-in cameras (speed is an important property of SSD over other models)
- If more than one face is detected, CM will use the largest

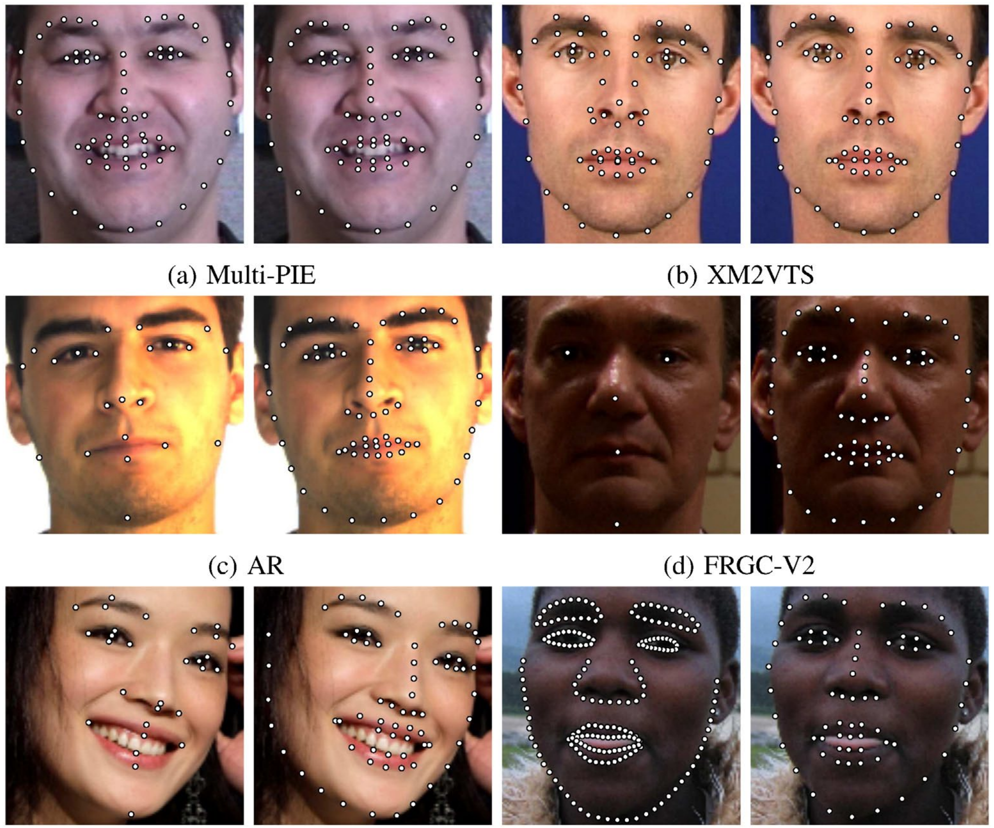


Fast Detection of 68 Facial Landmarks





- Pose estimation for face alignment: [Kasemi & Sullivan, 2014]
- Trained on [Sagonas et al., 2016]:



https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/

http://dlib.net/face_landmark_detection.py.html

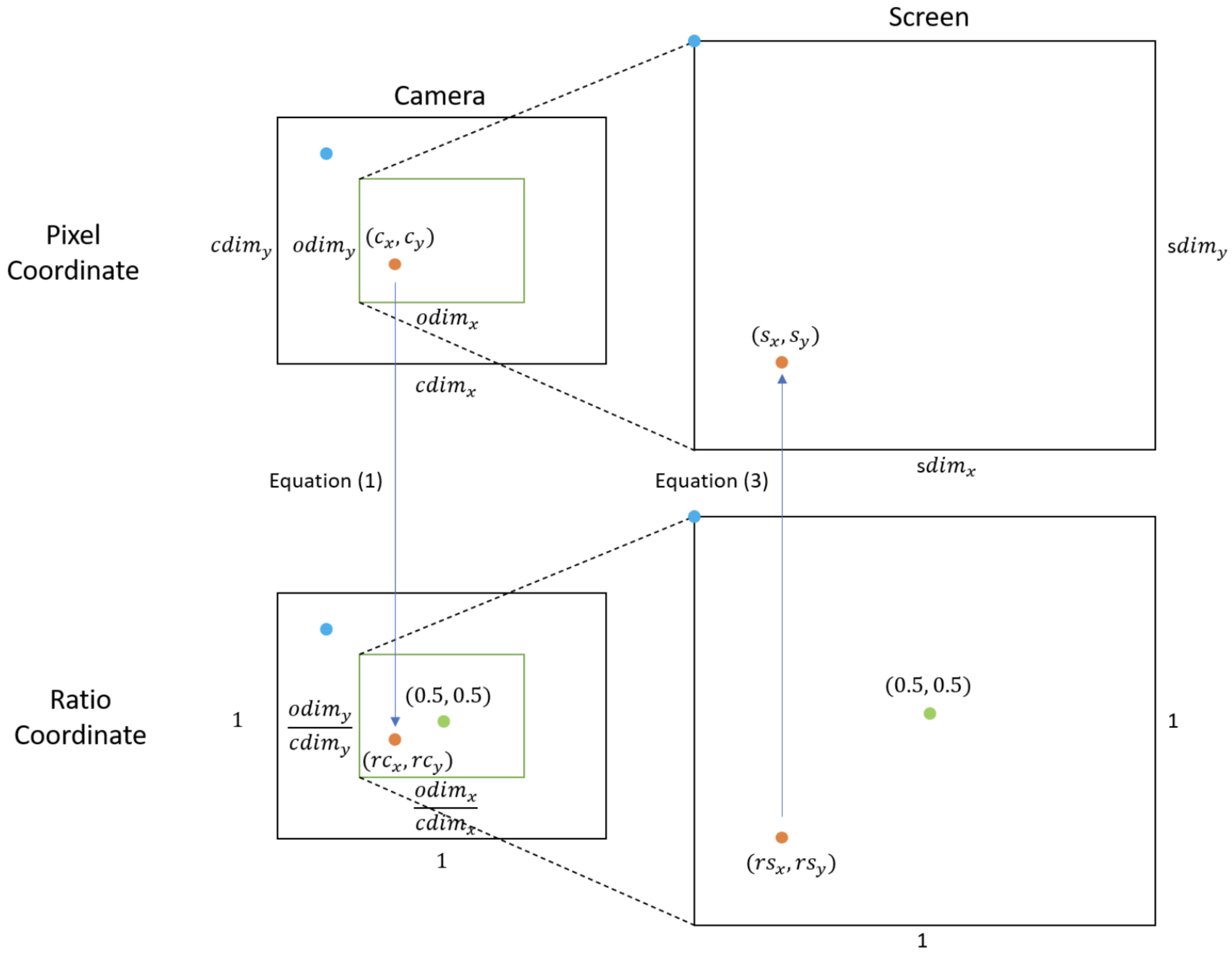
CM 2023: Mapping from Camera to Screen Coordinates

1. Conversion from camera dimensions into unit space: pixel coordinate (c_x,c_y) to ratio coordinate $(rc_x,rc_y)=(c_x/cdim_x,c_y/cdim_y)$
2. Mapping from (rc_x,rc_y) to (rs_x,rs_y)

   $rs_x = 0.5-(0.5-rc_x)/(odim_x/cdim_x)$ if in green box [0 or 1 if outside]

   rs_y similar

   $(odim_x,odim_y)$=green box dimensions
3. Conversion into screen dimensions: ratio coordinate $(s_x,s_y) = (rs_x\ sdim_x, rs_y\ sdim_y)$



CM 2013: Tracking

- Same as initial Camera Mouse: Template matching with the normalized correlation coefficient
- Track point drift detected by comparing the currently tracked point with the nose tip initialization.
- Re-initialize tracking if distance above a threshold

CM 2013 Gesture Commands

In addition to dwell-time clicking, we implemented:

Open-mouse clicking:

To detect an open mouth, we set a threshold for the mouse aspect ratio.

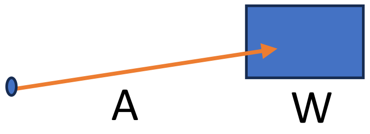
Eyebrow-raise clicking:

To detect raised eyebrows, we set a threshold for the distance of the edges of the eyebrows from the nose, accounting for distance of the face from the camera.

CM 2023: Ongoing User Experiments

3 Types of Experiments:

1. Movement Throughput Test
2. Browser navigation test
3. CM typing with a virtual keyboard

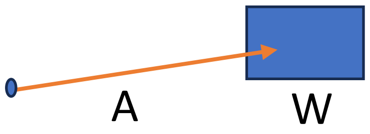


CM 2023: Ongoing User Experiments

3 Types of Experiments:

1. Movement Throughput Test

   Why throughput, not just speed?



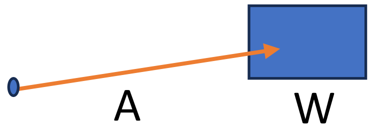
2. Browser navigation test
3. CM typing with a virtual keyboard

CM 2023: Ongoing User Experiments

3 Types of Experiments:

1. Movement Throughput Test:

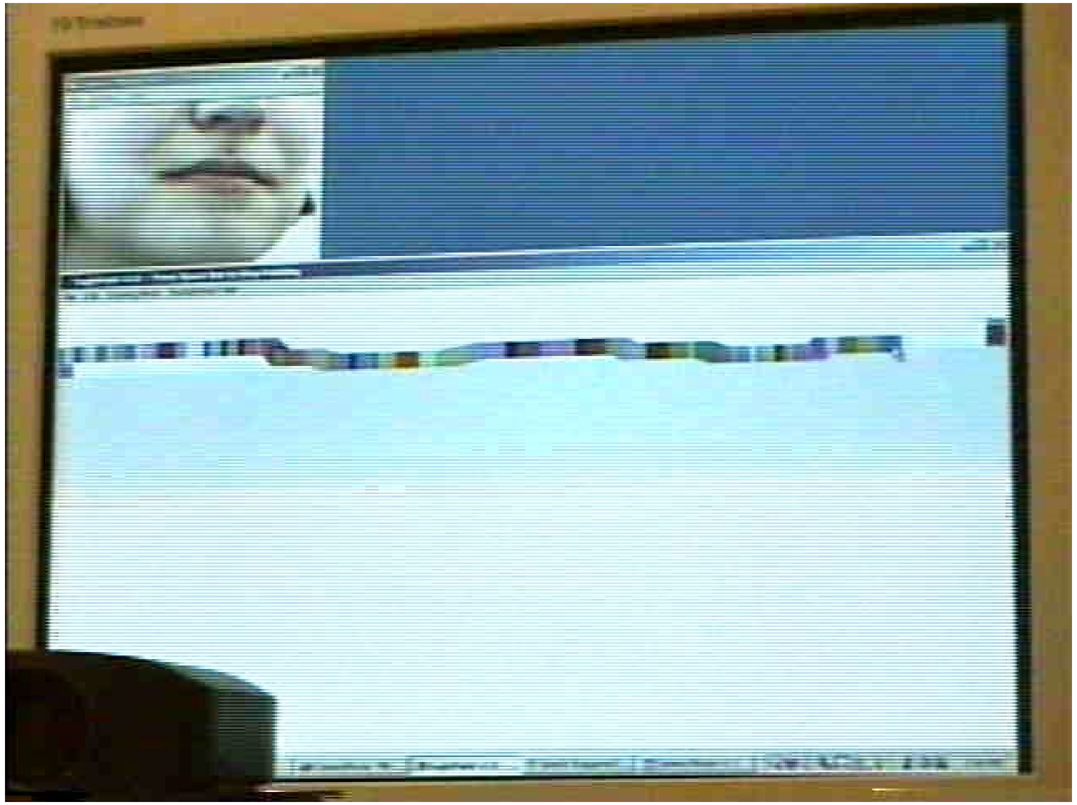
   To evaluate accuracy and speed of the CM, we measure the average throughput in [bits per second] of reaching targets and clicking, using Fitts' Law



2. Browser navigation test
3. CM typing with a virtual keyboard

Previous Experiment: Painting a Line

Subject was asked to draw a straight horizontal line from left to right



Painting a Line

User with Disabilities

		Nose	Lip	Eye
Horizontal Line	Trial 1	9 s		
Horizontal Line	Trial 2		21 s	
Horizontal Line	Trial 3			20 s
Horizontal Line	Trial 4			4 s
Horizontal Line	Trial 5			17 s
Vertical Line	Trial 1			11 s
Vertical Line	Trial 2			6 s
Vertical Line	Trial 3	28 s		
Vertical Line	Trial 4	55 s		

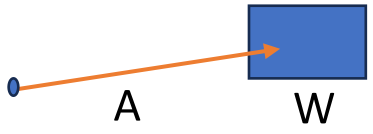
Average: 19 s, Std Deviation: 15.6 s

CM 2023: Ongoing User Experiments

3 Types of Experiments:

1. Movement Throughput Test:

To evaluate accuracy and speed of the CM, we measure the average throughput in [bits per second] of reaching targets and clicking, using Fitts' Law

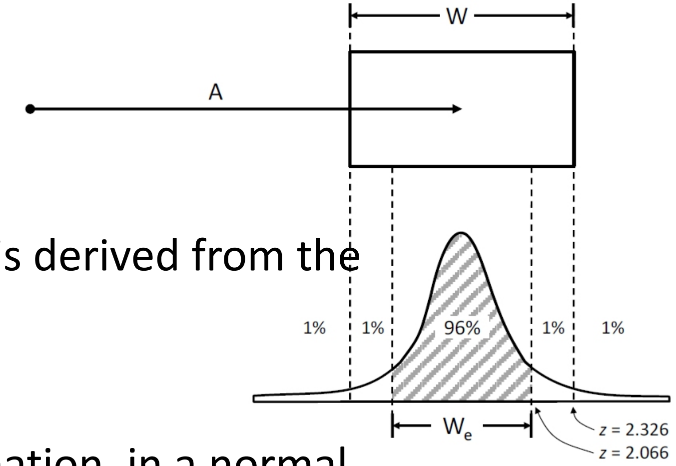


Background on Fitts' Law:

- Derived from Shannon's Law in Information Theory on "Information Capacity" of a communication channel = B $\log_2$ (S/N +1), where B = bandwidth, S = signal power, N=signal noise.
- Here the analogy is S=movement amplitude A, and N=target width W

2. Browser navigation test
3. CM typing with a virtual keyboard

CM 2023: Fitts' Law User Experiments



To evaluate accuracy and speed of the CM at the same time, we measure the average throughput in [bits per second] of reaching targets and clicking, using Fitts' Law:

Throughput =

[Effective index of difficulty] / [Movement time]

Effective index of difficulty [bits] = $\log_2$(([Movement amplitude A] / [Effective target width W_e]) + 1)

Effective target width W_e = 4.133 times [the standard deviation of the selection coordinate] (z in figure).

Typical throughput values in user experiments with the computer mouse range between 3.7 bps and 4.9 bps.

Effective target width is derived from the distribution of "hits:"

The entropy, or information, in a normal distribution is $\log_2((2\pi e)1/2\sigma) = \log_2(4.133\ \sigma)$, where σ is the standard deviation in the unit of measurement.

96% of hits fall between -2.066 and 2.066 of z

If the selection coordinates are normally distributed, W_e spans 96% of the distribution. If the observed error rate was 4% in the sequence of trials, then $W_e = W$. If the error rate was greater than 4%, $W_e > W$. If the error rate was less than 4%, $W_e < W$ (see figure). By using W_e rather than W, a Fitts' law model more closely reflects what users actually did, rather than what they were asked to do.

From https://www.yorku.ca/mack/hhci2018.html

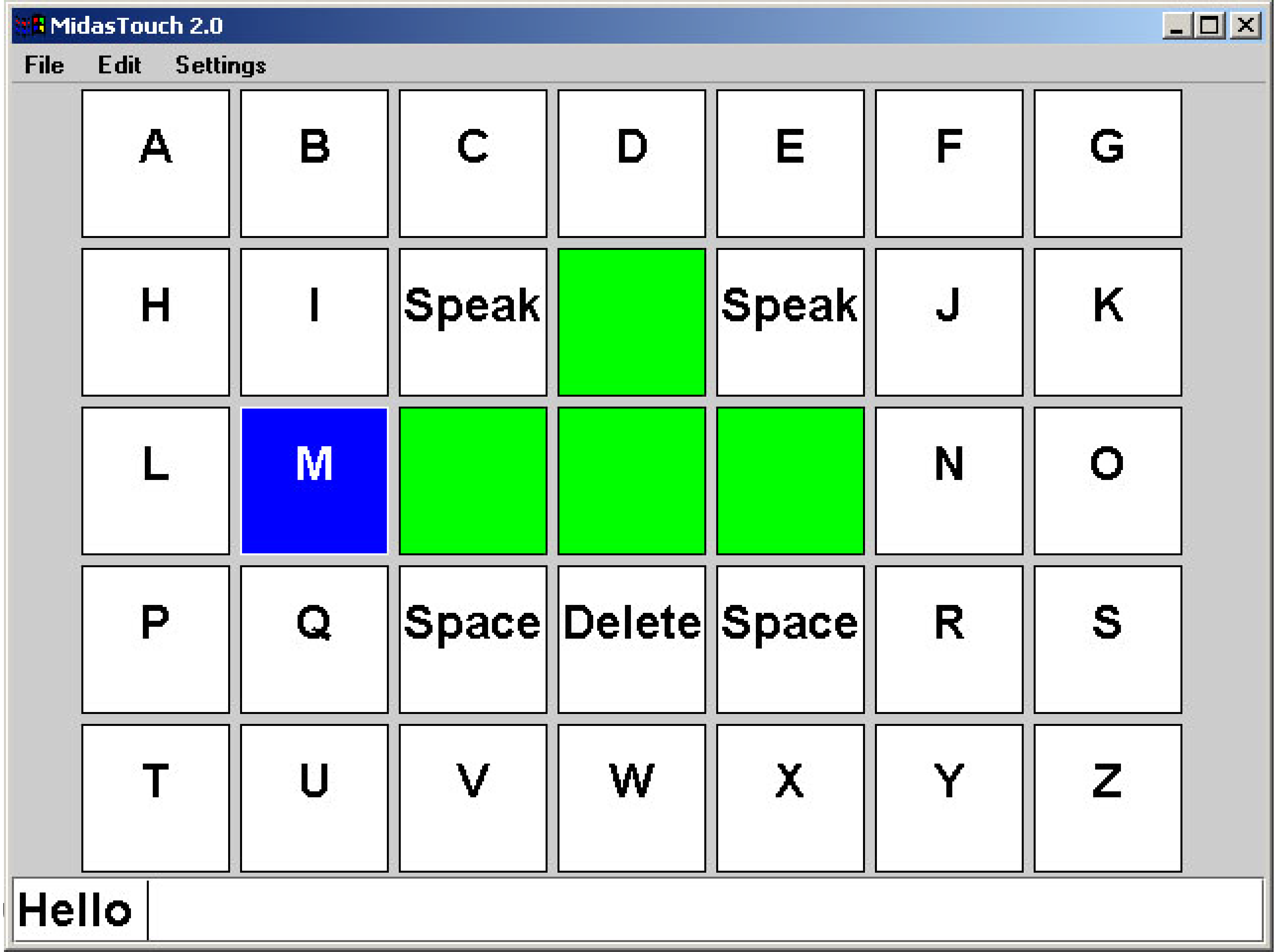
& https://en.wikipedia.org/wiki/Fitts%27s_law

Special Software for Camera Mouse Users

Special Software for Camera Mouse Users

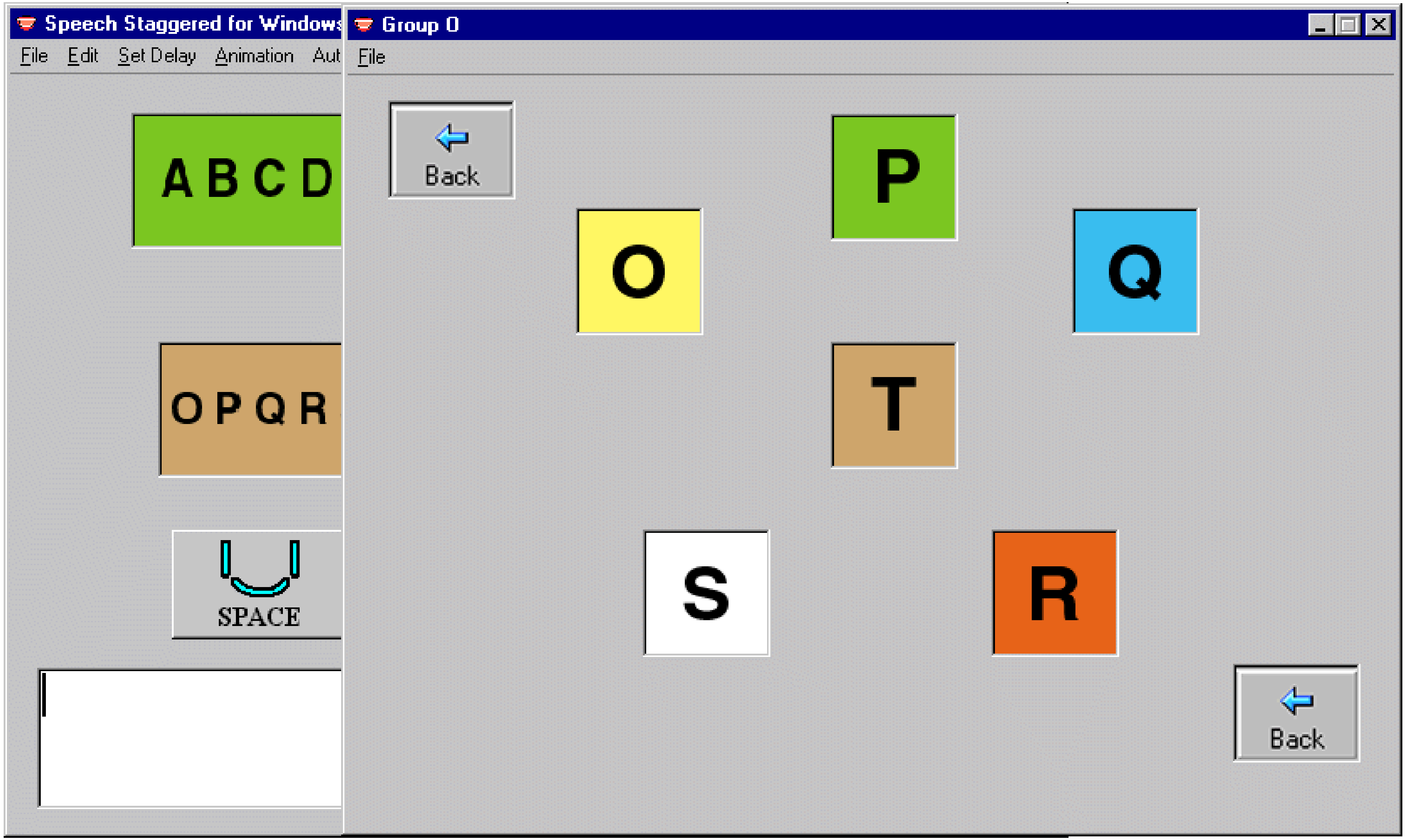
Test Entry

Entertainment



MidasTouch 2.0
File
Edit
Settings
A
B
C
D
E
F
G
H
I
Speak
Speak
J
K
L
M
N
O
P
Q
Space
Delete
Space
R
S
T
U
V
W
X
Y
Z
Hello

Two-level Text Entry



Experience

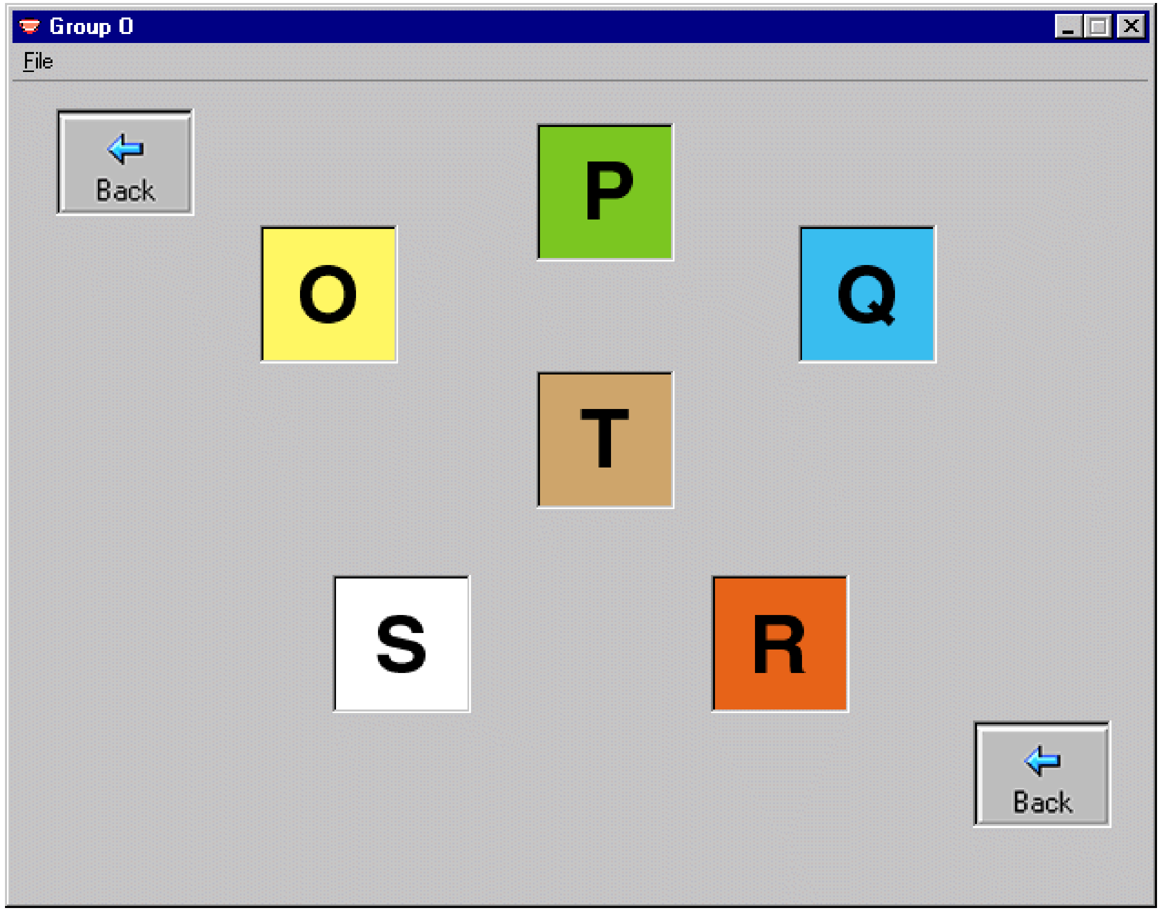
Each user was asked to spell 3 words:

"RAINING"

"MINIMAL"

"POOR"

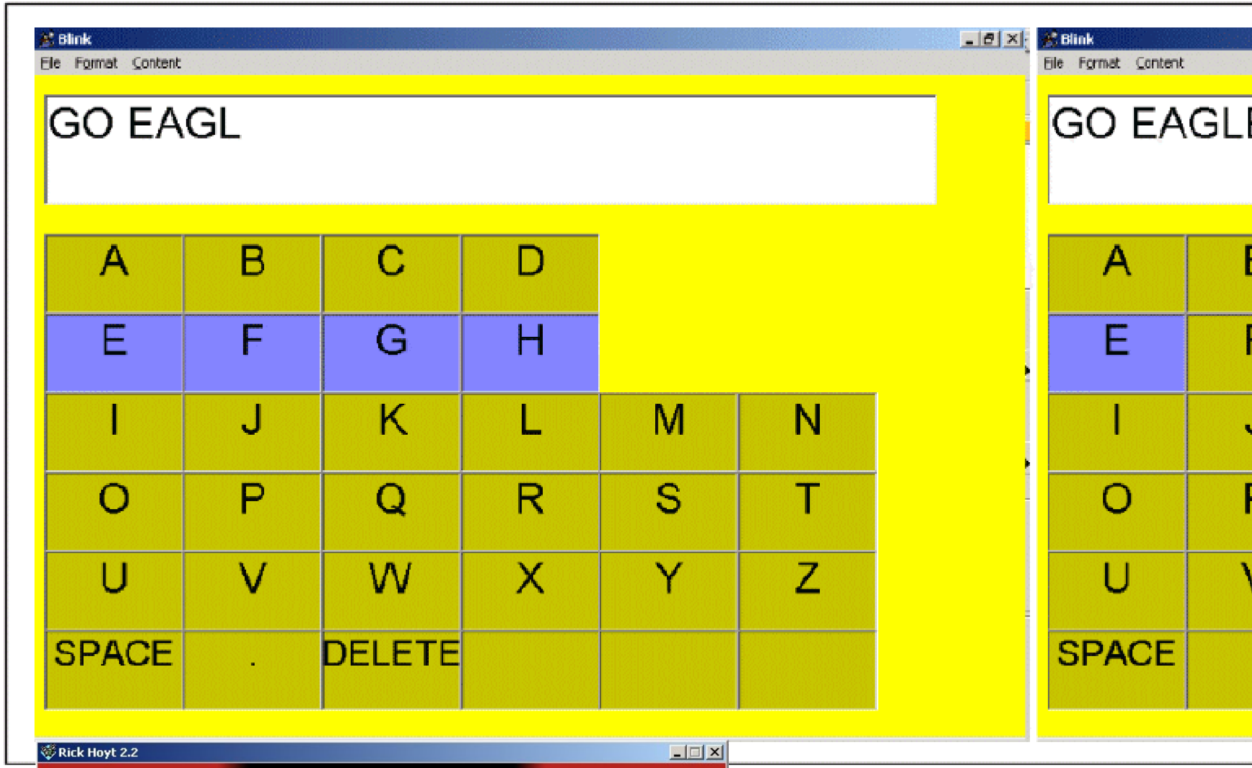
User with disabilities: "OPOTOTTR" instead of "POOR."



Text Entry

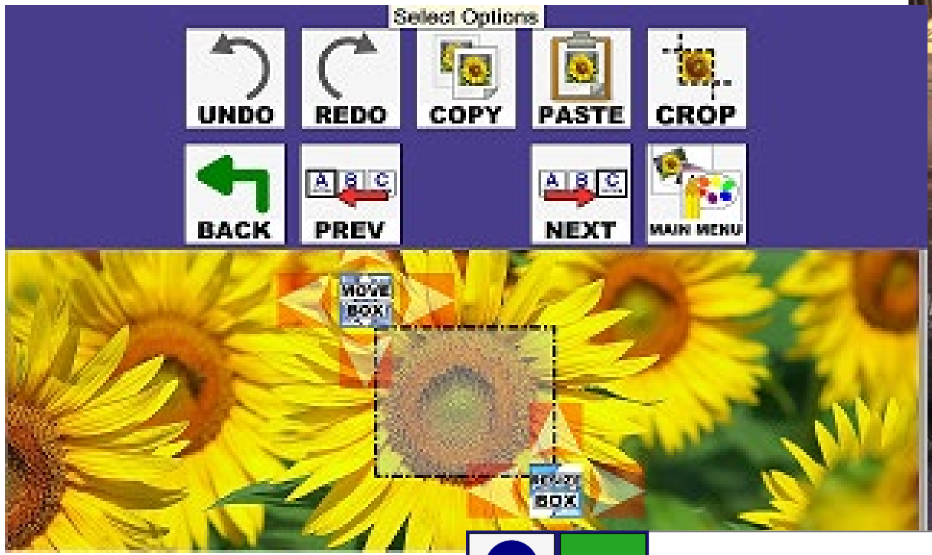
Two text entry programs used by people with severe motion impairments:

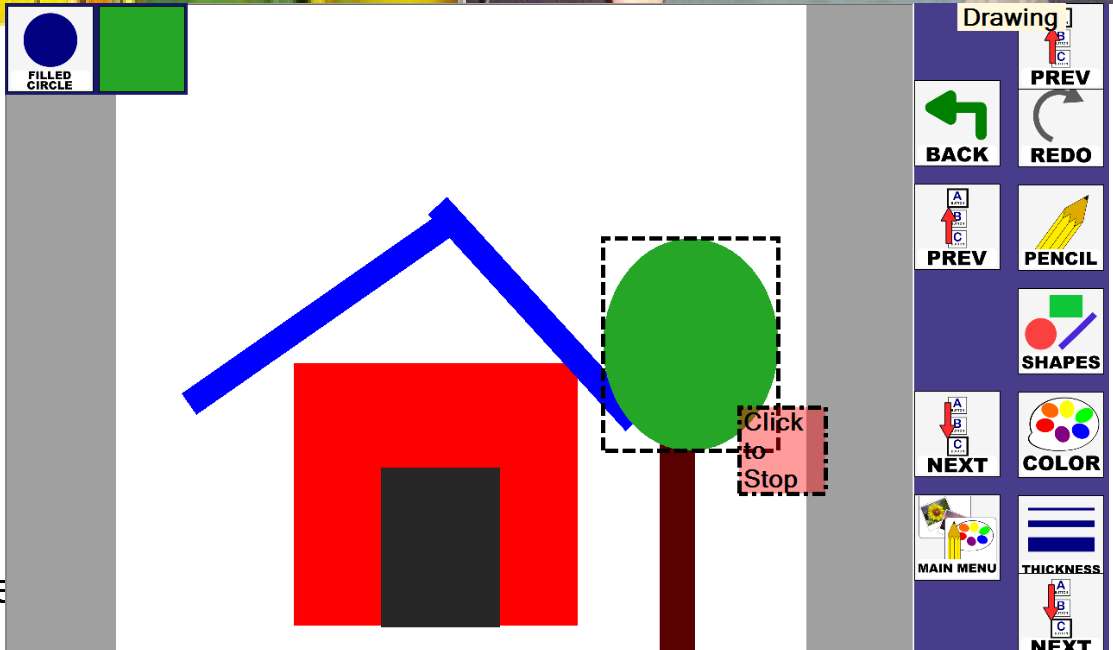
Commercial Scanning Program:



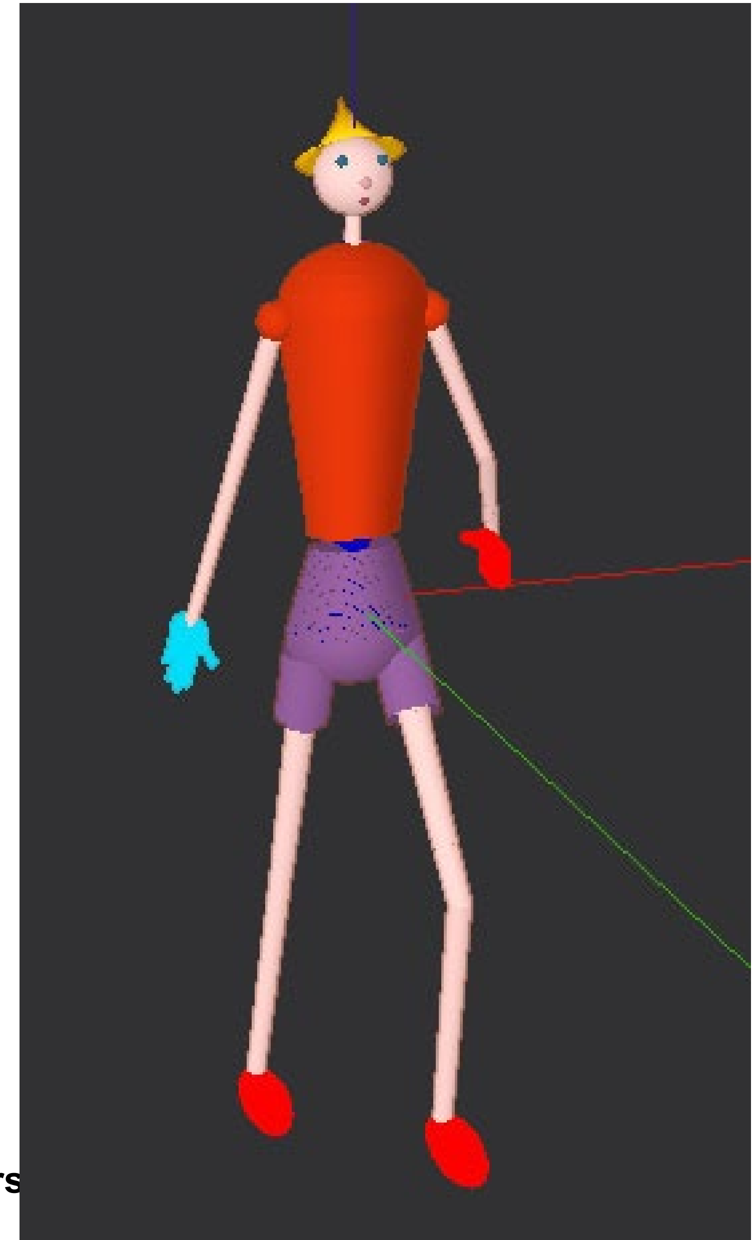
"Our" Rick Hoyt Speller:

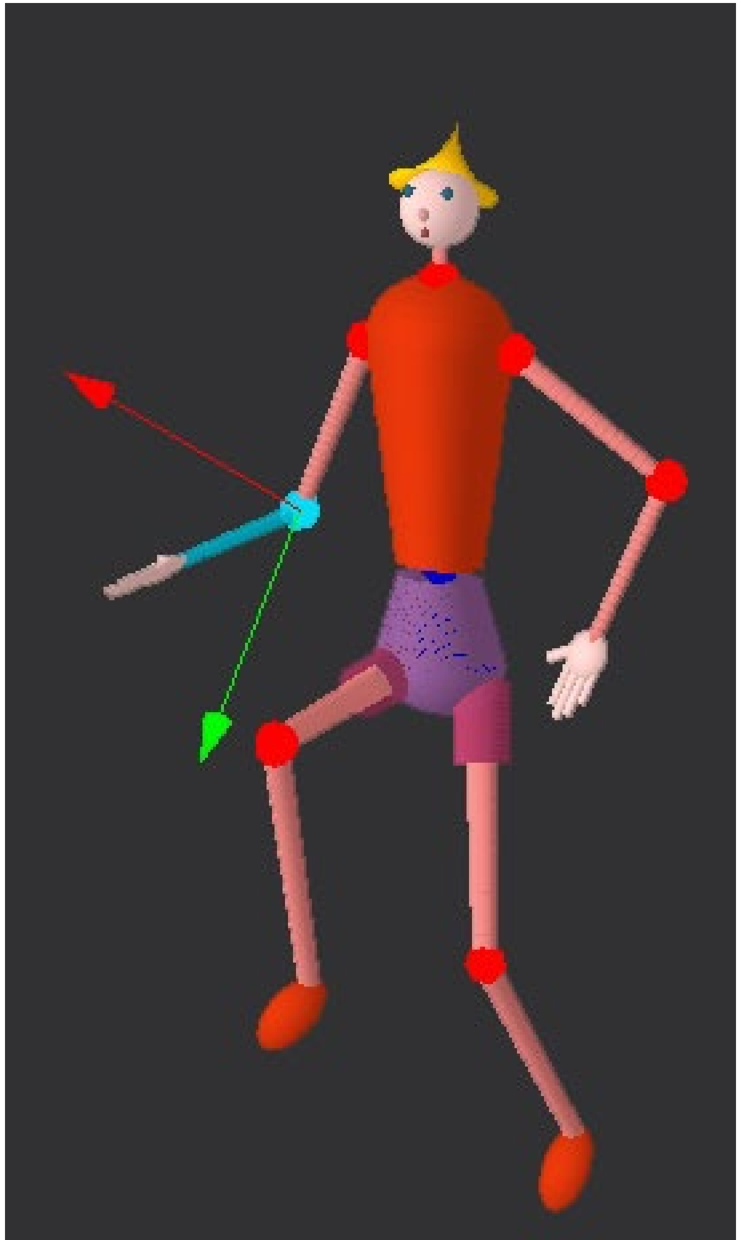
Camera Canvas



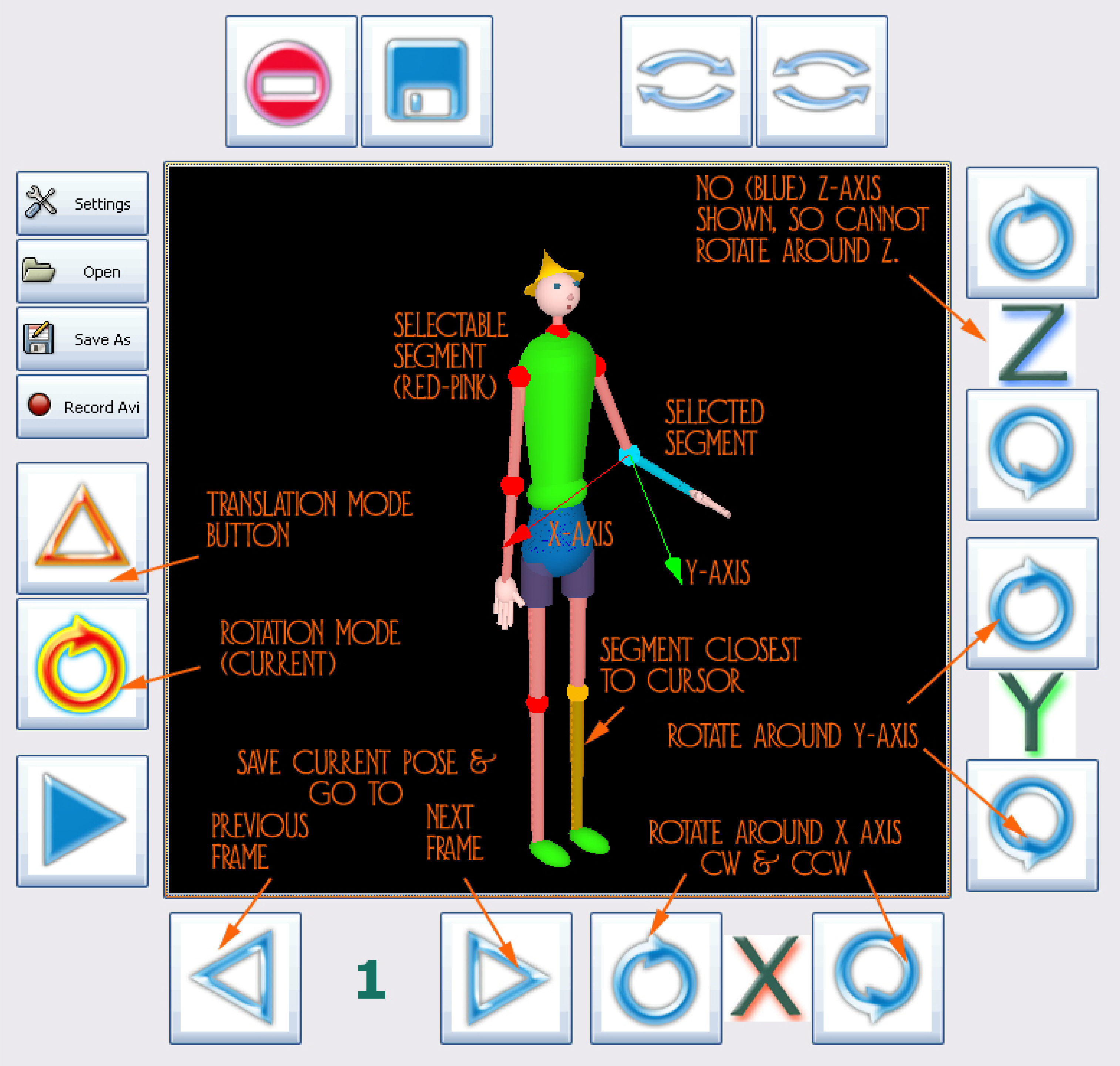


Animate!





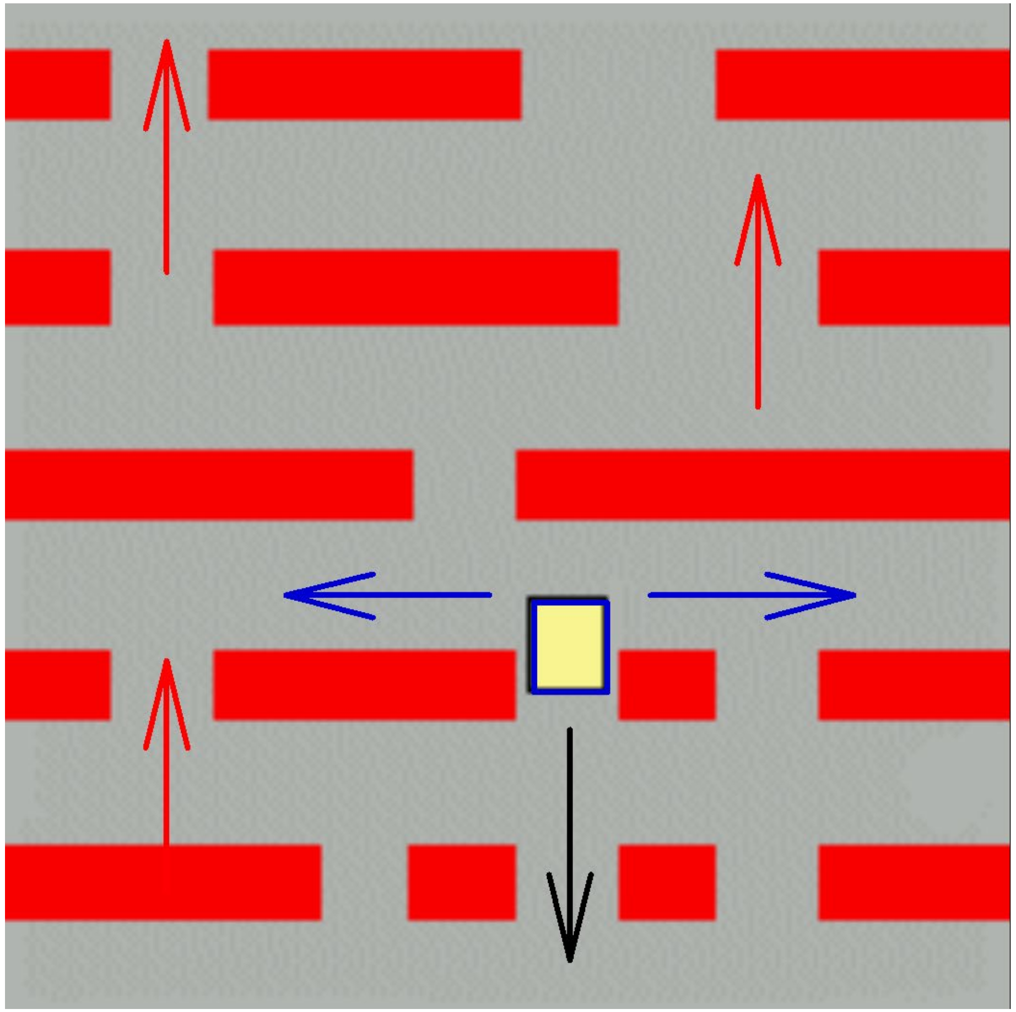
Animate!


Settings
Open
Save As
Record Avi
NO (BLUE) Z-AXIS SHOWN, SO CANNOT ROTATE AROUND Z.
Z
SELECTABLE SEGMENT (RED-PINK)
SELECTED SEGMENT
TRANSLATION MODE BUTTON
X-AXIS
Y-AXIS
ROTATION MODE (CURRENT)
SEGMENT CLOSEST TO CURSOR
Y
ROTATE AROUND Y-AXIS
SAVE CURRENT POSE & GO TO
PREVIOUS FRAME
NEXT FRAME
ROTATE AROUND X AXIS CW & CCW
1
X

Games

Memory

Block Escape



Automatic Movement

Walls move up

Block falls down

Some Lessons Learned

Designing Human-Computer Interface:

- Text entry applications need large, strategically placed areas for letters or words to reduce the problem of false selection
- “Rest” areas needed (“Midas Touch Problem”)

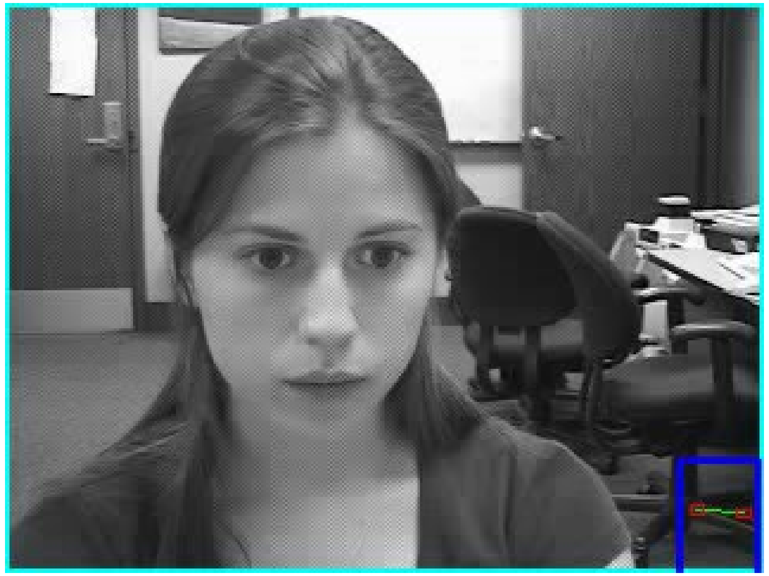
Using Human-Computer Interface:

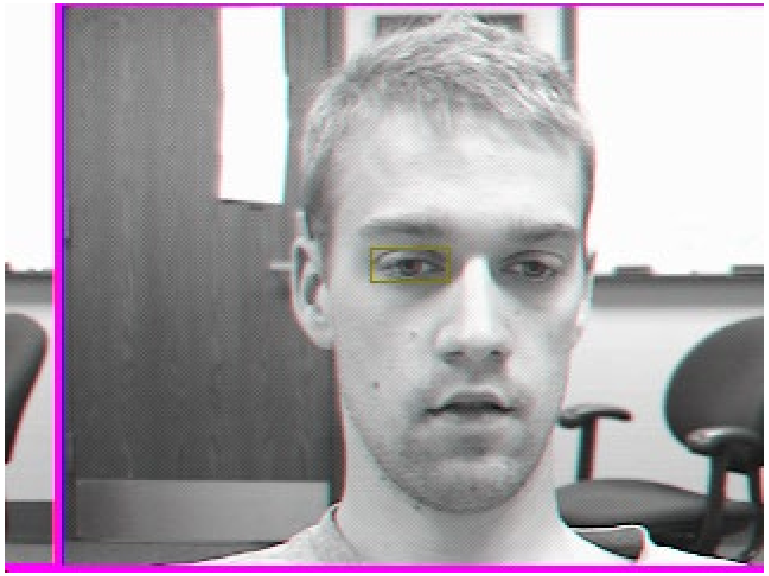
- Choice of tracking method should depend on application used
- Nose is most reliable feature for non-disabled users; allows fast and smooth motion
- No clear winning feature for users with disabilities

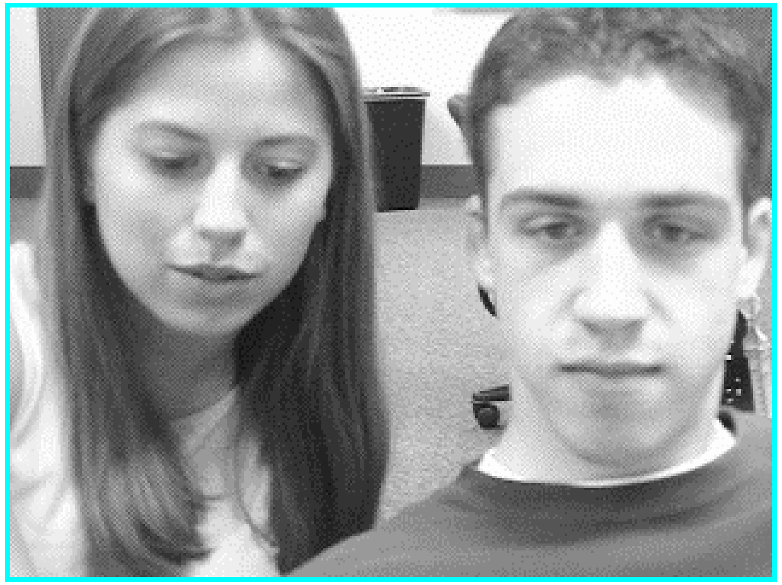
Other Input Mechanisms:

Blink Link
Eyebrow Raises
Open Mouth
Gaze Direction
Finger & Hand Motion
Facial Expressions

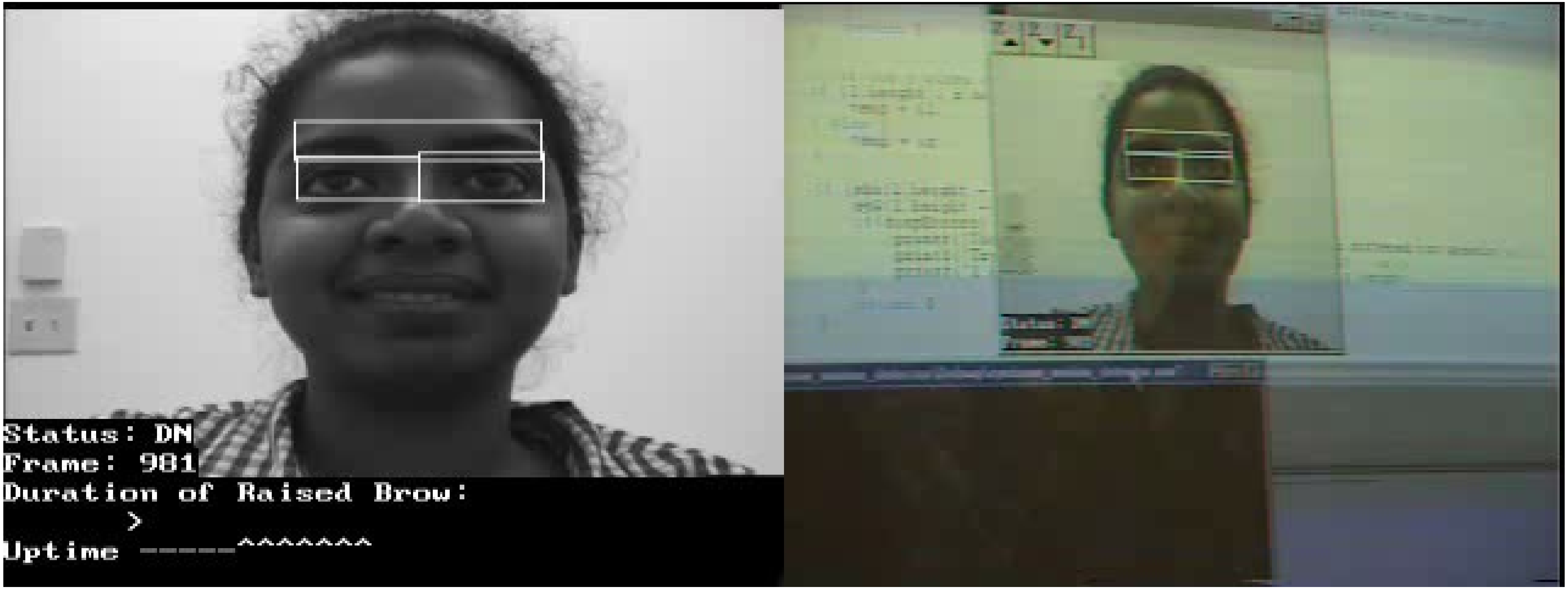
Blink Link





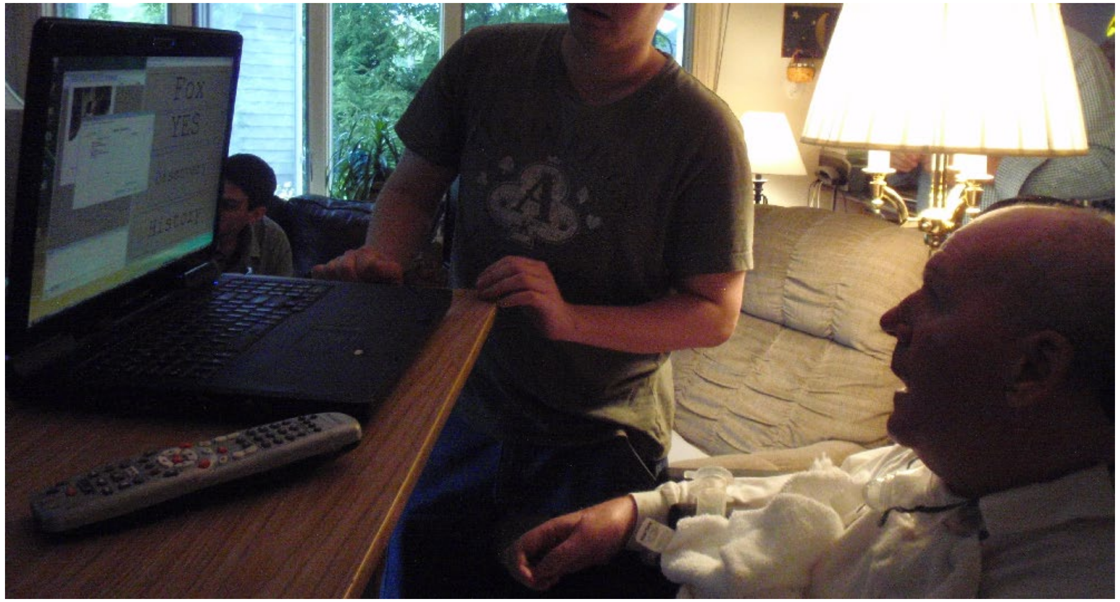


Eyebrow Clicker



User didn't like closing his eyes: We used the wide open eye input mechanism instead.

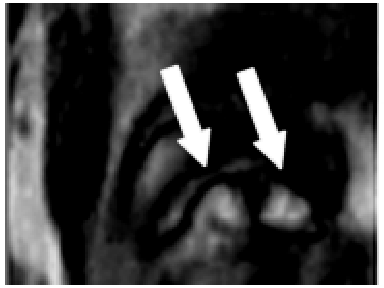
[TV.avi]



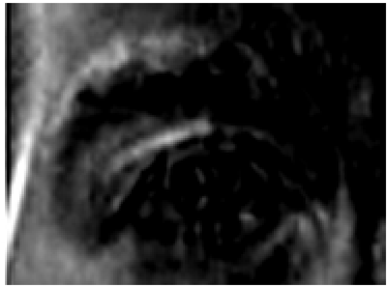
Gaze Analysis

Right Eye

Mirrored Left Eye



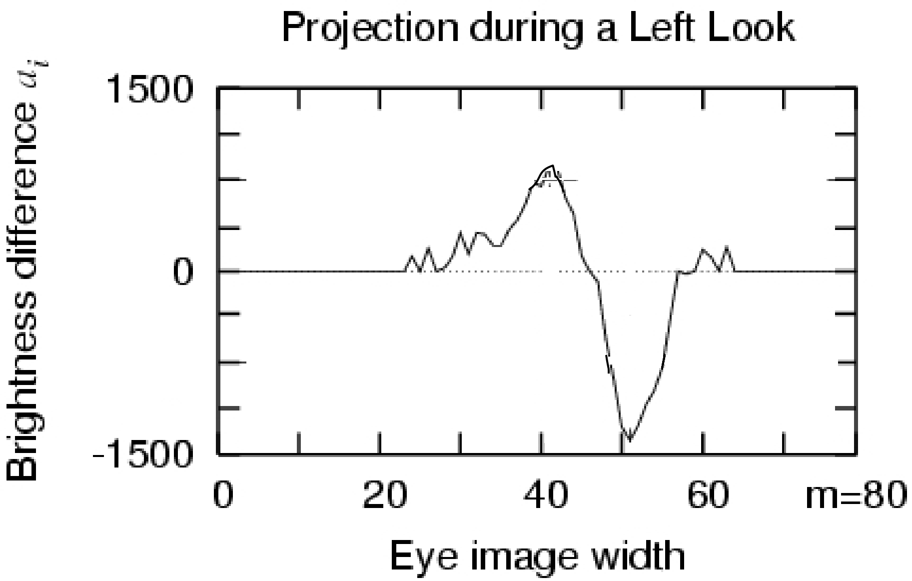
Looking Left



Looking Straight

Eye (m x n) image difference projected to x-axis:

$$a_i = \sum_{j=1}^{n} (I_r(i,j) - I_\ell(m-i,j)).$$

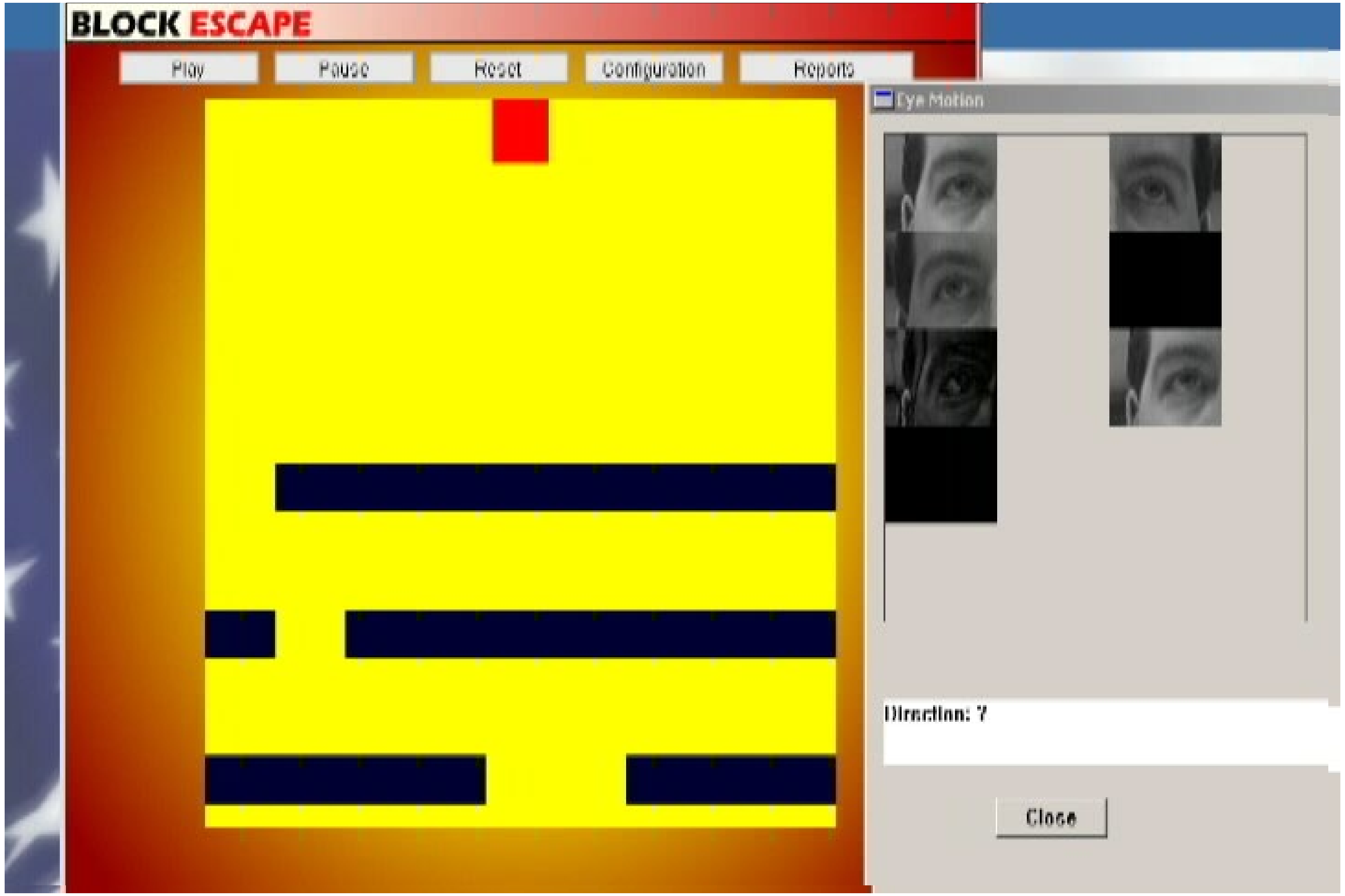


Experiment with *BlockEscape*

Methodology:

4 subjects playing a game with 3 interfaces:

- **EyeKeys**
- **Camera Mouse**
- **Keyboard**

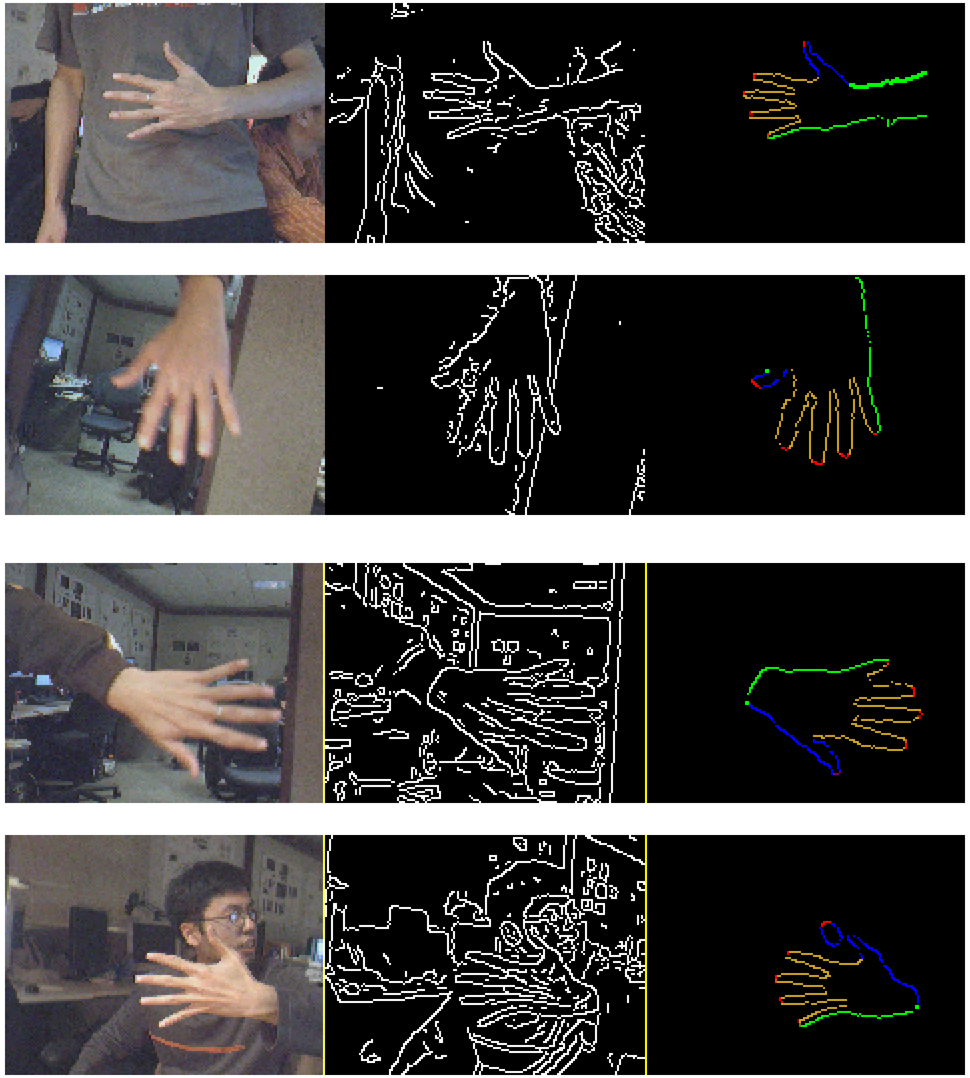


Results:

EyeKeys and Camera Mouse: 10/12 wins (83%)
Keyboard: 12/12 wins

The Finger Counter

Hand and Finger Tracking



Directed Study Idea

Design and implement a real-time "Empathic Painting" system with a modern image-generating AI model

Empathic Painting

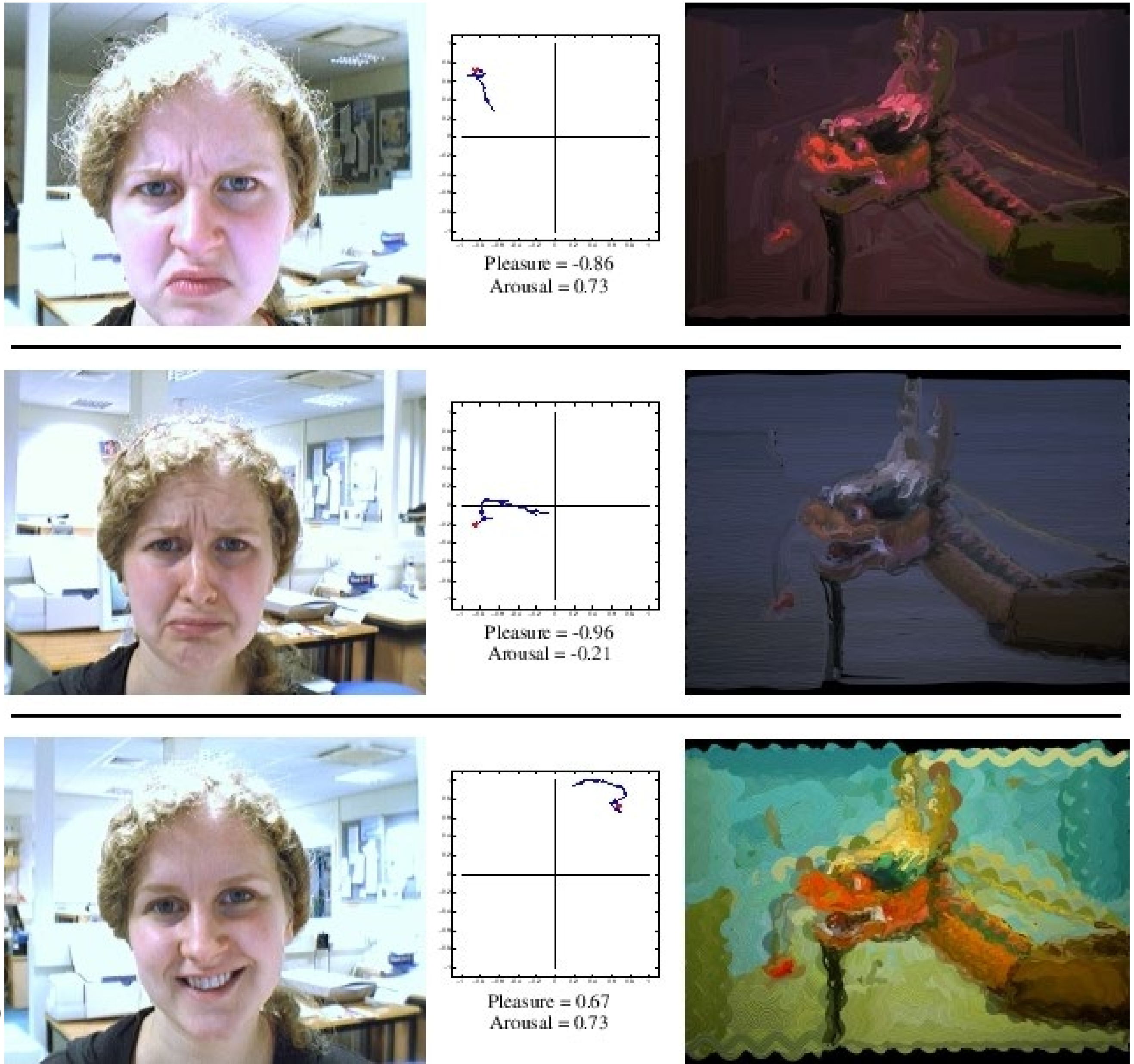
- Interactive stylization through observed emotional state
- Interactive painterly rendering whose appearance adapts in real time to reveal the perceived emotional state of the viewer

Shugrina, Betke, Collomosse, 2005

Example Input to Empathic Painting:

Photo of a dragon boat from Chinese festival on the Charles river

Empathic Painting

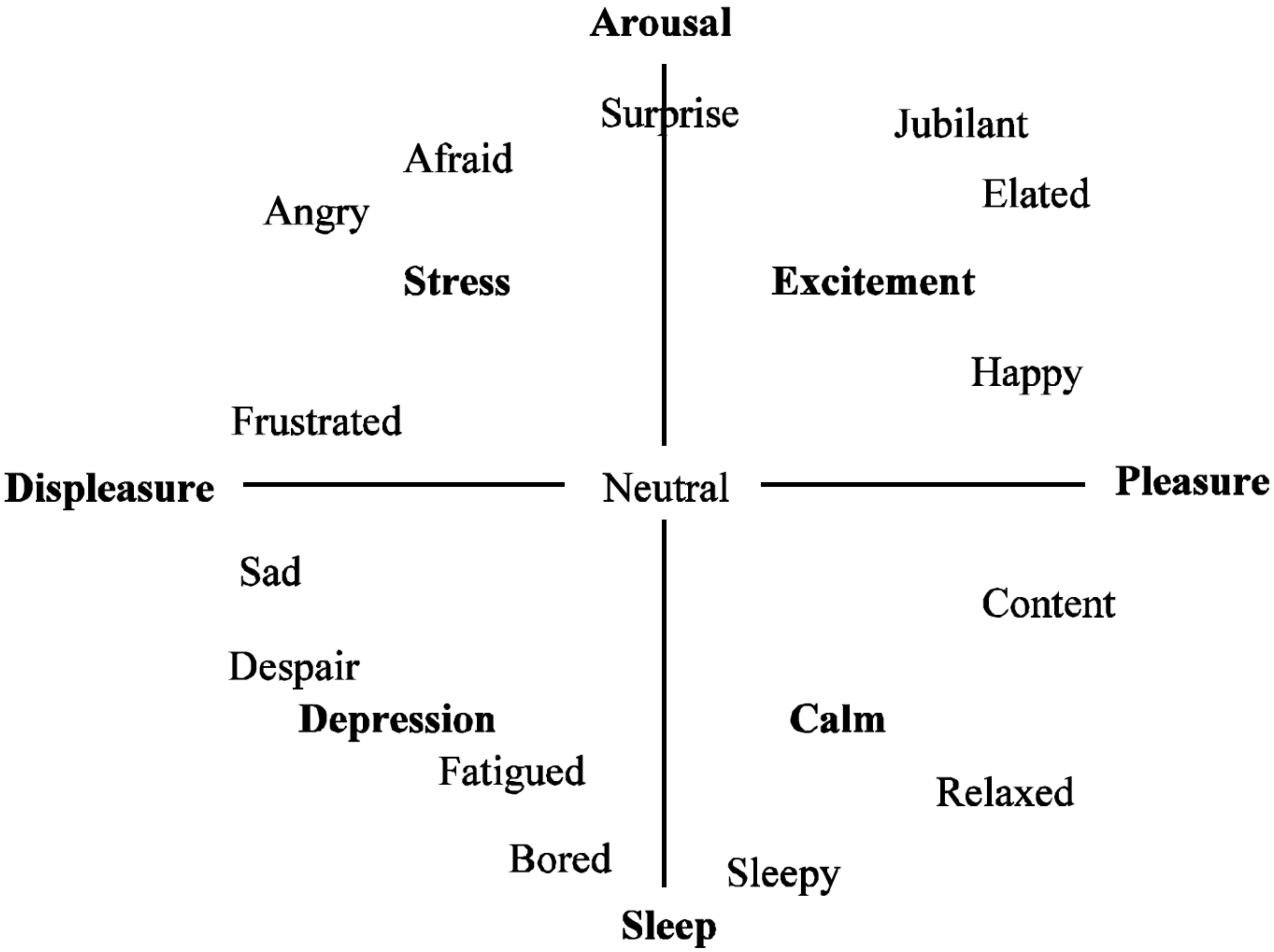


Facial Action Units

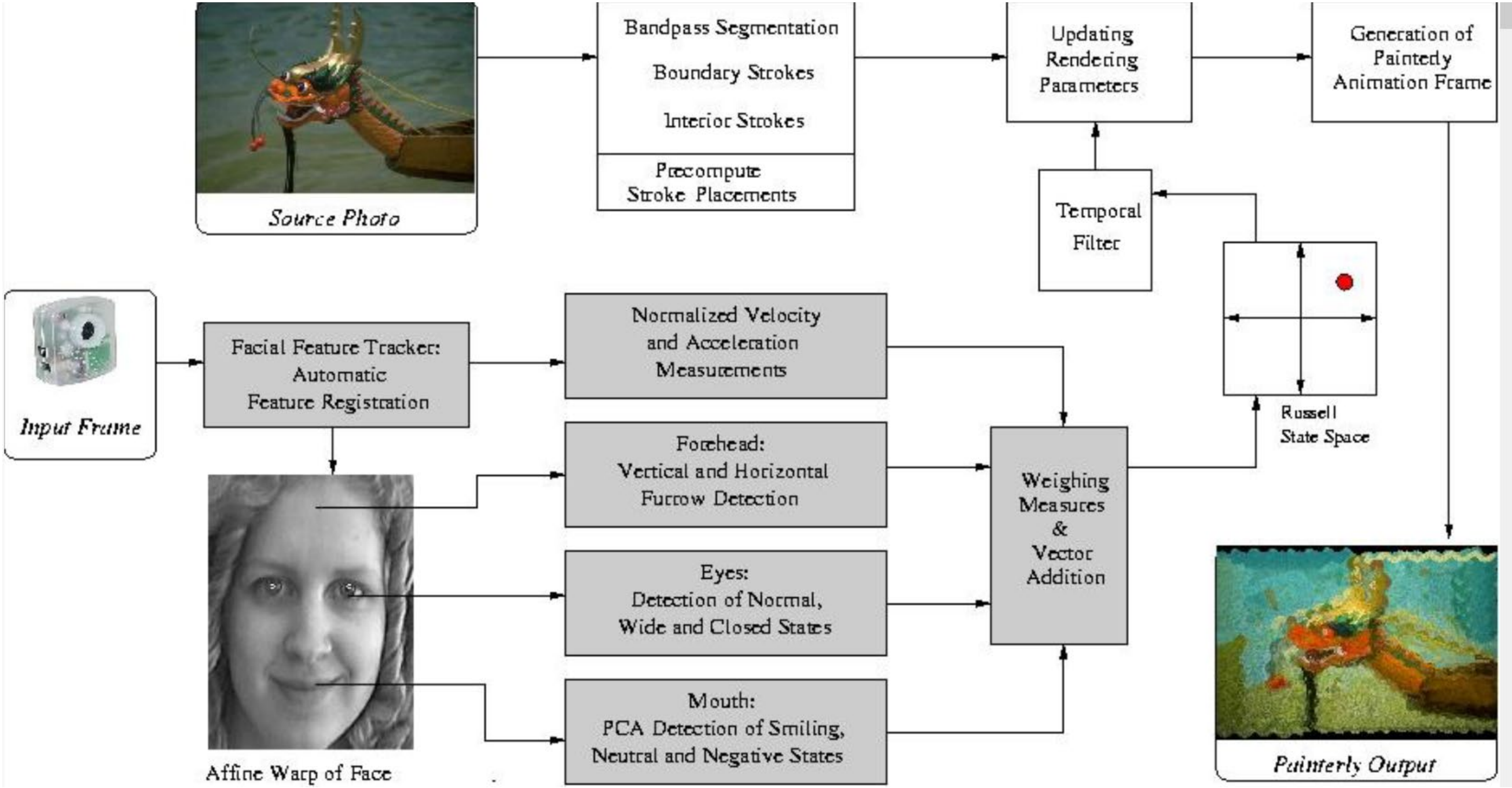
Vector	Our Action Unit	Ekman's AU	Pleasure	Arousal
1	Brow Fully Raised	1, 2	0	0.7
2	Inner Brow Raised	1	-0.5	-0.7
3	Brow Furrowed	4	-0.5	0.7
4	Negative Mouth	Various	-1.0	0
5	Positive Mouth Smile	12 25	1.0	0
6	Wide Eyes	5	0	0.4
7	Closed Eyes	41, 43, 45	0	-1.0
8	Agitation	N/A	0	0.5

See Paul Ekman

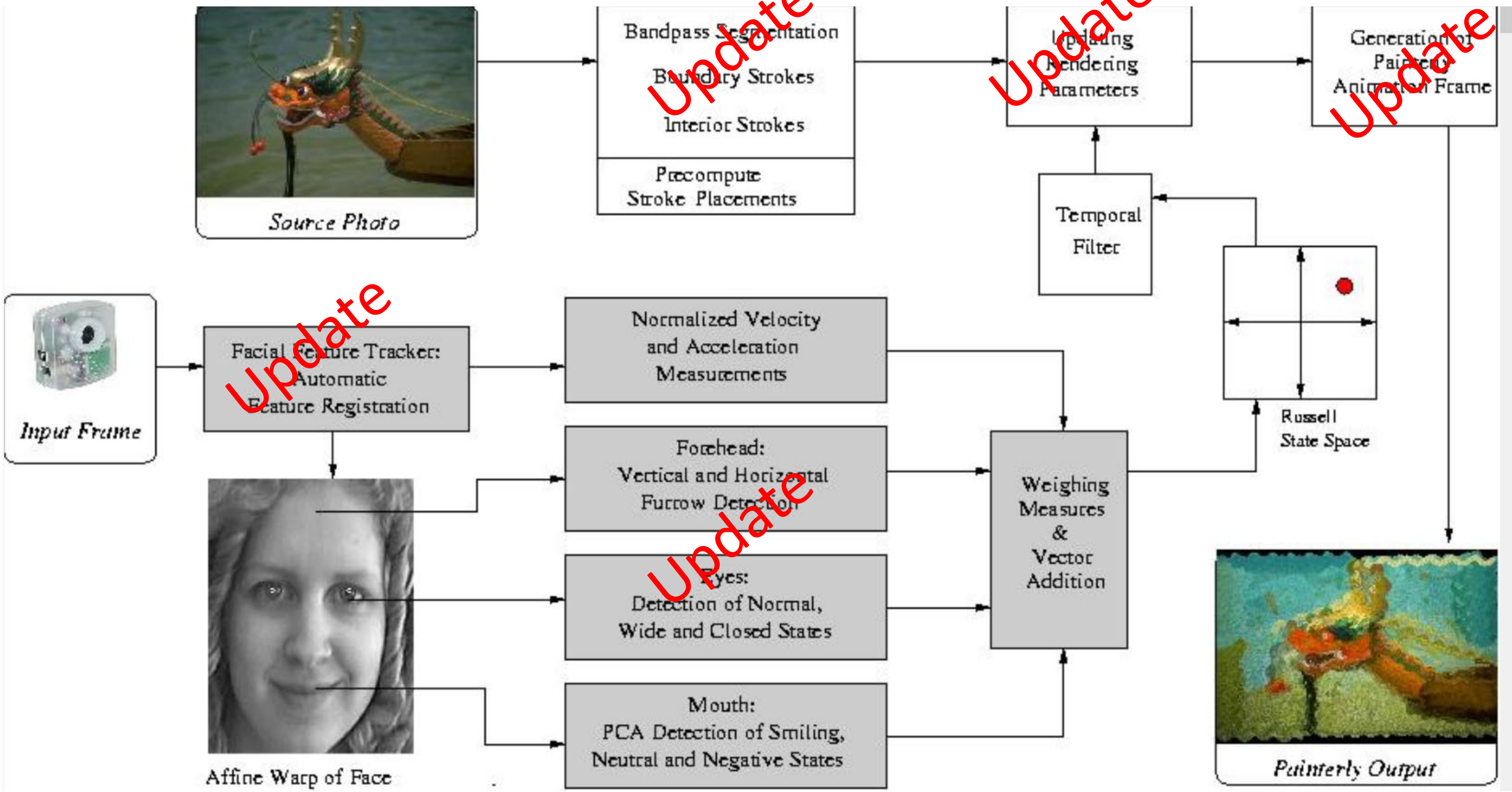
From Facial Expressions to Emotions

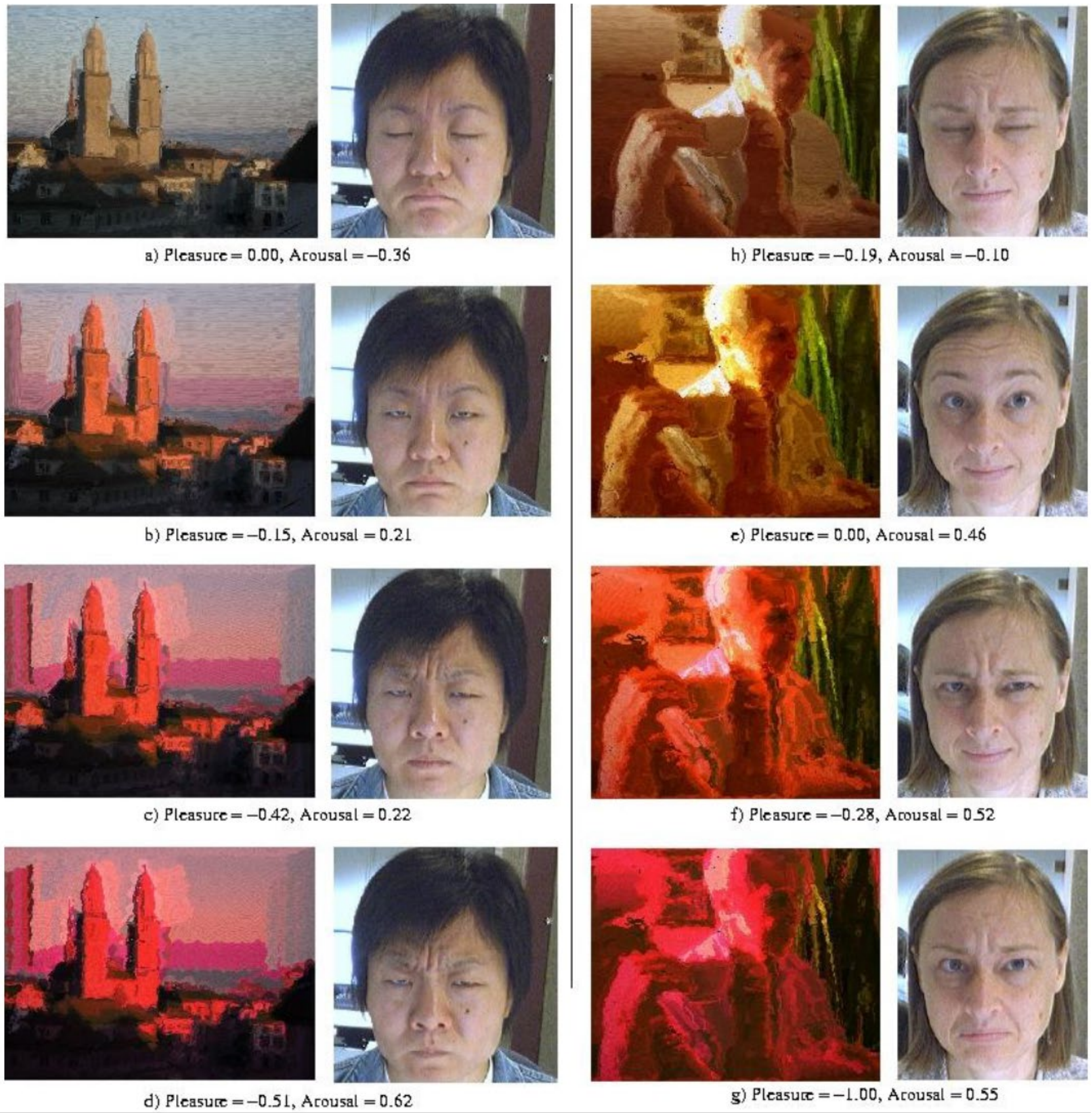


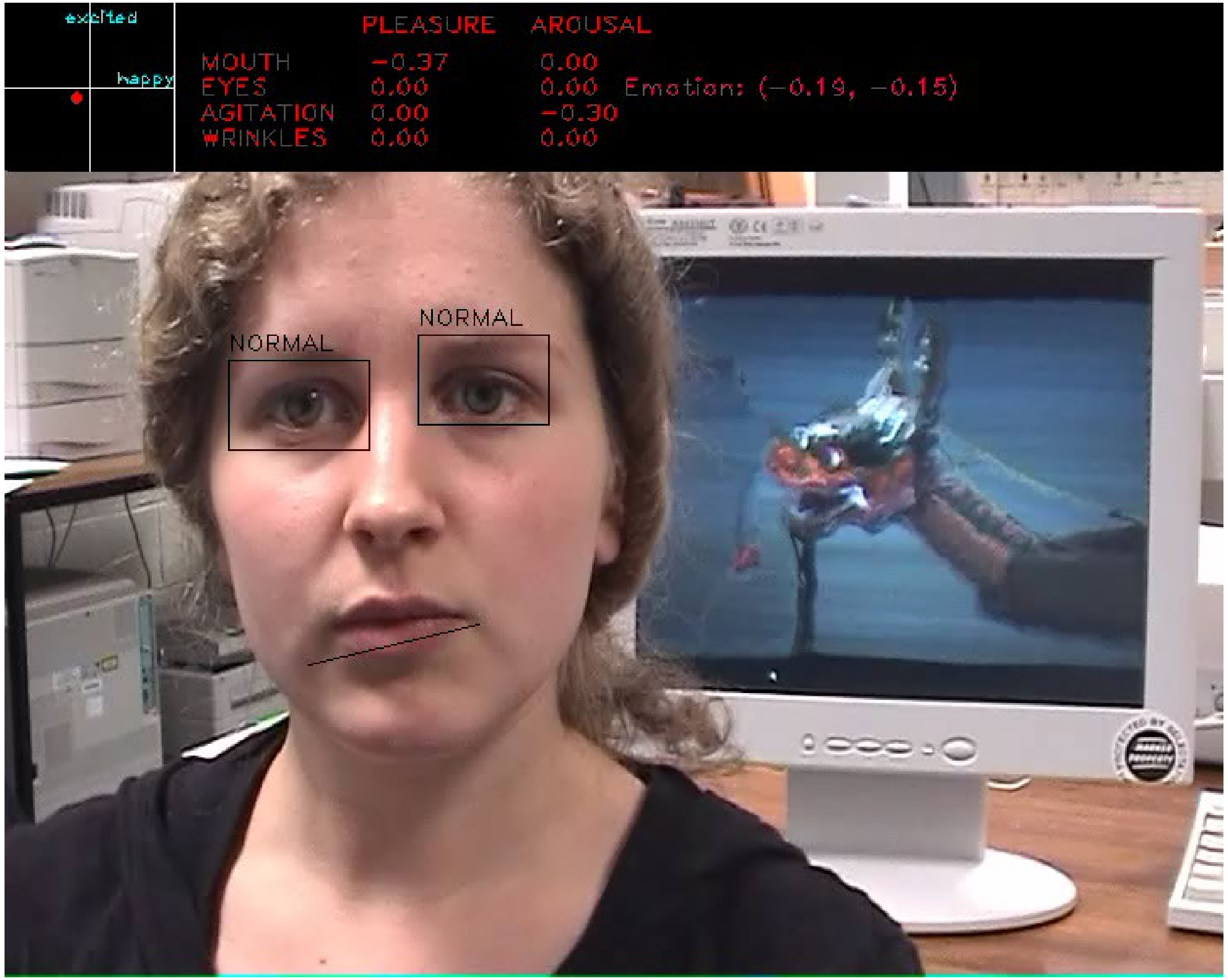
Empathic Painting System, 2005



Empathic Painting System 2024?




a) Pleasure = 0.00, Arousal = −0.36
b) Pleasure = −0.15, Arousal = 0.21
c) Pleasure = −0.42, Arousal = 0.22
d) Pleasure = −0.51, Arousal = 0.62
h) Pleasure = −0.19, Arousal = −0.10
e) Pleasure = 0.00, Arousal = 0.46
f) Pleasure = −0.28, Arousal = 0.52
g) Pleasure = −1.00, Arousal = 0.55


excited
happy
PLEASURE AROUSAL
MOUTH -0.37 0.00
EYES 0.00 0.00
AGITATION 0.00 -0.30
WRINKLES 0.00 0.00
Emotion: (-0.19, -0.15)
NORMAL
NORMAL

Acknowledgements

Funding by the National Science Foundation (IIS- 0713229, IIS-0308213, IIS-039009, IIS-0093367, P200A01031, and EIA-0202067), the Office of Naval Research (N000140110444), and The Whitaker Foundation.

More information:

www.cs.bu.edu/faculty/betke

Learning Outcomes of this Lecture

- Understand communication needs of users with severe physical disabilities
- Can explain how traditional and neural-network-based computer vision techniques can be used to detect and track facial features
- In particular, can explain correlation-based template tracking
- Understand mapping from camera coordinate system to screen coordinate system for converting facial feature movements to mouse pointer movements
- Can describe assistive software for people with motion disabilities
- Know about various input mechanisms: facial features, gaze direction, eyebrow raises, blinking, etc.
- Understand difficulties in conducting experiments with users with severe disabilities
- Can explain Fitts' law