Camera-based Interfaces for People with Severe Motion Impairments

John Magee
Margrit Betke
Boston University

Users in Need

 Millions of people with
  - Severe cerebral palsy
  - Traumatic brain injury
  - Stroke
  - Multiple sclerosis, muscular dystrophy
  - ALS (Lou Gehrig’s)

 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
Research Team

Prof. James Gips at Boston College

My students at Boston University:

Outline

- Motivation
- Camera Mouse demo
- Impact -- Camera Mouse users word-wide
- How does the Camera Mouse work?
- Assistive software
- Recent research
- Take-home message: Tasks for Future
**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

**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

- 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 Demo on Eagle Aliens

Free download at www.cameramouse.org

<table>
<thead>
<tr>
<th>Camera Mouse Users (2001)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>31</td>
</tr>
<tr>
<td>37</td>
</tr>
<tr>
<td>58</td>
</tr>
</tbody>
</table>
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
- Now 2,500 downloads per month; 35,000 in 2008

**cameramouse.org**

Traffic Statistics

Now about 2,500 downloads per month
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.

Sincerely

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 expensivetrakir 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 cursor 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" <jacqueline@optusnet.com.au>
To: <gips@bc.edu>
Subject: thank you

Hi Professor,

My name is Jacqui and I have severe Athetoid 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 developing 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.

http://www.jacquirogers.id.au

---

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

---

Tracked Features

---
**Correlation-based Tracking**

- Images showing correlation-based tracking.

**Method: Template Matching**

**Normalized Correlation Coefficient (NCC):**

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

- \( A = \text{number of pixels in template} \)
- \( \sigma_s = \sqrt{A \sum s(x, y)^2 - (\sum s(x, y))^2} \)
- \( \sigma_t = \sqrt{A \sum t(x, y)^2 - (\sum t(x, y))^2} \)

**Search area:**

- 50x50 possible new origin points
- 15x15 pixels
Correlation Surface

Tracking Methods

- **Correlation (NCC)**
- **Optical flow (Horn-Schunk)**
  \[ I(x+dx, y+dy, t+dt) = I(x, y, t) \]
- **Weighted optical flow (Lucas-Kanade)**
  \[ W(x, y) \cdot I(x, y, t) \]
- **Kalman filtering**
  
  state \( s(t|t) \): Position, Velocity, Acceleration

  Kalman filter minimizes Bayesian mean square error \( E[(s_{true}(t)-s(t|t))^2] \):

  Estimate update: \( s(t|t) = s(t|t-1) + K(t) \cdot (x(t) - s(t|t-1)) \)

  Prediction \( s(t|t-1) \) reduces search space for measurement \( x(t) \)
Summary of Tracking Results

- For slow, smooth motions, NCC with Kalman filtering worked best:
  - Tracking accuracy
  - Processing time
- For erratic movements, NCC without Kalman filtering worked better (LK tracker performed 2nd best)
- Horizontal motion easier to track than vertical motion
- Any time-consuming processing that results in skipping input frames reduces overall tracking accuracy
Evaluation Methods

1. Computer Vision:
   **Measure Tracking Accuracy:**
   True feature trajectory versus detected trajectory across different users and different tasks

2. Human-Computer Interaction:
   - Is the interface accurate enough to be operational?
   - What is the communication bandwidth?
     Fitts Law (accuracy/speed tradeoff)
   - What are the directions of movement most and least comfortable for users with disabilities?
   - Motivating test software for human subjects
Painting a Line

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

<table>
<thead>
<tr>
<th>Design</th>
<th>Trial</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Average</th>
<th>Std Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal Line</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal Line</td>
<td>Trial 1</td>
<td></td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal Line</td>
<td>Trial 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal Line</td>
<td>Trial 3</td>
<td></td>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal Line</td>
<td>Trial 4</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal Line</td>
<td>Trial 5</td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical Line</td>
<td>Trial 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical Line</td>
<td>Trial 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical Line</td>
<td>Trial 3</td>
<td></td>
<td></td>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical Line</td>
<td>Trial 4</td>
<td></td>
<td></td>
<td></td>
<td>55</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average: 19 s, Std Deviation: 15.6 s
Two-level Text Entry

Experience

Each user was asked to spell 3 words:

“RAINING”
“MINIMAL”
“POOR”

User with disabilities: “OPOTOTTR” instead of “POOR.”
Movement Evaluation Interface
Movement Evaluation Interface

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”)
- Trajectories most comfortable for a user with severe cerebral palsy were along diagonal axes

**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
- Automatic re-initialization needed when features lost
Assistive Software

- Text Entry
- Games
- Web mediators
  - GUI of browser
  - RefLink: automatic entity extraction
- Software for navigating music and videos
- Image editor “Camera Canvas”
- Animate! Making an anthropomorphic figure to dance

Rick Hoydt Speller

Group "A"?

YES

ABCD

YES

HELLO
Slime Volleyball

Animate!
Camera Canvas

- Image editing software package
- Works with Camera Mouse as mouse-substitution input system
- Sliding menu bar
- Selection area
Camera Canvas
An Interface that Enables People with Motion Impairments to Analyze Web Content and Dynamically Link to References

Navigate to the article
Copy and paste on search bar

Click on Search
Click on Result

Wikipedia Page
Search **without** using RefLink

1. Navigate to the article
2. Find a Term
3. Select Term and Copy
4. Click on Search
5. Click on the result

Search **without/with** using RefLink: 4 min/21 s

1. Navigate to the article
2. Find a Term
3. Select Term and Copy
4. Click on Search
5. Click on link of Term
**System Architecture**

- **Web cam**
- **Client Machine (with CameraMouse)**
- **Server**
- **Text-to-Speech**
- **Internet**
- **Taxonomy**

**RefLink Interface**

- **Video-Based Human-Computer Interfaces**
- **The Camera Mouse**
- **Control: Mouse**

**Head and neck**
- **Brain**
- **Eyes**
- **Nose**
- **Skin**

**Abdomen**
Dynamic Link to Wikipedia

Gesture Interpretation

Interpreting user-defined spatio-temporal patterns

Mapping patterns to arbitrary commands

Symbol Alphabet with Instances of Spatio-Temporal Patterns
System Overview

Gesture Interpretation

Preprocessor → Classifier → Interpreter

Preprocessing Spatio-Temporal Inputs
Interpreting Spatio-Temporal Patterns

Neural Net Output

Events at:
- t=12
- t=60
- t=88

Interarrival time between input events

Gesture Interpretation

Interpreting user-defined spatio-temporal patterns
Mapping patterns to arbitrary user-selected commands

5-11 commands: recognition accuracy 90%
Symbol Alphabet with Instances of Spatio-Temporal Patterns
Computer Vision Systems

- Face tracker
- *BlinkLink*
- Eyebrow clicker
- Gaze detector
- Facial feature analysis for non-photorealistic rendering

Eye Keys

Real-time video-based assistive device for people with severe disabilities

- only eye information used
- specialized hardware avoided
- works with inexpensive webcam
Interface Setup

Multilevel Image Analysis

Pyramid analysis of color, motion, and correlation with face template
**Face Tracking**

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

\[
a_i = \sum_{j=1}^{n} (I_r(i, j) - I_l(m - i, j)).
\]

**Gaze Analysis**

- **Right Eye**
- **Mirrored Left Eye**

**Looking Left**

**Looking Straight**

- **Projection during a Left Look**

- **Brightness difference:** $y = 1500$
- **Eye image width:** $m = 80$
Threshold Comparison

“Sufficient-motion threshold” $T_p$:

\[ a_{\text{min}} < -T_p \quad \text{and} \quad a_{\text{max}} > T_p. \]

“Direction-determination threshold” $T_d$:

\[ i_{\text{max}} - i_{\text{min}} > T_d \Rightarrow \text{‘right motion’} \]
\[ i_{\text{max}} - i_{\text{min}} < -T_d \Rightarrow \text{‘left motion’}. \]

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
Other Applications

- Web Browsing
- Text entry
- Other left-right games

Possible function mappings:

<table>
<thead>
<tr>
<th>Application</th>
<th>Left Look</th>
<th>Left Function</th>
<th>Right Look</th>
<th>Right Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>BlockEscape</td>
<td>Left arrow</td>
<td>Move block left</td>
<td>Right arrow</td>
<td>Move block right</td>
</tr>
<tr>
<td>Driving Game</td>
<td>Left arrow</td>
<td>Drive left</td>
<td>Right arrow</td>
<td>Drive right</td>
</tr>
<tr>
<td>Web browser</td>
<td>Tab</td>
<td>Go to next link</td>
<td>Enter</td>
<td>Follow current link</td>
</tr>
<tr>
<td>Rick Hoydt Speller</td>
<td>Click</td>
<td>Select</td>
<td>Click</td>
<td>Select</td>
</tr>
<tr>
<td>Other text entry</td>
<td>Enter</td>
<td>Select letter</td>
<td>Right arrow</td>
<td>Next letter</td>
</tr>
</tbody>
</table>

Text Entry

Two text entry programs used by people with severe disabilities:

Commercial Scanning Program:

Rick Hoydt Speller:
Scanning Games

Testing Experiences
Testing Experiences

Face Tracker Limitations:
- Head tilts and turns

Eye Analysis Limitations:
- Center of eyes

Using “Active Vision”
Using "Active Vision"

**Perfect Iris Detection**  **Perfect Iris Detection with Imperfect Face Detection**

**Detection through Glasses**  **False Iris Detection**

---

**Initial Zoom Sequence on Moving Subject**
Blink Link

Real-time eye blink detection and classification of blinks as voluntary or involuntary

- Long (voluntary) blinks generate mouse clicks
- Use with scanning software
- Requires no manual initialization, no special lighting
Blink & Wink Detection

Summary of Results

<table>
<thead>
<tr>
<th>Category</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall detection accuracy</td>
<td>95.6%</td>
</tr>
<tr>
<td>Long/Short classification accuracy</td>
<td>93.0%</td>
</tr>
<tr>
<td>Usability score as an input device</td>
<td>93.6%</td>
</tr>
<tr>
<td>Average frame rate</td>
<td>28 fps</td>
</tr>
</tbody>
</table>

Eric Missimer, 2010
Blink & Wink Detection

- For users who can wink with one eye while keeping their other eye visibly open:
  
  System allows complete use of a typical mouse, including left and right clicking, double clicking and click and drag

- For users who cannot wink but can blink voluntarily:
  
  System allows user to perform left clicks

- Online training in initialization phase to obtain open- and closed-eye templates:
  
  System does not require any training data to distinguish open eyes versus closed eyes

- Open/closed eye classification accomplished online during real-time interactions during the normalized correlation coefficient.

Feedback window used to indicate to user or caregiver the detected status of the eyes, here a closed left eye and an open right eye, which is interpreted as a command to click the left mouse button. The feedback window is positioned above the mouse pointer and follows the mouse pointer throughout the tracking. The window is semi-transparent to allow the user to see the interface below.
# Blink Detection

Results of classification tests: 8 test subjects

<table>
<thead>
<tr>
<th>Image Type</th>
<th>Number of Images</th>
<th>Successful Classifications</th>
<th>Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Eye Images</td>
<td>4,987</td>
<td>4,796</td>
<td>96.2%</td>
</tr>
<tr>
<td>Closed Eye Images</td>
<td>3,319</td>
<td>3,231</td>
<td>97.4%</td>
</tr>
<tr>
<td>All Images</td>
<td>8,306</td>
<td>8,027</td>
<td>96.6%</td>
</tr>
</tbody>
</table>

# Blink & Wink Detection

<table>
<thead>
<tr>
<th>Click Type</th>
<th>Wrong Click</th>
<th>Click outside of area</th>
<th>Failed to click in time</th>
<th>Successful Clicks</th>
<th>Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double Click</td>
<td>18</td>
<td>18</td>
<td>0</td>
<td>201</td>
<td>84.81%</td>
</tr>
<tr>
<td>Left and Right Click</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>63</td>
<td>81.82%</td>
</tr>
<tr>
<td>Dragging</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>80</td>
<td>95.24%</td>
</tr>
<tr>
<td>Blinking Left Click</td>
<td>N/A</td>
<td>18</td>
<td>0</td>
<td>201</td>
<td>90.95%</td>
</tr>
</tbody>
</table>
Eyebrow Clicker

Eyebrow template

Eye templates (left and right)

Relative Motion of Brows

At each frame, use adaptive template matching to locate eyes and eyebrows.

Measure distance from eyes to eyebrows.

If $\text{dist}_{\text{new}} > 1.25 \text{dist}_{\text{relaxed}}$ classify brows as raised.

If raised > 0.5 s, “Click”
Experiments

American Sign Language Recognition
The Finger Counter

Hand and Finger Tracking
3D Feature Tracking

Empathic Painting
Empathic Painting

Interactive stylization through observed emotional state

Shugrina, Betke, Collomosse, 2005

Interactive painterly rendering whose appearance adapts in real time to reveal the perceived emotional state of the viewer.
Take Home Messages

Research about assistive environments addresses lots of interesting computer vision, HCI, data mining, machine learning tasks

Develop assistive software that addresses needs of real people adapts to users

- immediately -- across-users learning
- through time – degenerative diseases

Acknowledgements

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

More information and PDFs of papers: www.cs.bu.edu/faculty/betke

Publications: various IEEE Transactions, UAIS, CVPR, ECCV, UA4All, HCII, etc.