Analyzing the Potential of Adapting Head-Mounted Eye Tracker Calibration to a New User

by Benedict Fehringer

Master-Seminar

Supervisor: Prof. Dr. Antonio Krüger
Advisor: Dr. Andreas Bulling (University of Cambridge, UK)

19.01.2012
Outline

- Introduction and Motivation
- Related Work
- Study Question
- Realization and Experimental Design
- First Results
- Conclusion
Introduction and Motivation
Some Basics

- Distinguish between two kinds of Eye Trackers:
  - static
  - mobile (head-mounted)

- **Eye tracker calibration** refers to finding the parameters of the eye model used to determine the point of gaze (POG), mostly the offset between visual and optical axis.
The „normal“ Calibration Procedure

→ „normal“ calibration is time-consuming
Assessment of time-consuming calibration

Burden of time-consuming calibration

- laboratory studies
- field studies / interaction scenarios

e – The aim is to reduce the calibration effort without losing (too much) accuracy
Related Work
Calibration-free system by reconstruction pupil ellipse

- Using stereo camera system for mobile eye tracking
- Reconstructing the original pupil ellipse from both projections
- The pupil surface’s normal vector is interpreted as visual axis
- Evaluation with 1 subject; average error of $\sim 1.2^\circ$

(Source: Kohlbecher et al.; 2008)
Automatic calibration procedure for static eye tracking systems

- Using stereo cameras for static eye tracking
- Recording both eyes
- Deriving the POG as the intersection of both visual axis
- Evaluating the system by simulation; average error nearly to $0^\circ$

(Source: Model, Guestrin & Eizenman, 2009)
Experiment results with 4 real subjects

<table>
<thead>
<tr>
<th>Subjects:</th>
<th>RMS error, (°)</th>
<th>UCF-REGT</th>
<th>1-point calib REGT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left Eye</td>
<td>Right Eye</td>
<td>Left Eye</td>
</tr>
<tr>
<td>1</td>
<td>1.2</td>
<td>0.8</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>1.4</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>4</td>
<td>1.6</td>
<td>1.7</td>
<td>0.9</td>
</tr>
</tbody>
</table>

‘UCF-REGT’ – PoG is estimated with the user-calibration-free REGT. ‘1-point calib REGT’ – PoG is estimated with the REGT system that uses one-point user-calibration procedure.

Average error of 1.3°

(Source: Model & Eizenman, 2010)
Estimating the horizontal angle between visual and optical axis

* Using stereo cameras for static eye tracking

* Assuming the POG between the both intersections between the display and the optical axis

* Ignores vertical angle between visual and optical axis

* Evaluation study with 20 subjects; average error 1.6°

(Source: Nagamatsu et al., 2010)
## Summary

<table>
<thead>
<tr>
<th></th>
<th>monocular</th>
<th>estimating visual axis</th>
<th>max error [°]</th>
<th>min error [°]</th>
<th>mean error [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupil Reconstruction</td>
<td>yes</td>
<td>no</td>
<td>2.2</td>
<td>0.4</td>
<td>~1.2</td>
</tr>
<tr>
<td>POG as intersection of both visual axis</td>
<td>no</td>
<td>yes</td>
<td>1.7</td>
<td>0.8</td>
<td>1.3</td>
</tr>
<tr>
<td>POG as mean between display and optical axis</td>
<td>no</td>
<td>yes (and no)</td>
<td>3.2</td>
<td>0.8</td>
<td>1.6</td>
</tr>
</tbody>
</table>
Problems of usage of the presented systems

- Measuring optical axis instead of visual axis
- Binocular assumption: Both eyes look at the same place

... and in general:

- Differences between models and reality
- Inaccuracy during measuring
Study Question
Basic Idea

- Trying to estimate the true POG (visual axis) with a monocular eye tracker

- Using the calibration from another person (foreign calibration) to estimate the POG

- Adapting the estimated („foreign“) POG with simple translation
Study’s Goal: Evaluation of the adapted foreign calibration

- Comparison between the *own error* and the *foreign error* (after adaption)

- *own error* refers to the error between the true POG and the estimated POG using the own calibration

- *foreign error* refers to the error between the true POG and the adapted estimated POG using the foreign calibration
Realization and
Experimental Design
Preliminary thoughts about the Study Design

- To know where the true POG is the participants have to look at predefined places.

- There is probably a difference if (only) the head or (only) the eyes are moving:
  
  - *Only Head*: the eye’s position is constant relative to the glasses → the correction based on the error would lead to perfect matching (beside of noise / statistical error)

  - *Only Eyes*: the eye’s position is **NOT** constant relative to the glasses → the error between true and estimated POD could change depending of the view angle (depending of the algorithm)
Difference between Only Head and Only Eyes (Example)

uncorrected error for:
- only eyes
- only head
Setup

- 3 conditions (Only Head, Only Eyes, Natural)
- the participants have to stand centered in front of the projection plane with a distance of 2 m
- in each condition the participants have to look at 25 targets (black points, indicated by the black crosses)
Apparatus
- Eye Tracker -

**Tobii Eye Tracker Glasses**

- mobile video-based eye tracker
- recording monocular gaze data from the right eye
- sampling rate 30Hz
- resolution 640x480 pixel
- recording angles 56° horizontal and 40° vertical

**Infrared-Marker (IR-Marker)**

- send unique Id that can detected by the Glasses (Where and Which)
**Apparatus**

- **Motion Tracker** -

**OptiTrack Arena System**

- 6 infrared cameras
- tracks objects in 3D by using reflecting markers
- sampling rate 100Hz
- delivers 3 rotation angles corresponding to the 3 Cartesian axis
- accuracy $\leq 0.1^\circ$

**2 intended purposes**

- indicating the head movement
- alternative for automatic analyzing
Adapting the foreign calibration

Using a translation with constant c:

* multi-point („optimal“ improvement):
  * c is equal to the mean deviation of all measured points

* 1-point:
  * c is equal to the deviation of the central point

* 0-point:
  * c is computed by the pupil position relative to the glasses
Analyzing Tool

- uses the Tobii Glasses SDK

supports features to:

- display estimated POG and IR-Markers
- move at certain video-frames (Play, Pause, revers, slider, speed changing, Time)
- automatically compute the distance between true and estimated POG (mark true POG, choose algorithm)
- exclude false recognized IR-Marker
- merge the created log-file with the OptiTrack’s log-file by using time stamps as synchronization link
- inform the user
First Results
Demographic Data

Test persons:

- \( N = 5 \) (2 female, 3 male)
- Age [years]: Range = 15-33; mean = 24.5; sd = 5.79
- 1 with contact lenses; 1 wears normally glasses (but not in the study)
- Tobii calibration quality scale (TCS) from 0 (worst) to 5 (best): Range = 2 - 5

1 person for foreign calibration

- male, 30 years, no glasses or lenses, 5 on the TCS
Some descriptive Data to differentiate the conditions

eye movement

head movement
Absolute deviations between the estimated and the true POG

<table>
<thead>
<tr>
<th>Participant</th>
<th>Method</th>
<th>X [°]</th>
<th>Y [°]</th>
<th>X [°]</th>
<th>Y [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>uncor</td>
<td>1.32</td>
<td>7.81</td>
<td>0.74</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>1.31</td>
<td>2.77</td>
<td>0.76</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>2.80</td>
<td></td>
<td>1.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>3.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>uncor</td>
<td>1.49</td>
<td>6.96</td>
<td>1.07</td>
<td>8.80</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>1.49</td>
<td>2.15</td>
<td>1.01</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>3.36</td>
<td></td>
<td>1.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>2.27</td>
<td></td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>uncor</td>
<td>5.32</td>
<td>9.81</td>
<td>3.28</td>
<td>10.85</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>5.28</td>
<td>6.42</td>
<td>3.30</td>
<td>7.39</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>8.77</td>
<td></td>
<td>7.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>7.79</td>
<td></td>
<td>8.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>0.93</td>
<td>2.43</td>
<td>0.50</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>2.57</td>
<td></td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>2.40</td>
<td></td>
<td>2.25</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>uncor</td>
<td>0.79</td>
<td>9.47</td>
<td>0.85</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>0.71</td>
<td>1.53</td>
<td>0.34</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>1.57</td>
<td></td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>2.07</td>
<td></td>
<td>4.76</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>uncor</td>
<td>0.74</td>
<td>5.45</td>
<td>0.52</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>0.74</td>
<td>2.72</td>
<td>0.47</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>2.80</td>
<td></td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>2.40</td>
<td></td>
<td>2.25</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant</th>
<th>Method</th>
<th>X [°]</th>
<th>Y [°]</th>
<th>X [°]</th>
<th>Y [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>uncor</td>
<td>1.88</td>
<td>9.55</td>
<td>0.50</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>0.43</td>
<td>0.65</td>
<td>0.43</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>uncor</td>
<td>1.47</td>
<td>10.73</td>
<td>1.36</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>0.71</td>
<td>0.62</td>
<td>0.21</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>0.62</td>
<td></td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0-point</td>
<td>4.27</td>
<td></td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>uncor</td>
<td>0.56</td>
<td>6.39</td>
<td>0.64</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td>multi-point</td>
<td>0.39</td>
<td>1.15</td>
<td>0.31</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>1-point</td>
<td>1.20</td>
<td></td>
<td>0.77</td>
<td></td>
</tr>
</tbody>
</table>

|                |                |        |        |        |        |
|                |                |        |        |        |        |
|                |                |        |        |        |        |
| Only eyes      | Only head      |        |        |        |        |
| mean           |                | 2.02   | 7.68   | 1.31   | 5.15   |
|               |                | 1.95   | 3.30   | 1.21   | 2.40   |
|               |                | 4.06   | 1.21   | 2.62   |       |
| mean           |                | 3.76   | 4.51   |        |        |

## Results from the other studies

<table>
<thead>
<tr>
<th></th>
<th>monocular</th>
<th>estimating visual axis</th>
<th>max error [°]</th>
<th>min error [°]</th>
<th>mean error [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pupil Reconstruction</strong></td>
<td>yes</td>
<td>no</td>
<td>2.2</td>
<td>0.4</td>
<td>~1.2</td>
</tr>
<tr>
<td><strong>POG as intersection of both visual axis</strong></td>
<td>no</td>
<td>yes</td>
<td>1.7</td>
<td>0.8</td>
<td>1.3</td>
</tr>
<tr>
<td><strong>POG as mean between display-optical axis</strong></td>
<td>no</td>
<td>yes (and no)</td>
<td>3.2</td>
<td>0.8</td>
<td>1.6</td>
</tr>
</tbody>
</table>
Limitations

* measuring error to determine the true POG‘s position because of false detected IR-markers

* rather small sample, although for each participant a large data set (up to 25 points * 150 frames/point = 3750 frames) was analyzed

* no cross-validation in 0-point

➔ the last two limits can be (partly) handled by analyzing the remaining subjects
Conclusion
Until now...

- Demonstrating of the feasibility of using a foreign calibration from one user to adapt it to a new user for monocular eye trackers

- Good results could be achieved by using only simple algorithms and one parameter

  ➔ with bigger samples and more various parameters further improvements should be possible

- Depending of the purpose a 0-point or 1-point adaption seems to be sufficient
Literature


Appendix
Adaption with 0-point

1. Measuring the eye position relative to the right glass
2. Computation of the $\Delta d_i$, the distance between the $i$th person and the foreign person for x and y coordinate.
3. Assuming the constant $c_{\text{multi},i}$ from multi-point as optimal adaption constant
4. Compute the relation $f_i$ between $\Delta d_i$ and $c_{\text{multi},i}$: $f_i = \frac{c_{\text{multi},i}}{\Delta d_i}$
   In the optimal case all $f_i$'s are equal
5. Compute $f_{\text{mean}}$ from all $f_i$'s as constant factor for each person
6. Based on the $\Delta d_i$'s for each person $i$ the constant $c_{\text{para},i}$ can be computed.
Analyzing Tool: Computation the position of the true POG

2 algorithms

* uses the mean position from all distances between each single IR-Marker to the true POG

* uses the mean position from the positions of the true POG in all possible planes / coordinate systems (each spanned by 3 IR-Markers)