Towards Gaze-Based Assistance on Maps: Recognizing Activities from Eye Movements

Peter Kiefer, Ioannis Giannopoulos, Martin Raubal

24 August 2013
ICC Pre-Conference Workshop
Eye-Tracking: Why, When, and How?
References

- The following slides are mainly based on

- Some content is also taken from
Mobile Eye Tracking
Geoinformation Engineering, ETHZ

• Main focuses
  – Outdoor wayfinding studies
  – Gaze-based interaction

Kiefer et al., TGIS (2013, in print)

Outdoor wayfinding studies
Gaze-Based Interaction

• Eye Tracking: Why?
  – Create «better» cartographic user interfaces and maps
  – Bidirectional view on interfaces

→ Gaze as input modality

• Hardware: available/coming
• Interaction concepts?

[Image of a woman wearing Google Glass with a link to a TechRadar article about Google patents for eye-tracking.]
Gaze-Based Interaction

A) Explicit Interaction
   – «What you look at is what you get»
   – Midas touch problem

Based on Kiefer et al., ACM SIGSPATIAL (2013, accepted)

Gaze-based zooming/panning (Stellmach, Dachselt, 2012)
Gaze-Based Interaction

B) Implicit Interaction

- User does not intend to trigger an action
- System assists proactively, based on the gaze track

• Example
  - GeoGazemarks
  - Map usage history based on fixations
  - Help user for orientation

GeoGazemarks
Giannopoulos et al. (2012)
Gaze-Based Activity Recognition

• «Eye movements depend on task (activity)»
  – Infer activity from eye movement measures?
  – Dependency of map content, map design, person?

• Engineering perspective
  – RQ: Is it possible to build a classifier that solves the problem with sufficient accuracy?
  – (... for one type of map)

Gaze-based activity recognition
Kiefer et al., ACM SIGSPATIAL (2013, accepted)
Gaze-Based Activity Recognition

• «Eye movements depend on task (activity)»
  – Infer activity from eye movement measures?
  – Dependency of map content, map design, person?

• Engineering perspective
  – RQ: Is it possible to build a classifier that solves the problem with sufficient accuracy?
  – (... for one type of map)

Gaze-based activity recognition for office activities
(Bulling et al., 2011)
Data Collection

• 6 Activities
  – Free exploration (6 stimuli)
  – Global search (9)
  – Route planning (8)
  – Focused search (5)
  – Line following (8)
  – Polygon comparison (4)
• Random order of stimuli
• Google Maps
• 17 participants (no experts)
• 587 valid recordings, each cut after 20 seconds

Data collection setup
Kiefer et al., ACM SIGSPATIAL (2013, accepted)
Do you see Mickhausen (top right) and Boos (bottom left)?
Please, search for the shortest route from Mickhausen to Boos.

Example for a route planning task
After a moment: «Please, fixate now on Mickhausen. The task begins in 3, 2, 1 ...»

Example for a route planning task
Example for a route planning task
Machine Learning

• 229 features based on
  – Blinks (5)
  – Fixations (17)
  – Saccades (11)
  – Saccadic direction (100)
  – Direction sequences (96)

• Leading to 587 instances of
  – \([f_1, \ldots, f_{229}, \text{true-activity}]\)
  – Training set (90%)
  – Test set (10%)
    \(\Rightarrow [\ldots, \text{true-activity}, \text{assigned-activity}]\)

• Support Vector Machine (SVM)
  – Standard machine learning approach
  – Toolboxes available online (LibSVM, RapidMiner)

• 10-fold cross validation
## Results

<table>
<thead>
<tr>
<th>predicted activity</th>
<th>true activity</th>
<th>precision (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>82</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>∑</td>
<td>102</td>
<td>131</td>
</tr>
<tr>
<td>recall (%)</td>
<td>80.4</td>
<td>85.5</td>
</tr>
<tr>
<td>accuracy =</td>
<td>77.7%</td>
<td></td>
</tr>
</tbody>
</table>

1: free exploration, 2: search, 3: route planning, 4: focused search, 5: line following, 6: polygon comparison.

(Kiefer et al., 2013, submitted)
# Map vs. Office Activities

<table>
<thead>
<tr>
<th>true activity</th>
<th>precision (%)</th>
<th>predicted activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 82 10 3 11 2 0</td>
<td>75.9</td>
<td></td>
</tr>
<tr>
<td>2 14 112 5 18 3 0</td>
<td>73.7</td>
<td></td>
</tr>
<tr>
<td>3 1 2 65 1 19 0</td>
<td>73.9</td>
<td></td>
</tr>
<tr>
<td>4 4 7 5 54 1 0</td>
<td>76.1</td>
<td></td>
</tr>
<tr>
<td>5 1 0 21 1 76 1</td>
<td>76.0</td>
<td></td>
</tr>
<tr>
<td>6 0 0 1 0 0 67</td>
<td>98.5</td>
<td></td>
</tr>
<tr>
<td>∑ 102 131 100 85 101 68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**recall (%)** 80.4 85.5 65.0 63.5 75.3 98.5

**accuracy = 77.7%**

1: free exploration, 2: search, 3: route planning, 4: focused search, 5: line following, 6: polygon comparison.

(Kiefer et al., 2013, submitted)

Bulling et al., 2011

Kiefer et al., ACM SIGSPATIAL (2013, accepted)
Outlook

- Other map designs
- Other tasks
- Segmentation
- Assistive system
  - Recognition in real-time
  - What is the minimum time needed? (currently 20 sec)
  - User-satisfaction?
- Combine with gaze-map matching
  - Kiefer and Giannopoulos (2012)
  - «Snap» fixations to vector features
  - e.g., which route is the person planning?
- Most discriminative features?
Thank you for your attention!