Automatic analysis of eye-tracking data using object detection algorithms

Stijn De Beugher
Younes Ichiche
Geert Brône
Toon Goedemé
Introduction

- Mobile eye-tracker:
  - Big potential: natural environment, beyond lab conditions
    - Supermarket
    - Sports court
    - On the road

- Problem: *manual* data analysis
  - Large amounts of data
  - Existing methods for automatic processing not
Introduction
Content

- Introduction
- Proposed approach
- Suitable object-detection algorithms
- Experimental results
- Conclusion
Proposed technique

- Object recognition algorithm automatically analyses video stream (with gaze data)

- Benefits:
  - Target of analysis is not restricted to a region
  - Objects can be moving
  - Manual labour limited

- Invariant region matching techniques:
  - Algorithm defines interest regions
  - Descriptor vectors invariantly describe visual content of regions
  - Features we use are invariant to translation, rotation and scale
Need for a reliable and distinctive features

- What is a good feature:
  - Satisfies brightness constancy
  - Has sufficient (but not too much) texture variation
  - Does not deform too much over time

- Features can be used to match objects between two images
Overview suitable techniques

- **SIFT** (Scale-Invariant Feature Transform) [Lowe99]
  - Finds local maximum of Difference of Gaussian in space and scale

- **ASIFT** (Affine Scale-Invariant Feature Transform)
  - Affine invariant implementation of SIFT

- **SURF** (Speeded Up Robust Features) [Bay2004]
  - Find local maximum of Hessian (approximation of Laplacian of Gaussian)
  - Uses integral images for major speed up
Overview suitable techniques

SIFT: 25 matches / Keypoints: 3189
ASIFT: 2270 matches / Keypoints: 34985
SURF: 259 matches / Keypoints: 3080
## Comparative experiments

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Rotation Invariant</th>
<th>Scale Invariant</th>
<th>Affine Invariant</th>
<th>Efficiency</th>
<th>Speed</th>
<th>#Keypoints</th>
<th>Correctness</th>
<th>Resistance to noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFT</td>
<td>X</td>
<td>X</td>
<td></td>
<td>+++</td>
<td>++</td>
<td>++</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>ASIFT</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>SURF</td>
<td>X</td>
<td>X</td>
<td></td>
<td>+++</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
</tr>
</tbody>
</table>
Overview approach
Clustering double keypoints

- Clustering of both inter- and intra-object double keypoints
Processing keypoint matches

- An object is recognized if:
  - It has the highest number of matches
  - This number exceeds a fixed threshold

- Statistics of each recognized object are stored:
  - How often the viewer fixated to that object
  - For how long it was fixated during the experiment
Graphical output
Conclusion

- Mobile eye-tracking hardware boom
  - Big potential: natural environment, beyond lab conditions
  - Datasets too large for manual analysis
  - IR-marker-based approaches not applicable in natural environments

- Proposed technique:
  - Object recognition algorithms for data analysis
  - Lots of benefits as compared to IR-markers

- Feasibility experiments
  - First results promising
  - Follow-up project started
Contact

- E-mail:
  - geert.brone@lessius.eu
  - stijn.debeugher@lessius.eu
  - toon.goedeme@lessius.eu
  - younes.ichiche@gmail.com

- Website:

- Questions?