Measuring and Inferring User Interest from Gaze

Yixuan Li
Cornell

Pingmei Xu, Dmitry Lagun, Vidhya Navalpakkam
Google
Inferring user interest from gaze
Eye-tracker

Gaze estimation using front-facing camera on mobile device

Training data
- Images
- Gaze points

Inference

TensorFlow
Current model accuracy

Radii (pixels):
- \( r_1 = 140 \)
- \( r_2 = 190 \)
Demo
Overview of eye-tracker backend

Server

Gaze Inference

Visualizing and Modeling

Calibration Points

Video Recording

Screenshot, Condition

User Interactions
Calibration
User study app is task based

Task Description

Task 2 (out of 33): [Start with calibration]. Suppose you are browsing a page of game collections. Pick up to 5 games you’d be interested in learning more about. Please hit ‘back’ button when you are done and fill out the response form. [1/5 collections]

Start Recording

Static Screenshot (scrollable)

Next Task
Collection page

- Playstore Collection Page
  - 5 Game collection pages

- Task: Ask user to browse the page and choose up to 5 items in mind.

- After browsing, ask user to mark interesting items (checkbox)

- Repeat calibration after each page.
Data collection

- Statistics
  - 2 weeks of lab study
  - 36 participants
  - 5 recordings per participants
  - 180 recording sessions
Gaze metrics

- **Page-level metrics**
  - Time on page (in seconds)
  - Number of scrolls
  - ...

- **Area of Interest (AOI) level metrics**
  - **Viewport**: Time on AOI (in seconds), Fraction of time on AOI, Time to first visit
  - **Gaze**: Time on AOI (in seconds), Fraction of time on AOI, Time to first visit
  - ...

Is there any preference for the left column?

![Comparison of gaze metrics graph with p<0.01 significance level]
Position effect on double column layout pages

Vertical position

Position effect on gaze dwell time
- left column
- right column

Gaze

Vertical Position
Position effect on double column layout pages

Vertical position

Position effect on gaze dwell time
- Left column
- Right column
Can we infer user’s interest from gaze?
Can we infer user interest from gaze?

**Comparison of gaze metrics**

- **Unrated**: 1.21 ± 0.03 (s)
- **Interested**: 2.05 ± 0.06 (s)

*** p<10e-7
Can we infer user interest from gaze?

Yes?
No?
Can we infer user interest from gaze?

**AUC (Area Under the ROC Curve):** better classifier gives AUC closer to 1

<table>
<thead>
<tr>
<th></th>
<th>Viewport</th>
<th></th>
<th>Gaze</th>
<th></th>
<th>Viewport &amp; Gaze</th>
<th></th>
<th>All Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/ Position</td>
</tr>
<tr>
<td><strong>SVM (RBF kernel)</strong></td>
<td>60.50</td>
<td>66.38</td>
<td>69.02</td>
<td>75.56</td>
<td>75.54</td>
<td>82.90</td>
<td>90.32</td>
</tr>
<tr>
<td><strong>Random Forest</strong></td>
<td>69.90</td>
<td>75.44</td>
<td>77.26</td>
<td>78.60</td>
<td>79.40</td>
<td>80.62</td>
<td>84.83</td>
</tr>
<tr>
<td><strong>Decision Tree</strong></td>
<td>64.85</td>
<td>66.49</td>
<td>72.13</td>
<td>72.57</td>
<td>74.05</td>
<td>73.62</td>
<td>80.62</td>
</tr>
</tbody>
</table>
Can we infer user interest from gaze?

**AUC (Area Under the ROC Curve):** better classifier gives AUC closer to 1

<table>
<thead>
<tr>
<th></th>
<th>Viewport</th>
<th>Gaze</th>
<th>Viewport &amp; Gaze</th>
<th>All Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
</tr>
<tr>
<td><strong>SVM (RBF kernel)</strong></td>
<td>60.50</td>
<td>66.38</td>
<td>69.02</td>
<td>75.56</td>
</tr>
<tr>
<td><strong>Random Forest</strong></td>
<td>69.90</td>
<td>75.44</td>
<td>77.26</td>
<td>78.60</td>
</tr>
<tr>
<td><strong>Decision Tree</strong></td>
<td>64.85</td>
<td>66.49</td>
<td>72.13</td>
<td>72.57</td>
</tr>
</tbody>
</table>
Can we infer user interest from gaze?

AUC (Area Under the ROC Curve): better classifier gives AUC closer to 1

<table>
<thead>
<tr>
<th></th>
<th>Viewport</th>
<th></th>
<th></th>
<th>Viewport &amp; Gaze</th>
<th>All Features</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
</tr>
<tr>
<td>SVM (RBF kernel)</td>
<td>60.50</td>
<td>66.38</td>
<td>69.02</td>
<td>75.56</td>
<td>75.54</td>
<td>82.90</td>
<td>90.32</td>
</tr>
<tr>
<td>Random Forest</td>
<td>69.90</td>
<td>75.44</td>
<td>77.26</td>
<td>78.60</td>
<td>79.40</td>
<td>80.62</td>
<td>84.83</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>64.85</td>
<td>66.49</td>
<td>72.13</td>
<td>72.57</td>
<td>74.05</td>
<td>73.62</td>
<td>80.62</td>
</tr>
</tbody>
</table>
Can we infer user interest from gaze?

AUC (Area Under the ROC Curve): better classifier gives AUC closer to 1

<table>
<thead>
<tr>
<th></th>
<th>Viewport</th>
<th>Gaze</th>
<th>Viewport &amp; Gaze</th>
<th>All Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
</tr>
<tr>
<td>SVM (RBF kernel)</td>
<td>60.50</td>
<td>66.38</td>
<td>69.02</td>
<td>75.56</td>
</tr>
<tr>
<td>Random Forest</td>
<td>69.90</td>
<td>75.44</td>
<td>77.26</td>
<td>78.60</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>64.85</td>
<td>66.49</td>
<td>72.13</td>
<td>72.57</td>
</tr>
</tbody>
</table>
Can we better personalize item order?
Can we better personalize feed relevance?

\[
\text{DCG}_p = r_{e1} + \sum_{i=2}^{p} \frac{r_{e_i}}{\log_2(i)}
\]

**Normalized DCG:** better ranking model gives NDCG score closer to 1.

<table>
<thead>
<tr>
<th>Position</th>
<th>Viewport</th>
<th>Gaze</th>
<th>Viewport &amp; Gaze</th>
<th>All Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
</tr>
<tr>
<td>NDCG@3</td>
<td>0.176</td>
<td>0.257</td>
<td>0.413</td>
<td>0.465</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.184</td>
<td>0.232</td>
<td>0.405</td>
<td>0.410</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.322</td>
<td>0.414</td>
<td>0.538</td>
<td>0.567</td>
</tr>
</tbody>
</table>
Can we better personalize feed relevance?

**DCG_p** = \( rel_1 + \sum_{i=2}^{p} \frac{rel_i}{\log_2(i)} \)

**Normalized DCG**: better ranking model gives NDCG score closer to 1.

<table>
<thead>
<tr>
<th>Position</th>
<th>Viewport</th>
<th>Gaze</th>
<th>Viewport &amp; Gaze</th>
<th>All Features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o Position</td>
<td>w/ Position</td>
<td>w/o Position</td>
<td>w/ Position</td>
</tr>
<tr>
<td>NDCG@3</td>
<td>0.176</td>
<td>0.257</td>
<td>0.322</td>
<td>0.413</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.184</td>
<td>0.232</td>
<td>0.281</td>
<td>0.405</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.322</td>
<td>0.414</td>
<td>0.434</td>
<td>0.538</td>
</tr>
</tbody>
</table>
Can we better personalize feed relevance?

\[ \text{DCG}_p = \text{rel}_1 + \sum_{i=2}^{p} \frac{\text{rel}_i}{\log_2(i)} \]

Normalized DCG: better ranking model gives NDCG score closer to 1.

<table>
<thead>
<tr>
<th>Position</th>
<th>Viewport w/o Position</th>
<th>Viewport w/ Position</th>
<th>Gaze w/o Position</th>
<th>Gaze w/ Position</th>
<th>Viewport &amp; Gaze w/o Position</th>
<th>Viewport &amp; Gaze w/ Position</th>
<th>All Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDCG@3</td>
<td>0.176</td>
<td>0.257</td>
<td>0.322</td>
<td>0.413</td>
<td>0.465</td>
<td>0.452</td>
<td>0.594</td>
</tr>
<tr>
<td>NDCG@5</td>
<td>0.184</td>
<td>0.232</td>
<td>0.281</td>
<td>0.405</td>
<td>0.410</td>
<td>0.406</td>
<td>0.513</td>
</tr>
<tr>
<td>NDCG@10</td>
<td>0.322</td>
<td>0.414</td>
<td>0.434</td>
<td>0.538</td>
<td>0.567</td>
<td>0.577</td>
<td>0.695</td>
</tr>
</tbody>
</table>
Summary

- Longer gaze time on interesting items vs. unrated ones.
- Purely attention based interest inference model is effective (AUC 90.32%).
- Improving feed relevance and personalization with gaze can be possible.
Implications for designing recommendation system

- Prioritizing items that are most likely of interest to user based on historical attention behavior.
- Prioritizing positions receiving significant portion of gaze attention.
Questions? Please contact

Yixuan Li: yli@cs.cornell.edu
Pingmei Xu: pingmeix@google.com
Dmitry Lagun: dlagun@google.com
Vidhya Navalpakkam: vidhyan@google.com