Lepton: Understanding Mobile Reading via Camera Based Gaze Tracking and Kinematic Touch Modeling

Wei Guo and Jingtao Wang
Mobile Interfaces and Pedagogical Systems (MIPS) Group http://mips.lrdc.pitt.edu
Department of Computer Science and LRDC, University of Pittsburgh
Pittsburgh, PA 15213, U.S.A.
{weg21, jingtaow}@cs.pitt.edu

Motivation

Mobile reading is experiencing rapid growth in the era of smartphones. The passive nature and the complicated reading context of mobile reading lead to declined attention and increased non-linear reading pattern. Mainstream solutions, click-through-rate (CTR) and dwell time, are too sparse and indirect to track and understand reading.

Our Approach

Lepton tracks the periodic lateral patterns, i.e. saccade, of users’ eye gaze via the front camera and infers their muscle stiffness during text scrolling via a Mass-Spring-Damper (MSD) based kinematic model for touch events on smart phone. Lepton combines a visual tracking channel and a kinematic channel to monitor and improve mobile reading activities.

Visual Channel

Captures periodic lateral pattern-based eye gaze features through face-detection and eye-tracking.

- Periodic lateral pattern based line detection
  - X-axis gaze signals
  - Pre-process
  - Detector: Peak/Valley
  - LivePulse
  - Merge Conservative Shorts

Algorithm framework for periodic lateral pattern detection

Kinematic Channel

Captures scrolling features through Mass-Spring-Damper (MSD) analysis on touch events:

- Mass (m): arm and finger(s)
- Spring (constant K): damper (coefficient C): muscles elements
- Linear predictive coding (LPC) to invert the input and output of the MSD model:

\[ x[n] = \sum_{k=1}^{\infty} a_k x[n-k] \]

\[ R(k) = R(0) \]

Damping frequency ω and damping ratio ζ calculated by taking the complex root:

\[ \omega = |\beta| \]

\[ \zeta = |\beta| \]

Abstract Mass-Spring-Damper (MSD) model

User study

We conducted a within-subject user study to understand the performance of Lepton. 25 participants (9 females) ranging from 19 to 35 years old (μ=26.32, σ=3.96) were recruited for the study

Phase 1: Periodic Lateral Pattern Detection

- Goal: Qualify the gaze periodic lateral pattern detection accuracy to prove the validity of periodic lateral pattern based gaze features

Phase 2: Cognitive State Inference

- Goal: Understand reading with the proposed two channel features and the traditional dwell time related features

Results

- Lepton 1) detects readers’ periodic lateral pattern with an acceptable accuracy; 2) predicts users’ comprehension, concentration, confidence, and interestingness more accurately compared to traditional dwell time related features.

Phase 1 – Line detection

<table>
<thead>
<tr>
<th>Method</th>
<th>Reread</th>
<th>Skip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.69</td>
<td>0.79</td>
</tr>
<tr>
<td>Y-only</td>
<td>0.58</td>
<td>0.31</td>
</tr>
<tr>
<td>X-line-counting</td>
<td>0.73</td>
<td>0.88</td>
</tr>
</tbody>
</table>

AUC under reread and skip condition

Phase 2 – Reading comprehension and engagements

<table>
<thead>
<tr>
<th>Features</th>
<th>Concentration</th>
<th>Confidence</th>
<th>Interestingness</th>
<th>Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF (1)</td>
<td>1.000 (1.11)</td>
<td>1.170 (5.33)</td>
<td>1.460 (0.99)</td>
<td>0.860 (0.17)</td>
</tr>
<tr>
<td>SF (16)</td>
<td>0.960 (3.2)</td>
<td>1.200 (2.9)</td>
<td>1.420 (2.27)</td>
<td>0.920 (2.1)</td>
</tr>
<tr>
<td>GF (3)</td>
<td>1.030 (0.05)</td>
<td>1.220 (1.13)</td>
<td>1.470 (0.07)</td>
<td>0.910 (0.07)</td>
</tr>
<tr>
<td>TP+SF</td>
<td>0.920 (4.0)</td>
<td>1.070 (4.7)</td>
<td>1.380 (3.4)</td>
<td>0.870 (3.2)</td>
</tr>
<tr>
<td>TP+GF</td>
<td>1.000 (1.44)</td>
<td>1.040 (3.9)</td>
<td>1.440 (1.4)</td>
<td>0.860 (2.0)</td>
</tr>
<tr>
<td>TP+SF+GF</td>
<td>0.980 (4.6)</td>
<td>1.040 (5.1)</td>
<td>1.370 (3.8)</td>
<td>0.840 (3.9)</td>
</tr>
</tbody>
</table>

The root mean square error (corresponding R² value) when linearly modeling reading via tradition features (TF), scrolling features (SF) and gaze features (GF).

Conclusions

Lepton captured and used the saccade patterns and muscle stiffness signals as expressive and effective features to better predict users’ comprehension (+318.18% in R-square), concentration (+54.55%), confidence (+322.22%), and interestingness (+129.41%).