Enhanced Representation of Web Pages for Usability Analysis with Eye Tracking

Raphael Menges, Hanadi Tamimi, Chandan Kumar, Tina Walber, Christoph Schaefer and Steffen Staab
Motivation

• Analysts estimate interface usability with eye tracking
• Usually performed for static stimulus (screenshot, image)
• Web pages: Dynamic and active stimulus (Blascheck et al.)¹

• How to enable efficient large-scale Web studies?

¹Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D. and Ertl, T. (2017), Visualization of Eye Tracking Data: A Taxonomy and Survey
Table of Contents

1. State-of-the-Art
2. Problems
3. Our Method
4. Evaluation
State-of-the-Art Representations

Video Recording
- Viewport position of users diverges through scrolling
  → Analysis must be performed per video, which makes the Video Recording method not scalable

Virtual Screenshot*
- Virtually extends browser viewport to capture entire page
- Maps screen-space gaze data onto Virtual Screenshot
  → Virtual Screenshot method is not accurate for analysis on viewport-relative elements…

*as used by sticky.ai, eyezag.de, realeye.io or Tobii Studio Pro
Problems of Virtual Screenshot Method

(P1) Infinite scrolling pages
(P2) Viewport-relative sized elements
(P3) Viewport-relative positioned elements

Idea: Combination of structural information and pixel data
Our Enhanced Representation Method

(a) Acquire information about fixed elements.

(b) Cropping of fixed elements from viewport screenshot.

(c) Integration into stitched screenshot.

(d) Combination of stitched screenshot and fixed elements.

Repeat over interaction with Web page

Final step
Our Enhanced Representation Method

The problems of the Virtual Screenshot are solved:

(P1) Infinite scrolling pages
→ Dynamic additions included

(P2) Viewport-relative sized elements
→ As displayed to the user

(P3) Viewport-relative positioned elements
→ Identified, cropped and correctly associated with gaze
Evaluation

Fixed elements and associated gaze data are placed either on top or bottom of stitched screenshot.

Hypotheses:

• **(H1) Accuracy**: The *Enhanced Representation* method supports the analysis of gaze data on fixed elements as accurate as the Video Recording.

• **(H2) Scalability**: For analyzing gaze data from multiple users, the *Enhanced Representation* method would be more efficient than a Video Recording.
Evaluation Setup

Dataset

We asked them to find a certain function placed in footer

2 Users → 4 Web pages with fixed elements* → 4 Enhanced Representations + 8 Video Recordings

Analysis

Analyze attention on fixed elements

5 Participants → WWW₁ → WWW₂ → WWW₃ → WWW₄

WWW₁: Video Recording
WWW₂: E
WWW₃: V
WWW₄: E

5 Participants

*diogg.com, jimdo.com, yelp.com and creativecommons.org (CC)
Evaluation Setup – Web Pages
Evaluation Results – Accuracy

Analysts report *Time to First Fixation (TTFF) and Total Fixations (TF)* within fixed element.

E = Enhanced Representation, V = Video Recording

|       | Digg.com | Jimdo.com | Yelp.com | CC     |
|-------|----------|-----------|----------|--------|
| **E** |          |           |          |        |
| TTFF  | 3.0 ± 6.2| 0.0 ± 0.0 | 1.2 ± 3.8| 0.0 ± 0.0|
| TF    | 9.4 ± 17.8| 31.3 ± 44.3| 2.5 ± 7.9| 1.0 ± 3.2|
| **V** |          |           |          |        |
| TTFF  | 15.3 ± 31.8| 0.0 ± 0.0 | 1.2 ± 3.8| 0.0 ± 0.0|
| TF    | 15.3 ± 17.1| 0.0 ± 0.0 | 2.5 ± 7.9| 0.0 ± 0.0|

Average absolute percentage errors

→ Validates hypothesis about accuracy (H1)
Evaluation Results – Task Completion Time

Box plot of task completion time.
E = Enhanced Representation, V = Video Recording

→ Supports our hypothesis (H2) about scalability
Evaluation Results – Temporal Demand

NASA-TLX Raw values.

- Significant difference ($p = .017$)
- Medium to high effect size (0.38)

→ Supports our hypothesis (H2) about scalability, too
Contribution and Future Work

• Our method allows a scalable and accurate analysis of attention on Web pages
  • As good as Virtual Screenshot for page-relative content
  • As good as Video Recording for viewport-relative content

→ Allows efficient analysis of high number of Web page users

• Future Work
  • Improve precision of fixed element cropping (e.g., shadows)
  • Cover dynamics on Web pages, like carousels, etc.

Thank you for your attention!

Raphael Menges, Hanadi Tamimi, Chandan Kumar, Tina Walber, Christoph Schaefer and Steffen Staab