Gaze Analytics Pipeline for Unity 3D Integration

Signal Filtering and Analysis

Andrew T. Duchowski[1] Sabarish V. Babu[1], Jeff Bertrand[1], and Krzysztof Krejtz[2]

[1] School of Computing, Clemson University, Clemson, SC, USA
[2] National Information Processing Institute and University of Social Sciences and Humanities, Warsaw, Poland

Abstract. A data processing pipeline is described for gaze analytics featuring velocity-based signal filtering leading to inferential statistics of fixation transitions. The approach is demonstrated on data collected from a virtual environment study conducted in Unity 3D testing the visibility of a wayfinding aid.

Keywords: virtual environments, wayfinding, gaze analytics

1 Introduction & Background

Analysis of human spatial cognition when navigating within physical or virtual environments can be bolstered by collecting eye movements with the help of an eye tracker. In physical reality, a head-mounted tracker can be worn during navigation with recorded gaze data often mapped to the screen coordinates of a forward-facing camera. In virtual reality, gaze data can be recorded with a so-called remote, or table-mounted eye tracker placed in front of the display screen on which the virtual environment is presented. In both cases gaze data will eventually need to be processed in order to infer insights about human visual attention to elements in the environment.

Application of inferential statistics to collected gaze data often relies on characterization of the raw data into fixations, usually derived from some form of filtering, e.g., dispersion-based or velocity-based. Unfortunately, most commercial software packages provide only a limited choice of fixation detection algorithms (i.e., filters), often hiding implementation details or filter parameters from the user. Some systems (e.g., Ogama) still rely on the dispersion-based “fixation pickers” [7] which have been shown to be less than reliable, particularly when evaluating data captured on different platforms, at different sampling rates [10]. Velocity-based filters, or “saccade pickers”, while perhaps more difficult to tune, offer a more reliable alternative. Beyond the lack of sufficient control over filtering parameters, commercial packages often do not include flexible means for statistical analysis. This is hardly surprising, however, since eye tracking vendors can hardly be expected to anticipate all possible experimental designs for which their devices are used.

In this paper we describe a gaze analytics pipeline through which raw gaze data is processed. The pipeline consists of the following steps:

1. denoising and filtering raw gaze data $g_i = (x_i, y_i, t_i)$, and classifying raw gaze into fixations $f_i = (x_i, y_i, t_i, d_i)$, where (x_i, y_i) coordinates indicate the position of the gaze

ET4S 2014, September 23, 2014, Vienna, Austria

(a) No wayfinding aid *(b) Wayfinding aid at bottom-right*

Fig. 1. Representative screenshots of interactive virtual environment rendered in Unity 3D [2].

point or centroid of the fixation, with t_i indicating the timestamp of the gaze point or fixation and d_i the fixation's duration,

2. collating fixation-related information for its subsequent statistical comparison,
3. interpreting and visualizing statistical tests conducted on processed data.

Visualization of the data at each stage of the pipeline is particularly helpful in fine-tuning parameters, such as threshold levels for velocity-based filtering.

We demonstrate the utility of the analytics pipeline on data collected from a wayfinding study in a virtual environment (see Fig. 1) [2]. Previously, analysis was conducted only on smoothed gaze data, with comparison of gaze time on, gaze transitions to, and proportion of gaze time over screen elements. Here, we revisit this data set and feed the raw data through the analytics pipeline, terminating in entropy transition matrix analysis of fixations captured on an Area Of Interest (AOI) grid overlaid atop the screen.

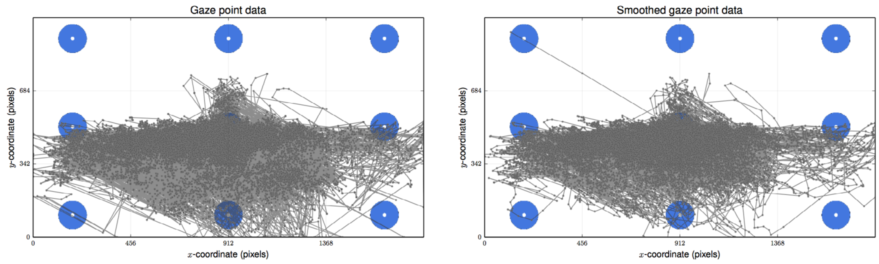
2 Fixation Filtering

The eye tracker outputs a stream of gaze points (x_i, y_i, t_i). Typically, this data is noisy and requires smoothing (see Fig. 2). Treating x_i or y_i independently, smoothing or differentiating (to order s) is achieved by convolving $2p+1$ inputs with filter $h_i^{t,s}$ and $2q+1$ (previous) outputs $\dot{x}_i$ or $\dot{y}_i$ with filter $g_i^{t,s}$ at midpoint i [6]:

$$\dot{x}_n^s(t) = 1/(\Delta t^s)\left(\sum_{i=-p}^{p} h_i^{t,s} x_{n-i} - \sum_{i=-q}^{q} g_i^{t,s} \dot{x}_{n-i},\right)$$

and similarly for y_i and $\dot{y}_i$, where n and s denote the polynomial fit to the data and its derivative order, respectively [5,10]. Based on prior work and evaluation of calibration data, we chose a 4^{th} order Butterworth filter to smooth the raw gaze data with sampling and cutoff frequencies of 60 and 6.15 Hz, respectively [3] (see Fig. 2(b)).

Following Andersson et al. [1] and Nyström and Holmqvist [9], a second-order Savitzky-Golay (SG) filter [11] is used to differentiate the (smoothed) positional gaze signal into its velocity estimate. The Savitzky-Golay filter fits a polynomial curve of order n via least squares minimization prior to calculation of the curve's s^{th} derivative



(a) Raw gaze data *(b) Smoothed gaze data*

Fig. 2. Representative (individual) gaze data smoothing (35,024 points, approx. 20 min time interval). Blue discs indicate positions of 9 calibration points used in the study and the rectangular grid shows the 4×3 rectangular AOIs used in the analysis.

(e.g., 1^{st} derivative ($s = 1$) for velocity estimation). We use a 6-tap (96 ms) SG filter with a threshold of ± 20 deg/s to produce fixations (see Fig. 3).

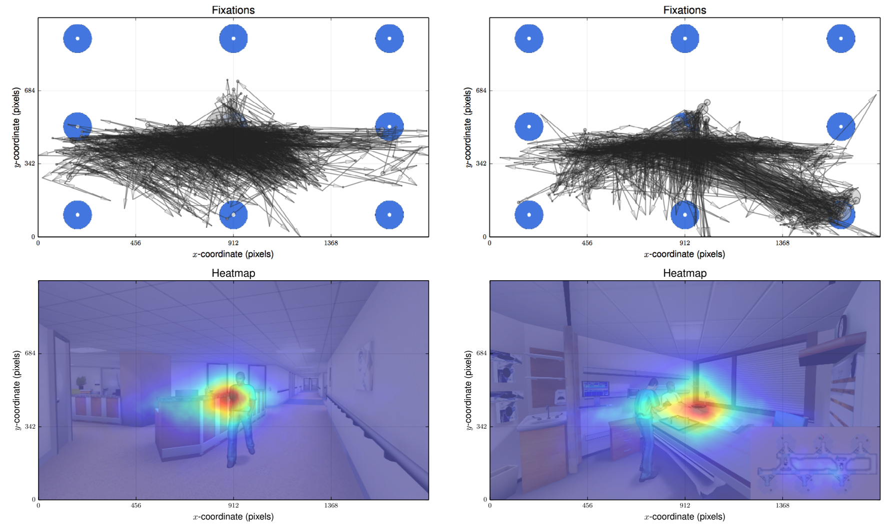
Fine-tuning of the velocity threshold in degrees per second depends on viewing distance and screen resolution (e.g., in dots per inch). In the exemplar fixations of Fig. 3, an 1824×1026 display (80″ diagonal) was viewed at 70″.

3 Statistical Comparison

Quantitative analysis of filtered fixation data generally depends on application of inferential statistics, e.g., comparison of means via analysis of variance (ANOVA). Typical eye movement metrics include number of saccades, saccade length and duration, saccadic amplitude, convex hull area, spatial density, number of fixations, fixation durations, and a fixation/saccade ratio [4]. We extend these analyses by considering entropy transition matrix analysis of fixations [8].

Assigning a character label to each of the 4×3 grid cells leads to a $\{a, b, c, d\} \otimes \{a, b, c\}$ labeling scheme (with cell *aa* at bottom-left in Fig. 3). Accumulating single fixation transitions between cells and normalizing to the source leads to a first-order Markov model of gaze transitions represented by transition matrices visualized in Fig. 4. The matrix representing the viewing condition with the wayfinding aid (Fig. 4(b)) shows a higher probability of transitions to the bottom-right cell where the aid was present (compare columns *da* in the matrices). A critical question is whether these transitions, on average, differed significantly under the given experimental conditions.

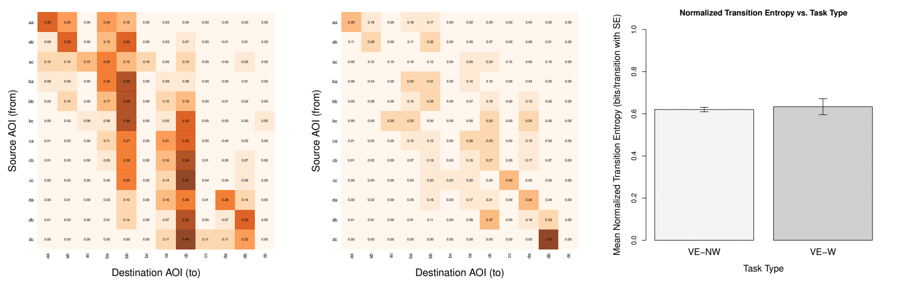
Considering the set of AOIs as $\mathcal{S} = \{1, \ldots, s\}$, transition matrices afford computation of Shannon's entropy $H_t = -\sum_{i \in \mathcal{S}} \pi_i \sum_{j \in \mathcal{S}} p_{ij} \log p_{ij}$ where p_{ij} denotes the probability of transitioning from the i^{th} to the j^{th} AOI, which in turn allows statistical comparison of matrices. In this particular instance, a Welch two sample t-test shows lack of significance of fixation transition entropies between the two wayfinding aid conditions ($t = -0.85$, $p = 0.41$, n.s.). Lack of significance is likely due to the relatively low number



(a) Data from session without wayfinding aid *(b) Data from session with wayfinding aid*

Fig. 3. Representative (individual) scanpaths from filtering of smoothed data. The scanpath captured with no wayfinding aid (1543 fixations) corresponds to the raw and smooth data in Fig. 2.

of participants. Presence of the wayfinding aid resulted in a higher mean transition entropy (M = 0.65, SD = 0.12) than when it was absent (M = 0.62, SD = 0.04), suggesting that users tended to make (slightly) more transitions with the wayfinding aid present than without it, e.g., without the aid, viewers tended to transition to the screen center (grid cells *bb* and *cb*, c.f. Figs. 4(a)–4(b) and Fig. 4(c)). Power analysis suggests that



(a) Transition matrix from session without wayfinding aid *(b) Transition matrix from session with wayfinding aid* *(c) Statistical comparison of transition matrix entropies*

Fig. 4. Fixation transition matrices showing empirical probabilities of transitions from source cell to destination and entropy comparison.

the effect of the wayfinding aid on gaze transitions may reach significance ($p < 0.05$) with about 400 participants (per between-subjects group).

4 Discussion & Conclusions

A data pipeline was described for processing raw gaze data through filtering and velocity-based fixation classification followed by collation of fixations into transitions for entropy-based statistical comparison.

With gaze data recorded in individual XML files, the entire processing pipeline, complete with data visualizations, is readily implemented in Python and R, the free software environment for statistical computing.

References

1. Andersson, R., Nyström, M., Holmqvist, K.: Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. Journal of Eye Movement Research 3(3), 1–12 (2010)
2. Bertrand, J., Ebrahimi, E., Wachter, A., Luo, J., Babu, S.V., Duchowski, A.T., Meehan, N., Gramopadhye, A.K.: Visual Attention to Wayfinding Aids in Virtual Environments. In: Proceedings of the Joint Virtual Reality Conference (JVRC) (11-13 December 2013)
3. Duchowski, A.T., Pelfrey, B., House, D.H., Wang, R.: Measuring Gaze Depth with an Eye Tracker During Stereoscopic Display. In: Applied Perception in Graphics & Visualization (APGV). ACM, Toulouse, France (August 27-28 2011)
4. Goldberg, J.H., Kotval, X.P.: Computer Interface Evaluation Using Eye Movements: Methods and Constructs. International Journal of Industrial Ergonomics 24, 631–645 (1999)
5. Gorry, P.A.: General Least-Squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method. Analytical Chemistry 62(6), 570–573 (1990), `http://pubs.acs.org/doi/abs/10.1021/ac00205a007`
6. Hollos, S., Hollos, J.R.: Recursive Digital Filters: A Concise Guide. Exstrom Laboratories, LLC., Longmont, CO (April 2014), iSBN: 9781887187244 (ebook), URL: `http://www.abrazol.com/books/filter1/`
7. Karn, K.S.: “Saccade Pickers” vs. “Fixation Pickers”: The Effect of Eye Tracker Choice on Research Findings (Panel Discussion). In: ETRA ’00: Proceedings of the 2004 Symposium on Eye Tracking Research & Applications. pp. 87–88. ACM, New York, NY (2000)
8. Krejtz, K., Szmidt, T., Duchowski, A.T., Krejtz, I.: Entropy-based Statistical Analysis of Eye Movement Transitions. In: Proceedings of the 2014 Symposium on Eye-Tracking Research and Applications. ETRA ’14, ACM, New York, NY (2014), `http://doi.acm.org/10.1145/2578153.2578168`
9. Nyström, M., Holmqvist, K.: An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. Behaviour Research Methods 42(1), 188–204 (2010)
10. Ouzts, A.D., Duchowski, A.T.: Comparison of Eye Movement Metrics Recorded at Different Sampling Rates. In: Proceedings of the 2012 Symposium on Eye-Tracking Research and Applications. ETRA ’12, ACM, New York, NY (March 28-30 2012)
11. Savitzky, A., Golay, M.J.E.: Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36(8), 1627–1639 (1964), `http://pubs.acs.org/doi/abs/10.1021/ac60214a047`