August 2020

Improved Fixation Filter For Eye Tracking On Small Devices

Anonymous

Improved Fixation Filter For Eye Tracking On Small Devices

ABSTRACT

When standard eye tracking techniques developed in the context of large-screen devices are applied to applications on mobile devices, the results are often erroneous. For example, when the true percentage of viewers (as identified manually) that look at a profile picture is 100%, standard eye techniques typically underestimate the eye fixation behavior, returning a much smaller value. Many mobile applications utilize a feed interface that involves moving targets as the user scrolls the feed. This disclosure describes techniques that use appropriate temporal, spatial, and velocity-based parameters to define fixation by area of interest (AOI) size for mobile-feed applications. The techniques enable accurate measurement and analysis of visual user behavior during mobile-feed viewing, such as eye-gaze patterns, fixation durations, time to first fixation, percentage of viewers fixated, etc.

KEYWORDS

- Eye-tracking
- Fixation filter
- Saccade
- Area of interest (AOI)
- Gaze area
- Eye gaze
- Small-screen device
- Smartphone gaze
- Visual user behavior
- Feed viewing
- Fixation duration
- Time to first fixation

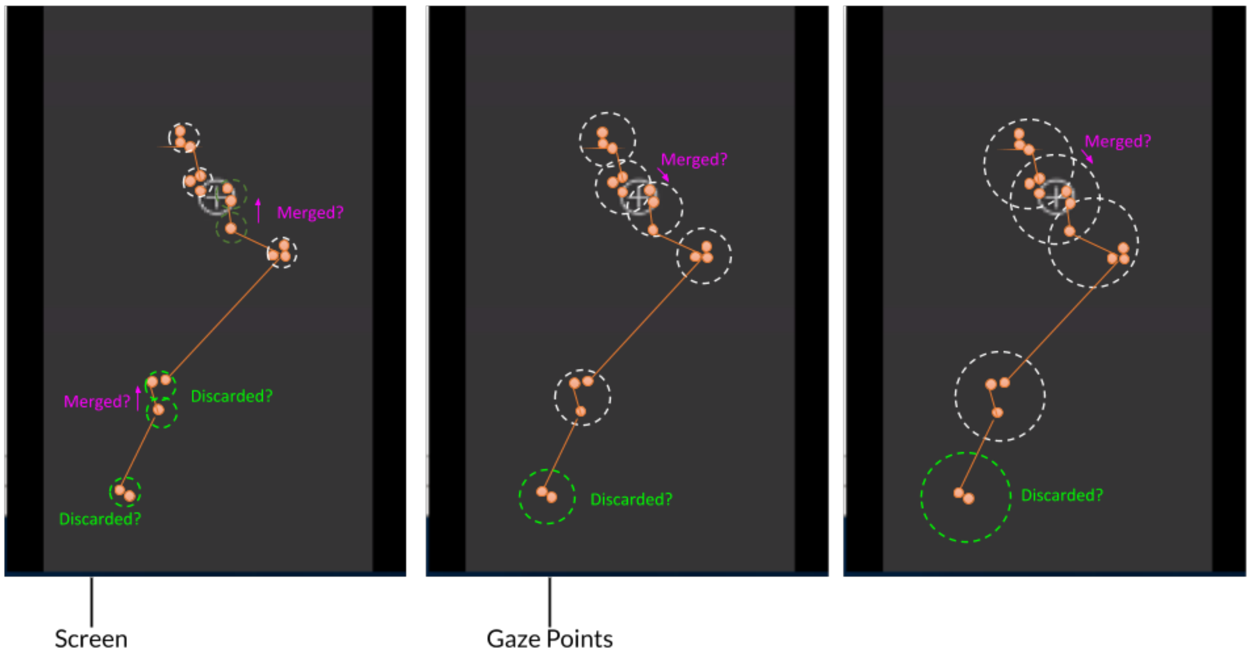
BACKGROUND

When standard eye tracking techniques developed in the context of large-screen devices are applied to applications on mobile devices, the results are often erroneous. For example, when the true percentage of viewers (as identified manually) that look at a profile picture in a social media post is 100%, standard eye techniques typically underestimate the eye fixation behavior, returning a much smaller value. A high rate of fixation is anticipated and correct since users typically determine the author, or voice, of the post by recognizing the face.

A reason for the inaccuracy of conventional eye-tracking metrics is that the algorithms to determine whether and how long someone is looking at an object, known as eye-fixation (or simply, fixation), were developed for large screens that are stationary at a distance of at least about 60 centimeters from the eye. Such is the case for televisions or laptops, where an area of interest (AOI) is relatively fixed in the visual field. The conventional eye-tracking metrics fail to produce accurate results for small, moving AOIs, such as those for a feed interface displayed on a mobile device that is about 30 cm from the eye.

DESCRIPTION

This disclosure describes techniques that use appropriate temporal, spatial, and velocity-based parameters to define fixation by area of interest (AOI) size. The parameters can be utilized to differentiate between a fixation (when the viewer is looking at an object), a saccade (when the viewer is gazing from point to point), and a smooth pursuit (when the viewer is following a moving target).



(a) (b) (c)

Fig. 1: Various thresholds for fixation filters: (a) A low threshold divides gaze points into large numbers of short fixations; (b) A medium threshold divides gaze points into fixations that accurately capture viewing behavior; (c) A high threshold merges gaze points into long fixations .

Fig. 1 illustrates the importance of setting the correct thresholds for eye-fixation filters. In Fig. 1(a), a low threshold divides the set of gaze points into a large number of short fixations, while in Fig. 1(c), a high threshold divides the gaze points into long fixations. A threshold of medium magnitude (Fig. 1(b)), divides the gaze points into fixations that accurately capture viewing behavior.

To determine the optimal settings for the analysis of eye-tracking data in conditions where the stimuli are presented on the screen of a smartphone (or other small screen device), the techniques described herein optimize three or more parameters, e.g., discarding of short-duration fixations, adjacent-fixation merging, and velocity threshold.

Discarding of short-duration fixations

Discarding short fixations is aimed at removing incorrectly classified fixations that are too short for information acquisition and processing. This parameter can be adjusted so that all fixations with duration below a specific threshold value are removed from the fixation data. In contrast with the 100-200 ms duration threshold for fixations on traditional, larger-screen AOIs, the temporal duration of fixation appropriate to small, moving AOIs on a mobile-device screen is found to be 60-80ms. Therefore, per the techniques, parameters are set to not discard short fixations from analysis for small, moving AOIs. However, on large AOIs, discarding fixations under 60 ms is appropriate.

Merging of adjacent fixations

Merging adjacent fixations is aimed at correcting for errors caused by noise and disturbances, due to which a single fixation is inappropriately split into multiple short fixations located close together. Although small AOIs on a small screen may be located in close proximity to each other, they may still require separate tabulations of fixation count and duration. Therefore, per the techniques, the spatial parameters do not merge adjacent fixations. However, on large AOIs, merging sequential fixations located within 0.2 degrees of each other is appropriate.

Velocity threshold

Velocity threshold is a parameter value based on which each data point is classified as being a part of a fixation or a saccade. A fixation comprises an unbroken chain of raw data samples that have the angular velocity below the velocity threshold. The coordinates of a fixation are computed as the arithmetic mean values of the coordinates of raw gaze samples that constitute the fixation. A saccade is a quick eye movement between fixation points.

Because the velocity of saccades between two nearby points is lower than saccades between two further apart points, the velocity threshold parameter is adjusted accordingly for small screens that are held at a close focal length to the eyes. The velocity threshold appropriate for mobile phone screens with small, moving targets is found to be 9-15 degrees/sec. However, on large AOIs, a velocity threshold of 30 degree/sec is appropriate.

Analytical procedure to determine parameters of the fixation filter

Fixation filter (FF)	**Actual gaze**	
	Target	**Noise**
Target (FF)	True positives (classifying target as target)	False positives (classifying noise as target)
Noise (FF)	False negatives (classifying target as noise)	True negatives (classifying noise as noise)

Table 1: Table of fixation filter predictions versus actual gaze values

The accuracy of different fixation filters is benchmarked on a metric known as total fixation duration (TFD), which is based on raw gaze. The TFD values based on raw data indicate the accumulated duration of all gaze points that landed on different areas of interest, e.g., a target area versus noise. As shown in Table 1, having the actual raw-gaze based values for TFD on the AOIs target and noise, the true positives, the true negatives, the false positives and false negatives can be computed. This approach allows to compare and contrast the output of different fixation filters with regards to the duration of fixations that are classified correctly or incorrectly.

False positives

An optimal gaze filter can be regarded as one that maximizes the proportion of correct classifications and minimizes the proportion of false classifications. However, when comparing TFD values based on fixation filter output and raw gaze samples, it is normal to have a certain proportion of false negatives in the data. TFD based on raw samples should generally be longer than the TFD based on fixation data, because raw gaze data also contains saccades and other types of noise, thereby representing a slight overestimation of the viewing time.

The situation is much worse when the data contains false positives. This means that the fixation filter merges together a number of gaze points that have actually landed outside the area of interest. False positives can be regarded as indicators of inaccuracy and unreliability of the fixation filter.

True positives and true negatives

To optimize the proportion of correct classifications, it is important to look into both the true positives and the true negatives. However, given the stimulus characteristics, e.g., a small target on the screen of a smartphone, the classification of true positives is likely more important than the classification of true negatives.

Fig. 2 illustrates an example of normalized true positives, false negatives, false positives, and true negatives as a function of velocity threshold in degrees per second for various discard/merge parameters. The dots are experimental observations while the solid lines are restricted maximum likelihood (REML) based least-squares regression fits with normalized true/false positives/negatives as dependent variables, and discard/merge, velocity threshold, and the interaction between the two as model predictors. The data of Fig. 2 was obtained with one target moving at 3 degrees/second speed. It is evident from Fig. 2 that adjustments to the velocity threshold do not have much impact on the true negatives, but the proportion of true positives increases as the threshold is increased. However, velocity thresholds starting from 12 degrees/sec introduce an increasing amount of false positives.

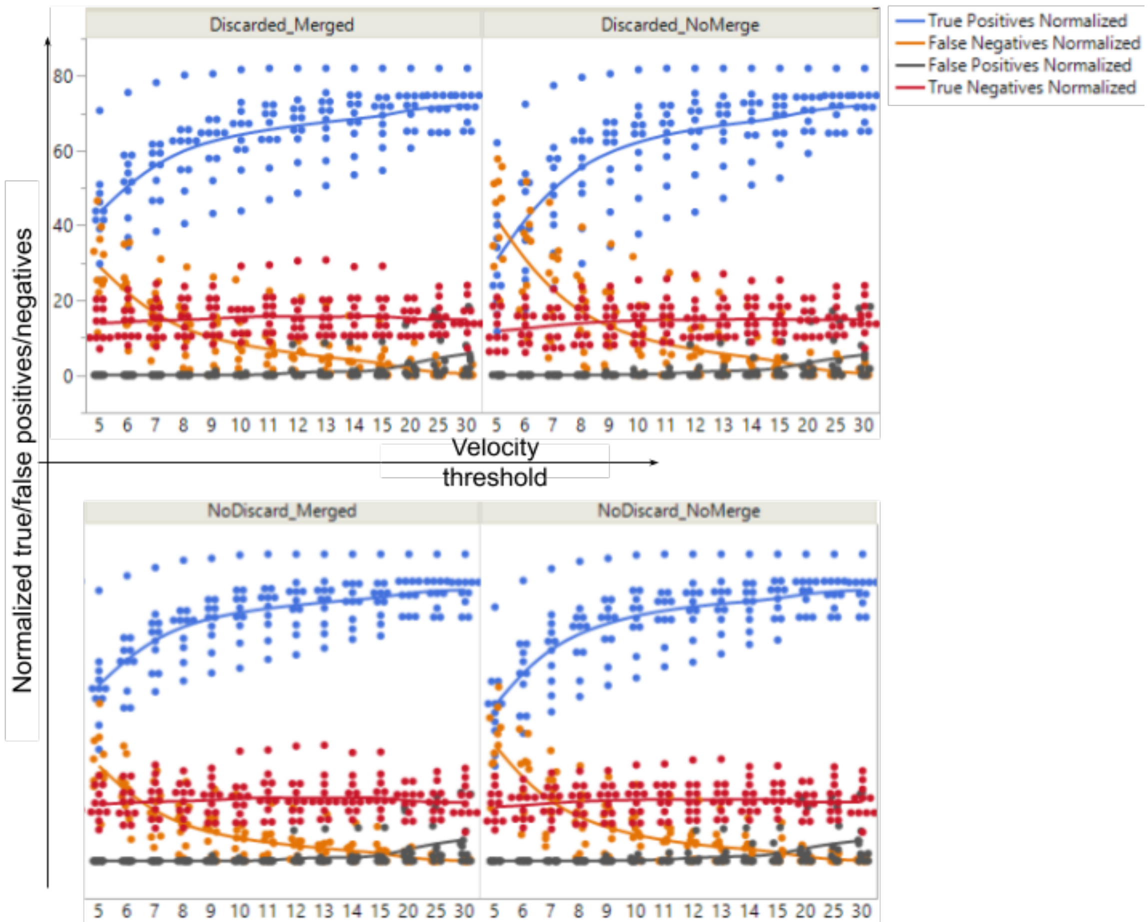


Fig. 2: Normalized true/false positives/negatives against velocity threshold in degrees/sec for various discarding and merging parameters

As mentioned earlier, true positives and false positives are important in determining the optimal fixation filter parameters. Through REML-based least-squares regression, it is found that discard/merge and velocity threshold are significant predictors of true positives, while for false positives, the velocity threshold is a significant predictor. In an example, an optimally-tuned fixation filter, e.g., with no discarding, adjacent fixations merged, and a velocity threshold of 11 degrees/sec, captures 93% of true positives while reducing false positives by 99.6%.

In this manner, the techniques of this disclosure enable the understanding, measurement, and analysis of visual user behavior during viewing content (e.g., a content feed) on a small screen device (e.g., a smartphone). Visual behavior can include eye-gaze patterns, fixation durations, time to first fixation, percentage of viewers fixated, etc. The techniques also enable the calculation of the gaze path taken by a viewer who that at posts in a feed; how long someone looked at elements of a post in a feed; how long it took before someone looked at an element once it came on screen; how many people looked at each element of a post; etc.

CONCLUSION

This disclosure describes techniques that use appropriate temporal, spatial, and velocity-based parameters to define fixation by area of interest (AOI) size for mobile-feed applications. The techniques enable accurate measurement and analysis of visual user behavior during mobile-feed viewing, such as eye-gaze patterns, fixation durations, time to first fixation, percentage of viewers fixated, etc.

REFERENCES

[1] Cheng, Shiwei, Jing Fan, and Anind K. Dey. "Smooth Gaze: a framework for recovering tasks across devices using eye tracking." *Personal and Ubiquitous Computing* 22, no. 3 (2018): 489-501.

[2] Kim, Jaewon, Paul Thomas, Ramesh Sankaranarayana, and Tom Gedeon. "Comparing scanning behaviour in web search on small and large screens." In *Proceedings of the Seventeenth Australasian Document Computing Symposium*, pp. 25-30. 2012.

[3] Agtzidis, Ioannis, Mikhail Startsev, and Michael Dorr. "A ground-truth data set and a classification algorithm for eye movements in 360-degree videos." *arXiv preprint arXiv:1903.06474* (2019).

[4] Al-Showarah, Suleyman, Al-Jawad Naseer, and Harin Sellahewa. "Effects of user age on smartphone and tablet use, measured with an eye-tracker via fixation duration, scan-path duration, and saccades proportion." In *International Conference on Universal Access in* Human-Computer Interaction, pp. 3-14. Springer, Cham, 2014.