Discrimination of Fixations and Smooth Pursuit Movements in High-Speed Eye-Tracking Data

Linnéa Larsson[1,2], Marcus Nyström[2], and Martin Stridh[1]

***Abstract*— A novel three-stage algorithm for detection of fixations and smooth pursuit movements in high-speed eye-tracking data is proposed. In the first stage, a segmentation based on the directionality of the data is performed. In the second stage, four spatial features are computed from the data in each segment. Finally, data are classified into fixations and smooth pursuit movements based on a combination of the spatial features and the properties of neighboring segments. The algorithm is evaluated under the assumption that the intersaccadic intervals represent fixations in data recorded when viewing images, and mainly smooth pursuit movements in data recorded when viewing moving dots. The results show that the algorithm is able to detect 94.3% of the fixations for image stimuli, compared to a previous algorithm with 80.4% detected fixations. For moving dot stimuli the proposed algorithm detects 86.7% smooth pursuit movements compared to 68.0% for the previous algorithm.**

I. Introduction

Measurement of eye movements is an important tool in basic research in, e.g., visual attention, perception [1], and cognition, as well as for studies investigating the functionality of the brain [2]. In these studies, the eye-tracking signal is used to investigate the different kinds of eye movements and their relationships to the underlying processes in the brain [3]. The two most common types of eye movements are the fixations and the saccades. Fixations are when the eyes are more or less still and visual information is gathered, and the saccades are the fast movements when the eyes are redirected from one position of interest to the next [4]. These types of eye movements are the most common ones when the stimuli are static, i.e., when images or texts are used. Recently, the interest in using dynamic stimuli, such as video clips, has been growing [5]. When there are moving objects in the stimuli, the eye movement smooth pursuit will also occur, corresponding to that the eyes are tracking the moving object. In order to be able to perform the eye movement smooth pursuit, a moving object is needed [6]. Since the focus for a long time has been on static stimuli, most algorithms are not developed to handle smooth pursuit movements. When these algorithms are applied to signals recorded during dynamic stimuli, the smooth pursuit movements will be spread into other types of movements and make the interpretation of these difficult, e.g., smooth pursuit movements may typically be erroneously classified as very long fixations with very short saccades in between [7]. Many of the measures that earlier have been used for analysis and interpretation of the eye movements during static stimuli, are based on characteristics of the fixations, e.g., the duration of fixations and the number of fixations [4]. When dynamic stimuli are used, these measures are still of interest and in order to be able to draw correct conclusions about the underlying processes, a robust algorithm for separation of fixations and smooth pursuit movements is needed.

Since the signal characteristics of fixations and smooth pursuit movements are overlapping [8], classification of fixations in the presence of smooth pursuit movements is a difficult task [5], [9]. A few algorithms developed for smooth pursuit movements detection in high-speed eye-tracking data have been proposed. In [9], three algorithms for detection of saccades, fixations, and smooth pursuit were evaluated. The outcome of the comparison was that an algorithm that uses a combination of velocity and dispersion (I-VDT), was performing best of the compared algorithms, and was less sensitive to changes in the parameter settings. An algorithm that makes use of the directional information in the data is the algorithm proposed by [10], which uses principal component analysis in combination with a velocity threshold to distinguish between saccades, fixations, and smooth pursuit movements. The algorithm was used to analyze saccades in humans and monkeys watching short video clips, but the detection performance was not further evaluated by the authors. There are a few algorithms developed for mobile eye-trackers, see [11], [12]. To our knowledge, there is today no commercial algorithm for separation between fixations and smooth pursuit movements.

In this paper, we propose a novel algorithm for separation between fixations and smooth pursuit movements for high-speed eye-tracking data recorded during static and dynamic stimuli, i.e., images and moving dots. The proposed algorithm is evaluated by comparing the results to that of the I-VDT algorithm, which was the best performing algorithm in a previous study [9].

II. Method

A. Preprocessing

The proposed algorithm is applied to the intersaccadic intervals resulting from the algorithm in [13], i.e., the intervals between the detected saccades, postsaccadic oscillations, and blinks. Since neither fixations nor smooth pursuit movements physiologically can have a velocity higher than 100°/s [14], the sample-to-sample velocities of the intervals are computed

[1]L. Larsson and M. Stridh are with the Department of Biomedical Engineering, Lund University, Sweden. (email: linnea.larsson@bme.lth.se, martin.stridh@bme.lth.se)

[2]L. Larsson and M. Nyström are with the Lund University Humanities Laboratory, Lund University, Sweden. (email: marcus.nystrom@humlab.lu.se)

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

and all samples in the beginning and/or in the end of the interval that are exceeding this threshold are removed.

B. Preliminary segmentation based on direction

Each intersaccadic interval is divided into windows, w_{i}, of length t_{w} ms, with an overlap of t_{o} ms. For all x- and y-coordinates contained in the window, the sample-to-sample directions are computed. The sample-to-sample direction, $\alpha(n)$, is the angle of the line between two consecutive coordinate pairs to the x-axis. In order to segment the signal based on the directional information, the sample-to-sample directions in each window are investigated using a Rayleigh test [15]. The p-value of the test, P_i, is computed for each window i. Since there is an overlap between the windows, each pair of x- and y-coordinates may belong to more than one window. The mean value of P_j, for all windows j which sample k belongs to is computed as,

$$P_{\mathrm{mean}}(k) = \frac{1}{K}\sum_{j=1}^{K} P_j \tag{1}$$

where K is the number of windows each sample belongs to, $k = 1, 2, ..., M$, and M is the number of samples in the intersaccadic interval. All consecutive samples in the interval satisfying either $P_{\mathrm{mean}}(k) \geq \eta_{\mathrm{P}}$ or $P_{\mathrm{mean}}(k) < \eta_{\mathrm{P}}$ are grouped into preliminary segments sharing similar properties in terms of consistent or non consistent directions. These preliminary segments are further analyzed in the next step.

C. Evaluation of spatial features in the position signal

For all preliminary segments with a duration longer than t_{min}, four parameters, p_{D}, p_{CD}, p_{PD}, and p_{R}, are calculated. These parameters describe the dispersion (D), the consistency in the direction (CD), the positional displacement (PD), and the range (R) of the segment. In order to measure the dispersion, Principle Component Analysis (PCA) is employed. The first principle component determines the direction in which the segment has its largest variance and the second principle component is chosen orthogonally to the first principle component. The parameter, p_{PD}, is calculated as the ratio between the lengths of the projections of the data to the first and the second principle component, d_{pc_1}, and d_{pc_2}, respectively.

$$p_{\mathrm{D}} = \frac{d_{\mathrm{pc}_2}}{d_{\mathrm{pc}_1}} \tag{2}$$

As proposed in [10], the parameter, p_{D}, measures if the segment is more spatially spread in one direction than in the other, i.e., a value of p_{D} close to one means that the segment is equally spread in both directions.

The second parameter, p_{CD}, measures if the segment has a consistent direction or not. It is determined by computing the Euclidean distance (ED) between the starting and ending positions of the interval, d_{ED}, and comparing it to d_{pc_1}.

$$p_{\mathrm{CD}} = \frac{d_{\mathrm{ED}}}{d_{\mathrm{pc}_1}} \tag{3}$$

Hence a value of p_{CD} lower than one corresponds to that the range of the data in the segment is much larger than the actual distance between the starting position and the ending position. The third parameter, p_{PD}, measures the relationship between d_{ED} and the trajectory length (TL) of the data in the segment, d_{TL}.

$$p_{\mathrm{PD}} = \frac{d_{\mathrm{ED}}}{d_{\mathrm{TL}}} \tag{4}$$

This third parameter reflects the trajectory length of the segment compared to the positional displacement of the interval, i.e., a straight line will have p_{PD} equal to one.

The fourth parameter, p_{R}, measures the absolute spatial range of the segment, and is computed as

$$p_{\mathrm{R}} = \sqrt{(\max x - \min x)^2 + (\max y - \min y)^2} \tag{5}$$

where x and y are the x- and y-coordinates in the segment.

The four parameters are calculated for each segment, and are compared to individual thresholds resulting in one criterion for each parameter.

1) Dispersion: $p_{\mathrm{D}} < \eta_{\mathrm{D}}$
2) Consistent direction: $p_{\mathrm{CD}} > \eta_{\mathrm{CD}}$
3) Positional displacement: $p_{\mathrm{PD}} > \eta_{\mathrm{PD}}$
4) Spatial range: $p_{\mathrm{R}} > \eta_{\mathrm{maxFix}}$

D. Fixation and smooth pursuit movement classification

The segments are divided into three different categories, depending on how many criteria the parameters are satisfying. All segments where none of the criteria are satisfied, are classified as fixations, all segments with all criteria satisfied are classified as smooth pursuit movements, and all segments where 1-3 criteria are satisfied are placed in a third category, containing uncertain segments. Consecutive segments belonging to the same category are grouped together.

The segments in the third category are segments which do not have properties that are only typical for fixations or only typical for smooth pursuit movements. The next step is to re-categorize these segments into either the fixation or smooth pursuit movement category, depending on which category the segments are most similar to, and which category the neighboring segments belong to.

The categories of the other segments in the interval gives information about if the uncertain segment is part of a larger fixational interval or a larger smooth pursuit interval. First, the uncertain segment is investigated, by evaluating the result of criterion 3). If criterion 3) is satisfied, the uncertain segment is most similar to a smooth pursuit movement and the spatial range, p_{R}, is recalculated as follows. The spatial range of the other segments in the interval that are classified as smooth pursuit movements and has a mean direction that does not differ more than ϕ to the mean direction of the uncertain segment, are added to the spatial range of the segment. If the merged segment has a $p_{\mathrm{R}} > \eta_{\mathrm{minSmp}}$, the segment is classified as a smooth pursuit and otherwise as a fixation. If, on the other hand, the segment is most similar to a fixation, i.e., criterion 3) is not satisfied, it is criterion 4) that will decided if the segment is classified as a fixation.

TABLE I

PARAMETER SETTINGS FOR THE PROPOSED ALGORITHM

Parameter	Value	Description
t_{w}	22 ms	window size
t_{o}	6 ms	overlap of the windows
η_{P}	0.01	significance level for the Rayleigh test
η_{D}	0.25	threshold for p_{D}
η_{CD}	0.8	threshold for p_{CD}
η_{PD}	0.30	threshold for p_{PD}
η_{maxFix}	$4.8°$	threshold for max spatial range for a fixation
η_{minSmp}	$1.2°$	threshold for min spatial range for a smooth pursuit movement
ϕ	$\frac{\pi}{4}$	max difference in mean direction for a smooth pursuit movement
$t_{\min}$	40 ms	minimum fixation duration

TABLE II

PERCENTAGE OF FIXATIONS AND SMOOTH PURSUIT MOVEMENTS IN THE INTERSACCADIC INTERVALS, FOR IMAGE AND MOVING DOT STIMULI, FOR PROPOSED ALGORITHM (PA) AND I-VDT ALGORITHM.

	Image		Moving dot	
Measure	PA	I-VDT	PA	I-VDT
% Fixations	94.3	80.4	13.3	32.0
% Smooth pursuit	5.67	19.6	86.7	68.0

If criterion 4) is not satisfied, the segment is classified as a fixation otherwise as a smooth pursuit movement.

III. EXPERIMENT AND DATABASE

The eye-tracking signals used in this paper were collected during an experiment described in [13], where a tower mounted eye-tracker with sampling frequency 500 Hz, from SensoMotoric Instruments (Teltow, Germany) was used. The experiment was designed specifically for evaluation of event detection algorithms when smooth pursuit movements are present. The experiment includes static images and short video clips as well as artificial stimuli, e.g., dots moving in different directions and speeds. In the present paper, eye movements of 14 participants recorded during image and moving dot stimuli are evaluated.

IV. RESULTS

All results presented in this section are generated using the settings shown in Table I. The performance of the algorithm is evaluated by computing the percentage of time in each type of event for the two different types of stimuli, images and moving dots. For the image stimuli, it is expected to have as close to 100% detected fixations as possible and in the moving dot stimuli the expectation is to have an as large amount of detected smooth pursuit movements as possible. Since it takes some time for the eye to start to move after the dot started to move, the first interval, before the first saccade, is removed from all moving dot recordings when calculating the percentage of time in smooth pursuit and fixations for moving dot stimuli. The results for the proposed algorithm and the I-VDT algorithm are presented in Table II, where the proposed algorithm detects 94.3% of the total number of samples as fixations in the image stimuli and 86.7% of the total number of samples as smooth pursuit in the moving dot stimuli. This can be compared to the results of the I-VDT algorithm with 80.4% and 68.0%, respectively. A more detailed description of the performance of the proposed algorithm is shown in Figs. 1 – 2, where the performance for the images and moving dots are shown, separately.

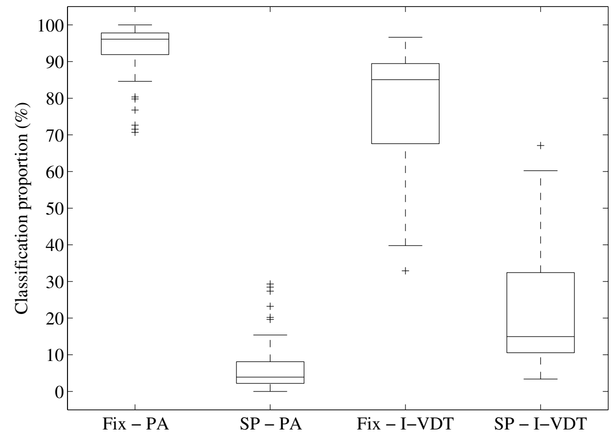

Fig. 1. Box plots of the classification proportion for image stimuli, for the proposed algorithm (PA) and I-VDT algorithm, for fixations (Fix) and smooth pursuit movements (SP). In the box, the middle mark is the median together with the 25th percentile and 75th percentile on the edges. Outliers are marked with (+).

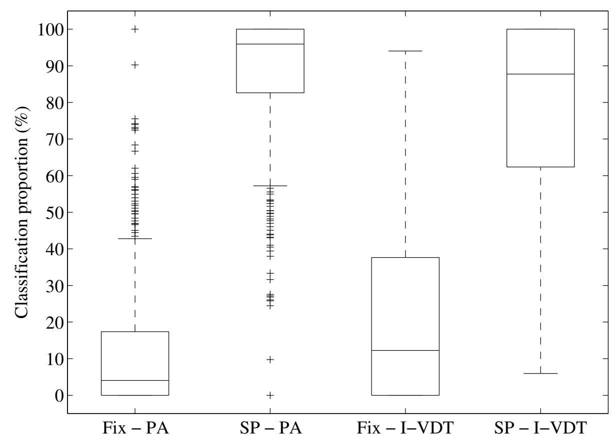

Fig. 2. Box plots of the classification proportion for moving dot stimuli, for the proposed algorithm (PA) and I-VDT algorithm, for fixations (Fix) and smooth pursuit movements (SP). In the box, the middle mark is the median together with the 25th percentile and 75th percentile on the edges. Outliers are marked with (+).

V. DISCUSSION

An algorithm for separation of fixations and smooth pursuit movements was developed. The algorithm has been evaluated using a database containing data recorded during both static and dynamic stimuli and the performance has been evaluated by comparing the detections of the algorithm to the detections of the I-VDT algorithm described in [9]. When comparing the results of the two algorithms, the proposed algorithm performs considerably better than the I-VDT algorithm. The proposed algorithm has a better discrimination performance and lower spread in classification proportions, for both image and moving dot stimuli, see Figs. 1– 2.

By using the four spatial features combined with information about the neighboring segments, the proposed algorithm takes more information into account when classifying the signal, compared to the I-VDT algorithm that only uses a dispersion threshold to distinguish between the fixations and the smooth pursuit movements. Future work will investigate if an algorithm that combines the four features, and optimize them in a four dimensional space will increase the performance of the algorithm.

The evaluation of the algorithms was divided into two groups, image stimuli and moving dot stimuli. By dividing the data into these two groups, the performance of the algorithm is evaluated by assuming that all intersaccadic intervals for the image stimuli are fixations and a majority of intersaccadic intervals for the moving dot stimuli are smooth pursuit movements. This is not always true, and especially not for the moving dot stimuli, where it is possible that the eye is still even though the dot is moving on the screen, e.g., in the beginning of the movement. By removing the first interval for each new dot that appeared, the percentage in smooth pursuit movements better reflects the actual movement of the eye in relation to the dot.

For the proposed algorithm and image stimuli, there were 5.67% samples detected as smooth pursuit movements in the intersaccadic intervals. These detections are caused by drift during fixations, small movements of the head, and remainders from the saccade and postsaccadic oscillation detection which are not distinguishable from smooth pursuit movements when only analyzing data from one eye. The 13.3% of detected fixations during moving dot stimuli could reflect that the eye does not move all the time, even though the dot is moving. Thus, it is expected with a certain amount of detected fixations when viewing moving objects.

There are several reasons to why separation between fixations and smooth pursuit movement is useful and important. Today, the separation has been mainly used in human-computer interaction using low speed eye-trackers, e.g., to stabilize the cursor during gaze control of a computer screen [16], and in interaction with information screens [17]. By having the possibility to separate between the two event types also for high-speed eye-trackers when dynamic stimuli are shown, is paving the way for studies where the properties of the two events can be investigated and compared. Two examples of such applications are, e.g., the difference between expert and novices in smooth pursuit characteristics when watching dynamic stimuli [18] and the amount of smooth pursuit when viewing natural stimuli as a diagnostic tool for neural disorders [19].

VI. Conclusions

Separation between fixations and smooth pursuit movements is a difficult task, since many of the signal characteristics of the two event types are similar. In this work, a novel algorithm for discrimination between fixations and smooth pursuit movements in high-speed eye-tracking data is developed and compared to an existing algorithm. The proposed algorithm performs considerably better than the compared algorithm for detection of the two events.

Acknowledgment

This work was supported by the Strategic Research Project eSSENCE, founded by the Swedish Research Council.

References

[1] J. Henderson and A. Hollingworth, "High-level scene perception," *Annu. Rev. Psychol.*, vol. 50, pp. 243 – 271, 1999.

[2] R. Leigh and D. Zee, *The Neurology of Eye Movements*. Oxford University Press, 2006.

[3] K. Rayner, K. Chace, T. Slattery, and J. Ashby, "Eye movements as reflections of comprehension processes in reading," *Sci. stud. read.*, vol. 10, no. 3, pp. 241–255, 2009.

[4] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, 2011.

[5] M. Dorr, T. Martinez, K. Gegenfurtner, and E. Barth, "Variability of eye movements when viewing dynamic natural scenes," *J. vision*, vol. 10, no. 28, pp. 1–17, 2010.

[6] E. Kowler, "Eye movements: The past 25 years," *Vision Res*, vol. 51, pp. 1457–1483, 2011.

[7] P. Mital, T. Smith, R. Hill, and J. Henderson, "Clustering of gaze during dynamic scene viewing is predicted by motion," *Cognitive computation*, vol. 3, no. 1, pp. 5–24, 2010.

[8] M. Vidal, A. Bulling, and H. Gellersen, "Analysing eog signal features for the discrimination of eye movements with wearable devices," in *Proceedings of the 1st international workshop on pervasive eye tracking & mobile eye-based interaction*. ACM, 2011, pp. 15–20.

[9] O. Komogortsev and A. Karpov, "Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades," *Behav. Res. Methods*, vol. 45, no. 1, pp. 203–215, 2013.

[10] D. J. Berg, S. E. Boehnke, R. A. Marino, D. P. Munoz, and L. Itti, "Free viewing of dynamic stimuli by humans and monkeys," *J. vision*, vol. 9, no. 5, pp. 1–15, 2009.

[11] M. Vidal, A. Bulling, and H. Gellersen, "Detection of smooth pursuits eye movement using shape features," in *ETRA '12 Proceedings of the Symposium on Eye Tracking Research and Applications*. New York: ACM, 2012, pp. 177–180.

[12] E. Tafaj, T. Kübler, G. Kasneci, W. Rosenstiel, and M. Bogdan, "Online classification of eye tracking data for automated analysis of traffic hazard perception," in *Artificial Neural Networks and Machine Learning – ICANN 2013*. Springer, 2013, pp. 442–450.

[13] L. Larsson, M. Nyström, and M. Stridh, "Detection of saccades and post-saccadic oscillations in the presence of smooth pursuit," *IEEE Trans. on Biomed. Eng.*, vol. 60, no. 9, pp. 2484–2493, 2013.

[14] C. Meyer, A. Lasker, and D. Robinson, "The upper limit of human smooth pursuit." *Vision Res.*, vol. 25, no. 4, pp. 561–563, 1985.

[15] P. Berens, "Circstat: A matlab toolbox for circular statistics," *J. Stat. Softw.*, vol. 31, no. 10, pp. 1–21, 2009.

[16] J. S. Agustin, "Off-the-shelf gaze interaction," Ph.D. dissertation, The IT University of Copenhagen, 2009.

[17] M. Vidal, A. Bulling, and H. Gellersen, "Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 2013, pp. 439–448.

[18] H. Jarodzka, K. Scheiter, P. Gerjets, and T. van Gog, "In the eyes of the beholder: How experts and novices interpret dynamic stimuli," *Learn. Instr.*, vol. 20, no. 2, pp. 146–154, 2010.

[19] M. Vidal, J. Turner, A. Bulling, and H. Gellersen, "Wearable eye tracking for mental health monitoring," *Comput. Commun.*, vol. 35, no. 11, pp. 1306–1311, 2012.