

# Measuring Focused Attention Using Fixation Inner-Density

Wen Liu, Mina Shojaeizadeh, Soussan Djamasbi, Andrew C. Trapp

User Experience & Decision Making Research Laboratory, Worcester Polytechnic Institute  
100 Institute Road, Worcester, MA 01609  
{wliu3, minashojaei, djamasbi, atrapp}@wpi.edu

**Abstract.** Examining user reactions via the unobtrusive method of eye tracking is becoming increasingly popular in user experience studies. A major focus of this type of research is accurately capturing user attention to stimuli, which is typically established by translating raw eye movement signals into fixations, that is, ocular events characterized by relatively stable gaze over a specific stimulus. Grounded in the argument that inner-density of gaze points within a fixation represents focused attention, a recent study has developed the fixation-inner-density (FID) methodology, which identifies fixations based on the compactness of individual gaze points. In this study we compare the FID filter with a widely used method of fixation identification, namely the I-VT filter. To do so we use a set of measures that investigate the distribution of gaze points at a micro-level, that is, the patterns of individual gaze points within each fixation. Our results show that in general fixations identified by the FID filter are significantly denser and more compact around their fixation center. They are also more likely to have randomly distributed gaze points within the square box that spatially bounds a fixation. Our results also show that fixation duration is significantly different between the two methods. Because fixation is a major unit of analysis in behavioral studies and fixation duration is a major representation of the intensity of attention, awareness, and effort, our results suggest that the FID filter is likely to increase the sensitivity of such eye tracking investigations into behavior.

**Keywords:** Eye Tracking, Fixation Identification, Fixation-Inner-Density, Fixation Micro-Patterns

## 1 Introduction

The study of eye movements in user experience research is becoming increasingly popular because eye tracking technology enables capturing the focus of a person's gaze on a visual display at any given time. Human gaze serves as a reliable indicator of attention because it represents effort in maintaining the eyes relatively steady to take foveal snapshots of an object for subsequent processing by the brain [1]. Hence, ex-

tracting relatively stable gaze points that are near in both spatial and temporal proximity, that is translating the raw gaze data into fixations, is essential in many eye tracking studies [2, 3]. One primary method for identifying fixations in a stream of raw eye movement data is the Velocity-Threshold Identification (I-VT) algorithm. The I-VT filter uses a fixed velocity threshold to identify whether individual gaze points qualify as a fixation point, or a saccade point.

Because a fixation is the collection of gaze points that are near to one another in both time and proximity, a denser collection of gaze points within a fixation represents higher level of focused attention, and thus higher level of cognitive processing [4]. Thus, a recent study [5] proposes a new way to group gaze points into fixations based on their inner-density property. Similar to the I-VT filter, this new Fixation Inner-density (FID) filter first uses a velocity threshold to identify a candidate set of gaze points that are slow enough to form a fixation. It then uses optimization-based techniques to identify a densest fixation of gaze points among all candidate points. Identifying fixations using the FID filter naturally eliminates those gaze points that are near to tolerance settings. *How* gaze points are dispersed in a fixation affects fixation metrics such as the duration and center location, and there is evidence that the FID filter reduces the possibility of skewing these metrics [5].

In this paper we translate raw gaze data into fixation using the I-VT and FID filters. We demonstrate that fixations processed by the FID filter are superior in terms of three key fixation *micro-patterns* than those that are processed by the I-VT filter. First, they are *denser*. Second, the extent to which points are dispersed within a fixation is *smaller*. Third, the points within a fixation are more likely to be uniformly distributed. This investigation is important because the compactness and the patterns of distribution of gaze points can directly affect fixation metrics, such as fixation duration and fixation center position, that are commonly used in eye-tracking studies to assess viewing behavior. This study is the first to investigate such fixation *micro-patterns* or properties of the distribution of gaze points within an individual fixation.

## 2 Background

Raw gaze data is a sequence of  $(x, y, t)$  triplets, where  $(x, y)$  represents the measured location of user gaze, and  $t$  is the time stamp. Common sampling rates in eye trackers range from 30 Hz to 1,000 Hz, and a gaze sequence can easily contain tens of thousands of triplets. Gaze data is often categorized into two common types: fixations and saccades. Fixations are pauses over informative regions during eye movement; in gaze data, a fixation is where gaze point triplets aggregate together. Fixation identification methods cluster those intensive gaze points into fixations to present focused attention and cognitive effort in eye tracking research [4].

One popular fixation identification algorithm is the I-VT filter. It identifies fixations by gaze point *velocity*. If the velocity exceeds the predefined threshold  $V$ , the corresponding gaze point is identified as a saccade, otherwise it is categorized as a fixation point. I-VT filter is efficient and practical; however, it has the drawback of ignoring the information about the spatial arrangement of individual gaze points within a distinct fixation. Some fixation metrics can express the distribution of points within a fixation.

One such metric is fixation inner-density, which was introduced by [4] and further refined in [5]. Fixation inner-density represents user focus, and [4] has validated that fixation inner-density is correlated with normalized fixation duration and average pupil dilation variation during fixation. The FID filter uses optimization-based techniques to optimize for inner-density, which means that it selects a set of candidate gaze points that guarantees there is no better set with respect to the objective function of maximizing fixation inner-density. Fixation inner-density improves upon previous fixation identification methods because it combines both the temporal and the spatial aspects of the fixation into a single metric that evaluates the compactness of a fixation.

As the problem of fixation identification is a type of time-series clustering, it shares the commonality that interpreting clustering results is somewhat subjective in nature. Hence, the choice of an appropriate metric will directly affect the formation of the clusters. While density and dispersion properties can be measured in various ways, they are inherently *positively related* to the number of gaze points in a fixation, and *negatively related* to the area occupied by the constituent points. We next discuss some important metrics to evaluate density and dispersion properties within fixations.

### 3 Methodology

We consider two representative ways of measuring fixation inner-density, both of which are advocated in [5]. Suppose a fixation identification algorithm locates fixations in a gaze data sequence with  $T$  gaze points. For any given fixation  $f$ , let  $n_f$  denotes the count of points inside  $f$ , and let  $i, j$  be any two points in  $f$ . We denote the Euclidean distance between  $i$  and  $j$  as  $d_{ij}$ , the minimum area box that spatially bounds the fixation as  $A_{sq}$ , and the minimum area rectangle box that spatially bounds the fixation as  $A_{rt}$ . The first density metric ( $D_1$ ) is the average pairwise distance between points within a fixation.

$$D_1 = \frac{\sum_{i=1}^{n_f} \sum_{j=i+1}^{n_f} d_{ij}}{\binom{n_f}{2}}. \quad (1)$$

The second density metric ( $D_2$ ) is the minimum area square bounding box surrounding the fixation divided by the number of fixation points it contains:

$$D_2 = \frac{A_{sq}}{n_f}. \quad (2)$$

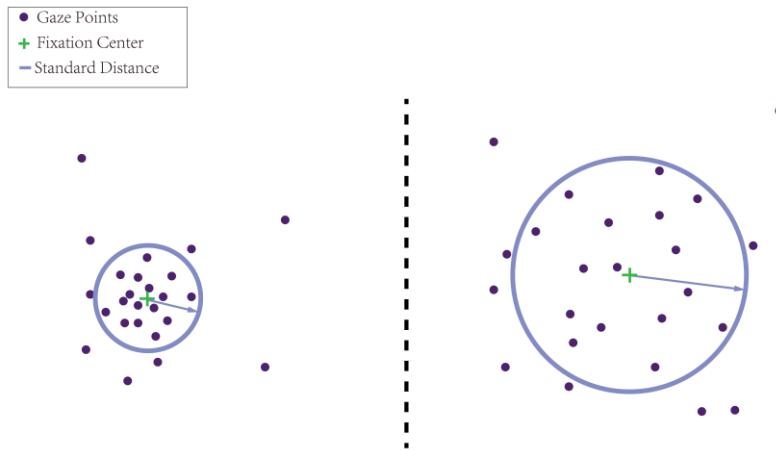
For both the  $D_1$  and  $D_2$  density metrics, small values imply greater density. A third metric, Standard Distance ( $SD$ ), measures the dispersion of gaze points around the fixation center.  $SD$  is a common metric in the Geographic Information System (GIS) literature, that evaluates how points are distributed around the fixation center [6]. Similar to standard deviation,  $SD$  quantifies the dispersion of a set of data values. Hence, the  $SD$  score is a summary statistic representing the compactness of point distribution. Smaller  $SD$  values correspond to gaze points that are more concentrated around the center ( $\bar{X}_f, \bar{Y}_f$ ) of fixation  $f$ , expressed as:

$$\bar{X}_f = \frac{\sum_{i=1}^{n_f} x_i}{n_f}, \bar{Y}_f = \frac{\sum_{i=1}^{n_f} y_i}{n_f}. \quad (3)$$

The standard distance of fixation  $f$ ,  $SD_f$ , is:

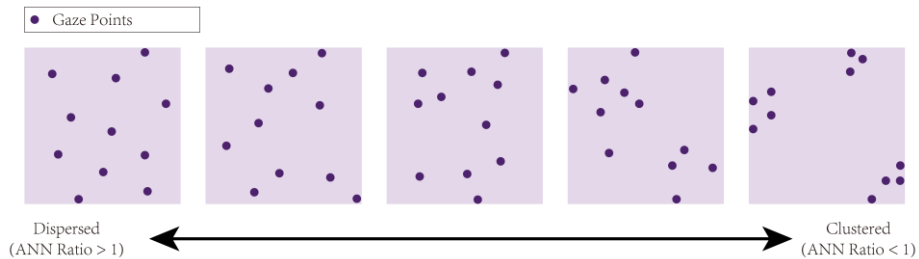
$$SD_f = \sqrt{\frac{\sum_{i=1}^{n_f} (x_i - \bar{X}_f)^2}{n_f} + \frac{\sum_{i=1}^{n_f} (y_i - \bar{Y}_f)^2}{n_f}}. \quad (4)$$

Figure 1. An illustrative depiction of standard distance,  $SD$ . When considering an identical number of gaze points,  $SD$  is smaller when points are more compactly distributed around the center (left); when they are more dispersed,  $SD$  becomes larger (right).



Spatial pattern analysis can also be examined in measuring the fixation gaze point distribution pattern. The Average Nearest Neighbor ( $ANN$ ) [6] is used to measure the degree to which fixation gaze points are *clustered*, versus *randomly distributed*, within a fixation bounding area. A fixation resulting from focused gaze toward a single area of interest would tend to exhibit a more uniformly distributed pattern, with greater  $ANN$  values. The  $ANN$  ratio is calculated as the average distance between each point and its nearest neighbor, divided by the expected average distance between points if a random pattern is assumed.  $ANN$  values greater than one imply that the fixation gaze points are *dispersed*; as this ratio decreases, fixation gaze points increasingly exhibit clustering.

Figure 2. Illustrating the  $ANN$  ratio as the distribution of gaze points change within an identical minimum square bounding box.



The four metrics  $D_1$ ,  $D_2$ ,  $SD$  and  $ANN$  will be used to evaluate three aspects of inner fixation patterns: fixation inner-density, fixation points dispersion, and their distribution. We expect fixations identified with the FID filter to be denser and more uniformly distributed than those identified with the I-VT filter. Our density assertion, which stems from the method of fixation identification, helps to test whether the FID filter does indeed more accurately group individual gaze points into focused attention. Our assertion that gaze points identified with the FID filter are more randomly distributed stems from the argument that if a fixation is compact, that is it has high inner-density, it is more likely to have a more uniform distribution around its center.

In addition to the above assertions, we also examine the impact of FID and I-VT filters on fixation duration and center location.

## 4 Experimental Evaluation

We begin this section by describing the specific context of our eye tracking datasets and experiments. We then compare the I-VT and FID filters with the aforementioned four metrics, and discuss our findings.

### 4.1 Dataset and Equipment

We perform our experiments on eye movement datasets obtained from a total of 28 university students who were assigned to read a text passage shown on a standard desktop computer monitor. Prior to the experiment, each participant completed a brief eye-calibration process lasting less than one minute. We used the Tobii X300 eye tracker [7] to collect participant’s eye-movements. The software version is 3.2.3 and the sampling rate was set to 300 Hz.

The 28 recordings were further analyzed using an Intel core i7-6700MQ computer with 3.40 GHz and 16.0 GB RAM running 64-bit Windows 10. Matlab 2016a and Python 2.7 were used for additional data analysis and processing.

### 4.2 Data Processing

For each eye tracking record, we used the Tobii Studio I-VT filter [8] to generate I-VT fixation identification results. The velocity threshold  $V$  was set to  $30^\circ/s$ , which is the recommended threshold in [8]. The minimum fixation duration is set to 100ms which is the theoretical minimum fixation duration suggested by other eye tracking studies [9, 10].

We further used the results of the I-VT fixation identification as the input data chunks for the mixed integer programming formulation (MIP) for minimizing square area of fixations from [5]. The Gurobi Optimizer 7.5.1 [11] is used as the solver. The FID filter is parametrized by a manually assigned constant  $\alpha$  that enables decision-makers to have fine-tuned control over the density. We varied  $\alpha$  from 0 to 1 by steps of 0.1 on one randomly selected eye tracking record and examined the fixation identification results manually. When  $\alpha = 0.1$ , the clustering result appeared the most reasonable, and averaging  $D_2$  values over all fixations yielded the smallest value, suggesting the

algorithm finds the (averaged) densest fixations at  $\alpha = 0.1$  comparing to other  $\alpha$  levels. Therefore, we set  $\alpha = 0.1$  when running the FID filter on the other 27 records. In the following evaluations, we discard the record used for selecting  $\alpha$  to avoid data snooping.

### 4.3 Experimental Results

After discarding the single record above, in this section we first report our statistical analyses from the point of view of a single record. Subsequently, we expand it to all 27 of the (remaining) records in our dataset.

### 4.4 Comparing I-VT and FID Filters for A Single Record

Fixation inner-density and the distribution of gaze points within an individual fixation are micro-patterns in gaze data. Such patterns are relatively difficult to evaluate by averaging over all eye tracking records. To more thoroughly investigate micro-patterns, we first illustrate the comparison results on the eye tracking record of one randomly selected participant. Toward the end of this section, the comparison summary over all recordings is also included.

For this gaze data record, there are 9,788 gaze points and 110 fixations. We calculated fixation inner-density metrics  $D_1$  and  $D_2$  on each individual fixation. The resulting average of both  $D_1$  and  $D_2$  from the I-VT filter is larger than that of FID, which indicates that fixations from the FID filter are denser than those in I-VT filter result. We performed a paired  $t$ -test with the following hypothesis:

$$\begin{aligned} H_0: \bar{D}_{I-VT} &= \bar{D}_{FID}, \\ H_a: \bar{D}_{I-VT} &> \bar{D}_{FID}. \end{aligned}$$

The  $t$ -test on both  $D_1$  and  $D_2$  returns a  $p$ -value smaller than 0.05, so at a 95% confidence level we reject  $H_0$ , which implies  $\bar{D}_{I-VT}$  is statistically larger than  $\bar{D}_{FID}$ .

Table 1. Comparison of fixation density for I-VT and FID filters.

Fixation Density	I-VT		FID ( $\alpha = 0.1$ )		t-test	
	Mean (pixel)	STD (pixel)	Mean (pixel)	STD (pixel)	$p$ -value	Result
$D_1$	6.769	2.382	5.994	1.961	<0.0001	Reject
$D_2$	7.690	10.450	5.112	3.920	0.0025	Reject

The  $SD$  metric measures the dispersion of fixation points around their center. Table 2 reveals that the  $SD$  mean and standard deviation for the I-VT filter are larger than that of the FID filter. We also performed a paired  $t$ -test when comparing the  $SD$  metric. The hypotheses are:

$$\begin{aligned} H_0: \overline{SD}_{I-VT} &= \overline{SD}_{FID}, \\ H_a: \overline{SD}_{I-VT} &> \overline{SD}_{FID}. \end{aligned}$$

With the same 95% confidence level as the previous test, the  $t$ -test result rejects the  $H_0$ . It indicates that the FID filter tends to identify fixations having points that are more

dispersed around the center. It further demonstrates that identifying fixations by optimizing for fixation inner-density yields fixations with more compact regions.

Table 2. Comparison of  $SD$  for I-VT and FID filters.

	I-VT		FID ( $\alpha = 0.1$ )		t-test	
	Mean (pixel)	STD (pixel)	Mean (pixel)	STD (pixel)	$p$ -value	Result
<b><math>SD</math></b>	5.5033	2.189	4.746	1.616	<0.0001	Reject

Finally, we perform a hypothesis test using the  $ANN$  ratio [6] to see if the gaze points are randomly distributed in a fixation region:

$$H_0: \text{gaze points are randomly distributed within fixation region,}$$

$$H_a: \text{gaze points are not randomly distributed within fixation region.}$$

If the hypothesis test results in a small  $p$ -value, we would reject the  $H_0$  because of the small probability that the fixation gaze points are randomly distributed in their fixation region.

The  $ANN$  hypothesis test is rather sensitive with respect to the bounding region used to cover all fixation points in an individual fixation. Therefore, we perform two experimental results using  $A_{sq}$  and  $A_{rt}$ , respectively, to represent fixation area. Table 3 reports the count of fixations (out of 110) for which  $H_0$  is rejected at 95% confidence level, implying that there is statistical evidence that fixation points are not randomly distributed. Table 3 reveals that, under both fixation regions, more fixations appear to not be randomly distributed when using the I-VT filter. Moreover, the difference between the I-VT and FID filters is greater under the  $A_{sq}$  region. This may be due to  $A_{sq}$  typically being larger than  $A_{rt}$ , as the FID filter specifically minimizes the square area of fixations.

Table 3. Comparison of  $ANN$  for I-VT and FID filters, reporting the count of fixations (out of 110) for which  $H_0$  is rejected.

	I-VT		FID ( $\alpha = 0.1$ )	
	$A_{sq}$	$A_{rt}$	$A_{sq}$	$A_{rt}$
<b># of Fixations Rejecting <math>H_0</math></b>	95	60	61	50

We now compare fixation duration and fixation center for the I-VT and FID filters. Fixation duration ( $FD$ ) is a commonly used metric in eye tracking research. We compare the average fixation duration on I-VT and FID filters with the hypotheses that

$$H_0: \overline{FD}_{I-VT} = \overline{FD}_{FID},$$

$$H_a: \overline{FD}_{I-VT} > \overline{FD}_{FID}.$$

The paired  $t$ -test result shows that  $\overline{FD}_{FID}$  is significantly smaller than  $\overline{FD}_{I-VT}$  at a 95% confidence level. This outcome may be due to the FID filter eliminating fixation points and refining the fixation region of each of the fixation chunks from the I-VT filter.

Table 4. Comparison of fixation duration for I-VT and FID filters.

	I-VT		FID ( $\alpha = 0.1$ )		t-test	
	Mean (second)	STD (second)	Mean (second)	STD (second)	p-value	Result
<b>Fixation Duration</b>	0.250	0.151	0.204	0.167	<0.0001	Reject

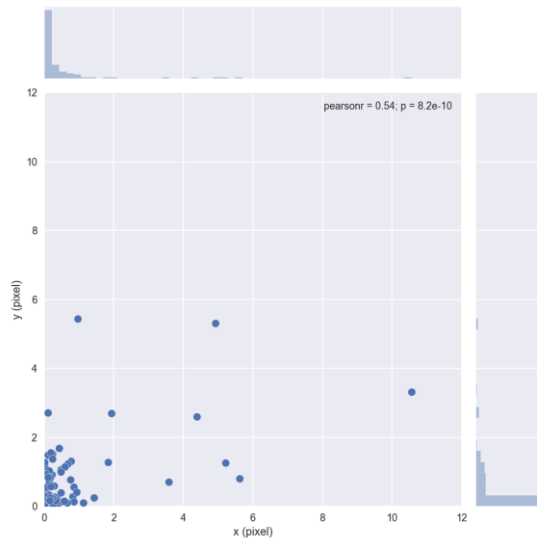
Fixation center is also a basic feature to represent fixation location, used in the depiction the scan path of eye movement. We introduce the center shift, which is the Euclidean distance between the fixation center of the I-VT filter and that of the FID filter. The 110 fixations within the eye tracking record generates mean and standard deviation (STD) of the center shift data as reported in Table 5.

Table 5. Statistics of fixation center shift between I-VT and FID filter.

	Mean (pixel)	STD (pixel)
<b>Center Shift</b>	0.881	1.617

When examining the mean and STD of center shift, it may be inferred that the difference of fixation center is negligible. The bivariate distribution of center shift depicted in Figure 3 displays the long tail distribution in both x and y axis. The 90% quantile of x, y is 0.922 and 1.308 respectively. It shows that while the refined results of the FID filter can skew some I-VT fixation centers, most of the time the center shift remains in a fairly small range.

Figure 3. The bivariate distribution of center shift in x, y coordinates.





#### 4.5 Comparing I-VT and FID Filters for All 27 Remaining Records

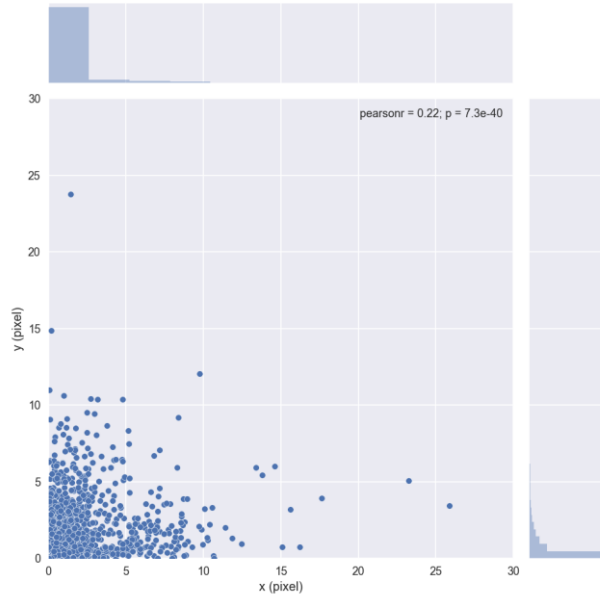
The results reported above were for a single eye tracking record. The average number of gaze points for all remaining 27 records is 10,959, and the average number of fixations is 127.7. Table 6 reports the results of the corresponding hypothesis tests for  $D_1$ ,  $D_2$ ,  $SD$  and fixation duration on all the 27 eye tracking records. We find that zero record does not reject the corresponding  $H_0$  in the  $t$ -test for  $D_1$ ,  $SD$  and fixation duration, and two for  $D_2$ . This analysis shows that the FID filter finds denser and more compact fixations than I-VT filter holds for most of eye tracking records in our dataset in terms of for  $D_1$ ,  $D_2$  and  $SD$ .

Table 6. Summary of hypothesis test results for 27 eye tracking records.

	$D_1$	$D_2$	$SD$	Fixation Duration
# of Records That Do Not Reject $H_0$	0	2	0	0

We calculate the center shift between all I-VT and FID filter fixation pairs; the bivariate distribution result is shown in Figure 4. The distribution on either  $x$  or  $y$  direction is again a long tail distribution. The 90% quantile value of  $x$ ,  $y$  is 2.095 and 2.411 respectively. Figure 4 shows only a few points that are far away from the origin, indicating that the FID filter identification results can indeed change the fixation center location, though this occurred relatively infrequently in our dataset.

Figure 4. The bivariate distribution of center shift for all fixations.



We also run the *ANN* hypothesis test on each recording and calculate the count of fixations (*FC*) for which the *ANN* hypothesis test  $H_0$  ( $FC - ANN$ ) is rejected over all recordings. The average is reported in Table 7. Both the mean and the standard deviation resulting from the FID filter are smaller than that of the I-VT filter.

We compare the *FC* results from the I-VT and FID filters by the paired *t*-test with 95% confidence level and the following hypotheses:

$$H_0: \overline{FC}_{I-VT} = \overline{FC}_{FID},$$

$$H_a: \overline{FC}_{I-VT} > \overline{FC}_{FID}.$$

The first row in Table 7 shows that when bounding the fixation region by  $A_{sq}$ ,  $\overline{FC}_{FID}$  is significantly smaller than  $\overline{FC}_{I-VT}$ . It indicates the general trend that the inner gaze points of fixations resulting from the FID filter tend to be randomly distributed. As for  $A_{rt}$ , the *t*-test result also reject  $H_0$ , implying that the same conclusion could be drawn on  $A_{rt}$ .

Table 7. Comparison of *FC - ANN* for I-VT and FID filters over all recordings.

Fixation Region	I-VT		FID ( $\alpha = 0.1$ )		<i>t</i> -test	
	Mean (count)	STD (count)	Mean (count)	STD (count)	<i>p</i> -value	Result
$A_{sq}$	109.1	40.0	70.2	27.4	<0.0001	Reject
$A_{rt}$	74.5	30.1	64.2	24.6	0.0002	Reject

## 5 Conclusions

Our results show that the FID filter, as compared to I-VT filter, does indeed identify fixations that are denser and more compact around the center, and more uniformly distributed patterns found in fixation bounding regions. These properties have major implications for two important fixation metrics that are widely used in eye tracking analysis: Fixation duration and location. Our results show that the two filters tend to result in significantly different fixation durations. The results displayed in Figure 3 and Figure 4 provide evidence that in some cases FID filter can result in quite different fixation centers comparing to I-VT filter. It is important to note that the data used in our study was gathered when users were reading an online text passage, which typically generates more focused fixations. Future investigation using different stimuli are needed to extend the generalizability of these results and to see whether the micro-level differences, including fixation duration and center location, observed in this study between FID and I-VT filters change for different tasks (e.g., reading more challenging text passages, viewing a picture, or browsing a website). For example, in this study we used a reading task which typically results in compact fixations. Using a browsing task may result in much larger differences in fixation center location, because gaze points within fixations in browsing tasks tend to more dispersed [5]. The metrics introduced in this study to compare fixations at a micro level serve to refine the analysis of eye movements to a deeper level. Future studies, however, are needed to validate and extend our findings.

The results of this study contribute in two ways to eye tracking studies that examine user behavior. First, they show that researchers can identify focused attention with the

FID filter and thereby improve the sensitivity of their analysis with regard to duration and center location of intense attention. Second, the micro-analysis introduced in this study provides a new way to compare gaze points within a fixation. This is important because it allows researchers to examine relationships between eye movements and behavior at a much smaller unit of analysis, namely fixation micro-patterns.

## 6 References

- [1] Djamasbi, S. (2014). Eye tracking and web experience. *AIS Transactions on Human-Computer Interaction*, 6(2), 37-54.
- [2] Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods*, 42(1), 188-204.
- [3] Salvucci, D. D., & Goldberg, J. H. (2000, November). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 71-78). ACM.
- [4] Shojaeizadeh, M., Djamasbi, S., & Trapp, A. C. (2016, July). Density of Gaze Points Within a Fixation and Information Processing Behavior. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 465-471). Springer International Publishing.
- [5] Trapp, A. C., Liu, W., & Djamasbi, S. (2017). New Density-based Optimization Formulations and Algorithms to Identify Fixations in Gaze Data. Presented in *INFORMS Annual Meeting, Houston, TX*. INFORMS, Hanover, MD.
- [6] Mitchell, A. (2005). The ESRI Guide to GIS analysis, Volume 2: Spatial measurements and statistics. *ESRI Guide to GIS analysis*.
- [7] Tobii. (2017). Tobii technology. <http://www.tobii.com>. Accessed: 2017-12-18.
- [8] Olsen, A. (2012). The tobii i-vt fixation filter. *Tobii Technology*.
- [9] Blygnaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, 71(4), 881-895.
- [10] Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., & Gowda, S. M. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering*, 57(11), 2635-2645.
- [11] Gurobi Optimization, Inc. (2017). Gurobi Optimizer 7.5.1 Reference Manual.