This is a pre-print.

Person Verification via Eye Movement-driven Text Reading Model

Evgeniy R. Abdulin, Oleg V. Komogortsev
Texas State University – San Marcos
601 University Drive, San Marcos TX 78666
{e_a146, ok11}@txstate.edu

Abstract

The paper presents a reading-based eye movement biometrics model. The model is able to process passages of text and extract metrics that represent the physiological and behavioral aspects of the eye movements in reading. When tested on a database of eye movements from 103 individuals, the model yielded the Equal Error Rate of 10.2%. The proposed method performed better in the template-aging scenario than comparable eye movement-driven biometrics methods.

1. Introduction

Security is a very important and desired aspect of human lives [1]. Security related to personal information is especially important considering the rapid development of the Internet and related technologies [2] that potentially give an opportunity to intruders to remotely get access to the large amounts of sensitive information [3-6], especially if the passwords guarding this information are weak. The science of biometrics strives to solve a problem that is presented by passwords – the necessity to remember potentially random pieces of data for authentication. User authentication via biometric technology occurs not based on what a person remembers (e.g., password) or what a person has (e.g., credit card) but based on the traits that essentially represent a person. Remarkably most sciences studying humans such as psychology [7] and sociology [8] try to understand and measure the similarities between people while the science of biometrics strives to find traits that are different between the individuals and make a recognition decision based on those traits [9].

Traits employed for biometric authentication can be loosely separated into two groups: a) physiological and b) behavioral. The examples of the physiological traits are [9] fingerprint, iris, hand, face, ear, DNA [10], and the examples of behavioral traits are [9] voice, gait, keystroke dynamics, and signature.

Eye movements contain both the physiological and the behavioral traits and can be collected on the equipment employed for the iris recognition [11]. Thus, they can be used to address some of the spoofing vulnerabilities of the cotemporary iris systems [12]. For example, related research demonstrates the robustness of eye movement driven biometrics to iris-print attacks [13].

Previous research indicated susceptibility of the eye movement biometrics to the template aging effects [14]. This is an important problem that should be solved if the time between the enrollment and the actual authentication is more than a few days. The text-reading model presented in this paper strives to improve the resistance of the eye movement driven biometrics to the template-aging effects by analyzing the eye movements and their characteristics in the direct response to text structure during the task of reading a text. The results indicate that the proposed approach provides better authentication accuracy in the template-aging scenario than the previously developed state-of-the-art approach.

1.1. Previous work

Eye movement driven biometrics methods can be loosely separated into two categories: a) based on the raw positional data [15, 16], b) based on the fixations/saccades and their characteristics [17-21]. Fixations are defined as eye movements that keep an eye stable toward the object of interest, thus providing high acuity visual data about the object to the brain. Saccades are rapid rotations of the eye that move it from one fixation spot to the next.

The works in the second group of eye movement-driven biometrics methods can be divided into the following sub-categories: a) methods based on the physiology of the eye muscles and the eye globe derived from the saccades via mathematical modeling of the eye, e.g., the Oculomotor Plant Characteristics (OPC) approach [18]; b) methods based on the deployment and the representation of the visual attention, given the specific type of stimulus, e.g., the Fixation Density Maps (FDM) approach [19]; c) methods based on the corrective eye movements, e.g., the Complex Oculomotor Behavior (COB), [20]; and d) mixed approach where a variety of fixation and saccade related characteristics are considered, e.g., the Complex Eye Movement (CEM-B) [21] approach.

The text reading approach proposed in this work is most similar to the CEM-B approach. Analogous to the CEM-B, the proposed method creates a biometric template from distributions of characteristics of fixations and saccades, but it puts into a biometric template only fixations and saccades that are exhibited as a direct response to the text structure. In addition, the proposed approach incorporates text specific metrics, i.e., related to the specific lines of text and words.

2. Text Reading Model (TRM)

2.1. Reading Text

When a person reads a piece of text, she or he tries to extract the meaning of the presented information, moving the eyes in a sequence of fixations and saccades that occur to overcome the anatomical limitation of the human vision, i.e., eyes can see with high quality only limited portions of the text. People normally read text line by line connecting pieces of information represented by the individual words into the global meaning represented by a text as a whole. Low-level cognitive processes are responsible for extracting the information from the individual words while higher-level cognitive processes are responsible for the formation of the overall meaning encoded in the lines of text [22].

Given a substantial amount of previous research regarding how people read text [22] we wanted to create a text reading model that employs metrics related to the text structure that potentially increase the amount of information that is extracted from the eye movements and stored in a biometric template representing a person.

To be able to take advantage of the text structure and the eye movements associated with it, it becomes very important to establish one to one correspondence between individual lines of text/words and to identify specific eye movements that correspond to these structures. Figure 1a illustrates the challenges that should be solved, i.e., fixations (represented by circles) are detected for the wrong lines of text and words. There are multiple reasons for such positional inaccuracies including but not limited to the calibration biases [23] caused by the eye image distortions due to squinting, excessive eye moisture, and specific eye shape.

The next subsection describes the approaches for processing the eye movements in response to the text stimulus including the commonly used Area-of-Interest (AOI) approach and the Pass-Based approach, which we propose in this work.

2.2. Processing Eye Movements for a Text

2.2.1 AOI-based approach

The most common approach employed in classifying eye movements in response to the text stimulus is Areas-of-Interest (AOI)-based [24], which is usually done in the following stages:

1. Divide the image of the text into a set of AOIs with each AOI corresponding to one word.
2. Associate the fixations detected within a specific AOI as reading fixations for the AOI represented word.

The disadvantage of the AOI-based approach is the requirement of the larger than normal font size for the words, larger than normal between the lines intervals, and eye movement records of high quality where fixations are directly positioned on the words that they correspond to.

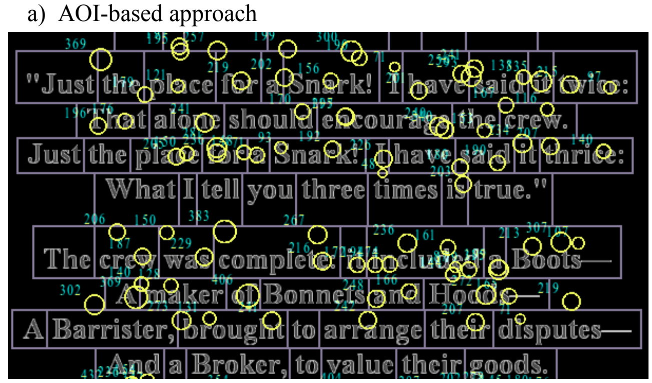
2.2.2 Pass-Based approach

However, every day reading conditions, where eye movement-driven biometric techniques could be potentially applied, frequently are presented with fonts that are not large and between line intervals that are not excessive. To cope with the issues that arrive as a result of the reading conditions that are relatively close to normal, we propose the Pass-Based approach.

Terminology

Pass is defined as a sequence of fixations and saccades that an individual makes during reading a line of a text. Passes create a robust structure that allows reliable eye movement analysis with respect to the lines of text even in the cases of smaller fonts and between the line intervals that is shown on Figure 1b where passes are presented by polylines.

a) AOI-based approach



b) Pass-based approach

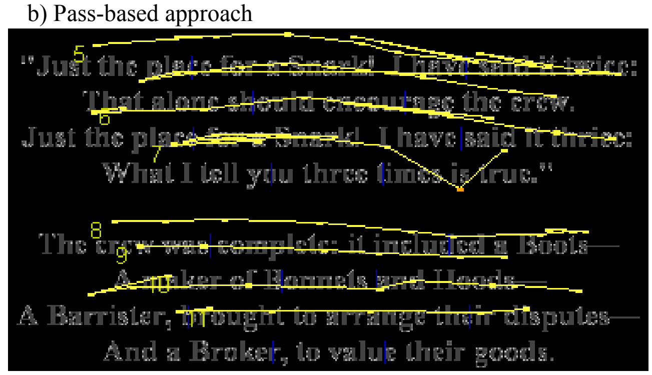


Figure 1.This is an example of the case when the AOI-based approach does not give a correct result, but the Pass-Based approach does. Both pictures are from the same record, calibration error: average 0.58 degrees, maximum 1.04 degrees. This case is common.

What is also important is that the pass structure allows processing even eye movement recordings where a vertical shift of the data is present (e.g., due the low positional accuracy) and the detected fixations are positioned on a wrong line of the text.

Local return saccade is a saccade that occurs when the eyes transition to the next line for reading. *Global return saccade* is a saccade that occurs when the eyes transition to the beginning of the page from the bottom of the page after the user is done reading a page of text. *Hop* is the movement of the eyes between two consequent fixations within a pass. Hop is usually represented by a single saccade; however it can be also a saccade toward the next word together with a subsequent corrective saccade in case the word was initially under/overshot by the eyes. There are multiple reasons for saccade dysmetria with fatigue being one of those reasons [25].

Obtaining passes

The passes are obtained by separating the eye movement record by local/global return saccades. The set of all of the fixations and saccades between two local return saccades or a local and global return saccade is a pass. Passes are associated with lines based on their chronological order with any shift from the intended position that the AOI method cannot do. Double reading passes are filtered out; they are recognized by being closer than a threshold value to each other in vertical coordinates. The threshold used in the presented work is 0.3 of the maximum vertical size of the line.

Once the pass is associated with an appropriate line of text, the fixations detected for this line are matched with the corresponding words by matching their horizontal coordinates with horizontal coordinates of a particular word in the same way as for the AOI-based approach. This means that the Pass-Based approach can suppress the vertical shift of fixations from corresponding words when the AOI-based approach cannot provide this feature. Due to words having a smaller vertical size compared to the horizontal size, the vertical shift causes fixation and word position mismatch with a higher probability than for the horizontal shift for the same level of positional accuracy.

Data filtering

Blinks and invalid eye positional samples are filtered out by a special blink filtering algorithm that is based on the blink detection algorithm employed by the EyeLink vendor [26]. The main difference from the original algorithm is the exclusion of any saccades and fixations that were involved in a blink. All eye movements that were not involved in any pass were also dropped out of the consideration of being a part of the biometric template.

2.3. TRM metrics

Following set of metrics depicting different aspects of the reading process represent the Text Reading Model (TRM) that we propose.

2.3.1 Hop-related metrics

The perceptual span of a person varies from one individual to another [22], so hops that reflects eye movements in reading from one location in the text to the next are directly affected by the size of the perceptual span.

M1. Hop Horizontal Amplitude – horizontal distance between two adjacent fixations (deg);

M2. Hop Vertical Amplitude – vertical distance between two adjacent fixations;

M3. Fixation Duration (ms);

M4. Hop Duration (ms);

M5. Hop Mean Horizontal Velocity – mean horizontal eye velocity during a hop (deg/s);

M6. Hop Mean Vertical Velocity – mean vertical eye velocity during a hop (deg/s);

M7. Hop Mean Vectorial Velocity – mean vectorial eye velocity during a hop (deg/s);

M8. Hop Peak Horizontal Velocity – peak horizontal eye velocity during a hop (deg/s);

M9. Hop Peak Vertical Velocity – mean vertical eye velocity during a hop (deg/s);

M10. Hop Peak Vectorial Velocity – mean vectorial eye velocity during a hop (deg/s);

M11. Ratio of the Fixation Duration to the duration of the antecedent Hop (ms);

M12. Ratio of the Fixation Duration to the duration of the subsequent Hop (ms);

2.3.2 Pass-related metrics

Metrics in this category aim to represent the individuality of the perceptual span, cognitive abilities related to the speed of reading, and individual features of the eye that can cause distortions and shifts after a calibration procedure. Moreover, fixation-based metrics are easier to obtain because fixations are longer than saccades – 200 ms vs. 80 ms on average, plus fixation can be reliably detected from the raw eye position data captured even from an inexpensive eye trackers.

M13. Ratio of Pass Amplitude (distance between the horizontal coordinates of left and right fixations) to the Length of Text Line corresponding to a pass;

M14. Ratio of Pass Total Length (summation of the amplitudes of all hops for a pass) to the Length of Text Line corresponding to the pass;

M15. Number of Fixations in a Pass;

M16. Total Time of a Pass (ms);

M17. Pass Offset – vertical distance between a Pass and the corresponding line of text (deg);

2.3.3 Fixation-related metrics

The potential of biometric performance of this set of metrics is based on the individuality of the perceptual span and cognitive abilities related to the speed of reading.

First fixation for a line of text

These metrics represent the positional accuracy of the first fixation for a line of text.

M18. First Fixation Horizontal Offset for a Line of Text – horizontal distance between the first fixation of a pass and the beginning of the first word in a line of text (deg).

M19. First Fixation Vertical Offset for a Line of Text (deg). Same as M18, but only vertical distance is considered.

First fixation for a word

These metrics represent the positional accuracy of the first fixation for a word of text and speed of reading.

M20. First Fixation Horizontal Offset for a Word – horizontal distance between the first fixation directed to a word and the beginning of that word (deg).

M21. First Fixation Vertical Offset for a Word (deg). Same as M20 only vertical distance is considered.

M22. First Word Fixation Duration – duration of the first fixation on a word (ms).

M23. Gaze Duration for a Word – sum of durations of all eyes fixations for a word (ms);

2.3.4 Return saccade-related metrics

This set contains metrics related to the local return saccades (LRS). The LRS is one of the longest saccades in reading. Saccades in general contain substantial amounts of biometric information [21].

M24. LRS Mean Horizontal Velocity - mean horizontal eye velocity during a return saccade (deg/s);

M25. LRS Mean Vertical Velocity (deg/s);

M26. LRS Mean Vectorial Velocity (deg/s);

M27. LRS Peak Horizontal Velocity (deg/s);

M28. LRS Peak Vertical Velocity (deg/s);

M29. LRS Peak Vectorial Velocity (deg/s);

M30. LRS Duration (ms);

Obviously, this set of metrics has some redundant information. For example, vectorial saccade velocity overlaps with the horizontal and vertical saccade velocities. To avoid degradation of the biometric performance by such redundancy, only a subset of features is taken in the final biometrics template. Section 3.5.1 describes the final feature selection process.

3. Experimental Methodology

3.1. Apparatus

An EyeLink 1000 eye-tracker in a tower mount setup that consisted of the camera hardware and chin and forehead rest was used to record the eye movements [26]. The eye tracker was set in a monocular mode (left eye was recorded) and 1000 Hz sampling frequency. The eye-tracker's Host PC was connected to a MS Windows Display PC with 22-inch LCD monitor with an active display area of 474x297mm and the resolution of 1680x1050pix. The distance from the participant's eyes to the display was 550mm. The primary eye position of subject corresponded to the center of the screen with a vertical offset of 35 mm above the eye level.

3.2. Stimulus & Procedure

During a recording day each participant was recorded in two sessions. The text for the reading stimulus was taken from Lewis Carroll's poem, "The Hunting of the Snark," chosen for its difficult and nonsensical content, forcing readers to progress slowly and carefully through the text, even in repeated readings. For each text recording, subjects were limited to 1 minute of reading. To reduce learning effect, subjects were given a different excerpt from the text for each recording session. Each excerpt contained 6 quatrains that consisted of 24 lines of text.

The text was displayed in Times New Roman 20 pt. size bold font and was single-spaced. The mean letter interval for each piece of text was approximately 0.50 degrees of the visual angle. The height of the line of the text was 0.92 degrees of the visual angle.

It must be noted that the text recordings for each session were a part of a larger experiment where subjects performed various eye movement tasks with several periods of rest to reduce possible fatigue effects. The time lapse between the text recordings conducted during the same day was approximately 20 minutes. Several other tasks and brief rest periods between those tasks were conducted in addition to text recordings. The total duration of all tasks and periods of rest during both recording sessions did not exceed 1 continuous hour.

3.3. Participants & Recorded Data Quality

3.3.1 Same Day test

Eye movement data was processed from 103 participants, 51 males and 52 females, ages 18 – 43 with an average age of 21.3±3.9, recorded over two sessions labeled as S1 and S2. The institutional review board approved the study, and all subjects provided informed consent. Data recorded from these participants had the average calibration error of 0.48°±0.17 and the maximum calibration error of 1.03°±0.46. The average recorded data validity was 95.5%±5.2.

3.3.2 Template Aging test

The same 103 subjects came back for a template aging experiment that took place approximately one month after the initial recording day.

Three major factors can potentially contribute to the template-aging effect: behavioral, short-term physiological, e.g. fatigue [25] and long-term physiological, e.g aging. There are evidences that aging does not affect saccadic behavior [27], which provides the highest biometric accuracy. Due to this reason we assume that actual muscle tissue aging does contribute to the degradation in performance found in our work. We consider one-month time interval between the recordings reasonable to see template-aging effects due to the behavioral and short-term physiological variability. Changes in the neuronal control signal sent to the extraocular muscles can be an example of short-term physiological variability.

Similar to the Same Day test experiment the participants were recorded over two sessions labeled S3 and S4 and were conducted on the same day. Data recorded on that day had the average calibration error of 0.44°±0.15 and the maximum calibration error of 0.93°±0.37. The average recorded data validity was 95.0%±6.7.

It must be noted that records employed in the Same Day test and Template Aging test experiments came from the larger experiment where 335 individuals came for the Same Day test recordings and only 103 came back for the Template Aging test. For consistency purposes, only data from these 103 individuals who participated in both the Same Day test and the Template Aging test days of recordings were considered in this work. To provide a sufficient data quality, furthermore only records that passed the following empirically found criterion, were analyzed for this study:

1. A record had to have a minimum number of valid passes present (4 in this work).
2. The validity of a pass was defined with min. amount of detected fixation points (3 in this work).
3. Max. distance between top and bottom fixation points in a pass (4 deg. of the visual angle in this work).

In Table 1, The participants proc. column presents the exact number of participants for each scenario that passed the above-mentioned criterion.

3.4. Processing data

3.4.1 Eye Movement Classification

A velocity threshold algorithm (I-VT) with documented accuracy [28] was employed to classify individual data points with a velocity greater than 30°/sec for TRM and 20°/sec for CEM-B as belonging to saccades, with all remaining points belonging to fixations. A micro-saccade filter re-classified saccades with an amplitude less than 0.5° for TAM and CEM-B as fixations, followed by a micro-fixation filter which re-classified fixations with a duration less than 50ms for TRM and 100ms for CEM-B as noise. These thresholds provided the best biometric performance for a particular model.

Eye movement velocity at 1000 Hz sampling frequency was computed as suggested by Bahill et al.'s [29] equation:

$$v(T) = \frac{x(T+3)-x(T-3)}{6T} \tag{1}$$

This equation to calculate saccade velocity was selected because it reduces the amount of noise in the obtained velocity profile when compared to the conventional two-point differentiation of the positional signal for the computation of velocity:

$$v(T) = \frac{x(T)-x(T-1)}{T} \tag{2}$$

3.4.2 Biometric Data Comparison

The biometric template construction phase included removing outliers from the distribution of the individual metrics to ensure meaningful statistical analysis. Outliers were defined as values exceeding 3 standard deviations from the mean. Average performance values for each metric are reported in the next section.

To be able to compare the biometric templates created by the TRM and CEM-B methods, it is important to employ statistical methods that allow comparing the distribution of values. Out of various methods we have investigated, the Two-Sample Kolmogorov-Smirnov (KS) Test performed the best for the TRM and the CEM-B both. For brevity only results from the KS are given.

Two-Sample Kolmogorov-Smirnov (KS) Test

The two-sample Kolmogorov-Smirnov test [30] tests the null-hypothesis that samples of two distributions are drawn from the same original distribution. The calculated p-value can be used as a continuous measure of the difference to get the numeric representation of how two distributed metric patterns are distinguished from each other.

3.4.3 Biometric Data Fusion

Eye movement recordings were analyzed as a set of partitions that were derived by splitting the original set of recorded data by subject into the training and the testing sets according to a uniformly random distribution with a ratio of 1:1, such that no subject had recordings in both the training and testing sets. The term 'training' is employed only in the relation to applied fusion algorithms to form coefficients and other necessary components of fusion process. The training/testing process was performed 100 times with subjects assigned to partitions randomly. Experimental results reported here are averages obtained by running such protocol.

Among different methods investigated in this work, we employed a fusion algorithm based on the Weighed Mean of Rank-1 Identification Rates (IR) [31] of each metric as an algorithm that provides the best balance of biometric accuracy vs. computation time; therefore, this method was employed for both TRM and CEM-B models.

IR Rank 1-based Weighted Mean fusion algorithm

This algorithm consists of the following stages:

1. Obtaining similarity matrices for each of the metrics;
2. Normalization of those matrices by MaxMin method;
3. Obtaining Rank-1 IR for each of the metrics;
4. Normalization of all Rank 1 IRs on a an interval [a, 1] with an equation:

$$nIR_i = \frac{IR_i - \min(IR)}{\max(IR) - \min(IR)} \cdot (1-a) + a \quad (3)$$

where a is a shift of the minimum value to avoid zeros. Zeros in IR values cause zeros in weights, which cause the information loss. Obviously, a=1 will cause a marginal case where all weights are equal. In the present work we used a=0.25, which was empirically found as the best tradeoff between the variability of weights and the prevention of the information losses.

5. Calculating weights for each metric by the following equation to provide a total sum of weights equal to 1:

$$W_i = \frac{IR_i}{\sum_{i=1}^{N} IR_i} \quad (4)$$

3.5. Results

3.5.1 Same Day test

To improve performance of the TRM method by removing redundant and low-performance metrics, two approaches – one based on the statistics and the other based on the performance – were applied to reduce the number of considered metrics.

The statistical approach was based on the multiple linear regression analysis. Biometric templates were analyzed metric by metric, except the fused score. First, the means were calculated for each metric within each record, so for N records and M metrics the $M\times N$ table was formed. Each of the M columns represented a variable. Values within each variable that exceeded 3 standard deviations from the variable mean were recognized as outliers, and L rows that had at least one outlier were discarded from further analysis, leaving $M\times(N-L)$ table. A multiple linear regression analysis was performed for each variable, calculating R^2 values for all given variables (dependent) related to the set of remaining variables (independent).

The variables removal process started with the whole set. The metric with the highest R^2 was removed from the subset. The process of removal continued until the information fusion of the remaining metrics increased the resulting EER instead of decreasing it.

The performance-based approach considered the metrics remaining from the filtering performed by the statistical approach. In case of a performance-based approach, each metric with the worst EER was removed and the EER that resulted from the fusion of the remaining metrics was evaluated. If the resulting EER was reduced the metric was removed. The process of metric removal continued until the minimum possible EER was achieved.

After statistical and performance-based filtering a final subset of the TRM metrics included the following 16, with the numbers in the brackets representing the average EER for this metric using the KS-test for the template matching (distribution comparison):

M1. Hop Horizontal Amplitude (30.3%);
M2. Hop Vertical Amplitude (34.5%);
M4. Hop Duration (22.9%);
M5. Hop Mean Horizontal Velocity (24.3%);
M8. Hop Peak Horizontal Velocity (19.8%);
M9. Hop Peak Vertical Velocity (23.9%);
M13. Ratio of Pass Amplitude to Length of Text Line (40.0%);
M14. Ratio of Pass Total Length to Length of Text Line (39.5%);
M18. First Fixation Horizontal Offset for Line of Text (35.3%);
M22 First Word Fixation Duration (26.3%);
M24. Return Saccade Mean Horizontal Velocity (30.2%);
M25. Return Saccade Mean Vertical Velocity (40.4%);
M26. Return Saccade Mean Vectorial Velocity (28.4%);
M27. Return Saccade Peak Horizontal Velocity (22.2%);
M28. Return Saccade Peak Vertical Velocity (43.3%);
M30. Return Saccade Duration (28.6%);
Fused EER (IR Rank 1-based Weighted Mean): 10.2%

The performance of this subset can be compared with the CEM-B model that does not process the eye movements based on the text's structure. As it was mentioned earlier, the comparison of metric distributions was done by the KS-test.

M1. Fixation start time (41.5%);
M2. Fixation duration (21.7%);
M3. Fixation horizontal centroid coordinates (35.1%);
M4. Fixation vertical centroid coordinates (38.3%);
M5. Saccade start time (42.7%);
M6. Saccade duration (16.9%);
M7. Saccade horizontal amplitude (23.8%);
M8. Saccade vertical amplitude (26.8%);
M9. Saccade horizontal mean velocity (17.1%);
M10. Saccade vertical mean velocity (25.6%);
M11. Saccade horizontal peak velocity (18.9%);
M12. Saccade vertical peak velocity (22.7%);
Fused EER (IR Rank 1-based Weighted Mean): 8.1%

3.5.2 Template-aging test

Table 1 presents the Template Aging results for the TRM and CEM-B models. To assess the difference in EER degradation for the TRM and CEM-B models by switching from Same Day to Template Aging tests, each Same Day value was subtracted from each Template Aging value. Each pair of values from different tests produced a new value, 4×2=8 values total per model. Both 8-number sets were normally distributed, and the Shapiro-Wilks normality test results were not statistically significant. A one-way ANOVA showed the significant difference between the EER of these two models $F(1,14) = 11.249$, $p < 0.05$.

To build a smooth Receiver Operator Characteristics (ROC) curve presented in Figure 2, parametric binormal regression was used [32].

	TRM		*CEM-B*		
Rec.	EER M	EER SD	EER M	EER SD	Particip-ants proc.
			Same Day test		
S1-S2	11.3	1.7	8.9	1.6	82
S3-S4	9.1	1.9	7.3	1.5	91
Mean	*10.2*	*1.8*	*8.1*	*1.5*	
			Template Aging test		
S1-S3	20.0	2.7	19.7	2.7	86
S1-S4	18.8	2.3	19.0	2.4	87
S2-S3	21.9	2.6	22.8	2.9	85
S2-S4	18.8	2.8	21.9	2.9	89
Mean	*19.9*	*2.6*	*20.8*	*2.7*	

Table 1. Results of Same Day and Template Aging tests for TRM and CEM-B

4. Discussion

Both the TRM and the CEM-B have similar metrics related to fixation duration and saccade velocities (M1-M9, M14-M22 of TRM, all metrics of CEM-B). Among the considered metrics, those that represent saccade characteristics yield the highest individual biometric accuracy than the metrics related to the fixational characteristics. Major difference between the TRM and CEM-B is that TRM contains also new metrics related to reading of a text (M13-M18). It also includes a new information fusion method based on the Weighed Mean of Rank-1 Identification Rates that can be also employed by other biometric model, e.g. the CEM-B as it was showed in the present work. In addition, the TRM analyses filtered each record of eye movements for the reading task only, omitting any eye movements that did not correspond to the text structure. The removal of the additional eye movement information might be the reason why the TRM yields lower accuracy in the Same Day scenario. However, in the Template Aging scenario where time-related changes of eye movement characteristics might be present, a structured approach such as the TRM yields better accuracy.

Analysis of the ROC curves provides additional information. The TRM allows getting a better trade-off for the higher TPR with a low FAR for the Same Day scenario whereas CEM-B is just slightly better from this prospective in the Template Aging scenario. This is an additional evidence that the TRM has a good potential for further research.

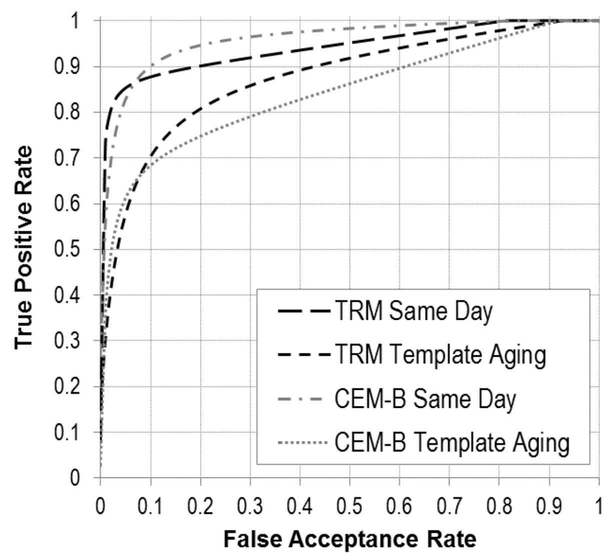


Figure 2. Joint Receiver Operator Characteristic

5. Acknowledgments

This work is supported in parts by NSF CAREER Grant #CNS-1250718 and NIST Grant #60NANB14D274.

6. Conclusion and future work

This paper presented an eye movement-driven biometric technique that is text-based in an effort to provide a framework that would be more robust to the template aging effects. The technique uses text structure to process and align the recorded eye movements to extract their eye movement characteristics that are directly relevant to the task of reading a text. Although the proposed technique did not outperform an existing eye movement-driven biometric method, in the Same Day recording scenario, it outperformed the unstructured method in a scenario where the passage of time might change characteristics of the captured eye movements, i.e., the Template Aging scenario. Thus, the proposed method might be more practical in real-life use, where the time interval between the enrollment and the subsequent uses of a biometrics system might be large.

Our future work will concentrate on the incorporation of the additional metrics into the Text Reading Model including performance characteristics related to the parafoveal-on-foveal effect [33], assessment properties of working memory, linguistic models, mindless reading [34] and mind wandering [35]. Other stimulus-structure methods would be investigated as well to determine their advantages.

References

[1] A. H. Maslow, R. Frager, J. Fadiman, C. McReynolds, and R. Cox, *Motivation and personality* vol. 2: Harper & Row New York, 1970.

[2] N. Jones. (2014, October 8th). *The Future Smart Home: 500 Smart Objects Will Enable New Business Opportunities*. Available: http://www.gartner.com/document/2793317

[3] T. Seals. (2014). *Russian Gang's Billions of Stolen Credentials Resurface in New Attack*. Available: http://www.infosecurity-magazine.com/news/russian-gangs-billions-of-stolen/

[4] J. Rose. (2014). *How Olivia Munn's Apple's iCloud Account Got Hacked*. Available: https://onemonth.com/hacking-olivia-munns-icloud-account

[5] J. J. Yan, A. F. Blackwell, R. J. Anderson, and A. Grant, "Password Memorability and Security: Empirical Results," *IEEE Security & privacy,* vol. 2, pp. 25-31, 2004.

[6] A. De Luca, E. von Zezschwitz, L. Pichler, and H. Hussmann, "Using fake cursors to secure on-screen password entry," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2399-2402.

[7] C. G. Jung, *Psychological Types, Collected Works of C.G. Jung* vol. 6. Princeton, N.J.: Princeton University Press, [1921] 1971.

[8] T. B. Bottomore and M. Rubel, *Karl Marx: Selected writings in sociology and social philosophy*: McGraw-Hill, 1964.

[9] A. K. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*: Springer, 2007.

[10] A.-C. Syvänen, A. Sajantila, and M. Lukka, "Identification of individuals by analysis of biallelic DNA markers, using PCR and solid-phase minisequencing," *American journal of human genetics,* vol. 52, p. 46, 1993.

[11] O. V. Komogortsev, A. Karpov, C. D. Holland, and H. P. Proença, "Multimodal ocular biometrics approach: A feasibility study," in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*, 2012, pp. 209-216.

[12] J. Galbally, A. Ross, M. Gomez-Barrero, J. Fierrez, and J. Ortega-Garcia, "Iris image reconstruction from binary templates: An efficient probabilistic approach based on genetic algorithms," *Computer Vision and Image Understanding,* vol. 117, pp. 1512-1525, 2013.

[13] I. Rigas and O. V. Komogortsev, "Gaze Estimation as a Framework for Iris Liveness Detection," in *IEEE International Joint Conference on Biometrics*, 2014.

[14] O. V. Komogortsev, C. D. Holland, and A. Karpov, "Template aging in eye movement-driven biometrics," in *SPIE Defense+ Security*, 2014, pp. 90750A-90750A-9.

[15] P. Kasprowski and J. Ober, "Eye Movements in Biometrics," presented at the European Conference on Computer Vision (ECCV), Prague, Czech Republic, 2004.

[16] R. Bednarik, T. Kinnunen, A. Mihaila, and P. Fränti, "Eye-movements as a biometric," in *Image analysis*, ed: Springer, 2005, pp. 780-789.

[17] Y. Zhang and M. Juhola, "On biometric verification of a user by means of eye movement data mining," in *IMMM 2012, The Second International Conference on Advances in Information Mining and Management*, 2012, pp. 85-90.

[18] O. V. Komogortsev, A. Karpov, L. Price, and C. Aragon, "Biometric Authentication via Oculomotor Plant Characteristic," in *IEEE/IARP International Conference on Biometrics (ICB)*, New Delhi, India, 2012.

[19] I. Rigas and O. V. Komogortsev, "Biometric recognition via fixation density maps," in *SPIE Defense+ Security*, 2014.

[20] C. D. Holland and O. V. Komogortsev, "Biometric Authentication via Complex Oculomotor Behavior," in *IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, 2013.

[21] C. D. Holland and O. V. Komogortsev, "Complex Eye Movement Pattern Biometrics: Analyzing Fixations and Saccades," in *IEEE/IARP International Conference on Biometrics (ICB)*, Madrid, Spain, 2013.

[22] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin,* pp. 372-422, 1998.

[23] K. Holmqvist, M. Nyström, and F. Mulvey, "Eye tracker data quality: what it is and how to measure it," in *Proceedings of the symposium on eye tracking research and applications*, 2012, pp. 45-52.

[24] 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.

[25] A. T. Bahill and L. Stark, "Overlapping saccades and glissades are produced by fatigue in the saccadic eye movement system," *Experimental Neurology,* vol. 48, 1975.

[26] "EyeLink User Manual For EyeLink models: EyeLink 1000, EyeLink 2000, EyeLink Remote Tower. Desktop, Arm and Primate Mounts. Version 1.4.0 " SR Research Ltd.2008.

[27] L. A. Abel, B. T. Troost, and L. F. Dell'Osso, "The effects of age on normal saccadic characteristics and their variability," *Vision research,* vol. 23, pp. 33-37, 1983.

[28] O. V. Komogortsev, S. Jayarathna, D. H. Koh, and S. M. Gowda, "Qualitative and quantitative scoring and evaluation of the eye movement classification algorithms," in *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 2010, pp. 65-68.

[29] A. T. Bahill, J. S. Kallman, and J. E. Lieberman, "Frequency limitations of the two-point central difference differentiation algorithm," *Biological Cybernetics,* vol. 45, pp. 1-4, 1982.

[30] D. A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises tests," *The Annals of Mathematical Statistics,* pp. 823-838, 1957.

[31] I. Rigas, Abdulin, E., Komogortsev, O., "Towards a multi-source fusion approach for eye movement-driven recognition," *Information Fusion Journal,* under review.

[32] K. H. Zou, A. J. O'Malley, and L. Mauri, "Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models," *Circulation,* vol. 115, pp. 654-657, 2007.

[33] E. R. Schotter, B. Angele, and K. Rayner, "Parafoveal processing in reading," *Attention, Perception, & Psychophysics,* vol. 74, pp. 5-35, 2012.

[34] E. D. Reichle, A. E. Reineberg, and J. W. Schooler, "Eye movements during mindless reading," *Psychological Science,* vol. 21, pp. 1300-1310, 2010.

[35] S. Feng, S. D'Mello, and A. C. Graesser, "Mind wandering while reading easy and difficult texts," *Psychonomic bulletin & review,* vol. 20, pp. 586-592, 2013.