This is a pre-print. The paper will be published in the proceedings of the IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS 2012)

Biometric Verification via Complex Eye Movements: The Effects of Environment and Stimulus

Corey D. Holland, Oleg V. Komogortsev
Texas State University – San Marcos
601 University Drive, San Marcos TX 78666
ch1570@txstate.edu, ok11@txstate.edu

Abstract

This paper presents an objective evaluation of the effects of stimulus type and eye tracking specifications on the accuracy of biometric verification based on complex eye movement patterns (CEM). Five stimulus types (simple, complex, cognitive, random, textual), six spatial accuracy tiers (0.5°, 1.0°, 1.5°, 2.0°, 2.5°, 3.0°), and six temporal resolution tiers (1000 Hz, 500 Hz, 250 Hz, 120 Hz, 75 Hz, 30 Hz) are evaluated to identify their effects. The results suggest the use of eye tracking equipment capable of 0.5° spatial accuracy and 250 Hz temporal resolution for biometric purposes, though biometric accuracy remains achievable for systems capable of at least 1.0° spatial accuracy and 30 Hz temporal resolution. While not conclusive, the complex and textual pattern stimuli provided the greatest accuracy, with little difference between the remaining stimuli.

1. Introduction

Biometrics is a far-ranging and complex discipline, spanning countless paradigms, techniques, and applications [1]. It is, however, a field of active development, in which continual improvements are necessary to maintain and improve upon the efficacy of existing techniques. The ideal biometric–from which an individual may be uniquely identified from his or her peers quickly, cheaply, and without error–has yet to be discovered.

Complex eye movement patterns (CEM) offer a novel method of biometric identification that may prove a panacea to these woes. The properties of human eye movements are determined by a neuronal control signal generated by the brain and the physical characteristics of the six extraocular muscles, eye globe, and surrounding tissue [2]. The dual aspect of physical and neurological components makes accurate replication of eye movements outside the original subject practically infeasible [3]. As well, the ability to perform eye tracking on an unmodified camera [4, 5] and process recorded eye movement data in real-time make it an attractive complement to conventional biometric techniques [6].

1.1. Background

The human eye exhibits several basic types of eye movement in response to various stimuli (both internal and external). In the field of human-computer interaction, fixations and saccades are of primary interest. Fixations occur when the eye globe is held in a relatively stable position such that the fovea remains centered on an object of interest, providing heightened visual acuity. Saccades occur when the eye globe rotates quickly between points of fixation, with very little visual acuity maintained during rotation. The term scanpath refers to the spatial path formed by a sequence of fixations and saccades.

1.2. Previous Research

Eye movements as a behavioral biometric are as of yet a largely underdeveloped branch of the biometric field, the basis for which was formed in 1971 when Noton and Stark [7] found that the general scanpath exhibited by a subject during the first viewing of a pattern was repeated in the initial eye movements of roughly 65% of subsequent viewings. There has been little research in this area [6, 8-11], enough to demonstrate the viability of eye movements as a biometric indicator, but too little to provide a realistic alternative to existing standards.

To our knowledge, Kasprowski and Ober [8] were the first to investigate the use of eye movements as a biometric indicator. The first 15 cepstral coefficients of the positional signal were used as the primary features of comparison, and information fusion was performed using naïve Bayes classifiers, C4.5 decision trees, SVM polynomials, KNN (k = 3 and k = 7). Average FAR and FRR were provided, but EER and detection error tradeoff were not reported.

Silver and Biggs [11] followed, investigating a larger range of possible features, including: the 8 most significant fixations in each recording, fixation count, mean fixation duration, mean saccade velocity, mean saccade duration, and mean vertical position. Information fusion made use of a neural network, but again EER and detection error tradeoff were not reported, in favor of average TPP, TNP, and ACC.

Holland and Komogortsev [6] investigated a wide range of 15 complex eye movement patterns for their viability as biometric indicators, including: fixation count, mean fixation duration, mean vectorial saccade amplitude, mean horizontal saccade amplitude, mean vertical saccade amplitude, mean vectorial saccade velocity, mean vectorial saccade peak velocity, velocity waveform indicator (Q), scanpath length, scanpath area, regions of interest, inflection count, amplitude-duration coefficient, and main sequence coefficient. The viability of each feature was evaluated, along with a pairwise distance comparison of fixation centroids, and a simple information fusion algorithm based on the weighted arithmetic mean of individual features. EER and detection error tradeoff were provided for each feature, with information fusion achieving a 27% equal error rate.

Komogortsev et al. [9] considered the fusion of complex eye movement patterns (CEM) and oculomotor plant characteristics (OPC) to enhance biometric accuracy. The OPC technique makes use of the saccadic eye movement signal to estimate the physical properties of the eye according to a mathematical model of human eye movements [10]. The combination of CEM and OPC biometrics provided a roughly 30% increase in authentication accuracy compared to the accuracy of individual techniques, achieving an equal error rate of 19%.

While it is obvious that eye movements are not yet in a position to replace existing biometric standards, such as fingerprint [12] and iris recognition [13], it is equally evident that eye movements are viable sources of biometric information, which even in their current state may be used in multibiometric systems to enhance the accuracy and specificity of existing techniques. In the current paper, it is our goal to expand this branch of the biometric field by investigating the previously unexplored effects of stimulus type and eye tracking specifications on the accuracy of biometric verification, to determine acceptable conditions under which to collect eye movement data.

2. Eye Movement-Based Biometrics

The biometric techniques considered in this paper follow those described by Holland and Komogortsev [3]. The eye movement signal provided by the eye tracking system is processed to identify, filter, and merge individual gaze points into meaningful units (fixations and saccades). Statistical features are extracted from the fixations and saccades of each recording and similarity scores are generated for pairs of recordings based on a comparison of their individual features. Information fusion is performed on the similarity scores for individual features in order to combine the biometric information inherent in each feature, providing a single score for biometric verification.

2.1. Processing

The eye movement signal provided by the eye tracking system is processed to identify the fixations and saccades that form the scanpath as the eye scans the screen [14]. At present, we employ a velocity threshold classification algorithm, followed by a micro-saccade filter, and a micro-fixation filter. Fixation and saccade groups are then merged. Fixation properties include: timestamp, duration, and centroid; saccade properties include: timestamp, duration, amplitude, velocity, and peak velocity.

Algorithm thresholds were determined empirically from previous eye movement research. Thresholds are provided in Section 3, and alternative processing techniques are considered in Section 5.

2.2. Feature Extraction

Fixations and saccades describe the scanpath of a recording. A variety of statistical features are calculated for each recording based on the properties of its unique scanpath. An explanation and rationale for each feature is provided in [3], considered features include:

1. Scanpath length.
2. Scanpath convex hull area.
3. Fixation count.
4. Average fixation duration.
5. Regions of interest.
6. Inflection count.
7. Average vectorial saccade amplitude.
8. Average horizontal saccade amplitude.
9. Average vertical saccade amplitude.
10. Average vectorial saccade velocity.
11. Average vectorial saccade peak velocity.
12. Slope of the amplitude-duration relationship.
13. Slope of the main sequence relationship.
14. Velocity waveform indicator (Q).
15. Pairwise distance comparison.

2.3. Feature Comparison

To determine the relative similarity between the features of separate recordings, a Gaussian cumulative distribution function (CDF) was applied, where x and μ are the features being compared and σ is the feature-specific standard deviation, according to Equation 1:

$$p = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t-\mu}{2\sigma^2}} \quad (1)$$

The standard deviation used in the CDF function was calculated as the average within-subject standard deviation of the specific feature, according to Figure 1:

```
1. For each subject:
   a. For each feature:
      i. Σ = Standard deviation across recordings.
2. For each feature:
   a. σ = Average Σ across subjects.
```

Figure 1: Feature-specific within-subject standard deviation.

The Gaussian CDF produces a probability value between 0 and 1, where a value of 0.5 indicates an exact match and a value of 0 or 1 indicates no match. This probability is converted into a more intuitive similarity score, where a value of 0 indicates no match and a value of 1 indicates an exact match, according to Equation 2:

$$p = 1 - |2p - 1| \quad (2)$$

2.4. Information Fusion

Information fusion combines the information obtained from different biometric traits to improve the overall accuracy of a biometric system [1]. At present, we employ a weighted mean for the information fusion stage, to improve the overall accuracy of identification while making allowance for the accuracy of individual features.

Metric-specific weighting values were determined using an iterative optimization algorithm to minimize the equal error rate. Algorithm details are given in Section 3, weighting values are provided Section 4, and alternative fusion techniques are considered in Section 5.

3. Methodology

Three experiment paradigms were employed to investigate the effects of environment and stimulus on the accuracy of biometric verification. The first experiment paradigm examined the effects of varied stimulus type, the second experiment paradigm examined the effects of varied spatial accuracy and temporal resolution, and the third experiment paradigm provided data recorded on low-cost eye tracking equipment for cross-validation purposes.

3.1. Participants

For the first experiment paradigm, eye movement data was collected for a total of 22 subjects (17 males, 5 females), ages 18 – 46 with an average age of 28 (SD = 8.7). 17 of the subjects performed 16 recordings each, 3 of the subjects performed 15 recordings each, and 2 of the subjects performed 8 recordings each, generating a total of 333 unique eye movement records.

For the second experiment paradigm, eye movement data was collected for a total of 32 subjects (26 males, 6 females), ages 18 – 40 with an average age of 23 (SD = 5.4). 29 of the subjects performed 4 recordings each, and 3 of the subjects performed 2 recordings each, generating a total of 122 unique eye movement records.

For the third experiment paradigm, eye movement data was collected for a total of 28 subjects (18 males, 10 females), ages 18 – 36 with an average age of 23 (SD = 4.6). 27 of the subjects performed 8 recordings each, and 1 of the subjects performed 7 recordings, generating a total of 223 unique eye movement records.

3.2. Apparatus & Software

For the first experiment paradigm, eye movements were recorded using a Tobii TX300 [15] eye tracking system running at 300 Hz with a vendor-reported spatial accuracy of 0.5° and average data validity of 65% (SD = 36%). Stimuli were presented on a flat screen monitor positioned at a distance of 565 millimeters from the subject, with screen dimensions of 550 × 240 millimeters, and resolution of 1920 × 1080 pixels.

For the second experiment paradigm, eye movements were recorded using an EyeLink 1000 [16] eye tracking system running at 1000 Hz with a vendor-reported spatial accuracy of 0.5°, average calibration accuracy of 0.7° (SD = 0.5), and average data validity of 66% (SD = 36%). Stimuli were presented on a flat screen monitor positioned at a distance of 685 millimeters from the subject, with screen dimensions of 640 × 400 millimeters, and resolution of 2560 ×1600 pixels.

For the third experiment paradigm, eye movements were recorded using a modified PlayStation Eye camera running at 75 Hz with an average calibration accuracy of 1.0° (SD = 0.5) and average data validity of 99% (SD = 4%)[1]. Stimuli were presented on a flat screen monitor positioned at a distance of 540 millimeters from the subject, with screen dimensions of 375 × 302 millimeters, and resolution of 1280 × 1024 pixels.

3.3. Procedure

For the first and third experiment paradigms, eye movement recordings were generated for four stimulus types, chosen to meet the following criteria: simple pattern, complex pattern, cognitive pattern, and textual pattern. For the second experiment paradigm only the textual pattern was considered, and for the third experiment paradigm the cognitive pattern was replaced with a random pattern.

The simple pattern stimulus (HSS) employed a technique commonly used in eye movement research to evoke a fixed-amplitude horizontal saccade at regular intervals [2]. A small white dot appears on a plain black background, the dot jumps back and forth across the screen eliciting a saccade for each jump. The distance between jumps was set to correspond to 20° of the visual angle, due in part to screen constraints, complications separating low-amplitude

[1]Average data validity may be skewed in this case, as it is not possible to detect when the eye tracker began tracking an area of the image other than the subject pupil (i.e. rim of glasses, eyelashes, hair, etc.).

saccades (less than 1°), and variation in the dynamics of high-amplitude (greater than 50°) saccades. Participants were instructed to follow the white dot with their eyes, with 100 horizontal saccades elicited per session. In the first experiment paradigm recordings were performed over eight sessions, while in the third experiment paradigm recordings were performed over two sessions.

The random pattern stimulus (RSS) was similar in presentation to the simple pattern stimulus. A small white dot appears on a black background, and jumps across the screen in a uniformly distributed random pattern to elicit saccades. Participants were instructed to follow the white dot with their eyes, with 100 randomly directed oblique saccades elicited per session. In the third experiment paradigm recordings were performed over two sessions.

The complex pattern stimulus (RIS) employed the Rorschach inkblots commonly used in psychological examination, in order to provide relatively clean patterns which were likely to evoke varied thoughts and emotions in participants. Inkblot images were selected from the original Rorschach psychodiagnostic plates and sized/cropped to fill the screen. Participants were instructed to examine the images carefully, with 3 rotations of 5 inkblots per session. In the first and third experiment paradigms recordings were performed over two sessions.

The cognitive pattern stimulus (CDS) was based loosely on the idea of visual passwords [17], with the intention that the user actively selects a pattern which represents their unique password. Each stimulus image contained 5 or 6 multi-colored dots on a black background, and all dots were visible throughout the stimulus presentation. Participants were instructed to form a pattern by looking at the dots in a specific order, and to remember the order, with 3 rotations of 5 patterns per session. In the first experiment paradigm recordings were performed over two sessions.

The textual pattern stimulus (RES) made use of various excerpts from Lewis Carroll's "The Hunting of the Snark." The poem was chosen for its difficult and nonsensical content, forcing readers to progress slowly and carefully through the text. Text excerpts were selected to ensure that reading required roughly 1 minute, line lengths and the difficulty of the material was consistent, and learning effects did not impact subsequent readings. In the first and second paradigms, participants were given a different excerpt for each of four recording sessions, and in the third experiment paradigm participants were given a different excerpt for each of two recording sessions.

For the second experiment paradigm, dithering and downsampling were applied (exclusively) to the eye movement recordings to artificially reduce spatial accuracy and temporal resolution. Dithering reduces spatial accuracy by adding uniformly distributed error to the recorded eye movement position; considered spatial accuracy tiers from a hardware base of 0.5° included: 0.5°, 1.0°, 1.5°, 2.0°, 2.5°, and 3.0°. Downsampling reduces the temporal resolution by removing data points to lower the average time between points; considered temporal resolution tiers from a hardware base of 1000 Hz included: 1000 Hz, 500 Hz, 250 Hz, 120 Hz, 75 Hz, and 30Hz.

Eye movement recordings were processed and classified to identify fixations and saccades. A velocity threshold algorithm classified individual points with a velocity greater than 50°/sec as saccades, where all remaining points were assumed to be fixations, a micro-saccade filter re-classified saccades with amplitude less than 0.5° as fixations, and a micro-fixation filter re-classified fixations with a duration less than 100 milliseconds as saccades.

Feature extraction was performed across all eye movement recordings, and feature comparison was performed across all possible combinations of eye movement recordings for a given stimulus and paradigm, leading to the following comparisons for each feature considered shown in Table 1:

Table 1. Feature comparisons.

Paradigm	Stimulus	Comparisons	Acceptance	Rejection
1	HSS	13,530	551	12,979
1	RIS	861	20	841
1	CDS	861	20	841
1	RES	3,486	122	3,364
2	RES	7,381	182	7,199
3	HSS	1,485	27	1,458
3	RIS	1,540	28	1,512
3	RSS	1,540	28	1,512
3	RES	1,540	28	1,512

Iterative optimization was then performed to identify suitable per-feature weighting for information fusion. The optimization goal was to minimize the equal error rate of biometric verification for each paradigm/stimulus according to the algorithm presented in Figure 2:

```
1. For each feature:
   a. α = Lower boundary.
   b. β = Upper boundary.
   c. χ = 0 and Σ = Inf.
   d. For φ = range(α, β):
       i. Feature weight = φ.
      ii. λ = Equal error rate.
     iii. If λ < Σ: χ = φ and Σ = λ.
```

Figure 2: Iterative weighting optimization.

For our purposes, the lower boundary was set to 0, and the upper boundary was set to 100. Using the optimized weighting values, information fusion was performed on the recordings of each paradigm/stimulus to obtain final similarity scores for comparisons between recordings. False acceptance rate, false rejection rate, and equal error rate were then calculated at various acceptance thresholds to evaluate the accuracy of each.

4. Results

False acceptance rate (FAR) is defined as the rate at which unauthorized individuals are accepted by the system as valid users, while false rejection rate (FRR) is defined as the rate at which authorized individuals are rejected by the system as invalid users.

4.1. Feature-Specific Weighting

Information fusion used a simple weighted mean to combine similarity scores from the various features. Feature-specific weighting values, shown in Table 2, were determined by an iterative algorithm, described in the previous section. The optimized weighting values provide insight into the eye movement features, described in Section 2, which obtained the highest biometric accuracy for each paradigm / stimulus combination.

Table 2. Feature-specific weighting values.

	Paradigm / Stimulus Combination								
	1				2	3			
Feature	***HSS***	***RIS***	***CDS***	***RES***	***RES***	***HSS***	***RIS***	***RSS***	***RES***
1	95	65	0	0	0	0	0	1	0
2	0	0	0	0	0	0	0	0	0
3	100	100	1	100	94	9	4	0	58
4	5	9	0	6	100	100	0	21	0
5	0	0	0	0	0	0	2	0	0
6	1	6	98	1	0	0	0	10	1
7	0	0	0	5	0	0	0	0	0
8	2	0	0	0	3	0	57	10	0
9	0	0	0	9	0	0	0	0	78
10	10	0	1	72	0	100	100	40	100
11	8	8	23	0	5	0	8	100	21
12	0	0	0	0	6	0	57	8	0
13	9	0	0	0	0	0	0	6	0
14	10	0	0	55	0	0	57	0	0
15	2	0	100	0	0	0	57	0	0

A one-way ANOVA indicated no significant main effect for the weighting values of different paradigm / stimulus combinations, $F(8, 126) = 0.14$, $p = 0.997$.

4.2. Detection Error Tradeoff

The detection error tradeoff (DET) curves plot FAR against FRR at varied acceptance thresholds. Figure 3 provides detection error tradeoff for the varied stimuli of the first experiment paradigm, Figure 4 provides detection error tradeoff for the varied stimuli of the third experiment paradigm, Figure 5 provides detection error tradeoff for the varied spatial accuracy of the second experiment paradigm, and Figure 6 provides detection error tradeoff for the varied temporal resolution of the second experiment paradigm.

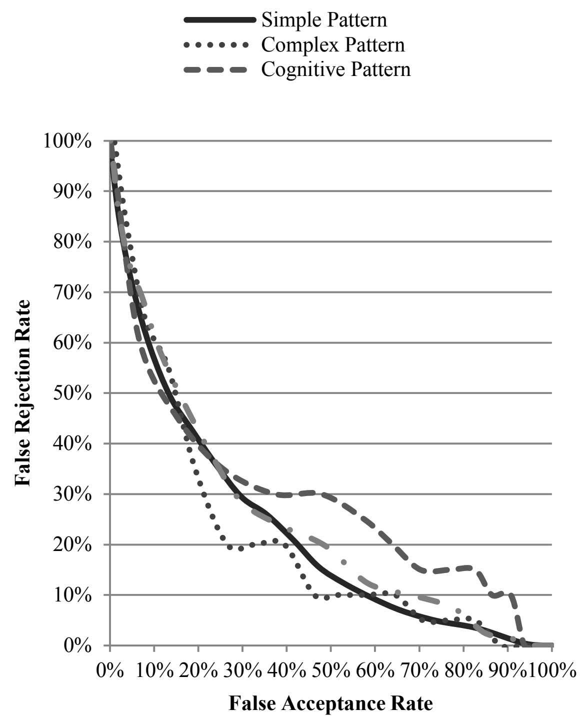

Figure 3: DET Curves: Varied stimulus type (Paradigm #1).

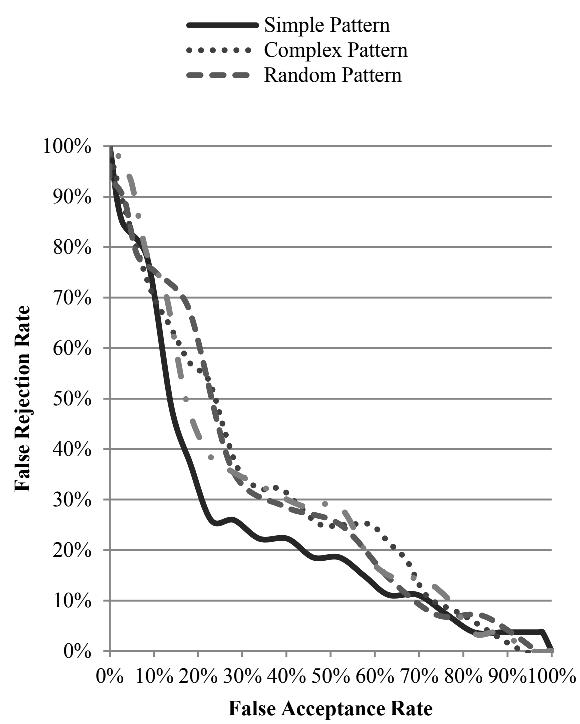

Figure 4: DET Curves: Varied stimulus type (Paradigm #3).

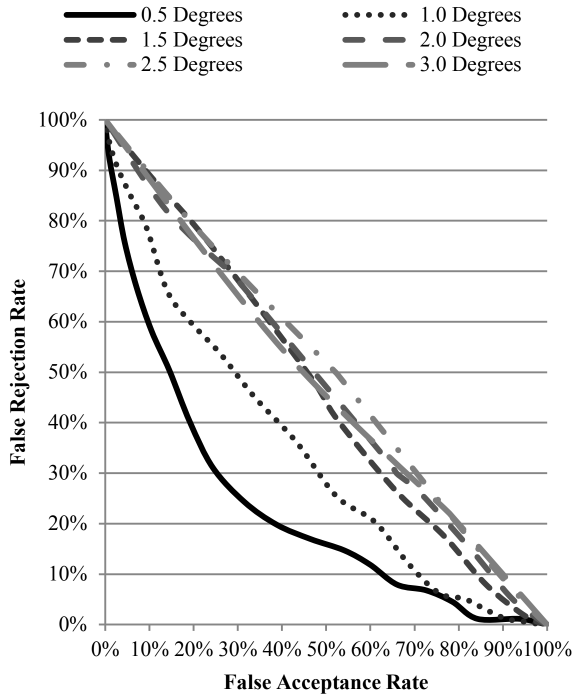

Figure 5: DET Curves: Varied spatial accuracy (Paradigm #2).

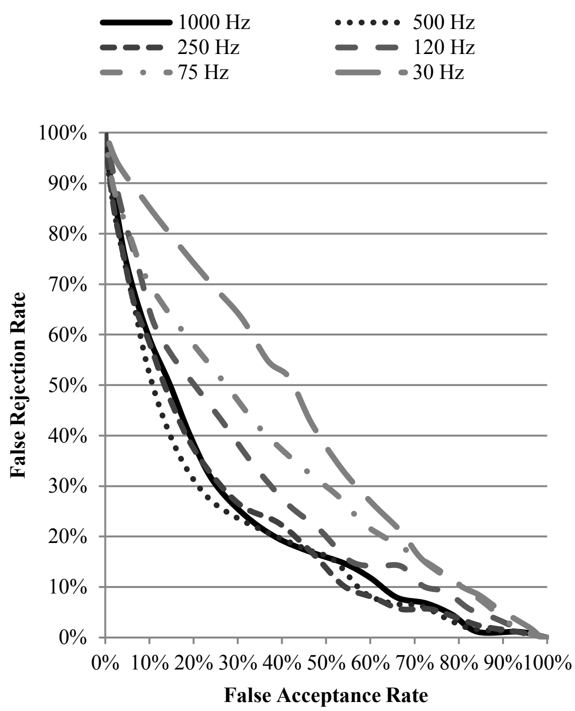

Figure 6: DET Curves: Varied temporal resolution (Paradigm #2).

4.3. Equal Error Rates

The equal error rate (EER) is the error rate at which FAR and FRR are equal. Table 3 provides EER obtained for each stimulus type, Table 4 provides EER obtained for varied spatial accuracy, and Table 5 provides EER obtained for varied temporal resolution.

Table 3. Equal error rates: Varied stimulus type.

	Paradigm / Stimulus Combination								
	1				**2**	**3**			
Feature	***HSS***	***RIS***	***CDS***	***RES***	***RES***	***HSS***	***RIS***	***RSS***	***RES***
1	32%	33%	42%	44%	43%	40%	43%	45%	43%
2	46%	44%	44%	47%	48%	48%	46%	49%	46%
3	32%	33%	38%	36%	34%	34%	39%	43%	40%
4	43%	34%	39%	40%	31%	33%	46%	39%	44%
5	43%	49%	45%	48%	49%	43%	53%	43%	58%
6	36%	39%	36%	39%	38%	51%	46%	44%	40%
7	41%	50%	43%	40%	42%	46%	46%	49%	43%
8	41%	50%	42%	40%	36%	46%	39%	46%	42%
9	48%	50%	50%	48%	46%	43%	43%	39%	38%
10	37%	35%	40%	36%	38%	33%	37%	38%	35%
11	38%	34%	37%	39%	44%	44%	42%	32%	39%
12	45%	54%	43%	45%	44%	40%	42%	40%	50%
13	36%	39%	39%	42%	49%	36%	45%	41%	46%
14	36%	50%	50%	38%	42%	50%	41%	46%	43%
15	41%	48%	35%	42%	44%	36%	42%	42%	50%
Fusion	**29%**	**23%**	**29%**	**29%**	**27%**	**29%**	**31%**	**27%**	**24%**

A one-way ANOVA indicated no significant main effect for varied stimulus type, $F(8, 135) = 0.46$, $p = 0.879$.

Table 4. Equal error rates: Varied spatial accuracy.

	Spatial Accuracy					
Feature	***0.5°***	***1.0°***	***1.5°***	***2.0°***	***2.5°***	***3.0°***
1	43%	42%	48%	N/A	N/A	N/A
2	48%	47%	49%	N/A	N/A	50%
3	34%	37%	49%	49%	N/A	N/A
4	31%	45%	50%	50%	50%	50%
5	49%	50%	49%	49%	N/A	N/A
6	38%	41%	50%	N/A	N/A	N/A
7	42%	47%	50%	43%	43%	48%
8	36%	49%	50%	45%	50%	50%
9	46%	46%	47%	46%	46%	48%
10	38%	47%	50%	49%	47%	46%
11	44%	47%	47%	48%	47%	47%
12	44%	49%	50%	50%	51%	N/A
13	49%	50%	49%	47%	50%	N/A
14	42%	49%	48%	48%	51%	46%
15	44%	45%	48%	N/A	N/A	N/A
Fusion	**27%**	**39%**	**48%**	**49%**	**51%**	**47%**

A one-way ANOVA indicated a significant main effect for varied spatial accuracy, $F(5, 76) = 5.55$, $p < 0.001$.

Table 5. Equal error rates: Varied temporal resolution.

Feature	Temporal Resolution					
	1000 Hz	*500 Hz*	*250 Hz*	*120 Hz*	*75 Hz*	*30 Hz*
1	43%	45%	46%	44%	43%	48%
2	48%	48%	50%	50%	47%	48%
3	34%	33%	36%	36%	39%	44%
4	31%	29%	30%	37%	41%	44%
5	49%	50%	48%	51%	50%	46%
6	38%	36%	40%	42%	40%	N/A
7	42%	39%	41%	45%	42%	50%
8	36%	36%	37%	43%	37%	50%
9	46%	43%	44%	45%	45%	50%
10	38%	38%	38%	43%	38%	50%
11	44%	41%	39%	39%	39%	50%
12	44%	49%	48%	47%	45%	50%
13	49%	46%	44%	42%	44%	45%
14	42%	40%	40%	41%	38%	50%
15	44%	44%	44%	47%	43%	45%
Fusion	**27%**	**26%**	**29%**	**34%**	**39%**	**45%**

A one-way ANOVA indicated a significant main effect for varied temporal resolution, $F(5, 89) = 4.09$, $p < 0.003$.

5. Discussion

Though differences in weighting between trials are not statistically significant, several features stand out for various reasons. Fixation count, average fixation duration, and average vectorial saccade velocity provided a high degree of accuracy across multiple trials, while scanpath length, inflection count, average vertical saccade amplitude, average vectorial saccade peak velocity, velocity waveform indicator (Q), and the pairwise distance comparison provided accurate indicators for different stimuli. Scanpath convex hull area, regions of interest, average vectorial saccade amplitude, and the slope coefficient of the main sequence relationship provided no discernible benefit as biometric indicators, and the remaining features provided only minor accuracy gains.

Differences in equal error rates among stimulus types are not significant; however, several conclusions can be drawn from the distribution of equal error rates and DET curves produced for each stimulus. There is little difference in the accuracy provided by the simple, cognitive, and random pattern stimuli, while the complex and textual pattern stimuli provide slightly better accuracy under certain conditions. An obvious anomaly is the differences in weighting and accuracy of the RES trials, which may have resulted from the differences in subject pool, screen dimensions, or eye tracking specification.

It is apparent that differences in spatial accuracy have a substantial effect on biometric accuracy, with a 44% increase in equal error rate from 0.5° to 1.0° spatial accuracy, and spatial accuracy above 1.5° providing no relevant biometric information. Varied temporal resolution has a much less pronounced effect, with little difference in biometric accuracy from 1000 Hz to 250 Hz. As temporal resolution reduces below 250 Hz, accuracy reduces gradually until it becomes unusable around 30 Hz.

The results of the three experiment paradigms largely serve to reinforce these findings. The equal error rate of 29% obtained with the RES stimulus of the first experiment paradigm collected at 300 Hz correlates with the equal error rate obtained with the RES stimulus of the second experiment paradigm downsampled to 250 Hz. Of particular note is the biometric accuracy obtained on the relatively low-quality, low-cost eye tracking system of the third experiment paradigm. With eye tracking specifications equivalent to 1.0° spatial accuracy and 75 Hz temporal resolution, the results of the second experiment paradigm would suggest that biometric accuracy will degrade below 39% equal error rate; however, in practice we were able to obtain equal error rates not exceeding 31%. This suggests that the error introduced by dithering and downsampling may be exaggerated over natural effects.

5.1. Limitations

The considered techniques are obviously limited in their practical application due to the relatively high error rates in comparison to accepted biometric standards. The focus of this paper has been on determining acceptable conditions under which to collect viable eye movement data, and as such we have done little to improve the biometric accuracy over previous research. Due to the underdeveloped nature of eye movement-based biometrics, future research will likely investigate not only improvements in biometric accuracy, but the techniques and practices used for collection, processing, and analysis of human eye movements.

It is also worth noting that the relatively smaller number of acceptance comparisons to rejection comparisons results in false rejection rates that are statistically less sound than false acceptance rates. This is a common issue in biometrics, however, which results from the constraints on same-subject experimentation. To achieve equivalent amounts of acceptance and rejection comparisons, it is necessary for each participant to perform a number of trials greater than the total number of participants, which becomes increasingly prohibitive as the number of participants increases.

As well, the dithering approach applied to reduce spatial accuracy may not accurately model the spatial accuracy of specific individuals and systems. There exists no current literature which mathematically describes the distribution of eye tracking error across the screen. As such, we have employed a uniform distribution of random noise as an approximation.

5.2. Future Research

Processing is a crucial step in eye movement-based biometrics, and is essential to the transformation of the continuous eye movement signal into discrete units of attention. There are a wide range of classification algorithms and post-classification filters that may be applied to identify fixations and saccades within the recording, each with unique benefits. As well, the thresholds ascribed to these algorithms may have a substantial effect on the accuracy of classification. In the current paper, we have applied a simple velocity threshold algorithm, followed by micro-saccade and micro-fixation filters to improve classification accuracy. This combination was chosen for its speed and simplicity, though more accurate classification may, in turn, improve biometric accuracy. It is likely that future research will investigate alternative classification algorithms/thresholds to identify more effective solutions.

Information fusion is another crucial step in eye movement-based biometrics, and is necessary to combine biometric information provided by individual features. While it has provided sufficient accuracy for the focus of this paper, the use of a simple weighted mean as the primary source of information fusion is a striking weakness of the current implementation. More advanced fusion techniques will be necessary to achieve a level of accuracy that meets current biometric standards. Future research will investigate alternative information fusion techniques.

6. Conclusion

This paper has presented an objective evaluation of the effects of stimulus type and eye tracking specifications on the accuracy of eye movement-based biometric verification. Five stimulus types (simple, complex, cognitive, random, textual), six spatial accuracy tiers (0.5°, 1.0°, 1.5°, 2.0°, 2.5°, 3.0°), and six temporal resolution tiers (1000 Hz, 500 Hz, 250 Hz, 120 Hz, 75 Hz, 30 Hz) were evaluated to determine acceptable conditions under which to collect viable eye movement data for biometric accuracy.

Based on the results, we can conclude that complex and textual pattern stimuli are preferable to the simple, cognitive, and random pattern stimuli examined in this work. As well, eye tracking systems with spatial accuracy of less than 0.5° and greater than 250 Hz are recommended for biometric purposes, due to degradation in accuracy as specifications are reduced beyond this point. Eye tracking systems with greater than 1.0° spatial accuracy or less than 30 Hz temporal resolution are not likely to produce viable biometric information.

References

[1] A. K. Jain, P. Flynn and A. A. Ross. Handbook of biometrics. Springer US, 1-556, 2007.

[2] R. J. Leigh and D. S. Zee. The neurology of eye movements. Oxford University Press, 1-776, 2006.

[3] O. V. Komogortsev and J. I. Khan. Eye movement prediction by kalman filter with integrated linear horizontal oculomotor plant mechanical model. Proceedings of Eye Tracking Research & Applications (ETRA) Symposium, 229-236, 2008.

[4] C. D. Holland and O. V. Komogortsev. Eye tracking on unmodified common tablets: Challenges and solutions. Proceedings of Eye Tracking Research & Applications (ETRA) Symposium, 1-4, 2012.

[5] W. Sewell and O. V. Komogortsev. Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network. Proceedings of Conference on Human Factors in Computing Systems (CHI), 3739-3744, 2010.

[6] C. D. Holland and O. V. Komogortsev. Biometric identification via eye movement scanpaths in reading. Proceedings of International Joint Conference on Biometrics (IJCB), 1-8, 2011.

[7] D. Noton and L. W. Stark. Scanpaths in eye movements during pattern perception. Science, 171(3968):308-311, 1971.

[8] P. Kasprowski and J. Ober. Eye movements in biometrics. Proceedings of European Conference on Computer Vision (ECCV), 248-258, 2004.

[9] O. V. Komogortsev, A. Karpov and C. D. Holland. Cue: Counterfeit-resistant usable eye-based authentication via scanpaths and oculomotor plant characteristics. Proceedings of SPIE Defense, Security, and Sensing, 1-10, 2012.

[10] O. V. Komogortsev, A. Karpov, L. R. Price and C. R. Aragon. Biometric authentication via oculomotor plant characteristics. Proceedings of International Conference on Biometrics (ICB), 1-8, 2012.

[11] D. L. Silver and A. J. Biggs. Keystroke and eye-tracking biometrics for user identification. Proceedings of International Conference on Artificial Intelligence (ICAI), 344-348, 2006.

[12] R. Cappelli, D. Maio, D. Maltoni, J. L. Wayman and A. K. Jain. Performance evaluation of fingerprint verification systems. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(1):3-18, 2006.

[13] V. Dorairaj, N. A. Schmid and G. Fahmy. Performance evaluation of non-ideal iris based recognition system implementing global ica encoding. Proceedings of International Conference on Image Processing (ICIP), 285-288, 2005.

[14] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. Proceedings of Eye Tracking Research & Applications (ETRA) Symposium, 71-78, 2000.

[15] Tobii Technology. Tobii TX300 Eye Tracker. Available: http://www.tobii.com/.

[16] SR Research. EyeLink 1000 Eye Tracker. Available: http://www.sr-research.com/.

[17] K. Renaud and A. d. Angeli. Visual passwords: Cure-all or snake oil? Communications of the ACM, 52(12):135-140, 2009.