Complex Eye Movement Pattern Biometrics: The Effects of Environment and Stimulus

Corey D. Holland, *Student Member*, *IEEE* and Oleg V. Komogortsev, *Member*, *IEEE*

***Abstract*—This paper presents an objective evaluation of the effects of eye tracking specification and stimulus presentation on the biometric viability of complex eye movement patterns (CEM). Six spatial accuracy tiers (0.5°, 1.0°, 1.5°, 2.0°, 2.5°, 3.0°), six temporal resolution tiers (1000 Hz, 500 Hz, 250 Hz, 120 Hz, 75 Hz, 30 Hz), and five stimulus types (simple, complex, cognitive, textual, random) are evaluated to identify acceptable conditions under which to collect eye movement data. The results suggest the use of eye tracking equipment capable of at least 0.5° spatial accuracy and 250 Hz temporal resolution for biometric purposes, while stimulus had little effect on the biometric viability of eye movements.**

***Index Terms*—Biometrics, eye movements, pattern analysis, security and protection.**

I. Introduction

YOU are unique. Physically, mentally, emotionally, there is no one on earth quite like you. There are more than 1.4 million variable nucleotides in the human genome, with trillions of possible variations [3]. As humans, we pick up on these differences, and assign identity based on them. The shape of his nose, the color of her eyes, we notice and respond to these traits, often subconsciously.

Attempts to systematize the collection and comparison of these identifying characteristics (biometrics) can be traced back as early as 1858, when a civil service employer in India began recording handprints to identify workers [5]. Over the past century this practice has expanded considerably, encompassing fingerprints, palm prints, hand geometry, facial structure, signature, speech, and iris patterns [6]; however, automated biometric systems have only begun to see widespread production and use in the past several decades [7].

In a very short time, biometric systems have experienced far-reaching adoption in the fields of law enforcement, criminal justice, and corporate and personal security. Suspect identification, criminal conviction, personalized interfaces, and access restriction constitute only a subset of the many and varied applications of biometrics in modern society.

Manuscript received January 22, 2014. This work is supported in part by NSF CAREER Grant #CNS-1250718 and NSF GRFP Grant #DGE-11444666, and NIST Grants #60NANB10D213 and #60NANB12D234. Special gratitude is expressed to Katie Holland for her aid with technical illustrations.

C. D. Holland is with the Computer Science Department, Texas State University, San Marcos, TX 78666 USA (e-mail: ch1570@txstate.edu).

O. V. Komogortsev is with the Computer Science Department, Texas State University, San Marcos, TX 78666 USA (phone: 512-245-0349; fax: 512-245-8750; e-mail: ok11@txstate.edu).

As humans, we assess these identifying features qualitatively, and rather vaguely: “he has a big nose” or “she has green eyes.” When we attempt to automate these assessments, we are forced to formalize our definitions of these traits, assigning quantitative values based on measurable characteristics. Unfortunately, something is often lost in translation, and automated systems are incapable of assessing an entire individual in the same manner that we are able to assess each other.

This leads to a number of factors that must be considered when implementing an automated biometric system: accuracy, counterfeit-resistance, speed, and cost. For example, a human observer might review facial photographs with perfect accuracy, but too slowly to be considered useful; or, a fingerprint scanner may be implemented with acceptable accuracy and speed, but easily fooled by duplicate images.

The human visual system offers a number of unique and novel properties that make it an interesting candidate in this respect. Eye movements are dependent upon both the physical properties of the oculomotor plant and the neurological properties of the brainstem control [8]. This dual aspect makes the accurate replication of eye movements practically infeasible (if not impossible) outside of a living subject. This, in and of itself, provides inherent counterfeit-resistance and liveness detection. Further, the ability to record and process eye movements in real-time using an unmodified camera [9] makes their application both cheap and efficient.

Open dataset competitions have shown fingerprint recognition [2] capable of equal error rates approaching 2% under the effects of skin distortion and rotation, while iris recognition [10] has shown itself capable of equal error rates approaching 1% under multiple environments and sessions. Though both of these systems suffer from known spoofing vulnerabilities [11], eye movement biometrics can be easily integrated into existing iris recognition systems to improve accuracy and enhance counterfeit-resistance without the need for additional equipment, cost, processing time, or reduction in usability [12].

Aside from their biometric applications, the study of complex eye movement patterns has implications in the fields of clinical health and psychology. Eye movements have been used to diagnose neurological illness for decades [8], and identifying the variability of these metrics in normal human subjects is necessary to establish the limits of normal and abnormal ocular behavior.

A. 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 [13] 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, enough to demonstrate the viability of eye movements as a biometric indicator, but too little to provide an alternative to existing standards.

To the best of our knowledge, the first investigation of eye movements as a biometric indicator occurred in 2004, when Kasprowski and Ober [14] examined the first 15 cepstral coefficients of the positional eye movement signal, a technique commonly applied in voice recognition. Information fusion was performed using naïve Bayes classifiers, C4.5 decision trees, SVM polynomials, and KNN (k = 3 and k = 7). On a dataset of 9 subjects, the described techniques achieved an average 1% FAR and 23% FRR.

Silver and Biggs [15] followed in 2006, 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 ROC curves were not reported, in favor of average TPP, TNP, and ACC. On a dataset of 21 student participants, the described techniques achieved an average 66% TPP and 98% TNP.

Holland and Komogortsev [16] investigated a wide range of complex eye movement patterns (CEM-P), in 2011, 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. With a subject pool of 32 participants, the considered techniques achieved 27% EER.

Komogortsev et al. [17] examined the use of mathematical models of the oculomotor plant (OPC), in an attempt to extract the anatomical constants unique to a given individual from the observable properties of human eye movements. Over a subject pool of 59 participants, the considered techniques achieved a minimum HTER of 19%.

Rigas et al. [18] applied graph-based matching techniques to the positional eye movement signal in 2012, comparing minimum spanning trees using a multivariate Wald-Wolfowitz runs test. On a dataset of 15 subjects, the proposed techniques were able to achieved 70% rank-1 IR and 30% EER.

Most recently, in 2013, Holland and Komogortsev [19] improved upon the complex eye movement pattern (CEM-B) biometrics originally developed in 2011 [16], comparing the distribution of fixations and saccades with statistical techniques such as the two-sample t-test, the Ansari-Bradley test, the two-sample Kolmogorov-Smirnov test, and the two-sample Cramér-von Mises test. On a dataset of 32 subjects, the proposed techniques achieved 83% rank-1 IR and 17% EER.

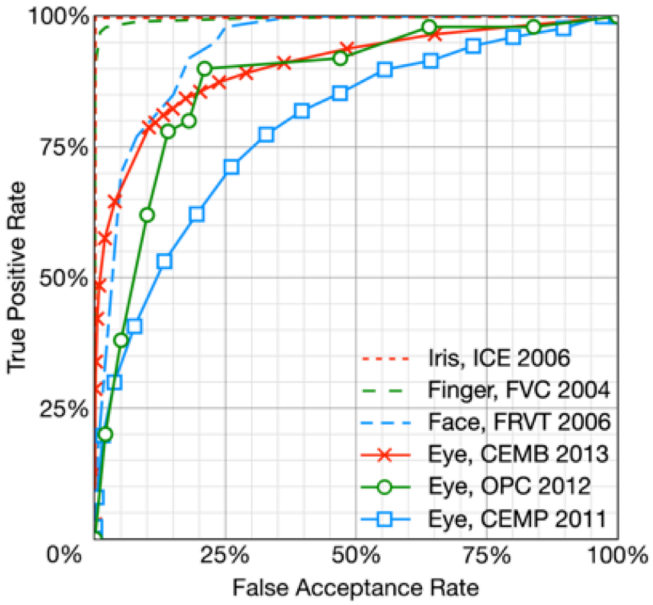


Figure 1. Receiver operating characteristics, reconstructed from the relevant literature. Iris [1], Finger [2], and Face [4] represent the best performing biometric algorithms from the most recently published NIST competition.

B. Eye Movement Biometrics

Having existed for less than a decade, the field of eye movement biometrics is still in its infancy. Despite this, and perhaps because of it, the achievable accuracy and robustness has increased at an exponential rate. In just the past year, equal error rates have seen a reduction of 63%, from 27% EER to 17% EER, while rank-1 identification rates have seen an increase of 157%, from 53% rank-1 IR to 83% rank-1 IR.

As a behavioral—rather than physical—biometric, it is expected that eye movements may never achieve the level of accuracy afforded by physical traits, such as fingerprints and iris patterns; however, when considered in the context of behavioral biometrics, eye movements are quite promising. For example, gait recognition was proposed in the mid-1970s [20], but did not begin to achieve reasonable accuracy until the early 2000s, with equal error rates ranging from 18-25% [21] and rank-1 identification rates ranging from 30-70% [22, 23], depending largely on the angle and speed of gait. Similarly, face recognition was proposed in the 1960s [24], but was largely ignored in a biometric context until the mid-1990s, with early work being highly susceptible to aging effects, achieving equal error rates of 1-7% [25] and rank-1 identification rates of 80-90% [26] for images captured within a single recording session, which became dramatically reduced to equal error rates of 12-23% [25] and rank-1 identification rates of 30-60% [26] after as little as 1 week between enrollment and authentication.

While eye movements are not yet capable of competing with existing biometric standards in standalone systems, as is evident in Figure 1, they possess a number of qualities which make them ideal for multi-biometric systems. First, diverse eye movement biometric techniques can be readily combined to increase the overall accuracy of the system; Komogortsev et al. [27] showed that a combination of CEM-P and OPC techniques increased authentication accuracy by 30% when

compared to individual techniques. Second, because eye movement data can be collected from a single image sensor, eye movement biometrics can be easily incorporated into multi-biometric systems to increase overall accuracy without increasing cost or reducing speed; Komogortsev et al. [12] showed that incorporating eye movement biometrics into an iris recognition system could reduce error rates by up to 19%. Finally, the difficulty of accurately replicating human eye movements provides inherent liveness detection capabilities; Komogortsev and Karpov [28] have shown the OPC technique is capable of correctly classifying 80-93% of recordings as human or spoof.

Unfortunately, in much the same way that smudged fingerprints and off-angle facial images reduce recognition accuracy in their respective systems, the quality of the recorded eye movement signal has been shown to reduce the accuracy of eye movement biometrics [29]; and in much the same way that speed of gait can affect the accuracy of gait recognition, it is unclear if the particular pattern of eye movements, invoked by a specific stimulus, has a noticeable effect on the accuracy. It is for this reason that, in the current paper, we examine the effects of environment and stimulus on the accuracy of eye movement biometrics, specifically the CEM-P technique [16].

C. *Human Visual System*

The oculomotor plant consists of the eye globe and six extraocular muscles: the lateral and medial recti are primarily responsible for horizontal rotation; the superior and inferior recti are primarily responsible for vertical rotation; and the superior and inferior obliques are primarily responsible for torsion of the eye globe. The brainstem control involves a neuronal control signal, generated by the brain and sent to individual extraocular muscles, and is responsible for the muscle contractions and relaxations that produce the many and varied types of human eye movement [8].

Among the various eye movement types exhibited by the human visual system, fixations and saccades are of particular 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. Scanpaths have been identified with higher-level cognition and applied to the study of usability and visual search [30].

D. *Complex Eye Movement Patterns*

Complex eye movement patterns (CEM-P), presented graphically in Figure 2, represent the cognitive strategies employed by the brain throughout the guidance of visual attention. The human eye is connected to and controlled by a complex network of brain regions, sub-regions, and neural pathways [8, 31]. Information is transmitted from region to region along neural pathways in the form of neural signals, which may convey visual field information from the eye or control information from the brain.

The firing rate of individual neural signals (which occur in sustained bursts) is dependent on the physical properties of the involved neurons and surrounding brain tissue. This neural activity may be influenced by the task being performed, which can cause variation in baseline firing rates, firing rate profiles, and modulations of neuronal activity related to particular stimuli and behavioral responses [32].

Generally, eye movements are controlled by three major brain regions, the thalamus, superior colliculus, and posterior parietal cortex [33], where the thalamus is responsible for engaging visual attention, the superior colliculus is responsible for relocating visual attention, and the posterior parietal cortex is responsible for disengaging visual attention; however, this is a major simplification of an extremely complex process. The human visual system is functionally dependent on the superior colliculus, frontal eye fields, lateral intra parietal, posterior parietal cortex, and visual cortex areas V1 – V5 [34]. In addition, the human visual system is affected by the passive properties of tissue surrounding the eye globe, and a myriad of subsidiary neural regions and pathways.

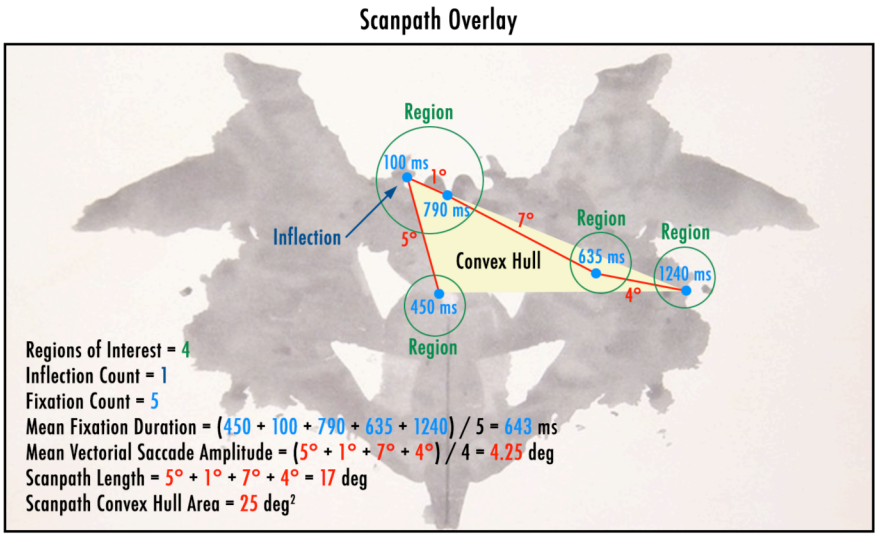
Further, the neuronal control signals responsible for innervation of individual extraocular muscles during a saccade are generated in two physically separate regions of the brain [8]. Specifically, the horizontal saccade component is generated by the paramedian pontine reticular formation, while the vertical and torsional saccade components are generated by the rostral interstitial nucleus of the medial longitudinal fasciculus. As well, the omnipause neurons responsible for fixation are located in the nucleus raphe interpositus, in the midline of the pons.

At a higher level of abstraction, learned behaviors and subconscious memory mechanisms are involved in the coordination of eye movements over time, as evidenced by the inhibition of return and scanpath theory phenomena. Inhibition of return refers to the marked tendency to avoid re-fixation on previously examined features during visual search, effecting both oculomotor programming and target detection [35], while scanpath theory describes the phenomenon in which individuals tend to repeat certain scanpath trajectories during repeated viewings of a given pattern [13].

E. *Motivation & Hypothesis*

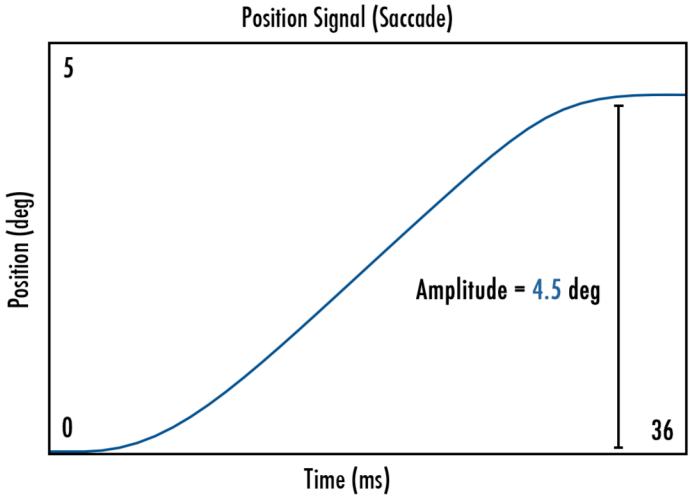
Despite their limited accuracy, eye movement biometrics have been utilized in their current state to increase the accuracy and counterfeit-resistance of multi-biometric systems [12]. As a relatively unexplored branch of the biometric field, whose initial investigations began in 2004 and have received little attention from the biometric community since, it is first necessary to identify and standardize a set of techniques and practices for the collection, processing, and analysis of human eye movements as biometric indicators.

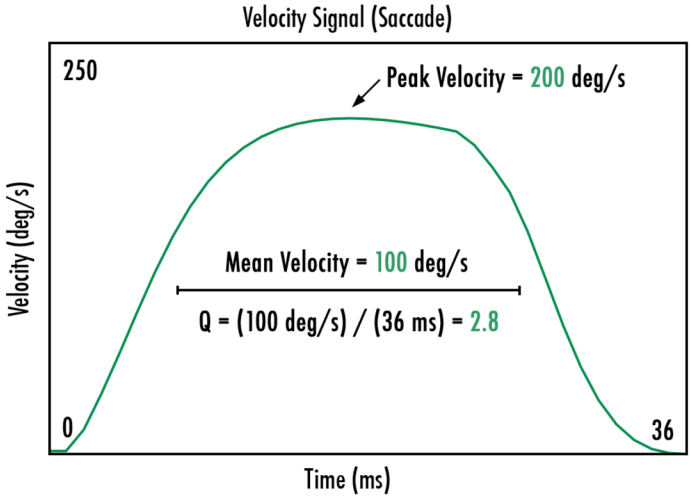
In this paper, we hypothesize that stimulus presentation and eye tracking specifications, such as spatial accuracy and temporal resolution, may affect the biometric viability of complex eye movement patterns. We examined these issues previously [29], providing detection error tradeoff and equal error rates in a verification scenario. The current paper expands greatly on our previous work, adding: an increased subject pool; genuine vs. impostor match score distributions; d-prime; receiver operating characteristic in a verification scenario; cumulative match characteristic and identification rate in an identification scenario.

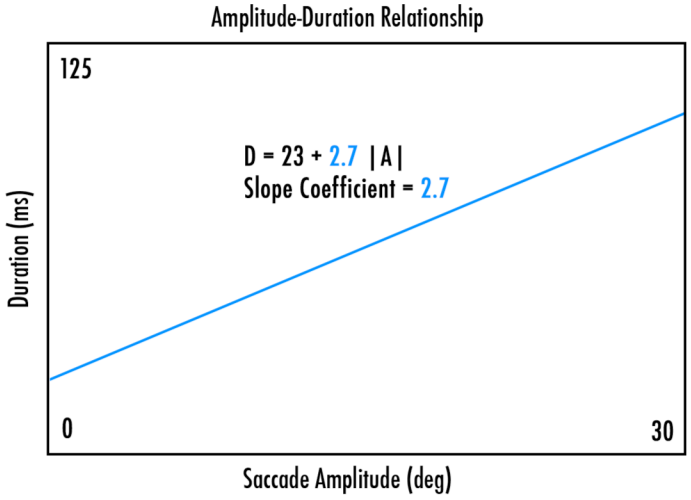


CEM BIOMETRIC FEATURES:

1. Fixation Count
2. Mean Fixation Duration
3. Mean Vectorial Saccade Amplitude
4. Mean Horizontal Saccade Amplitude
5. Mean Vertical Saccade Amplitude
6. Mean Vectorial Saccade Velocity
7. Mean Vectorial Saccade Peak Velocity
8. Slope of the Amplitude-Duration Relationship
9. Slope of the Main Sequence Relationship
10. Velocity Waveform Indicator (Q)
11. Scanpath Length
12. Scanpath Convex Hull Area
13. Regions of Interest
14. Inflection Count







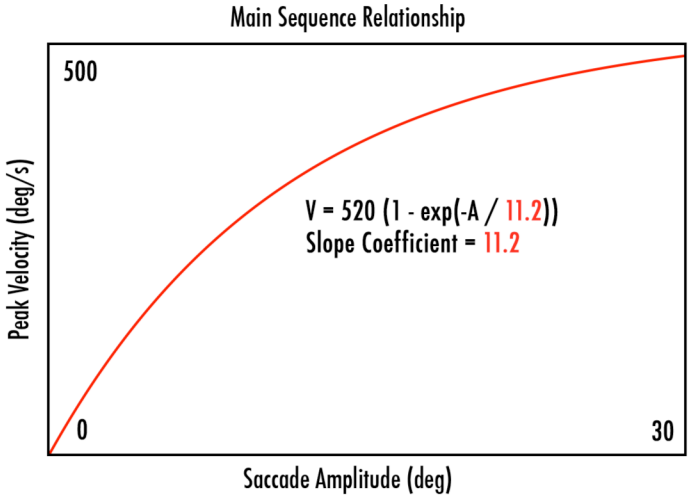


Figure 2. Complex Eye Movement Pattern (CEM-P) Biometrics.
The Scanpath Overlay demonstrates an idealized scanpath segment, where dots indicate points of fixation, and lines indicate saccades.
The Position Signal (Saccade) and Velocity Signal (Saccade) demonstrate the respective positional and velocity signals for a 4.5° saccade.
The Amplitude-Duration Relationship demonstrates the tendency for saccade duration (D) to increase linearly with amplitude (A).
The Main Sequence Relationship demonstrates the tendency for saccadic peak velocity (V) to increase exponentially with amplitude (A).

II. Methodology

The biometric techniques considered in this paper extend from those described by Holland and Komogortsev [29]. Three experiments were conducted to investigate the effects of environment and stimulus on the biometric viability of complex eye movement patterns. The first experiment examined the effects of varied stimulus type, the second experiment examined the effects of varied spatial accuracy and temporal resolution, and the third experiment provided data recorded on low-cost eye tracking equipment for cross-validation purposes. The collected eye movement datasets are available as part of the EMDB v2 database [36].

A. Participants

For the first experiment, 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, generating a total of 333 unique eye movement recordings.

For the second experiment, 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 recordings.

For the third experiment, eye movement data was collected for a total of 173 subjects (117 males, 56 females), ages 18 – 49 with an average age of 23 (SD = 5.3). 158 of the subjects performed 8 recordings each, 6 of the subjects performed 7 recordings each, 7 of the subjects performed 6 recordings each, 1 of the subjects performed 3 recordings, and 1 of the subjects performed 2 recordings, generating a total of 1353 unique eye movement recordings.

The subject pools for each experiment did not overlap. Each subject performed multiple recording sessions for each stimulus. The number of recording sessions for each subject, per stimulus, is presented in the following sub-sections. Texas State University's institutional review board approved the study, and all subjects provided informed consent.

B. Apparatus & Software

For the first experiment, eye movements were recorded using a Tobii TX300 eye tracking system, with a temporal resolution of 300 Hz, 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 mm from each subject, with dimensions of 550×240 mm, and screen resolution of 1920×1080 pixels.

For the second experiment, eye movements were recorded using an EyeLink 1000 eye tracking system, with a temporal resolution of 1000 Hz, vendor-reported spatial accuracy of 0.5°, average calibration accuracy of 0.7° (SD = 0.5), and average data validity of 95% (SD = 5%). Stimuli were presented on a flat screen monitor positioned at a distance of 685 mm from each subject, with dimensions of 640×400 mm, and screen resolution of 2560×1600 pixels.

For the third experiment, eye movements were recorded using a modified version of the open-source ITU Gaze Tracker software [9] and PlayStation Eye Camera, with a temporal resolution of 75 Hz and average calibration accuracy of 1.1° (SD = 0.8). Average data validity is the percentage of gaze points reported by the eye tracking system that contained valid eye movement data, and is unreportable in this instance as it was 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.). Stimuli were presented on a flat screen monitor positioned at a distance of 540 mm from each subject, with dimensions of 375×302 mm, and screen resolution of 1280×1024 pixels.

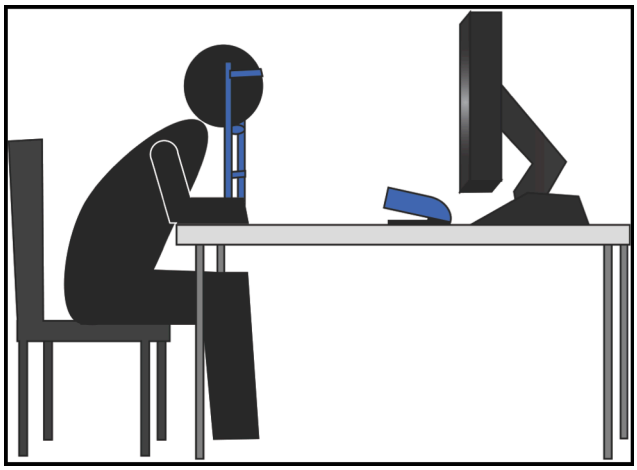


Figure 3. Experimental Setup.

In all cases, the pupil was illuminated by infrared LED to improve eye tracking accuracy, and a chin rest was employed to improve stability, as shown in Figure 3. Stimulus presentation was consistent across all experiments, with only minor changes required for varied screen dimensions.

C. Procedure

For the first experiment, eye movement recordings were generated for four stimulus types and recorded with a standard, commercial-grade eye tracking system. Stimuli were selected to meet the following criteria: simple pattern, complex pattern, cognitive pattern, and textual pattern.

In theory, the eye movements evoked by each pattern would be influenced by different aspects of the human visual system. Responses to the simple pattern would be influenced by the physical structure of the human visual system; responses to the complex pattern would be influenced by subconscious search strategies; responses to the cognitive pattern would be influenced by conscious search strategies; and responses to the textual pattern would be influenced by both conscious and subconscious search strategies.

The simple pattern (SIM) employed a technique used in eye movement research to evoke a fixed-amplitude horizontal saccade at regular intervals [8]. A small white dot jumps back and forth across a plain black background, eliciting a saccade for each jump. The distance between jumps was set to correspond to 30° of the visual angle, due in part to screen constraints, complications separating low-amplitude saccades (less than 1°), and variation in the dynamics of high-amplitude saccades (greater than 50°). Subjects were instructed to follow the white dot with their eyes, with 100 horizontal saccades elicited per session, over 8 recording sessions per subject.

The complex pattern (COM) employed the Rorschach inkblots commonly used in psychological examination, in order to provide relatively clean patterns that were likely to evoke varied thoughts and emotions in participants. Inkblots were selected from the original Rorschach psychodiagnostic plates [37] and sized/cropped to fill the screen. Subjects were not required to provide verbal feedback, but were instructed to examine the images carefully, with 3 rotations of 5 inkblots per session, over 2 recording sessions per subject.

The cognitive pattern (COG) was based loosely on the idea of visual passwords [38], with the intention that the user would actively select a pattern that 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. Subjects 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, over 2 recording sessions per subject.

The textual pattern (TEX) 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 material was consistent, and learning effects did not impact subsequent readings. Subjects were given different excerpts for each session, over 4 recording sessions per subject.

For the second experiment, eye movement recordings were generated for a single stimulus type and recorded with a high-quality, commercial-grade eye tracking system. Again the textual pattern stimulus was employed, over 4 recording sessions per subject. 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 30 Hz.

For the third experiment, eye movement recordings were generated for four stimulus types and recorded with a low-cost eye tracking system. The simple pattern, complex pattern, and textual pattern stimuli of the first experiment were employed to validate previous results, while the worst performing stimulus of the first experiment, the cognitive pattern, was replaced with a random pattern stimulus. Each stimulus was recorded over 2 recording sessions per subject.

The random pattern (RAN) was similar in presentation to the simple pattern stimulus. A small white dot jumps across a plain black background in a uniformly distributed random pattern, eliciting a saccade for each jump. Subjects were instructed to follow the dot with their eyes, with 100 randomly directed oblique saccades elicited per session.

Following the collection of eye movement data, eye movement recordings were processed according to the biometric framework described previously by Holland and Komogortsev [29]. A velocity threshold algorithm (I-VT), with documented accuracy [39], classified gaze points with a velocity greater than 50°/sec as saccades, a micro-saccade filter re-classified saccades with an 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 for the biometric features described in Figure 1. Matching was performed for each feature, using a Gaussian CDF normalized to a scale of the unit interval to compare features between recordings, where σ is the standard deviation of the feature within the training set, and both t and μ represent the feature values of separate recordings:

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

Information fusion applied a weighted average of the scores from individual features [40], to improve the overall accuracy while making allowance for the accuracy of specific features. Iterative optimization was performed on the training set to identify suitable per-feature weighting, using the following forward search algorithm:

Algorithm *ForwardSearch*
1. $N \leftarrow$ Number of biometric features;
2. *Weight* ← Zero-filled array of size N;
3. **for** $i \leftarrow 1$ **to** N
4. **do** $W \leftarrow$ *Weight*;
5. $E \leftarrow$ +Inf;
6. **for** $j \leftarrow 0$ **to** *100*
7. **do** $W[i] \leftarrow j$;
8. $T \leftarrow$ Equal error rate with feature weights W;
9. **if** $T < E$
10. **then** *Weight[i]* ← j;
11. $E \leftarrow T$;

III. Results

For each experiment, database records were partitioned into training and testing sets by subject, according to a uniformly random distribution with a ratio of 1:1. Results presented in the following section are averaged over 20 random partitions.

A. Component Weighting

Information fusion made use of a weighted mean to combine the match scores generated by individual features with weighting calculated on a scale of 0 – 100 according to the algorithm described in Section III. Average component weightings are provided in Table 1, and signify the relative contribution of CEM features for each experiment.

A one-way ANOVA indicated no significant main effect for weighting across stimuli, $F(8, 117) = 0.11$, $p = 0.999$; however, there was a significant main effect for weighting across features, $F(13, 112) = 6.55$, $p < 0.001$.

B. Randomness & Entropy

In order to assess the stability of biometric matching, three statistical tests were performed across the match scores generated for each feature, including: Kolmogorov-Smirnov tests for uniformity and normality, and Wald-Wolfowitz runs tests for randomness. In addition, the Shannon entropy was calculated for the match scores generated for each metric, as a measure of information density.

	Experiment / Stimulus Combination								
	1				2	3			
CEM	*SIM*	*COM*	*COG*	*TEX*	*TEX*	*SIM*	*COM*	*RAN*	*TEX*
1	**49**	**32**	9	**41**	**53**	6	**32**	4	**31**
2	1	**17**	3	**11**	**87**	**10**	**39**	**38**	**33**
3	3	1	2	6	1	7	2	1	2
4	7	1	9	1	**25**	3	3	2	3
5	1	2	**57**	1	2	2	1	2	8
6	**30**	**12**	**10**	**40**	**31**	**100**	**45**	**64**	**43**
7	**14**	**40**	**22**	6	0	**48**	**65**	**70**	**56**
8	1	1	4	0	2	8	0	3	2
9	5	2	1	0	2	1	4	9	2
10	**34**	**52**	**14**	**35**	2	**24**	2	**44**	8
11	**60**	**17**	**10**	0	3	1	**10**	1	1
12	2	1	**16**	1	1	3	2	0	1
13	1	1	1	1	2	2	1	0	2
14	7	**11**	**16**	**16**	**17**	2	**13**	1	3

Table 1. Component Weighting.

	Experiment / Stimulus Combination								
	1				2	3			
CEM	*SIM*	*COM*	*COG*	*TEX*	*TEX*	*SIM*	*COM*	*RAN*	*TEX*
1	**0.90**	**1.06**	**0.80**	**0.91**	**0.80**	0.34	**0.69**	0.49	**0.52**
2	0.22	**0.65**	0.06	0.12	**0.85**	0.47	0.20	**0.57**	**0.58**
3	0.34	0.14	**0.59**	**0.50**	0.42	**0.53**	0.41	0.34	0.30
4	0.36	0.29	**0.60**	**0.52**	**0.58**	**0.53**	0.40	0.23	0.32
5	0.18	0.31	**0.62**	0.10	0.30	0.18	0.30	0.18	0.30
6	**0.67**	**0.73**	**0.53**	**0.63**	**0.64**	**0.79**	**0.60**	**0.71**	**0.57**
7	0.35	0.48	0.38	0.38	0.20	**0.59**	0.49	**0.52**	0.45
8	0.07	0.01	0.14	0.14	0.10	0.48	0.13	0.29	0.11
9	0.24	0.40	0.28	0.16	0.14	0.20	0.32	0.34	0.17
10	**0.52**	**0.65**	0.36	0.41	0.27	0.42	0.33	0.44	0.38
11	**0.78**	**0.79**	**0.66**	0.41	0.37	0.14	**0.58**	0.20	0.34
12	0.20	0.25	0.27	0.21	0.17	0.25	0.27	0.16	0.26
13	0.35	0.04	0.23	0.11	0.01	0.20	0.02	0.12	0.07
14	**0.81**	**0.73**	**0.70**	**0.74**	**0.72**	0.13	**0.62**	0.23	0.43
Fusion	**0.95**	**0.73**	**0.51**	**0.61**	**1.05**	**0.85**	**0.66**	**0.70**	**0.63**

Table 2. d-Prime Values.

	Experiment / Stimulus Combination								
	1				2	3			
CEM	*SIM*	*COM*	*COG*	*TEX*	*TEX*	*SIM*	*COM*	*RAN*	*TEX*
1	**32%**	**31%**	**35%**	36%	**35%**	41%	37%	42%	39%
2	43%	37%	39%	43%	**31%**	40%	41%	39%	38%
3	42%	48%	41%	40%	40%	41%	42%	46%	42%
4	42%	42%	40%	40%	36%	40%	42%	45%	41%
5	48%	48%	**35%**	47%	45%	46%	44%	46%	42%
6	37%	**33%**	37%	36%	37%	**35%**	37%	36%	38%
7	39%	**35%**	40%	41%	43%	37%	37%	36%	38%
8	45%	50%	52%	45%	49%	41%	48%	42%	47%
9	52%	39%	44%	46%	46%	44%	44%	42%	47%
10	36%	**27%**	43%	38%	41%	40%	41%	37%	42%
11	**33%**	**34%**	38%	43%	42%	48%	38%	46%	43%
12	46%	43%	46%	47%	47%	43%	43%	47%	44%
13	43%	51%	44%	49%	49%	45%	50%	48%	48%
14	**35%**	37%	**35%**	38%	37%	48%	39%	48%	42%
Fusion	**31%**	**33%**	37%	37%	**28%**	**34%**	36%	**35%**	36%

Table 3. Equal Error Rates.

	Experiment / Stimulus Combination								
	1				2	3			
CEM	*SIM*	*COM*	*COG*	*TEX*	*TEX*	*SIM*	*COM*	*RAN*	*TEX*
1	18%	**31%**	2%	**29%**	16%	1%	2%	1%	1%
2	22%	6%	10%	**32%**	18%	2%	3%	2%	1%
3	16%	6%	20%	15%	14%	2%	1%	2%	1%
4	13%	8%	6%	18%	11%	2%	2%	1%	1%
5	16%	12%	20%	5%	9%	2%	1%	1%	2%
6	20%	12%	**25%**	14%	10%	3%	2%	0%	2%
7	22%	7%	21%	20%	12%	1%	5%	1%	2%
8	18%	0%	4%	7%	8%	2%	1%	1%	0%
9	19%	10%	11%	16%	16%	3%	0%	2%	1%
10	**29%**	14%	9%	23%	15%	1%	2%	1%	1%
11	23%	12%	9%	13%	10%	0%	2%	1%	1%
12	12%	1%	9%	8%	6%	1%	2%	2%	2%
13	12%	4%	10%	5%	5%	1%	1%	0%	2%
14	23%	10%	6%	17%	14%	1%	2%	1%	1%
Fusion	**53%**	22%	19%	**31%**	**38%**	7%	5%	5%	4%

Table 4. Rank-1 Identification Rates.

For reference, randomness and entropy values were averaged over 1000 iterations for 1000 uniformly and normally distributed random numbers. For uniform random numbers:

- Kolmogorov-Smirnov test for uniformity: $p = 0.5021$.
- Kolmogorov-Smirnov test for normality: $p < 0.0001$.
- Wald-Wolfowitz runs test for randomness: $p = 0.5247$.
- Shannon entropy = 7.8019

For normal random numbers:

- Kolmogorov-Smirnov test for uniformity: $p < 0.0001$.
- Kolmogorov-Smirnov test for normality: $p = 0.4996$.
- Wald-Wolfowitz runs test for randomness: $p = 0.5079$.
- Shannon entropy = 3.9496

For all testing partitions, all experiments, all stimuli, and all features, the calculated match scores are highly non-random. In all cases:

- Kolmogorov-Smirnov test for uniformity: $p < 0.0001$.
- Kolmogorov-Smirnov test for normality: $p < 0.0001$.
- Wald-Wolfowitz runs test for randomness: $p < 0.0001$.

Shannon entropy was slightly more variable, ranging from 5.1 – 7.9, but altogether relatively stable, with average entropy of 7.52 (SD = 0.50). A one-way ANOVA indicated no significant main effect for Shannon entropy across stimuli, $F(8, 126) = 1.67$, $p = 0.112$; however, there was a significant main effect for across features, $F(14, 120) = 4.88$, $p < 0.001$.

C. *d-Prime*

The d-prime value (d′) is a measure of the separation of two distributions, and may be applied to the distributions of genuine and impostor match scores as a measure of achievable error rate tradeoff. The d-prime value of genuine and impostor match score distributions for each experiment are given in Table 2, according to the following equation, where μ_1 and σ_1 indicate the mean and standard deviation, respectively, of genuine match scores, and μ_2 and σ_2 indicate the mean and standard deviation, respectively, of impostor match scores:

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}} \tag{2}$$

A one-way ANOVA indicated no significant main effect for d-prime across stimuli, $F(8, 126) = 0.44$, $p = 0.895$; however, there was a significant main effect for d-prime across features, $F(14, 120) = 12.08$, $p < 0.001$.

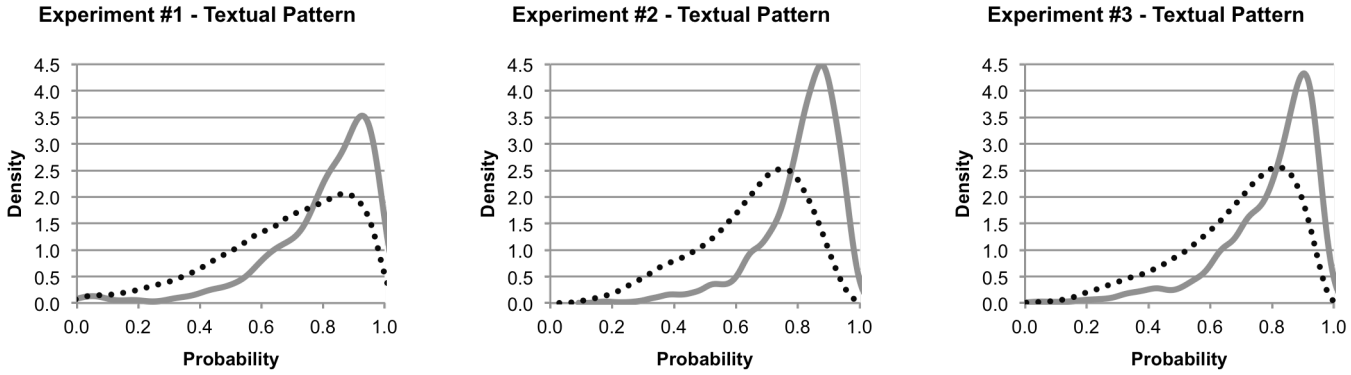


Figure 4. Genuine vs. impostor match score distributions. Probability density functions are based on the kernel smoothing density estimate obtained by MATLAB's ksdensity function. The solid line indicates the genuine distribution, and the dotted line indicates the impostor distribution.

D. *Match Score Distribution*

The distribution of genuine and impostor match scores were averaged across all random partitions and smoothed with a kernel smoothing density estimate. Figure 4 illustrates representative match score distributions for the information fusion of biometric features across the various experiments.

E. *Verification Scenario*

True positive rate (TPR) is defined as the percentage of genuine match scores that exceed the acceptance threshold, false acceptance rate (FAR) is defined as the percentage of impostor match scores that exceed the acceptance threshold, false rejection rate (FRR) is defined as the percentage of genuine match scores that fall below the acceptance threshold, and equal error rate (EER) is the percentage at which FAR and FRR are equal. The receiver operating characteristic (ROC), shown in Figure 5, plots TPR against FAR, and Table 3 provides EER for each experiment.

A one-way ANOVA indicated no significant main effect for EER across stimuli, $F(8, 126) = 0.53$, $p = 0.829$; however, there was a significant main effect across features, $F(14, 120) = 11.76$, $p < 0.001$.

F. *Identification Scenario*

Rank-k identification rate (IR) is the percentage of individuals for which the correct match is found within the top k matches. The cumulative match characteristic (CMC), shown in Figure 6, plots IR by rank, and Table 4 provides rank-1 identification rates for each experiment.

A one-way ANOVA indicated a significant main effect for rank-1 identification rate across stimuli, $F(8, 126) = 21.22$, $p < 0.001$; however, there was no significant main effect across features, $F(14, 120) = 1.63$, $p = 0.079$.

IV. Discussion

In the current paper, we are interested in several related topics, the viability of CEM features and techniques in a biometric context and the effects of stimulus type, spatial accuracy, and temporal resolution on the biometric viability of complex eye movement patterns.

A. *Biometric Viability*

From the distribution of genuine vs. impostor match scores, we identify a large overlap in all cases. While they represent distinct distributions, it is difficult to separate the two. Despite this overlap, match scores were highly non-random, as indicated by Kolmogorov-Smirnov tests for uniformity and normality, and the Wald-Wolfowitz runs test.

When comparing entropy and d-prime, there appears to be a tendency for metrics with lower Shannon entropy to have a correspondingly low d-prime. In fact, Shannon entropy and d-prime were strongly correlated, $r(133) = 0.39$, $p < 0.001$. Low entropy suggests clustering of match scores around some common values, and this clustering may reduce the distance between genuine and impostor distributions.

We notice that the d-prime of information fusion is often lower than for individual features. This is a result of feature weighting being selected on the training set, and indicates that the considered information fusion algorithm does a poor job at predicting component weights based on a subset of features.

When compared to the state-of-the-art in iris (0.0011% EER [1]), fingerprint (2.07% EER [2]), and facial geometry (15% EER [4]) biometrics, it is obvious that eye movement biometrics are not yet suitable for standalone applications; however, it is important to remember that, at the time or writing, this course of study has existed for less than a decade, while many of todays biometric standards have existed for over a century [5].

B. *The Effects of Stimulus Type*

While there were obvious differences between CEM features, there was no discernible effect from the stimuli used to evoke eye movements. This is confirmed by ANOVA results from component weighting, Shannon entropy, d-prime, and equal error rates. Receiver operating characteristics and cumulative match characteristics lead to similar conclusions.

In considering the receiver operating characteristics and cumulative match characteristics, it is important to note that in the 1st experiment the number of recording sessions for each stimulus was inconsistent. This lead to variation in the number of acceptance/rejection comparisons during the verification scenario, and maximum rank during the identification scenario. This issue was rectified in the 3rd experiment, and the effects are obvious in the respective plots.

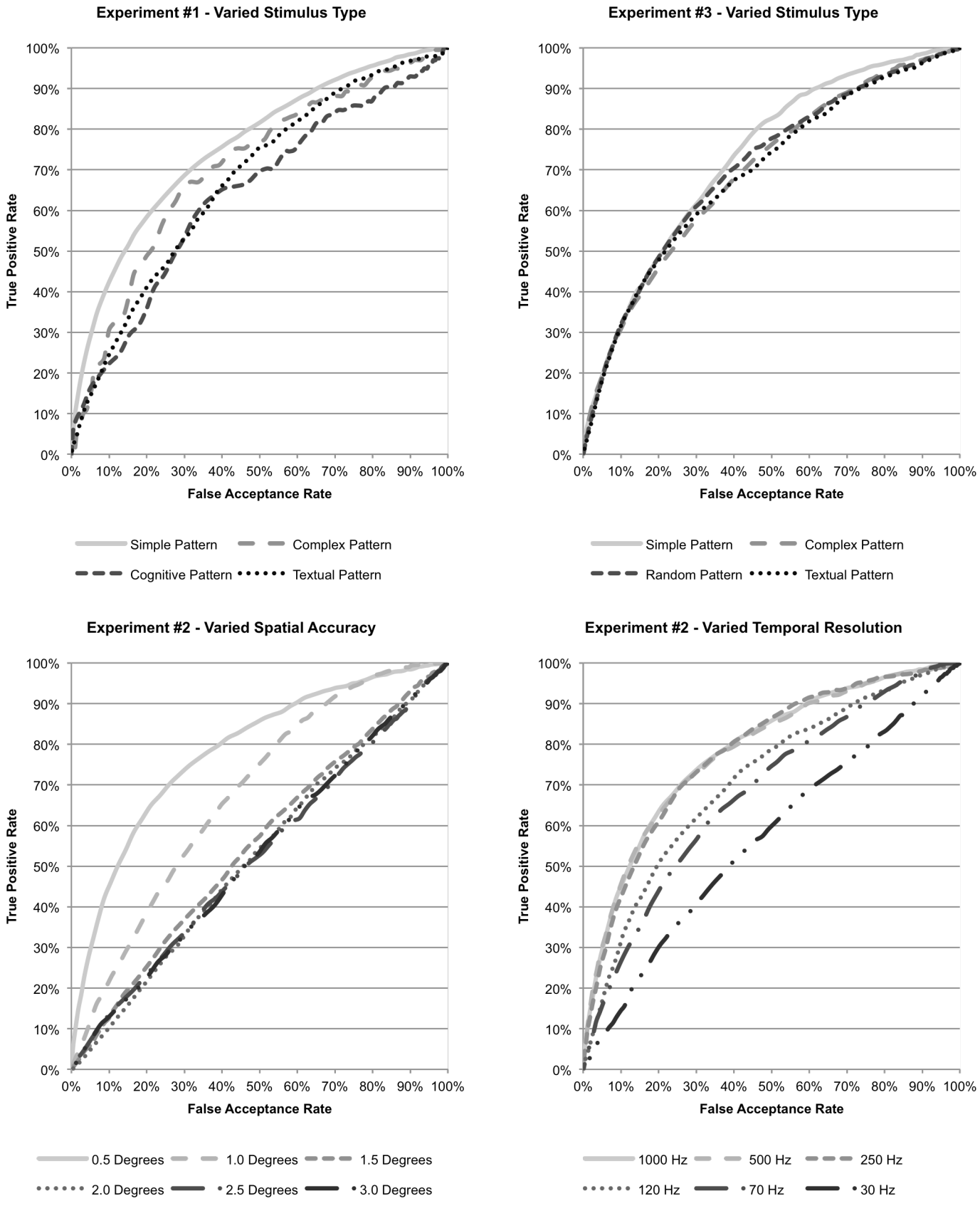


Figure 5. Receiver operating characteristic (ROC) for varied stimulus type, spatial accuracy, and temporal resolution. True positive rate is defined as the percentage of genuine match scores that exceed the acceptance threshold, and false acceptance rate is defined as the percentage of impostor match scores that exceed the acceptance threshold.

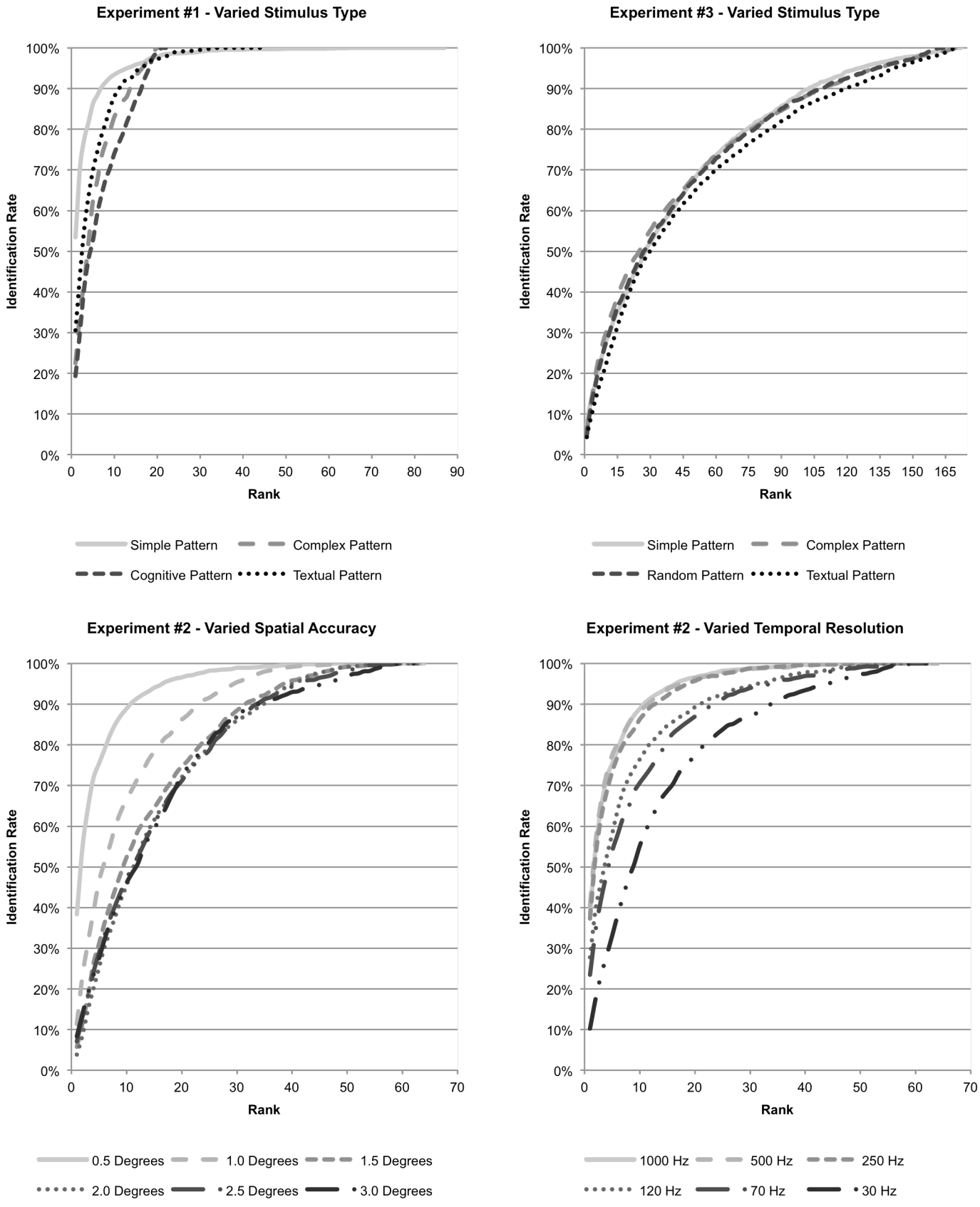


Figure 6. Cumulative match characteristic (CMC) for varied stimulus type, spatial accuracy, and temporal resolution. Rank-k identification rate is the percentage of individuals for which the correct match is found within the top k matches.

C. The Effects of Spatial Accuracy

There is a stark contrast between spatial accuracy tiers that is immediately obvious from the ROC and CMC curves. From the baseline of 0.5° to the next accuracy tier of 1.0°, biometric accuracy is substantially reduced in both verification and identification scenarios, with equal error rate increasing from 28% to 38% and rank-1 identification rate reduced from 38% to 11%. At all spatial accuracy tiers beyond 1.0°, biometric accuracy in both verification and identification scenarios is virtually indistinguishable, and is essentially random.

It should be noted that 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 that 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.

D. Further Analysis

Further analysis was conducted on the eye movement recordings generated in the 3rd experiment (due to the larger subject pool) to identify any discernible effect of age or gender. In the case of gender, there was no obvious difference in the error rates obtained by male (34% EER, 7% IR) and female subjects (36% EER, 10% IR). For age, the recordings were split into two groups of approximately equal size, subjects 20 years of age and under, and subjects older than 20 years of age. In this case, the results suggest that the considered techniques produced more accurate identification on the older age group (32% EER, 9% IR) than on the younger age group (36% EER, 7% IR).

To assess the liveness detection properties of the considered techniques, artificial eye movement recordings were generated as in Komogortsev and Karpov's assessment of OPC biometrics [28]. An SVM was applied at the feature level to classify individual recordings as human or spoof, utilizing leave-one-out cross-validation to generate liveness scores. On the eye movement recordings of the 2nd experiment, CEM-P was able to achieve 0% false live rejection rate, 0% false spoof acceptance rate, and 100% classification accuracy. For comparison, on the same dataset, with the same eye model, the OPC technique was only able to achieve 27% false live rejection rate, 4% false spoof acceptance rate, and 85% classification accuracy [28].

In order to estimate the computational performance of the proposed techniques, profiling was performed and averaged over 1000 iterations on a dataset of 100 subjects, with 2 sessions per subject (a subset of the data from the 3rd experiment). The Sensor module required an average 11.5 seconds to load and parse all 200 recordings; that is, 57.5 milliseconds per recording. The Feature Extraction module required 10.3 seconds to classify and merge fixations and saccades across all 200 recordings; or, 51.5 milliseconds per recording. The Matching module required 36.8 seconds to perform comparison, information fusion, and calculate match scores for all 19900 recording combinations; or, roughly 1.8 milliseconds per comparison. It should be noted that these results were obtained using single-threaded execution on a 2.7 GHz processor with 16 GB of RAM; however, these algorithms are easily parallelized for increased performance.

E. Future Research

Perhaps the most important step toward improving the accuracy of eye movement biometrics is large-scale data collection. The NIST and related organizations have spent years collecting and distributing thousands of biometric templates for standards such as fingerprint and iris, providing a common benchmark from which to measure the accuracy of new algorithms. One of the major disadvantages of eye movements is the need to collect data over time. This property effectively limits the number of biometric templates that can be captured in a given time period, and has lead to a number of disjoint datasets collected by various researchers. Access to large, open, and freely available datasets is essential to the advancement of eye movement biometrics.

Another major concern is the continued development of mathematical models of human eye movement. At present, there do not exist any models of the oculomotor plant that are capable of reproducing all forms of human eye movement with absolute accuracy. As our mechanical understanding of the human visual system increases, so too will the biometric accuracy of OPC-based techniques that attempt to estimate physical constants from measurable properties; however, at the same time, as these models become more sophisticated it will become necessary to invest additional effort into detecting and identifying spoofed recordings.

Beyond this, the most immediate necessity is the refinement of existing biometric techniques, and the development of further eye movement features. For example, it may be possible to improve biometric accuracy by utilizing the properties of: corrective eye movements, such as glissades and saccadic dysmetria; task-specific eye movements, such as within-word fixation distributions and regressive saccades in reading; and more complex eye movement types, which are typically more difficult to reproduce and identify, such as smooth pursuits and optokinetic reflex.

There are, further, a number of aspects to biometric performance that must be examined. Among these, the effects of fatigue, caffeine, tobacco, and alcohol on eye movements are well documented [8], and additional research will be necessary to quantify their impact on the biometric viability of eye movements. Similarly, template aging and permanence studies are still necessary to identify the rate and magnitude of degradation over time for the biometric signatures produced by eye movement biometrics. The development of quality metrics, with which to identify and ignore unsuitable recordings, is still necessary to ensure that enrollment and authentication are not skewed by noise; and more rigorous liveness detection research is necessary to confirm the accuracy, or inaccuracy, of eye movements in this regard.

Finally, we must examine the hardware itself. As the frame rate and resolution of general-purpose cameras increases, so will the accuracy and efficiency of video-oculography techniques, thereby lower the cost and increasing the accuracy of eye movement biometrics. As face and iris detection are already key components of most video-oculography systems, it is only a short leap to fast and cost-effective multi-modal systems that incorporate face, iris, and eye movement biometrics through a single sensor.

V. Conclusion

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

Based on the results, we conclude that stimulus type has little, if any, effect on the accuracy of eye movement-based biometrics. Eye tracking systems with spatial accuracy of less than 0.5° and greater than 250 Hz temporal resolution 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] Survey, "NIST Test Results Unveiled," *Biometric Technology Today,* vol. 15, pp. 10-11, 2007.

[2] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "FVC2004: Third Fingerprint Verification Competition.," presented at the International Conference on Biometric Authentication (ICBA), Hong Kong, China, 2004.

[3] I. H. G. S. Consortium, "Initial Sequencing and Analysis of the Human Genome," *Science,* vol. 409, pp. 860-921, 2001.

[4] P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe, "FRVT 2006 and ICE 2006 Large-Scale Experimental Results," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, pp. 831-846, 2010.

[5] P. D. Komarinski, *Automated Fingerprint Identification Systems (AFIS)*, 1 ed.: Academic Press, 2004.

[6] A. K. Jain, P. Flynn, and A. A. Ross, Eds., *Handbook of Biometrics*. Springer US, 2007, pp. 1-556.

[7] D. D. Zhang, *Automated Biometrics: Technologies and Systems*, 1 ed.: Springer, 2000.

[8] R. J. Leigh and D. S. Zee, *The Neurology of Eye Movements*, 4 ed. Oxford, NY, USA: Oxford University Press, 2006.

[9] J. S. Agustin, H. Skovsgaard, J. P. Hansen, and D. W. Hansen, "Low-cost Gaze Interaction: Ready to Deliver the Promises," in *Conference on Human Factors in Computing (CHI)*, Boston, MA, USA, 2009, pp. 4453-4458.

[10] ITIRT, "Independent Testing of Iris Recognition Technology: Final Report," International Biometric Group2005.

[11] C. M. Roberts, "Biometric Attack Vectors and Defences," *Computers & Security,* vol. 26, pp. 14-25, 2006.

[12] O. V. Komogortsev, A. Karpov, C. D. Holland, and H. Proença, "Multimodal Ocular Biometrics Approach: A Feasibility Study," in *Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Washington DC, USA, 2012, pp. 1-8.

[13] D. Noton and L. W. Stark, "Scanpaths in Eye Movements during Pattern Perception," *Science,* vol. 171, pp. 308-311, 1971.

[14] P. Kasprowski and J. Ober, "Eye Movements in Biometrics," in *European Conference on Computer Vision (ECCV)*, Prague, Czech Republic, 2004, pp. 248-258.

[15] D. L. Silver and A. J. Biggs, "Keystroke and Eye-Tracking Biometrics for User Identification," in *International Conference on Artificial Intelligence (ICAI)*, Las Vegas, NV, USA, 2006, pp. 344-348.

[16] C. D. Holland and O. V. Komogortsev, "Biometric Identification via Eye Movement Scanpaths in Reading," in *International Joint Conference on Biometrics (IJCB)*, Washington, D.C., 2011, pp. 1-8.

[17] O. V. Komogortsev, A. Karpov, L. R. Price, and C. R. Aragon, "Biometric Authentication via Oculomotor Plant Characteristics," in *International Conference on Biometrics (ICB)*, New Delhi, India, 2012, pp. 1-8.

[18] I. Rigas, G. Economou, and S. Fotopoulos, "Biometric Identification Based on the Eye Movements and Graph Matching Techniques," *Pattern Recognition Letters,* vol. 33, pp. 786-792, 2012.

[19] C. D. Holland and O. V. Komogortsev, "Complex Eye Movement Pattern Biometrics: Analyzing Fixations and Saccades," presented at the IAPR International Conference on Biometrics (ICB), Madrid, Spain, 2013.

[20] J. E. Cutting and L. T. Kozlowski, "Recognizing Friends by Their Walk: Gait Perception Without Familiarity Cues," *Bulletin of the Psychonomic Society,* vol. 9, pp. 353-356, 1977.

[21] A. F. Bobick and A. Y. Johnson, "Gait Recognition Using Static, Activity-Specific Parameters," presented at the Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, HI, USA, 2001.

[22] A. Kale, A. N. Rajagopalan, N. Cuntoor, and V. Krüger, "Gait-based Recognition of Humans Using Continuous HMMs," presented at the International Conference on Automatic Face and Gesture Recognition (FGR), Washington, DC, USA, 2002.

[23] L. Lee and W. E. L. Grimson, "Gait Analysis for Recognition and Classification," presented at the International Conference on Automatic Face and Gesture Recognition (FGR), Washington, DC, USA, 2002.

[24] H. Chan and W. W. Bledsoe, "A Man-Machine Facial Recognition System; Some Preliminary Results," Panoramic Research Inc., Palo Alto, CA, USA1965.

[25] S. A. Rizvi, P. J. Phillips, and H. Moon, "The FERET Verification Testing Protocol for Face Recognition Algorithms," presented at the International Conference on Automatic Face and Gesture Recognition (FGR), Nara, Japan, 1998.

[26] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, pp. 1090-1104, 2000.

[27] O. V. Komogortsev, A. Karpov, and C. D. Holland, "CUE: Counterfeit-resistant Usable Eye-based Authentication via Scanpaths and Oculomotor Plant Characteristics," in *SPIE Defense, Security, and Sensing*, Baltimore, Maryland, USA, 2012, pp. 1-10.

[28] O. V. Komogortsev and A. Karpov, "Liveness Detection via Oculomotor Plant Characteristics: Attack of Mechanical Replicas," presented at the IEEE/IAPR International Conference on Biometrics (ICB), Madrid, Spain, 2013.

[29] C. D. Holland and O. V. Komogortsev, "Biometric Verification via Complex Eye Movements: The Effects of Environment and Stimulus," in *Fifth Internation Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Washington DC, USA, 2012, pp. 1-8.

[30] J. H. Goldberg and X. P. Kotval, "Computer Interface Evaluation using Eye Movements: Methods and Constructs," *International Journal of Industrial Ergonomics,* vol. 24, pp. 631-645, 1999.

[31] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 2 ed. London, UK: Springer-Verlag, 2007.

[32] W. F. Asaad, G. Rainer, and E. K. Miller, "Task-Specific Neural Activity in the Primate Prefrontal Cortex," *Journal of Neurophysiology,* vol. 84, pp. 451-459, 2000.

[33] S. E. Palmer, *Vision Science: Photons to Phenomenology*, 1 ed. Cambridge, MA, USA: MIT Press, 1999.

[34] S. Zeki, *A Vision of the Brain*. Cambridge, MA, USA: Blackwell Scientific Publications, 1993.

[35] T. Ro, J. Pratt, and R. D. Rafal, "Inhibition of Return in Saccadic Eye Movements," *Experimental Brain Research,* vol. 130, pp. 264-268, 2000.

[36] O. V. Komogortsev. (2012). *Eye Movement Biometric Database v2*. Available: http://cs.txstate.edu/~ok11/embd_v2.html

[37] H. Rorschach, *Psychodiagnostics: A Diagnostic Test Based on Perception*, 5 ed.: Verlag Hans Huber, Bern, 1942.

[38] K. Renaud and A. d. Angeli, "Visual Passwords: Cure-All or Snake Oil?," *Communications of the ACM,* vol. 52, pp. 135-140, 2009.

[39] O. V. Komogortsev, D. V. Gobert, U. K. S. Jayarathna, D. H. Koh, and S. M. Gowda, "Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors," *IEEE Transactions on Biomedical Engineering,* vol. 57, pp. 2635-2645, 2010.

[40] A. A. Ross and A. K. Jain, "Information Fusion in Biometrics," *Pattern Recognition Letters,* vol. 24, pp. 2115-2125, 2003.