Complex Eye Movement Pattern Biometrics: Analyzing Fixations and Saccades

Corey D. Holland and Oleg V. Komogortsev

Abstract

This paper presents an objective evaluation of previously unexplored biometric techniques utilizing patterns identifiable in human eye movements to distinguish individuals. The distribution of primitive eye movement features are compared between eye movement recordings using algorithms based on the following statistical tests: the Ansari-Bradley test, the Mann-Whitney U-test, the two-sample Kolmogorov-Smirnov test, the two-sample t-test, and the two-sample Cramér-von Mises test. Score-level information fusion is applied and evaluated by: weighted mean, support vector machine, random forest, and likelihood ratio. The accuracy of each comparison/fusion algorithm is evaluated, with results suggesting that, on high resolution eye tracking equipment, it is possible to obtain equal error rates of 16.5% and rank-1 identification rates of 82.6% using the two-sample Cramér-von Mises test and score-level information fusion by random forest, the highest accuracy results on the considered dataset.

1. Introduction

Biometrics is, in essence, a science of people. Yet, instead of studying what brings us together, biometrics is concerned with what sets us apart [1]. The systematic collection of physical and behavioral characteristics for the purposes of identification dates as far back as 1858, with the collection of handprints to identify workers [2], and has since expanded to include such features as: fingerprints [3], palm prints, action, hand geometry, facial structure [4], signature, speech, gait, and iris patterns [5].

Despite this long history, automated biometric systems have only become feasible in the recent past [6], and are limited by several factors: accuracy, counterfeit-resistance, speed, and cost. The human visual system offers a number of unique properties that make it an interesting candidate in this respect. Eye movements are highly dependent upon both physical and neurological properties [7], making them highly counterfeit-resistant, and may be recorded and processed by an unmodified camera in real-time [8], making their application both cheap and efficient.

1.1. Human Visual System

The human visual system is composed of two major functional units, the oculomotor plant and brainstem control [7]. The oculomotor plant consists primarily of the eye globe, surrounding tissue, and six extraocular muscles: the lateral and medial recti, responsible for horizontal rotation; the superior and inferior recti, responsible for vertical rotation; and the superior and inferior obliques, responsible for torsional rotation. The brainstem control innervates extraocular muscles with burst/omnipause neurons, causing muscle contractions and relaxations.

The neuronal control signal is ultimately responsible for producing the many and varied human eye movements [9]. Among the various eye movement types, 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. Over time, fixations and saccade form a spatial path over a region, known as the scanpath.

1.2. Previous Research

In 2004, Kasprowski and Ober [10] lead the investigation of eye movements as biometric indicators. Applying techniques commonly used in voice recognition, Kasprowski and Ober examined the first 15 cepstral coefficients of the positional eye movement signal using naïve Bayes classifiers, C4.5 decision trees, SVM polynomials, and KNN (k = 3 and k = 7). Average false acceptance rate and false rejection rate were reported for each technique, but the lack of equal error rates and identification rates made it difficult to draw conclusions. The best results obtained an average 1% false acceptance rate and 23% false rejection rate.

In 2006, Silver and Biggs [11] investigated a set of higher-level eye movement features in conjunction with keystroke biometrics. Match scores were derived from fixation count, the 8 most significant fixations, average fixation duration, average saccade velocity, average

saccade duration, and average vertical position, combined using a probabilistic neural network. The best results obtained an average 66% true positive rate and 98% true negative rate, equal error rates and identification rates were not provided.

In 2011, Holland and Komogortsev [12] examined a much larger set of high-level eye movement features, including: fixation count, average fixation duration, average vectorial saccade amplitude, average horizontal saccade amplitude, average vertical saccade amplitude, average vectorial saccade velocity, average vectorial saccade peak velocity, velocity waveform indicator, scanpath length, scanpath convex hull area, mean-shift regions of interest, inflection count, amplitude-duration coefficient, and amplitude-peak velocity coefficient. Features were compared using a Gaussian kernel and combined with a weighted mean. The best results obtained a 27% equal error rate, identification rates were not provided.

In 2012, Rigas et al. [13] evaluated the use of graph-based matching on the positional signal of an eye movement recording, comparing minimum spanning trees using the multivariate Wald-Wolfowitz runs test. The best results obtained 70% rank-1 identification rate and 30% equal error rate.

1.3. Motivation & Hypothesis

Biometric recognition via eye movements is still in its infancy, and has not yet reached the level of accuracy required to compete with accepted biometric standards. While this may be the case, the body of evidence suggests that eye movements do convey biometric information that may be used to uniquely identify an individual. The challenge, then, is to identify those traits that provide the highest level of biometric accuracy, and refine the techniques used to process them.

The current work expands on the complex eye movement (CEM) framework proposed by Holland and Komogortsev [12], which made use of averages and aggregate features, and explored only basic fusion techniques. We hypothesize that the distribution of primitive features inherent in basic eye movements can be utilized to uniquely identify a given individual. In examining this hypothesis we evaluate several comparison algorithms based on statistical tests for comparing distributions, including: the two-sample t-test, the Ansari-Bradley test, the Mann-Whitney U-test, the two-sample Kolmogorov-Smirnov test, and the two-sample Cramér-von Mises test. We proceed to investigate information fusion techniques including score-level fusion by: weighted mean, support vector machine, random forest, and likelihood ratio.
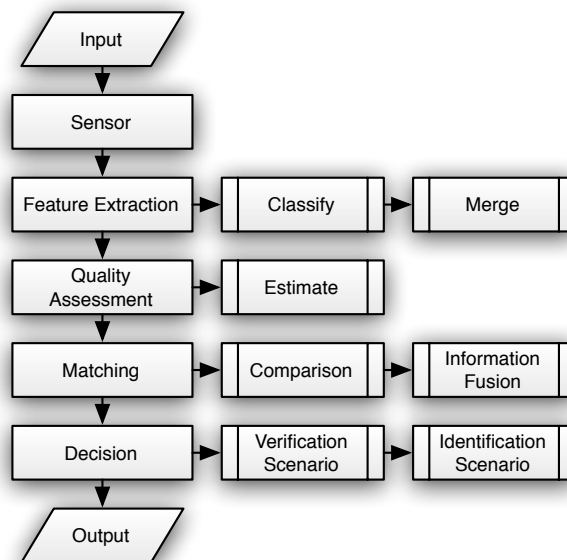

Figure 1: Biometric framework diagram.

2. CEM Biometric Framework

The biometric techniques described in this paper, shown in Figure 1, extend from those described by Holland and Komogortsev [14]. The Sensor module processes the eye movement signal, the Feature Extraction module identifies, filters, and merges individual gaze points into fixations and saccades, the Quality Assessment module assesses the biometric viability of each recording, the Matching module generates training/testing sets and compares individual recordings, and the Decision module calculates error rates under biometric verification and identification scenarios [1].

2.1. Sensor Module

The Sensor module parses individual eye movement recordings, combining available left/right eye coordinates and removing invalid data points from the eye movement signal. Eye movement recordings are stored in memory as an eye movement database, with the eye movement signal linked to the experiment, trial, and subject that generated the recording.

2.2. Feature Extraction Module

The Feature Extraction module generates feature templates for each record in the eye movement database. Eye movement features are primarily composed of fixations and saccades. The eye movement signal is parsed to identify fixations and saccades using an eye movement classification algorithm [15], followed by micro-saccade and micro-fixation filters.

Fixation and saccade groups are merged, identifying fixation-specific and saccade-specific features. Fixation features include: start time, duration, horizontal centroid, and vertical centroid. Saccade features include: start time, duration, horizontal amplitude, vertical amplitude, average horizontal velocity, average vertical velocity, horizontal peak velocity, and vertical peak velocity.

2.3. Quality Assessment Module

The Quality Assessment module identifies the biometric viability of the generated feature templates. In this context, we utilize the fixation quantitative score, ideal fixation quantitative score, fixation qualitative score, and saccade quantitative score [16] as tentative measure of the quality of features obtained from the recording. It should be noted, however, that these measures are currently untested in this context, and their evaluation is beyond the scope of the current paper.

2.4. Matching Module

The Matching module compares individual records, generating match scores for various metrics using comparison algorithms that operate on feature templates. In this case, comparison algorithms operate to compare the distribution of fixation- and saccade-based features throughout each record. Match scores from each comparison algorithm are then combined into a single match score with an information fusion algorithm [17].

The Matching module partitions records, splitting the database into training and testing sets by subject, according to a uniformly random distribution. Comparison and information fusion thresholds and parameters are generated on the training set, while error rates are calculated on the testing set.

2.5. Decision Module

The Decision module calculates error rates for comparison and information fusion under biometric verification and identification scenarios. Under the verification scenario, each record in the testing set is compared to every other record in the testing set exactly once, and false acceptance rate and true positive rate are calculated at varied acceptance thresholds. Under the identification scenario, every record in the testing set is compared to every other record in the testing set, and identification rates are calculated from the largest match score(s) from each of these comparison sets.
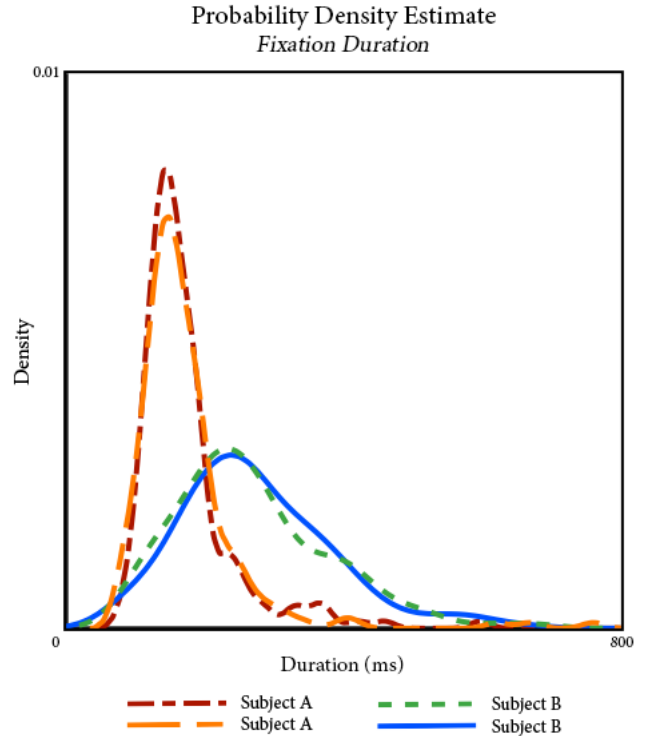

Figure 2: Comparative distribution of fixation duration across multiple recording sessions.

3. CEM Biometrics

In this paper, we are primarily interested in the biometric viability of primitive eye movement features:

M1. Start time (fixation)
M2. Duration (fixation)
M3. Horizontal centroid (fixation)
M4. Vertical centroid (fixation)
M5. Start time (saccade)
M6. Duration (saccade)
M7. Horizontal amplitude (saccade)
M8. Vertical amplitude (saccade)
M9. Horizontal mean velocity (saccade)
M10. Vertical mean velocity (saccade)
M11. Horizontal peak velocity (saccade)
M12. Vertical peak velocity (saccade)

These features accumulate over the course of a recording, as the scanpath is generated. By analyzing the distribution of these features throughout each recording, as shown in Figure 2, we are effectively examining the behavior of the scanpath as a whole. At the same time, by considering the fixations and saccades that compose the scanpath, we remove much of the unavoidable signal noise from the raw eye movement signal, and reduce the dataset to a computationally manageable size.

In order to compare the distribution of primitive eye movement features, we employ several statistical tests, detailed in Section 3.1, which operate on different aspects of the distribution. These statistical tests are applied as a comparison algorithm to the distributions of each feature separately, generating 12 match scores altogether. The information fusion algorithms, described in Section 3.2, are applied to the match scores generated by each comparison algorithm to produce a single match score used for biometric authentication.

3.1. Comparison Algorithms

(C1) Two-Sample t-Test

The two-sample t-test measures the probability that observations from two recordings are taken from normal distributions with equal mean and variance.

(C2) Ansari-Bradley Test

The Ansari-Bradley test measures the probability that observations from two recordings with similar median and shape are taken from distributions with equivalent dispersion.

(C3) Mann-Whitney U-Test

The Mann-Whitney U-test measures the probability that observations from two recordings are taken from continuous distributions with equal median.

(C4) Two-Sample Kolmogorov-Smirnov Test

The two-sample Kolmogorov-Smirnov test measures the probability that observations from two recordings are taken from the same continuous distribution, measuring the distance between empirical distributions.

(C5) Two-Sample Cramér-von Mises Test

The two-sample Cramér-von Mises test measures the probability that observations from two recordings are taken from the same continuous distribution, measuring the goodness-of-fit between empirical distributions.

3.2. Information Fusion Algorithms

(F1) Weighted Mean

The weighted mean algorithm combines the match scores produced for individual metrics into a single match score on the interval [0, 1]. The genuine and imposter match score vectors of the training set are used to select per-metric weighting which minimizes equal error rate via iterative optimization, and the weighted mean produces a single match score as a linear combination of the match scores for each metric.

(F2) Support Vector Machine

The support vector machine algorithm classifies the match scores produced for individual metrics into a single match score in the set {0, 1}. The support vector machine builds a 7th order polynomial on the genuine and imposter match score vectors of the training set, and match scores are classified by dividing them into categories separated by the polynomial on an n-dimensional hyperplane [18].

(F3) Random Forest

The random forest algorithm combines the match scores produced for individual metrics into a single match score on the interval [0, 1]. An ensemble of 50 regression trees is built on the genuine and imposter match score vectors of the training set, and the random forest calculates the combined match score based on a set of conditional rules and probabilities [19].

(F4) Likelihood Ratio

The likelihood ratio algorithm combines the match scores produced for individual metrics into a single match score on the interval [0, ∞). The genuine and imposter match score vectors of the training set are modeled using Gaussian mixture models, and the likelihood ratio is calculated as the ratio of the genuine probability density over the imposter probability density [20].

4. Methodology

Previous research has shown that minor variations in eye tracking specifications, such as spatial accuracy and temporal resolution, can have a substantial impact on the biometric viability of eye movements [14]. Therefore, to properly evaluate the proposed techniques, it was deemed necessary to examine biometric accuracy on both high- and low-resolution eye tracking systems. Existing eye movement datasets, collected by Komogortsev et al. [14, 21], were utilized for comparative evaluation, with collection methodology in the following subsections. The high-resolution dataset is freely available via [21]; the low-resolution dataset is not currently available to the research community, but there are plans to release it in the near future.

4.1. Apparatus & Software

High-resolution eye movements were recorded using an EyeLink 1000 eye tracking system [22], 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 66% (SD = 36%). Stimuli were presented on a flat screen monitor positioned at a distance of 685 millimeters from each subject, with screen dimensions of 640×400 millimeters, and screen resolution of 2560×1600 pixels.

Low-resolution eye movements were recorded using a modified version of the open-source ITU Gaze Tracker software [23] and PlayStation Eye Camera [24], with a temporal resolution of 75 Hz and average calibration

accuracy of 1.1° (SD = 0.8). Average data validity is not reportable, 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 the glasses, eyelashes, hair). Stimuli were presented on a flat screen monitor positioned at a distance of 540 millimeters from each subject, with screen dimensions of 375×302 millimeters, and screen resolution of 1280×1024 pixels.

In both cases, the pupil was illuminated by infrared LED to improve eye tracking accuracy, and a chin rest was employed to improve stability. Stimulus presentation was consistent across experiments, with only minor changes required for varied screen dimensions. All algorithms and data analysis were implemented and performed in MATLAB.

4.2. Participants

High-resolution 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.

Low-resolution 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). 169 of the subjects performed 2 recordings each, and 3 of the subjects performed 1 recording each, generating a total of 341 unique eye movement recordings.

4.3. Procedure

Eye movement recordings were generated on both high-resolution and low-resolution eye tracking systems using a textual stimulus pattern. The text of the stimulus was taken from Lewis Carroll's poem, "The Hunting of the Snark," chosen for its difficult and nonsensical content, forcing readers to progress slowly and carefully through the text.

For each recording session, subjects were limited to 1 minute of reading. To reduce learning effects, subjects were given a different excerpt from the text for each recording session and each excerpt was selected to ensure that line lengths and the difficulty of material were consistent. As well, excerpts were selected to require approximately 1 minute of active reading.

Eye movements were processed with the biometric framework described in Section 2, with eye movement classification thresholds: velocity threshold of 20°/sec, micro-saccade threshold of 0.5°, and micro-fixation threshold of 100 milliseconds. Feature extraction was performed across all eye movement recordings, while matching and information fusion were performed according to the methods described in Section 3. To assess biometric accuracy, error rates were calculated under both verification and identification scenarios.

5. Results

Eye movement recordings were partitioned, by subject, into training and testing sets according to a uniformly random distribution with a ratio of 1:1, such that no subject had recordings in both the training and testing sets. Experimental results are averaged over 80 random partitions for each metric, and 20 random partitions for each fusion algorithm. Scores for the best performing algorithms are highlighted for readability.

5.1. Verification Scenario

False acceptance rate is defined as the rate at which imposter scores exceed the acceptance threshold, false rejection rate is defined as the rate at which genuine scores fall below the acceptance threshold, and true positive rate is defined as the rate at which genuine scores exceed the acceptance threshold. The equal error rate, given in Table 1, is the rate at which false acceptance rate and false rejection rate are equal, and the receiver operating characteristic, shown in Figure 3 for the best performing fusion methods, plots true positive rate against false acceptance rate.

5.2. Identification Scenario

Identification rate is defined as the rate at which enrolled subjects are successfully identified as the correct individual, where rank-k identification rate is the rate at which the correct individual is found within the top k matches. The rank-1 identification rate, given in Table 2, is the rate at which the correct individual has the highest match score, and the cumulative match characteristic, shown in Figure 4 for the best performing fusion methods, plots identification rate by rank across all ranks, where the maximum rank is equivalent to the available comparisons.

6. Discussion

Of the considered comparison algorithms, the Kolmogorov-Smirnov and Cramér-von Mises tests have a clear advantage in terms of accuracy. While the two are practically equivalent on the high-resolution dataset, the Kolmogorov-Smirnov test appears to hold up better on the low-resolution dataset, based on receiver operating and cumulative match characteristics. Despite this, the Cramér-von Mises test provided the overall lowest equal error rate (16.5%) on the high-resolution dataset.
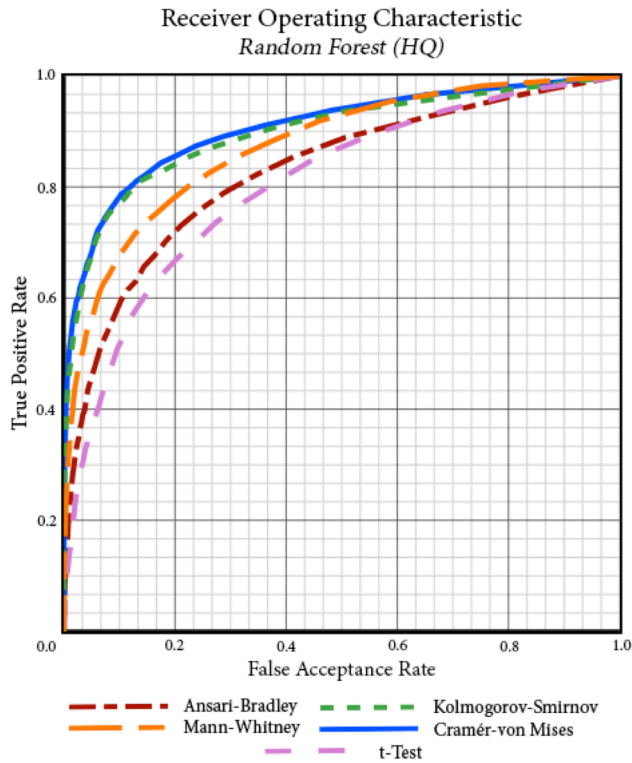
Figure 3a: Receiver Operating Characteristic (High-Resolution).
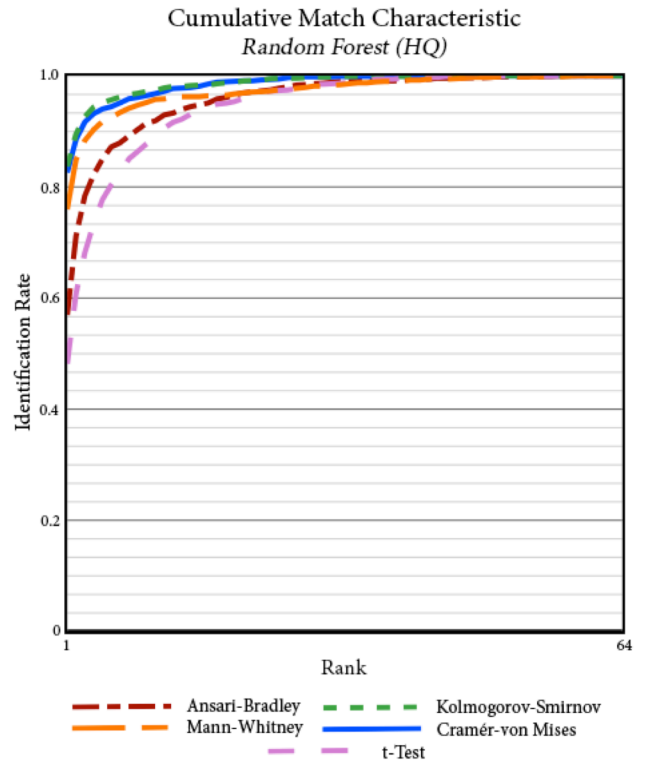

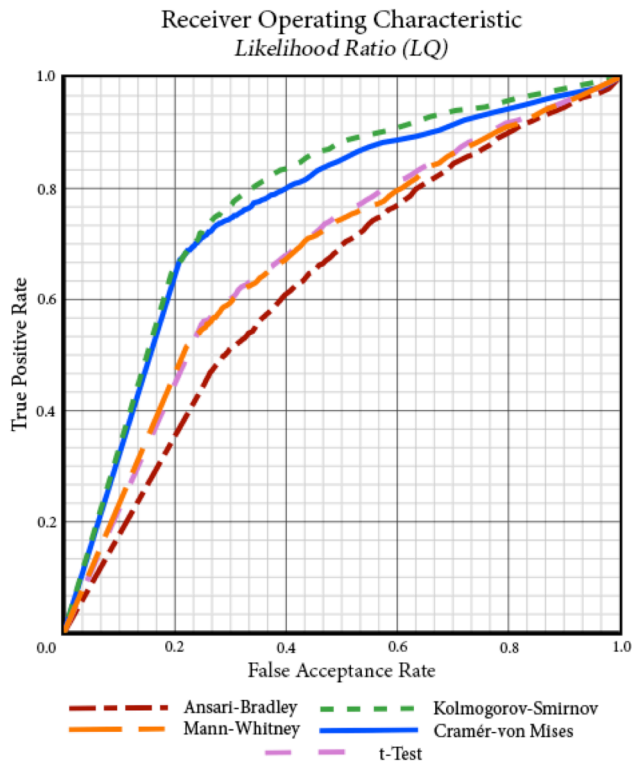Figure 4a: Cumulative Match Characteristic (High-Resolution).

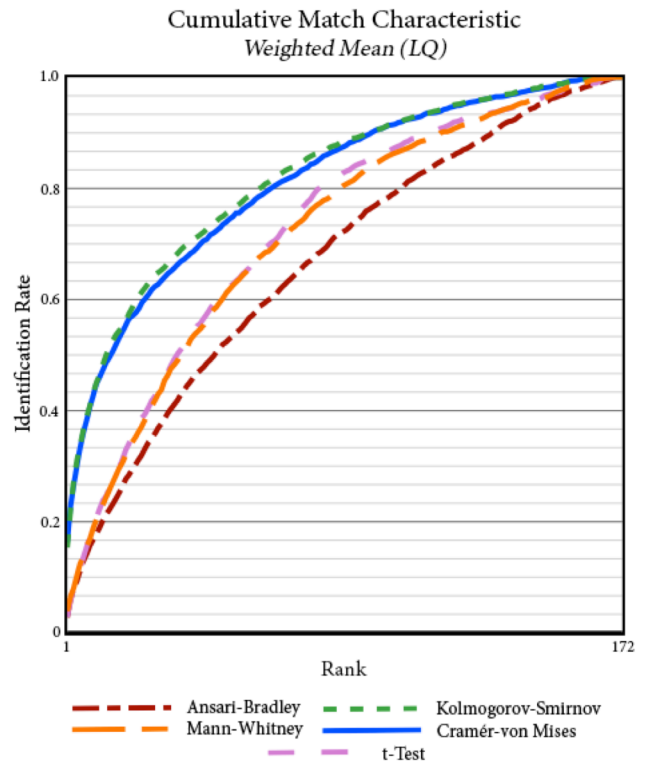Figure 3b: Receiver Operating Characteristic (Low-Resolution).

Figure 4b: Cumulative Match Characteristic (Low-Resolution).

	High-Resolution Dataset				
Metric	C1	C2	C3	C4	C5
M1	33.5%	34.0%	34.8%	**32.9%**	34.3%
M2	27.6%	37.3%	24.5%	23.0%	**22.4%**
M3	42.7%	47.1%	42.4%	**38.5%**	39.4%
M4	42.8%	47.6%	41.9%	44.0%	**41.7%**
M5	33.9%	33.8%	34.8%	**33.3%**	33.9%
M6	43.3%	38.1%	31.9%	**27.4%**	27.6%
M7	44.8%	30.8%	38.5%	26.3%	**25.7%**
M8	43.2%	38.3%	36.5%	**30.5%**	30.9%
M9	36.4%	35.1%	35.8%	29.1%	**27.2%**
M10	45.5%	40.2%	39.0%	**33.2%**	33.7%
M11	42.0%	50.0%	29.6%	25.9%	**25.5%**
M12	45.1%	40.3%	37.4%	**28.8%**	30.4%
F1	30.6%	35.9%	24.4%	21.5%	**19.4%**
F2	39.9%	35.4%	30.4%	26.6%	**25.6%**
F3	26.9%	24.3%	21.2%	17.9%	**16.5%**
F4	30.6%	30.3%	25.9%	**18.2%**	20.1%
	Low-Resolution Dataset				
Metric	C1	C2	C3	C4	C5
M1	45.8%	**42.8%**	46.4%	43.6%	44.5%
M2	35.5%	41.7%	34.6%	33.8%	**33.2%**
M3	50.0%	**41.6%**	50.0%	50.0%	44.5%
M4	**44.3%**	46.3%	46.0%	44.4%	46.0%
M5	47.2%	**41.8%**	47.6%	43.7%	44.7%
M6	42.8%	45.1%	39.2%	37.5%	**36.8%**
M7	34.7%	41.3%	35.1%	30.7%	**30.2%**
M8	43.4%	42.2%	39.4%	**34.3%**	35.2%
M9	33.9%	41.4%	34.7%	31.5%	**30.0%**
M10	43.7%	42.8%	38.4%	39.2%	**36.8%**
M11	41.9%	40.6%	36.1%	**30.3%**	30.8%
M12	44.2%	39.1%	41.3%	32.0%	**31.7%**
F1	33.7%	39.5%	34.8%	**26.7%**	28.2%
F2	37.5%	38.6%	33.2%	**25.6%**	26.7%
F3	39.7%	41.8%	38.6%	32.6%	**31.8%**
F4	35.6%	39.5%	35.8%	**26.2%**	27.0%

Table 1: Equal Error Rates.

	High-Resolution Dataset				
Metric	C1	C2	C3	C4	C5
M1	15.4%	13.8%	14.6%	15.4%	**16.6%**
M2	17.6%	6.6%	19.6%	41.9%	**48.7%**
M3	10.9%	7.8%	5.9%	10.7%	**14.9%**
M4	9.8%	7.7%	9.1%	**12.5%**	10.4%
M5	**16.9%**	10.2%	16.4%	15.8%	15.6%
M6	12.9%	9.9%	10.0%	37.9%	**37.9%**
M7	5.7%	12.3%	6.4%	**36.7%**	34.7%
M8	6.8%	13.7%	14.3%	28.3%	**31.5%**
M9	12.2%	14.3%	14.5%	**44.0%**	42.2%
M10	6.5%	11.2%	11.2%	20.2%	**23.7%**
M11	5.6%	12.6%	14.4%	43.3%	**48.5%**
M12	4.4%	12.1%	13.2%	27.1%	**27.7%**
F1	24.2%	17.4%	33.4%	61.4%	**64.8%**
F2	11.8%	13.3%	16.7%	29.1%	**29.9%**
F3	48.2%	57.0%	76.0%	**83.7%**	82.6%
F4	34.3%	32.0%	47.1%	**72.0%**	70.6%
	Low-Resolution Dataset				
Metric	C1	C2	C3	C4	C5
M1	1.4%	1.2%	1.0%	**2.2%**	2.0%
M2	1.8%	2.0%	0.5%	3.2%	**3.4%**
M3	0.0%	1.6%	0.2%	**1.9%**	0.9%
M4	0.2%	0.5%	**0.7%**	0.4%	0.5%
M5	1.5%	0.9%	1.3%	**2.2%**	2.0%
M6	0.7%	1.5%	1.6%	**2.8%**	1.2%
M7	2.0%	1.2%	2.6%	5.4%	**6.7%**
M8	1.8%	0.6%	1.0%	4.3%	**5.5%**
M9	3.5%	1.2%	1.5%	**8.5%**	8.2%
M10	2.8%	1.2%	1.4%	2.6%	**4.2%**
M11	1.0%	2.5%	2.0%	**9.9%**	9.4%
M12	0.6%	1.2%	1.3%	4.9%	**5.8%**
F1	2.9%	2.8%	3.9%	15.4%	**16.3%**
F2	0.9%	1.1%	1.5%	2.0%	**2.8%**
F3	2.8%	2.0%	4.7%	**14.2%**	12.9%
F4	6.1%	3.2%	4.9%	**14.0%**	12.6%

Table 2: Rank-1 Identification Rates.

Of the considered information fusion algorithms, a clear distinction is less evident. On the high-resolution dataset, the random forest algorithm provided both the lowest equal error rate (16.5%) and the highest rank-1 identification rate (83.7%), with a clear advantage in every category. Unfortunately, the random forest algorithm did not maintain this advantage when tested on the low-resolution dataset; and, in fact, was clearly less accurate than fusion by weighted mean and likelihood ratio in terms of equal error rate. It should be noted that, because the support vector machine applied classification rather than regression, equal error rates in this case were rough approximations; even so, the support vector machine was the worst performing fusion method.

While these results do not yet represent a realistic alternative to existing biometric standards, such as fingerprint recognition, they do provide a substantial improvement in the biometric viability of eye movements, as compared to existing research. Further, less accurate eye movement biometrics have been employed to improve the accuracy and liveness detection of multi-modal systems [25]. The fairest comparison with existing work is to the results reported by Holland and Komogortsev [14], which made use of equivalent datasets, obtaining equal error rates of 27% on the high-resolution dataset and 24% on the low-resolution dataset; however, these results may suffer from overfitting, as training and testing were conducted on the full dataset, without partitioning.

For comparison, error rates were recalculated using the techniques described in [14], and the updated biometric framework described in Section 2. With random splitting by subject, averaged over 20 partitions, previous CEM techniques achieved a 28.5% equal error rate and 37.5% rank-1 identification rate on the high-resolution dataset, and 32.9% equal error rate and 3.7% rank-1 identification rate on the low-resolution dataset. The techniques proposed in the current paper reduce equal error rate by approximately 42% on the high-resolution dataset and 22% on the low-resolution dataset, and increase rank-1 identification rate by 223% on the high-resolution dataset and 441% on the low-resolution dataset. This is unquestionably a massive improvement.

Differences in eye tracking specifications and subject pool make it difficult to compare these results with similar work in eye movement biometrics. Moreover, to our knowledge, much of the early work in this context did not provide equal error rates or rank-1 identification rates to compare against. Recently, Rigas et al. [13] achieved 30.0% equal error rate and 70.2% rank-1 identification rate on a pool of 15 subjects, using an eye tracking system with 0.25° spatial accuracy and a sampling rate of 50 Hz. Previous research [14] has shown that spatial accuracy plays a much greater role in biometric accuracy than does sampling rate, and for this reason we are inclined to consider the recordings as being high-resolution, but realistically this dataset does not fit either category neatly.

It should be noted, that while the recent EMVIC competition [26] has produced higher identification rates, there are several issues with the data collection procedures and database partitioning that make the results suspect. Specifically, the use of uncalibrated eye tracking data in two of the four major datasets may have introduced artifacts into the eye movement recordings, and the use of a single non-random partition with subject crossover may have caused overfitting. For reference, the uncalibrated datasets achieved rank-1 identification rates of 97.6% and 95.1% on a mid-resolution eye tracker (250 Hz, unreported accuracy), while the calibrated datasets achieved rank-1 identification rates of 58.6% and 66.7% on a high-resolution eye tracker (1000 Hz, 0.5° accuracy).

7. Conclusion

This paper has presented an objective evaluation of previously unexplored biometric techniques utilizing patterns identifiable in human eye movements to distinguish individuals. The distribution of primitive eye movement features were compared between eye movement recordings using algorithms based on the following statistical tests: the two-sample t-test, the Ansari-Bradley test, the Mann-Whitney U-test, the two-sample Kolmogorov-Smirnov test, and the two-sample Cramér-von Mises test. Score-level information fusion was applied and evaluated by: weighted mean, support vector machine, random forest, and likelihood ratio.

Based on the results, we found that on high-resolution eye tracking equipment, it is possible to obtain equal error rates of 16.5% and rank-1 identification rates of 82.6% using the two-sample Cramér-von Mises test and score-level fusion by random forest. Unfortunately, these techniques scale poorly for low-resolution eye tracking equipment, receiving an equal error rate of 26.2% and rank-1 identification rate of 14.0% using the two-sample Kolmogorov-Smirnov test and fusion by likelihood ratio.

References

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

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

[3] 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,* vol. 28, pp. 3-18, 2006.

[4] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms," *IEEE Transactions on Neural Networks,* vol. 14, pp. 195-200, 2003.

[5] V. Dorairaj, N. A. Schmid, and G. Fahmy, "Performance Evaluation of Non-ideal Iris Based Recognition System Implementing Global ICA Encoding," in *International Conference on Image Processing (ICIP)*, Genoa, Italy, 2005, pp. 285-288.

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

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

[8] W. Sewell and O. V. Komogortsev, "Real-time Eye Gaze Tracking with an Unmodified Commodity Webcam Employing a Neural Network," in *Conference on Human Factors in Computing Systems (CHI)*, Atlanta, GA, USA, 2010, pp. 3739-3744.

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

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

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] D. D. Salvucci and J. H. Goldberg, "Identifying Fixations and Saccades in Eye-tracking Protocols," in *Eye Tracking Research & Applications (ETRA) Symposium*, Palm Beach Gardens, FL, USA, 2000, pp. 71-78.

[16] 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.

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

[18] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning,* vol. 20, pp. 273-297, 1995.

[19] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, pp. 5-32, 2001.

[20] K. Nandakumar, Y. Chen, S. C. Dass, and A. K. Jain, "Likelihood Ratio-based Biometric Score Fusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 30, pp. 342-347, 2008.

[21] O. V. Komogortsev. (2012). *Eye Movement Biometric Database v1*. Available: http://cs.txstate.edu/~ok11/embd_v1.html

[22] EyeLink. *EyeLink 1000 Eye Tracker*. Available: http://www.sr-research.com/

[23] 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.

[24] Sony. *PlayStation Eye Camera*. Available: http://us.playstation.com/ps3/accessories/playstation-eye-camera-ps3.html

[25] 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.

[26] P. Kasprowski, O. V. Komogortsev, and A. Karpov, "Eye Movement Verification and Identification Competition at BTAS 2012," presented at the Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), Washington DC, 2012.