Template Aging in Eye Movement-driven Biometrics

Oleg V. Komogortsev*, Corey D. Holland, Alexey Karpov
Texas State University, Dept. of Computer Science, San Marcos, TX USA 78666-4684

ABSTRACT

This paper presents a template aging study of eye movement biometrics, considering three distinct biometric techniques on multiple stimuli and eye tracking systems. Short-to-midterm aging effects are examined over two-weeks, on a high-resolution eye tracking system, and seven-months, on a low-resolution eye tracking system. We find that, in all cases, aging effects are evident as early as two weeks after initial template collection, with an average 28% (±19%) increase in equal error rates and 34% (±12%) reduction in rank-1 identification rates. At seven months, we observe an average 18% (±8%) increase in equal error rates and 44% (±20%) reduction in rank-1 identification rates. The comparative results at two-weeks and seven-months suggests that there is little difference in aging effects between the two intervals; however, whether the rate of decay increases more drastically in the long-term remains to be seen.

Keywords: eye tracking, biometrics, template aging

1. INTRODUCTION

Biometrics is a far-reaching and ever-changing field[1], an amalgamation of bits and pieces of numerous disciplines, working together to answer a single question: What makes us unique? This question has led to the study of a wide range of human characteristics, physical and behavioral, including: fingerprints, iris patterns, speech, handwriting, walking gait, and facial structure.

Many of these characteristics have been under active research in a biometric context for decades, some tracking initial investigations as far back as the mid to late 19th century[2]. More recently, study of the human visual system has shown that eye movements provide a number of novel and unique properties that make them desirable as a behavioral biometric.

Template aging is the term used to address the degradation in biometric accuracy that occurs as physical and behavioral traits are altered by the growth and decay of the human body[3]. Aging effects such as bone growth, reduced skin elasticity, reduced muscle strength, reduced sight/hearing, disease, and illness may affect biometric modalities to different extents[4].

Template aging effects have been shown to reduce biometric accuracy in many standard biometric modalities, including: face[5], iris[6], fingerprint[7], and voice[8] recognition. Algorithms have been devised to address these effects[9], correcting for the effects of age on biometric templates, though in this regard there is still much work to be done.

As a relatively recent behavioral biometric, the effects of template aging on eye movement biometrics has not been investigated in any capacity. It is likely that, like the face and iris, eye movements are not invariant to the effects of age; however, the degree and extent to which aging affects the biometric viability of eye movements is an as of yet unstudied topic. We hypothesize that the effects of template aging on eye movement biometrics may be noticeable yet insignificant in the short-to-midterm (i.e. weeks / months), and attempt to quantify the rate of re-enrollment required to maintain biometric viability.

2. EYE MOVEMENT BIOMETRICS

In this paper, we consider three techniques for the application of eye movement biometrics, defined previously in the relevant literature: complex eye movement patterns[10], complex eye movement behavior[11], and oculomotor plant characteristics[12]. Each of these techniques concerns itself with fundamentally different eye movement properties and distinct methods of comparison.

* ok11@txstate.edu; phone: 1 512 245-0349; fax 1 512 245-8750; cs.txstate.edu/~ok11/

2.1 Human Visual System

The human visual system is composed of two primary components, the oculomotor plant and brainstem control. The oculomotor plant encompasses the eye globe, six extraocular muscles, and the surrounding tissues and ligaments that provide both viscous and elastic properties to the whole. The brainstem control is responsible for the generation and transmission of the neuronal control signal, a blanket term for the omnipause and burst neurons that cause contraction and relaxation of extraocular muscles

Together, these components produce the many and varied types of human eye movement, including: fixation, saccade, smooth pursuit, vergence, optokinetic reflex, and vestibule-ocular reflex. Of these, fixations and saccades are of particular interest due to their highly stereotypical and reproducible behavior. 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.

2.2 Complex Eye Movement Patterns

Complex eye movement pattern (CEM-P) biometrics was the result of a first attempt at relating high-level neurological processes to basic and aggregate patterns found in the properties of human eye movements[10]. This early work placed a focus on scanpath theory[13] and a number of traits that had previously proven useful indicators of software usability[14].

For a given eye movement recording, a biometric feature vector is generated to include the following statistics: scanpath length, scanpath convex hull area, fixation count, average fixation duration, regions of interest count, inflection count, average vectorial saccadic amplitude, average horizontal saccadic amplitude, average vertical saccadic amplitude, average vectorial saccadic velocity, average vectorial saccadic peak velocity, slope coefficient of the amplitude-duration relationship, slope coefficient of the amplitude-peak velocity relationship, and average velocity waveform indicator (Q).

These statistics are compared individually, between recordings, using a Gaussian kernel to generate biometric match scores, and the scores for each feature are combined using an information fusion algorithm. In the original work on this topic[10], a weighted mean was employed to perform information fusion, using a forward-search algorithm to select weighting; however, in the current paper, we employ a random forest for this purpose[15].

2.3 Complex Eye Movement Behavior

Complex eye movement behavior (CEM-B) biometrics is a natural extension of the ideas expressed in the formulation of CEM-P biometrics[11]; however, the two techniques are fundamentally different in their approach. While CEM-P examines the distance between sums and averages of quantifiable eye movement patterns, CEM-B considers the overall distribution of fixations and saccades throughout a recording.

For each recording, fixations and saccades are identified and quantified according to the following properties, where: fixations are described by start time, duration, horizontal centroid, and vertical centroid; and saccades are described by start time, duration, horizontal amplitude, vertical amplitude, average horizontal velocity, average vertical velocity, horizontal peak velocity, and vertical peak velocity.

The distribution of these features is compared between recordings using the Cramér-von Mises statistical test, an extension of the Kolmogorov-Smirnov test used to compare the goodness of fit of two empirical distributions. This produces a biometric match score for each eye movement property (4 for fixations and 8 for saccades), which are then combined into a single match score using information fusion by random forest[15].

2.4 Oculomotor Plant Characteristics

Oculomotor plant characteristic (OPC) biometrics arose as an attempt to estimate the physical properties of the oculomotor plant from the properties exhibited by saccadic eye movements[12]. A mathematical model of the oculomotor plant is employed to simulate saccadic eye movements, and an optimization problem is devised to estimate parameters of the model that most closely fit the simulated saccade trajectory to the trajectory of a measured saccade.

In the current paper, we employ this technique as it was originally applied in a biometric context[12], utilizing a two-dimensional linear homeomorphic oculomotor plant model[16] with the following parameters: series elasticity, length-tension relationship, force-velocity relationship, passive viscosity*, tension slope, inertial mass*, activation time, deactivation time, tension intercept*, neural pulse, and neural pulse width*. These parameters are referred to as oculomotor plant characteristics (OPC), and except where noted (*), these parameters are applied separately for agonist/antagonist muscle pairs.

The Nelder-Mead simplex algorithm[17] is applied to estimate model parameters that minimize the absolute difference between measured and simulated saccade trajectories. For each recording, OPC are estimated for the horizontal component of each saccade separately (OPC-H), ignoring any saccade of less than 4° amplitude, less than 20 milliseconds duration, or containing abnormal trajectory artifacts. Further, OPC estimation by the Nelder-Mead algorithm is constrained to a maximum of 13 minutes per saccade. The result is a biometric feature vector containing OPC values for each saccade in a recording, the distributions of which are compared between recordings using the multivariate Hotelling T2 test to generate a biometric match score[18].

3. METHODOLOGY

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.[19, 20], were utilized for comparative evaluation, with collection methodology in the following subsections.

3.1 Participants

Low-resolution eye movement data was collected for a total of 33 subjects (21 males, 12 females), ages 18 – 34 with an average age of 21 (SD = 3.3). 30 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 per stimulus. Subjects were given a 20-minute break between the 1st and 2nd recording, with 7 month between the 2nd and 3rd recording, and 20 minutes between the 3rd and 4th recording.

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 per stimulus. Subjects were given a 20-minute break between the 1st and 2nd recording, with 2 weeks between the 2nd and 3rd recording, and 20 minutes between the 3rd and 4th recording.

3.2 Apparatus & Software

Low-resolution eye movements were recorded using video-oculography techniques and a common web camera. For these purposes, modified versions of the open-source ITU Gaze Tracker software[21] and PlayStation Eye Camera[22] were employed for monocular gaze tracking, providing a temporal resolution of 75 Hz and average calibration accuracy of 1.0° (SD = 0.5°). 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.

High-resolution eye movements were recorded using a high-performance commercial eye tracking system. For these purposes, the EyeLink 1000 was employed for binocular gaze tracking, providing a temporal resolution of 1000 Hz and average calibration accuracy of 0.7° (SD = 0.5°). 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.

In both cases, a chin rest was employed to improve stability. All algorithms and data analysis were implemented and performed in MATLAB, and run on a 3.1 GHz quad-core CPU with 16 GB memory.

3.3 Procedure

Eye movement recordings were generated for three distinct stimuli, including: a horizontal saccade stimulus (HSS), a random saccade stimulus (RSS), and a reading saccade stimulus (RES). For both high- and low-resolution eye tracking systems, stimulus presentation was roughly equivalent, with only minor variation required to accommodate different screen sizes.

The horizontal saccade stimulus (HSS) employed a technique commonly used to evoke a fixed-amplitude saccade at regular intervals[23]. A small white dot jumped back and forth on a plain black background, eliciting a 30° horizontal saccade with each jump. The 30° amplitude was chosen due to screen constraints and the complications associated with separating low-amplitude saccades (less than 1°). Subjects were instructed to follow the white dot with their eyes, with 100 saccades elicited per recording.

The random saccade stimulus (RSS) applied the same techniques as the horizontal saccade stimulus. A small white dot jumped across a plain black background in a uniformly distributed random pattern, eliciting an oblique saccade of random amplitude with each jump. Subjects were instructed to follow the white dot with their eyes, with 100 saccades elicited per recording.

The reading saccade stimulus (RES) presented a textual excerpt of moderate length and difficulty, in order to evoke eye movements as they naturally occur during reading. Textual excerpts were selected to ensure that readying required roughly 1 minute, line lengths and the difficulty of material was consistent, and learning effects did not impact subsequent readings. Subjects were given 1 minute for each textual excerpt, and a different excerpt was displayed for each session.

Eye movement recordings were parsed to remove invalid data points. Recordings were stored in an eye movement database, with each record linked to the stimulus, subject, and session that generated the recording. The recordings were then processed and classified into fixations and saccades using an eye movement classification algorithm[24], followed by micro-saccade and micro-fixation filters respectively. The initial classification algorithm (I-VT) classified individual data points with a velocity greater than 20°/sec as belonging to a saccade, with all remaining points belonging to fixations. The micro-saccade filter re-classified saccades smaller than 0.5° amplitude as fixations and the micro-fixation filter re-classified fixations of less than 100 milliseconds as saccades. Fixation and saccade groups were then merged, identifying fixation and saccade-specific features.

For ease of reference the recording sessions were labeled alphabetically and grouped into subsets, where the 1st session was labeled A, the 2nd session labeled B, the 3rd session labeled C, and the 4th session labeled D. A subset consists of eye movement recordings from two sessions that represents a distinct time interval between recordings; for example, subset AB would include eye movement recordings from the 1st and 2nd sessions, and is representative of a 20-minute time interval. For both the high- and low-resolution eye tracking systems, recording subsets AB and CD represent a 20 minute time interval; for the high-resolution eye tracking system, recording subsets AC, AD, BC, and BD represent a 2 week time interval; and for the low-resolution eye tracking system, recording subsets AC, AD, BC, and BD represent a 7 month time interval.

Recording subsets were partitioned into training and testing sets according to a uniformly random distribution, and biometric match scores were generated according to the techniques described in Section 2, with algorithm thresholds and parameters selected on the training set(s) and error rates calculated on the testing set(s).

Error rates were calculated for biometric verification and identification scenarios. In the verification scenario, each recording in the testing set was compared to every other recording in the testing set exactly once, and error rates were calculated from those comparisons. In the identification scenario, every recording in the testing set was compared to every other recording in the testing set, and identification rates were calculated from the largest match score(s) from each of the comparison sets.

4. RESULTS

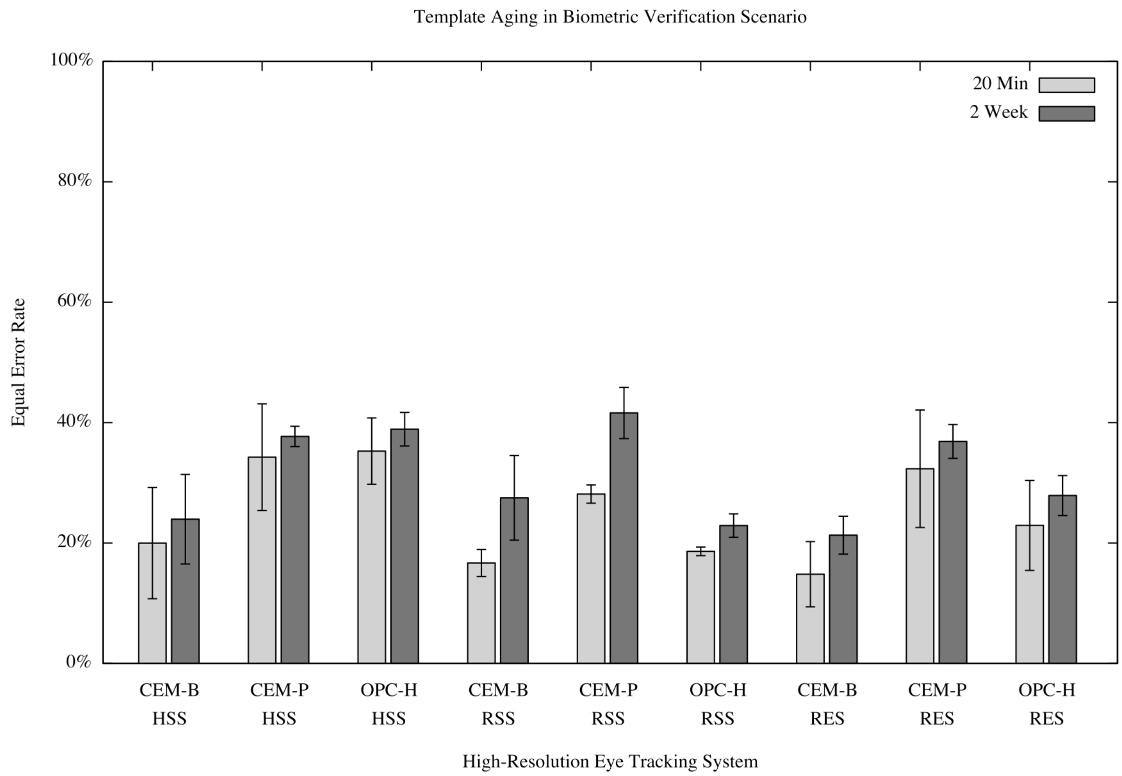
Eye movement recordings were partitioned, by subject, into training and testing sets with a ratio of 1:1, such that no subject had recordings in both the training and testing sets. Experimental results were averaged over 20 random partitions for each recording subset, and error rates of recording subsets (AB, BC, etc.) in the same time interval (20 minutes, 2 weeks, 7 months) were similarly averaged.

4.1 Verification Scenario

False acceptance rate (FAR) is defined as the rate at which imposter match scores exceed the acceptance threshold and false rejection rate (FRR) is defined as the rate at which genuine match scores fall below the acceptance threshold. The equal error rate (EER), shown in Figure 1, is the rate at which false acceptance rate and false rejection rate are equal.

4.2 Identification Scenario

Identification rate (IR) 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, shown in and Figure 2, is the rate at which the correct individual has the highest match score.



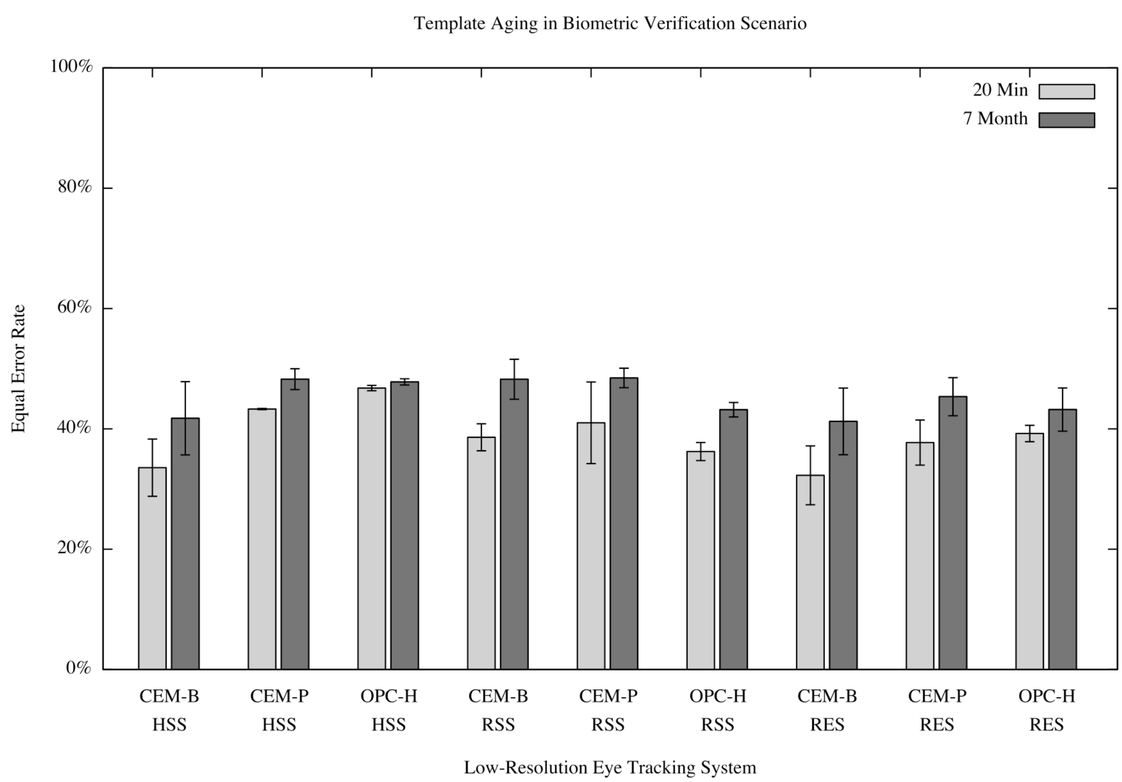
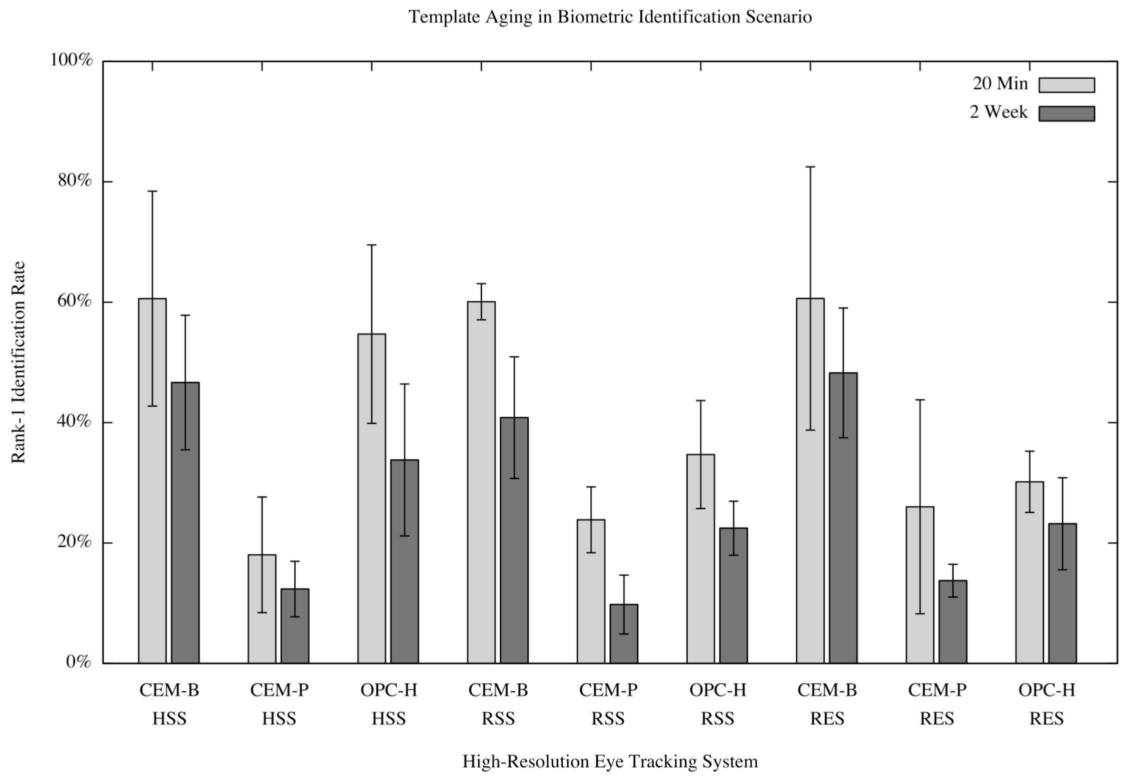


Figure 1. Comparative equal error rates by system, technique, stimulus, and age group.



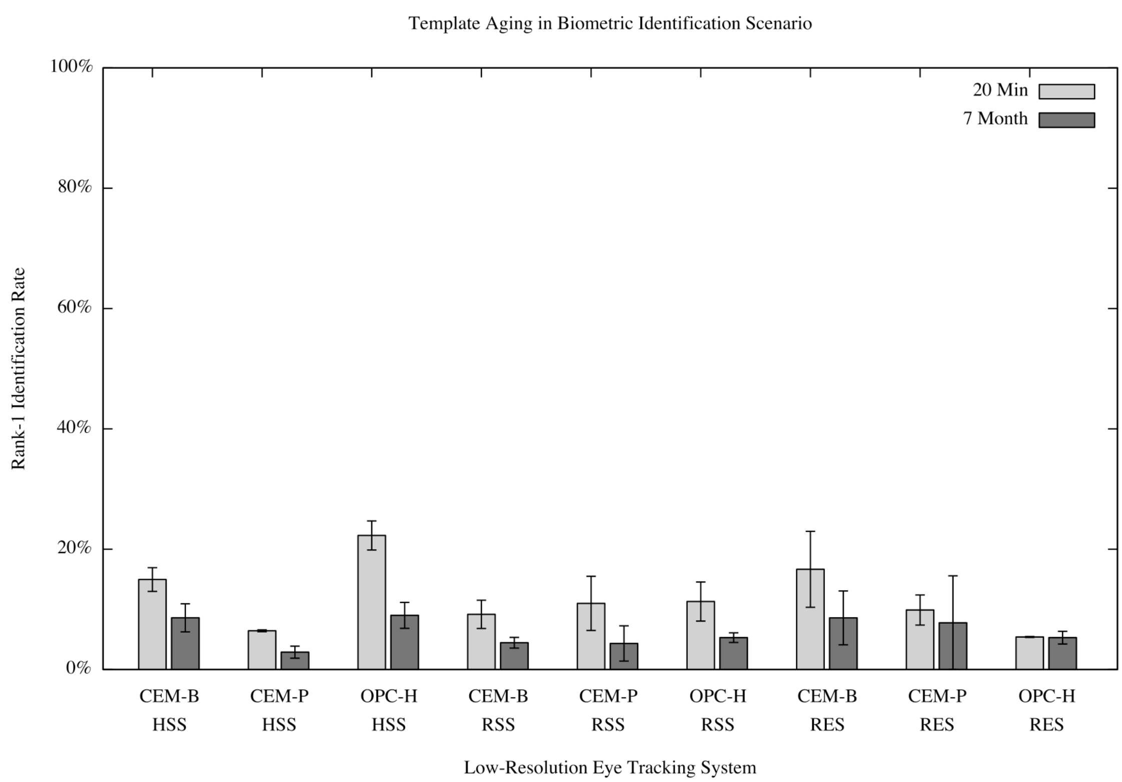


Figure 2. Comparative rank-1 identification rates by system, technique, stimulus, and age group.

4.3 Aging Effects

To obtain a quantitative measure of aging effects, we measure the percentage change in equal error rates and rank-1 identification rates, as shown in Table 1, from the 20-minute interval to the two-week and seven-month intervals.

Table 1. Percentage change in biometric accuracy with aging.

Interval	Stimulus	Biometric	EER Change	IR Change
2 Weeks	HSS	CEM-B	+20%	-23%
2 Weeks	HSS	CEM-P	+10%	-32%
2 Weeks	HSS	OPC-H	+10%	-38%
2 Weeks	RSS	CEM-B	+65%	-32%
2 Weeks	RSS	CEM-P	+48%	-55%
2 Weeks	RSS	OPC-H	+23%	-35%
2 Weeks	RES	CEM-B	+44%	-20%
2 Weeks	RES	CEM-P	+14%	-47%
2 Weeks	RES	OPC-H	+22%	-23%
7 Months	HSS	CEM-B	+24%	-43%
7 Months	HSS	CEM-P	+11%	-55%
7 Months	HSS	OPC-H	+2%	-60%
7 Months	RSS	CEM-B	+25%	-51%
7 Months	RSS	CEM-P	+18%	-61%
7 Months	RSS	OPC-H	+19%	-53%
7 Months	RES	CEM-B	+28%	-49%
7 Months	RES	CEM-P	+20%	-22%
7 Months	RES	OPC-H	+10%	-2%

5. DISCUSSION

Aging effects are immediately obvious from the results. Even as early as two weeks after initial enrollment, there is a noticeable increase in equal error rates and decrease in rank-1 identification rates from the baseline 20-minute interval. This tendency holds true for all considered eye tracking systems, stimuli, and biometric techniques.

The differences in aging effects between the two-week and seven-month intervals are less obvious. At two weeks, equal error rates showed an average increase of 28% and rank-1 identification rates showed an average decrease of 34%; while at seven months, equal error rates showed an average increase of 18% and rank-1 identification rates showed an average decrease of 44%.

Since it is safe to assume that error rates do not improve with time, as per the effects of equal error rate at two- weeks and seven-months, we must instead draw another conclusion. Applying a two-tailed Student's t-test revealed no significant difference in the percent change in equal error rates ($t(8) = 2.06$, $p = 0.074$) or rank-1 identification rates ($t(8) = 1.50$, $p = 0.173$) at two-weeks and seven-months. As a result, we cannot draw useful conclusions from these differences, though it is possible to speculate that the differences in the rate of decay of biometric templates may be due to the differences in eye tracking specification (i.e. high- vs. low-resolution).

In general, the biometric application of eye movements performed much poorer on the low-resolution eye tracking system. It is possible, and even likely, that the poor error rates on the low-resolution system left less room for degradation due to template aging, and may be responsible for the peculiarities evident in the differences between the two-week and seven-month intervals.

With regard to individual stimuli, a two-way ANOVA indicated that there was no significant interaction effect between template aging effects and stimulus in either equal error rate ($F(2, 17) = 1.47$, $p = 0.269$) or rank-1 identification rate ($F(2, 17) = 1.73$, $p = 0.219$). Similarly, no significant interaction effect was identified between template aging effects and biometric algorithm, with: $F(2, 17) = 0.25$, $p = 0.784$, and, $F(2, 17) = 0.65$, $p = 0.5413$, respectively.

Regarding, more generally, error rates without respect to template aging effects, we observe obvious interaction between biometric technique and stimulus, and substantial differences in biometric accuracy by technique. A two-way ANOVA indicated significant interaction effects between biometric technique and stimulus on the: high-resolution verification scenario, $F(4, 53) = 4.6$, $p < 0.003$; low-resolution verification scenario, $F(4, 53) = 2.65$, $p < 0.046$; and low-resolution identification scenario, $F(4, 53) = 3.62$, $p < 0.012$. This contradicts previous findings [19], which indicated that stimulus had no discernable effect on the accuracy of eye movement biometrics, specifically the CEM-P technique.

At the algorithm level, CEM-B consistently achieved the lowest equal error rates and the highest rank-1 identification rates of the considered techniques, while CEM-P consistently performed the worst. It is interesting to note that, in terms of identification rate, OPC-H performed best on the horizontal stimulus and worst on the reading stimulus; and while we did not include the vertical OPC component due to time constraints, it is likely that it could be utilized to improve biometric error rates through information fusion.

6. CONCLUSION

This paper has presented a template aging study of eye movement biometrics, considering three distinct biometric techniques on multiple stimuli and eye tracking systems. Short-to-midterm aging effects were examined over two-week and seven-month intervals, on high- (1000 Hz) and low- (75 Hz) resolution eye tracking equipment.

Based on the results, we find that aging effects are evident as early as two weeks after initial template collection, with an average 28% increase in equal error rates and 34% reduction in rank-1 identification rates. At seven months, we observe an average 18% increase in equal error rates and 44% reduction in rank-1 identification rates. The comparative results at two-weeks and seven-months suggest that there is little difference in aging effects between the two intervals; however, whether the rate of decay increases more drastically in the long-term remains to be seen. Future work in this area will likely consider long-term aging effects, on the order of years, and the possible application of corrective algorithms that may reduce the impact of aging.

7. ACKNOWLEDGEMENTS

This work is supported in part by NSF CAREER Grant #CNS-1250718 and NSF GRFP Grant #DGE-11444666, and NIST Grants #60NANB10D213 and #60NANB12D234.

REFERENCES

[1] Jain, A. K., Flynn, P. and Ross, A. A., "Handbook of Biometrics," 1-556 (2007).
[2] Komarinski, P. D., [Automated Fingerprint Identification Systems (AFIS)], Academic Press, 1-312 (2004).
[3] Mansfield, A. J. and Wayman, J. L., "Best Practices in Testing and Reporting Performance of Biometric Devices," 1-36 (2002).
[4] Lanitis, A. and Tsapatsoulis, N., "Quantitative Evaluation of the Effects of Aging on Biometric Templates," Computer Vision 5(6), 338-347 (2011).
[5] Patterson, E., Sethuram, A., Albert, M., Ricanek, K. and King, M., "Aspects of Age Variation in Facial Morphology Affecting Biometrics," IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS), 1-6 (2007).
[6] Fenker, S. P. and Bowyer, K. W., "Analysis of Template Aging in Iris Biometrics," IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 45-51 (2012).
[7] Modi, S. K., Elliott, S. J. and Kim, H., "Impact of Age Groups on Fingerprint Recognition Performance," IEEE Workshop on Automatic Identification Advanced Technologies, 19-23 (2007).
[8] Kelly, F. and Harte, N., "Effects of Long-Term Ageing on Speaker Verification," Biometrics and ID Management, 113-124 (2011).

[9] Kekre, H. B. and Bharadi, V. A., "Ageing Adaptation for Multimodal Biometrics using Adaptive Feature Set Update Algorithm," IEEE International Advance Computing Conference (IACC), 535-540 (2009).
[10] Holland, C. D. and Komogortsev, O. V., "Biometric Identification via Eye Movement Scanpaths in Reading," International Joint Conference on Biometrics (IJCB), 1-8 (2011).
[11] Holland, C. D. and Komogortsev, O. V., "Complex Eye Movement Pattern Biometrics: Analyzing Fixations and Saccades," IAPR International Conference on Biometrics (ICB), 1-8 (2013).
[12] Komogortsev, O. V., Karpov, A., Price, L. R. and Aragon, C. R., "Biometric Authentication via Oculomotor Plant Characteristics," International Conference on Biometrics (ICB), 1-8 (2012).
[13] Noton, D. and Stark, L. W., "Scanpaths in Eye Movements during Pattern Perception," Science 171(3968), 308-311 (1971).
[14] Goldberg, J. H. and Kotval, X. P., "Computer Interface Evaluation using Eye Movements: Methods and Constructs," International Journal of Industrial Ergonomics 24(6), 631-645 (1999).
[15] Breiman, L., "Random Forests," Machine Learning 45(1), 5-32 (2001).
[16] Komogortsev, O. V. and Jayarathna, U. K. S., "2D Oculomotor Plant Mathematical Model for Eye Movement Simulation," International Conference on BioInformatics and BioEngineering (BIBE), 1-8 (2008).
[17] Lagarias, J. C., Reeds, J. A., Wright, M. H. and Wright, P. E., "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," SIAM Journal of Optimization 9(1), 112-147 (1998).
[18] Hotelling, H., "The Generalization of Student's Ratio," The Annals of Mathematical Statistics 2(3), 360-378 (1931).
[19] Holland, C. D. and Komogortsev, O. V., "Biometric Verification via Complex Eye Movements: The Effects of Environment and Stimulus," Fifth Internation Conference on Biometrics: Theory, Applications and Systems (BTAS), 1-8 (2012).
[20] "Eye Movement Biometric Database v1," 2012, <http://cs.txstate.edu/~ok11/embd_v1.html%3E.
[21] Agustin, J. S., Skovsgaard, H., Hansen, J. P. and Hansen, D. W., "Low-cost Gaze Interaction: Ready to Deliver the Promises," Conference on Human Factors in Computing (CHI), 4453-4458 (2009).
[22] "PlayStation Eye Camera," <http://us.playstation.com/ps3/accessories/playstation-eye-camera-ps3.html%3E.
[23] Leigh, R. J. and Zee, D. S., [The Neurology of Eye Movements], Oxford University Press, Oxford, NY, USA, 1-776 (2006).
[24] Salvucci, D. D. and Goldberg, J. H., "Identifying Fixations and Saccades in Eye-tracking Protocols," Eye Tracking Research & Applications (ETRA) Symposium, 71-78 (2000).