IMPROVING EYE MOVEMENT BIOMETRICS USING REMOTE REGISTRATION OF EYE BLINKING PATTERNS

Sherif Seha, Georgios Papangelakis, Dimitrios Hatzinakos

Electrical and Computer Engineering Dept.,
University of Toronto, Canada
{sherif.seha, georgios.papangelakis}@mail.utoronto.ca,
dimitris@comm.utoronto.ca

Ali Shahidi Zandi, Felix JE Comeau

Alcohol Countermeasure Systems Inc.,
Toronto, Canada
{aszandi, fjecomeau}@acs-corp.com

ABSTRACT

In this paper, the biometric potential of eye movement and eye blinking for human recognition task is investigated. These modalities might be useful for specific biometric applications like driver authentication for law enforcement. For this purpose, a database of 22 subjects was build where eye movements and blinks were recorded using Gazepoint GP3 while users were watching real driving sessions. Eye movement features were extracted from eye fixations and saccades separately. Eye blinking features include the blink pattern, its speed and acceleration patterns, and time delineation features. Evaluation of each modality was investigated first, then, both modalities are combined in a multi-modal setup for performance improvement. Although the employment of eye movement or eye blinking separately as a biometric trait might not be secure enough, the fusion of both traits achieves higher levels of identification which are comparable to that of other conventional biometric traits like fingerprint.

1. INTRODUCTION

Authentication systems based on biometric traits have been discussed extensively for the past few decades for their inherent uniqueness and consistency. Originally, static biometric features such as iris, fingerprint, and facial recognition systems are considered viable. However, these all possess the inherent issue that they can be fooled by using a fake copy of the feature such as a picture or a model [1, 2]. Biometric signals are also unique and consequently much harder to replicate as each signal is caused by subconscious processes. They have not seen widespread usage due to their invasive behavior caused by the medical equipment needed to capture the signal. This paper attempts to solve this problem by creating a non-invasive continuous authentication system based on biometric signals caused by eye movements and blinking. Such a device would be very applicable, as it could remotely authenticate the identity of a subject every few minutes. The example used in this paper is the application of eye movement and eye blinking biometrics in smart cars, to prevent theft as well as to increase the effectiveness of intoxication prevention devices such as Breathalyzers.

Most research in eye blinking and eye movement biometrics requires the subjects to be in a controlled environment where they cannot move their heads, as well as be exposed to specific predetermined stimulation such as a moving dot on a screen. The objective of this paper is to apply these modalities in a more practical scenario for continuous driver authentication taking advantage of eye tracking device that are already installed in smart cars to measure distraction levels. Moreover, in order to improve the eye movement based biometric setup, the biometric potential of eye blinking patterns will be investigated first, then, combined with eye movements in a multi-modal setup.

The database in this paper is collected using a commercial and cost-effective eye tracking device (Gazepoint GP3). The device uses Infra-Red (IR) light to detect the eyes, as well as the gaze of the subject. The light intensity profile of the subjects eye's is extracted from the recording in order to get the blinking data. As infrared light is absorbed more by the iris and pupil than the eyelid. Consequently, the eyelid will reflect more light causing positive spikes in the intensity profile. This method has already been implemented in past experiments with great success [3].

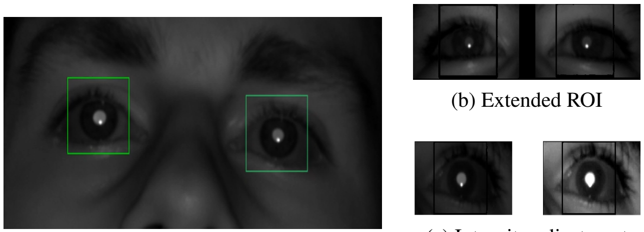
The rest of this paper is organized as follows. Previous works employing eye movements and blinking for biometric authentication are briefly discussed in Section 2. These works are compared to this experiment in order to justify and evaluate the results. The methods for pre-processing and feature extraction for the eye movement and eye blinking patterns as well as their fusion will be discussed in Section 3. Section 4 summarizes the results of the experiments. Finally, main conclusion and future work are discussed in Section 5.

2. RELATED WORK

Eye movements have been heavily investigated for biometric research, however, in this brief discussion, we focus on the most recent published work, specifically speaking, the 2015 BioEye competition [4, 5]. Usually eye movement patterns are recorded in response to a visual task like reading text or following a randomly moving target. Features from eye movements can be extracted directly (from time or frequency domain) using signal processing techniques like Fourier Transform (FT), Dynamic Time Warping (DTW), or Gaussian Mixture Models (GMM) [6]. The more common approach is to classify eye movement into fixations and saccades, then, extract features from these profiles [7, 8]. The best system's performance in this competition was achieved by A. George and A. Routray [7] where features like duration, amplitude, velocity and acceleration were extracted from fixations and saccades. There proposed system achieved a correct recognition rate of 89.54% over a population of 153 subjects. In this paper, the same features were adopted for our proposed eye movement system.

On the other side, eye blinking based biometrics is a new research topic where different registration techniques for eye blinking patterns were followed. The first attempt to use eye blinking patterns for human recognition tasks was conducted by S. Seha et al. [9] us-

Thanks to NSERC Canada and Alcohol Countermeasure Systems (ACS) Inc. for their support



(a) Gazepoint eye detection

(b) Extended ROI

(c) Intensity adjustment

Fig. 1: Eye detection and intensity adjustment

ing ElectroOculoGraphy (EOG). Using time delineation features of the eye blinks, a recognition rate of 93.75% and equal error rate of 7.45% were achieved over a database of 40 subjects. A different technique for registering eye blinking patterns was followed by J. Espinosa el al. [3], where eye blinking patterns were registered remotely using the intensity of light reflect from the eyes. Again, time delineation features from blinking patterns were adopted along with speed, acceleration and power patterns. A correct recognition rate up to 99.7% was achieved in identification mode. In addition, eye blinking patterns were found successful in improving performance and security of other biometric systems in a multi-modal setup. Examples of these multi-modal systems are EEG-based biometric systems [10, 11] and facial recognition systems [12].

3. PROPOSED APPROACH

3.1. Recording protocol

The database used in this paper was collected after having an approval from the research ethics board of the University of Toronto (Protocol # 00035655). Eye movement and blinking recordings are done simultaneously using the Gazepoint GP3 device from 22 participants. The stimulation attempts to create a similar environment to that of a driver. Consequently, the simulation is a video clip of a real driving session in downtown Toronto [1]. The video is trimmed into three 5 min clips. This is done so that there are 2 videos for training and 1 for testing (Section 4). The subject is seated about half a meter away from the Gazepoint device, with it looking up towards the subject at an angle of about 30 degrees. Before each video is played, Gazepoint is first calibrated in order to get a better gaze estimate. A 15 sec introduction with instructions as well as a cross sign for the subjects to center their gaze are displayed before each video begins. Since Gazepoint software does not provide us with the raw frames captured from the IR camera, we use a third party software (OBS studio) to record a video of Gazepoint control window at the same frame rate of the GP3 device (60 fps) [2]. The recorded videos are used to extract the eye blinking patterns as discussed in the next section.

3.2. Eye blinking registration and pre-processing

Gazepoint has a built-in feature, where a 200×155 green pixel box is placed around each eye (Fig. 1a). As only the data around the eye is desired, the video is cropped around the green box for each eye. However, the green box does not cover the corners of the eye, so the region acquired is elongated (about 70-pixel lengths in each x-direction as shown in Fig. 1b). The pre-processing part for eye blinking involves enhancing the contrast between the flat signal, and the peaks caused by blinking; in order to more consistently separate the two for feature extraction. The pre-processing steps are as follows.

The image is already gray scaled by Gazepoint, making the image very dark and there is not a large contrast between the different intensities of the pixels. Consequently, an adjustment function is used in the Region of Interest (ROI) that saturates the top and bottom 1% of all pixel values. This strengthens the contrast, making the different regions of the eyes more apparent (Fig. 1c). The pixels in this region is then summed every frame to create an intensity profile.

After creating an intensity profile, eye blinking patterns can be easily extracted and validated using Gazepoint blinking validation flag (BKID). The flag has a non-zero value if a blink occurs. To take advantage of this flag, the video frames (from OBS) and BKID signal (from Gazepoint) are synchronized as shown in Fig. 2. After locating the frames on the intensity profile where the blink flag is on, a threshold is used to determine if there is a peak in these frames, which would result from a blink. Once the frame with the peak of a blink is found, 15 frames before and 30 frames after the peak are extracted as well to create a consistent 46 frame long signal capturing the blinking pattern.

Finally after extracting the blinking profiles, they are smoothed using a Spline smooth curve (Fig. 3a). This process is done by creating continuous piecewise 3^{rd} degree polynomials to create a smoothed curve representing each blink.

3.3. Eye blinking feature extraction

After extracting the blinks from each subject, the features for comparison are then extracted. The features used are similar to those used in [9, 3]. The first feature is the signal itself (Fig. 3a), as well as the first and second derivative of the signal which are the speed and acceleration profiles of the blinks (Fig. 3b and 3c). Time delineation features are also extracted from the blink pattern. This includes the onset and offset (onset is the closing of the eyelids, offset is the opening of the eyelids) area, duration, slope, and energy (the sum of the square of the amplitude), and the strength of the blink.

3.4. Eye movements pre-processing

The Gazepoint device provides information about the gaze position on the screen plane (2D plane) along with validation flags. The average gaze position (BPOGX, BPOGY) is used as raw data in this work. The pre-processing stage comprises filling gaps, smoothing, and eye movement classification into fixations and saccades as described hereafter.

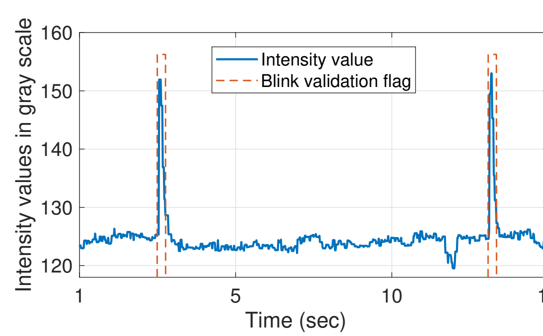


Fig. 2: The intensity profile synchronized with the blink validation flag in order to locate blinks in the signal

[1]Driving session video: `https://www.youtube.com/watch?v=rM8dbiH0kfY`

[2]OBS studio: `https://obsproject.com/download`

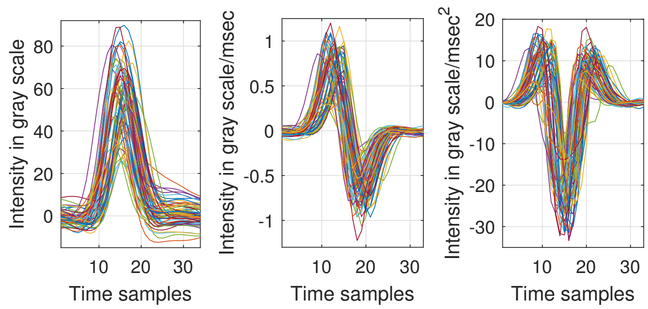


Fig. 3: The intensity profile of subject #10 during blinking (a), as well as their speed (b) and acceleration profiles (c)

For some time instances, eyes are not detected due to system issues like light reflection or algorithm failure, or user-related issues like blocking their eyes or blinking [13]. The gaze validation flag (BPOGV) is used to detect these time instances and the gaze position in these times is linearly interpolated using the gaze data before and after these time points. After gap filling, the 5-minute gaze data are filtered using Savitzky-Golay filter for signal smoothing. The polynomial order and window size used are 3 and 17 samples. To classify eye movements into fixation and saccades; the velocity based method (I-VT) [13, 14] is employed. The velocity from the filtered gaze position in the 2D plane (screen plane) is estimated first as follows:

$$V = \frac{\sqrt{(X_f(i+1) - X_f(i))^2 + (Y_f(i+1) - Y_f(i))^2}}{F_t(i+1) - F_t(i)} \quad (1)$$

where V, X_f, Y_f, and F_t denote velocity, filtered gaze in X and Y directions, and frame time, $i = 1, 2, \ldots, N-1$ denotes the frame index and N is the total frame number. Since the angular X and Y position are not available from Gazepoint, a threshold of unit $pixels/sec$ is used for eye movement classification. A velocity greater than 1.8 $pixels/msec$ is classified as saccade and below this value is a fixation. To reject short fixation, any detected fixation with a duration less than 100 $msec$ is considered as saccade [7].

3.5. Eye Movement feature extraction

After eye movement classification, features are extracted from saccades and fixation separately similar to those used in the biometric system proposed by A. George and A. Routray [7]. The features based on the angular position of the eye in [7] were not used here since angular position data is not provided by Gazepoint. A total of 12 and 35 features are extracted from fixation and saccades, respectively, comprising duration, amplitude, and statistical features (mean, median, maximum, standard deviation, skewness and kurtosis). These two sets of features are concatenated to form an eye movement feature vector with a dimension of 47.

3.6. Multi-modal system: fusion of eye Movements and blinking

3.6.1. Feature-level fusion

For feature fusion, two techniques are adopted; feature concatenation (F_{CAT}) and Canonical Correlation Analysis (CCA - F_{CCA}). Feature concatenation is done by simply concatenating the training features from eye movement and eye blinking together before feeding to the classifier. Since, features are of different scales (units), z-score normalization was applied on the features before training the classifier. The same is done for the testing features; features are concatenated and then normalized by the mean and standard deviation values (estimated in the training phase).

For CCA, the two sets of features are linearly projected onto another space where the correlation between these two sets is maximized [15, 16]. Using CCA, a new feature space is estimated where similar information from both datasets are preserved. Moreover, in the new space, information gain from features is maximized by minimizing correlation of the canonical variates. The pair of projecting matrices, W_X and W_Y, are obtained from the training feature matrices of the eye movement and blinking data. Then, the new pair of projected feature sets $\hat{X} = W_X^T X$ and $\hat{Y} = W_Y^T Y$ are concatenated together to form the fused feature set (parallel fusion) before feeding to the classifier. Then, the testing feature sets are projected using the W_X and W_Y matrices (estimated during training) and then concatenated together.

3.6.2. Score-level fusion

For score level fusion, three simple fusion techniques are adopted which are: sum (S_{SUM}), product (S_{PROD}), and max (S_{MAX}) score fusion techniques. Sum score fusion simply average the score from the eye movement and blinking biometric systems. In case of product score fusion, since the fused score for each class does not sum to 1, the fused score is normalized by dividing it by the sum of matching scores to each class.

3.7. Classification

The classifier used is Linear Discriminant Analysis (LDA). Other classifiers tested were multi-class SVM, and RBF Neural Networks, however, their performance was very subpart compared to the LDA classifier. LDA attempts to project the feature space into a smaller subspace, while maximizing the distance between the classes [17]. This ensures that the new feature space will separate each subject as far apart as possible, making it easier to identify the test feature examples. It also assumes every subject's (class - c) feature is Gaussian distributed with a mean μ_c and a standard deviation Σ_c, while assuming every class has the same covariance. It then uses the optimum Bayes rule to make the classifier decision, in order to maximize the logarithm of the posterior probability given by

$$\log(P(c|x_s)) = -\frac{1}{2}(x_s - \mu_c)^T \Sigma_c^{-1}(x_s - \mu_c) + \log(P(c)) \quad (2)$$

where $c = 1, 2, ..., N_s$. In the equation above, $\Sigma_c = \Sigma$ for LDA. N_s is the total number of subjects, where $P(c|x_s)$ is the probability of matching the features from an unknown subject x_s to a class c. $P(c)$ is the prior probability. The equation itself represents a linear equation where the boundaries between the different classes is determined by as straight line.

4. EXPERIMENTAL SETUP AND RESULTS

In this paper, the system's performance is evaluated using hold one out cross-validation. In other words, one trial (out of three) is used for building the test pool of features and the other two are used for building the training features pool. This is conducted for three runs and a new test trial is selected every run. Moreover, since eye movement and blinking patterns are varying in time even for the same subject, features from multiple patterns, selected randomly, are extracted and averaged to form one testing or training example. Based

Table 1: Evaluation of eye movement and blinking biometric system in single-modality setup (white and shaded cells show CRR and EER, respectively, in the form *mean (std) %*)

T_A (sec)	30	60	90	120
Eye movement	62.1 (2.5)	74.9 (1.7)	80.6 (1)	**83.5 (2.2)**
	14.2 (1.3)	11.3 (1.6)	9.8 (1.7)	**9.3 (2)**
Eye blinking	76.6 (4.9)	80.9 (4)	82.8 (3.3)	**83.8 (3.3)**
	11.2 (1.7)	10.3 (1.7)	**10 (1.6)**	**9.9 (1.9)**

on the database recorded, the average number of eye movements (fixations and saccades) and blinks per minute are 50 and 15, respectively. The Correct Recognition Rate (CRR) and the Equal Error Rate (EER) are used as a performance metric in identification and verification modes of authentication, respectively.

In each run, 500 examples per subject are computed by averaging N_{avg} random feature vectors from the training pool. Similarly, 50 examples per subject are computed by averaging N_{avg} random feature vectors from the testing pool. To achieve a better estimate of the system's performance, the random sub-sampling is performed 10 times and the averaged performance metric is reported. The N_{avg} value reflects the total time required for user authentication. In this paper, the system was evaluated using different values of N_{avg}; for eye movements $N_{avg} = [25, 50, 75, 100]$, and for eye blinking, $N_{avg} = [8, 15, 23, 30]$. Based on our database, these values represent an average authentication recording time (T_A) of 30, 60, 90, and 120 seconds, respectively. For subjects with number of blinks per trial is less than 30, their blinking data are doubled by adding random noise to their original blinks.

Table 1 summarizes the results for the eye movement and eye blinking biometric system (single modality setup) in identification and verification modes for different authentication time lengths T_A. Based on the achieved results, averaging larger number of features (longer recording time) provides better estimate of the eye movement and blinking behavior for each subject, hence, improving the performance of the biometric authentication systems. In general, both modalities achieved similar performance in the two modes of authentication. For eye movements, the highest CRR and the lowest EER achieved were 83.5% and 9.3% using an authentication time of 2 min ($N_{avg} = 100$). Similarly, for eye blinking-based biometric system, the best CRR and the lowest EER achieved are 83.8% and 9.9%, respectively. Moreover, for a shorter authentication time, e.g. 30 sec., the eye blinking based biometric system outperforms the eye movement biometric system, however, both of them converges to a similar performance after an authentication time of 2 min.

Our proposed system for eye movement and blinking biometrics (single-modality setup) achieved lower CRR and higher EER values compared to previous works [7, 3]. However, our database used an eye tracking device with a lower resolution (60 Hz) and simulates a real scenario (driver authentication). Databases in [7, 3] were built in a controlled environment were subjects have to rest their head on a chin-rest. Moreover, the experimental setup in [3] is questionable as the eye blinking data are shared between the averaged training and testing samples.

Table 2 shows the performance of the multi-modal system for the five fusion techniques discussed in Section 3.6. A noticeable improvement in the identification and verification modes is achieved for all the recording time lengths in comparison to a single modality using eye movement or eye blinking. Out of the five fusion techniques, F_{CCA} achieved the best identification and verification rates

Table 2: Evaluation of eye movement and blinking biometric system in multi-modal setup (white and shaded cells show CRR and EER, respectively, in the form *mean (std) %*)

T_A (sec)	30	60	90	120
$\mathbf{F_{CAT}}$	89.2 (4.1)	92.7 (4.6)	93.7 (4.4)	94.5 (4.5)
	6.4 (0.9)	5.3 (1.6)	7 (1.9)	8.3 (1.8)
$\mathbf{F_{CCA}}$	90.1 (4.1)	93.7 (4.5)	95 (4.2)	**95.8 (4.1)**
	5.3 (0.6)	3.3 (0.7)	2.6 (0.8)	**2.2 (0.7)**
$\mathbf{S_{SUM}}$	83.2 (4.6)	88.9 (3.9)	91.2 (2.8)	92.4 (2.6)
	8 (1)	6.5 (1.1)	5.7 (1.3)	5.2 (1.4)
$\mathbf{S_{PROD}}$	89.2 (4.2)	92.6 (4.5)	93.7 (4.3)	94.5 (4.5)
	6.4 (0.5)	5.2 (0.8)	6.7 (1)	8.5 (1.5)
$\mathbf{S_{MAX}}$	82.4 (4.7)	88.6 (3.9)	91 (2.9)	91.9 (4)
	9 (0.9)	6.9 (0.9)	6 (1.1)	5.4 (1.3)

of 95.8% and 2.2% respectively. Fusing eye movement and blinking features using CCA shows approximately 12% improvement in the identification mode over single modality system. On the other hand, CCA feature fusion lowers the EER in the verification mode by about 8% making it comparable to the equal error rates of conventional biometric traits like fingerprint [18].

5. CONCLUSION AND FUTURE WORK

The experimental results in this paper showed that both eye movement and eye blinking biometrics on its own had an underperforming success rate with both having an identification rate of around 83.5%, and verification error rate of about 9.5%. Although not as good as the other leading experiments, this project achieved these scores with a much more realistic environment, and non-specialized equipment. By combining the two system, a noticeable improvement in the recognition performance was achieved. A success rate in identification of 95.8% was obtained, and for verification, the error rate was 2.2%. This proves the viability of biometric authentication in real life scenarios such as driver authentication, by using the fusion of eye movement and eye blinking biometrics as a continuous remote authentication system. Furthermore, the employment of these modalities may be extended to other applications such as security.

The next step for this work would be to conduct the same experiment with a camera attached to the front of a car's dashboard, facing towards the driver. As there are different (and changing) lighting conditions and more movements done by the driver, there may be more outliers and variation in the data that may alter the results. An improvement that can be addressed in future work is to use a more accurate eye tracking device with a higher temporal resolution. The Gazepoint camera has a very low frame rate compared to the other cameras used in similar work [7] which has a frame rate of 1000 Hz. This attributed to much of the error in our experiment, which could have been avoided. With a higher frame rate, a more accurate estimation of eye movement and blinking patterns will be achieved resulting in an improved performance.

6. REFERENCES

[1] Javier Galbally, Sébastien Marcel, and Julian Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014.

[2] J Galbally-Herrero, J Fierrez-Aguilar, JD Rodriguez-Gonzalez, Fernando Alonso-Fernandez, Javier Ortega-Garcia, and M Tapiador, "On the vulnerability of fingerprint verification systems to fake fingerprints attacks," in *Carnahan Conferences Security Technology, Proceedings 2006 40th Annual IEEE International*. IEEE, 2006, pp. 130–136.

[3] Julián Espinosa, Begoña Domenech, Carmen Vázquez, Jorge Pérez, and David Mas, "Blinking characterization from high speed video records. application to biometric authentication," *PloS one*, vol. 13, no. 5, pp. e0196125, 2018.

[4] Oleg V Komogortsev and Ioannis Rigas, "Bioeye 2015: Competition on biometrics via eye movements," in *Biometrics Theory, Applications and Systems (BTAS), 2015 IEEE 7th International Conference on*. IEEE, 2015, pp. 1–8.

[5] Ioannis Rigas and Oleg V Komogortsev, "Current research in eye movement biometrics: An analysis based on BioEye 2015 competition," *Image and Vision Computing*, vol. 58, pp. 129–141, 2017.

[6] Pawel Kasprowski and Katarzyna Harezlak, "Using dissimilarity matrix for eye movement biometrics with a jumping point experiment," in *Intelligent Decision Technologies 2016*, pp. 83–93. Springer, 2016.

[7] Anjith George and Aurobinda Routray, "A score level fusion method for eye movement biometrics," *Pattern Recognition Letters*, vol. 82, pp. 207–215, 2016.

[8] Narishige Abe, Shigefumi Yamada, and Takashi Shinzaki, "A novel local feature for eye movement authentication," *Biosig 2016*, 2016.

[9] Sherif N Abbas and M Abo-Zahhad, "Eye blinking eog signals as biometrics," in *Biometric Security and Privacy*, pp. 121–140. Springer, 2017.

[10] Mohammed Abo-Zahhad, Sabah M Ahmed, and Sherif N Abbas, "A new multi-level approach to eeg based human authentication using eye blinking," *Pattern Recognition Letters*, vol. 82, pp. 216–225, 2016.

[11] Qunjian Wu, Ying Zeng, Chi Zhang, Li Tong, and Bin Yan, "An eeg-based person authentication system with open-set capability combining eye blinking signals," *Sensors*, vol. 18, no. 2, pp. 335, 2018.

[12] Santosh Tirunagari, Norman Poh, David Windridge, Aamo Iorliam, Nik Suki, and Anthony TS Ho, "Detection of face spoofing using visual dynamics," *IEEE transactions on information forensics and security*, vol. 10, no. 4, pp. 762–777, 2015.

[13] Anneli Olsen, "The tobii I-VT fixation filter: Algorithm description," Available Online: `https://www.tobiipro.com/siteassets/tobii-pro/learn-and-support/analyze/how-do-we-classify-eye-movements/tobii-pro-i-vt-fixation-filter.pdf/?v=2012`, Accessed: 2018-08-23.

[14] Dario D Salvucci and Joseph H Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proceedings of the 2000 symposium on Eye tracking research & applications*. ACM, 2000, pp. 71–78.

[15] Mohammad Haghighat, Mohamed Abdel-Mottaleb, and Wadee Alhalabi, "Fully automatic face normalization and single sample face recognition in unconstrained environments," *Expert Systems with Applications*, vol. 47, pp. 23–34, 2016.

[16] Quan-Sen Sun, Sheng-Gen Zeng, Yan Liu, Pheng-Ann Heng, and De-Shen Xia, "A new method of feature fusion and its application in image recognition," *Pattern Recognition*, vol. 38, no. 12, pp. 2437–2448, 2005.

[17] Sebastian Raschka, "Linear discriminant analysis, bit by bit," Available Online: `https://sebastianraschka.com/Articles/2014\_python\_lda.html#%E2%80%93-bit-by-bit`, 2014, Accessed: 2018-08-20.

[18] Raffaele Cappelli, Matteo Ferrara, Annalisa Franco, and Davide Maltoni, "Fingerprint verification competition 2006," *Biometric Technology Today*, vol. 15, no. 7-8, pp. 7–9, 2007.