Personal Authentication with Eye Movement Features During PIN Input

Shion Tagawa* and Hironobu Takano†
* Toyama Prefectural University, Toyama, Japan
E-mail: tagawa@neu.pu-toyama.ac.jp
† Toyama Prefectural University, Toyama, Japan
E-mail: takano@pu-toyama.ac.jp

***Abstract*—In recent years, personal authentication is used in various situations. Among them, the personal authentication using biometric information is becoming widespread. Conventional biometrics have a spoofing problem. Personal authentication using eye movement is considered to be more difficult to spoof. Biometrics research using eye movement has been conducted in the case where the eye movement range is relatively wide. In this study, we aim at development of the personal authentication method using eye movement in the case of narrow range of eye movement such as smartphone and ATM operation. In this paper, we investigated effective features of eye movement in the case where the range of eye movement is narrow. The authentication accuracy was evaluated using equal error rate (EER). In addition, we examined whether the authentication accuracy could be improved by score level fusion. From the experimental results, it was found that complex eye movement (CEM) had the best authentication accuracy when the range of eye movement was narrow, and the authentication accuracy was improved by score level fusion.**

I. Introduction

In recent years, personal authentication is used in various situations. Personal authentication includes passwords, ID cards, biometric authentication, etc. The personal identification using password and ID cards has risks of loss, forgetting, and shoulder-surfing. Instead of them, personal authentication by biometrics has become widespread. The biometric authentication using face, fingerprint, and iris that are static and physical features are used in smartphones and ATMs. However, biometrics using physical features has a problem of spoofing. If once biometric information is spoofed, it is difficult to change biometric features which are registered in an authentication system. Therefore, biometric authentication using behavior characteristics that are difficult to spoof has become necessary. Eye movement is one of biometric authentication with behavioral features. Biometric authentication using eye movement has advantages that are robustness against spoofing, continuous authentication, and unconscious authentication without a specific action by user [1]. Conventional authentication methods using eye movement adopt the Mel-frequency Cepstrum Coefficients (MFCC) [1] – [2], Complex Eye Movement (CEM) such as saccades or fixations [3], scan paths using the locus of gaze [4], etc. In addition, the effective features and recognition methods for eye movement authentication have been actively investigated [5] – [10]. Competitions for eye movement authentication has been also conducted [11] – [14]. Previous studies for eye movement authentication were deal with the wide range of eye movement compared with that in the case of smartphone or ATM operation. The viewing angle of a typical PC display is 43.6° in the horizontal direction and 24.8° in the vertical direction. On the other hand, the viewing angle of the smartphone (iPhone7) is 5.6° in the horizontal direction and 9.9° in the vertical direction. We aim at developing personal authentication method using eye movement in the case of narrow range of eye movement such as smartphone and ATM operation. In this study, we focused on the development of eye movement authentication during PIN input. The personal authentication with PIN code includes the risk of leakage by shoulder surfing. The eye movement authentication combining with PIN has an advantage that imposters cannot release the lock even if the PIN code is leaked. In this paper, we investigated effective features in the case where the range of eye movement was narrow. For improvement of personal authentication accuracy, we also evaluated score level fusion using degrees of similarity obtained from three types of features.

II. Experimental Method

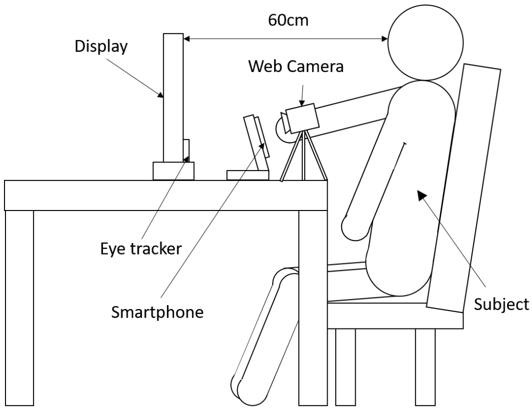
In this experiment, gaze data was obtained as a time-series waveform of X and Y coordinates. However, the gaze cannot be measured during blinking. The absence of gaze data was compensated by interpolating with piecewise cubic Hermite interpolating polynomial. Authentication accuracy was evaluated using Equal Error Rate (EER). The EER is the ratio that the FAR (False Acceptance Rate) and FRR (False Rejection Rate) become same. The FAR is the ratio that the biometric security system incorrectly accepts an access attempt by an imposter. The FRR is the ratio that the biometric security system incorrectly rejects an access attempt by an authorized user. The lower value of EER indicates higher accuracy of personal authentication. This section describes a measuring method of gaze and a feature extraction method. In this study, dynamic time warping (DTW) is used recognition method as proposed method. The CEM and features based on cepstrum coefficients are used as conventional method. In addition, weighted fusion and fusion by logistic regression are used as the method of score level fusion.

A. Measuring Methods of Eye Movement

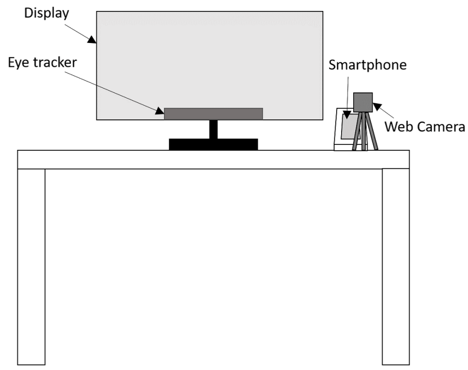
An experimental system for measuring gaze is shown in Fig.1. In the experiment, iPhone7 was used for entering PIN code. For measuring gaze during iPhone7 operation, a moving image obtained by capturing the screen of iPhone7 with a webcam was presented on a display of a desktop PC. The screen presented on the iPhone7 is shown in Fig.2. Gaze measurement was performed using Tobii Pro X2-30 which has a sampling rate of 30 Hz, an accuracy of 0.4°, and a precision of 0.32° [15]. The eye tracker was attached to the lower part of the PC display. The distance from the subject to the PC display was about 60 cm. The PIN code was set as "1065" for all subjects. The subject operated the fixed smartphone with the right hand and input the PIN code. At this time, subjects entered the PIN code while looking at the input screen of iPhone7 presented on the PC display. Three sessions were conducted for each subject. One session consists of 5 trials. The subject input the PIN code one time in one trial. Eye tracker was calibrated for measuring the eye movement before staring each experiment. The gaze measurement was started when the START button was input. Then, the measurement was finished after the END button was input. The subject were 11 healthy students in his/her twenties which are 9 males and 2 females. These experiments were approved by the Ethics Committee of Toyama Prefectural University. In addition, written informed consent was obtained from each participant.

B. DTW

The DTW was used for matching of time-series waveform of x and y coordinates of gaze position. The distance calculated by DTW is shown in (1) – (3). The R and A in (1) are registration and authentication information which are composed of the x and y components of the gaze, respectively. In addition, we define $\boldsymbol{R_i} = (x_{ri}, y_{ri}), \boldsymbol{A_j} = (x_{aj}, y_{aj})$ using x and y coordinate values of gaze. The DTW distance was normalized to 1 by (4). The registration information is an averaged waveform which is calculated with all gaze data of one session (5 trials). Authentication information is an averaged waveforms of gaze in other one session (5 trials). The gaze data was resampled to 400 points for obtaining the averaged waveform. The legitimate registrant and non-registrant were identified by changing the decision threshold for every 0.001.



(a) side



(b) front

Fig. 1: Experimental System.

Fig. 2: Experimental screen on iPhone7.

$$R = \begin{pmatrix} x_{r1} & y_{r1} \\ x_{r2} & y_{r2} \\ \vdots & \vdots \\ x_{ri} & y_{ri} \\ \vdots & \vdots \\ x_{rn} & y_{rn} \end{pmatrix}, A = \begin{pmatrix} x_{a1} & y_{a1} \\ x_{a2} & y_{a2} \\ \vdots & \vdots \\ x_{ai} & y_{ai} \\ \vdots & \vdots \\ x_{an} & y_{an} \end{pmatrix} \quad (1)$$

$$d(\boldsymbol{R_i}, \boldsymbol{A_j}) = \sqrt{(x_{ri} - x_{aj})^2 + (y_{ri} - y_{aj})^2} \quad (2)$$

$$\begin{cases} g(\boldsymbol{R_1}, \boldsymbol{A_1}) = 2d(\boldsymbol{R_1}, \boldsymbol{A_1}) \\ g(\boldsymbol{R_i}, \boldsymbol{A_j}) = min[g(\boldsymbol{R_{i-1}}, \boldsymbol{A_{j-1}}) + 2d(\boldsymbol{R_i}, \boldsymbol{A_j}), \\ \qquad g(\boldsymbol{R_i}, \boldsymbol{A_{j-1}}) + d(\boldsymbol{R_i}, \boldsymbol{A_j}), \\ \qquad g(\boldsymbol{R_{i-1}}, \boldsymbol{A_j}) + d(\boldsymbol{R_i}, \boldsymbol{A_j})] \\ D_{DTW}(R, A) = g(R_I, A_J) \end{cases} \quad (3)$$

$$D_n(R, A) = \frac{1}{1 + D_{DTW}(R, A)} \quad (4)$$

C. CEM

In previous studies, the CEM used the peak velocity and the standard deviation of duration of saccades, the average and the standard deviation of the duration of fixation, etc. [3]. The CEM used in this study are shown in Table I. The f_5 and f_6 in Table I are the dispersion D shown in (5) which is obtained by adding differences between maximum and minimum values of x and y coordinate of gaze position [3]. The feature vector $\boldsymbol{F_{CEM}}$ used for registration and authentication is shown in (6). When the standard deviation is 0 or eye movement data cannot be obtained, the element of feature vector is dealt as a missing value. After standardization transformation for each element of the feature vector, F was normalized from 0 to 1 using (7). For decision criteria of fixation, the fixation time is 100 ms and the spatial range of fixation is within 70 pixels. The registration information is an average value of feature vectors in one session (5 trials). The authentication information is an average value of feature vectors in other one session (5 trials). The legitimate registrant and non-registrant were identified by changing the decision threshold for every 0.001.

$$D = max(X) - min(X) + max(Y) - min(Y) \quad (5)$$

$$\boldsymbol{F_{CEM}} = \left(\; f_1, f_2, \cdots, f_{10} \; \right)^T \quad (6)$$

$$\boldsymbol{F'_{CEM}} = \frac{\boldsymbol{F_{CEM}}}{||\boldsymbol{F_{CEM}}||} \quad (7)$$

TABLE I: Features of CEM.

f_1	Average of Fixation Duration
f_2	Average of Saccade Duration
f_3	Standard Deviation of Fixation Duration
f_4	Standard Deviation of Saccade Duration
f_5	Dispersion of Fixation
f_6	Dispersion of Saccade
f_7	Standard Deviation of Fixation (X)
f_8	Standard Deviation of Fixation (Y)
f_9	Peak Velocity of Saccade (X)
f_{10}	Peak Velocity of Saccade (Y)

D. Cepstrum-based Features

The MFCC is an effective feature of eye movement authentication [1] – [2]. Abe et al. has proposed the cepstrum based feature that exclude the mel-scale conversion [1]. The capstrum based feature is obtained from (8) – (10). N represents the number of FFT points which is set as 256 in this study. First, power spectrums of eye movement waveforms are obtained by the fast Fourier transform. Next, the logarithmic power spectrums (P_x, P_y) are calculated. Finally, discrete cosine transformation is performed to obtain the capstrum based feature (C_x, C_y). According to Abe et al., the capstrum based feature are characterized by 12 coefficients representing low frequency components [1]. Therefore, 12 points of low frequency components are used in this study. As shown in (11), the cepstrum based feature is obtained by combining x and y components of 24 cepstrum coefficients. The registration information is an average value of capstrum based feature obtained in one session (5 trials). The authentication information is an average value of capstrum based feature in other one session (5 trials). The legitimate registrant and non-registrant were identified by changing the decision threshold for every 0.001.

$$\begin{cases} F_x(n) = \sum_{t=0}^{N-1} x(t) e^{(-j2\pi \frac{n}{N} t)} \\ F_y(n) = \sum_{t=0}^{N-1} y(t) e^{(-j2\pi \frac{n}{N} t)} \end{cases} \quad (8)$$

$$\begin{cases} P_x(n) = \log |F_x(n)|^2 \\ P_y(n) = \log |F_y(n)|^2 \end{cases} \quad (9)$$

$$\begin{cases} C_x(n) = \sum_{t=0}^{N-1} P_x(t) \cos(\frac{\pi}{N}(n + \frac{1}{2})t) \\ C_y(n) = \sum_{t=0}^{N-1} P_y(t) \cos(\frac{\pi}{N}(n + \frac{1}{2})t) \end{cases} \quad (10)$$

$$\begin{aligned} \boldsymbol{F_{CEPS}} = (&C_x(1), C_x(2), \cdots, C_x(12), \\ &C_y(1), C_y(2), \cdots, C_y(12))^T \end{aligned} \quad (11)$$

E. Score Level Fusion

Score level fusion was performed using three types of the degree of similarities that are obtained from matching results for DTW, CEM, and cepstrum based feature. Two types of score level fusions with weighted summation and logistic regression are evaluated.

The score level fusion by weighted summation is shown in (12). In (12), $\omega_{w1} \sim \omega_{w3}$ indicate weights. S_{wD}, S_{wC}, and S_{wM} indicate scores obtained with DTW, CEM, cepstrum based feature, respectively. S_{wF} is the fused score obtained by weighed summation. The weight was changed for every 0.01 under the condition that summation of $\omega_{w1} \sim \omega_{w3}$ is equal to 1.

$$S_{wF} = \omega_{w1} S_{wD} + \omega_{w2} S_{wC} + \omega_{w3} S_{wM} \tag{12}$$

Next, score level fusion using logistic regression is shown in (13). The objective variable of logistic regression is a score S_{lF} after fusion, and the regression coefficient is weights $\omega_{l1} \sim \omega_{l3}$. S_{lD}, and S_{lC}, S_{lM} which are scores of DTW, CEM, and cepstrum based feature are used as explanatory variables. The b in (13) is a constant term. The S_{lD}, S_{lC}, and S_{lM} were randomly selected for calculation of regression coefficients.

$$S_{lF} = \frac{1}{1 + \exp(-(b + \omega_{l1} S_{lD} + \omega_{l2} S_{lC} + \omega_{l3} S_{lM}))} \tag{13}$$

III. Experimental Results

A. Authentication Results of Each Feature

Figure 3 shows ROC curves obtained using each feature. The horizontal axis is FRR and vertical axis is FAR. The yellow line is a ROC curve by DTW, the blue line is a ROC curve by CEM, and the red line is a ROC curve by cepstrum based feature. The EERs obtained from Fig.3 are shown in Table II. From this result, it was found that CEM has the best authentication accuracy when the range of eye movement is narrow.

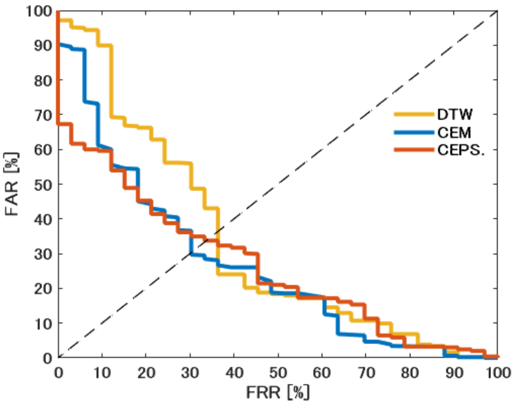


Fig. 3: ROC curves obtained using DTW, CEM, and capstrum based feature.

TABLE II: EERs obtained using DTW, CEM, and capstrum based feature.

	EER [%]
DTW	36.36
CEM	30.10
CEPS.	33.54

B. Authentication Results of Score Level Fusion

Figure 4 shows ROC curves obtained using score level fusion. The horizontal axis is FRR and the vertical axis is FAR. The yellow and blue lines are ROC curves obtained by score level fusion of weighted summation and logistic regression, respectively. The EERs obtained from Fig.4 are shown in Table III. This result indicates that the accuracies of personal authentication are improved by both methods of score level fusion.

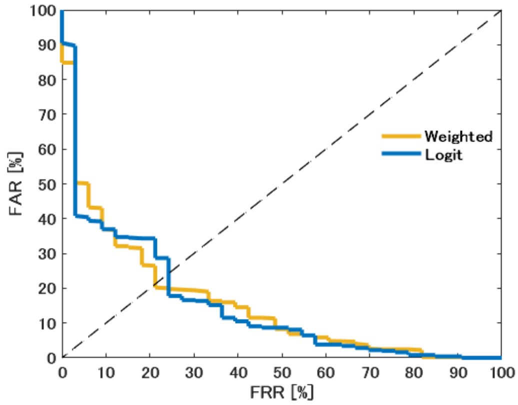


Fig. 4: ROC curves obtained using score level fusions.

TABLE III: Weights and EER obtained using score level fusions.

	EER [%]	ω_1	ω_2	ω_3	b
Weighted	20.61	0.68	0.16	0.16	–
Logit	24.04	62.36	24.77	11.72	-34.23

IV. Discussion

From experimental results shown in Fig.3 and Table II, the authentication accuracy of CEM is the highest among three types of features. However, the accuracy of authentication was lower than that of previous studies. The sufficient eye movement data for extracting CEM cannot be obtained because of the low sampling rate of eye tracker and a short measurement time (about 8 to 12 seconds).

The score level fusion provides the improvement of accuracy of personal authentication using eye movement. Thus, the score level fusion is effective when the measurement time of gaze is short. The authentication accuracy of score level fusion with weighted summation was higher than that with logistic regression. Furthermore, the weights of both score level fusions indicated that DTW especially contributed to personal authentication with gaze information. However, since the number of matching data for learning of logistic regression is insufficient, the personal authentication accuracy for score level fusion with logistic regression should be evaluated with gaze data obtained from many more subjects.

V. Conclusions

We focused on biometrics authentication method using eye movement when the eye movement range was narrow. In this paper, we estimated effective features for eye movement authentication when the eye movement range was narrow and measuring time was short. The accuracies of eye movement authentication using DTW, CEM, and cepstrum based feature were evaluated with EER. The experimental results indicated that EER of CEM was the lowest among three types of features. In addition, we evaluated the authentication accuracy of the score level fusion of weighted summation or logistic regression that used scores obtained from three types of features. The experimental result indicated that the authentication accuracies of both fusions were improved compared with those before fusion. Moreover, the authentication accuracy of weighted summation was higher than that of logistic regression. In future work, we will investigate the relationship between the authentication accuracy and sampling rate of eye tracker or measuring time of eye movement with many more subjects.

References

[1] Narishige Abe, Shigefumi Yamada, and Takashi Shinzaki, " Novel Local Feature for Eye Movement Authentication," Proc. International Conference on the Biometrics Special Interest Group(BIOSIG) 2016.

[2] Nguyen Viet Cuong, Vu Dinh, and Lam Si Tung Ho, "Mel-frequency Cepstral Coefficients for Eye Movement Identification," Proc. IEEE 24th International Conference on Tools with Artificial Intelligence, pp.253-260, 2012.

[3] Corey D. Holland and Oleg V. Komogortsev, "Complex Eye Movement Pattern Biometrics: Analyzing Fixations and Saccades," Proc. ICB 6th International Conference on Biometrics, pp.1-8, 2013.

[4] Corey D. Holland and Oleg V. Komogortsev, "Biometric Identification via Eye Movement Scanpaths in Reading," Proc. 2011 International Joint Conference on Biometrics, 2011.

[5] Andrey V. Lyamin and Elena N. Cherepovskaya, "An Approach To Biometric Identification by Using Low-Frequency Eye Tracker," IEEE Transactions on Information Forensics and Security, Vol.12, No.14, pp.881-891, 2017.

[6] Ioannis Rigas, George Economou, and Spiros Fotopoulos, "Biometric identification based on the eye movements and graph matching techniques," Pattern Recognition Letters, pp.786-792, 2012.

[7] Anjith George and Aurobinda Routray,"A Score-level Fusion Method for Eye Movement Biometrics," Pattern Recognition Letters, pp.207-215, 2016.

[8] Oleg Komogortsev, Alexey Karpov, and Corey Holland, "Oculomotor Plant Characteristics: The Effects of Environment and Stimulus," IEEE Transactions on Information Forensics and Security, pp.621-632, 2016.

[9] Pawel Kasprowski and Katarzyna Harezlak, "Fusion of Eye Movement and Mouse Dynamics for Reliable Behavioral biometrics," Pattern Analysis and Applications, Vol.21, pp.91-103, 2018.

[10] Hong-Jun Yoon, Folami Alamudun, Kathy Hudson, Garnetta Morin-Ducote, and Georgia Tourassi, "Deep Gaze Velocity Analysis During Mammographic Reading for Biometric Identification of Radiologists," Journal of Human Performance in Extreme Environments, Vol.14, 2018.

[11] EMVIC2012, BTAS 2012, 2012, http://www.kasprowski.pl/emvic2012/

[12] EMVIC2014, IJCB 2014, 2014, http://www.emvic.org/

[13] BioEye2015, BTAS 2015, 2015, https://bioeye.cs.txstate.edu/

[14] Look And Click, ICB, 2016, http://www.kasprowski.pl/lac/results_final.php

[15] Tobii pro, urlttps://www.tobiipro.com/product-listing/tobii-pro-x2-30/