Continuous biometric authentication for smartphones considering usage environments

Yuka Watanabe and Yasushi Yamazaki
The University of Kitakyushu, Fukuoka, Japan
E-mail: c1mca013@eng.kitakyu-u.ac.jp, y-yamazaki@kitakyu-u.ac.jp Tel/Fax: +81-93-695-3259

***Abstract*— With the rapid spread of smartphones, there is a growing need for them to have a user authentication function. However, conventional user authentication methods using PINs, passwords, pattern locks, etc. have a problem in that user authentication is not performed continuously after the first authentication success; therefore, there is a risk that the authenticated smartphone might be used improperly by unauthorized individuals. Given the above background, we propose a new continuous authentication method for smartphones that uses biometric information considering the variation in usage environments and evaluated its effectiveness.**

I. Introduction

With the rapid spread of smartphones, opportunities to deal with information related to user privacy are increasing; in particular, there is a growing need to implement a user authentication function in smartphones. The conventional user authentication methods for smartphones are personal identification number (PIN), password, and pattern lock; they are popular because they are easy to use. However, they are not performed continuously after the first authentication success; therefore, there is a risk that the authenticated terminal will be used improperly by unauthorized individuals. One solution to this problem is to continue to verify the identity of the user after the initial authentication, which is commonly referred to as continuous authentication. However, it is not convenient for the user to perform the continuous authentication by using the conventional user authentication methods. As an alternative, continuous authentication using biometric information [1] obtained from the built-in sensors of the smartphones is attracting attention as a user authentication method that improves security while maintaining usability.

For example, Crouse *et al.* [2] proposed a continuous authentication method using input patterns on a touch screen. In their method, two types of text, that is predefined text and free text, are used for authentication. In particular, authentication using the free text can cope with arbitrary input patterns, which enables continuous authentication. Another example, proposed by Alshehri *et al.* [3], is continuous authentication method based on face matching. In this method, continuous authentication is executed by integrating different types of data acquired from an accelerometer, a gyroscope, and a magnetic sensor as well as a camera sensor on the device in order to maintain the accuracy of continuous authentication.

However, since continuous authentication assumes that the user does not consciously attempt to authenticate him/herself, in other words, the authentication is done implicitly, the quality of the biometric data to be acquired may be low depending on the usage environment. This problem remains unsolved in most of the previous research. Therefore, in this research, we deal with the problem by devising a new continuous authentication method for smartphones that considers the user's usage environment and evaluating its effectiveness.

II. Continuous Authentication System Considering The Variation In Usage Environment

A. Authentication Algorithm

As a solution to the problem that authentication accuracy is affected by the user's usage environment, we apply the concept of context-awareness-based multi-factor authentication [5], which recognizes the usage environment and selects biometric data for authentication that are suitable for that environment. An overview of the proposed authentication system is shown in Fig. 1(a), and its details are in Fig. 1(b).

An overview of the proposed authentication system is shown in Fig. 1(a). The proposed method first verifies the user's identity by using conventional user authentication methods such as pattern lock, and when the user's identity is confirmed, the terminal is unlocked ("authentication success") and continuous authentication starts. Here, the user is continuously authenticated by referring to the acquired usage environment information and biometric information in such a way that the user is not aware of the authentication. On the other hand, when the user's identity cannot be confirmed ("authentication failure") at this time, the conventional user authentication is repeated. Moreover, when the user's identity can be confirmed during the continuous authentication ("authentication success"), the continuous authentication is continued. When the user's identity cannot be confirmed ("authentication failure") at this time, the system locks the device and returns to the conventional user authentication.

The details of the above continuous authentication are shown in Fig. 1(b). In the continuous authentication phase, the following usage environment information and biometric information are acquired for a certain period of time.

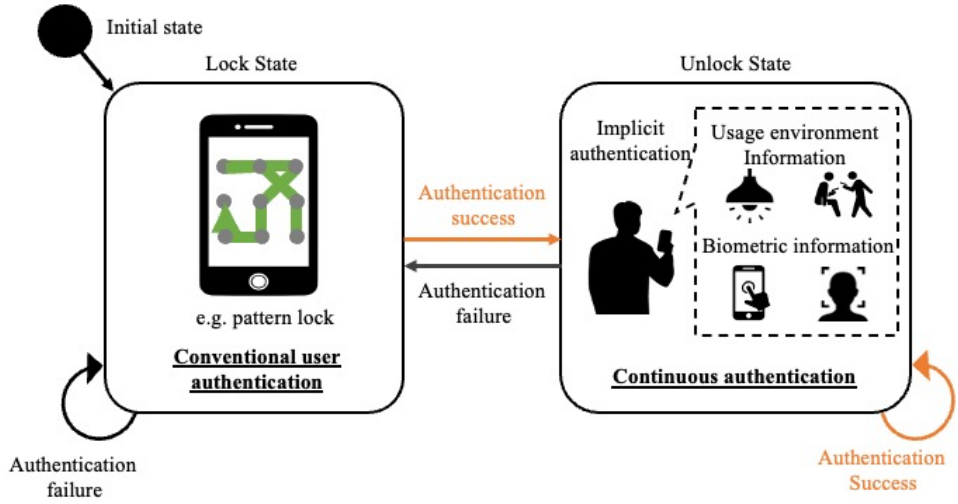

Fig. 1(a): Overview of the proposed authentication system

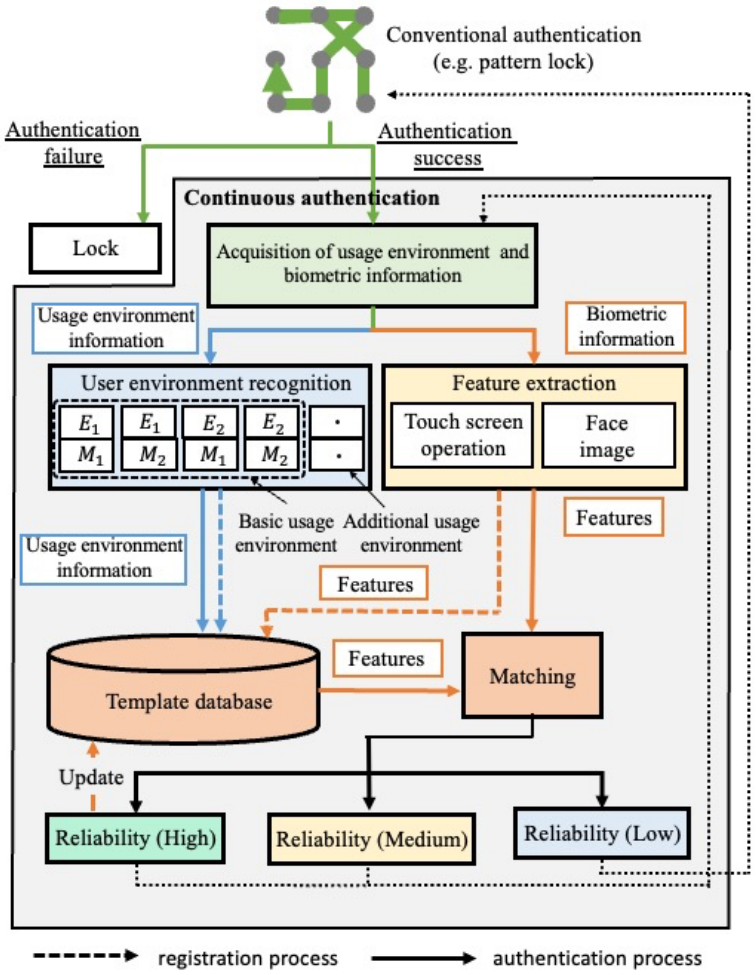

Fig. 1(b): Continuous authentication using the usage environment recognition

- Usage environment information:
  Gravitational acceleration (g_x, g_y, g_z), angular acceleration (ω_x, ω_y, ω_z), position information, brightness, and present time
- Biometric information:
  Contact point (x, y) and contact area of the touch screen, and face image

After acquiring the above information, the usage environment is recognized and features are extracted.

For usage environment recognition, we define the lighting conditions E1 and E2 around the terminal, and the user's behavior conditions M1 and M2 in advance. Based on these conditions, the usage environment is classified into four types: (E1, M1), (E1, M2), (E2, M1), and (E2, M2), which serve as the basic usage environments. On the other hand, in the feature extraction, features are extracted from the acquired biometric information. The proposed method uses face images and touch screen operations as biometric information.

The proposed authentication algorithm can be divided into two processes: a registration process and an authentication process.

In the registration process, the usage environment obtained from the usage environment recognition and the features obtained from the feature extraction are paired and registered as a template in the template database for each usage environment. As described later, the template is updated as needed when the identity is determined to be sufficient.

In the authentication process, the reliability of the user's identity is evaluated by matching the extracted features and the features in the template database corresponding to the obtained usage environment.

As shown in Table I, when the reliability level is high, the template is updated with the features used for matching, and the acquisition of the usage environment and biometric information is repeated. When the reliability level is medium, the template is not updated and the acquisition of usage environment and biometric information is repeated. When the reliability level is low, the template is not updated, the continuous authentication is aborted and the initial process of user authentication such as using pattern lock is resumed. In this way, depending on the reliability of the user's identity, the system decides whether to update the template or not and restricts the number of authentication attempts, thereby improving convenience and security.

TABLE I
RELIABILITY LEVEL

Reliability level	Description
High	User's identity is sufficiently reliable, and no additional authentication is required.
Medium	User's identity is reliable, however, additional authentication is required when a secure transaction is performed (e.g. making payments on the Web).
Low	User's identity is not reliable, and additional authentication is required.

B. Basic Usage Environment

Preliminary experiments were conducted to define the basic usage environment described in Section *A*. with the usage environment shown in Table II.

TABLEII
ENVIRONMENT FOR PRELIMINARY EXPERIMENTS

	Location	Time	Behavior
A	Indoor, with lighting	Daytime	Sitting on a chair
B	Outdoor, shade	Daytime	Walking
C	Outdoor, sunny	Daytime	Standing
D	Indoor, no lighting	Evening	Lying down
E	Indoor, with lighting	Evening	Sitting on a chair
F	Indoor, with lighting	Evening	Going down the stairs
G	Outdoor	Night	Standing
H	Indoor, with lighting	Night	Going up the stairs
I	Indoor, with lighting	Night	Sitting on a chair

(a) Brightness state

Regarding the brightness state that can be obtained from a brightness sensor, we defined a normalized value of 0.8 or higher as "light," 0.2 or higher but less than 0.8 as "medium," and less than 0.2 as "dark." This classification is based on the brightness measurements obtained for each usage environment (see Fig. 2). In Fig. 2, the vertical axis represents normalized brightness, and the horizontal axis represents time. From this figure, (C) can be classified as "light," (A, B, E, F, H, I) as "medium," and (D, G) as "dark". The Adaboost-based cascade discriminator [6] was used to detect faces in face images captured in multiple usage environments with different brightness states. The face detection worked on face images captured in usage environments classified as a "light" or "medium" brightness state. However, face detection was not possible for face images captured in a "dark" brightness state. Therefore, on the basis of these results, two brightness states, "light" (re-defined as a combination of "light" and "medium") and "dark", were used as the basic usage environments.

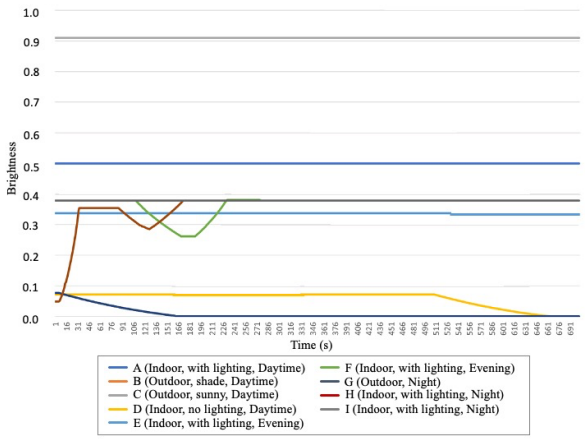

Fig. 2: Brightness in each usage environment

(b) User's behavioral state

We calculated the mean, variance, maximum, minimum, and frequency from time series data of gravitational acceleration and angular acceleration measured by the accelerometer and classified the usage environment by using random forests (RF). The classification accuracies were compared using a heat map that shows the percentage of each usage environment classified. This heat map is shown in Fig. 3. In this figure, the usage environment when stationary (A, C, D, E, G, I) was mostly not classified into the usage environment when moving (B, F, H). Therefore, we found that it is possible to distinguish between "stationary" and "moving".

We classified the experimental usage environments into two categories: stationary (A, C, D, E, G, I) and moving (B, F, H); the heat map for these environments is shown in Fig. 4. From these results, the usage environment can be classified fairly accurately between stationary and moving. Accordingly, we defined two types of usage environments, "stationary" and "moving," as the basic usage environments for the user's behavioral state.

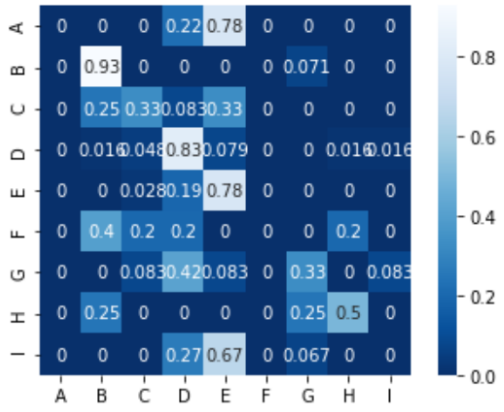

Fig. 3: Classification of user's behavioral states (A, B, C, D, E, F, G, H, I)

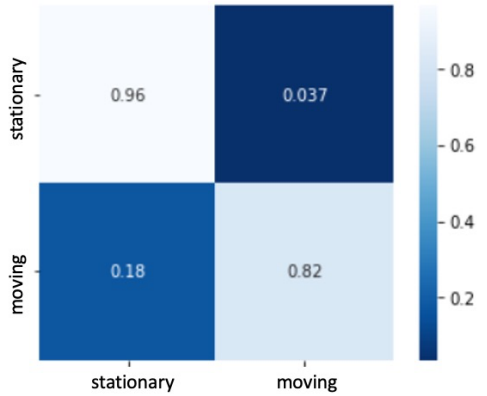

Fig. 4: Classification of user's behavioral states (Stationary and moving)

C. Features

The proposed method uses face images and touch screen operations as modalities of biometric information. For the face images, we used Dlib [7], which is a machine learning library for face detection and recognition by using feature points of the face. On the other hand, for the touch screen operation, we referred to the literature [8] and extracted the features shown in Table III from the terminal.

TABLE III
FEATURE OF TOUCH SCREEN OPERATION

	Feature	Description
1	X displacement	Displacement of X coordinate from the moment the touch screen is touched until the touch stops.
2	Y displacement	Displacement of Y coordinate from the moment the touch screen is touched until the touch stops.
3	Swipe distance	Length of trajectory from the moment the touch screen is touched until the touch stops.
4	Curvature	Degree of curvature during swipe operation.
5	Swipe speed	Speed from the moment the touch screen is touched until the touch stops.
6	Swipe angle	Angle of swipe (deviation from the center)
7	Contact area	Average contact area from the moment the touch screen is touched until the touch stops.

III. Simulation Experiments

We evaluated the reliability of the proposed continuous authentication system in a simulation experiment using actual biometrics data. In this experiment, we asked four subjects to use a comic viewer application that can acquire usage environment information and biometric information for one week. We classified the usage environment information obtained by the application into four types: (light, moving), (light, stationary), (dark, moving), and (dark, stationary), and evaluated the reliability of continuous authentication by performing face recognition and touch-screen operation authentication using the biometric information obtained in the same usage environment. The specifications of the experiment are shown in Table IV. We assumed that the system is able to correctly recognize the usage environment, since the usage environment was manually classified in this experiment. All of the experiments in this paper were conducted under a protocol approved by the University of Kitakyushu with informed consent from the subjects.

A. Face Recognition

We authenticated a user by using his/her face images acquired while he/she used the comic viewer application. Since the face images acquired in the experiment were taken without the user being consciousness of it, not all of the images contained faces. Therefore, we performed face detection on the acquired face images by using Dlib and investigated the face detection rate in each usage environment. The results are shown in Table V, which indicates that the face detection rate is higher for face images taken in the light condition than in the dark condition. In addition, the face

detection rate is higher when the face images are taken in the stationary condition than in the moving condition.

Next, we performed face recognition using the detected face images. We used half of the detected images as test images and the other half as training images and compared them using Dlib. In the experiment, we used the equal error rate (EER) as the performance criteria. Table VI shows the resulting of the EERs for the case where all face images were compared with the same template without usage environment recognition (referred to "None" in the same table) and the case where each face image was compared with the template in the same usage environment.

The results show that the EERs are generally low, suggesting that the continuous authentication using face images is capable of detecting unauthorized use by others. In addition, compared with the case without usage environment recognition, the EER is lower when the authentication is performed for each usage environment other than the case of the (light, moving) environment, suggesting that changing the template according to the usage environment is effective. On the other hand, the reason for the high EER in the (light, moving) environment is thought to be that many blurred images were included in the detected face images. In the case. of the (dark, stationary) environment, which has the lowest EER, the face detection rate is low, but there are few blurred images in the detected face images. This suggests that the authentication accuracy may be improved if only good-quality face images are detected during the face detection process.

TABLE IV
SPECIFICATIONS OF THE EXPERIMENT

Devices used	iPhone X, iPhone 11, iPhone 11 Pro
Number of subjects	4
Basic usage environment	(light, moving), (light, stationary), (dark, moving), (dark, stationary)
Usage environment	・Gravitational acceleration of device (g_x, g_y, g_z) ・Angular acceleration of device (ω_x,ω_y,ω_z) ・Brightness
Biometric information	・Face image ・Time stamp・Contact area ・Contact position on the touch screen (x,y) ・Contact status (1:with contact, 0:w/o contact)

TABLE V
FACE DETECTION RATE

Usage environment	(light, moving)	(light, stationary)	(dark, moving)	(dark, stationary)
Face detection rate (%)	77.5	82.6	58.5	61.7

TABLE VI
FACE RECOGNITION ACCURACY

Usage environment	None	(light, moving)	(light, stationary)	(dark, moving)	(dark, stationary)
EER(%)	13.2	21.2	5.1	8.3	0.0

B. *Touch-screen Operation Authentication*

In the comic viewer application used in this experiment, swiping to the right turns the page forward, and swiping to the left turns the page backward. We used the features extracted from the right and left swipe operations (Table VII) for authentication using the touch-screen operation. Half of the swipe operation data acquired for each usage environment were used as test data and the other half as training data, and comparisons were made using RF and Support Vector Machine (SVM). Table VII shows the resulting EER for the case where all the swipe operation data are compared with the same template without usage environment recognition and the case where each set of the swipe operation data is compared with the template in the same usage environment.

From Table VII, it can be seen that EER is generally high when SVM is used and that EER is low when RF is used. Therefore, in the continuous authentication using touch-screen operation, it is considered that RF is suitable for detecting unauthorized use by others. Since the EER is lower when the user is stationary than in the case without usage environment recognition (referred to as "None" in the same table), touch- screen authentication by recognizing the usage environment when the user is stationary is considered to be effective. On the other hand, the EER is higher when the user is moving than when the user is stationary, and it is considered that touch-screen operation authentication while moving is not effective. Therefore, it is necessary to take measures not to perform touch-screen authentication when the user is moving.

In addition, the EER of the touch screen operation authentication was generally higher than that of the pattern lock with biometric information [4], which is one of the authentication methods using the same touch screen. This may be due to the fact that the writing trajectory of the touch-screen operation is shorter than that of the pattern lock. Therefore, there is a possibility that the authentication accuracy can be improved by obtaining the features from data of multiple touch-screen operations combined, instead of the features from a single touch-screen operation.

TABLE VII
TOUCH-SCREEN OPERATION AUTHENTICATION ACCURACY

Usage environment			None	(light, moving)	(light, stationary)	(dark, moving)	(dark, stationary)
EER (%)	Right swipe	SVM	31.9	32.4	31.5	37.4	28.0
		RF	10.25	15.0	6.6	24.9	0.0
	Left swipe	SVM	32.2	38.9	14.0	40.7	32.7
		RF	8.0	24.3	0.0	29.9	9.3

IV. Conclusions

We proposed a new continuous authentication system for smartphones that takes into account the user's usage environment, and evaluated its effectiveness in simulation experiments. As future work, it will be necessary to evaluate the security and convenience of the proposed method when it is implemented on an actual device in an experiment with a larger number of subjects, and also evaluate the amount of resources consumed by the authentication process.

Acknowledgment

We would like to express our deepest gratitude to Mr. Daiki Izumoto, who was engaged in this research during his studies and helped us to compile this paper. Part of this work was supported by a JSPS Grant-in-Aid for Scientific Research, JP120K11814.

References

[1] P. A. Tresadern, C. McCool, N. Poh, P. Matejka, A. Hadid, C. Levy, T. F. Cootes, and S. Marcel, "Mobile Biometrics: Combined Face and Voice Verification for a Mobile Platform," IEEE Pervasive Computing, 12, 1, pp. 79-87, 2013.

[2] D. Crouse, H. Han, D. Chandra, B. Barbello, and A. K. Jain, "Continuous Authentication of Mobile User: Fusion of Face Image and Inertial Measurement Unit Data," Proc. of 2015 International Conference on Biometrics (ICB), pp. 135-142, 2015.

[3] A. Alshehri, F. Coenen, and D. Bollegala, "Iterative Keystroke Continuous Authentication: A Time Series Based Approach," KI - Künstliche Intelligenz, 32, 4, pp. 231-243, 2018.

[4] D. Izumoto and Y. Yamazaki, "A Study on Effective Features for Pattern Lock Authentication with Biometric Information," IEICE D, J103-D, 10, pp. 698-699, 2020.

[5] Y. Matsubara, H. Nishimura, and T. Samura, "A New Biometrics Technique with Flick Operation on Electronic Device," IEICE technical report, BioX2015-39, pp. 91-96, 2015.

[6] R. Okabe, T. Higashi, and Y. Yamazaki, "A Study on Biometric-bit-string Generation Method for Flick Input on Smart Phones," IEICE technical report, 116, 263, pp. 73-76, 2016.

[7] D. E. King, "Dlib-ml: A Machine Learning Toolkit," The Journal of Machine Learning Research, 10, pp. 1755-1758, 2009.