Mobile phone user authentication with grip gestures using pressure sensors

Murao, Kazuya ; Tobise, Hayami ; Terada, Tsutomu ; Iso, Toshiki ; Tsukamoto, Masahiko ; Horikoshi, Tsutomu

(Citation)
International Journal of Pervasive Computing and Communications,11(3):288-301

(Issue Date)
2015

(Resource Type)
journal article

(Version)
Accepted Manuscript

Mobile Phone User Authentication with Grip Gestures using Pressure Sensors

Kazuya Murao
College of Information Science and Engineering, Ritsumeikan University
Shiga, Japan
murao@cs.ritsumei.ac.jp

Hayami Tobise
Graduate School of Engineering, Kobe University
Hyogo, Japan
tobise@stu.kobe-u.ac.jp

Tsutomu Terada
Graduate School of Engineering, Kobe University
Hyogo, Japan
tsutomu@eedept.kobe-u.ac.jp

Toshiki Iso
Research Laboratories, NTT DOCOMO
Kanagawa, Japan
isot@nttdocomo.co.jp

Masahiko Tsukamoto
Graduate School of Engineering, Kobe University
Hyogo, Japan
tuka@kobe-u.ac.jp

Tsutomu Horikoshi
Research Laboratories, NTT DOCOMO
Kanagawa, Japan
hirikoshi@nttdocomo.co.jp

ABSTRACT
User authentication is generally used to protect personal information such as phone numbers, photos, and account information stored in a mobile device by limiting the user to a specific person, e.g. the owner of the device. Authentication methods with password, PIN, face recognition, and fingerprint identification have widely been used, however these methods have problems of difficulty in one-handed operation, vulnerability to shoulder hacking, and illegal access using fingerprint with super glue or facial portrait. From viewpoints of usability and safety, strong and uncomplicated method is required. In this paper we propose a user authentication method based on grip gestures using pressure sensors mounted on the lateral and back of a mobile phone. Grip gesture is an operation of grasping a mobile phone, which is assumed to be done instead of conventional unlock procedure. Grip gesture can be performed with one hand. Moreover, it is hard to imitate grip gestures since finger movements and grip force during a grip gesture are hardly seen by the others. We experimentally investigated the feature values of grip force and evaluated our proposed method from viewpoint of error rate. From the result, our method achieved 0.02 of EER (equal error rate), which is equivalent to face recognition.

Article Classification
Research paper

Keywords
User authentication, Grip gesture, Pressure sensor.

Professional Biography
Kazuya Murao: Kazuya Murao is an Assistant Professor at College of Information Science and Engineering, Ritsumeikan University, Japan. He received B.Eng, M.Info.Sci., and Ph.D. degrees from Osaka University in 2006, 2008, and 2010, respectively. Prof. Murao is investigating the wearable computing, ubiquitous computing, and context aware systems. He is a member of IEEE and IPSJ.

Hayami Tobise: Hayami Tobise is in SONY Corporation. He received B.Eng and M.Eng from Kobe University in 2012 and 2014, respectively. Mr. Tobise had been interested in music and mobile computing.

Tsutomu Terada: Tsutomu Terada is an Associate Professor at Graduate School of Engineering, Kobe University, Japan. He received B.Eng., M.Eng., and Ph.D. degrees from Osaka University in 1997, 1999, and 2003, respectively. Prof. Terada is working on wearable computing, ubiquitous computing, and entertainment computing. He is a member of IEEE, IPSJ, and IEICE.

Toshiki Iso: Toshiki Iso is a Researcher at NTT DOCOMO Research Laboratories, Japan. He received Ph.D. degree from Waseda University in 2008. Dr. Iso is working on image processing. He is a member of IEICE.

Masahiko Tsukamoto: Masahiko Tsukamoto is a Professor at Graduate School of Engineering, Kobe University, Japan. He received B.Eng., M.Eng., and Ph.D. degrees from Kyoto University in 1987, 1989, and 1994, respectively Prof. Tsukamoto is working on wearable computing and ubiquitous computing. He is a member of ACM, IEEE, and IPSJ.

Tsutomu Horikoshi: Tsutomu Horikoshi is a Senior Researcher at NTT Research Laboratories, Japan. He received Ph.D. degree from Keio University. Dr. Horikoshi is working on computer vision and pattern recognition. He is a member of IEICE.

1. INTRODUCTION
A lot of important information is stored in our smartphones and tablets, as mobile devices have been high-powered and sophisticated in these years, therefore the users have to protect it from malicious users. Currently, user authentication using password, personal identification number (PIN), or stroke pattern is generally used. Safety of a password is decided by the number of combinations of alphanumeric characters and symbols in the password. Simple passwords derived from personal information like date of birth and phone number are easily guessed by the others. In addition, password and stroke-pattern on touchscreen are easy to leak by shoulder hacking. Though setting a long and complex password consolidates the authentication, users hardly remember and type it. In most of OS, screen is locked for several tens of seconds if the password is incorrectly typed several times, taking a long time to be authenticated with secure password.

Biometrics are also used for authentication. Biometric identifiers are often categorized as physiological and behavioral characteristics; physiological characters are derived from the body

such as iris, voiceprint, and fingerprint. Behavioral characteristics are derived from the pattern of behavior such as gait, penmanship, and keyboard stroke. Users do not have to carry nor remember biometric identifiers. Since biometric identifiers are unique to individuals, they are more reliable in verifying identity than password. Physiological characteristics, however, involve the risks of aged deterioration and replication. Users hardly change their physiological characteristics like a password once they are leaked. In addition, the collection of physiological characteristics raises privacy concerns about the ultimate use of this information. In contrast, behavioral characteristics are hardly imitated nor copied unless the user's behavioral model is known. Even if the model is leaked, users can change the behavior used for authentication like password. Behavioral characteristics are reliable in verifying identity than password and are robust against leakage than physiological characteristics. From viewpoint of usability, however, conventional behavioral characteristics have drawbacks. Penmanship authentication requires touch stroke on screen with one hand holding a device and the other hand stroking a screen. Authentication with gesture actions like drawing a circle in the air is not realistic done in a public space.

In this work, we focus on actions holding a device, which is naturally performed during use of a device, and propose a user authentication method based on pressure distribution of grip gesture using pressure sensors mounted on the both lateral and back of a mobile phone. Grip gesture is one of the behavioral characteristics and is performed naturally in a series of operations using a mobile phone. It does not require any complicated operations. For example, just after a mobile phone is taken out of a pocket or bag, the user can unlock the device by gripping the phone instead of typing a password, therefore the user can use the device smoothly. In addition, it is difficult for the others to peep the grip gesture since grip force changes on a large scale without moving fingers clearly. In this paper, we constructed a user authentication method with grip gesture and evaluated it from viewpoints of error rate.

The remainder of this paper is organized as follows: Section 2 introduces related work, section 3 explains the proposed system, and section 4 evaluates the system and discusses the results. Finally, section 5 concludes this work.

2. RELATED WORK

As an example of user authentication using behavioral characteristics, Ohta et al. proposed a method that recognizes a movement of the mobile phone with an accelerometer in a mobile phone [1]. This method achieves low false acceptance rate (FAR) and false rejection rate (FRR), however, they conducted an evaluation in the environment where the user is stationary at a stable and vibration-free location, therefore it is not clear that this method works in the train or while walking.

Systems that recognize grip patterns of mobile phone with capacitive touchscreen have been proposed [2, 3, 4, 5]. In particular, the system in [2] recognizes grip patterns of a mobile phone such as taking pictures with both hands and typing a mail in one hand, and automatically runs corresponding applications to the grip patterns. This system, however, does not identify individuals, therefore, it cannot be used for user authentication.

Iso et al. proposed a user authentication method for mobile devices based on grip force during using a device [6]. The grip force is captured during natural use without user's intended input. However, this method captures grip force for few tens of seconds while using a phone, then authenticates the user, taking a long time to unlock the device.

In this paper we construct a user authentication system for mobile phones that performs accurate identification and smooth input by specific grip gestures users can freely select.

3. PROPOSED METHOD

This section explains the proposed user authentication system using pressure distribution of grip gestures obtained from pressure sensor arrays.

3.1 Sensor Hardware

This work uses a mobile phone with pressure sensor arrays attached on the both lateral and back sides of the phone as shown in Figure 1. The size of one pressure sensor in the array is 2.5*2.5 [mm] and Figure 2 shows the snapshot of the sensors. The number of sensors is 226 in total, all of which are rubber-covered. The measurement range is 0 to 100 [kPa], resolution is 0.024 [kPa], and the sampling rate is 100 [Hz]. The pressure values of each sensor are stored in the storage of the mobile device due to hardware limitation, and the authentication process is conducted on PC after exporting the data.

Figure 1. Experimental mobile phone with pressure sensors.

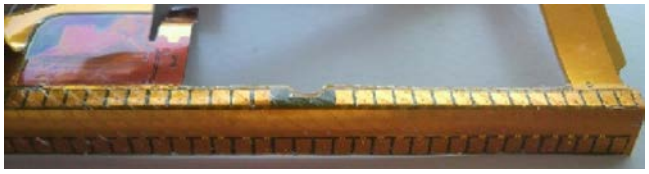

Figure 2. Pressure sensor array.

3.2 Authentication Process

Figure 3 shows the structure of the proposed system. The system extracts feature values from raw pressure data obtained from the sensors. Then, the system compares the extracted feature values with training data which are registered in advance. Finally, the system unlocks the device if both feature values of input data and training data are closer than predefined threshold, otherwise requests retry. The procedure consists of the following five phases:

1. Acquisition of pressure data
2. Calculation of time-difference data
3. Gesture spotting
4. Feature extraction and distance calculation
5. Authentication decision.

A grip gesture for authentication is conducted at the same timing as password authentication, therefore we assume a grip gesture is performed with one hand and finishes in couple of seconds.

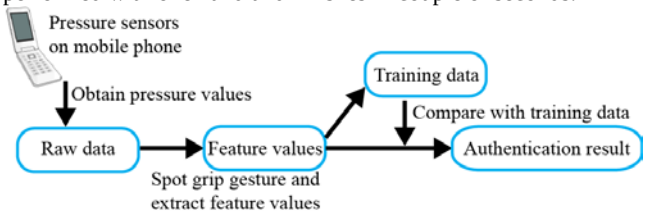

Figure 3. System structure.

3.2.1 Acquisition of pressure data

Pressure data during grip gesture is captured with the pressure sensor array mounted on the device. Figure 4 shows gray-scale pressure values when a user grips the device with four fingers in the order from the index finger to the little finger. The vertical axis shows time and the horizontal axis shows index of the sensors. White cell indicates large pressure. 226-dimensional data is obtained per one sampling.

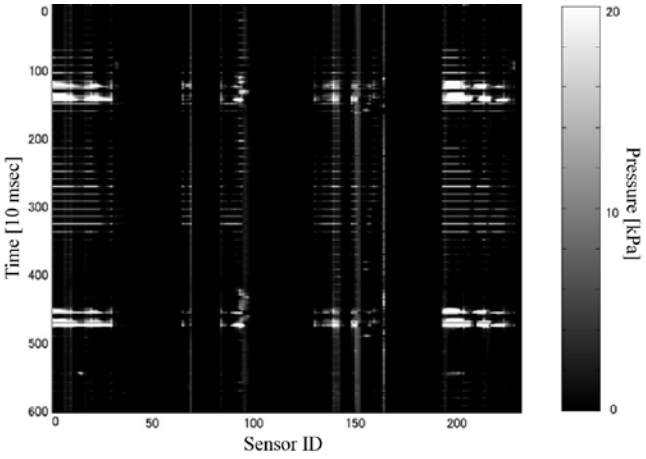

Figure 4. Pressure sensor data.

3.2.2 Time-difference data calculation

Pressure values are captured through the pressure sensor arrays mounted on the device. Pressure sensors used in our study are soft and easily deformative, causing unstable output values for grip gestures and non-zero values when the device is not held. In order to mitigate the distortion, our system uses time-difference values instead of raw pressure values. Let the raw pressure values at time t be $\boldsymbol{f}(t)=(f_1(t), f_2(t),\ldots, f_N(t))^T$, where N is the number of pressure sensors, i.e. N=226 in this work, $f_i(t)$ $(1<i<N)$ is the pressure value of the ith sensor, and $(\bullet)^T$ is the transposed matrix of $(\bullet)$. Then, time-difference value

$$\boldsymbol{g}(t)=(g_1(t), g_2(t),\ldots, g_N(t))^T=\boldsymbol{f}(t)-\boldsymbol{f}(t-1)$$

is obtained, where $g_i(t)$ is the time difference value of the ith sensor at time t.

3.2.3 Gesture spotting

Pressure data while the user is actually doing a grip gesture have to be extracted from stream of pressure data. The proposed method judges that a grip gesture has started when the number of sensors whose value is larger than the threshold exceeds the threshold. After the predetermined interval has elapsed from the starting point, our system judges that the gesture has finished when the number of sensors whose value is larger than the threshold becomes less than the threshold.

Detailed algorithm is as follows. The system counts the number of sensors N' out of N=226 that satisfy $|g_i(t)|>g_{Th}$ $(1<i<N)$, where g_{Th} is a predefined threshold for the time-difference data set to 10 [kPa] in this work from our preliminary experiment. Then, N' is compared with the threshold number N_{Th}. If $N'<N_{Th}$ is met, the system starts recording $\boldsymbol{g}(t)$ and defines time $t_a=t$ as a starting time of a gesture. After the starting point is found, the system finish recording data when $N'<N_{Th}$ is consecutively satisfied for B_{Th} samples and defines time $t_b=t$ as an ending time of the gesture.

If the length of the gesture t_b-t_a is longer than A_{Th}, the recorded data is extracted and used for authentication, otherwise the data is discarded since it is too short, where A_{Th} and B_{Th} are the threshold values set in advance. A_{Th} is set in order not to extract short gesture data. Length of the data becomes not shorter than A_{Th}. In addition, B_{Th} is set in order not to finish extraction when grip force temporarily becomes small. In this paper, we set A_{Th}=10 and B_{Th}=5, which correspond to 1 second and 0.5 second, respectively, and N_{Th} is set to 10. These values are decided through our preliminary experiment. The extracted data becomes $\boldsymbol{G}(t)=(\boldsymbol{g}(t_a),\ldots, \boldsymbol{g}(t_b))$.

3.2.4 Feature extraction and distance calculation

Our system extracts feature values for calculating distances between training data and input data. There are three elements for grip gestures: grip position, grip timing, and grip force. In this paper, we employed the four combinations of these elements:

A) Position + Timing + Force

B) Timing + Force

C) Position + Force

D) Force

Combinations that do not include "Force" feature such as "Position + Timing" are not employed since the pressure values are treated as binary and pressure sensor works just like touchscreen. The following part explains the feature calculation and distance calculation between input data and training.

3.2.4.1 A) Position + Timing + Force

This combination is the extracted data $\boldsymbol{G}$ as it is. Distance between the gestures whose grip position, grip timing, and grip force are similar becomes small. The feature vector used for distance calculation $\boldsymbol{Y}$ is defined as

$$\boldsymbol{Y} = (\boldsymbol{y}(1), \boldsymbol{y}(2),\ldots, \boldsymbol{y}(l)) \text{ and}$$

$$\boldsymbol{y}(i) = (g_1(i), g_2(i),\ldots, g_N(i))^T,$$

where l is the length of the gesture data. After the feature extraction, distance between two feature vectors of training data $\boldsymbol{Y}_{train}=(\boldsymbol{Y}_{train}(1),\ldots, \boldsymbol{Y}_{train}(l_{train}))$ and testing data $\boldsymbol{Y}_{test}=(\boldsymbol{Y}_{test}(1),\ldots, \boldsymbol{Y}_{test}(l_{test}))$ is calculated as

$$d(Y_{train}, Y_{test}) = \sum_{i=1}^{l} \left\| Y_{train}(i) - Y_{test}(i) \right\|,$$

where $l=min(l_{train}, l_{test})$ and ||•|| is the Euclidean norm of vector $(\bullet)$.

3.2.4.2 B) Timing + Force

This combination is a *l*-length vector consists of histograms of pressure values of the *N* sensors. Distance between the gestures whose grip force and grip timing are similar becomes small. The feature vector used for distance calculation $\boldsymbol{Y}$ is defined as

$$\boldsymbol{Y} = (\boldsymbol{y}(1), \boldsymbol{y}(2), \ldots, \boldsymbol{y}(l)) \text{ and}$$

$$\boldsymbol{Y}(i) = (y_1(i), y_2(i), \ldots, y_K(i))^T,$$

where *l* is the length of the gesture data. $\boldsymbol{Y}(t)$ is a histogram of pressure values at time *t* whose range is *R* and the number of classes is *K*. $y_j(i)$ is the frequency of class *j* and the width of class is $H=R/k$.

Let the feature vectors of training data be $\boldsymbol{Y}_{train}=(\boldsymbol{Y}_{train}(1),\ldots, \boldsymbol{Y}_{train}(l_{train}))$ and of test data be $\boldsymbol{Y}_{test}=(\boldsymbol{Y}_{test}(1),\ldots,\boldsymbol{Y}_{test}(l_{test}))$. The distance between $\boldsymbol{Y}_{train}$ and $\boldsymbol{Y}_{test}$ is calculated as

$$d(Y_{train}, Y_{test}) = \sum_{i=1}^{l} \left\| Y_{train}(i) - Y_{test}(i) \right\|.$$

3.2.4.3 C) Position + Force

This combination is a *N*-length vector consists of histograms of *l* samples of pressure values. Distance between the gestures whose grip force and grip position are similar becomes small. The feature vector used for distance calculation $\boldsymbol{Y}$ is defined as

$$\boldsymbol{Y} = (\boldsymbol{y}(1), \boldsymbol{y}(2), \ldots, \boldsymbol{y}(N)) \text{ and}$$

$$\boldsymbol{y}(i) = (y_1(i), y_2(i), \ldots, y_K(i))^T,$$

where *N* is the number of sensors. $\boldsymbol{y}(i)$ is *i*th histogram of pressure values whose range is *R* and the number of classes is *K*. $y_j(i)$ is the frequency of class *j* and the width of class is $H=R/k$. Let the feature vectors of training data be $\boldsymbol{Y}_{train}=(\boldsymbol{Y}_{train}(1),\ldots, \boldsymbol{Y}_{train}(N))$ and of test data be $\boldsymbol{Y}_{test}=(\boldsymbol{Y}_{test}(1),\ldots, \boldsymbol{Y}_{test}(N))$. The distance between $\boldsymbol{Y}_{train}$ and $\boldsymbol{Y}_{test}$ is calculated as

$$d(Y_{train}, Y_{test}) = \sum_{i=1}^{l} \left\| Y_{train}(i) - Y_{test}(i) \right\|.$$

3.2.4.4 D) Force

This combination is a scalar consists of sum of all the pressure values. Distance between the gestures whose grip force in total are similar becomes small. The feature vector used for distance calculation *Y* is defined as

$$Y = \sum_{i=1}^{l} \sum_{j=1}^{N} \left| g_i(j) \right|,$$

where $g_i(j)$ is the pressure value of *i*th sensor at the time *j*. Let the feature vectors of training data be Y_{train} and of test data be Y_{test}. The distance between Y_{train} and Y_{test} is calculated as

$$d(Y_{train}, Y_{test}) = \left| Y_{train} - Y_{test} \right|.$$

3.2.5 Authentication decision

The system calculates the distance d for all the training data and chooses the smallest one d_{min} of all, then compares the distance with predefined threshold value d_{Th}. If $d_{min} < d_{Th}$ is satisfied, the system unlocks the device, otherwise rejects the authentication and request retry.

4. EVALUATION

In order to evaluate the performance of the proposed user authentication method, we conducted experiments in three cases where grip gestures for authentication are chosen from a predefined list and the grip gestures are known by the others (case 1), where grip gestures for authentication are chosen from a predefined list and the grip gestures are guessed by the others who know the list by peeping the authentication scene (case 2), and where the grip gestures for authentication are freely set and the grip gestures are guessed by the others peeping the authentication scene (case 3). The case 3 is the most realistic condition and our proposed method performed well under the condition.

4.1 Case1: Grip gesture is chosen from list and known by the others

At first, we assume that the grip gestures for authentication is chosen from a list and the gestures are known by the others. This is the easiest case of the three for the malicious others who intend to pass the authentication.

4.1.1 Experimental procedure

Data on 10 kinds of grip gestures listed in Table 1 were captured 10 times for each gesture for 10 male subjects aged 20s. A total of 1,000 samples were obtained. These grip gestures differ on how to use four fingers to grip a mobile phone except for thumb. How to use the thumb was not specifically instructed. The flow of acquisition of data is described below. Firstly, the user picks up the mobile phone on the desk with his/her right hand. Then, the user performs a grip gesture, and puts the device on the desk. This flow is the same as the flow of the actual authentication operation. However, putting the device on the desk is performed for the repetition. Interval of each trial is approximately 10 seconds in order to mitigate the influence of the distortion of the sensors. The acquired data were exported to PC, then the system delimits the actual grip gestures in the data and extracts four kinds of feature values. Thereafter, the system conducts the authentication process, and we examined false acceptance rate (FAR), false rejection rate (FRR), and equal error rate (EER) by changing the threshold of distance d_{Th}. FAR is an error rate that third party incorrectly authenticated by the system. FRR is an error rate that the owner of the device is incorrectly unauthorized by the system. EER is an error rate at the point where FAR and FRR are equal. Increasing the threshold value d_{Th}, FAR increases and FRR drops, and EER takes a minimum value at a certain point. Generally, the performance of authentication is assessed by EER. Then, we investigated the performance of the each feature and gesture from view point of these indices. For the calculation of FRR, one sample of a gesture is given as a training data and the other nine samples are tested. This procedure is conducted 10 times by changing the training data. For the calculation of FAR, one sample of a gesture is given as a training data and 90 samples of the others are tested. This procedure is conducted 10 times by changing the training data. The above procedures were conducted for the 10 kinds of gestures.

Table 1. Grip gestures.

ID	Gesture
1	Grip with four fingers from the little finger to the index finger.
2	Grip three times with four fingers from the little to the index finger.
3	Grip with four fingers in the order from the index finger to the little finger. Relieve a finger then grasp with the next finger.
4	Grip with four fingers in the order from the index finger to the little finger. Fingers remain grasping while the gesture.
5	Grip with four fingers in the order from the little finger to the index finger. Relieve a finger then grasp with the next finger.
6	Grip with four fingers in the order from the little finger to the index finger. Fingers remain grasping while the gesture.
7	Grip with the index finger and the middle finger for a second.
8	Grip with the ring finger and the little finger for a second.
9	Grip with the index finger and the little finger for a second.
10	Grip with the middle finger and the ring finger for a second.

4.1.2 Results and consideration

Table 2 shows EER for each gesture and feature combination. With respect to feature value, "Position + Force" shows lowest EER of the four feature values. The reason adding "Timing" feature increased the EER is that the reproducibility of timing is low. Focusing on grip gesture, EER of gesture #3 and #8 are high, therefore it can be said that these gestures are not appropriate for authentication. Further investigations on why EER increased, and the tendency of gestures which are not suitable for authentication are our future work. In contrast, gestures #2, #4, #5, #6, #7, and #9 showed relatively low EER that are 0.19 to 0.22. However, even these values of EER mean that the owner of the device is rejected and others accepted once five trials. It is considered that these gestures are not practicable for authentication.

For these results, it is considered that the cause of unpractical EER is that evaluation experiments were conducted on the assumption that gestures for authentication had been known to the others. As a reference, FAR-FRR curves of gestures #6 and #8 with "Position + Force" feature are shown in Figure 5 and Figure 6 as examples of a gesture whose EER is low and high, respectively. Seeing from the figures, FRR-FAR curves of gesture #8 are overlapping closely, resulting in high EER.

Table 2. EER when grip gesture is known.

Feature	Gesture									
	1	2	3	4	5	6	7	8	9	10
P+T+F	0.34	0.36	0.45	0.42	0.42	0.42	0.34	0.41	0.39	0.38
T+F	0.32	0.28	0.40	0.36	0.29	0.29	0.28	0.40	0.34	0.28
P+F	0.26	0.22	0.38	0.20	0.20	0.20	0.19	0.44	0.20	0.26
F	0.32	0.32	0.43	0.32	0.36	0.36	0.31	0.44	0.60	0.24

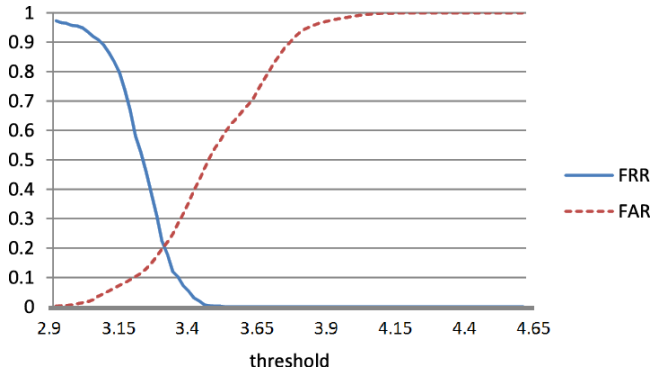

Figure 5. FRR-FAR curve of gesture #5 with "Position + Force" feature (case 1, EER=0.20).

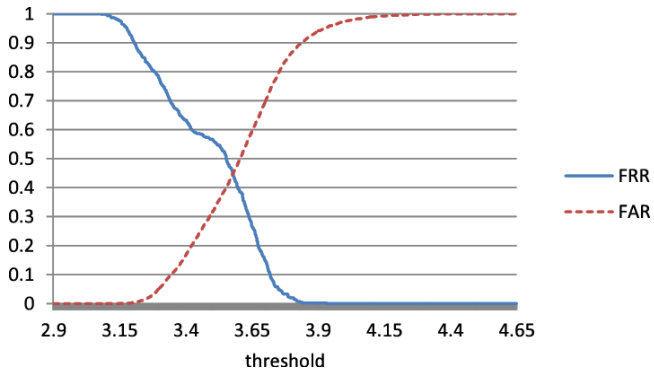

Figure 6. FRR-FAR curve of gesture #8 with "Position + Force" feature (case 1, EER=0.44).

4.2 Case 2: Grip gesture is chosen from list and peeped by the others

We assume that the grip gestures for authentication is chosen from a list same as case 1 but the gestures are guessed by peeping by the others. This is more realistic case than case 1.

4.2.1 Experimental procedure

At first, we investigated the accuracy of shoulder hacking the grip gestures. The subjects were the ten people same as case 1. One subject performed ten kinds of grip gestures in a random order as though authenticating. Then, nine other subjects peeped the gestures and answered the grip gesture they guessed. This procedure is repeated for all the grip gestures in a round-robin fashion, and concealing rate for gesture g is calculated with

$$ConcealingRate(g) = \frac{C(all) - C(correct)}{C(all)},$$

where *C(all)* is the total number of answers for each grip gesture and *C(correct)* is the number of correct answers. Table 3 shows the concealing rate of each grip gesture. Except for the gesture #2, concealing rates are approximately 0.50, which means that the grip gesture for authentication is shoulder hacked by one out of two people. The reason for the low concealing rate would be that the authentication operation can easily be seen. In the actual environment, concealing rate is expected to become higher because the actual authentication operation is less visible than the experiment. Reason the concealing rate of gesture #2 was low would be that the movement of the finger was bigger than the other gestures.

We evaluated the performance of the proposed method when the gesture is guessed by peeping by using the concealing rate. In the previous evaluation, the surrounding people knew the gesture for authentication and same gesture data were used as testing. This evaluation used the gestures they answered as testing for FAR calculation. Then, we investigated performance of the each gesture same as in the previous experiment. Results and consideration.

Table 4 shows EER for each gesture and feature combination. EER for the gestures other than #2 was slightly improved since the concealing rate for gesture #2 was quite low as shown above. EER for "Timing + Force" was the highest of all same as that in case 2. Focusing on grip gesture, EER for gestures #5, #6, #7, #9 became 0.14. Comparing the results of "Position + Force" and "Force", position feature much improved the EER, which indicates that grip positions were significantly different over the subjects. However, this performance is still too bad to be used in practical as reported that EER of face recognition is 0.012 [7].

Table 3. Concealing rate of grip gestures.

Gesture	Concealing rate
1	0.45
2	0.14
3	0.48
4	0.42
5	0.44
6	0.46
7	0.46
8	0.52
9	0.52
10	0.52

Table 4. EER when the grip gesture is guessed by peeping.

Feature	Gesture									
	1	2	3	4	5	6	7	8	9	10
P+T+F	0.32	0.36	0.44	0.42	0.42	0.35	0.28	0.39	0.38	0.38
T+F	0.32	0.27	0.38	0.32	0.20	0.32	0.26	0.33	0.29	0.24
P+F	0.19	0.21	0.38	0.19	0.14	0.14	0.14	0.39	0.14	0.18
F	0.30	0.32	0.43	0.32	0.36	0.27	0.28	0.41	0.30	0.22

4.3 Case 3: Grip gesture is freely set and peeped by the others

In this evaluation, a test subject registers a grip gesture as he/she likes and the other subjects try to peep the grip gesture.

4.3.1 Experimental procedure

Five male subjects aged twenties who are different to the subjects in case 1 and 2 joined the experiment. Each subject defined one grip gesture for authentication as they like and registered it five times as a training data. Then, they had seated around a table. In the experiment, one of the subjects unlocks the device five times, and the other four subjects tried to peep the grip gesture. At last, they tried to unlock the device five times. This round is repeated five times by shifting the position and role of the subjects clockwise. The subjects were asked to register a grip gesture as they like, but to finish the gesture in couple of seconds. Flow of data collection is as follows. Firstly, the subject picks up the mobile phone on the table with the right hand. All the subjects are right-handed. Then, the subject performs a grip gesture, and puts the device on the table. The action of putting the device on the desk is just for the next trial. The subjects did not talk each other during the experiment and concentrated to steal the grip gesture. Training data and testing data are collected through this flow.

Table 5 shows the grip gestures that the subjects defined. Each subject defined a sequence of three to five grip actions. For example, subject 1 applied force to the device with the index finger, then released the index finger and applied force with the little finger, release the little finger and applied force with index finger, release the index finger and applied force with little finger. Lastly, the little finger is released and force is applied with middle finger and ring finger together. During the grip gesture, unused fingers touch the device, but did not apply force. We investigated the performance of the proposed system for the feature values and the gestures by changing the number of training data from one to five.

Table 5. Grip gestures set by the subjects.

Subject	Order of fingers used for grip gesture
1	Index finger->little finger->index finger->little finger->middle finger and ring finger
2	index finger->ring finger->middle finger->little finger
3	ring finger and little finger->index finger and middle finger->index finger and little finger
4	middle finger->ring finger->little finger->index finger
5	index finger->little finger->ring finger->middle finger

4.3.2 Results and consideration

Table 6 shows EER for each gesture and for the number of training data. Average EER for "Position + Force" feature achieved 0.04, which means that third party hardly unlocks the device. In addition, "Position + Force" feature shows the lowest average EER 0.02 when four training data were used. As a reference, [7] reported EER of face recognition is 0.012 without considering a risk of social engineering. Average EER for "Position + Timing + Force" and "Timing + Force" improved to 0.05 and 0.14 by increasing the number of training data. There features, however, show worse EER than that of "Position + Force" feature, which means "Timing" feature deteriorates the performance. From these results, users cannot reproduce the timing of grip gestures as mentioned in the results of case 1 and 2. Average EER for "Force" feature showed high EER and is not improved as the number of training data increases, which means that "Force" feature is not significant since information is compressed too much.

Focusing on the grip gestures, EER for Subject 3 and Subject 4 reach zero when the number of training data is one, and EER for Subject 2 and Subject 5 reach zero when the number of training data is four. On the other hand, EER for Subject 1 is 0.15 even when the number of training data is five. To investigate the reproducibility of the gestures, we calculated the distance among owner's data. Table 7 shows average and variance of the distances between two samples out of the owner's five samples.

The variance of Subject 1 is larger than that of the others. This would be caused by complex finger actions of grip gesture of Subject 1, which deteriorates the reproducibility, resulting in high

EER. Performance of our system would be improved and stabilized by rejecting gestures whose training data show large variance of distance at registration.

Table 6. EER for each gesture and for the number of training data.

# training data	Features	Ave.	1	2	3	4	5
1	P+T+F	0.17	0.36	0.20	0.04	0.00	0.24
	T+F	0.40	0.25	0.56	0.42	0.18	0.56
	P+F	0.04	0.18	0.01	0.00	0.00	0.03
	F	0.21	0.44	0.20	0.08	0.00	0.36
2	P+T+F	0.13	0.28	0.13	0.02	0.00	0.24
	T+F	0.25	0.26	0.36	0.28	0.14	0.22
	P+F	0.02	0.08	0.02	0.00	0.00	0.02
	F	0.21	0.41	0.20	0.07	0.00	0.38
3	P+T+F	0.10	0.15	0.12	0.03	0.00	0.16
	T+F	0.20	0.22	0.28	0.28	0.08	0.12
	P+F	0.02	0.09	0.00	0.00	0.00	0.00
	F	0.20	0.40	0.24	0.06	0.00	0.32
4	P+T+F	0.05	0.08	0.04	0.0	0.00	0.08
	T+F	0.16	0.16	0.20	0.35	0.04	0.04
	P+F	0.02	0.12	0.00	0.00	0.00	0.00
	F	0.22	0.46	0.32	0.04	0.00	0.30
5	P+T+F	0.05	0.10	0.05	0.00	0.00	0.10
	T+F	0.14	0.10	0.20	0.40	0.00	0.00
	P+F	0.03	0.15	0.00	0.00	0.00	0.00
	F	0.22	0.40	0.40	0.00	0.00	0.30

Table 7. Average and variance of the distance among owner's data.

Feature		Gesture				
		1	2	3	4	5
P+T+F	Ave.	16	3.7	2.9	0.84	3.7
	Var.	680	1.2	0.44	0.041	0.61
T+F	Ave.	0.53	0.61	0.53	0.45	0.58
	Var.	0.016	0.020	0.0045	0.0066	0.013
P+F	Ave.	3.4	3.25	3.26	3.39	3.30
	Var.	0.021	0.0016	0.0051	0.0030	0.0093
F	Ave.	0.19	0.0096	0.0067	0.0024	0.017
	Var.	0.13	<0.001	<0.001	<0.001	<0.001

5. CONCLUSION

We constructed the system that performs personal authentication based on pressure distribution of grip gesture using a pressure sensor mounted on the both laterals and back of a mobile phone. We investigated the characteristics of grip gestures from view point of accuracy. We have confirmed that the average EER was 0.02 and we the proposed method is effective when the gesture was freely set.

We plan to investigate more effective feature values for authentication, extraction method of gestures, calculation method of distance between the data, other types of gestures.

6. REFERENCES

[1] M. Ohta, E. Namikata, S. Ishihara, and T. Mizuno: Individual Authentication for Portable Devices using Motion Features, Proc. of the 1st International Conference on Mobile computing and Ubiquitous networking (ICMU 2004), pp. 100-105 (2004).

[2] K. Kim, W. Chang, S. Cho, J. Shim, H. Lee, J. Park, Y. Lee and S.Kim: Hand Grip Pattern Recognition for Mobile User Interfaces, Proc. of the 18th Conference on Innovative Applications of Artificial Intelligence (IAAI'06), vol. 2, pp. 1789-1794 (2006).

[3] W. Chang, K.E. Kim, and H. Lee: Recognition of Grip-Patterns by Using Capacitive Touch Sensors, Proc. of the IEEE International Symposium on Industrial Electronics (ISIE 2006), pp. 2936-2941 (2006).

[4] T. Feng, Z. Liu, K. Kwon, W. Shi, B. Carbunar, Y. Jiang, and N. Nguyen: Continuous Mobile Authentication using Touchscreen Gestures, Proc. of the IEEE Conference on Technologies for Homeland Security (HST 2012), pp. 451-456 (2012).

[5] B. Taylor, V. Michael Bove, Jr.: Graspables: Grasp-recognition as a User Interface, Proc. of the 27th ACM Annual Conference on Human Factors in Computing Systems (CHI 2009), pp. 917-926 (2009).

[6] T. Iso and T. Horikoshi: Statistical Approaches for Personal Feature Extraction from Pressure Array Sensors, Proc. of the 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2013), pp. 129-133 (2013).

[7] Q. Tao and R. Veldhuis: Biometric Authentication for a Mobile Personal Device, Proc. of the 3rd Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services (Mobiquitous 2006), pp. 1-3 (2006).