A 3-D Hand Gesture Signature Based Biometric Authentication System for Smartphones

Ziwen Sun
School of Internet of Things Engineering, Jiangnan University.
No. 1800 Lihu Avenue, Wuxi,Jiangsu,214122, CHINA
(86)13915355548
sunziwen@jiangnan.edu.cn

Yao Wang
School of Internet of Things Engineering, Jiangnan University.
No. 1800 Lihu Avenue, Wuxi,Jiangsu,214122, CHINA
(86)18352513475
6121905045@vip.jiangnan.edu.cn

Gang Qu
Department of Electrical & Computer Engineering and Institute for Systems Research.
University of Maryland, College Park
College Park,Maryland 20742,USA
(1)301-405-6703
gangqu@umd.edu

Zhiping Zhou
School of Internet of Things Engineering, Jiangnan University.
No. 1800 Lihu Avenue, Wuxi,Jiangsu,214122, CHINA (86)13861867502
zzp@jiangnan.edu.cn

ABSTRACT

Most of the smart phones are equipped with user authentication mechanism such as entering a 4-digit password or drawing a simple pattern. These system are easy to be hacked and more importantly, they authenticate the password or the pattern, not the real user. In this paper, we describe the design and implementation of a 3-D hand gesture signature (HGS) based biometric authentication system. We take advantage of the on-phone accelerometer to capture the 3-D acceleration information when user holds the phone and makes a gesture in order to gain access to the phone. This data will go through a sequence of signal processing, namely data smooth, gesture spotting, sequence alignment and interpolation. Then the processed data will be compared with the genuine user's registered pattern to determine whether access to the phone should be granted to the user.

We have implemented the proposed 3-D HGS authentication system on real smart phones of different brands and recruited volunteers to perform on-phone experiments to test the performance of the system. The authentication process is tested for a total of 76,520 times by 19 users. The results show very low false acceptance (0.28%) and false rejection rates (5.13%). The system is user friendly (the acceptance/rejection decision is made instantaneously) and does not require any addition hardware on the phone. We further export the real gesture samples from the phones to a desktop PC, where we implement two existing gesture based authentication systems similar to ours. The simulation results reveal that our system has the best authentication accuracy.

Categories and Subject Descriptors

K.4.4 [**Electronic Commerce**]: Electronic Commerce-*Security*.

General Terms

Measurement, Performance, Experimentation, Security.

Keywords

Authentication, Gesture recognition, Pattern matching.

BICT 2014, December 01-03, Boston, United States

DOI 10.4108/icst.bict.2014.257951

1. INTRODUCTION

As smartphones become a computing and communication platform for emerging applications such as online shopping, mobile banking, electronic voting, and a vehicle for social media, user authentication for smartphone usage is playing a more and more important role. Authentication is a process to determine whether a user or a device is what it is claimed to be, or to confirm its identity. The pair of username and password is the most common and popular authentication mechanism for computers and networks. However, its security weakness has been well-documented, particularly because of the fact that what it verifies is the username and password pair, not the user himself. As a comparison, biometric authentication methods identify a person by combining computer technology, biological sensors and biological principles[1][2][3]. These methods authenticate individuals based on their inherent physiological or behavioral characteristics. They offer the following advantages over the traditional non-biometrics based authentication mechanisms:

- Biometrics is inherently more reliable and more capable in differentiating between an authorized person and a fraudulent imposter than non-biometrics methods for verifying an owner's identity rather than simply confirming the user's knowledge or possession of a particular secret [2].
- Biometric identifiers are uniquely and permanently associated with their owners, hence they can efficiently prevent identity theft and unauthorized access to mobile terminal resources [4].
- The price of biometrics sensors continues to fall, several major providers have announced the availability of biometrics-enabled mobile devices [4][5].

Depending on the type of features being used to identify a person, biometrics techniques are usually categorized as either physiological (or static) or behavioral[6]. Physiological biometric techniques are based on biological measurements and inherent characteristics that a person owns such as face, iris, teeth, voice, fingerprint, hand geometry and heart rate [4][7][8][9]. Behavioral biometric techniques are related to something that a person is able to perform and repeat in a unique manner, such as handwriting signature, gait, and keystroke dynamics [5][7][10].

There are some drawbacks in using physiological biometrics for authentication on the mobile phone platform. First, it is not user-friendly because one has go through certain specific measurement procedure, some may be uncomfortable, to collect data such as face, teeth, voice, and signature [7]. Second, physiological biometrics can be vulnerable to criminal penetration such as spoof attacks and be suffer from replay attacks. Artificial

fingerprint, face photo, and voice recording can pass the authentication procedure with 80% or higher successful rate [11]. Finally, the hardware required for data collection and analysis are not available on current smartphones [10][12].

Behavioral biometrics can overcome most of these shortcomings[5]. The main reasons are: (*i*) Behavioral biometrics are user-friendliness duo to the fact that the behavior characters can be captured unobtrusively and continuously. (*ii*) Behavioral biometrics are security due to the fact that the behavior characters of an individual is difficult to mimic due to being related to specific actions and bring no digitally duplicate.(*iii*) Behavioral biometrics do not require hardware for inertial sensor increasingly being embedded in commercial smartphones.

In this paper, we propose a novel 3-D hand gesture signature (3-D HGS) based biometric authentication system to enhance the security of mobile phones. Our method exploits the micro sensors available on smartphones to capture acceleration signals when users perform hand gesture signatures. These signals are processed following a sequence of operations that include data smooth, gesture spotting, sequence alignment and interpolation. The absolute distances between gesture signature sequences are used as the measure to determine the genuine user.

We implement and test our 3-D HGS authentication system on mobile phones (from Huawei, Samsung, and Sony) and build a database of hand gestures from 19 volunteers. We design various simulation scenarios to validate our proposed method, including genuine user's attempt under different conditions and the spoofing attack where an impostor mimics genuine user's gesture. Simulation results show that our system has lower error rates than two recently proposed similar biometrics authentication methods, DTW (dynamic time warping) [13] and GSA (global sequence alignment) [14].

The remainder of the paper is structured as follows. In section 2, we survey the most related works on this subject and analyze their limitations. In section 3, we highlight the requirements and challenges in designing biometrics authentication system and the key concept of the proposed 3-D HGS system. In section 4, the core algorithm and mathematical foundations of the 3-D HGS authentication system are elaborated. In section 5, we describe the simulation setup and report the performance compared with other similar approaches. In section 6, we conclude the paper with discussion on our contributions in this paper and future work.

2. RELATED WOK

3-D hand gesture biometric authentication focus on identifying the person against impostor users using identifying gesture [15],[16] by drawing one's handwritten signature in the air using a phone incorporating an accelerometer [13]. In spite of different variations in acceleration signals among repeating identifying gestures, an intrinsic part of the gestures remaining invariant can be extracted as characters to identify individual [16].

Several different pattern recognition techniques could be used to cope with user authentication by means of identifying 3-D hand gesture. To compare performance among three different techniques, the signatures measured by three acceleration signals have been captured with a sampling rate of 100Hz using iPhone [13]. The simulation experiments conducted in Matlab on a computer showed Average Dynamic Time Warping was considered to be the best accuracy/computation cost relationship among three methods of Hidden Markov Models, Bayes classifiers and DTW. To eliminate the difference between repeating identifying gestures, the Longest Common Subsequence (LCS) were applied to find the invariant information in enrolled temple sequence and identified sequence[16]. The experiments, conducted under sampling the 3-D hand gesture at a rate of 50Hz. To correct slight differences between gestures, a strategy of score optimization and template updating was used. Two 300 points acceleration signals were generated with gesture signal lasting less than 6s and a sampling rate of 50 Hz[15].The 300 points were insured by a global sequence alignment algorithm and optional interpolation of aligned signals. Moreover, a template updating strategy was presented to get more stable performance [14].

Despite the above techniques fulfilling properties of minimality and circumvention[15],[17] , there are some issues having not been solved completely:

- **Performance:** The performances are very poor and unpractical. The average DTW algorithm obtained the best accuracy with an equal error rate (EER) of 2.12% [13]. The LCS algorithm obtained an optimal EER value of 3.58±0.78 (%) [16]. Hand gesture recognition obtained EER of 2.01% in a zero-effort and 4.82% in an active impostor attack with Mac Computer and iPhone[15].The score optimization and template updating method obtained the best performance with an False Acceptance Rate (FAR) 1.45%±1.83%, False Rejection Rate (FRR) 5.65%±3.37% [14].
- **Permanence:** Gestures do not long exactly the same due to any repetition being different from each other [16]. The repeatability of 3-D gesture signature maybe change over time and this kind of variability will lead to bad performance [13], the variability problem need to be solved by updating the template over time.

3. 3-D HGS: Requirements and Challenges

In this section, we first summarize the requirements for a biometric authentication systems, then briefly describe our proposed 3-D hand gesture signature based approach and discuss the advantages and challenges to implement the 3-D HGS system.

Characteristics of user behavioral biometrics should satisfy the following properties for user authentication:

- **Universality:** Each user should possess the characteristics to be identified. The vast majority of mobile phone owners should have the characteristic to be measured.
- **Uniqueness:** No two users should have identical characteristics. The identifying characteristic should be distinctive among different individuals.
- **Permanence:** The characteristics should be sufficiently invariant over a period of time with no significantly change, environment conditions or other variables.
- **Collectability:** The characteristics should be ready for collection and easy to quantify.
- **Performance:** The biometric authentication system built on the characteristics should have high performance in terms of accuracy, speed, robustness, resource requirements, and other operational or environmental factors.
- **Acceptability:** Users will be willing to accept the system, feeling safe, secure, and comfortable when the biometric characteristics are extracted.
- **Circumvention:** Successful forgery of the behavioral biometric characteristics should be hard, if not impossible.
- **Minimality:** Additional hardware and software required to collect the characteristics and perform the authentication should be kept at the minimal level.

In this paper, we propose the 3-D acceleration values of human hand gesture as the characteristics for a behavioral biometric authentication system. The proposed 3-D hand gesture signature (3-D HGS) based biometric authentication system will extract distinctive and measurable patterns of mobile phone users from the information captured by the 3-axis accelerometer embedded in the smartphone. It consists of two components: user signature enrollment phase and gesture verification phase.

In the signature enrollment phase, the genuine user's biometric gestural templates are registered when the user performs his 3-D gesture signature while holding the phone. In the gesture verification phase, the user who attempts to access the phone will make the gesture. The system will collect the input gesture and verify with the registered gestural template for authentication purpose. If a match is found, user will be authenticated. Otherwise, user has to repeat the gesture just like re-entering password in the username-password authentication protocol.

Clearly the proposed 3-D HGS authentication system satisfies the requirements of universality, collectability, acceptability, and minimality (accelerometer is available on all major brand smartphones). In the remaining of the paper, we will elaborate the proposed 3-D HGS biometric authentication system and provide simulation results based on real human subjects to evaluate the other requirements, namely *uniqueness, permanence, performance*, and *circumvention*. Here we give the two main error metrics for such evaluation: FAR and FRR.

FAR measures the likelihood that an impostor can pass the authentication system and gain access to the smartphone belonging to a genuine user. This metric will be used to evaluate *circumvention*, *uniqueness,* and *performance*. FRR, on the other hand, measures the likelihood when the genuine user's attempts to access the mobile phone is rejected. This metric will be used to evaluate *permanence* and *performance*. In the case of authentication for smartphones, we argue that it is the top priority to find the behavioral characteristics and build the corresponding authentication system that can achieve zero or extremely low FAR. This is because preventing the phone from being illegitimately accessed (a low FAR does) is more important than authenticating the genuine user at the first attempt (a low FRR does).

There are two major challenges to design the proposed 3-D HGS based authentication systems. First, how to extract information from user's 3-D hand gesture that can lead us to low FAR an FRR. Second, how to evaluate FAR and FRR to validate the system.

4. 3-D HGS AUTHENTICATION SYSTEM

This section elaborates the signature enrollment phase and the gesture verification phase in the proposed 3-D hand gesture signature based biometric authentication system.

4.1 Data Acquisition

The gesture signal taken by the powerful 3-axis accelerometer from a mobile device is represented by a vector of the current acceleration of the controller in all three dimensions [18]. A 3-D gesture can be described by a locus of hand motion recorded in a sequence of accelerometer signals. To precisely characterize user's 3-D gesture signature, we use the acceleration values along the three axis (x, y, z) measured by the accelerometer and denoted as (x_t, y_t, z_t) shown below:

$$A^k = \left\{ \left(x_1^k, y_1^k, z_1^k\right), \cdots, \left(x_t^k, y_t^k, z_t^k\right), \cdots, \left(x_n^k, y_n^k, z_n^k\right) \right\}. \quad (1)$$

where $k, t, n \in Z$, $1 \le t \le n$, k represents the k-th gesture repetition, t represents time points and n represents the end time point of the gesture duration. At least 300 points signature data are required for each axis for authentication.

4.2 Signal Preprocessing

Data Smoothing. The accelerometer is a very sensitive device that can capture the inevitable tiny shiver occurred when the user performs the gesture while holding the phone. The shiver can affect the quality of the 3-D gesture signature and can be modeled as environmental noise. We first use a smooth denoising technique, moving average (or moving mean) method, to eliminate the interference of shiver to reduce its influence. Formula (2) shows how the acceleration values after smoothing is obtained:

$$\left(sx_t^k, sy_t^k, sz_t^k\right) = \sum_{j=t-\Delta}^{t+\Delta} \left(x_j^k/(2\Delta+1), y_j^k/(2\Delta+1), z_j^k/(2\Delta+1)\right). \quad (2)$$

The parameter Δ is a key factor which determines the smoothing effect. If the value of Δ is too small, the data smoothing effect will not be significant; if the value of Δ is too large, it may cause data distortion. According to different conditions, the suggested value of Δ is generally between 2 and 7 from some references. Now we describe the triaxial acceleration sequence after smooth denoising processing:

$$SA^k = \left\{ \left(sx_1^k, sy_1^k, sz_1^k\right), \cdots, \left(sx_t^k, sy_t^k, sz_t^k\right), \cdots, \left(sx_n^k, sy_n^k, sz_n^k\right) \right\}. \quad (3)$$

Data Normalization. The purpose of data normalization processing is to make it easy to compare the enrolled 3-D identifying gesture and the 3-D verifying gesture. The average-standard deviation normalization method, shown in Formula (4), is adopted for this purpose:

$$\left(nx_t^k, ny_t^k, nz_t^k\right) = \left(\frac{sx_t^k - \overline{sx^k}}{\sigma x^k}, \frac{sy_t^k - \overline{sy^k}}{\sigma y^k}, \frac{sz_t^k - \overline{sz^k}}{\sigma z^k} \right). \quad (4)$$

where $\overline{sx^k}$, $\overline{sy^k}$, $\overline{sz^k}$ represents the average values of each axis acceleration data after smoothing respectively, σx^k、σy^k、σz^k represent the standard deviation values of each axis acceleration data after smoothing respectively. The normalized triaxial acceleration sequence of 3-D hand gesture can be expressed as:

$$NA^k = \left\{ \left(nx_1^k, ny_1^k, nz_1^k\right), \cdots, \left(nx_t^k, ny_t^k, nz_t^k\right), \cdots, \left(nx_n^k, ny_n^k, nz_n^k\right) \right\}. \quad (5)$$

Gesture Spotting. The signals of a gesture generally experience three phases: preparation, stroke and retraction. During the preparation phase and retraction phase, the 3-D acceleration signals are relatively stable due to the mobile phone being relatively static. The acceleration data captured during these phases are not suitable for user's gesture signature (see Fig. 1). Gesture spotting seeks to find the starting and ending time of the stroke phase. We adopt the following threshold-based spotting algorithm to determine the boundaries of the stroke phase:

$$ps = \min\left\{ xs, ys, zs \mid nx_{xs}^k > \eta_x, ny_{ys}^k > \eta_y, nz_{zs}^k > \eta_z \right\}. \quad (6)$$

$$pe = \max\left\{ xe, ye, ze \mid nx_{xe}^k < \eta_x, ny_{ye}^k < \eta_y, nz_{ze}^k < \eta_z \right\}. \quad (7)$$

The gesture's starting point ps (ending point pe) is the first (last) time when normalized acceleration value along any one of

the three axis becomes larger (smaller) than the threshold vector along the same axis. The threshold vector is given by:

$$\left(\eta_x,\eta_y,\eta_z\right)=\left(\left(\sum_{t=1}^{sm} nx_t^k \Big/ sm\right)\pm\varepsilon,\left(\sum_{t=1}^{sm} ny_t^k \Big/ sm\right)\pm\varepsilon,\left(\sum_{t=1}^{sm} nz_t^k \Big/ sm\right)\pm\varepsilon\right). \quad (8)$$

where sm is a number of stable points, and the parameter ε is a constant between 0.05 and 0.5. The normalized 3-D gesture data $CA^k=\{CX^k,CY^k,CZ^k\}$ during the stroke phase [ps , pe] represent the spotted acceleration sequences along the three axes, where for example $CX^k=\left(nx_{ps}^k,\cdots,nx_{pe}^k\right)$.

4.3 Sequence Alignment and Interpolation

Sequence Alignment. Inconsistency exits among data collected from the same user making the same gesture multiple times due to the inability of human's imperfect physiological system to repeat the same gesture. This means that there are different points and amplitudes between the gesture sequences made by the same user. In order to reduce FRR, we use DTW to align different acceleration sequences to align them:

$$D\left(t_n,r_m\right)=d\left(t_n,r_m\right)+\min\left(D\left(t_{n-1},r_m\right),D\left(t_{n-1},r_{m-1}\right),D\left(t_{n-1},r_{m-2}\right)\right). \quad (9)$$

$$d\left(t_n,r_m\right)=\sqrt{\left(t_n-r_m\right)^2},1\le n\le N,1\le m\le M. \quad (10)$$

where t_n represents the n-th point value of test template, r_m represents the m-th point value of reference template, $d\left(t_n,r_m\right)$ represents the Euclidean distance between the 3-D acceleration vectors of points t_n and point r_m , $D\left(t_n,r_m\right)$ represents the minimum cumulative distance from point $\left(t_1,r_1\right)$ to point $\left(t_n,r_m\right)$. The main object of DTW is to find an optimal path from back to front, as a result, both signals are extended or compressed to be aligned and the points on this optimal path can be used to reconstruct two aligned sequences, denoted by $T\left(N\right)=\left(t_1,t_2,\cdots,t_N\right)$ and $R\left(M\right)=\left(r_1,r_2,\cdots,r_M\right)$, respectively.

Sequences Interpolation. The interpolation processing make two reconstruct aligned sequences having the same number of corresponding points through inserting zero values between the points of two reconstructed aligned sequences respectively to compensate the mismatching of the reconstructed aligned sequences. The inserted zero values must be modified to reduce the error between the two sequences introduced by the interpolation processing.

Rule 1 (interpolation rule): Assume $\left(t_p,r_q\right)$ is a pair of aligned sequence points, $t_p\left(1\le p\le N\right)$ is from $T\left(N\right)=\left(t_1,t_2,\cdots,t_N\right)$, and $r_q\left(1\le q\le M\right)$ is from $R\left(M\right)=\left(r_1,r_2,\cdots,r_M\right)$. If $p>q$, then insert $\left(p-q\right)$ zero points between t_p and t_{p+1} and $\left(p-q\right)$ zero points before r_q ; otherwise, insert $\left(q-p\right)$ zero points between r_q and r_{q+1} and $\left(q-p\right)$ zero points before t_p. Two sequences are created after interpolation, denoted as $T'\left(l\right)=\left(t_1',t_2',\cdots,t_i',t_{i+1}',\cdots,t_{i+m}',\cdots t_l'\right)$ and $R'\left(l\right)=\left(r_1',r_2',\cdots,r_i',r_{i+1}',\cdots,r_{i+m}',\cdots r_l'\right)$.

Rule 2 (zero value modification rule): The first and last inserted zero points will be modified as one half of value of their neighbor point; other inserted zero points will be modified as the average of the point before it and the next non-inserted point. That is:

$$\begin{cases} t_k'=t_2'/2,k=1 \\ t_k'=t_{l-1}'/2,k=l \\ t_k'=\left(t_{k-1}'+t_{i+m}'\right)/2,\forall k=i+1,\cdots,i+m-1, \\ \qquad \exists t_{i+1}'=t_{i+2}'=\cdots=t_{i+m-1}'=0\wedge t_i'\ne 0\wedge t_{i+m}'\ne 0 \end{cases}. \quad (11)$$

4.4 The Similarity Criteria

The difference $\delta'\left(T',R'\right)$ between two sequences $T'=\left(t_1',t_2',\ldots,t_l'\right)$ and $R'=\left(r_1',r_2',\ldots,r_l'\right)$ is quantified by absolute distance (or L^1 norm):

$$\delta'\left(T',R'\right)=\sum_{j=1}^{l}\left|t_j'-r_j'\right|. \quad (12)$$

Similar sequences will have smaller distance $\delta'\left(T',R'\right)$. The similarity criteria of two sequences is:

Rule 3 (similarity rule): if the absolute distance between two sequences is smaller than a given threshold, $\delta'\left(T',R'\right)<\lambda$, the two sequences are classified as similar, otherwise, the two sequences are classified as non-similar.

4.5 User Authentication Process

Gesture Template Enrollment. During this phase，the genuine user is required to enroll his 3-D identifying gesture templates by repeating the gesture three times so the accelerometer can capture three acceleration sequences A^1 , A^2 , A^3 . After the processing described above in subsections 4.2 to 4.4, we can obtain:

$$\begin{aligned} \delta\left(CA^1,CA^2\right)&=\left[\delta'\left(CX^1,CX^2\right)+\delta'\left(CY^1,CY^2\right)+\delta'\left(CZ^1,CZ^2\right)\right]/3 \\ \delta\left(CA^1,CA^3\right)&=\left[\delta'\left(CX^1,CX^3\right)+\delta'\left(CY^1,CY^3\right)+\delta'\left(CZ^1,CZ^3\right)\right]/3 \\ \delta\left(CA^2,CA^3\right)&=\left[\delta'\left(CX^2,CX^3\right)+\delta'\left(CY^2,CY^3\right)+\delta'\left(CZ^2,CZ^3\right)\right]/3 \end{aligned}. \quad (13)$$

where $\delta\left(CA^1,CA^2\right)$, $\delta\left(CA^1,CA^3\right)$ and $\delta\left(CA^2,CA^3\right)$ are distances between each pair of gestures as defined in subsection 4.4. We store A^1 , A^2 , A^3 and the average template distance μ (defined below) in the mobile phone as the genuine user's template biometric characteristic data.

$$\mu=\left[\delta\left(CA^1,CA^2\right)+\delta\left(CA^1,CA^3\right)+\delta\left(CA^2,CA^3\right)\right]/3. \quad (14)$$

User Identity Verification. Let A^v be the verifying gesture sequence collected when a user performs the 3-D verifying gesture in order to gain access to the phone. We compute the absolute distances $\delta\left(CA^v,CA^1\right)$, $\delta\left(CA^v,CA^2\right)$ and $\delta\left(CA^v,CA^3\right)$ as described above and the average verifying distance ψ :

$$\psi=\left[\delta\left(CA^v,CA^1\right)+\delta\left(CA^v,CA^2\right)+\delta\left(CA^v,CA^3\right)\right]/3. \quad (15)$$

Rule 4 (identity verification rule): For a given threshold θ , if $\psi/\mu\le\theta$, the user is authenticated; otherwise, authentication fails. The value of the threshold θ determines the accuracy of the authentication. Small θ results in high FRR while large θ gives high FAR. We will discuss the selection of appropriate value for θ in the next section.

5. EXPERIMENTAL RESULTS

From the design of the proposed 3-D hard gesture signature (HGS) based authentication system, we see that it meets four of the eight properties that user behavioral biometrics should satisfy: universality, collectability, acceptability, and minimality. The goal of our experimentation is to validate the other four: ***uniqueness, permanence, circumvention, and performance (accuracy and speed)***. For this purpose, we have implemented the 3-D HGS on smart phones of different brands and recruited volunteers to perform in-field experiments. In this section, we describe our experimentation and report results.

5.1 Experimentation Setup

The proposed 3-D HGS biometric authentication system has been developed (roughly 7200 lines of code in Java and Matlab) and implemented on smart phones of three brands: Huawei, Samsung, and Sony. Table 1 shows the important features of these phones. In the rest of this section, when we report the performance of our authentication system on these phones, we will use generic names PhoneA, PhoneB, PhoneC with no particular order to avoid any direct comparison among different brands.

Table 1. Key parameters of smartphones.

Brand Name	Model	Memory Size	Operating System	Accelero-meter Model
Huawei	Ascend P6	2 GB	Android OS 4.2	LIS303DLHC
Samsung	I9192	1.5 GB	Android OS 4.2	MPU-6K
Sony	L36h	2 GB	Android OS 4.1	BMA250

We have recruited 19 volunteers to perform various tests on our system,including 11 males and 8 females, the range of their ages is from 18 to 35 with average around 25. We randomly select 5 of them as the genuine users and the rest will only play the role of impostors. In the enrollment phase, each genuine user makes his/her specific hand gesture signature being different from each other three times for each brand of the phones. It is not required to halt between each trial as long as the user can return his/her hand to the normal position. The accelerometer will capture the gesture with a sampling rate of 50 Hz; our on-phone 3-D HGS will process the data and enroll the signature in the phone. Everyone, including the impostors, observes this phase.

The verification phase is tested as follows: each genuine user repeats his/her registered hand gesture 1,000 times on each brand of phones (except User 2 who accidentally repeats 1,500 times). For the signature of each genuine user on each brand of phones, each of the 18 users (the 14 impostors and 4 other genuine users) is asked to immediately imitate the genuine user's signature for a minimum of 200 and a maximum of 300 times (except in the case of User 2 who repeated his signature 500 times more than required and did not leave enough time to the impostors). The authentication process was tested for a total of 76,520 times, where 16,500 are performed by genuine users on their genuine phones and the other 60,020 are imitation attempts from either impostors or genuine users on other's phone. Each time, the on-phone 3-D HGS authentication system will either deny or grant the attempt to access the phone. User can see this decision right after he/she completes the hand gesture, without any noticeable delay. This, combined with the fact that no waiting time in the enrollment phase, confirms the ***speed part of the performance property***. In the rest of this section, we will analyze the data to validate other properties: *uniqueness, permanence, circumvention, and performance (accuracy)*.

Finally, to compare the performance with existing approaches, we have also implemented two other similar authentication systems, DTW[13] and GSA [14] in Matlab, and test them on a subset of the collected data exported from the smartphones.

5.2 On-Phone Performance Evaluation

Test 2 reports the FAR and FRR for each of the five genuine users on each of the three phones. The values of FAR are obtained by dividing the number of successful imitation by the number of imitation (INUM); the values of FRR are obtained by dividing the number of genuine user's failed authentication by the number of attempts (UNUM). Here we discuss several interesting findings related to the *uniqueness, permanence, circumvention, and performance/accuracy* of the proposed 3-D HGS based biometric authentication system.

First, as shown in the last row of Table 2, the average FAR among the five genuine users on the three phones are 0.28%, 0.30%, and 0.28%; and the average FRR are 4.18%, 4.92%, and 4.90%. This shows that our proposed 3-D HGS based biometric authentication system is robust and platform independent. That is, it does not perform perfectly on one brand and terribly on the other. The reason of different results on different smartphones is that the performance is related with the stable of the working frequency of an accelerometer, unfortunately, the frequency of a smartphone is not stability and different in different smartphones.

Second, at individual user level, we can also observe similar robustness. Take User 3 for example, the FAR values on three phones are 0.30%, 0.33%, and 0.40%. The average FAR values are reported in the second column from the right. These values vary from 0.14% in User 1 who has a relatively complicated gesture with fast speed and changeable behavior, to 0.52% in User 5 whose gesture is relatively simple with slow speed and no-changeable behavior. They measure the difficulty of faking a genuine gesture.

However, the nature of the gesture does not correlate well with FRR. For instance, both User 2 and User 4 have an average FAR of 0.22%, but their FRR values are 7.75% and 2.30%, respectively. We believe this is caused by user's different ability in performing the gesture consistently on different models. For example, User 4 is the most consistent one with 2.40%, 2.90%, and 1.60% FRR on three different phone; User 3 also has consistent FRR values of 6.70%, 5.40%, and 5.20%, but these high values show the user's inconsistence in repeating his gesture. Meanwhile, the other three users seem to perform badly on one of the phones, User 1 on PhoneC (FRR=7.00%), User 2 on PhoneB (FRR=10.90%), and User 5 on PhoneA (FRR=5.10%).

Finally we mention that in most cases, we have very low FAR and relatively high FRR. This is due to the design requirement for such authentication systems, it is ok to occasionally deny the access of a genuine user, but the damage of giving impostors access could be devastating. In our 3-D HGS based biometric authentication system, this can be controlled by the selection of threshold θ defined in section 4.5. In our experiments, we have set θ=1.80. A detailed discussion about the selection of this threshold value can be found in section 5.3.3.

Table 2. The test results on different smartphones. INUM: number of total imitations; UNUM: number of attempts by the genuine user; FAR: false acceptance rate (in %); FRR: false rejection rate (in %); TOTAL: overall of the three brands of phones.

User	PhoneA				PhoneB				PhoneC				TOTAL			
	INUM	UNUM	FAR	FRR	INUM	UNUM	FAR	FRR	INUM	UNUM	FAR	FRR	INUM	UNUM	FAR	FRR
User 1	4600	1000	**0.20**	**3.60**	4000	1000	**0.20**	**3.60**	3600	1000	**0.00**	**7.00**	12200	3000	**0.14**	**4.73**
User 2	3500	1500	**0.20**	**5.47**	3200	1500	**0.28**	**10.90**	3500	1500	**0.17**	**6.87**	10200	4500	**0.22**	**7.75**
User 3	4000	1000	**0.30**	**6.70**	4000	1000	**0.33**	**5.40**	4000	1000	**0.40**	**5.20**	12000	3000	**0.34**	**5.77**
User 4	4860	1000	**0.25**	**2.40**	4080	1000	**0.27**	**2.90**	5320	1000	**0.17**	**1.60**	14260	3000	**0.22**	**2.30**
User 5	3700	1000	**0.49**	**5.10**	3960	1000	**0.43**	**3.70**	3700	1000	**0.65**	**2.50**	11360	3000	**0.52**	**3.77**
Average	_ _	_ _	**0.28**	**4.18**	_ _	_ _	**0.30**	**4.92**	_ _	_ _	**0.28**	**4.90**	_ _	_ _	**0.28**	**5.13**

5.3 Performance Simulation

In order to show the performance of our proposed 3-D HGS based biometric authentication system with existing similar systems, we have implemented our systems and two other popular systems, DTW and GSA, in MATLAB7.11.0. We also use the Matlab tool to show the effectiveness of our proposed system.

5.3.1 3-D HGS Data Processing Operations

We first report the Matlab simulation of the data processing operations described in Section 4. The four figures in Figure 1 illustrated a sample data collected by the accelerometer on a Huawei phone and the effect of the preprocessing of data smoothing, normalization and spotting.

The upper left figure shows the acceleration values along the three axes. From top to bottom at the time t=0, the three curves represent the 3-D acceleration along z-axis, x-axis, and y-axis, respectively. Clearly we can see that this hand gesture started around t=3,000 ms, lasted about 6 seconds, and finished around t=9,000 ms. These curves are oscillating due to noise.

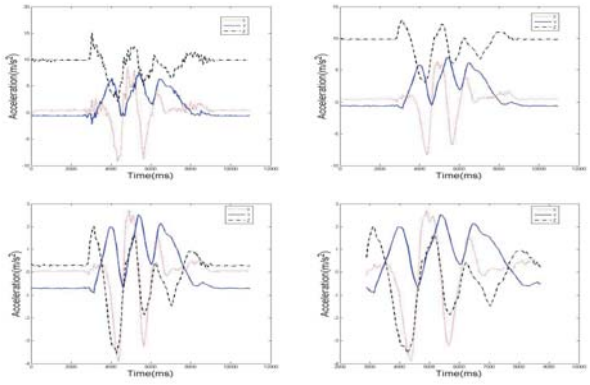

Figure 1. Original data sample and data after the data processing operations described in Section 4. x-axis: time (ms); y-axis: acceleration value (m/s^2). Upper left: original sample data; upper right: smoothed data; lower left: normalized data; lower right: spotted data.

After applying formula(2), all the curves become smooth as shown in the upper right figure. Notice that in these two figures on top, the acceleration values are from around -8.5 m/s^2 to around 14.0m/s^2. We normalize the smoothed data with formula(4) to obtain the lower left figure. The normalization preprocessing will reduce the amount of calculation of gestures matching algorithm and improve the efficiency of the system. Finally, after spotting process, the three acceleration curves are shown in the lower right figure which eliminate the non-gesture part of the data.

5.3.2 Sequence Alignment and Interpolation

We now show the effectiveness of the sequence alignment and interpolation methods in Section 4.3. We only analyze the data processing of the 3-D hand gesture along x-axis. The y-axis and z-axis gesture signal have similar behavior.

Figure 2 shows the sequence processing results of the two groups of acceleration data when the same user performs the same gesture twice. There are some mismatch on time points between the two similar sequences before being preprocessed, and this mismatch will be eliminated after being alignment and interpolation processed. It is obviously that the two sequences are aligned and time points matched pretty well.

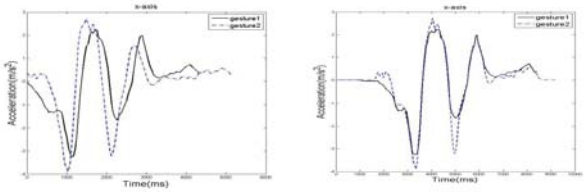

Figure 2. Preprocessed (left) and aligned and interpolated (right) same gesture sequences performed by same user.

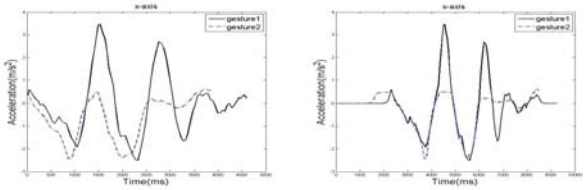

Figure 3. Preprocessed (left) and aligned and interpolated (right) same gesture sequences performed by different users.

Figure 3 shows the sequence processing results of the same gesture performed by two different users. The first user is genuine user and the second user mimics the genuine user's gesture. It is obviously the time points are mismatch between the genuine user's acceleration data and the imposter's imitation acceleration data, and two sequences are not aligned as well as that in Figure 2.

The comparison between Figure 2 and Figure 3 shows that the differences between the same gesture performed by different users

are larger than that performed by the same user. This confirms that the proposed method is suitable for authentication.

5.3.3 Impact of Threshold θ and Its Value Selection

Recall that in the user identity verification phase, a threshold value θ is used to determine whether the user is genuine. This threshold plays a vital role in the proposed 3-D HGS based biometric authentication system. Therefore, the value of θ has to be selected carefully.

To show this impact quantitatively, we export 400 randomly selected data samples from the smartphone, 200 are from the genuine user and the other 200 are from the impostors. We process the data for user identify verification, and use different threshold value to decide whether the data comes from the genuine user or some impostor. The results are plotted in Figure 4.

First, as the threshold value θ starts going up from 1.0, the FRR drops dramatically. This is because a small valued θ requires the gesture sample to have a very high similarity with genuine user's 3-D HDS template for authentication, accordingly to rule 4. When θ becomes 2.2 or larger, the FRR almost drops to 0. However, we can also see from Figure 4 the opposite trend for FAR. When the threshold value θ is small, it is hard for an impostor to pass the authentication system.

When we use the standard EER criterion, we find that the two curves meet when θ is approximately 1.86, which means FAR=FRR, or it is equally likely for the system to accept an impostor or to reject the genuine user. This provides a significant information as to which value should be chosen at the point to accept or reject in a real application. When the application is in favor of having a low FAR, we should choose θ to be smaller than the EER value, 1.86 in this case; when the application is user friendly and prefers a low FRR, θ needs to be set with a larger value.

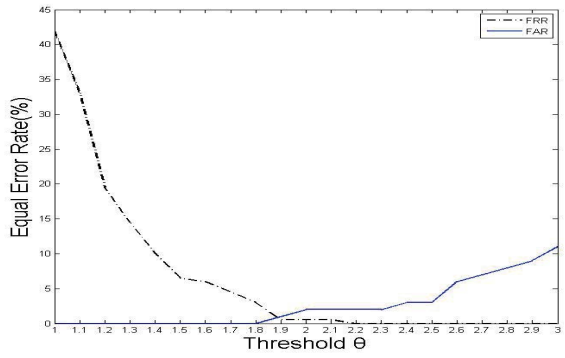

Figure 4. Error rates under different threshold values. The solid line on the right: FAR; dashed line on the left: FRR.

In our case when user authentication is performed to decide whether a user can access a smart phone, we believe that low FAR is more important. This is because that the genuine user can repeat his hand gesture to authenticate himself in case of false rejection, which will not bother most of the users. Just think about how many times we have to re-enter a well-designed password. But when the authentication system mistakenly accepts an impostor, the system (smartphones in this case) will be accessed by the wrong person and the damage can be huge. Simulate results indicate that FAR becomes close to zero when θ = 1.80 and if we further reduce the value of θ, FAR remains almost unchanged but FRR increases dramatically. Therefore, we have set θ = 1.80 throughout the experimentation and simulation.

5.3.4 Comparison with Other Similar Systems

To validate the performance of our proposed 3-D HGS based authentication system, we compare it with two popular gesture based authentication systems: DTW and GSA. We implement all three systems in Matlab and test them on the real gesture data we have collected.

Among the total 76,520 gesture samples we have collected from 5 genuine users and 14 impostors, we randomly choose 3 genuine users and 10 impostors. Then we export 300 samples from each of these 3 genuine users (a total of 900) and 120 samples from each impostor when he/she mimics each of the genuine users (a total of 120*10*3=3,600) from the smart phones to a desktop PC where the three authentication systems are implemented.

Table 3 reports the accuracy of the three systems, measured by FAR and FRR. As we have analyzed earlier, FAR is the most important factor for authentication systems. The proposed method 3-D HGS achieves an average FAR of 0.22%, while GSA[14] has 1.01% and DTW[13] has 0.59%. Furthermore, one can see that our system indeed has zero FAR for both User 1 and User 2. User 3 here is the user who has a relatively simple hand gesture. At the same time, our method also has a lower FRR, 3.67% comparing to 4.63% and 4.93% from the other two systems. The 3-D HGS can obtain higher performance than GSA and DTW due to the minute differences between sequences being extracted by the sequence alignment with 3-D HGS.

Table 3. The result of simulation

User	FAR			FRR		
	3-D HGS	GSA	DTW	**3-D HGS**	GSA	DTW
User 1	**0.00**	0.57	0.28	**2.80**	4.10	5.00
User 2	**0.00**	0.67	0.50	**3.40**	4.60	4.20
User 3	**0.67**	1.80	1.00	**4.80**	5.20	5.60
Average	**0.22**	1.01	0.59	**3.67**	4.63	4.93

6. CONCLUSIONS AND FUTURE WORK

Despite the numerous reported research activities and commercial products on every major brand, the security of mobile phone remains a challenging problem due to the limited resource (memory, energy, computation and communication power).

In this paper, we focus on the problem of user authentication for smart phones and make the following contributions: we first list a set of requirements that good user behavioral biometrics should satisfy; then we propose to build an authentication system based on user's 3-D hand gesture signature (HGS) that can be conveniently captured by on-phone accelerometers. We discuss how our 3-D HGS based system meets the requirements by design and demonstrate its performance by carefully designed field experiments with real users and real phones on which we have implemented the 3-D HGS authentication system. The results show that our system can achieve very low FAR and FRR, and outperforms two other similar authentication systems. The field experiment also shows the usability of the proposed system.

Encouraged by these promising experimental results, we plan to continue our investigation and development of the 3-D HGS based biometric authentication system in the following directions:

1. *More comprehensive in-field testing.* Although we have recruited 19 users and collected a total of 76,520 3-D hand gestures on three different brand of smart phones, we feel that more empirical study is still needed. More specifically,
   - Follow up the genuine users who will use the system on a daily basis to collect more data.
   - Recruit volunteers from a more diversified group. For example, the gesture of young smart phone users may not be as stable as adults. It will be interesting to see how our system performance on this group of users.
   - Perform test on more phones, and if possible, use more different brands. Or more accurately, test on different brand of accelerometers.
2. *Additional research issues.* There are several challenges that need more than just in-field testing to address.
   - The selection of the verification threshold value θ. As we have discussed earlier, this value plays a critical role in the accuracy of the authentication system. We believe that the make of the accelerometer and the genuine user's ability to repeat his gesture are the two major factors to determine θ. To commercialize our system, collaboration with the accelerometer manufacture and scientists on human social behavior will be needed.
   - Automatic updating of the registered gesture. People's gesture changes slowly as they grow or age. We plan to design an adaptive mechanism that will follow genuine user's attempts, particularly those that barely meet/fail the verification criterion, and learn how the same gesture evolves and update the gesture pattern.
   - To avoid replay attacks, we will research template encryption algorithm to protect the template. And even more, we will do some works for normalizing the signature data to get a sort of ground truth.
   - We did not measure the speed and power consumption of our authentication system. In on-phone experiments, the acceptance/rejection decision is normally made instantaneously after the user's hand gesture. It will be helpful to measure and optimize these performance metrics when the system becomes more sophisticated.

7. ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant No.61373126，the National Natural Science Foundation of China under Grant No.61228204, the Natural Science Foundation of Jiangsu Province of China under Grant No.BK20131107， and The Scientific Research Foundation project of Ministry of Education of China-China Mobile (MCM20122062) for sponsoring this project. This paper is the result of the research findings of this project. We express our sincere thanks to them.

8. REFERENCES

[1] Jain, A., Bolle, R., and Pankanti, S. 1996. *Introduction to biometrics*. Biometrics, Springer US, 1-41.

[2] Jain, A., Hong, L., and Pankanti, S.2000.Biometric identification. *J. Communications of the ACM*.43, 2(Feb. 2000), 90-98.

[3] Gafurov, D.2007.A Survey of Biometric Gait Recognition: Approaches, Security and Challenges. In *Proceedings of the Annual Norwegian Computer Science Conference*. 119-130.

[4] Abeni, P., Baltatu, M. and D'Alessandro, R.2006.Nis03-4: Implementing Biometrics-Based Authentication for Mobile Devices. In *Proceedings of the Global Telecommunications Conference*.1-5.

[5] Derawi, M. and Bours, P.2013. Gait and activity recognition using commercial phones. *J. Computers & Security*. 39(Nov. 2013), 137-144.

[6] Drosou, A., Ioannidis, D., Moustakas, K. and Tzovaras, D.2012.Spatiotemporal analysis of human activities for biometric authentication. *J. Computer Vision and Image Understanding*. 116, 3(mar. 2012), 411–421.

[7] Kim, D.J., Chung, K.W. and Hong, K.S.2010. Person Authentication using Face, Teeth and Voice Modalities for Mobile Device Security. *J. IEEE Transactions on Consumer Electronics*. 56, 4(Nov. 2010), 2678-2685.

[8] Crawford, H., Renaud, K. and Storer, T.2013. A framework for continuous, transparent mobile device authentication. *J. Computers & Security*. 39(Nov. 2013), 127-136.

[9] Nymi. *http://www.getnymi.com/*, 11-19. 2013.

[10] Conti, M., Zachia-Zlatea, I. and Crispo, B.2011. Mind how you answer me: transparently authenticating the user of a smartphone when answering or placing a call. In *Proceedings of the sixth ACM Symposium on Information, Computer and Communications Security*.249-259.

[11] Manabe, H., Yamakawa, Y., Sasamoto, T. and Sasaki, R. 2009. Security Evaluation of Biometrics Authentications for Cellular Phones. In *Proceedings of the fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*.34-39.

[12] Gascon,H., Uellenbeck, S. , Wolf, C., Rieck, K.2014. Continuous Authentication on Mobile Devices by Analysis of Typing Motion Behavior. In *Proceedings of GI Conference.* 1-12.

[13] Bailador, G., Sanchez-Avila, C., Guerra-Casanova, J. and de Santos Sierra, A.2011. Analysis of pattern recognition techniques for in-air signature biometrics. *J. Pattern Recognition*. 44(Nov. 2011), 2468-2478.

[14] Guerra-Casanova, J., Sánchez-Ávila, C., de Santos Sierra, A. and et al 2011.Score optimization and template updating in a biometric technique for authentication in mobiles based on gestures. *J. Journal of Systems and Software*. 84, 11(Nov. 2011), 2013-2021.

[15] Guerra-Casanova, J., Sánchez-Ávila, C., Bailador, G. and de Santos Sierra, A.2012. Authentication in mobile devices through hand gesture recognition. *J. International Journal of Information Security*.11, 2(Apr. 2012), 65–83.

[16] Guerra-Casanova, J., Sánchez-Ávila, C., Bailador-del Pozo, G.and de SantosSierra, A.2011. Application of LCS Algorithm to Authenticate Users within Their Mobile Phone Through In-Air Signatures. *Advanced Biometric Technologies*. 265-280.

[17] Jain, A.K., Ross, A. and Prabhakar, S.2004. An introduction to biometric recognition. *J. IEEE Transactions on Circuits and Systems for Video Technology*. 14, 1(Jan. 2004), 4-20.

[18] Lee-Cosioa, B.M., Delgado-Mataa, C. and Ibanezb, J. 2012. ANN for Gesture Recognition using Accelerometer Data. *J. Procedia Technology*. 3(May 2012), 109-120.