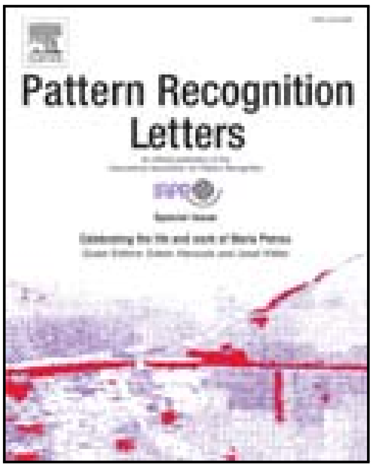

Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures

Ankita Jain , Vivek Kanhangad

PII: S0167-8655(15)00205-6
DOI: 10.1016/j.patrec.2015.07.004
Reference: PATREC 6271

To appear in: *Pattern Recognition Letters*

Received date: 15 January 2015
Accepted date: 1 July 2015

Please cite this article as: Ankita Jain , Vivek Kanhangad , Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures, *Pattern Recognition Letters* (2015), doi: 10.1016/j.patrec.2015.07.004

Research Highlights (Required)

- A novel approach for user verification in smartphones using touchscreen gestures.
- Performance of the approach is evaluated on relatively large dataset of 104 users.
- Orientation sensor information outperforms other features considered in this work.
- Proposed approach achieves 0.31% EER for score level fusion of all gestures.

Pattern Recognition Letters
journal homepage: www.elsevier.com

Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures

Ankita Jain[a] and Vivek Kanhangad[a]*

[a]*Discipline of Electrical Engineering, Indian Institute of Technology Indore, India*

ABSTRACT

In this paper, we propose an approach for user authentication in smartphones using behavioral biometrics. The approach involves analyzing behavioral traits while the user performs different gestures during his interaction with the device. In addition to the commonly employed features such as x - y coordinate information and finger area, the proposed approach utilizes the information from orientation sensor for each of the seven gestures considered in this study. The feature set is further enriched with features such as accelerometer sensor reading, curvature of the swipe. Matching of corresponding features is performed using the modified Hausdorff distance (MHD). Performance evaluation of the proposed authentication approach on a dataset of 104 users yielded promising results, suggesting that the readings from orientation sensor carry useful information for reliably authenticating the users. In addition, experimental results demonstrate that consolidating multiple features results in performance improvement. The proposed method outperforms dynamic time warping (DTW) based matching for all gestures considered in this study, with significant reduction in EER from 1.55% to 0.31% for score level fusion of all gestures. In addition, the performance of the proposed algorithm is ascertained on a dataset of 30 subjects captured using another smartphone.

Keywords: Behavioral biometrics; smartphones; modified Hausdorff distance; personal authentication; touchscreen gestures; orientation sensor reading.

* Corresponding author. Tel.: +91-732-4240-757; fax: +91-731-2361-482; e-mail: kvivek@iiti.ac.in

1. Introduction

With advances in computing technology, smartphones have become feature rich, affordable and one of the widely used electronic gadgets. It is quite common to see people performing banking transactions, accessing confidential or sensitive information such as corporate data, email and social media accounts using their smartphones. Apart from this, smartphones are often used as a means for storing important and sensitive information. Therefore, it is extremely important that these devices are able to perform user authentication, even more so, in the event of lost or stolen phones in order to prevent access to user's personal data by impostors.

Traditionally, personal computers have been secured by password based authentication, which has the inherent limitation that the passwords can be stolen or forgotten. Behavioral biometrics such as keystroke and mouse dynamics provide a more secure and reliable alternative to password based authentication in computers [1]. Keystroke dynamics performs authentication based on the way the user operates the keyboard [2], while mouse dynamics [3, 4] is an emerging behavioral biometric trait that performs authentication based on user's mouse operating patterns. In general, user authentication in computers can be achieved in two ways; static and continuous authentication. Static authentication is a one-time user verification performed at the time of unlocking the computer [2]. The major limitation of this approach is that once the system is unlocked, it does not continue to provide any security against unauthorized access. In other words, an unauthorized person can operate the system as it has already been unlocked. On the other hand, continuous authentication [5] overcomes this problem by constantly authenticating the user and thereby, preventing unauthorized access to the system by an impostor.

The rest of the paper is organized as follows: Section 2 describes existing work in the area of biometrics authentication in portable devices. The proposed approach is detailed in Section 3, in which firstly, an overview of the approach is provided followed by detailed descriptions of techniques employed for feature extraction, score computation, score normalization and fusion of scores. Section 4 presents experimental results and discussion. This section also presents a description of the datasets used for performance evaluation. Finally, conclusions are presented in Section 5.

2. Related work and motivation

In recent years, with rapid increase in the usage of smartphones and subsequent increase in the security issues associated with it, researchers have focused their effort on developing methodologies for static as well as continuous user authentication. Static authentication in smartphones is achieved using traditional or biometrics authentication techniques. Traditional techniques include swipe or click to unlock, creating a specific pattern on the screen to unlock and password based authentication, whereas commonly employed biometric techniques for static authentication include face [6, 7, 8], voice [9, 10, 11], fingerprint [12, 13] and keystroke [14, 15, 16] based authentication. In addition to this, researchers have explored other physiological biometric traits such as finger knuckle [17], iris [18] and palmprint [19, 20, 21]. Multimodal biometric solutions [22, 23, 24, 25, 26, 27, 28] have also been proposed for mobile user authentication. However, it may be noted that these biometrics traits are more suited for providing static authentication in mobile phones. A detailed survey of various biometric approaches proposed for user authentication in mobile phones can be found in [29].

Keystroke based authentication in computers inspired researchers to explore similar approaches for touchscreen devices. Authors in [30] performed user authentication employing the information acquired from a touchpad, which records finger pressure, finger position, hold time and inter-key time. The approach achieved 1% equal error rate (EER) by employing only finger pressure information. However, their approach was evaluated on a dataset of only 10 users. Luca et al. [31] presented an approach for behavioral biometrics based static authentication while the user performs a password pattern on touchscreen. As discussed earlier, static user authentication has inherent limitations and therefore, continuous authentication approaches are more desirable in smartphones. Most of the existing approaches for continuous authentication are based on behavioral characteristics acquired during users' continuous interaction with the device.

Frank *et al.* [32] proposed an approach for continuous user authentication based on 30 behavioral features extracted from the touchscreen input. Authors explored two classifiers, namely k-nearest neighbors (k-NN) and support vector machine and reported EER between 0-4% for different experimental scenarios. However, authors in their study considered only up-down and left-right scrolling. Feng *et al.* [33] developed a glove sensor that captures linear and angular acceleration of finger movement. By combining the information acquired from the touchscreen and the sensor glove, their approach yielded false accept rate (FAR) of 4.66% and false reject rate (FRR) of 0.13%. However, the use of additional sensor to improve the accuracy is the major drawback. In another work, Feng *et al.* [34] proposed to maintain separate templates for each application. The proposed approach was evaluated on a dataset of only 23 users. In addition, the increased memory requirements due to separate templates might limit its applicability in mobile phones. Authors in [35] presented an approach based on 21 features, which are fed as input to a neural network classifier. Further, they employed particle swarm optimization (PSO) to optimize the neural network and reported an improved EER of 2.92%. Authors in [36] acquired 22 multi-touch gestures such as drag, swipe, pinch and user-defined gestures on an iPad. They employed dynamic time warping (DTW) for matching and achieved an average EER of 7.88%. Antal *et al.* [37] gathered data from 71 users for horizontal and vertical scrolling using 8 different tablets and mobile phones. Authors explored k-NN and random forest algorithm for classification and reported more than 95% accuracy. However, authors considered only horizontal and vertical scrolling in this work. Zhao *et al.* [38, 39] proposed a novel way of user authentication, in which trace of the points (swipe gesture) are converted into image. They represented the trace movement and pressure as shape and intensity values in a 2D image. On a dataset of 30 users, their approach achieved 2.62% EER by combining six gestures [38].

The primary objective of our study is to explore additional information available from the built-in sensors that can potentially improve the performance of user authentication in smartphones. The major contributions of the work reported in this paper can be summarized as follows:

1) We propose a new approach for user authentication in smartphones using modified Hausdorff distance for matching.

2) This study investigates usefulness of information acquired from orientation sensor for user authentication.

3) Performance evaluation of the proposed approach is performed on a relatively large dataset of 104 users.

4) Performance evaluation on a second dataset of 30 users acquired using another smartphone to study the impact of device on the verification performance.

3. Proposed approach

Fig. 1 shows a complete overview of the proposed approach. In this approach, an android application is employed to capture

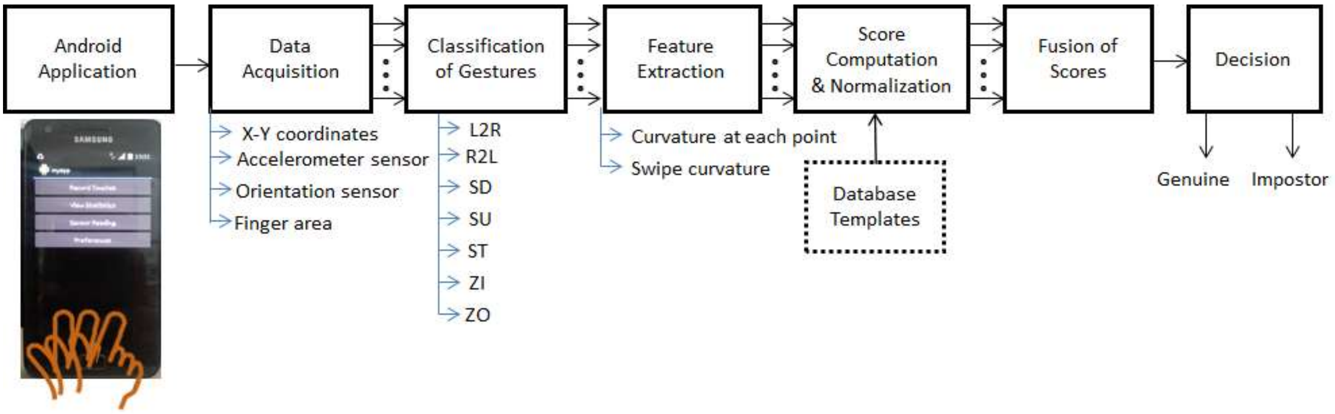

Fig. 1. Overview of the proposed approach for user authentication

the behavioral data during user's interaction with the device. Specifically, a set of behavioral data is captured at every touch point when the user performs a specific gesture on touchscreen. In this study, we consider the following seven gestures: left to right swipe (L2R), right to left swipe (R2L), scroll up (SU), scroll down (SD), zoom in (ZI), zoom out (ZO) and single tap (ST). The set of data captured by our application for each of the above gestures include x - y coordinates, accelerometer sensor readings, orientation sensor readings and the area covered by the finger on the screen. In the next step, the acquired data is processed to identify the category of the gesture. This is done in order to facilitate the matching of feature templates of the corresponding gestures in the matching stage.

In addition to the behavioral data, which we also refer to as raw data, captured by our application, two additional features are computed by using the x - y coordinate information. One of the features is the curvature at each touch point of swipe and the other one is the curvature of swipe. The computation of these features is detailed in Section 3.2. During the enrolment phase, the behavioral data, along with the two extracted features are stored as feature templates in the database. In this way, the database contains feature templates corresponding to each of the seven gestures performed by every user. During the authentication stage, the android application running on the smartphone captures user's behavioral data. In the first step of processing, the type of gesture performed by the user is identified and the two additional features are extracted to form a query feature set. This is followed by matching of the query feature set with the corresponding feature templates from the database, generating multiple matching scores. These matching scores are then normalized and combined by fusion of scores technique to obtain the final score. Finally, the decision stage utilizes the final score to determine whether the user is genuine or impostor.

3.1. Classification of gestures

In this work, we considered the following seven most commonly used gestures in touchscreen devices: left to right swipe (L2R), right to left swipe (R2L), scroll up (SU), scroll down (SD), zoom in (ZI), zoom out (ZO) and single tap (ST). L2R and R2L are commonly used for performing tasks such as unlocking the phone, browsing photos and switching between the home pages. SD and SU gestures are often performed while reading a document and browsing internet. ZI and ZO gestures are used for viewing specific content in images and documents, while ST is used to write a message or mail and more generally, to select an option [38]. As discussed in the previous section, the identification of the gesture performed by the user is an important task in the proposed approach for user authentication. For this purpose, we developed a simple and efficient heuristic method based on the x - y coordinate information captured by the application. In this method, L2R, R2L, SU and SD gestures are identified by computing differences between x and y coordinates of the first and last points of the gesture. If the magnitude of the difference of x -coordinates is greater than that of y -coordinates, the gesture is either L2R or R2L. Further, based on the sign of the difference of x -coordinates, the gesture can be classified as L2R or R2L. In a similar way, SU and SD are identified based on the difference of y -coordinates. ZI and ZO gestures are first separated from other gestures using the finger count on touchscreen. Further, the distance between the start point of one finger and the start point of second finger is calculated. Similarly, the distance between the two end points is also calculated. If the distance between the start points is greater than that of the end points, the gesture is ZO, otherwise the gesture is classified as ZI. The identification of ST is trivial as it contains only a single point.

3.2. Feature Extraction

The feature set in the proposed approach comprises a set of raw behavioral data acquired by the android application and the two additional features extracted from the x - y coordinate information. The rest of this section provides detailed descriptions [40] of the constituents of our feature set:

1. x - y coordinates: It indicates the position of finger on touchscreen. Specifically, when user performs a gesture, this provides the x - y coordinate values of finger position at each time event.

2. Finger area: This provides an approximation of area of the finger touched on the screen. It returns a scalar value at a given time event. The actual value of touch area in pixels is normalized for the device's explicit range and scaled to a value between 0 and 1.

3. Accelerometer: Reading from the accelerometer sensor provides the amount of acceleration on the device applied by the user. Acceleration is measured along three axes x , y and z of the device at each event time.

4. Orientation: Orientation sensor records the position of device with respect to the earth's frame of reference. At each event time, it provides orientation values in the following three directions:

Azimuth: It indicates the rotation around z -axis. Specifically, it is the angle between magnetic north and device's y -axis.

Pitch: It indicates the rotation around x -axis. The pitch value will be positive when the positive z -axis moves towards positive y -axis and will be negative in the opposite direction.

Roll: It indicates the rotation around y -axis. The roll value will be positive when the positive z -axis rotates towards positive x -axis and will be negative in the opposite direction.

Fig. 2 and 3 show patterns of orientation along the roll for two users acquired at two time instances. It may be observed from these figures that the two patterns belonging to a user are very similar (high intra-class similarity). It may also be observed that there is hardly any inter-class similarity in this case. This has motivated us to explore the effectiveness of information from the built-in orientation sensor for user authentication.

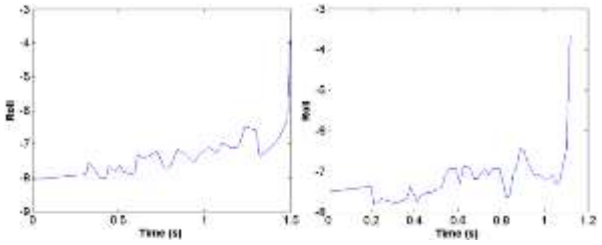

Fig. 2. Orientation along roll of a user at two time instances.

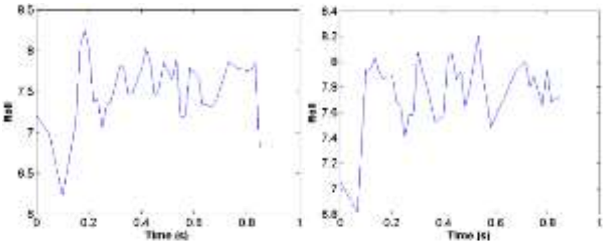

Fig. 3. Orientation along roll of another user at two time instances.

In addition to the set of behavioral data described above, we computed the following two additional features for every gesture, except for the ST gesture. Since the ST gesture contains only a single point for x - y coordinates, the following features cannot be computed.

1. Point curvature: Curvature at each point specifies the slope formed by the user at each successive point. At time t_i, curvature can be computed by [35]:

$$P_i = \tan^{-1}\left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right) \tag{1}$$

where x_i , y_i are the x and y coordinates of the sample point at event time t_i and x_{i-1} , y_{i-1} are the x and y coordinates of the sample point at the previous time event (t_{i-1}). Slope at the first point is considered to be zero.

2. Swipe curvature: Curvature of swipe specifies the slope formed by the user while performing a particular gesture. Curvature of swipe can be calculated by [34]:

$$S_i = \tan^{-1}\left(\frac{y_{end} - y_{start}}{x_{end} - x_{start}}\right) \tag{2}$$

where x_{start} and y_{start} are the x and y coordinates of the start point of gesture, x_{end} and y_{end} are the x and y coordinates of the end point of gesture.

3.3. Score computation and normalization

During enrolment, feature templates that contain the above mentioned features are created for each of the gestures. In a similar way, a query template is formed when the user performs a gesture on the touchscreen device. During verification, matching of the corresponding feature sets is performed using the modified Hausdorff distance (MHD) [41]. It has been observed through experiments that [41] the MHD consistently outperforms other possible Hausdorff based distance measures, when used to measure similarity or dissimilarity between two sets of points. Therefore, in this work, we employ MHD for computation of matching scores. The modified Hausdorff distance between two sets $A = \{a_1, a_2, \ldots, a_p\}$ and $B = \{b_1, b_2, \ldots, b_q\}$ is defined as

$$H(A,B) = \max(h(A,B), h(B,A)) \tag{3}$$

where

$$h(A,B) = \frac{1}{p}\sum_{a \varepsilon A}\left(\min_{b \varepsilon B}\|a - b\|\right) \tag{4}$$

$$h(B,A) = \frac{1}{q}\sum_{b \varepsilon B}\left(\min_{a \varepsilon A}\|a - b\|\right) \tag{5}$$

In the above equations, $\|.\|$ represents the L_2 norm between the two points of sets A and B. Also, $h(A,B)$ is the forward Hausdorff distance and $h(B,A)$ is the reverse Hausdorff distance. Essentially, the average of minimum distances from every point of one set with the other set is computed in both forward and reverse directions. The MHD is then computed by finding the maximum of the forward and the reverse Hausdorff distance.

The proposed approach for matching features using MHD generates matching scores for each of the constituents of our feature set. In the case of a scalar feature, the swipe curvature in our case, the feature matching using the MHD effectively reduces to matching the corresponding features using the Euclidean distance. As the information acquired from the accelerometer and orientation sensor is measured in three directions, the matching process generates a total of six scores. The matching of rest of the features that include x - y coordinates, finger area, point curvature and swipe curvature generates a score each. Therefore, a comparison or a match between a query and reference template in the database results in a total of 10 matching scores for each of the gestures, except for ST. As discussed earlier, since point curvature and swipe curvature cannot be computed for ST, the number of scores generated for a comparison between corresponding ST features is only 8.

The next step in our approach is to normalize multiple matching scores generated to a common domain. There are various techniques available in literature for score normalization [42, 45]. In this work, we have investigated the following techniques for score normalization.

1. min-max: min-max normalization technique linearly transforms matching scores into the range of 0 to 1. Normalized matching scores are computed as follows:

$$s_k^{'} = \frac{s_k - \min}{\max - \min} \tag{6}$$

where $\min$ and $\max$ are the minimum and the maximum values of matching scores, respectively.

2. z-score: In this technique, normalized matching scores are computed as follows:

$$s_k^{'} = \frac{s_k - \mu}{\sigma} \tag{7}$$

where μ and σ are the mean and standard deviation of matching scores, respectively.

3. tanh-estimator: The mathematical expression for score normalization using tanh-estimator is given as follows:

$$s_k^{'} = \frac{1}{2}\left\{\tanh\left(0.01 \times \left(\frac{s_k - \mu_G}{\sigma_G}\right)\right) + 1\right\} \tag{8}$$

where μ_G and σ_G are the mean and standard deviation of genuine scores. In the above equations, s_k and $s_k^{'}$ are the matching score and the normalized matching score, respectively.

4. w-score: The w-score scheme is originally proposed for score normalization in the recognition framework [45], and the approach uses distribution of scores generated by matching a probe to all gallery samples. Since we performed experiments in the verification scenario, which is more appropriate for user authentication in smartphones, overall distribution of genuine and impostor scores is used for score normalization. This is consistent with other score normalization techniques considered in our work.

3.4. Fusion of Scores

The score normalization technique transforms the scores into a common domain, so that the scores can be combined using a fusion method [43]. There are numerous ways by which scores can be combined. Some of them are sum of scores, maximum score, minimum score, weighted sum and product of scores [44]. In this work, we have explored the following rules for combination of matching scores.

1. Min Rule: In this rule, the combined score is the minimum of the set of scores being combined

$$S = \min(s_1, s_2, \ldots\ldots s_n) \quad (9)$$

2. Max Rule: The combined score is the maximum of the set of scores being combined

$$S = \max(s_1, s_2, \ldots\ldots s_n) \quad (10)$$

3. Product Rule: Product rule is mathematically expressed as follows:

$$S = \prod_{k=1}^{n} s_k \quad (11)$$

4. Sum Rule: Mathematically, the combined score is computed as follows:

$$S = \sum_{k=1}^{n} s_k \quad (12)$$

4. Experimental results and discussion

4.1. Dataset-I

Since there is no publicly available database that contains all the behavioral data that we have explored in this work, we developed a database of 104 users (Dataset-I). For the purpose of data acquisition, we developed an android application on IntelliJ IDEA platform and ran it on Samsung Galaxy S-II GT-I9100 android phone. Out of 104 users, 82 users were having prior experience of operating touchscreen phones. Participants in the data collection process conducted at our institute primarily included students aged between 19-36 years. Specifically, the dataset comprises 9 users between the age 31-36 years, 40 users between 26-30 years and 55 users between 19-25 years. These participants were asked to perform the following gestures on the touchscreen: left to right swipe (L2R), right to left swipe (R2L), scroll up (SU), scroll down (SD), zoom in (ZI), zoom out (ZO) and single tap (ST). Each of these gestures was performed three times by every user. As described in Section 3, the raw data captured by our application consists of x - y coordinates, finger area, accelerometer and orientation sensor readings from the touchscreen.

The following sections present results from a set of experiments carried out to evaluate the performance of the proposed approach. In our initial experiments, we evaluated various combinations of score normalization and fusion techniques. Results from these experiments are presented in Section 4.2. Based on these results, the best techniques for score normalization and fusion are identified and these techniques are investigated further.

4.2. Performance of different score normalization and fusion techniques

The objective of this set of experiments is to evaluate the performance of different score normalization and fusion techniques for matching gestures in our dataset. Table 1 summarizes EERs obtained for score-level fusion of all gestures using various score normalization and fusion techniques. As can be seen in this table, tanh-estimator normalization approach consistently provides the best matching performance, except for the case in which fusion is performed with min rule. According to the observations in [42], tanh-estimator is highly efficient and robust as it is less sensitive to outliers in the matching scores. It can also be noted that EER of w-score normalization with max rule is quite high as compared to its performance with other fusion rules. The major problem with the max rule is that if any of the individual scores being combined is 1 (upper limit of normalized scores), then the fused score gets confined to 1. In our experiments, we observed that w-score normalization is more likely (than tanh-estimator) to yield a score with value 1 and this leads to high verification error as majority of the fused scores have a value of 1. This explains the poor performance of the max rule based fusion with w-score normalization.

Table 1. EERs (%) obtained with different score normalization and fusion techniques

Score normalization technique	Score-level fusion technique			
	Min	Max	Product	Sum
min-max	12.84	47.14	50.14	24.12
z-score	64.29	56.09	50.02	50.64
w-score	11.76	50	5.43	1.92
tanh-estimator	27.80	3.40	0.32	0.31

In addition, the combination of tanh-estimator for score normalization and sum rule for fusion provides the best matching performance among different combinations considered. Fig. 4, 5, 6 and 7 show receiver operating characteristics (ROC) of score-level fusion techniques with different score normalization schemes. It may be noted that for the product and max rule based fusion schemes in Fig. 5 and 6, ROC curves corresponding to the w-score normalization are not plotted. This is due to the nature of distribution of fused matching scores. Specifically, range of values of genuine scores is quite high with majority of scores having low values and a few of them having very high values. This necessitates a very small increment in threshold to plot a smooth ROC curve. However, a small step size (threshold increment) results in high computational complexity and leads to memory error in MATLAB.

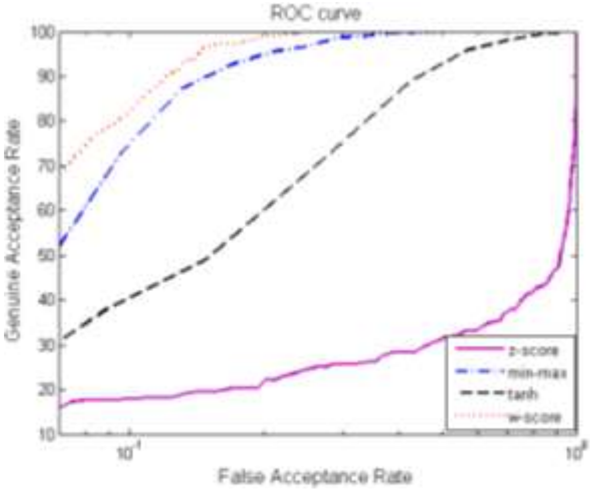

Fig. 4. ROCs for score-level combination of gestures with min rule and different normalization techniques

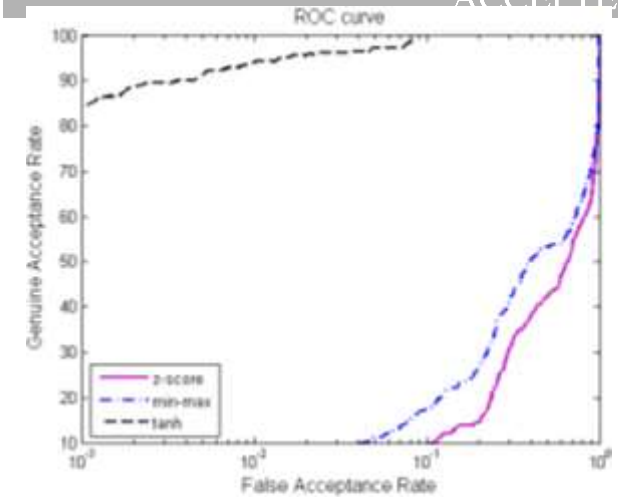

Fig. 5. ROCs for score-level combination of gestures with max rule and different normalization techniques

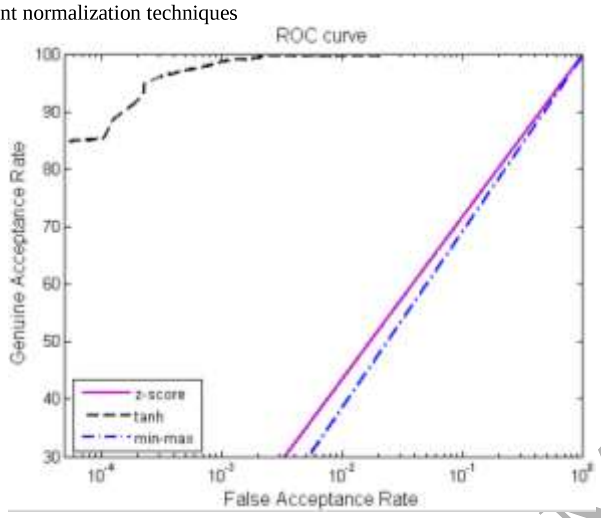

Fig. 6. ROCs for score-level combination of gestures with product rule and different normalization techniques

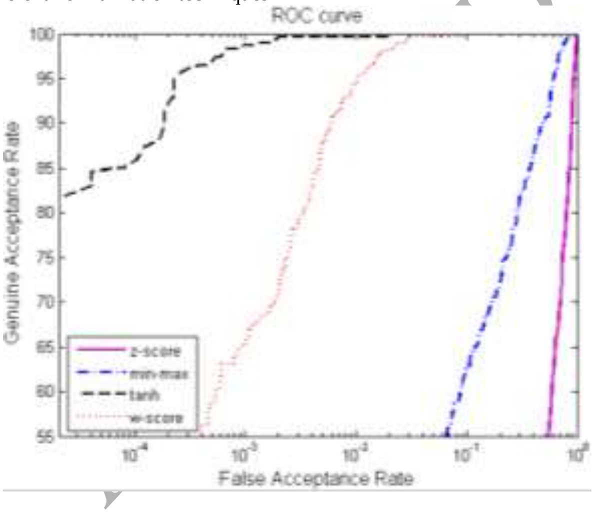

Fig. 7. ROCs for score-level combination of gestures with sum rule and different normalization techniques

4.3. Performance of score normalization techniques with sum rule for fusion

In this section, sum rule is employed for fusion of scores and performance (in terms of EERs) of different score normalization techniques are investigated for matching individual gestures. Here, matching of gestures is performed by considering the entire feature set. A set of genuine and impostor scores generated using the MHD based matching of corresponding features are normalized using min-max, z-score, w-score and tanh-estimator techniques. These normalized scores are then combined using sum rule. Experimental results are presented in Table 2, which shows EERs of individual gestures and their combination. It may also be noted from Tables 1 and 2 that w-score normalization scheme performs significantly better than z-score. A similar trend has been observed in [45] for recognition. However, more important observation is that tanh-estimator scheme for score normalization clearly outperforms min-max, z-score and w-score techniques. This may be due to fact that min-max technique uses minimum and maximum value, while the z-score method employs mean and standard deviation of matching scores; hence both these techniques are quite sensitive to outliers in the matching scores and cause performance degradation. This observation is also consistent with the results presented in [42] for sum rule based fusion. Superior performance of tanh-estimator on our dataset as compared to other schemes considered in this study is probably due to its robustness to noisy data [42, 45].

Table 2. EERs (%) for gesture matching with sum rule based fusion

Gesture	min-max	z-score	w-score	tanh-estimator
L2R	36.53	48.98	12.03	**3.42**
R2L	20.96	48.37	9.44	**3.69**
SD	32.03	47.73	13.44	**4.68**
SU	33.27	48.39	11.83	**3.06**
ST	25.86	46.14	9.96	**7.47**
ZI	23.43	45.46	10.83	**3.10**
ZO	24.26	45.83	7.02	**1.61**
Combined	24.12	50.64	1.92	**0.31**

4.4. Performance of fusion rules with tanh-estimator for score normalization

In this set of experiments, matching scores are normalized using tanh-estimator technique and the performance of four commonly used score-level fusion rules is evaluated. Table 3 presents EERs achieved while matching individual gestures with this experimental setting. This table also presents EERs for score-level fusion of all gestures. The corresponding ROCs for fusion of all gestures are depicted in Fig. 8. It can be observed from the table that sum and product rules clearly outperform (with significant reduction in EERs) min and max rules for score-level fusion. Interestingly, EERs achieved with the product rule are quite comparable to that of the sum rule. However, a closer observation of the ROCs (in Fig. 8) of two of these score-level fusion techniques reveals that sum rule based fusion yields consistently higher genuine acceptance rates (GAR) for the same range of false acceptance rates (FAR). This experimental observation shows that sum rule performs better than other fusion schemes is consistent with the observation in [44], in which authors reported experimental results from extensive evaluation of different combination schemes.

Table 3. EERs (%) for gesture matching with tanh-estimator for score normalization

Gesture	Min	Max	Product	Sum
L2R	28.19	5.03	3.48	3.42
R2L	29.32	6.31	3.64	3.69
SD	37.74	7.51	4.68	4.68
SU	33.51	3.95	3.05	3.06
ST	27.99	8.40	7.54	7.47
ZI	45.34	4.71	3.23	3.10
ZO	24.39	2.32	1.60	1.61
Combined	27.80	3.40	0.32	0.31

For the rest of the experiments in this section, tanh-estimator and sum rule are employed for normalization and fusion of matching scores, respectively. This is based on our observation from the above set of experiments that these techniques for normalization and fusion of matching scores achieve better performance as compared to other techniques.

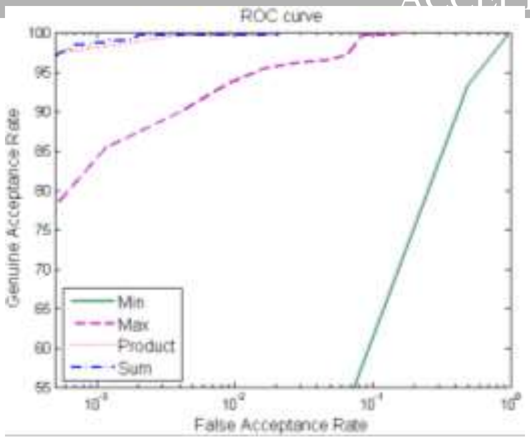

Fig. 8. ROCs for score-level combination of gestures with tanh-estimator for score normalization

As a comparative study, we have also implemented dynamic time warping (DTW) based matching, as it has been found very effective [31, 34, 36] for user authentication based on touchscreen gestures. In order to compare the performance of the proposed approach with DTW, we performed two sets of experiments in the verification mode. The objective of the first set of experiments was to evaluate the effectiveness of individual features, while in the second, we ascertained the matching performance of individual gestures and their combination.

In the first set of experiments, we considered one feature at a time for each of the gestures. In order to compute EER for each of the features, a set of genuine scores are generated by comparing a gesture with the rest of the corresponding gesture samples of the same user. This is repeated for all the gestures considered in this work. This experimental setting assumes that users perform all gestures during authentication. Multiple matching scores generated from matching of these gestures are normalized using tanh-estimator technique and then combined using sum rule to obtain a consolidated score. Since we have collected three samples of each gesture from 104 users, the total number of genuine scores generated is 312. Similarly, a set of impostor scores are generated by matching gestures of a user with the corresponding gestures of all other users in the dataset, resulting in 48204 impostor scores.

The results from this set of experiments are presented in Table 4. As it can be observed from this table, the orientation feature achieves EER of 0.56% and outperforms all other features considered in this study. Fig. 9 shows distribution of genuine and impostor matching scores for orientation feature. It can be noticed from the figure, there is only a little overlap between the genuine and impostor scores, which explains the promising results obtained for gesture matching using orientation feature.

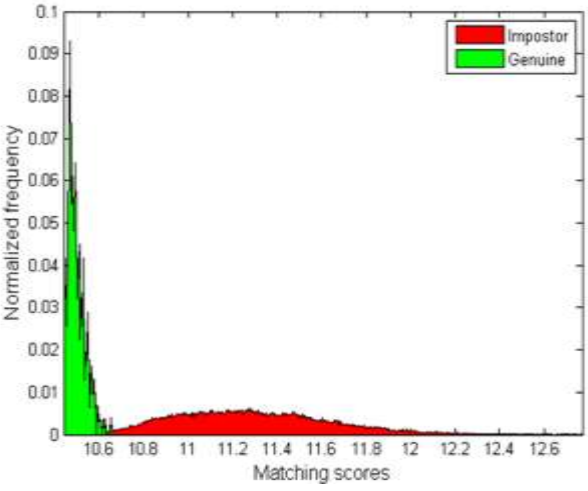

Fig. 9. Distribution of genuine and impostor scores for orientation feature

It is also important to note from Table 4 that the proposed approach outperforms the DTW based matching for three of the features namely, orientation sensor, x - y coordinates and the point curvature, with significant improvement in performance. For the rest of the features, the EERs of the two methods are quite comparable, with DTW approach achieving marginal improvement over our approach. Fig. 10 shows performance of the features individually in terms of the ROC for DTW and MHD based matching.

Table 4. EER (%) of individual features

Features	DTW	MHD
Accelerometer sensor	**2.91**	3.35
Orientation sensor	1.67	**0.56**
x-y coordinates	17.03	**11.24**
Finger area	**13.06**	14.12
Point curvature	27.36	**19.70**
Swipe curvature	**21.83**	22.70

In second set of experiments, firstly, we investigated the performance of proposed approach when user performs only one of the gestures considered in this study. Secondly, we also investigated the performance for score level combination of all gestures. In order to evaluate and compare the performance of the gestures, we considered one gesture at a time.

A set of genuine scores are then generated by matching a query gesture with the rest of the corresponding gesture samples of the same user. As described in Section 3.3, matching a pair of gestures involves matching of the corresponding features in the feature set. Multiple matching scores generated in the process are then combined using sum rule to get a consolidated score. In the same way, the impostor scores are obtained by comparing a query gesture with corresponding gestures of other users. The process is repeated for every gesture and corresponding EERs are calculated. Table 5 presents EERs obtained for gesture matching using the proposed and the DTW based method. It can be seen that ZO gesture provides the best authentication performance in our dataset. Also, it can be observed that performance of ST gesture is relatively poor. This is quite expected as the ST gesture, being a single point gesture, does not contain much information. More importantly, it may also be observed that our approach outperforms DTW based matching consistently for all gestures. To further investigate the combined performance of all gestures, the scores generated from matching of individual gestures are combined using the sum rule. Table 5 also shows EERs obtained for this experiment. The proposed approach achieves EER of 0.31%, a significant improvement (of 80%) over the DTW based method. In other words, the total number of falsely accepted impostor and falsely rejected genuine samples has reduced from 5 to 1. Fig. 11 shows comparison of ROCs for the proposed and the DTW based matching for score level fusion of all gestures.

Table 5. EER (%) of individual gestures and their combination

Gesture	DTW	MHD
L2R	6.22	**3.42**
R2L	4.87	**3.69**
SD	8.98	**4.68**
SU	3.86	**3.06**
ST	8.04	**7.47**
ZI	13.20	**3.10**
ZO	4.88	**1.61**
Combined	1.55	**0.31**

Table 6 shows average time taken to calculate matching scores for individual gestures using DTW and MHD based matching. It can be observed from this table that the proposed approach, as compared to DTW, requires significantly less time to perform

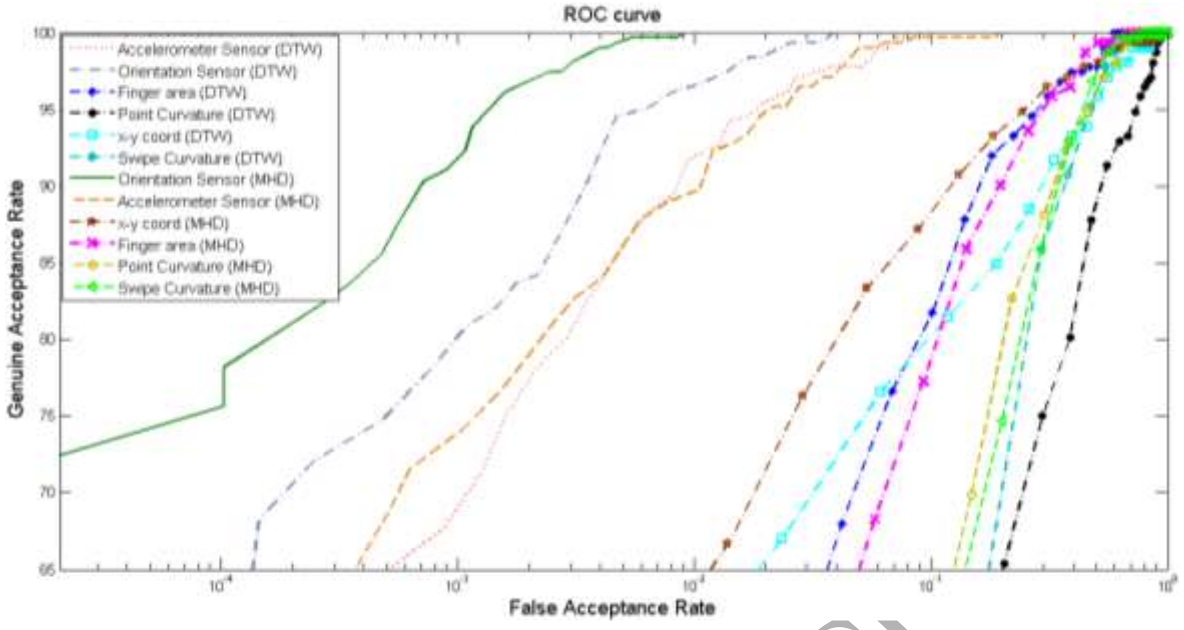

Fig. 10. ROCs of individual features with DTW and the proposed MHD based matching

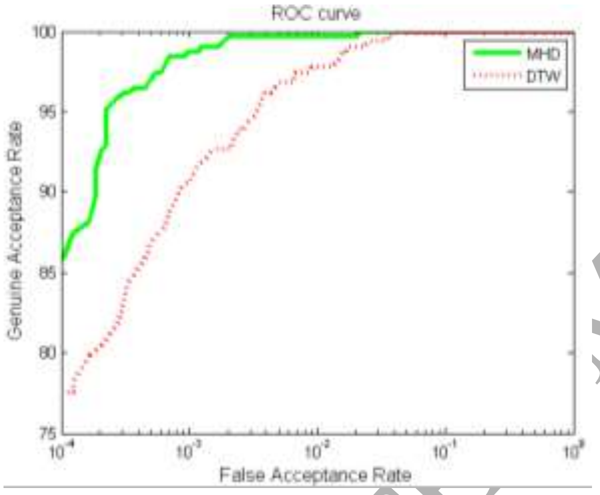

Fig. 11. ROCs for score level combination (sum rule) of all gestures

matching of gestures. It can also be noted that average time taken to match ST gestures is very less as compared to other gestures. It is because the ST gesture contains single point for matching. In addition, two of the features namely, the point curvature and swipe curvature features are not employed for matching ST gestures. Matching of ST gestures using MHD achieves highest performance improvement of 80% over DTW. For score level combination of all gestures, our approach achieves an improvement of 18.84% over DTW based matching. From these experimental results, it is evident that the proposed approach outperforms DTW based matching in terms of both error rate and the computational time.

Table 6 Time (ms) required for matching individual gestures

Gesture	DTW	MHD	Performance improvement (%)
L2R	8.66	5.96	31.18
R2L	9.24	6.52	29.44
SD	11.34	8.74	22.93
SU	11.50	8.79	23.56
ST	1.30	0.26	80
ZI	20.76	19.98	3.76
ZO	14.58	12.55	13.92
Combined	77.38	62.8	18.84

It may be noted that, in our experiments, we have not considered weighted combination (sum) of the matching scores as we do not have adequate data to train the system on an independent training subset to obtain optimal values for weights. We believe that if optimal weights are employed for combination of matching scores, it may be possible to further improve the performance of the system.

4.5. Performance of orientation, accelerometer versus the rest

In this section, the individual performance of orientation feature is compared with rest of the features. The ROC curves corresponding to the orientation feature and score-level combination of rest of the features using sum rule are shown in Fig. 12. In terms of EERs, user authentication based solely on orientation feature yielded 0.56% EER, while the sum rule combination of rest of the features yielded an EER of 1.2%.

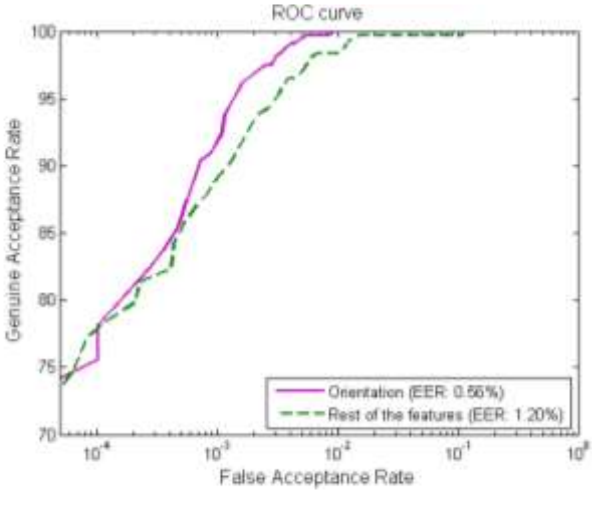

Fig. 12. Performance comparison of orientation feature with the sum rule based score level combination of rest of the features

Furthermore, we have performed experiments to compare performance of the proposed user authentication approach using combination of orientation and accelerometer features with the combination of rest of the features. Fig. 13 shows ROC curves from this set of experiments. It may be noted from this figure that the combination accelerometer and orientation features yielded EER of 0.64%, while the combination of rest of the features yielded EER of 6.26%, suggesting that combination of accelerometer and orientation features clearly outperform combination of rest of the features. The above experimental results indicate that the orientation and accelerometer sensor readings carry significant discriminatory information for user authentication in smartphones. It may be noted that the data acquired from built-in orientation sensor not only provides information on how the user holds the device, but also provides measurement of the continuous changes (on all three axes) in the orientation of the device while the user performs a gesture. The accelerometer sensor also operates in a similar manner providing linear acceleration of the device along the three axes. On the other hand, gesture features such as x-y coordinates, point curvature and swipe curvature rely mainly on the pattern (shape) of the gesture performed on the touchscreen. Intra-class variability of these gesture patterns causes erroneous matches resulting in higher error rates. This probably explains why these features carry limited discriminatory information as compared to orientation and accelerometer sensor reading.

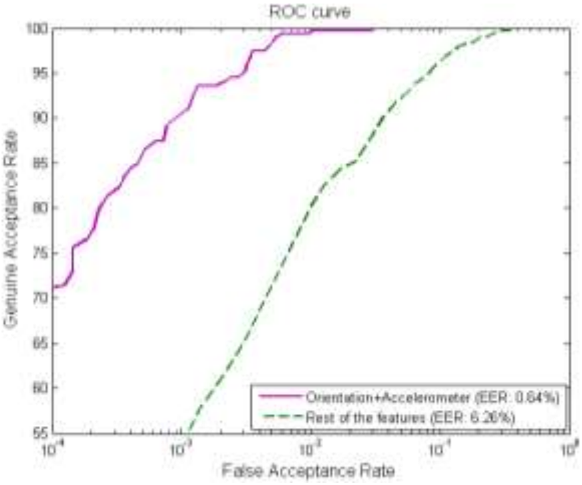

Fig. 13. Performance comparison of combination of orientation and accelerometer features with combination of rest of the features

4.6. Dataset-II

To investigate if the device used for data acquisition has any impact on the performance of the proposed authentication approach, we evaluated performance on a new dataset, which we refer to as Dataset-II. This dataset contains data from 30 subjects acquired using another smartphone-Samsung Galaxy Note-II N7100. Out of the 30, only two users were not having prior experience of operating touchscreen phones. Experimental setting for data acquisition remained the same as that of Dataset-I. The performance of the proposed authentication algorithm for different features on Dataset-II is summarized in Table 7. The key observation in Table 7 is that, although the individual EERs corresponding to different features have changed, their performance trend remains the same as in dataset-I. More importantly, it can be observed that the orientation sensor reading offers the best discrimination, followed by the accelerometer. It may be noted that a one-to-one comparison of EERs reported in Tables 4 and 7 cannot be made as these performance statistics are obtained on two different datasets with Dataset-II having considerable number (about 50%) of new users. The above results suggest that the device has no considerable impact on the performance of the proposed approach for mobile user authentication.

Table 7. EER (%) of individual features for dataset-II

Features	EER
Accelerometer sensor	4.55
Orientation sensor	0.99
x-y coordinates	7.63
Finger area	19.45
Point curvature	23.07
Swipe curvature	23.17
Combined	0.03

5. Conclusions

In this paper, we presented an approach for user authentication in smartphones based on behavioral biometrics. The android application that we have developed in this work runs on the smartphones and acquires a set of behavioral data when the user interacts with the device. The modified Hausdorff distance (MHD) is used for matching of features of the corresponding gestures, after identifying the category of the gesture performed by the user. Performance evaluation on a relatively large dataset (Dataset-I) of 104 users show that the proposed MHD based matching achieves better performance (in terms of both error rate and computational time) than the approach based on DTW. Our approach achieves the lowest EER of 0.31% when all the gestures performed by the users are combined at the score level for user authentication. Our experimental results also show that the information acquired from the built-in orientation and accelerometer sensor carry high discriminatory information for user authentication. This also indicates that the way user performs gestures on mobile phones is unique to some extent and the information can be exploited for user authentication. The performance of the proposed algorithm is also ascertained on a dataset (Dataset-II) of 30 subjects captured using another smartphone. The experimental results show a performance trend similar to the one on the Dataset-I. It can be inferred that changing the device does not affect the performance of the proposed algorithm. The framework for user authentication developed in this work is very well suited for continuous authentication in smartphones.

As part of our future work, we plan to expand our dataset by acquiring more samples of gestures per user and also, by increasing the number of users in our database. This will also help us in evaluating the performance of our approach on a dataset collected over a period of time.

References

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

[2] L. C. F. Araujo, L. H. R. Sucupira, M. G. Lizarraga, L. L. Ling, J. Yabu-Uti, User authentication through typing biometrics features, IEEE Transactions on Signal Processing 53 (2) (2005) 851-855.

[3] A. A. E. Ahmed, I. Traore, A New Biometric Technology based on Mouse Dynamics, IEEE Transactions on Dependable and Secure Computing 4 (3) (2007) 165-179.

[4] C. Shen, Z. Cai, X. Guan, Y. Du, R. A. Maxion, User authentication through mouse dynamics, IEEE Transactions on Information Forensics and Security 8 (1) (2013) 16-30.

[5] A. A. Ahmed, I. Traore, Biometric recognition based on free-text keystroke dynamics, IEEE Transactions on Cybernetics 44 (4) (2014) 458-472.

[6] K. Venkataramani, S. Qidwai, B. VijayaKumar, Face authentication from cell phone camera images with illumination and temporal variations, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 35 (3) (2005) 411-418.

[7] P. Abeni, M. Baltatu, R. D. Alessandro, Implementing biometrics based authentication for mobile devices, In Proc. IEEE Global Telecommunications Conference (GlobeCom), 2006, pp. 1-5.

[8] S. Chen, A. Pande, P. Mohapatra, Sensor-assisted facial recognition: an enhanced biometric authentication system for smartphones, In Proc. MobiSys, 2014, pp. 109-122.

[9] M. Baloul, E. Cherrier, C. Rosenberger, Challenge-based speaker recognition for mobile authentication, In Proc. BIOSIG, 2012, pp. 1-7.

[10] M. Kunz, K. Kasper, H. Reininger, M. Mobius, J. Ohms, Continuous speaker verification in realtime, In Proc. International Conference of the Biometrics Special Interest Group, 2011, pp. 79-87.

[11] E. Miluzzo, C. Cornelius, A. Ramaswamy, T. Choudhury, Z. Liu, A. T. Campbell, Darwin phones: the evolution of sensing and inference on mobile phones, In Proc. MobiSys, 2010, pp. 5-20.

[12] M. O. Derawi, B. Yang, C. Busch, Fingerprint Recognition with Embedded Cameras on Mobile Phones, In Proc. of MobiSec, 2012, pp. 136-147.

[13] X. Chen, J. Tian, Q. Su, X. Yang, F. Y. Wang, A secured mobile phone based on embedded fingerprint recognition systems, In Proc. ISI, 2005, pp. 549-553.

[14] N. L. Clarke, S. M. Furnell, B. M. Lines, P. L. Reynolds, Keystroke dynamics on a mobile handset: A feasibility study, Information Management and Computer Security 11 (4) (2003) 161-166.

[15] P. Campisi, E. Maiorana, M. L. Bosco, A. Neri, User authentication using keystroke dynamics for cellular phones, IET Signal Processing 3 (4) (2009) 333-341.

[16] S. Hwang, S. Cho, S. Park, Keystroke dynamics-based authentication for mobile devices, Computers & Security 28 (1–2) (2009) 85-93.

[17] K. Cheng, A. Kumar, Contactless finger knuckle identification using smartphones. In Proc. of BIOSIG. 2012, pp. 1-6.

[18] K. R. Park, H. A. Park, B. J. Kang, E. C. Lee, D. S. Jeong, A Study on Iris Localization and Recognition on Mobile Phones, EURASIP Journal on Advances in Signal Processing, 2008, pp. 1-12.

[19] N. F. Moco, I. S. Tecnico, I. Telecomunicacoes, P. L. Correia, Smartphone-based palmprint recognition system, 21st International Conference on Telecommunications (ICT), 2014, pp. 457-461

[20] Y. Han, T. Tan, Z. Sun, Y. Hao, Embedded palmprint recognition system on mobile devices, In Proc. International Conference on Biometrics, 2007, pp. 1184-1193.

[21] M. Franzgrote, C. Borg, B. J. T. Ries, S. Bussemaker, X. Jiang, M. Fieseler, L. Zhang, Palmprint verification on mobile phones using accelerated competitive code, In Proc. ICHB, 2011, pp. 1-6.

[22] D. J. Kim, K. W. Chung, K. S. Hong, Person authentication using face, teeth and voice modalities for mobile device security, IEEE Transactions on Consumer Electronics 56 (4) (2010) 2678-2685.

[23] P. Tresadern, T. F. Cootes, N. Poh, P. Matejka, A. Hadid, C. Levy, C. McCool, S. Marcel, Mobile biometrics: combined face and voice verification for a mobile platform, IEEE Pervasive Computing 12 (1) (2013) 79-87.

[24] A. Hadid, J. Y. Heikkila, O. Silven, M. Pietikainen, Face and eye detection for person authentication in mobile phones, International Conference on Distributed Smart Cameras, 2007, pp. 101-108.

[25] D. J. Kim, K. S. Hong, Multimodal biometric authentication using teeth image and voice in mobile environment, IEEE Transactions on Consumer Electronics 54 (4) (2008) 1790-1797.

[26] E. Khoury, L. E. Shafey, C. McCool, M. Gunther, S. Marcel, Bi-modal biometric authentication on mobile phones in challenging conditions, Image and Vision Computing 32 (12) (2013) 1147-1160.

[27] S. H. Khan, M. A. Akbar, F. Shahzad, M. Farooq, Z. Khan, Secure biometric template generation for multi-factor authentication, Pattern Recognition 48 (2) (2015) 458-472.

[28] C. McCool, S. Marcel, A. Hadid, M. Pietikainen, P. Matejka, J. Cernocky, N. Poh, J. Kittler, A. Larcher, C. Levy, et al., Bi-modal person recognition on a mobile phone: using mobile phone data, IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2012, pp. 635-640.

[29] W. Meng, D. S. Wong, S. Furnell, J. Zhou, Surveying the development of biometric user authentication on mobile phones, IEEE Communications Surveys & Tutorials, 2014.

[30] H. Saevanee, P. Bhattarakosol, Authenticating user using keystroke dynamics and finger Pressure, 6th IEEE Consumer Communications and Networking Conference, 2009, pp. 1-2.

[31] A. D. Luca, A. Hang, F. Brudy, C. Lindner, H. Hussmann, Touch me once and I know it's you! implicit authentication based on touch screen patterns, ACM Conference on Human Factors in Computing Systems, 2012, pp. 987-996.

[32] M. Frank, R. Biedert, E. Ma, I. Martinovic, D. Song, Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication, IEEE Transactions on Information Forensics and Security 8 (1) (2013) 136-148.

[33] T. Feng, Z. Liu, K. A. Kwon, W. Shi, B. Carbunar, Y. Jiang, N. Nguyen, Continuous mobile authentication using touchscreen gestures, IEEE Conference on Technologies for Homeland Security, 2012, pp. 451-456.

[34] T. Feng, J. Yang, Z. Yan, E. M. Tapia, W. Shi, TIPS: context-aware implicit user identification using touch screen in uncontrolled environments, ACM HotMobile, 2014.

[35] Y. Meng, D. S. Wong, R. Schlegel, L. Kwok, Touch gestures based biometric authentication scheme for touchscreen mobile phones, in: M. Kutyowski and M. Yung (Eds.), Springer Information Security and Cryptology, Springer Berlin Heidelberg, 2013, pp. 331-350.

[36] N. Sae-Bae, N. Memon, K. Isbister, K. Ahmed, Multitouch gesture-based authentication, IEEE Transactions on Information Forensics and Security 9 (4) (2014) 568-582.

[37] M. Antal, L. Z. Szabo, Z. Bokor, Identity information revealed from mobile touch gestures, 10th Joint Conference on Mathematics and Computer Science, 2014.

[38] X. Zhao, T. Feng, W. Shi, Continuous mobile authentication using a novel Graphic Touch Gesture Feature, IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2013, pp. 1-6.

[39] X. Zhao, T. Feng, W. Shi, I. Kakadiaris, Mobile user authentication using statistical touch dynamics images, IEEE Transactions on Information Forensics and Security 9 (11) (2014) 1780-1789.

[40] Package Index, 2014 [online]. Available: http://developer.android.com/reference/packages.html

[41] M. P. Dubuisson, A. K. Jain, A modified Hausdorff distance for object matching, Proceedings of the 12th International Conference on Pattern Recognition, 1994, pp. 566-568.

[42] A. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, Pattern Recognition 38 (12) (2005) 2270-2285.

[43] A. Ross, A. Jain, Information fusion in biometrics, Pattern Recognition Letters 24 (13) (2003) 2115–2125.

[44] J. Kittler, M. Hatef, R. P. Duin, J. G. Matas, On combining classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (3) (1998) 226–239.

[45] W. J. Scheirer, A. Rocha, R. Micheals, T. E. Boult. Robust fusion: Extreme value theory for recognition score normalization. In ECCV, 2010.