On-line Gesture Based User Authentication System Robust to Shoulder Surfing

Suman Bhoi, Debi Prosad Dogra and Partha Pratim Roy

Abstract People often prefer to preserve a lot of confidential information in different electronic devices such as laptops, desktops, tablets, etc. Access to these personalized devices are managed through well known and robust user authentication techniques. Therefore, designing authentication methodologies using various input modalities received much attention of the researchers of this domain. Since we access such personalized devices everywhere including crowded places such as offices, public places, meeting halls, etc., the risk of an imposter gaining one's identification information becomes highly feasible. One of the oldest but effective form of identity theft by observation is known as shoulder surfing. Patterns drawn by the authentic user on tablet surfaces or keys typed through keyboard can easily be recognized through shoulder surfing. Contact-less user interface devices such as Leap Motion controller can be used to mitigate some of the limitations of existing contact-based input methodologies. In this paper, we propose a robust user authentication technique that has been designed to counter the chances of getting one's identity stolen by shoulder surfing. Our results reveal that, the proposed methodology can be quite effective to design robust user authentication systems, especially for personalized electronic devices.

Keywords Contact-less authentication · Shoulder surfing · Personalized device authentication · Gesture recognition

S. Bhoi (✉) · D.P. Dogra
School of Electrical Sciences, Indian Institute of Technology, Bhubaneswar, India
e-mail: sb31@iitbbs.ac.in

D.P. Dogra
e-mail: dpdogra@iitbbs.ac.in

P.P. Roy
Department of Computer Science and Engineering,
Indian Institute of Technology, Roorkee, India
e-mail: proy.fcs@iitr.ac.in

B. Raman et al. (eds.), *Proceedings of International Conference on Computer Vision and Image Processing*, Advances in Intelligent Systems and Computing 460,
DOI 10.1007/978-981-10-2107-7_50

1 Introduction

The process by which a system recognizes a user or verifies the identity of a user trying to access it, is known as authentication. Installing a robust authentication technique that prevents impersonation, is of utmost importance for any personalized system since it plays a major role to defend against unauthorized access of the system. The procedure for establishing the identity of a user can be broadly branched into three categories [1]:

1. Proof by Knowledge—A user's identity can be authenticated with the help of information which is known only to the actual user. (e.g. Password)
2. Proof by Possession—Here the authentication is done with the help of an object specific to and in possession of the real user. (e.g. Smart Card)
3. Proof by Property—The user's identity is validated by measuring certain properties (e.g. Biometrics) and comparing these against the claimed user's original properties (e.g. Fingerprint)

Majority of the research in this domain mainly focuses on the proof by knowledge domain. Here, the validation is done with the use of password, PIN or pattern based techniques. These authentication schemes are mainly victim of shoulder surfing as shown in Fig. 1. It is a form of spying to gain knowledge of one's password or identity information. Here, the forger or imposter may observe or glance at the password, PIN or pattern being entered during authentication and may use it to impersonate a valid user. Extensive research is going on in this field to aid various applications such as authentication to prevent e-financial incidents [7], etc. Most of these applications use keystroke patterns [8] or biometrics [9, 11] or password entry [10] for authentication. The visual feedback provided by above mentioned techniques make them vulnerable to theft of identity. A possible solution to this may be to exploit the fact that the field of view of the valid user will be different as compared to the impersonator while shoulder surfing. Combining minimization of visual feedback with the above possible solution is likely to create a robust system resistant to user impersonation.

This paper proposes a novel authentication technique to avoid identity theft mainly caused by shoulder surfing. Here, we have used pattern based authentication technique without visual feedback (unlike pattern based authentication used in touch enabled devices) where Leap Motion device serves as the sensor to capture input signal. The device's interface has been used to create patterns with the help of on-air gestures. Leap Motion sensor[1] is a recent release by Leap Motion Inc. It can capture real-time movement of fingers and it can track precise movement of hand and fingers in three-dimensional space. It has a tracking accuracy of 0.01 millimetre. The device is currently being used for various gesture based applications like serious gaming [13], human computer interface, augmented reality, physical rehabilitation [12], etc. It is a low-cost device that is small in size. It supports a number of frameworks and is fairly accurate. These features of the device makes it a good choice as compared

[1] https://www.leapmotion.com/.

Fig. 1 An instance portraying authentication by a legitimate user while an imposter is applying shoulder surfing

to other similar devices such as Microsoft's Kinect or Intel's RealSense. For proper tracking of hand or fingers, a user should place his/her hand in the field of view of the device. Its range is about 150° with the distance constrained to less than a meter. The device comprises of a pair of infra-red cameras and three LEDs providing a frame rate varying from 20 to 200 fps. Information regarding the position of fingers, palm, or frame time-stamp can be obtained from each frame.

We have developed a methodology to use this for authentication on personalized devices. We start with partitioning the 2D screen or display into non-overlapping rectangular blocks and map it with the 3D field of view of the device. Assuming each of these blocks represent one character or symbol of the alphabet, users are asked to draw patterns on air. However, during the process, we do not provide any visual feedback to the user. Therefore, no cursor movement is seen on the screen. Then the task of recognising these patterns can be done by classifiers such as Hidden Markov Model (HMM) [5], Support Vector Machine (SVM), Conditional Random Field (CRF), etc. Here we have used HMM as the classifier due to its ability to model sequential dependencies and its robustness to intra-user variations. We train independent HMM for each distinct pattern in the training set and then a given sequence is verified against all trained model. The model having maximum likelihood is assumed to be the best choice.

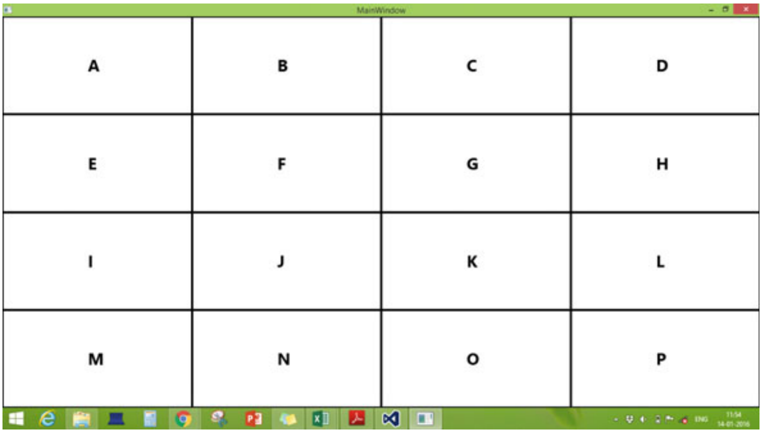
Rest of the paper is organized as follows. In Sect. 2, proposed methodology is presented. Results obtained using large set of samples collected in laboratory involving several volunteers, are presented in Sect. 3. We conclude in Sect. 4 by highlighting some of the possible future extensions of the present work.

2 Proposed Methodology of Authentication

This section describes about the signal acquisition, field of view mapping, training and testing the authentication methodology.

2.1 Device Mapping and Feature Extraction

First, we divide the whole screen or display into non-overlapping rectangular boxes and label each of those boxes. As an example, the screen can be arranged as a matrix of size 4 × 4 labelled "A" to "P" as depicted in Fig. 2. Using the finger and hand tracking utility of the Leap Motion device, we track movement of a user's index finger while performing the gesture during authentication. Initially, we provide a visual feedback to the user in the form of a visible cursor that helps the user to get an idea of the initial position of his/her finger with respect to the screen. Therefore, before drawing the authentication pattern, the user first executes a predefined gesture (e.g. circle gesture) that is used as a marker to start the authentication pattern and thereafter we hide the cursor. Therefore, the visual feedback is removed and the user draws the pattern through sense and anticipation. We have tested various patterns such as swipe, screen-tap, key-tap or circle to understand the best choice for the start marker. Circle gesture was found to be the most suitable and comfortable by the volunteers. Based on their feedback and the fact that the execution of the gesture should facilitate the knowledge of the finger position on screen before the cursor is hidden, circle gesture was used.



A	B	C	D
E	F	G	H
I	J	K	L
M	N	O	P

Fig. 2 Partitioning of the display into non-overlapping blocks and assignment of symbols or alphabets

Next, we present the method to map the field of view of the Leap Motion device to the display screen (e.g. computer screen). Since the display screen is rectangular in shape, therefore, instead of mapping the entire inverted pyramid interaction-space of the device (3D) to the 2D screen, we create an interaction box within the field of view to ease the movement and mapping of fingers. The height of interaction box can be set according to the user's preference of interaction height. Respective coordinate systems of display screen and Leap Motion are shown in Fig. 3. From the figure, it is evident that we need to flip the Y-axis of Leap Motion to map the coordinates correctly to the display screen. We normalize the real world position of the finger so that the coordinates lie between 0 and 1 and then translate the coordinates to the screen position as described in (1) and (2). This helps us to localize the position of the finger-tip on the display screen segment towards which the finger is pointing. We have not included Z-axis of the real-world position (with respect to the device) of the finger since we want to portray the movement on the 2-D screen.

$$X_s = (X_n)W_s \tag{1}$$

$$Y_s = (1 - Y_n)H_s \tag{2}$$

where,

X_s, Y_s represent X and Y coordinate of the finger position mapped on to the screen, respectively. X_n and Y_n represent the normalized X and Y coordinate of the finger-tip within the field of view of the device. W_s and H_s represent width and height of the screen.

Next, acquisition of authentication patterns with respect to the above mentioned mapping is described. Suppose, a user wants to draw a pattern "AEIJKL" as depicted in Fig. 4. The user needs to move his/her finger over the device's field of view to traverse the labelled boxes in the following order of sequence, A, E, I, J, K, L. To accom-

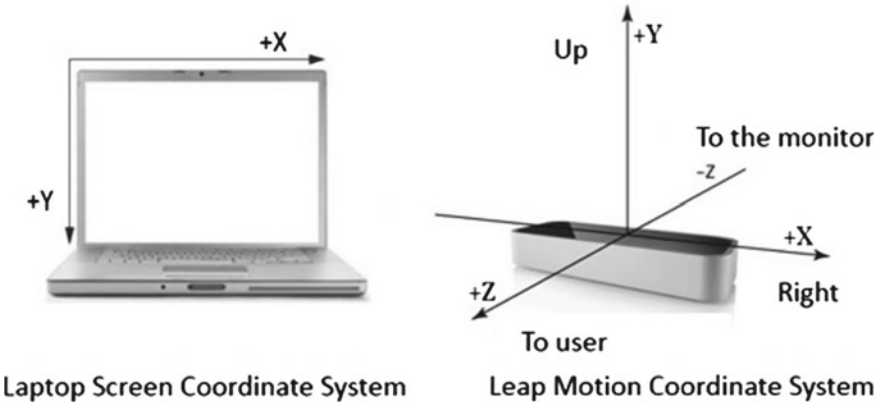

+X
+Y
Laptop Screen Coordinate System
Up
+Y
To the monitor
-Z
+X
Right
+Z
To user
Leap Motion Coordinate System

Fig. 3 Respective coordinate systems and possible mapping

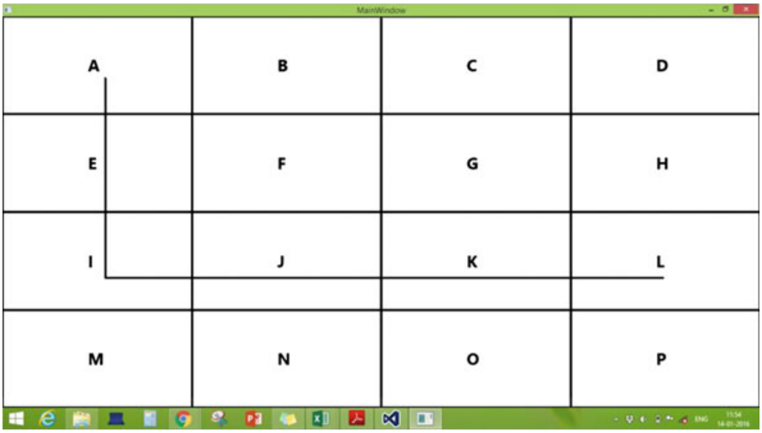


Fig. 4 A sample pattern "AEIJKL" drawn over the field of view of the device and its corresponding 2D mapping onto the screen

plish this, the user must bring his/her finger within the device's field of view and try to point box "A". After making a small circle gesture on box "A" (as described earlier), the user needs to traverse the other boxes in the above mentioned order. Although there is no visual feedback, position of the finger-tip in each frame is recorded. This information is used for generating the pattern. A pattern of such movement can be represented as follows,

$$p = (x_1, y_1), \ldots \ldots .., (x_k, y_k) \tag{3}$$

where, p represents the pattern under consideration, (x_k, y_k) represents the coordinate of the finger-tip with respect to the screen-space in the kth frame. Figure 5 depicts some of the patterns engaged in this experiment.

2.2 Training of Hidden Markov Model and Recognition

In this section, we present a methodology to implement the authentication protocol. We have applied Hidden Markov Model (HMM) based stochastic sequential classifier to train our system and classify test patterns. In our authentication scheme, users were asked to register their favourite patterns or secret sequence of symbols.

HMM is a preferred choice for such pattern classification tasks because of its ability to model sequential dependencies. An HMM can be defined by initial state probabilities π, state transition matrix $A = [a_{ij}]$, $i, j = 1, 2, \ldots, N$, where a_{ij} denotes the transition probability from state i to state j, and output probability $b_j(O)$ is modelled with discrete output probability distribution with S number of states. After several experiments, we find that, $S = 5$ provides optimum results. Vector quantization with 16 clusters has been used to discretize the input patterns or sequences. We perform

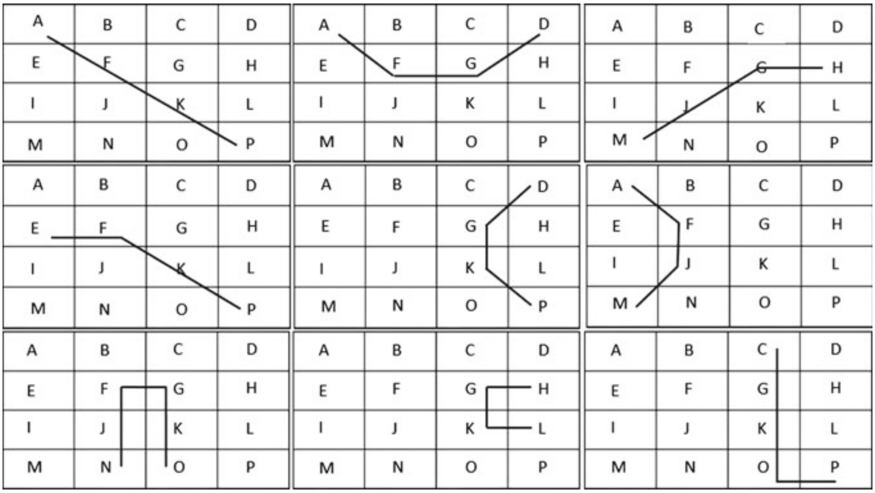


Fig. 5 Different test patterns involved in the study

the task of recognition using the Viterbi decoding algorithm [2–4]. We assume that the observation variable depends only on the present state. Therefore, a first order left to right Markov model has been presumed to be correct in the present context. The estimation of maximum likelihood parameters is carried out using Baum-Welch training algorithm. It uses EM technique for maximization of the likelihood where $\theta = (A, b_j, \pi)$ describes the Hidden Markov chain. The algorithm finds a local maximum of θ for a given set of observations (Y) as depicted in (4), where Y represents the observation sequence. More on the method can be found in Rabiner's pioneering documentation on HMM [5].

$$\theta^* = \max_{\theta} P(Y|\theta) \tag{4}$$

The parameter θ that maximizes the probability of the observation can be used to predict the state sequence for a given vector [6]. We compute the probability of observing a particular pattern ($p_j \in S$) using (5), where θ_i represents the parameters of the ith HMM that are learned through training and X denotes the hidden state sequence. Finally, given a test pattern we can classify it into one of the classes using (6) assuming there are C such distinct patterns in the dataset.

$$P(p_j, \theta_i) = \sum_{X} P(p_j|X, \theta_i)P(X, \theta_i) \tag{5}$$

$$\arg\max_{\theta_i} P(p_j, \theta_i) \;\; i = 1, 2, \ldots, 10 \tag{6}$$

Using the normalized coordinate vector representing all samples including training and testing patterns, the approach seems fairly robust to intra-user variations. In addition to that, since HMMs are scale-invariant in nature, the recognition process works fairly well regardless of the size of the coordinate vector. The procedure is summarized in Algorithm 1.

Algorithm 1 Recognition of 2D patterns using HMM

```
Input: $p \in P_{test}$ = Set of test sequences, $P_{train}$ = Set of training sequences, $C$ = Number of classes or distinct patterns, $S$ = Number of states, $O$ = Number of observation symbols.

Output: $c_j$(Class of $p$)where $c_j \in C$.
1: Create codebook using all training data ($P_{train}$).
2: Training: Vector quantize all training samples of a user with chosen feature set using the codebook.
3: Initialize $\pi, A, b_j \in B$, where $B$ is the observation matrix.
4: Train and fix the model for the present pattern or a user's sequence.
5: Repeat steps 2 to 4 for all $C$ patterns.
6: Recognition: Pass a test pattern ($p$) through all trained models and find the model with $\theta^*$.
7: Repeat step 6 for all the test samples $p \in P_{test}$.
8: Return $c_j$ for each test pattern.
```

3 Results

This section presents the results obtained during experiment involving 10 unbiased volunteers. To test the robustness of the proposed system, we have selected 10 varying authentication patterns (simple as well as complex patterns). Users were asked to mimic these patterns. Each volunteer was involved for the data acquisition phase where they were given a short demonstration to make them familiar with the Leap Motion device. A total of 1000 patterns were collected and 80 % of this data was used for training and remaining 20 % was used for testing.

A total of 10 models were created, one for each of the 10 unique patterns (essentially represent 10 distinct users). These models were trained (HMM) following the procedure described in Algorithm 1. Out of 1000 samples, 800 patterns were used for training and 200 patterns were used for testing. Confusion matrix of the classification is presented in Table 1. It is evident from the results that, accuracy is quite high for majority of these patterns except a few patterns. For example, it may be noted that, single instance of two of the patterns, namely 7 and 9, often getting confused with 9 and 6, respectively. 9 (“HGKL”) is being recognized as 6 (“DGKP”). This is due to the fact that, while the user was trying to draw “HGKL” pattern, he/she might have traversed the path representing “DGKP” as depicted in Fig. 6. Therefore,

Table 1 Confusion matrix depicting accuracy of test pattern recognition (authentication)

Pattern	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
P_1	20	0	0	0	0	0	0	0	0	0
P_2	0	20	0	0	0	0	0	0	0	0
P_3	0	0	20	0	0	0	0	0	0	0
P_4	0	0	0	20	0	0	0	0	0	0
P_5	0	0	0	0	20	0	0	0	0	0
P_6	0	0	0	0	0	20	0	0	0	0
P_7	0	0	0	0	0	0	19	0	1	0
P_8	0	0	0	0	0	0	0	20	0	0
P_9	0	0	0	0	0	1	0	0	19	0
P_{10}	0	0	0	0	0	0	0	0	0	20

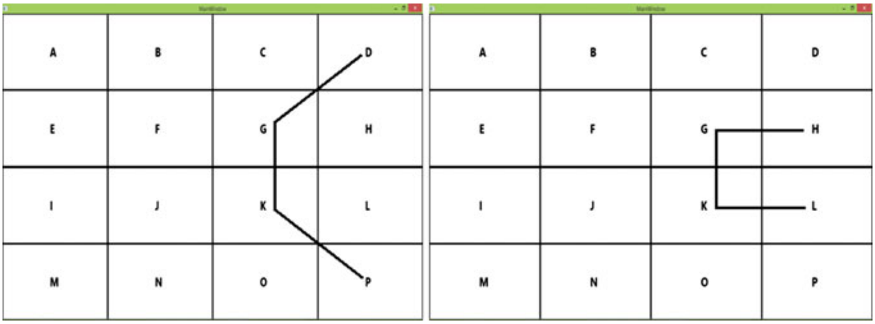


Fig. 6 Illustration of closely matching patterns “DGKP” and “HGKL”

unintentionally visiting nearby blocks during the gesture may cause failure in the logging-in procedure. However, our experiments reveal that only in two cases, we got a mismatch. Remaining cases were detected correctly with an overall accuracy of 99 %.

4 Conclusion

The paper proposes a novel technique for personalized device authentication via patterns without visual feedback. Here, we can conclude that, if the visual feedback is eliminated during authentication, the process becomes robust. However, existing touch-less or touch-based systems rely on visual feedbacks. On the contrary, the proposed Leap Motion based interface is robust against shoulder surfing attacks. This happens due to the difference in the field of views of the authentic user and the imposter.

The proposed system can be used for designing robust authentication schemes for personalized electronic devices. This will mitigate some of the limitations of existing contact-based or visual feedback based authentication mechanisms. However, the proposed system needs to be tested against real-imposter attacks and experiments need to be carried out to test its protection potential against such attacks.

References

1. Jansen, W.: Authenticating users on handheld devices. In: Proceedings of the Canadian Information Technology Security Symposium, pp. 1–12 (2003)
2. Iwai, Y., Shimizu, H., Yachida, M.: Real-time context-based gesture recognition using HMM and automaton. In: International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, pp. 127–134 (1999)
3. Rashid, O., Al-Hamadi, A., Michaelis, B.: A framework for the integration of gesture and posture recognition using HMM and SVM. In: IEEE International Conference on Intelligent Computing and Intelligent Systems, vol. 4, pp. 572–577 (2009)
4. Shrivastava, R.: A hidden Markov model based dynamic hand gesture recognition system using OpenCV. In: 3rd IEEE International Conference on Advance Computing, pp. 947–950 (2013)
5. Rabiner, L.: A tutorial on hidden Markov models and selected applications in speech recognition. In: Proceedings of the IEEE, vol. 77, no. 2, pp. 257–286 (1989)
6. Yamato, J., Ohya, J., Ishii, K.: Recognizing human action in time-sequential images using hidden Markov model. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 379–385 (1992)
7. Seo, H., Kang Kim, H.: User Input Pattern-Based Authentication Method to Prevent Mobile E-Financial Incidents. In: Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops (ISPAW), pp. 382–387 (2011)
8. Sheng, Y., Phoha, V. V., Rovnyak, S. M.: A parallel decision tree-based method for user authentication based on keystroke patterns. In: IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 35, no. 4, pp. 826–833 (2005)
9. Mengyu, Q., Suiyuan, Z., Sung, A. H., Qingzhong, L.: A Novel Touchscreen-Based Authentication Scheme Using Static and Dynamic Hand Biometrics. In: 39th Annual IEEE conference on Computer Software and Applications, vol. 2, pp. 494–503 (2015)
10. Syed, Z., Banerjee, S., Qi, C., Cukic, B.: Effects of User Habituation in Keystroke Dynamics on Password Security Policy. In: IEEE 13th International Symposium on High-Assurance Systems Engineering (HASE), pp. 352–359 (2011)
11. Frank, M., Biedert, R., Ma, E., Martinovic, I.; Song, D.: Touchalytics: On the Applicability of Touchscreen Input as a Behavioral Biometric for Continuous Authentication. In: IEEE Transactions on Information Forensics and Security, vol. 8, no. 1, pp. 136–148 (2013)
12. Vamsikrishna, K., Dogra, D. P., Desarkar, M. S.: Computer Vision Assisted Palm Rehabilitation With Supervised Learning. In: IEEE Transactions on Biomedical Engineering, DOI:10.1109/TBME.2015.2480881 (2015)
13. Rahman, M., Ahmed, M., Qamar, A., Hossain, D., Basalamah, S.: Modeling therapy rehabilitation sessions using non-invasive serious games. In: Proceedings of the IEEE International Symposium on Medical Measurements and Applications, pp. 1–4 (2014)