A Hand Gesture Approach to Biometrics

Nahumi Nugrahaningsih, Marco Porta(✉), and Giuseppe Scarpello

Dip. di Ingegneria Industriale e dell'Informazione,
Università di Pavia, Via Ferrata 5, 27100 Pavia, Italy
{nahumi.nugrahaningsih01,
giuseppe.scarpello01}@universitadipavia.it,
marco.porta@unipv.it

Abstract. In this paper we present a biometric technique based on hand gestures. By means of the Microsoft Kinect sensor, the user's hand is tracked while following a circle moving on the screen. Both 3D data about the position of the hand and 2D data about the position of the screen pointer are provided to different classifiers (SVM, Naive Bayes, Classification Tree, KNN, Random Forest and Neural Networks). Experiments carried out with 20 testers have demonstrated that the method is very promising for both identification and verification (with success rates above 90%), and can be a viable biometric solution, especially for soft biometric applications.

Keywords: Hand gesture biometrics · Soft biometrics · Vision-based biometrics · Microsoft Kinect

1 Introduction

Hand gestures, and the way they are performed, can be exploited for biometric identification and authentication. To date, relatively few biometric approaches exploit this principle (see Section2), in spite of its potentials.

In the general context of hand gesture recognition, an important distinction is that between *explicit* (or *control*) gestures — when their purpose is to provide some form of input to the computer, such as a command — and *implicit* gestures — when they are exploited to obtain indirect information about the user and his or her environment, such as activity recognition. A further distinction is that between *dynamic* gestures (associated with hand movements) and *static* gestures (characterized by specific hand positions and/or postures).

Two main approaches exist to hand gesture recognition, depending on the kind of device employed for data acquisition: techniques which require the user to move some physical object (e.g. an accelerometer) and vision-based methods in which hand shifts are detected by means on one or more cameras. The second case is more complex, but fortunately there are now sensors, such as Microsoft Kinect, which greatly simplify the tasks of hand recognition and tracking. Microsoft Kinect, in particular, can merge 2D with 3D data, thus allowing to work in the "RGBZ" (color + distance) space.

Using Microsoft Kinect, in this paper we propose a novel hand gesture based biometric approach in which the user has to "follow" a circle moving on the screen.

V. Murino et al. (Eds.): ICIAP 2015 Workshops, LNCS 9281, pp. 51–58, 2015.
DOI: 10.1007/978-3-319-23222-5_7

The obtained results are very encouraging (in particular for soft biometric applications), and suggest further research in this direction.

The article is organized as follows. In Section 2 we briefly present some previous works related to hand gestures used for biometric purposes. In Section 3 we describe our approach and the experimental setting. In Section 4 we illustrate the results of our tests. In Section 5, lastly, we draw some conclusions and provide hints for future research.

2 Related Work

In this section we provide a short overview of some relevant works in which dynamic hand gestures have been exploited for biometric purposes.

Okumura et al. [1] collected acceleration data from 22 participants, who were asked to shake an accelerometer. The data were then analyzed using the squared error of Euclidean distance, Error of Angle and DP-matching. Matsuo et al. [2], from the same research group, proposed a template update approach to overcome the long term stability problem from their previous study.

Zaharis et al. [3] used seven features from hand signature gestures obtained using the Nintendo WiiMote remote device to verify four participants. The features were the elapsed time of gesture completion, maximum and minimum acceleration values per axis per time segments, starting and ending sensor positions, and maximum and minimum overall acceleration per axis.

In [4] and [5], Liu et al. developed a verification system called *uWave*. uWave adopted eight predefined simple gestures from a Nokia research study [6], and applied the Dynamic Time Warping (DTW) algorithm on time series data (acquired through Nintendo WiiMote) of each one of the three axes.

Also Guna et al. [7] exploited the Nintendo WiiMote to implement an identification system based on three natural gestures, i.e. making a signature in the air, picking up the device, and shaking the device. The DTW algorithm was applied to the data obtained from 10 participants.

Tian et al. [8] developed an authentication system called *Kin-Write*. This study involved 18 participants. Participants were asked to perform a hand signature in the air. In this case a vision-based approach was used, exploiting the Microsoft Kinect device to capture hand movements, and the DTW algorithm was applied on six features, namely: hand position and position differences between frames, velocity, acceleration, slope angle, path angle, and log radius of curvature.

Other techniques (e.g. [9]) are based on measurements of the user's hand pose (i.e. static gestures) in hand sign language. There are also biometric systems not specifically based on hand movements but which use the Kinect as a source of data. For example, Lai et al. [10] focused on the body silhouette, while Wu et al. [11] and Ball et al. [12] exploited the body skeleton.

3 Methodology

3.1 Apparatus and Participants

Hand movements were captured using a *Kinect for Windows v1* sensor, and the stimuli were displayed on a 30'' monitor with 2048×1536 screen resolution. The sampling rate of the RGB camera was 30 fps. Twenty testers were involved as participants (8 females and 12 males), aged between 14 and 50.

3.2 Procedure and Stimuli

The principle of our biometric approach is as follows. Randomly, a circle appears (sequentially) in five predefined positions of the screen (the four corners and the center, as shown in Fig. 1), where it is displayed for 3.8 seconds. The task of the user is simply to move the hand so that the screen pointer is always over the circle: when the circle disappears from a position and (almost instantaneously) appears in another position, the user has to immediately move the cursor to the new location, as fast as he or she can. All 20 paths between all couples of positions are covered (Fig. 1, bottom), which means that, in total, the circle is displayed in each position four times.

In our experiments, participants stood at a distance of 150 cm from the monitor, over which the Kinect was placed at a height of 120 cm from the floor (Fig. 2). These distances were chosen after some trials to obtain adequate precision. The diameter of the circle was 350 pixels and its color blue (displayed on a white background). Also the display time of 3.8 seconds was selected after preliminary assessments aimed at finding a good compromise between task duration and significance of the acquired data.

Each participant repeated the test 20 times, with the interval between two tests varying between two hours and two weeks. In total, 400 data recordings were therefore collected.

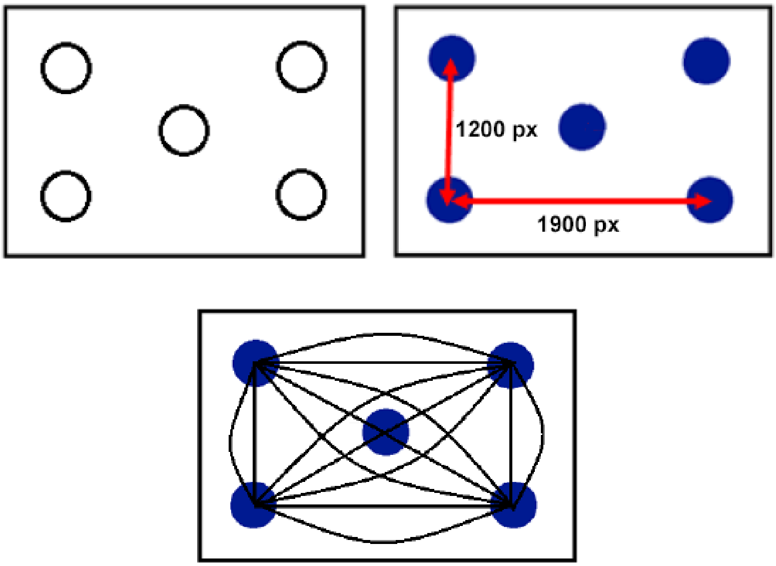

Fig. 1. Stimuli

Fig. 2. Experimental setting

3.3 Data Processing

This study used two kinds of raw data acquired by the Kinect: *hand position* in the 3D space and *pointer position* in the 2D screen space.

From these raw data, 11 features were derived, namely: 1) the average position of the hand (x, y and z coordinates) when the screen pointer is within a circle; 2) the average position of the pointer (x, y coordinates) within a circle; 3) the total time spent by the pointer inside a circle; 4) the user's "reaction time", i.e. the interval between the disappearance of the circle from a position and the time when the hand starts moving to shift the pointer to the new position; 5) the pointer's "travel time", i.e. the interval between the disappearance of the circle from a position and the instant when the pointer enters the circle in the new position; 6) the average speed of the hand (over the 20 paths) when moving the pointer from one position to another; 7) the average speed of the pointer (over the 20 paths) when it is shifted from one position to another; 8) the highest vertical position (y coordinate) reached by the hand; 9) the lowest vertical position reached by the hand; 10) the average vertical position of the hand; and 11) the variance of the pointer position when it is within a circle. After applying the Greedy Backward feature selection method [13], we excluded features 10 and 11, as they were not significant.

3.4 Classifiers

The first nine features described in the previous section became the input to six different classifiers, namely: K-Nearest Neighbors (KNN), Naive Bayes, Support Vector Machines (SVM), Classification Tree, Neural Network, and Random Forest.

Both random sampling and cross validation were used. With random sampling, we tried two different compositions of the training and test sets' data: 1) 70% training and 30% test, and 2) 50% training and 50% test. The final result was calculated from the average of 100 trials. In the cross validation method, we tried with 10- and 20-fold cross validation.

4 Results

In the classification process, the configurations described in Section 3.4 were applied to both the identification and verification cases.

4.1 Identification

As can be seen from Table 1, the accuracy of SVM was higher than that of the other classifiers in all sampling methods (and did not significantly decrease when the training dataset was reduced from 70% to 50%).

Table 1. Results of Identification

	Classification Accuracy			
	Random sampling 70% : 30%	Random sampling 50% : 50%	Cross validation 10 fold	Cross validation 20 fold
SVM	0.9452	0.9268	0.9475	0.955
Naive Bayes	0.8745	0.8697	0.8725	0.88
Classification Tree	0.7337	0.6976	0.7275	0.715
KNN	0.8321	0.8237	0.8425	0.85
Random Forest	0.9072	0.8922	0.9175	0.9275
Neural Network	0.9201	0.9006	0.92	0.925

4.2 Verification

In the verification scenario, we considered three performance metrics, namely: 1) sensitivity (true positive rate); 2) specificity (true negative rate); and 3) accuracy. From Tables 2 to 5, we can see that the classifiers' performance varied over the sampling methods. Considering as a reference the best performer in each metrics category, we can briefly summarize results as follows.

In random sampling with 70%:30% composition, SVM, Classification Tree, KNN, and Random Forest obtained (similar) good results with accuracy, sensitivity, and specificity. In random sampling with 70%:30% composition, Naive Bayes had generally a good performance, even though it was the lowest in terms of specificity.

In ten-fold cross validation, Random Forest and Neural Network exhibited very low values of specificity; conversely, the other three classifiers were overall comparable. In 20-fold cross validation, KNN showed the best global results, with 0.99, 1, and 0.9895 for accuracy, sensitivity, and specificity, respectively.

Table 2. Results of Verification: random sampling, 70% training and 30% test

	Accuracy	Sensitivity	Specificity
SVM	0.9968	0.9367	1
Naive Bayes	0.8823	1	0.8761
Classification Tree	0.997	0.98	0.9979
KNN	0.9987	0.9733	1
Random Forest	0.9987	0.9333	1
Neural Network	0.9935	0.87	1

Table 3. Results of Verification: random sampling, 50% training and 50% test

	Accuracy	Sensitivity	Specificity
SVM	0.9647	0.474	0.9905
Naive Bayes	0.8351	0.994	0.8267
Classification Tree	0.9434	0.41	0.9715
KNN	0.9566	0.628	0.9739
Random Forest	0.95	0.008	0.9996
Neural Network	0.9449	0.134	0.9876

Table 4. Results of Verification: cross validation, 10-fold

	Accuracy	Sensitivity	Specificity
SVM	0.975	0.9921	0.65
Naive Bayes	0.8275	0.8184	1
Classification Tree	0.97	0.9868	0.65
KNN	0.9625	0.9763	0.7
Random Forest	0.9525	1	0.05
Neural Network	0.9475	0.9842	0.25

Table 5. Results of Verification: cross validation, 20-fold

	Accuracy	Sensitivity	Specificity
SVM	0.995	0.9	1
Naive Bayes	0.8025	1	0.7921
Classification Tree	0.9475	0.35	0.9789
KNN	0.99	1	0.9895
Random Forest	0.97	0.55	0.9921
Neural Network	0.9875	0.75	1

5 Conclusion and Future Work

In this paper we have proposed a biometric approach based on hand gestures. The technique is mainly conceived for soft biometrics, and the encouraging results obtained are surely a spur to continue in this research direction.

In this preliminary study we have considered all 20 paths as the input of the classification system. In a real identification or verification scenario, however, performing so many gestures would probably be too demanding for the user. As a further study, we will therefore consider (random) subsets of the 20 paths, so as to reduce the user's effort and the duration of the procedure. In particular, we will focus on finding the shortest path length that guarantees a good performance.

A further interesting development could be the integration of the proposed verification/authentication method with other approaches (e.g. facial recognition or general body movements), to implement a multimodal biometric system. Moreover, to improve precision, the new Kinect sensor (version *v2*) could be tried, as well as the joint use of two Kinect devices in parallel.

References

1. Okumura, F., Kubota, A., Hatori, Y., Matsuo, K., Hashimoto, M., Koike, A.: A study on biometric authentication based on arm sweep action with acceleration sensor. In: International Symposium on Intelligent Signal Processing and Communications (ISPACS 2006), pp. 219–222 (2006)
2. Matsuo, K., Okumura, F., Hashimoto, M., Sakazawa, S., Hatori, Y.: Arm swing identification method with template update for long term stability. In: Lee, S.-W., Li, S.Z. (eds.) ICB 2007. LNCS, vol. 4642, pp. 211–221. Springer, Heidelberg (2007)
3. Zaharis, A., Martini, A., Kikiras, P., Stamoulis, G.: ``User authentication method and implementation using a three-axis accelerometer''. In: Chatzimisios, P., Verikoukis, C., Santamaría, I., Laddomada, M., Hoffmann, O. (eds.) MobiLight 2010. LNICST, vol. 45, pp. 192–202. Springer, Heidelberg (2010)
4. Liu, J., Wang, Z., Zhong, L., Wickramasuriya, J., Vasudevan, V.: uWave: accelerometer-based personalized gesture recognition and its applications. In: IEEE International Conference on Pervasive Computing and Communications, PerCom 2009, pp. 1–9 (2009)
5. Liu, J., Zhong, L., Wickramasuriya, J., Vasudevan, V.: User evaluation of lightweight user authentication with a single tri-axis accelerometer. In: 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, p. 15 (2009)
6. Kela, J., Korpipää, P., Mäntyjärvi, J., Kallio, S., Savino, G., Jozzo, L., Marca, D.: Accelerometer-based gesture control for a design environment. Pers. Ubiquitous Comput. **10**(5), 285–299 (2005)
7. Guna, J., Humar, I., Pogacnik, M.: Intuitive gesture based user identification system. In: 35th International Conference on Telecommunications and Signal Processing (TSP), pp. 629–633 (2012)
8. Tian, J., Qu, C., Xu, W., Wang, S.: KinWrite: handwriting-based authentication using kinect. In: 20th Annual Network & Distributed System Security Symposium (2013)
9. Fong, S., Zhuang, Y., Fister, I., Fister, I.: A biometric authentication model using hand gesture images. BioMedical Engineering OnLine **12**(1), 111 (2013)

10. Lai, K., Konrad, J., Ishwar, P.: Towards gesture-based user authentication. In: IEEE 9th International Conference on Advanced Video and Signal-Based Surveillance (AVSS), pp. 282–287 (2012)
11. Wu, J., Konrad, J., Ishwar, P.: Dynamic time warping for gesture-based user identification and authentication with kinect. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2371–2375 (2013)
12. Ball, A., Rye, D., Ramos, F., Velonaki, M: Unsupervised clustering of people from 'skeleton' data. In: 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 225–226 (2012)
13. Caruana, R., Freitag, D.: Greedy attribute selection. In: 11th International Conference on Machine Learning, pp. 26–36 (1994)