Hindawi Publishing Corporation
Discrete Dynamics in Nature and Society
Volume 2007, Article ID 70756, 8 pages
doi:10.1155/2007/70756

Research Article

Forgery Detection in Dynamic Signature Verification by Entailing Principal Component Analysis

Shohel Sayeed, S. Andrews, Rosli Besar, and Loo Chu Kiong

Received 21 July 2007; Accepted 6 October 2007

The critical analysis of the data glove-based signature identification and forgery detection system emphasizes the essentiality of noise-free signals for input. Lucid inputs are expected for the accuracy enhancement and performance. The raw signals that are captured using 14- and 5-electrode data gloves for this purpose have a noisy and voluminous nature. Reduction of electrodes may reduce the volume but it may also reduce the efficiency of the system. The principal component analysis (PCA) technique has been used for this purpose to condense the volume and enrich the operational data by noise reduction without affecting the efficiency. The advantage of increased discernment in between the original and forged signatures using 14-electrode glove over 5-electrode glove has been discussed here and proved by experiments with many subjects. Calculation of the sum of mean squares of Euclidean distance has been used to project the advantage of our proposed method. 3.1% and 7.5% of equal error rates for 14 and 5 channels further reiterate the effectiveness of this technique.

1. Introduction

The emerging trend insignature verification is utilizing the benefits of signal processing over the conventional image processing techniques. The benefits of involving signal processing for this purpose include reduction of equipment complexity, increased robustness against forgery, independence of media, and signature concealment.

In the conventional image processing method [1–3], a strong impression of the signature with prescribed ink on a stipulated medium is essential, and it is the key component for further processing. The signature image is then scanned into digital representation and this digital image is used for key construction. The process continues to search the

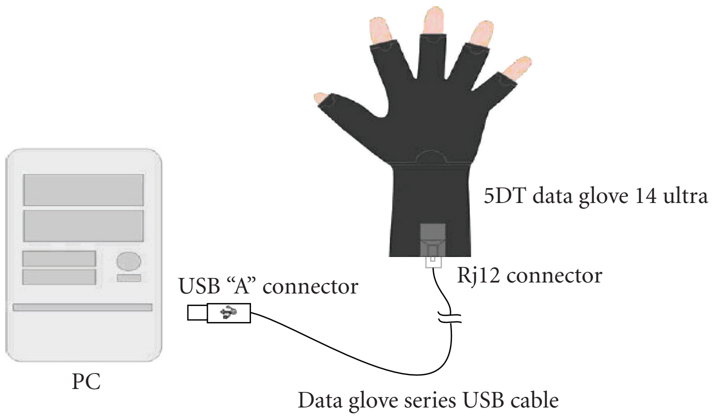

Figure 2.1. The data glove.

database for the nearest match and the result determines the authentication decision [4]. Since the signature represents the only key component here, it is perceptible and available for any misuse like forging. Moreover, the image processing involves volumes of data, and any reduction in volume needs a compromise with the efficiency of the system.

The continuous enhancement to this scenario leads to improving two major criteria. One is to protect the signature from public view and the second is to reduce the volume of data involved, and hence the speed of the process can be improved. Using a data glove is a paradigm shift from imaging to signal processing [5], and the first problem is immediately solved since there is no significance given to ink or paper.

However, the data glove used for this purpose,which the subject wears while signing for access, consists of many electrodes in various positions. These electrodes produce continuous signals during the signing process, resemble the image processing in data volume, and need to be reduced to overcome the drawback. PCA which is a popular technique [6] used for source separation has been introduced in this quandary case for both noise and volume reduction. The PCA is found to be competent in solving the problem by the improved results.

2. Methods

2.1. The data glove. We used a hand glove of 5DT Data Glove 14 Ultra model. Figure 2.1 shows the 14 fully enclosed fiber optic bend sensors spread twice per finger as well as abduction between fingers. The data glove interfaces with the computer via a cable to the platform independent USB port. This structure can be further simplified by interfacing with the computer wirelessly by means of Bluetooth technology with up to 20 m distance. This glove is made up of flexible material like lycra to fit to many hand sizes unreservedly.

The data captured using this glove is of 8-bit flexure resolution, at the sampling rate of minimum 75 Hz. The data glove is designed as a 3D input device, suitable for a broad range of applications like control and manipulation of virtual worlds, gesture and cognitive media, physiotherapy rehabilitation, control device for artists in remote controlled

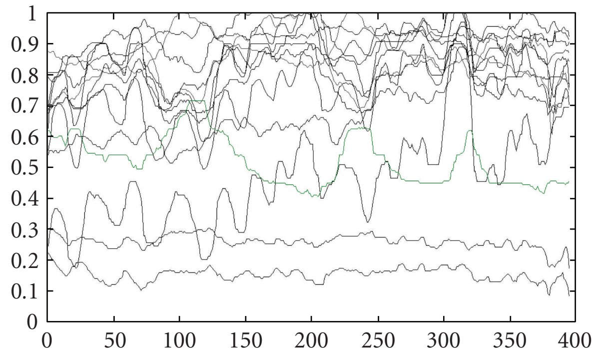

Figure 2.2. Signals from 14-electrode channels.

robotics, and so forth. The angle sensors have an 8-bit resolution, with 256 angular values per joint.

Each sensor is read by a 12-bit A/D converter, and it delivers a value throughout the range of 0–255 according to the joint's bending from open to closed position. The glove can be calibrated for the individual user differently. The five electrode gloves having the similar specifications except the number of electrodes are restricted to one per finger.

2.2. Signal capturing. We adopted a scheme of repeating all the data collections for both the 14-electrode and the 5-electrode combinations. The first category collects data from all the fourteen channels, and the second category consists of five channels, one per each finger. The recording is done along with the timing information in two ways. The coordinate value "x" of each sensor channel "n" at time "t" is recorded when the subject starts signing:

$$x(n)_t, \tag{2.1}$$

where $t = 0,\ldots,T, n = 1,\ldots,M$, and $M \subset \{14,5\}$. The total time "T" taken to complete one signature is also recorded. The data (see (2.1) per every signature is then stored as a data matrix x:

$$x = x_{(i,j)}, \tag{2.2}$$

where "i" represents the number of channels and "j" represents the number of data points captured per signature. The "j" varies from signature to signature along with "T." "T" and "j" may vary to inter- and intrapersonnel signatures. Hence, a straightforward comparison is difficult due to the variable j and T. Here, we overcome the problem by the adoption of PCA concepts to reduce not only noise but also the volume of data representing the entire signals.

2.3. Experimental settings. We collected three different versions of data, namely, reference, original, and forgery for both gloves as detailed in Tables 2.1 and 2.2.

Table 2.1. Signature database organization for training.

Dataset	Subjects	Signature/subject	Total signatures
Reference	10	10	100
Original	10	5	50
Forgery	10	25	250

Table 2.2. Signature database organization for test.

Dataset	Subjects	Signature/subject	Total signatures
Reference	30	10	300
Original	30	5	150
Forgery	50	150	7500

The subjects involved in this experiment are with an average age of 37.4. We collected 30 average reference signatures for test data from thirty subjects signing ten times each to derive the average. The original data is collected from the same thirty subjects signing their own signatures five times against their respective reference signatures. Forged data is collected from fifty subjects trying to forge the 30 average reference signatures for five times each. The subjects from the forging group are allowed to familiarize and practice the target signatures with an unlimited number of trials for forging. The data is recorded once they are confident about forging the authentic reference signatures. The test data includes the training data as a subset of the whole data collected.

In each signature data sample $x(i, j)$ if the value of "j" is high, the dimension of the matrix also becomes larger, which increases the computing load and processing time [7]. In fact, all the data points from a signature are not necessary to identify the key components involved. But at the same time, it is difficult to declare where these key components are present during the whole signing process.

Another important problem in this setup is the noise factor. The raw signal from the electrode x is a mixture of pure signals "S" and the background noise "N" is present in it due to minor extra movements and variance in equipment stability:

$$x = \mathrm{S} + N. \tag{2.3}$$

These noises vary from time to time, from place to place, and from equipment to equipment, which may affect the genuineness of the signal, and in turn the system.

2.4. Noise removal using PCA. Filtering of raw signals captured fromthe data glove may remove the noise contamination. But most of the time it also removes the significant signal components that degrade the original signal.

This problem can be rectified by introducing a suitable method for noise removal, like principal component analysis on the raw signals [8]. In PCA, the contaminated signal X was then normalized to zero mean and unit variance:

$$x = \frac{x - \mu(x)}{\sum(x)}, \tag{2.4}$$

where μ = mean (x) and $\sum$ = Std(x).

The PCA is used to extract perfect hand movement signals S from the noisy signal X. The correlation coefficient of the signal X was computed using

$$R = E(xx^T), \tag{2.5}$$

where x^T is the transpose of x.

Let F be the orthogonal matrix of eigenvectors of R and D, and let it be the diagonal matrix of its eigenvalues, $D = \text{diag}\,(d_1, \ldots, d_n)$. Then, the principal components could be computed by

$$Y = F^T x^T. \tag{2.6}$$

The PCs with larger value and feeble value represent the original signals S and the background white noise, respectively [8]. The PCs are selected in such a way that the eigenvalues are greater than 1. The selections of only the signal representing PCs from all available PCs were carried out to reduce the volume of operational data by PCA. The range of PCs taken for operation is 1–12.

These selected PCs were then used as the key dataset to represent the voluminous signature data and stand for the Euclidian test in distinguishing the forged signature from the original one. The chosen PCs were considered as the key dataset for calculation and comparisons. The same PCs are used in the reconstruction of the noise-free data signal S to confirm the clarity. The rebuilding was done using

$$x = FF^T YY^T, \tag{2.7}$$

where FF and YY correspond to the selected eigenvectors and PCs.

2.5. Signature verification. The average reference signatures of subjects are matched with five individual authentic signatures of the same subjects, by means of calculating the Euclidian distance [1, 4]. Considering the same average reference, pursuiting the 250 forged signatures from the forging group per reference was also carried out. The Euclidian distance for every forged signature $F(f_1, f_2, f_3, \ldots, f_n)$ with thereference signature $R(r_1, r_2, r_3, \ldots, r_n)$ is calculated. Euclidean distance "d" is specified as

$$d = \sqrt{\sum_{i=1}^{l} |f_i - r_i|^2}. \tag{2.8}$$

The noise-free signals extracted using PCA reduce the chance of forged signals getting into acclamation by precisely distinguishing the originals with the distinct Euclidean distance.

TABLE 3.1. The comparative equal error rates.

Type of channel	Equal error rate (%)
14 channels	3.1
5 channels	7.5

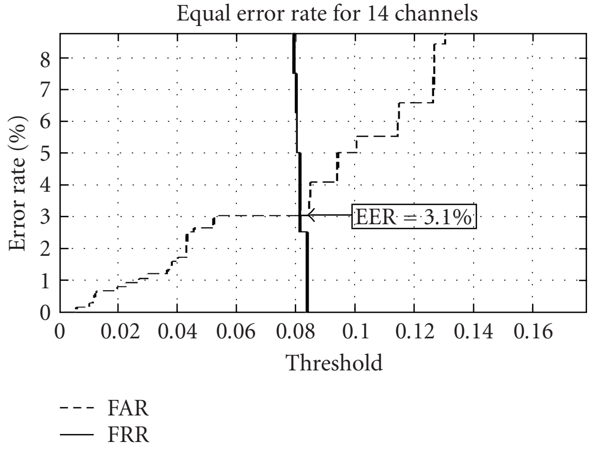

FIGURE 3.1. EER for 14-channel data.

3. Results and discussion

The average values of Euclidean distance between the mean reference signatures for the 150 authentic trials are calculated for every dataset from 14- and 5-sensor channels. Following the similar way, the average Euclidean distance for 50 forging subjects against the 30 authentic signatures is calculated. The number of forged signatures considered here is 7050.

Evaluation of results shows that riddance of forged signatures from the authentic signature can be easily identified using our PCA-based approach.

The false acceptance rate (FAR) and false rejection rate (FRR) are calculated for the normalized threshold values ranging from 0 to 1. FAR and FRR are calculated by

$$\mathrm{FAR} = \frac{\text{Total number of accepted forgeries}}{\text{Total number of tested forgeries}} \times 100, \tag{3.1}$$

$$\mathrm{FRR} = \frac{\text{Total number of rejected genuine}}{\text{Total number of tested genuine}} \times 100. \tag{3.2}$$

The comparative EERs for both 14 and 5 channels are shown in Table 3.1.

From the experimental results, we have achieved the EERs of about 3.1% and 7.5% for 14- and 5-electrode data, respectively, with the thresholds of 0.083 and 0.020. Figures 3.1 and 3.2 show the resulting EERs of 14- and 5-electrode data, respectively.

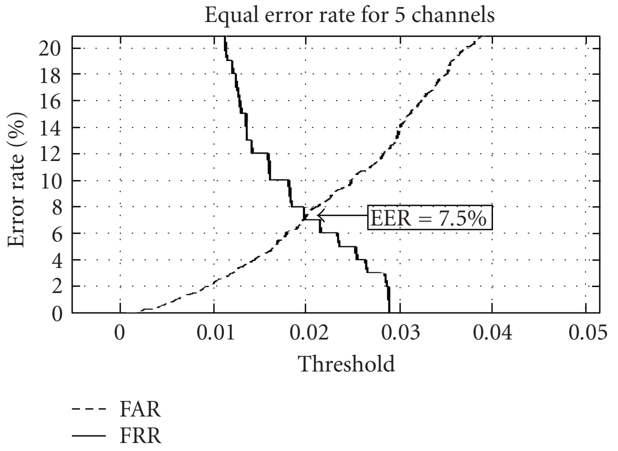

Figure 3.2. EER for 5-channel data.

4. Conclusion

Here, we demonstrated a new real-time technique for the easy recognition of handwritten signature. The technique is based on linearly projecting the signature space of data glove into a low-dimensional and noise-free space, through the use of PCA. The resulting projectionsmaximize the total scatter across all classes, that is, across all signals of all signatures, and result in a much simpler and efficient approach for signature recognition and verification. This work may be extended to further increase the credibility by involving an artificial neuralnetwork classifier.

References

[1] E. J. R. Justino, A. El.Yacoubi, A. F. Bortolozzi, and R. Sabourin, "An off-line signature verification system using hidden markov model and cross-validation," in *Proceedings of the IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing (NNSP '00)*, vol. 2, pp. 859–868, Sydney, Australia, December 2000.

[2] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63–84, 2000.

[3] V. S. Nalwa, "Automatic on-line signature verification," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 215–239, 1997.

[4] B. Fang and Y. Y. Tang, "Reduction of feature statistics estimation error for small training sample size in off-line signature verification," in *Proceedings of the 1st International Conference on Biometric Authentication (ICBA '04)*, vol. 3072 of *Lecture Notes in Computer Science*, pp. 526–532, Hong Kong, July 2004.

[5] S. Sayeed, R. Besar, and N. S. Kamel, "Dynamic signature verification using sensor based data glove," in *Proceedings of the 8th International Conference on Signal Processing*, vol. 3, pp. 2387–2390, Beijing, China, 2006.

[6] I. T. Jolliffe, *Principal Component Analysis*, Springer Series in Statistics, Springer, New York, NY, USA, 1986.

[7] D. L. Donoho and C Grimes, "Hessian eigen maps, new locally linear embedding techniques for high dimensional data," *National Academy of Sciences of the United States of America*, vol. 102, no. 21, pp. 7426–7431, 2005.

[8] S. Andrews, R. Palaniappan, and N. Kamel, "Single trial VEP source separation by selective eigen rate principal components," *Enformatika Journal*, vol. 7, pp. 330–333, 2005.

Shohel Sayeed: Faculty of Information Science and Technology, Multimedia University,
Melaka 75450, Malaysia
Email address: shohel.sayeed@mmu.edu.my

S. Andrews: Faculty of Information Science and Technology, Multimedia University,
Melaka 75450, Malaysia
Email address: andrews.samraj@mmu.edu.my

Rosli Besar: Faculty of Engineering and Technology, Multimedia University,
Melaka 75450, Malaysia
Email address: rosli@mmu.edu.my

Loo Chu Kiong: Faculty of Engineering and Technology, Multimedia University,
Melaka 75450, Malaysia
Email address: ckloo@mmu.edu.my