On-Line Signature Verification Using Most Discriminating Features and Fisher Linear Discriminant Analysis (FLD)

Muhammad Talal Ibrahim[1], Matthew Kyan[2] and Ling Guan[3]
Ryerson Multimedia Research Lab
Ryerson University, Toronto, Canada.
[1]muhammadtalal.ibrahi@ryerson.ca,
[2]mkyan@ee.ryerson.ca, [3]lguan@ee.ryerson.ca

Abstract

In this work, we employ a combination of strategies for partitioning and detecting abnormal fluctuations in the horizontal and vertical trajectories of an On-line generated signature profile. Alternative partitions of these spatial trajectories are generated by splitting each of the related angle, velocity and pressure profiles into two regions representing both high and low activity. The overall process can be thought of as one that exploits inter-feature dependencies by decomposing signature trajectories based upon angle, velocity and pressure - information quite characteristic to an individual's signature. In the verification phase, distances of each partitioned trajectory of a test signature are calculated against a similarly partitioned template trajectory for a known signer. Finally, these distances become inputs to Fisher's Linear Discriminant Analysis (FLD). Experimental results demonstrate the superiority of our approach in On-line signature verification in comparison with other techniques.

1. Introduction

Automatic Signature verification can be classified into two categories: On-line signature verification and Off-line signature verification. An Off-line signature verification system uses the camera or the scanner to capture the image of the signature and then uses the shape of signature for verification. The shape-related features include $x - y$ coordinates, height-to-width ratio, direction histogram, and curvature [7]. The main application of Off-line signature verification systems is in automatic verification of signatures found on bank cheques and documents.

An On-line signature verification system uses a digitizer for data acquisition purpose that generates the dynamic features such as velocity, acceleration and pressure [14], [10], [12]. Moreover total signature time, RMS speed, instantaneous velocity at sampling points are also considered as dynamic features [1]. It is always harder for the forger to imitate both the dynamic features and the shape of a signature than just imitating the shape alone [10], [5]. Our proposed system falls into the On-line signature verification system.

In this paper, we provide a novel approach which decomposes the shape ($horizontal$ and $vertical\ trajectories$) of a signature on the basis of angle, pressure and velocity profiles, and uses horizontal and vertical trajectories as separate features during verification. Angle, Velocity and Pressure profiles are partitioned into low and high regions, and underlying horizontal and vertical trajectories are extracted on the basis of these partitions. This process works to exploit a quite strong, inherent inter-feature dependance between the shapes of very specific regions of an individual's authentic signature, and the subtle velocities, angles and pressure changes the genuine signer has used to formulate that signature. By splitting into regions of high and low activity, we boost the sensitivity of our verification system as it attempts to flag cases in which there is excessive deviation in the "style" with which a $forgery$ has been formed, regardless of how similar the final shape might appear to be to the genuine signature.

Variations are evaluated on a per-partition basis, by assessing the distance between trajectories for each partition against a template for the original signer. The template itself is generally constructed from only 3 and 5 sample signatures, and the distances of each test signature's partitioned trajectories from the corresponding partitioned template trajectories are used as a set of inputs to Fisher's Linear Discriminant Analysis (FLD). FLD then attempts to establish the best separation between the test and the template forming genuine signatures. If no good separation is established, the signature is automatically deemed a forgery.

The paper is organized as follows. The second section deals with the acquisition of signature data and preprocessing steps. The third section is dedicated to the design and structure of our proposed system and finally in the last section, experimental results and concluding remarks are presented.

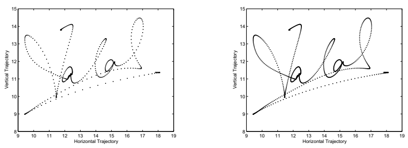
2. Data Acquisition And Preprocessing

For the experimental evaluation of our proposed system, we have used a subset of database used in [5], [4], [6] containing 25 genuine signatures and 25 skilled forgeries for 25 different signers.

2.1. Preprocessing

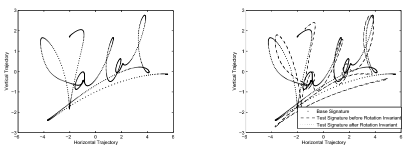
Every time a signer signs, he signs it in a different way. Because of this, some factors like velocity variation, different scales, different rotations or different places within the data acquisition device have to be taken into account in order to get a representation of the signature independent from these factors. For this purpose, we have followed a number of preprocessing steps as suggested in [5] to have a signature database independent from the above mentioned factors.

978-0-7695-3454-1/08 $25.00 © 2008 IEEE
DOI 10.1109/ISM.2008.115



(a) (b)

Figure 1. Signature a) before smoothing, b) after smoothing.



(a) (b)

Figure 2. a) Base Signature, b) another genuine signature of the same signer. It is obvious from the (b) that after rotation invariance eigen vectors of base-signature and test are at same angle and they are translation invariant.

2.1.1 Smoothing using Cubic Spline

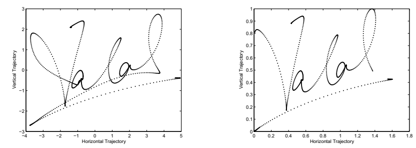
Horizontal and vertical trajectory of each signature should be smoothed to avoid jagged trajectories which are produced due to low sampling rate of our data acquisition device i.e. 100 samples/sec. In order to avoid the jagged trajectories [3], we have used cubic spline which not only smooths out the trajectories independently but also provides us with another dynamic feature i.e velocity which is the first derivative of cubic spline [3], [5]. Figure 1 shows signature before and after smoothing.

2.1.2 Translation, Rotation and Scaling

Translation invariance is achieved by subtracting the mean of each independent trajectory (horizontal,vertical) of each jth signature of ith signer from its respective trajectory. Rotation invariance is achieved by then rotating the Principal Component (PC) of each signature to the angle of the PC of base signature for each signer. Result of translation and rotation invariance is shown in Figure 2 b). These PCs are calculated by using PCA which is not only the dimensionality reduction transform but also works for making images rotation invariant [2]. Now the need of making all the signatures of ith signer scale invariant is achieved by using the following pair of equations:

$$\begin{aligned} Ratio_j^i &= \frac{max(x_j^i)}{max(y_j^i)} \\ y_j^i &= \frac{y_j^i - min(y_j^i)}{max(y_j^i) - min(y_j^i)} \\ x_j^i &= \left(\frac{x_j^i - min(x_j^i)}{max(x_j^i) - min(x_j^i)}\right) \times Ratio_j^i \end{aligned} \tag{1}$$

After making all the signature scale invariant, the vertical trajectories of each signature will be scaled from 0 to one and the horizontal trajectories will be scaled from zero to the ratio between the un-scaled vertical and horizontal trajectories as shown in Figure 3 b).



(a) (b)

Figure 3. a) before scale invariance and zero pressure removal, b) after scale invariance zero pressure removal

2.1.3 Zero Pressure Removal

Now there is a need to remove spatial areas of a signature which have falsely became part of signature. Actually, these spatial areas are generated due to high sensitivity of the acquisition tablet, which is capable of capturing the data even when pen tip is close to its surface. These spatial areas correspond to sample points that are captured between the pen-up and pen-down. The spatial areas, velocity and pressure of signature corresponding to such regions are named as zero pressure regions. . By the experimental results, we selected a threshold value $zero_j^i$ based on the pressure profile of each signature of ith signer. The spatial areas, velocity and pressure corresponding to the pressure profile below this threshold value were considered as zero pressure region and hence were removed. Results of all the above mentioned preprocessing steps are shown in Fig. 3 b). Mathematically this threshold value can be calculated as

$$\begin{aligned} std_j^i &= \sqrt{\frac{1}{M}\sum_{m=1}^{M}(z_j^i(m) - \frac{1}{M}\sum_{m=1}^{M} z_j^i(m))^2}, \\ zero_j^i &= \frac{1}{M}\sum_{m=1}^{M} z_j^i(m) - std_j^i, \end{aligned} \tag{2}$$

where M is total number of samples in a pressure profile z_j^i of jth signature of ith signer.

2.1.4 Dynamic Time Warping

Dynamic Time Warping (DTW) is used to establish a point-to-point correspondence between the base-signature b^i and all the signatures including genuine and forgeries of the ith signer. In our proposed system, DTW is performed between the base velocity v_b^i and velocity profiles v_j^i of jth signature of signer i as described in [5]. The reason of choosing the velocity profile for doing DTW over pressure profile is because velocity profile is considered to be richer in detail than the pressure profile [10]. In order to make the length of warping equal to the length of base vector (v_b^i), the one-to-many relationship present in the warping has been eliminated as suggested in [5], [4]. It is achieved by discarding all the repeated values in the warping path of v_b^i. Now the corresponding indices are used to retrieve x_j^i, y_j^i and z_j^i for all signatures of ith signer.

3. Proposed System

Normally, each Biometric identification system can be decomposed into two major stages, one is $Training$ and other is $Verification$:

3.1. Training

For the training of our proposed system, we have only used 3 and 5 genuine signatures to verify the discriminating power

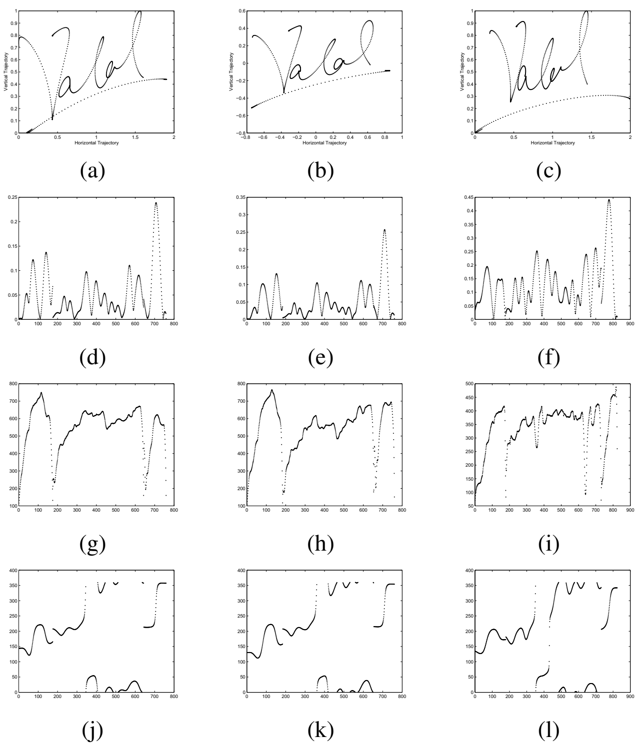


Figure 4. a), b) are shape, d), e) are velocity , g), h) are pressure profiles, j), k) are angle profiles of genuine signatures, whereas c), f), i) and l) are shape, velocity, pressure and angle profiles of forgery. It is obvious from the figure that decision on basis of only shape can be misleading, but a composite feature set containing shape along with dynamic features (velocity, pressure and angle) can lead us to more reliable decision.

of our system. Our system is based on the exploitation of inter-feature dependencies between the dynamic features of signature. We have only used three dynamic features i.e. velocity, pressure and angles. But, we have exploited the inter-feature dependencies between these feature by partitioning them into low and high.

3.1.1 Partitioning of Signatures

Feature extraction has its fundamental importance in biometric systems. In case of On-line signature verification system, features are mainly grouped into two classes, namely as Global features and Local features. Global features are easy to compute but have less discriminating power than the local features [7]. Normally an On-line signature verification system uses signature shape as a single feature during the verification process [10, 12]. But recently evolved techniques for signature verification are mostly based on exploitation of inter-feature dependencies [3, 5, 6, 4]. It is clear from Fig. 4 that decision on basis of only shape can be misleading, but a composite feature set containing shape along with dynamic features (velocity, pressure and angle) can lead us to more reliable decision. In our proposed system, we have used angle, velocity and pressure to form a higher dimensional discriminating feature set which facilitates the verification phase more than using velocity alone. To accomplish this task, we have partitioned the horizontal x_j^i and vertical y_j^i trajectories separately into two partition on the basis of angle and the same procedure was done on the basis of velocity and pressure profile as well. We have partitioned the horizontal x_j^i and vertical y_j^i trajectories of signer i by calculating the absolute angle of the horizontal trajectory x^i and vertical trajectory y^i of the base-signature as given below:

$$\begin{aligned} x_b^i &= x_b^i - \tfrac{1}{M}\textstyle\sum_{m=1}^{M} x_b^i(m) \\ y_b^i &= y_b^i - \tfrac{1}{M}\textstyle\sum_{m=1}^{M} y_b^i(m) \end{aligned} \tag{3}$$

$$ang_b^i = mod(arctan(y_b^i, x_b^i) + 2\pi, 2\pi) \tag{4}$$

where $x_b^i(m)$, $y_b^i(m)$ represents mth sample of horizontal trajectory x^i and vertical trajectory y^i of the base signature corresponding to ith signer respectively. After calculating the ang_b^i, the next step is to partition the ang_b^i into $low\ angles\ angl_b^i$ and $high\ angles$ using the following criteria as proposed in [4]

$$\begin{array}{ll} angl_b^i & \text{if } 0 \le ang_b^i \le \pi \\ angh_b^i & \text{otherwise.} \end{array} \tag{5}$$

After partitioning the ang_b^i into $angl_b^i$ and $angh_b^i$, next step is to segment each profile x_j^i and y_j^i into two partitions corresponding to the indices of $angl_b^i$ and $angh_b^i$.

We have also partitioned the x_j^i and y_j^i trajectories into low-velocity, high-velocity, low-pressure and high-pressure regions by using the following algorithm.

1. First step is to compute mean velocity mvb^i and mean pressure mzb^i for the base-signature by using the following equation,

$$mvb^i = \frac{1}{M}\sum_{m=1}^{M} vb^i(m), \tag{6}$$

$$mzb^i = \frac{1}{M}\sum_{m=1}^{M} zb^i(m), \tag{7}$$

   where M is total number of samples in v_b^i and z_b^i .

2. Velocities are declared as $low\ velocities\ vlb^i$ if and only if v_b^i is less than or equal to mvb^i. And velocities which are greater than low velocities are declared as $high\ velocities\ vhb^i$. Same is done to z_b^i for partitioning the pressure profile into $low\ pressure\ zlb^i$ and $high\ pressure\ zhb^i$. Then for each signer i, we have picked up the spatial areas (x_j^i, y_j^i) of signature corresponding to vlb^i, vhb^i, zlb^i and zhb^i respectively. Spatial areas corresponding to our proposed partitions of genuine and forgeries are shown in Fig. 5, Fig. 6 and Fig. 7. It is clear from the Fig. 5, Fig. 6 and Fig. 7, that our proposed partitioning method correspond to same spatial areas in case of genuine signatures while there is a significant difference in forgeries in case of angles, velocity and pressure.

After partitioning the ang_b^i, v_b^i and z_b^i into partitions, next step is to segment each profile x_j^i, y_j^i into two partitions correspondingly. Twelve partitions created after the partitioning process are named as:

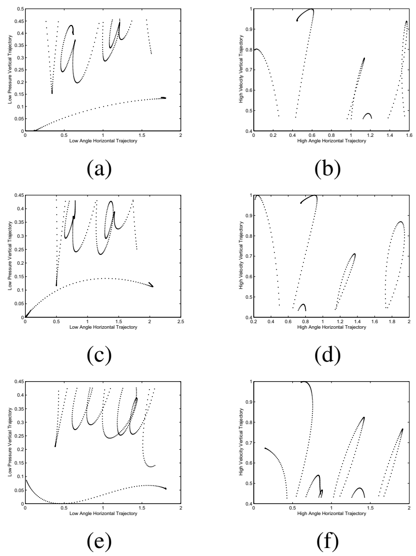


Figure 5. a),b),c) and d) are spatial areas of genuine signatures corresponding to low and high angles, e) and f) are spatial areas of forgeries. We see a lot of similarity among four genuine signature areas, whereas forgery signature areas are quite different from genuine.

1. Horizontal Trajectory:

   Low, High-Angle (xal^i_j, xah^i_j)

   Low, High-Velocity (xvl^i_j, xvh^i_j)

   Low, High-Pressure (xzl^i_j, xzh^i_j)

2. Vertical Trajectory:

   Low, High-Angle $(yal^i_j, yah^i_j$

   Low, High-Velocity (yvl^i_j, yvh^i_j)

   Low, High-Pressure (yzl^i_j, yzh^i_j)

After exploiting the inter-feature dependencies, next step is to generate the template of each partition by using only 3 and 5 genuine signatures by using the following equation.

$$tem^i_{abc} = \frac{1}{T}\sum_{t=1}^{T} abc^i_t, \tag{8}$$

where $a \in \{horizontal, vertical\}$, $b \in \{angle, velocity, pressure\}$, $c \in \{low, high\}$ and T represents total number of signatures used for training of ith signer. The reason for using horizontal and vertical trajectories in isolation is because of the radial nature of Euclidean Distance. The shape of an On-line signature is basically formed due to the wrist and fingers movements where the wrist movement is represented by the horizontal trajectory and the movement of the fingers is represented by vertical trajectory. As the On-line signature is formed due to the combination of two movements that are independent of each other so it will be more useful to use them as two separate discriminating features.

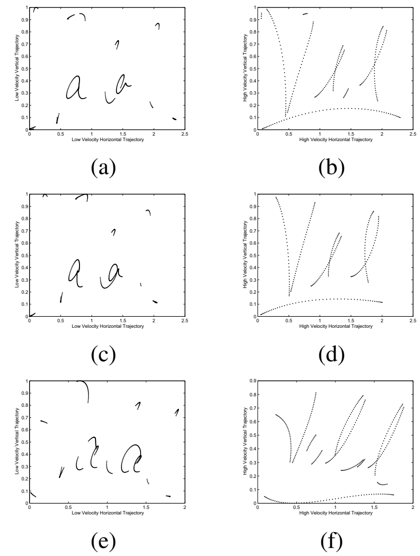


Figure 6. a),b),c) and d) are spatial areas of genuine signatures corresponding to low and high celocities, e) and f) are spatial areas of forgeries. We see a lot of similarity among four genuine signature areas, whereas forgery signature areas are quite different from genuine.

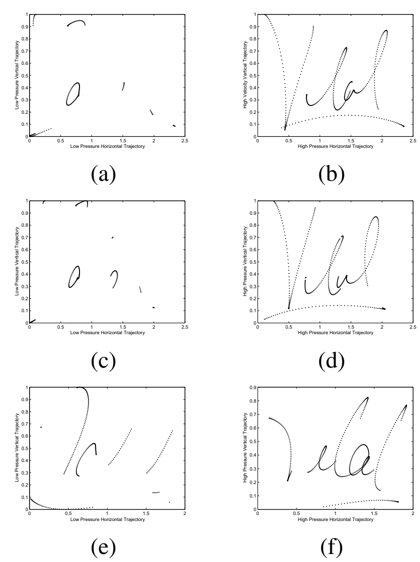
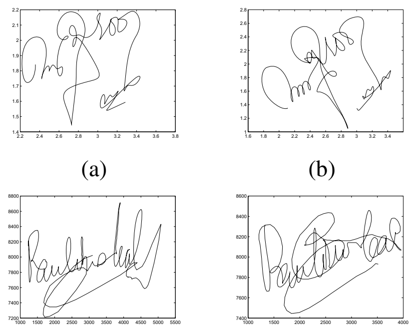


Figure 7. a),b),c) and d) are spatial areas of genuine signatures corresponding to low and high pressure, e) and f) are spatial areas of forgeries. We see a lot of similarity among four genuine signature areas, whereas forgery signature areas are quite different from genuine.



(a) (b) (c) (d)

Figure 8. Sample signatures, a) is a genuine signatures from database used in [5], [4], [6], whereas c) is a genuine signatures from MCYT database, b) is a forgery from database used in [5], [4], [6] and d) is a forgery from MCYT database. Here we see that skilled forgeries are quite closer in shape to their genuine counterparts. The need of dynamic features like angle, velocity and pressure are evident for proper discrimination.

3.2. Verification

In verification phase, a test signature of ith signer is declared as "genuine" or "forgery". For this purpose, first test signature ($test^i$) is passed through all necessary preprocessing steps, then its horizontal and vertical trajectories are decomposed into the above mentioned partitions and Euclidean distance is calculated between the each partitioned trajectory of test signature ($test^i$) with its corresponding template trajectory. In essence, FLD treats the test signature distances as a multi-dimensional feature representing an external class to that of the known authentic sample signatures (also multi-dimensional) used to generate the template. These are used to train FLD which works to maximize the between-class distance while minimizing within-class distance. A zero threshold is used to assess whether or not the test is likely to be genuine, yielding a simple yes/no response verifying the test signature. In other words, in the event of good separation found between the test and the template, the signature is automatically deemed a forgery. So, instead of choosing any stable feature based on standard deviation in the training phase [4], FLD helps to find the linear combination of features which best separate two classes. The resulting combination, used as a linear classifier can be considered a form of dimensionality reduction prior to classification. FLD helps us in reducing the dimensionality from 12 to only one while keeping the most of the variance.

4. Experimental Results and Conclusion

For the experimental results,we have used a database used in [5], [4], [6]. We have only used a subset of this database, containing only 25 genuine and 25 forgery signatures for each signer. We have trained our system using only 3 and 5 genuine signatures which has become a standard for training On-line Signature Verification System. We have also used one of the most famous public database MCYT-Signature Database [11]. This database contains 100 signer and for each signer 25 genuine and 25 forgery signatures are captured. All of the forgeries are highly skilled forgeries, as the static image of the signature was available to the forger and they were allowed to practice until they feel confident. These forgeries are made by 5 users for each signer. Fig. 8 shows some genuine signatures along with their respective skilled forgeries.

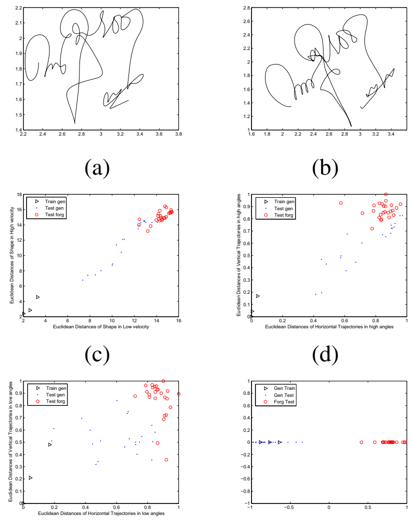
No. of Training Signatures	Avg. EER of [6]	Avg. EER of [5]	Avg. EER of our Proposed System
3	1.4000	0.8800	0.1875
5	0.9400	0.5800	0.0425

Figure 9. Avg. Equal Error Rates (EER) in percent for 25 signers belonging to our signature database.

No. of Training Signatures	Avg. EER of [6]	Avg. EER of [5]	Avg. EER of our Proposed System
3	8.680	6.800	0.200
5	7.200	5.600	0.080

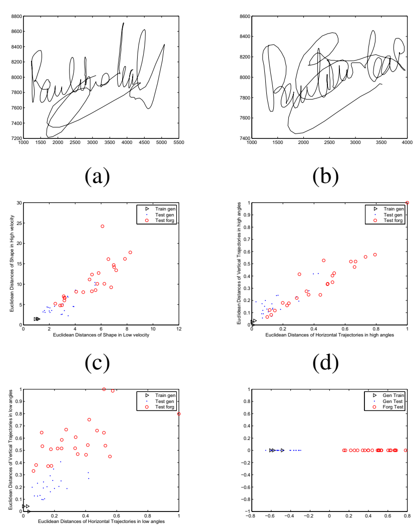
Figure 10. Avg. Equal Error Rates (EER) in percent for 100 signers belonging to MCYT signature database.

From the Fig. 11, Fig. 12, it is clear that our proposed system works better even if we are doing training only 3 genuine signature for each signer as shown in Fig. 9 and Fig. 10. This paper presents a novel approach to exploit the inter feature dependencies in a signature by decomposing the signatures into partitions based on angle, velocity and pressure profile of base-signature of ith signer and further using the partitioned independent trajectories i.e horizontal and vertical trajectories independently and finally giving the distance of each partitioned independent trajectory as an input to FLD, which seeks maximal separation between a test signature and template forming signatures. If achieved reasonably, then a forgery is detected. Here, we have not used the conventional majority rule for the decision fusion process, as the majority rule gives equal weights to all the decisions and does not account for correlation among features [8] and also we have not declared any partition as the most stable partition based on the standard deviation during the training phase [4]. The average Equal Error Rate (EER) is used to verify the validity of our proposed method. It is clear from the results that our proposed system performs better even if training is done by using only 3 and 5 genuine signatures. The results on the MCYT database is also satisfactory where we have achieved the EER of .0020 when only 3 genuine signatures were used for training and EER of .0008 when only 5 genuine signatures were used, where as average EER in [13] and [9] is 0.0120, 0.0375 with 5 training signatures respectively.



(a) (b) (c) (d) (e) (f)

Figure 11. a) genuine signatures from database used in [5], [4], [6], whereas b) is forgery, c), and d) represents the separation between genuine and forgery by using the method proposed in [4] when training was done by using only 3 genuine signature, e) represents the separation between genuine and forgery by using the method proposed in [5] when training was done by using only 3 genuine signature and finally f) represents the plot of our proposed system.



(a) (b) (c) (d) (e) (f)

Figure 12. a) genuine signatures from database used in [5], [4], [6], whereas b) is forgery, c), and d) represents the separation between genuine and forgery by using the method proposed in [4] when training was done by using only 3 genuine signature, e) represents the separation between genuine and forgery by using the method proposed in [5] when training was done by using only 3 genuine signature and finally f) represents the plot of our proposed system.

References

[1] G. Dimauro, S. Impedovo, M. G. Luccchese, R. Modugno, and G. Pirlo. Recent advancement in automatic signature verification. In *IEEE. 9th International Workshop on Frontiers in Handwriting Recognition*, pages 179–184, Oct. 2004.

[2] R. C. Gonzalez and R. E.Woods. *Digital Image Processing (2nd Edition)*. Pearson Education, Inc., Singapore, 2002.

[3] T. Hastie, E. Kishon, M. Clark, and J. Fan. A model for signature verification. In *IEEE. International Conference on Systems, Man and Cybernetics*, volume 1, pages 600–604, July 1991.

[4] M. T. Ibrahim, K. Saleem, M. A. Khan, and I. A. Taj. Creation and selection of most stable discriminating features for on-line signature verification. In *IEEE International Conference on Machine Vision*, pages 97–101, Dec. 2007.

[5] M. A. U. Khan, M. K. Khan, and M. A. Khan. Velocity-image model for online signature verification. *IEEE Trans. Image Processing*, 15(11):3540–3549, Nov. 2006.

[6] M. K. Khan, M. A. Khan, M. A. U. Khan, and I. Ahmad. On-line signature verification by exploiting inter-feature dependencies. In *18th International Conference on Pattern Recognition*, volume 2, pages 796–799, 2006.

[7] A. A. Kholmatov. Biometric identity verification using on-line and off-line signature verification. Master's thesis, Graduate School of Engineering and Natural Sciences, Sabanci University, 2003.

[8] L. L. Lee, T. Berger, and E. Aviczer. Reliable on-line human signature verification systems. *IEEE. Trans. on Pattern Analysis and Machine Intelligence*, 18:643–647, Jun. 1996.

[9] O. Miguel-Hurtado, L. Mengibar-Pozo, M. G. Lorenz, and J. Liu-Jimenez. On-line signature verification by dynamic time warping and gaussian mixture models. In *IEEE 41st Annual Carnahan Conference on Security Technology*, pages 23–29, Oct. 2007.

[10] W. Nelson and E. Kishon. Use of dynamic features for signature verification. In *IEEE International Conference on Systems, Man and Cybernetics*, pages 201–205, Oct. 1991.

[11] J. Ortega Garcia, J. Fierrez Aguilar, D. Simon, J. Gonzalez, M. Faundez Zanuy, V. Espinosa, A. Satue, I. Hernaez, J. Igarza, C. Vivaracho, D. Escudero, and Q. Moro. Mcyt baseline corpus: a bimodal biometric database. 150(6):395–401, December 2003.

[12] T. Qu, A. E. Saddik, and A. Adler. A stroke based algorithm for dynamic signature verification. In *Canadian Conference on Electrical and Computer Engineering*, volume 1, pages 461–464, May 2004.

[13] Z.-H. Quan and K.-H. Liu. Online signature verification based on the hybrid hmm/ann model. *International Journal of Computer Science and Network Security*, 7:313–322, March. 2007.

[14] J. Yi, C. Lee, and J. Kim. Online signature verification using temporal shift estimated by phase of gabor filter. *IEEE. Trans. on Signal Processing*, 53:776–783, Feb. 2005.