ijergs
Open Access
International Journal of Engineering Research and Generic Science (IJERGS)
Available Online at www.ijergs.in
ISSN: 2455 - 1597
Volume -3, Issue-1, January – February 2017, Page No. 14 - 25

Signature Verification Using Envelope and Histogram Feature Set

P.Asha[1], S.Swarna Latha[2]

[1]S.V University, Department of E.C.E, Tirupati, Chittoor District, A.P

E-Mail Id: polapalliasha@gmail.com

[2] Associate, S.V. University, Department of E.C.E, Tirupati, Chittoor District, A.P

E-Mail Id: swarnasvu09@email.com

Abstract

In this work an offline signature verification method is proposed which is based on two sets of features extracted from the signature image. First set of feature consists of envelope features which are obtained by scanning the signature from all the four sides and recording their curvature in a vector. Second set of feature will have histogram details of the signature image which is used to record the variations of gray level in the signature image which replicates the confidence levels of the signer. Before the features are extracted, the signature is subjected to different preprocessing stages which include extraction of exact size of the signature, resizing, binarization, erosion and box removal. Finally Euclidean distance is used as the classifier to decide whether the signature under test is genuine or forged. The parameters like False Acceptance Ratio (FAR), False Rejection Ratio (FRR), Equal Error Rate (EER) and Total Success Rate (TSR) were calculated and were compared with the previous works.

Keywords: Binarization, Euclidean, False Acceptance Ratio (FAR), False Rejection Ratio (FRR), Equal Error Rate (EER) and Total Success Rate (TSR)

1. Introduction

Biometric recognition, or simply biometrics, offers a natural and more reliable solution to the problem of person recognition. Since the biometric identifiers are inherent to an individual, it is more difficult to manipulate, share, or forget these traits. Hence, biometric traits constitute a strong and reasonably permanent link between a person and his identity. Formally, biometric recognition can be defined as the science of establishing the identity of an individual based on the physical and/or behavioral characteristics of the person either in a fully automated or a semi-automated manner.

The fundamental task in identity management is to establish the association between an individual and his personal identity. One must be able to determine a person's identity or verify the identity claim of an individual whenever required. This process is known as person recognition. A person can be recognized based on the following three basic methods

(a) What he knows,

(b) What he possesses extrinsically, and

(c) Who he is intrinsically.

2. Signature Recognition

Handwritten signature verification has been extensively studied & implemented. Its many applications include banking, credit card validation, security systems etc. In general, handwritten signature verification can be categorized into two kinds.

1. online verification and
2. Off–line verification.

A. Steps in Signature Recognition: Signature Recognition Systems need to preprocess the data. The major steps are as follows

1. Data acquisition: The signatures to be processed by the system should be in the digital image format. We need to scan the signatures from the document for the verification purpose.

2. Signature pre-processing: We have to normalize the signature, resize it to proper dimensions, remove the background noise, and thin the signature.

3. Feature extraction: We are using various feature extraction algorithms. The feature set includes the conventional global features of signature as well as new features. The new features include gray level histogram, Local Binary Patterns (LBP) and Wavelet coefficients etc.

4. Enrollment & training: The extracted features are stored in to database. We train the system using a training set of signature obtained from a person. Designing of a classifier is a separate area of research. The decision thresholds required for the classification are calculated by considering the variation of features among the training set. Separate set of thresholds (user Specific) is calculated for each person enrolled, some system also use common threshold form all users.

5. Performance evaluation: The performance of system depends on how accurately the system can classify between the genuine and fraud signatures. The forgeries involved in handwritten signatures have been categorized based on their characteristic features.

B.Levels of Forgeries: Various kinds of forgeries are

1. Random forgery: The signer uses the name of the victim in his own style to create a forgery known as the simple forgery or random forgery. This forgery accounts for the majority of the forgery cases although they are very easy to detect even by the naked eye.

2. Unskilled forgery: The signer imitates the signature in his own style without any knowledge of the spelling and does not have any prior experience. The imitation is preceded by observing the signature closely for a while.

3. Skilled forgery: Undoubtedly the most difficult of all forgeries is created by professional impostors or persons who have experience in copying the signature. For achieve this one could either trace or imitate the signature by hard way. Figure 1 shows the different types of forgeries and how much they are varies from original signature.

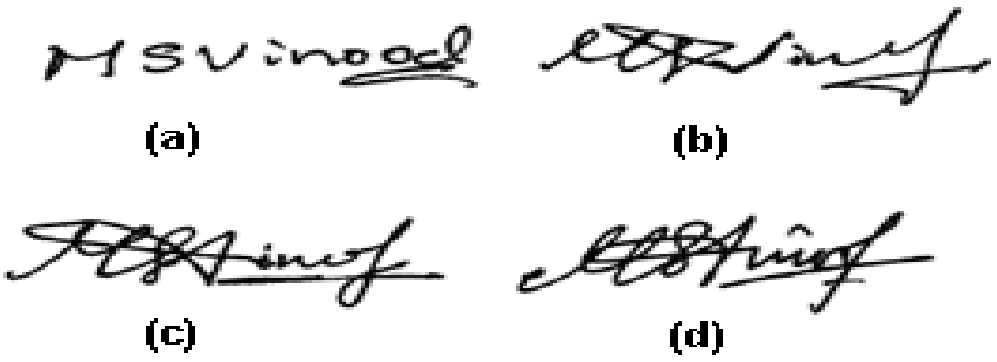

Figure 1: Different types of forgeries (a) Original Signature (b) Random forger (c) Unskilled forgery (d) Skilled forgery

3. Performance Parameters in Biometrics

1. False Rejection Rate (FRR): It is a measure of the biometric security system, that incorrectly reject an access attempt by an authorized user. A FRR is the ratio of the number of false rejections to the total number of identification attempts is given by equation 1.

$$FRR = \frac{\text{Number of Persons falsely rejected}}{\text{Total Number of Persons}} \quad \text{---------- (1)}$$

2. False Acceptance Rate (FAR): It is a measure of the biometric security system that incorrectly accepts an access attempt by an unauthorized user. A FAR is the ration of the number of false acceptance to the total number of identification attempts is given by equation 2.

$$FAR = \frac{\text{Number of Persons falsely accepted}}{\text{Total Number of Persons}} \quad \text{---------- (2)}$$

3. Total Success Rate (TSR): The number of test faces matched with the appropriate person accurately. It is the ratio of correctly matched persons to the total number of person in the database, and is given by equation 3.

$$TSR = \frac{\text{Number of Persons matched correctly}}{\text{Total Number of Persons}} \quad \text{---------- (3)}$$

4 Threshold/Decision Thresholds: The acceptance or rejection of a data is dependent on the match score falling above or below the threshold. The threshold is adjustable so that the system can be made more or less strict.

4. PROPOSED METHOD

The signature identification is used to recognize a person. The envelope features and Histogram features are fused for better performance. The signature samples are preprocessed and features are extracted using envelope and histogram techniques. The block diagram of proposed model is given in Figure 2.

1. Signature database: The GPDS 300 signature database is considered. Signatures are obtained from persons on a blank white paper at different timings depending upon the mood and stress levels and are scanned to get images of 96 dpi resolution in png format to create the database. The genuine signature is forged after sufficient training. The database consists of a collection of genuine and forged signature samples for three hundred persons. Each person consists of twenty four genuine signature samples and twenty seven forged signature samples.

2. Preprocessing: For any signature verification technique before extracting the features some set of operations should be performed in order to improve the quality of the image. In the current work we consider the exact signature area, Resizing, binarization and erosion and box removal by morphological operations for filling up the discontinuities in the signature image.

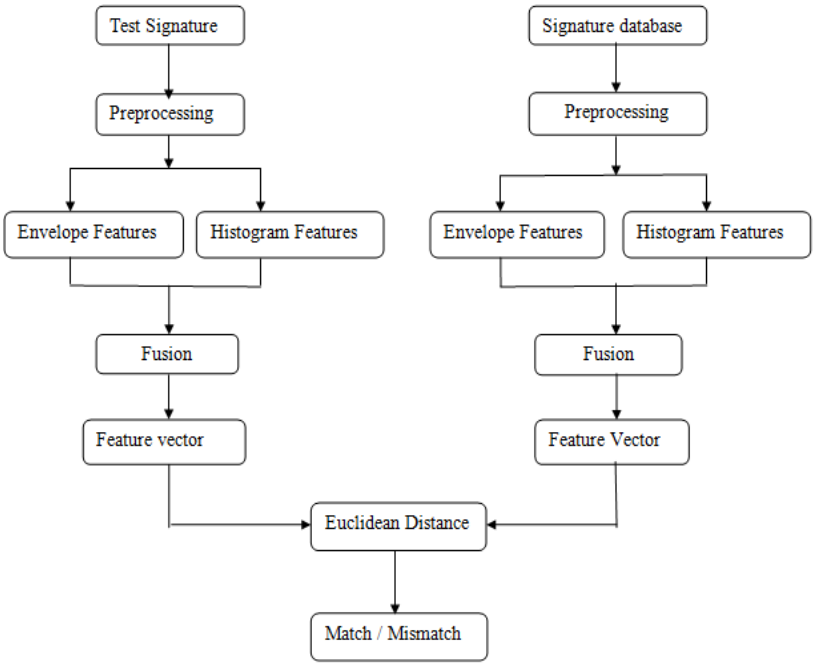

Figure 2: Block diagram of the proposed Model

3. Exact signature area: The signature image might not be present on the entire frame. So, the exact signature area is considered in the skeletonized image for further analysis. This reduces verification time and is cost effective. The signature image is first binarized and then scanned horizontally from top to get the first black pixel row a1 of the image and is the value of the row variable *i* corresponding to first black pixel. The signature image is scanned from the bottom to get the last black pixel row a2 of signature and is the value of the row variable *i* corresponding to last black pixel. The horizontal scanning for finding the top row and bottom row of the exact signature area is given by the Equations 4 and 5.

$$\sum_{i=1}^{M}\sum_{j=1}^{N} I_s(i,j) = 0 \dots\dots\dots\dots\dots\dots (4)$$

$$\sum_{i=1}^{M}\sum_{j=1}^{N} I_s(N+(1-i),j) = 0 \dots\dots (5)$$

$$\sum_{i=1}^{M}\sum_{j=1}^{N} I_s(i,j) = 0 \dots\dots\dots\dots\dots\dots (6)$$

$$\sum_{i=1}^{M}\sum_{j=1}^{N} I_s(i,M+(1-j)) = 0 \dots\dots (7)$$

The signature image is scanned vertically from left to get the first black pixel column *a3* of the image and is the value of the column variable *j* corresponding to first black pixel. The signature image is scanned vertically from right to get the last black pixel column *a4* of the signature and is the value of the column variable *j* corresponding to last black pixel. The vertical scanning for finding exact signature area is given by the Equations 6 and 7.

4. Resizing: In order to standardize the inputs to the system all the images in the database are resized to a uniform size of 200 x 200. This operation is applied on the image after extracting the exact size of the image. The same size is also applied on the test image while comparison.

5. Binarization: The next step is to convert the grayscale image into a binary image. The output image replaces all pixels in the input image with luminance greater than level with the value 1 (white) and replaces all other pixels with the value 0 (black). Specify level in the range [0,1]. This range is relative to the signal levels possible for the image's class. Therefore, a level value of 0.5 is midway between black and white, regardless of class. But in order to avoid the discontinuities in the image a threshold of 0.8 was considered which gives the optimum levels of discontinuities and noise in the image.

6. Erosion: While binarizing an image due to the optimization of the noise in the image and the discontinuities in the image some of the pixels in the signature will be removed leading to the unavoidable discontinuities. In order to eliminate those discontinuities, the signature image is subjected to morphological operations. The one morphological operation that has been considered in this work is erode. Erode operation fills up the missing pixels and thickens the image so that the discontinuities will be eliminated. The images before and after the erosion were shown in the figure 3 (a) and 3 (b) respectively.

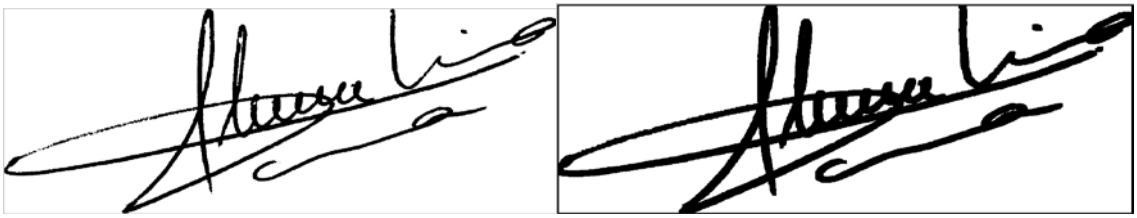

Figure 3: Binarized signature image (a) before erosion (b) after erosion

7. Box removal: If the figure 3 (b) is observed carefully a box is formed at the edges of the image. The box is actually formed due to the erosion operation. The proposed algorithm cannot be applied if the box exists at the edges of the image. So, using manual algorithms the box should be removed. So the Box removal should be done as many times as the erosion operation is applied. The images before and after box removal were shown in the figures 4(a) and 4(b).

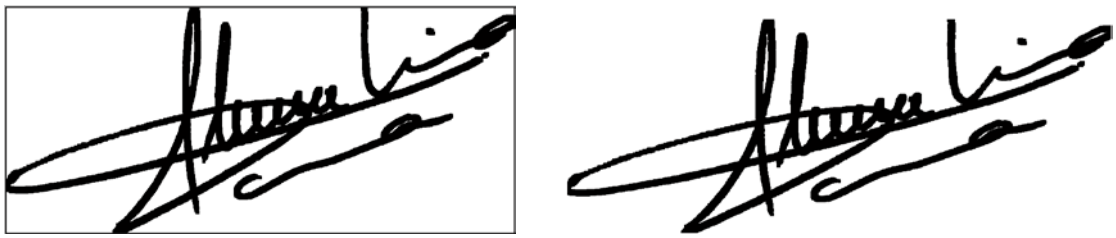

Figure 4: (a) Signature image before box removal (b) after box removal

8. Feature extraction: In the proposed work the feature extraction was done in two different phases. In the first phase the envelope signals were extracted. In the second phase the gray level histogram of the image is considered. The process involved in extracting the two different features is explained as follows.

A. Envelope features: In order to get the envelope signals the preprocessed signature image is scanned in four different phases.

1. Top to bottom: Since the signature is considered as a matrix, it is scanned starting from the first column. The respective row numbers where the first black pixel is encountered were stored in a vector. The length of the vector will be equal to the number of columns in the signature image.

2. Left to right: For getting the envelope from left to right, the image is scanned from the left end. The scanning is started from the first row and the respective column numbers where the first black pixel is encountered is stored in a vector. The length of the vector will be equal to the number of rows in the image.

3. Bottom to top: For getting the bottom to top envelope, the image is scanned from the bottom starting from the first column. The respective row numbers where the first black pixel is encountered will be stored in a vector. The length of the vector will be equal to the number of columns in the image.

4. Right to left: For getting the envelope from right to left, the image is scanned from the right end. The scanning is started from the first row and the respective column numbers where the first black pixel is encountered is stored in a vector. The length of the vector will be equal to the number of rows in the image.

After taking all the envelopes all of them were concatenated into a single vector. It can be observed that the overall length of the envelope feature vector depends on the size of the image. The length of the feature vector will be equal to twice the sum of the dimensions considered i.e., if the image is of size a x b, then the length of the feature vector will be 2 (a+b). As the size of the image considered in this work is 200x200 the length of the feature vector will be equal to 800. Then the feature vector is adjusted to zero for nullifying the position offset of the signature image. The final envelope signals before and after adjusting to zero offset were shown in the figures 5(a) and 5(b).

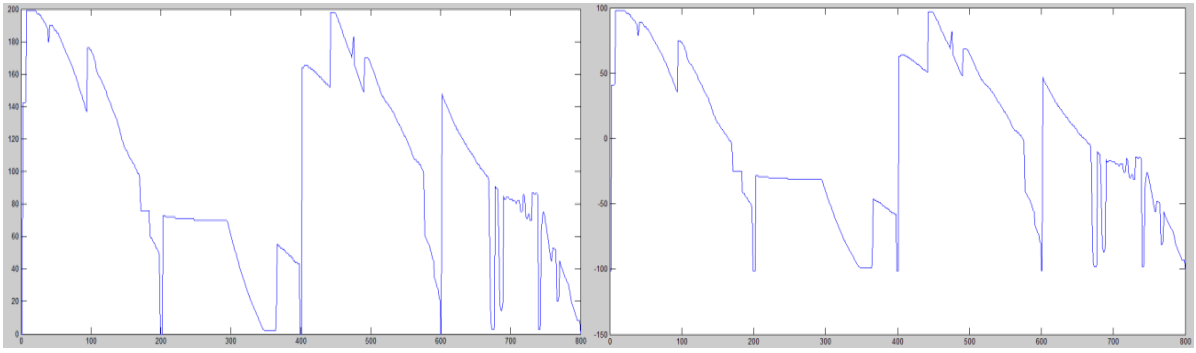

Figure 5: (a) Envelope signal before zero adjusting (b) Envelope signal after zero adjusting

The envelope features of two different signatures were shown in the figure 6. By observing the figure below one can easily deduce the difference between them. The same concept was used in the algorithm in which the dissimilarity is considered to verify whether the two signatures belong to a same person or not.

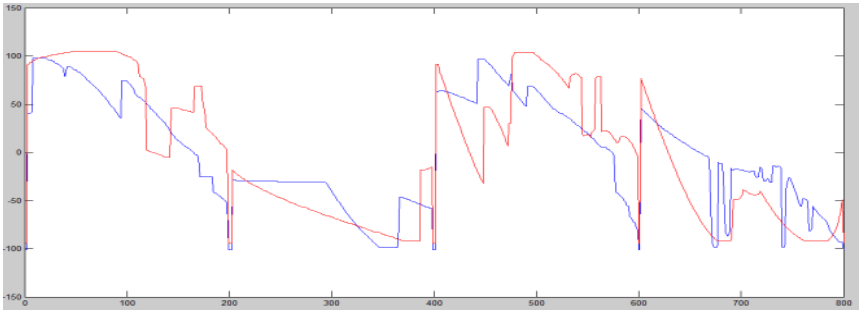

Figure 6: Envelope signals of two signatures made by different persons

B. Histogram features: The second set of features is extracted by considering the histogram of the gray levels in the image. The concept of histogram can be explained by considering a data set with 'I' disjoint categories, let n_i indicate the individual disjoint category and let k_i denote number of samples belonging to the category n_i. Then the total histogram vector having length 'I' can be given by

$$m_i = k_i n_i \qquad \text{---------- (8)}$$

where i = 1,2,3……I.

In general mathematical sense, a histogram is a function m_i that counts the number of observations that fall into each of the disjoint categories (known as bins), whereas the graph of a histogram is merely one way to represent a histogram. Thus, if we let n be the total number of observations and 'I' be the total number of bins, the histogram m_i meets the following conditions.

$$n = \sum_{i=1}^{I} m_i \qquad \text{---------- (9)}$$

In the context of an image processing the histogram of a gray level image is calculated by considering two hundred and fifty six gray levels as the disjoint categories and the pixels as the data set. The histogram is calculated by counting the number of pixels existing for every disjoint gray level ranging from zero to two hundred and fifty five (0 to 255). Thus a vector of length two hundred and fifty six elements is formed as the total gray level categories are two hundred and fifty six. The image and its histogram plot are shown in the figure 7.

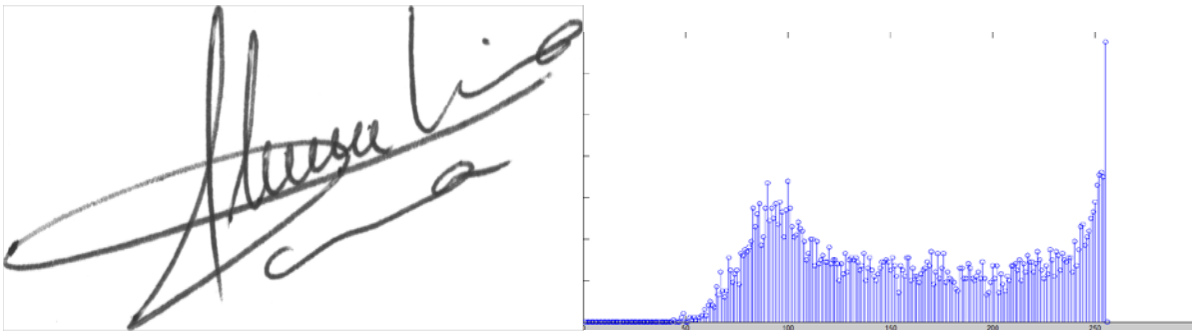

Figure 7: (a) Signature and (b) its histogram

The histogram vectors plotted for a genuine signature and a forged signature were shown in the figure 8. The plot for genuine is represented in blue and the plot for forged is represented in red.

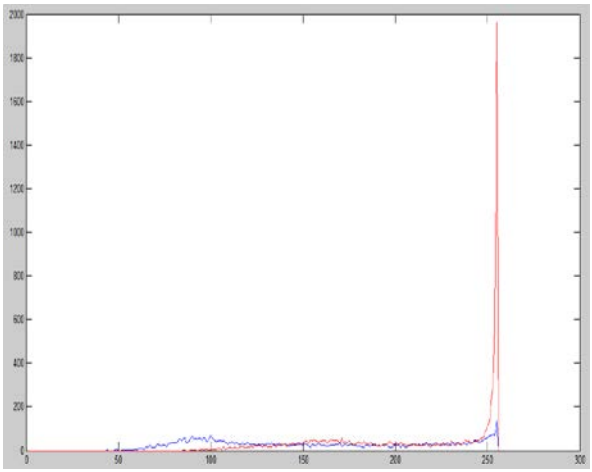

Figure 8: Histograms plotted for a genuine signature

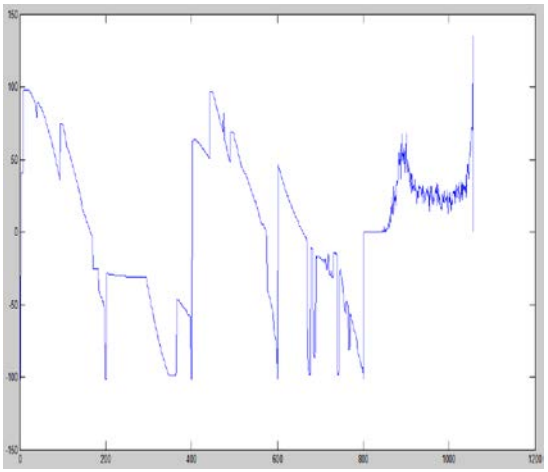

Figure 9 : Final feature set (blue) and a forged signature (red)

Finally the feature set for a signature image is obtained by concatenating the envelope signal and histogram vector. The final feature vector is shown in the figure 9.

9. Matching: Matching in biometric security system is the process of comparing a biometric sample with a stored reference template and subsequently assigning a score based on the level of similarity. Matching allows us to verify whether the person is in database or not. For this, we make use of Euclidean distance. Test image is taken from the database and Euclidean distance is calculated by comparing the feature vector of one test image and feature vectors of all images in the database. Euclidean distance value and position of the image in the database for which Euclidean distance is minimum, is noted and is used to calculate Person number. Euclidean distance value is compared with the threshold value. If the Euclidean distance value is less than the threshold, we have to check whether the person from the database and test image of a person is same. If it is same, then the match count is incremented. If it is not same, then the mismatch count is incremented. If the Euclidean distance value is greater than threshold then the false rejection rate count is incremented indicating the image in database is falsely rejected. Likewise, The biometric system then issues an accept or reject decision based on the results of the matching.

Euclidean distance is used to verify whether the person is in database or not, by comparing final features vector set of database image with final feature vector set of test images.

The Euclidean distance is calculated via equation (10)

$$d(p,q) = \sqrt{(p_i - q_i)^2 + (p_j - q_j)^2} \qquad \text{---------- (10)}$$

Where (p_i, p_j).The feature values of database image. (q_i, q_j) - The features value of test image.

The range of Euclidean distances for all the comparisons of test signatures vs. the database signatures were considered as threshold. Varying the threshold from minimum to maximum the performance parameters of the system were calculated.

The signature under test is considered to be matching when the Euclidean distance of that particular test image and the database image is less than the considered threshold value. If it is greater than the threshold value then the test signature is considered to be mismatched.

5. Results

The performance parameters like FAR, FRR, EER, TSR are calculated using the Euclidean distances between the final feature vectors of the test and database signatures.

Table 1: FAR, FRR, TSR for different thresholds calculated for 20 persons

Threshold	FAR	FRR	TSR
363	0.00	1.00	0.00
494	0.00	0.90	0.10
559	0.00	0.55	0.45
690	0.05	0.35	0.65
821	0.05	0.00	1.00
952	0.20	0.00	1.00
1083	0.40	0.00	1.00
1214	0.65	0.00	1.00
1345	0.80	0.00	1.00
1411	0.90	0.00	1.00
1607	0.95	0.00	1.00

The database is created by considering 20 persons from GPDS 300 with twenty genuine signatures per person, i.e., four hundred signatures are available in the database. In the test section genuine signatures are considered to compute FRR and TSR. The forged signatures are considered in the test section to compute FAR. The values of FAR, FRR and TSR for twenty persons are tabulated in table1 As threshold value increases FAR and TSR increases, whereas FRR decreases. The FAR, FRR and TSR values plotted by considering 20 persons were shown in the figure 10

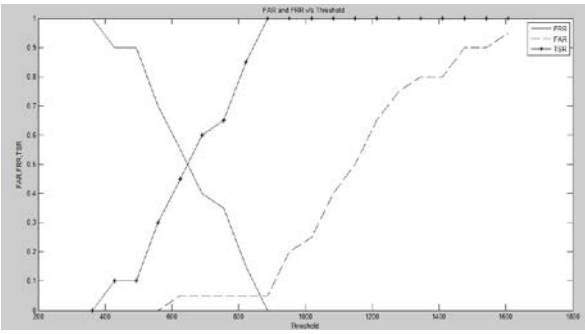

Figure 10: FAR, FRR, TSR vs. Threshold graph

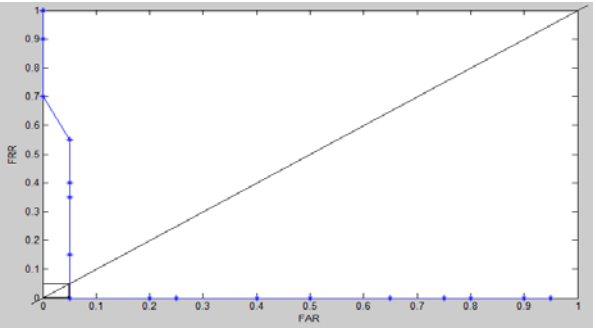

Figure 11: Plot of FAR vs. FRR for 20 persons for 20 persons

The value of EER is 0.05 or 5% .The TSR at the optimum threshold of 830 is obtained as 0.95 or 95%. The plot of FAR vs. FRR for 20 persons is shown in the figure 11. The other way of indicating the EER is using the graph of FAR vs. FRR as shown in the figure 5.2. In the figure the EER is the point on the graph where the line x = y coincides with the graph

drawn between FAR and FRR. In the figure 11 it can be observed that the graph coincides with the line x=y at a point where both FAR and FRR are equal to 0.05. Hence, the EER is 0.05 or 5%.

Table 2: FAR, FRR, TSR for different threshold values for 50 persons

Threshold	FAR	FRR	TSR
0	0.00	1.00	0.00
200	0.00	1.00	0.00
400	0.00	0.94	0.06
600	0.08	0.68	0.32
800	0.20	0.20	0.78
1000	0.70	0.08	0.88
1200	0.92	0.04	0.90
1400	1.00	0.00	0.92
1600	1.00	0.00	0.92
1800	1.00	0.00	0.92
2000	1.00	0.00	0.92

The FAR, FRR and TSR plotted by considering 50 persons were shown in the figure 13.

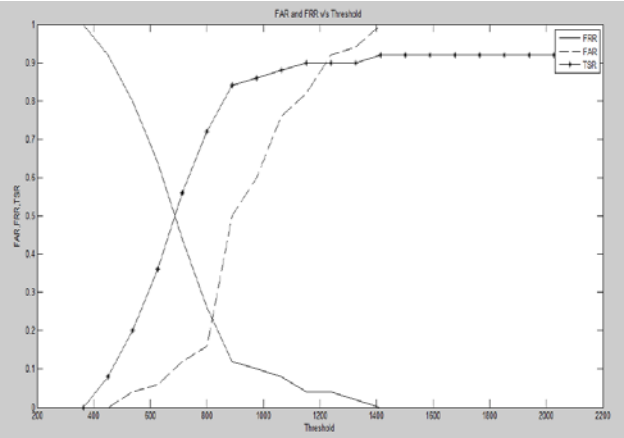

Figure 12: FAR, FRR and TSR vs. Threshold for 50 persons

The database is created by considering 50 persons from GPDS 300 with twenty genuine signatures per person, i.e., one thousand signatures are available in the database. In the test section genuine signatures are considered to compute FRR and TSR. The forged signatures are considered in the test section to compute FAR. The values of FAR, FRR and TSR for fifty persons are tabulated in table 2. As threshold value increases FAR and TSR increases, whereas FRR decreases. The FAR, FRR and TSR values plotted by considering 50 persons were shown in the figure 12. The value of EER is 0.20 or 20% .The TSR at the optimum threshold of 800 is obtained as 0.78 or 78%. The plot of FAR vs. FRR for 50 persons is shown in the figure 13 .The other way of indicating the EER is using the graph of FAR vs. FRR as shown in the figure 13.

In the figure the EER is the point on the graph where the line x = y coincides with the graph drawn between FAR and FRR. In the figure 14. It can be observed that the graph coincides with the line x=y at a point where both FAR and FRR are equal to 0.2. Hence, the EER is 0.2 or 20%.

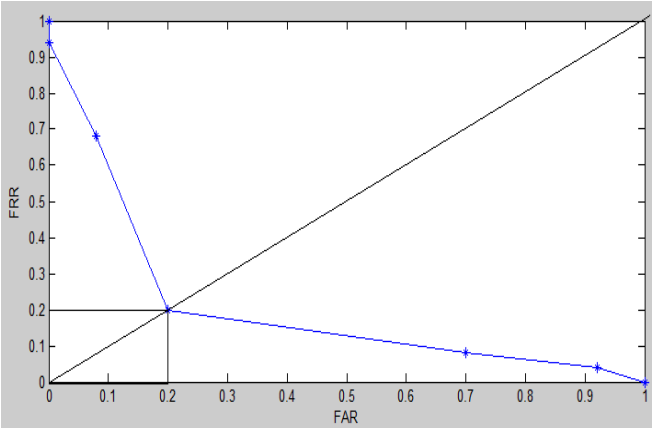

Figure 13: Plot of FAR vs. FRR for 50 persons

A. comparison with the previous results:

Table 3: Comparison with previous work for 20 persons

Method(previous and present)	EER	TSR
Using gabor wavelet transform	----	89%
Using graphometric feature set	10.4%	93.2%
Proposed Method	5%	95%

Table 4: Comparison with previous work for 50 persons

Method	EER	TSR
Using Gabor wavelet transform	21.2%	-----
Using graphometric feature set	21.9%.	*78.1%*
Proposed method	20%	*78%*

From the tables 3 and 4 it can be observed that the proposed model yields better EER and TSR when compared with the previous methods.

6. Conclusion And Future Work

Signature verification using envelope and histogram features has been described. It is developed using a feature set comprising the envelope and histogram features of the image. The envelope features will differentiate the signatures of different persons whereas the histogram features will differentiate the genuine and forged signature of the person. So their combination is used for verification of the signature. The results have been tabulated and have been shown that it gives better EER and TSR when compared to the previous works.

In future the results are expected to be further improved with the use of neural networks or SVM (Support Vector Machines) in the place of Euclidean distance classifier.

7. References

[1]. Mohamad Hoseyn Sigari, Muhammad Reza Pourshahabi and Hamid Reza Pourreza "Offline Handwritten Signature Identification and Verification Using Multi-Resolution Gabor Wavelet", International Journal of Biometrics and Bioinformatics (IJBB), pp. 234-248, Volume (5): Issue (4) : 2011.

[2]. Jing Wen, BinFang, Y.Y.Tang and TaiPing Zhang "Model-based signature verification with rotation invariant features", Pattern Recognition 42, Elsevier, pp.1458 – 1466 : 2009.

[3]. D. Bertolini, L.S.Oliveirab, E.Justino and R.Sabourin "Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers", Pattern Recognition 43, Elsevier, 387 – 396 : 2010.

[4]. Ms.Pallavi Patil and Ms.Archana Patil "Offline Signature Recognition Using Global Features", International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 1, pp. 408 – 411, January 2013.

[5]. Kekre, H.B. and Bharadi, V.A "Signature Recognition using Cluster Based Global Features", Advance Computing Conference,. IACC 2009. IEEE International , pp. 1323 – 1329 : 2009.

[6]. Vahid Malekian, Alireza Aghaei, Mahdie Rezaeian and Mahmood Alian "Rapid Off-line Signature Verification Based on Signature Envelope and Adaptive Density Partitioning", First Iranian Conference onPattern Recognition and Image Analysis (PRIA), pp. 1 – 6 : 2013.

[7]. Vaibhav Shah, Umang Sanghavi and Udit Shah "Off-line Signature Verification Using Curve Fitting Algorithm with Neural Networks" International Conference on Advances in Technology and Engineering (ICATE), pp. 1 – 5 : 2013.

[8]. Suhail M. Odeh and Manal Khalil "Off-line signature verification and recognition: Neural Network Approach" International Symposium on Innovations in Intelligent Systems and Applications (INISTA), pp. 34 – 38 : 2011.