



Offline Signature Verification System with Gaussian Mixture Models (GMM)

Charu Jain¹, Priti Singh², Preeti Rana³

¹Research Scholar, Amity University Haryana

² Department of Electronics and Communication, Amity University Haryana

³Banasthali Vidhyapith

Abstract: Gaussian Mixture Models (GMMs) has been proposed for off-line signature verification. The individual Gaussian components are shown to represent some global features such as skewness, kurtosis, etc. that characterize various aspects of a signature, and are effective for modeling its specificity. The learning phase involves the use of Gaussian Mixture Model (GMM) technique to build a reference model for each signature sample of a particular user. The verification phase uses three layers of statistical techniques. The first layer involves computation of GMM-based log-likelihood probability match score, second layer performs the mapping of this score into soft boundary ranges of acceptance or rejection through the use of z-score analysis and normalization function, thirdly, threshold is used to arrive at the final decision of accepting or rejecting a given signature sample. The focus of this work is on faster detection of authenticated signature as no vector analysis is done in GMM. From the experimental results, the new features proved to be more robust than other related features used in the earlier systems. The FAR (False Acceptance Rate) and FRR (False Rejection Rate) for the genuine samples is 0.15 and 0.19 respectively.

Keywords:

Off-line Signature Verification, Feature Extraction, Gaussian Mixture Model (GMM), Z-score analysis, False Acceptance Rate(FAR), False Rejection Rate(FRR)



Council for Innovative Research

Peer Review Research Publishing System

Journal: INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY

Vol 10, No 6

editor@cirworld.com

www.cirworld.com, member.cirworld.com



1. INTRODUCTION

Signature has been a distinguishing feature for person identification through ages. Even today an increasing number of transactions, especially financial, are being authorized via signatures; hence methods of automatic signature verification must be developed if authenticity is to be verified on a regular basis. Approaches to signature verification fall into two categories according to the acquisition of the data: On-line and Off-line [1,2]. On-line data records the motion of the stylus while the signature is produced, and includes location, and possibly velocity, acceleration and pen pressure, as functions of time. Off-line data is a 2D image of the signature and processed off-line [3]. Difficulty lies in the fact that it is hard to segment signature strokes due to highly stylish and unconventional writing styles. The non-repetitive nature of variation of the signatures, because of age, illness, geographic location and perhaps to some extent the emotional state of the person, accentuates the problem. All these coupled together cause large intra-personal variation. A robust system has to be designed which should not only be able to consider these factors but also detect various types of forgeries. The system should neither be too sensitive nor too coarse. It should have an acceptable trade-off between a low False Acceptance Rate (FAR) and a low False Rejection Rate (FRR). Numerous approaches have been proposed for handwritten signature identification, recognition and authentication systems [5][6].

J. F. Vargas et.al [4] proposed an offline signature verification system based on grey level information using texture features. They analyzed the co-occurrence matrix and local binary pattern and used as features. Genuine samples and random forgeries were used to train an SVM model. Random and skilled forgeries were used for testing. Daksina Ranjan Kisku et.al [8] proposes a technique Support Vector Machines (SVM) to fuse multiple classifiers for an offline signature system. From the signature images, global and local features are extracted and the signatures are verified with the help of Gaussian empirical rule, Euclidean and Mahalanobis distance based classifiers. SVM is used to fuse matching scores of these matchers. Finally, recognition of query signatures is done by comparing it with all signatures of the database. The proposed system is tested on a signature database contains 5400 offline signatures of 600 individuals and the results are found to be promising. Ali Karouni et.al [12] provides a method for offline verification of signatures using a set of simple shape based geometric features like area, center of gravity, eccentricity, kurtosis and skewness. Before extracting the features, preprocessing of a scanned image is necessary to isolate the signature part and to remove any spurious noise present. The system is initially trained using a database of signatures obtained from those individuals whose signatures have to be authenticated by the system. Artificial Neural Network was used to classify and verify the signatures and a classification of about 93% was obtained under a threshold of 90%.

Gaussian Mixture Models (GMM) is generally used as a parametric model of the probability distribution of continuous features in speaker verification system [9]. In this paper, we propose a method, where GMM's yields a better performance in offline signature verification system. We approach the problem in two steps. Initially, the scanned signature image is preprocessed to be suitable for extracting features. Then, the preprocessed image is used to extract relevant parameters that can distinguish signatures of different persons. Section 2 deals with the methodology used for verification. Implementation results and conclusion are listed in Section 3.

2 METHODOLOGY

The algorithm used for the implementation of offline signature verification systems consist of five major modules [7]

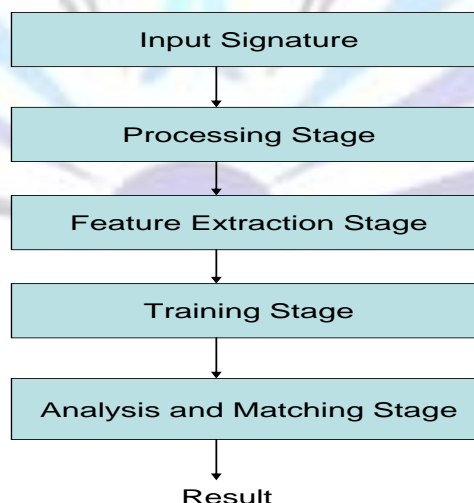


Figure 1: Block Diagram of Proposed Offline Signature Verification System

2.1 Data Acquisition

The first step in the design of a static signature verification system is data acquisition. Handwritten signatures are collected from different individuals and some unique features are extracted from them to create a knowledge base for each individual. The system has been tested for its accuracy and effectiveness on data from 25 users with 10 specimens of

each making up a total of 250 signatures. The proposed verification algorithm is tested on both genuine and forged signature sample counterparts. So we developed a signature database which consists of signatures from all the age groups. Our database is also language independent and also it consists of signatures done with different pens with different colors. 10 users were asked to provide genuine signatures, 5 were asked to do skilled forgeries, 5 provide casual forgeries and 5 did random forgeries. A scanner is set to 300-dpi resolution in 256 grey levels and then signatures are digitized. For further working we cut and pasted scanned images to rectangular area of 3 x 4 cm or 200 x 200 pixels and were each saved separately in files.

2.2 Preprocessing

The purpose of preprocessing [13] is to make signature standard and ready for feature extraction. It is also applied to improve the efficiency and performance of the verification system. Preprocessing steps done are: - converting RGB image to gray, noise removal, cropping in image, binarisation, thinning. The result of preprocessing is as follows:

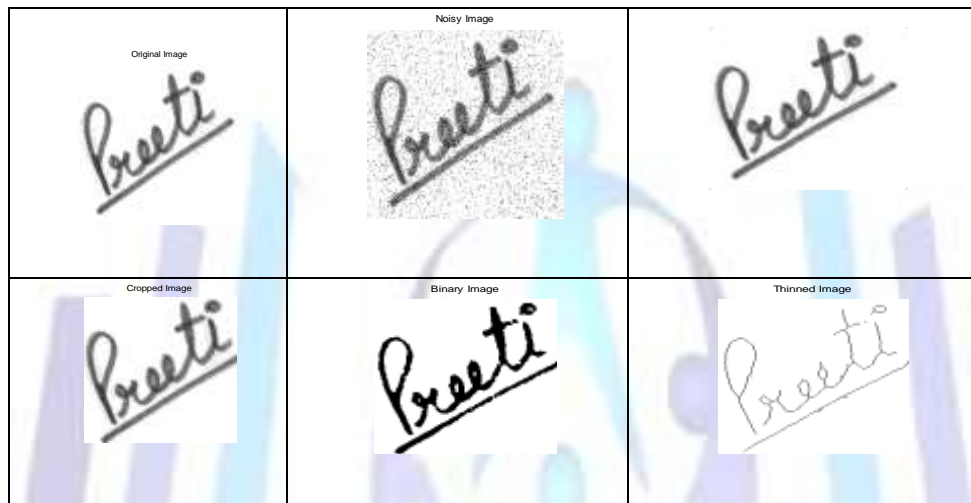


Fig1: All Preprocessing Steps

2.3 Feature Extraction

Features Extraction [14] is the key to develop an offline signature recognition system. Features extracted for offline signature verification can be broadly divided into two main types:

Local features are extracted from a portion or a limited area of the signature image. Local features are applied to the cells of a grid virtually super imposed on a signature image or to particular elements obtained after signature segmentation. These features are calculated to describe the geometrical and topological characteristics of local segments, such as position, tangent direction, and curvature.

Global features depict or categorize the signature as a whole. These features are usually extracted from all the pixels that lie within the region circumscribing the signature image such as the length, width or baseline of the signature [15]. Global Features that are extracted are summarized as follows:-

Skewness, Kurtosis, Signature area (Signature Occupancy Ratio, Height [25], Width [25], Horizontal projection, Vertical projection, Width to Height Ratio (Aspect Ratio), Centre of Gravity, Density of thinned Image, Density of Smoothed Image, Normalized area of black pixels

First, getting all features values of reference signatures as shown in Table 1, then, the mean, variance and standard deviation are calculated as shown in Table 2.



Features	Feature Vector (Mean of all features)
Skew(F1)	-7.1761
Kurtosis(F2)	52.4966
Area(F3)	5.22e+04
Height(F4)	231
Width(F5)	229
Aspect Ratio(F6)	0.9913
nBlack(F7)	51928
nWhite(F8)	971
Centre of gravity(F9)	116.2857,115.0397
Density of thinned image(F10)	0.9816
Density of smoothed image(F11)	220.209
Normalized area of black pixels(F12)	0.9956

Table 1: Values of features that have been extracted

2.4 GMM For Training

Gaussian Mixture Model(GMM) has been used for training our samples. The Gaussian Mixture Models technique is a statistical method that can be used for clustering low dimensional data with the help of several multidimensional Gaussian probability distributions [10] [11]. GMM is a weighted sum of M component Gaussian densities as given by equation (1)

$$p(x_i | \lambda) = \sum_{i=1}^M w_i g(x_i | \mu_i, \Sigma_i) \tag{1}$$

Where X_i is a D-dimensional continuous valued features, w_i , $i=1, \dots, M$, are the mixture weights & $g(x_i | \mu_i, \Sigma_i)$, $i=1, \dots, M$, are the component Gaussian densities.

The construction of a GMM starts with a random initialization of cluster centers (means of Gaussian distributions) and their shapes (covariance matrices). The number of clusters (Gaussians) is a parameter that can be specified by the users. Once the model is trained, it can be used to measure the correspondence between sample data (test set) and the model.

Since good verification results have been achieved in online signature verification using GMMs, we have used them in the training phase. The mean and variance of each feature is estimated. Table for mean, variance and standard deviation for each signature is given below (Table 2)

Features	Mean	Std Dev	Variance
F1	-7.242	0.110782	0.012273
F2	53.45483	1.611712	2.597615
F3	5.24E+04	5317.258	28273236
F4	228.6667	4.041452	16.33333
F5	232.3333	20.2726	408.3333
F6	1.015433	0.075068	0.005635
F7	52211.67	5302.194	28113262
F8	958	73.36893	5383
F9	115.0732	2.020903	4.084049
F10	0.981933	0.000493	2.43E-07
F11	217.0005	3.436678	11.81075
F12	0.995567	0.000153	2.33E-08

Table 2: Mean, Variance and Standard Deviation of a Reference Signatures for a user



Extracted feature values have very different distributions due to which it is required to normalize the feature values before attempting to match the feature vectors. To perform normalization we have used Z –score normalization [16]. This makes use of arithmetic mean & standard deviation of the data.

A Normalized score is produced by the equation (2)

$$s' = \frac{s-\mu}{\sigma} \quad (2)$$

Where μ is the mean & σ is the standard deviation of the matching score distribution. This technique does not guarantee a common range for the normalized scores but ensure that the distributions of each feature will have a mean of Zero & standard deviation of 1.

This transformation only preserves the original distribution of it is Gaussian, due to fact that the mean & standard deviation are the optimal location & scale parameters for Gaussians. The Assumptions that features have Gaussian distribution is acceptable with our application. For each feature, z-score or standard score is calculated.

In order to obtain the threshold of the global features classifier for a specific user, the probability score of the reference signatures is calculated as shown in Table 3.

Signatures	Probability Score
Sig 1	0.68848
Sig2	0.68925
Sig3	0.688832
Sig4	0.688548
Sig5	0.53869

Table 3: Signatures and Their Probability Score

Threshold of a specific user is the minimum values among all reference signatures, and it is kept in his profile. In the test phase, all steps are repeated until getting the probability score value; if the current score is greater than or equal to the threshold value stored in the user profile, the user is accepted as genuine, otherwise is rejected as a forgery.

3 Conclusion

This paper uses Gaussian mixture models for offline signature verification and provided a performance evaluation which showed them to be effective. The algorithm used simple features to characterized signatures that effectively served to distinguish signatures of different persons. The approach was good for continuous pattern. The system was robust and could detect random, simple and semi-skilled forgeries but the performance deteriorates in case of skilled forgeries. A larger database can reduce false acceptances as well as false rejections. Using a higher dimensional feature space and also incorporating dynamic information gathered during the time of signature can also improve the performance. It gives fast output response and needs low storage requirements. The experimental results have shown the ability of the proposed system against all kinds of forgeries.

Results for the proposed system are shown in Table (4).

	FAR	FRR
Genuine	0.15	0.11
Simple Forgery	1.7	1.9
Random Forgery	2.01	1.99
Skilled Forgery	4.77	3.97

Table 4: Simulation Results

4. References

- [1] Plamondon, R. and Lorette, G.: Automatic signature verification and writer identification: the state of the art, Pattern Recognition, 22, 1 07- 13 1, 1989
- [2] Leclerc, F. and Plamondon, R.: Automatic signature verification - the state of the art 1989-1993, Int. J. Patt. Rec. & Art. Int., 8,643-659, 1994.



- [3] Almudena Gilperez, Fernando Alonso-Fernandez, Susana Pecharroman, Julian Fierrez, Javier Ortega-Garcia, "Off-line Signature Verification Using Contour Features", Biometric Recognition Group – ATVS, 2008.
- [4] J.F. Vargas, M.A.Ferrer , C.M.Travieso, and J.B.Alonso, "Off-line signature verification based on grey level information using texture features" Pattern Recognition(Elsevier), vol. no. 44, pp. 375-385, 2011.
- [5] A.K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," IEEE Trans. on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4–20, January 2004.
- [6] Emre Özgündüz, Tülin Şentürk and M. Elif Karslıgil, "OFFLINE SIGNATURE VERIFICATION AND RECOGNITION BY SUPPORT VECTOR MACHINE", 13th European Signal Processing Conference (EUSIPCO 2005), 4-8 September 2005, Antalya, Turkey.
- [7] M. Banshider, R .Y Santhosh and B .D Prasanna (2006): Novel features for off-line signature verification" International Journal of Computers, Communications & Control ,Vol. 1 , No. 1, pp. 17-24.
- [8] Dakshina Ranjan Kisku, Phalguni Gupta, and Jamuna Kanta Sing, "Offline Signature Identification by Fusion of Multiple Classifiers using Statistical Learning Theory", Proceeding of International Journal of Security and It's Applications, vol.4, no.3, pp. 35-44 , July 2010.
- [9] D. Reynolds, T. Quatieri, and R. Dunn. Speaker verification using adapted Gaussian mixture models. Digital Signal Processing, 10(1{3}):19{41, 2000.
- [10] Jonas Richiardi And Andrzej Drygajlo, "Gaussian Mixture Models For Online Signature Verification" November 8, 2003, Berkeley, California, USA.
- [11] G. Xuan, W. Zhang, and P. Chai. EM algorithms of Gaussian Mixture Model and Hidden Markov Model. In Proc. International Conference on Image Processing 2001, pages 145{148, 2001.
- [12] Ali Karouni, Bassam Daya, Samia Bahlak, "Offline signature recognition using neural networks approach", Procedia Computer Science, Elsevier ,vol. no. 3, pp. 155-161, 2 011.
- [13] Bharadi V A and Kekre H B, "Off-Line Signature Recognition Systems", International Journal of Computer Applications vol. 1, no. 27, pp. 975-980, 2010.
- [14] Dr. Daramola Samuel, Prof. Ibiyemi Samuel , "Novel Feature Extraction Technique For Off-Line Signature Verification System" International Journal of Engineering Science and Technology Vol. 2(7), 2010, 3137-3143
- [15] Deepthi Uppalapati, Thesis on "Integration of Offline and Online Signature Verification systems," Department of Computer Science and Engineering, I.I.T., Kanpur, July 2007.
- [16] Sharifah Mumtazah Syed Ahmad, Asma Shakil, Mustafa Agil Muhamad Balbed, "Offline Signature Verification System using Hidden Markov Model in MATLAB Environment", 7th WSEAS Int. Conf. on Applied Computer & Applied Computational Science (ACACOS2008), ,Hangzhou, China, April, 2008