An Efficient Offline Signature Verification System

M. Arathi and A. Govardhan, *Member, IACSIT*

Abstract—**Offline signature verification is one of most challenging area of pattern recognition. Many methods have been introduced in literature to find whether a given signature is genuine or forgery. In the proposed work, the signature image is converted into time series data using linear scanning method and then time series shapelets are identified to distinguish genuine signatures from forged ones. The shapelets are time series subsequences which are maximally representative of a class. To compare the time series data, the proposed method uses Mahalanobis distance measure. The experimental results show that the method has great reduction in equal error rate.**

Index Terms—**Mahalanobis distance measure, time series data, time series shapelets, signatures.**

I. INTRODUCTION

Notwithstanding efforts toward the dematerialization of documents, the need for fast and accurate paper-based document authentication is still growing in our society. The field of biometrics is an important area of study as it offers many advantages over more commonly used authentication methods such as photo ID cards, magnetic strip cards etc. Nowadays, biometric technologies are increasingly and more frequently being used to ensure identity verification. Signatures often incorporate complex geometric patterns that make them a relatively secure means for authorization for high security environments. For historical reasons, the handwritten signature continues to be the most commonly accepted form of transaction confirmation, as well as being used in civil law contracts, acts of volition, or authenticating one's identity. Some other offline signature verification applications include the authentication of bank checks, ID personal cards, administrative forms, formal agreements, acknowledgement of services received, etc. Signature verification has been a topic of intensive research during the past several years due to the important role it plays in numerous areas, including in financial applications.

Considering the large number of signatures verified daily through visual inspection by people, the construction of a robust and accurate automatic signature verification system has many potential benefits for ensuring authenticity of signatures and reducing fraud and other crimes. The goal of an automatic signature verification system is to be able to verify the identity of an individual, based on the analysis of his or her signature through a process that discriminates a genuine signature from a forgery. The verification of human signatures is particularly concerned with the improvement of the interface between human beings and computers.

Depending on data acquisition mechanism, there are two methods of signature verification - Online or Dynamic and Offline or Static. Online method requires special set of devices and instruments to capture the pen movements and pressure over the paper at the same time of the writing. On the other hand, the offline approach uses an optical scanner in order to obtain a digital representation of the signature composed of $M \times N$ pixels. In offline signature verification, the signature image is considered as a discrete 2D function $f(x, y)$, where $x = 0, 1, 2, …, M$ and $y = 0, 1, 2, …, N$ denote the spatial coordinates. The value of f in any (x, y) corresponds to the grey level in that point [1]. Processing is done on the scanned images.

Offline signature recognition is more difficult than online as dynamic information are not available and it is difficult to recover them from the offline images. But requirement of acquiring the signature on some special arrangement makes the online method unsuitable for many of the practical uses. Offline has the advantage of using it in the same way as the existing manual recognition method.

II. DEFINITIONS

A. Types of Forgeries

In signature verification systems, forgeries may be classified into three basic types [2]:

Random forgery: The forger nor has access to the genuine signature neither has any information about the author's name. Forger reproduces a random signature.

Simple forgery: The forger has no access to the sample of the signature but he/she knows the author's name and the forger produces the signature in his/her own style.

Skilled forgery: The forger has access to the samples of the genuine signature and thus he/she is able to reproduce it.

B. Error Rate

In signature verification systems, the performance is evaluated in terms of error rates [2]. There are two types of errors: False Rejection and False Acceptance. Also, there are two types of error rates: False Rejection Rate (FRR) and False Acceptance Rate (FAR). The false rejection rate (FRR) is related to genuine signatures that were rejected by the system; that is, classified as forgeries, whereas the false acceptance rate (FAR) is related to forgeries that were misclassified as genuine signatures. FRR is known as type 1 and FAR is known as type 2 error. The Average Error Rate (AER) is the average of type 1 and type 2 errors [3]. Another factor that determines the efficiency of the system is the Equal Error Rate (EER). The EER is the location on a ROC

or Detection Error Trade-off curve where the FAR and FRR are equal. Smaller the value of EER, better is the performance of the system.

III. RELATED WORK

To improve the efficiency of the signature verification systems, researchers have tried different methods with various approaches. The offline signature verification system proposed in [4] combines some statistical classifiers. This signature verification system consisted of three steps – the first step is to transform the original signatures using the identity and four Gabor transforms, the second step is to intercorrelate the analyzed signature with the similarly transformed signatures of the learning database and then in the third step verification of the authenticity of signatures by fusing the decisions related to each transform.

An automatic off-line signature verification system presented in [5] is built with several statistical techniques. They used Hidden Markov Modeling (HMM) technique to build a reference model for each local feature.

Another system proposed in [6] was based on global, grid and texture features. For each one of the feature sets a special two stage Perceptron OCON (one-class-one-network) classification structure was implemented. In the first stage, the classifier combined the decision results of the neural networks and the Euclidean distance obtained using the three feature sets. The results of the first-stage classifier feed a second-stage radial base function (RBF) neural network structure, which made the final decision.

A proposed system in [7] is based on a contour matching algorithm. They used the geometrical properties of the signature and considered the inevitable intrapersonal variations for the user set.

To tackle the problem of detecting skilled forgeries in off-line signature verification, [8] proposed an approach based on a smoothness criterion. They observed that skilled forgery signatures consisting of cursive graphic patterns are less smooth on a detailed scale than the genuine ones.

Another system was proposed in [9] that has a Cross-validation for Graph Matching based offline Signature Verification (CGMOSV) algorithm. The dissimilarity measure between two signatures in the database was determined by (i) constructing a bipartite graph (ii) obtaining complete matching in and (iii) finding minimum Euclidean distance by Hungarian method.

Another work is based on the total energy that a writer uses to create his/her signature as a global feature [10]. They extracted energy information from the boundary of the whole signature image.

Another work is based on three different kinds of feature extractors - wavelet, curvelet and contourlet transform [11]. The curvature and orientation of a signature image was used as feature. They utilized Support Vector Machine (SVM) as a tool to evaluate the performance of the proposed methods.

In another paper [12], an off-line signature verification and recognition system using the global, directional and grid features of signatures is proposed. Global features used were Signature area, Aspect Ratio of the signature, Maximum horizontal histogram and maximum vertical histogram, Horizontal and vertical center of the signature, Local maxima numbers of the signature and Edge point numbers of the signature. SVM was used for classification.

An offline signature verification system based on two neural networks classifier and three features (global, texture and grid) was proposed in [13]. The first NN classifier they used was three Back Propagation NNs and the second classifier consisted of two Radial Basis Function NNs.

The authors have designed a multi algorithmic signature recognition system [14] considering the conventional features like Number of pixels, Picture Width, Picture Height, Horizontal max Projections, Vertical max Projections, Dominant Angle-normalized, Baseline Shift etc.

Another system is based on score level fusion of distance and orientation features of centroids [15]. The proposed method used symbolic representation of offline signatures using bi-interval valued feature vector.

An offline signature verification and forgery detection proposed in [16] is based on fuzzy modeling that used a model called the "Takagi–Sugeno (TS) model". The TS model involved structural parameters in its exponential membership function.

From results obtained by the researchers in the field of offline signature verification, it is noticed that the statistical approach, (HMMs, Bayesian etc) can detect causal and skilled forgeries. An acceptable recognition rate was reported to be achieved using SVM [14]. Template matching is the simplest and easiest approach, but it is rigid. So, it cannot detect skilled forgery. Still it is suitable for detecting casual forgeries from genuine signatures. When the signature image is considered as a whole entity, the structural approach is useful. But computational complexity in this approach is very high as it requires large training sets. The performance is reported to be better when number of training set is sufficiently large. One additional advantage in using structural pattern recognition is that this approach also provides a description of the given pattern. Signatures with most of the Indian scripted languages are usually long in nature. For such long signatures, spectrum analysis approach (like Curvelet, Contourlet or Wavelet transform) can be better. In [12] authors have found that the contourlet transform could extract better features. Neural Networks based approaches have the advantages of being flexible and adaptive [15]. Neural network based approaches are unified approaches for feature extraction and classification and flexible procedures for finding good, moderately nonlinear solutions. Due to the advent of new learning algorithms and seemingly low dependence on domain specific knowledge, neural network is becoming more popular in the field of pattern recognition. In addition, existing feature extraction and classification algorithms can also be mapped on neural network architectures for efficient (hardware) implementation.

The closest work is that of time series classification using shapelets [17]. Here, the authors classify the time series data using shapelets. A shapelet is a subsequence of time series data which represent a particular class. The algorithms that are based on shapelets are interpretable, accurate and faster than state-of-the-art classifiers. Shapelets are the local features of time series data. Because shapelets are small in size compared to the original data, algorithms that use shapelets for classification, results in less time and space

complexity. For classification with shapelets, decision trees (binary) are used, where each nonleaf node represents a shapelet and leaf nodes represent class labels. To know how well the shapelet classifies the data, information gain [18] is used.

IV. OFFLINE SIGNATURE VERIFICATION PROCESS

Offline signature verification is a pattern recognition problem and a typical pattern recognition system has the following steps [1], [19]: i) Data Acquisition ii) Preprocessing iii) Feature Extraction iv) Classification v) Performance Evaluation.

A. Data Acquisition

For offline signature verification system, images of the signatures are scanned using a digital scanner. Scanned images are stored digitally for offline processing.

B. Preprocessing

The purpose of pre-processing phase is to make signatures standard and ready for feature extraction. The pre-processing stage primarily involves some of the following steps: [3], [19].

1. Noise reduction: A noise filter is applied to remove the noise caused during scanning

2. Resizing: The image is cropped, to the bounding rectangle of the signature

3. Binarization: Transformation from color to grayscale, and then to binary

4. Thinning: The goal of thinning is to eliminate the thickness differences of pen by making the image one pixel thick.

5. Clutter Removal: Any unconnected black dots are removed before processing. This is done by masking.

6. Skeletonization: Skeletonization is used to remove selected foreground pixels from the binary image. So the outcome is a representation of a signature pattern by a collection of thin arcs and curves.

C. Feature Extraction

An ideal feature extraction technique extracts a minimal feature set that maximizes interpersonal distance between signature examples of various persons while minimizing intrapersonal distance for those belonging to the same person [1]. Features extracted for off-line signature verification can be broadly divided into three main categories [20]:

- Global Features
- Local Features
- Geometric Features

Global features: The signature is viewed as a whole and features are extracted from all the pixels confining the signature image.

Local features: Local features are extracted from a portion or a limited area of the signature image [10]. These features are calculated to describe the geometrical and topological characteristics of local segments, such as position, tangent direction, and curvature [21]. These features are generally derived from the distribution of pixels of a signature, such as local pixel density or slant.

Geometric features: These features describe the characteristic geometry and topology of a signature and preserve their global as well as local properties.

D. Classification

It evaluates the evidence presented in the values of the features obtained from feature extraction and makes a final decision for classification

E. Performance Evaluation

The efficiency of the signature verification system is evaluated. It is done based on False Acceptance Rate and False Rejection Rate.

V. PROPOSED METHOD

A. Converting Images to Time Series Data

To create the time series data for signature image, start at the left of the image and consider each column of pixels in turn. The value at each time is just the number of dark pixels in that column. This process is known as *linear scanning*.

Time series analysis is only sensitive to an object's shape. It is invariant to colors and internal features. These properties make time series analysis good for comparing rigid objects, such as skulls, leaves, and handwriting. These shapes do not change over time, so they will have similar time series no matter when they are measured. Time series analysis will not work on objects that can change their shapes over time.

B. Finding Best Shapelet

Initially, all possible shapelets are generated ranging from length 3 to length of the time series data [17]. To find best shapelet, which classifies the data into genuine and forgery, information gain is used. Once shapelet is found, it is inserted into the decision tree. And the above process is repeated. Because one shapelet is not sufficient to classify the data, a number of shapelets are used which clearly distinguishes one class from other. Each shapelet has its corresponding distance threshold value, which divides the data into two sets. Hence, the decision tree which is used as classifier is a binary decision tree. The non-leaf nodes of the decision tree specify shapelet and distance threshold; and leaf nodes specify the class label. To classify a time series data, it is fed into decision tree classifier, which moves it from the root node to leaf node, which in turn gives the predicted class label.

C. Mahalanobis Distance Measure

In order to compare two time series data, the Mahalanobis distance measure is used. The measure is a descriptive statistic that provides a relative measure of a data point's distance from a common point. It was introduced by P. C. Mahalanobis in 1936 [22]. It is a unitless measure. It takes into account the correlations of the data set and is scale-invariant. Hence, it is a suitable measure for classification.

Given a time series $x^{(k)}$, let the ith data point be $x_i^{(k)}$. First, compute the(sample) covariance matrix $C = (c_{ij})$ of a family of time series $x^{(1)}$, $x^{(2)}$,..., $x^{(N)}$ of lengths n by

$$c_{ij} = \frac{1}{N-1}\Sigma_{k=1}^{N}(x_i^{(k)} - \bar{x}_i)(x_j^{(k)} - \bar{x}_j)$$ where N is the

number of instances and where $\bar{x}_i$ is the average of the ith

data point of the time series ($\bar{x}_i = \frac{1}{N} \sum_{k=1}^{N} x_i^{(k)}$).

The Mahalanobis distance measure is a special case of the generalized ellipsoid distance measure $D_M(x, y) = (x - y)^T M (x - y)$ where M is proportional to the inverse of the covariance matrix i.e., $M \alpha C^{-1}$. Though the Mahalanobis distance measure is often defined by setting M to the inverse of the covariance matrix ($M = C^{-1}$), it is convenient to normalize it when possible so that the determinant of the matrix M is one:

$$M = (\det(c))^{\frac{1}{n}} c^{-1}$$ where n is the length of the time series. The Mahalanobis distance measure minimizes the sum of distances between time series $\sum_{x,y} D_M(x, y)$ subject to a regularization constraint on the determinant ($\det(M) = 1$). In this sense, it is optimal.

VI. Experimental Results

The experiments are conducted on dataset containing 1287 questioned signatures and 646 reference signatures. The questioned signatures contain both genuine and forged signatures whereas reference signatures contain only genuine signatures. The proposed method has shown Equal Error Rate of 5.8%. As time series data along with shapelets has been used for classification and Mahalanobis distance measure is used for comparing two time series data, a good result for the experiment has been observed.

VII. Conclusion

The offline signature verification system has been developed by converting the images to time series data using linear scanning and then the time series dataset is classified using shapelets. The shapelets are time series subsequences and are highly representative of a class. In order to compare two time series data, Mahalanobis distance measure is used. Mahalanobis distance measure is a good choice for classification as it takes the correlation of data items into consideration and is scale in-variant. Hence, it is obvious that Mahalanobis distance measure will give more accurate results. The experimental results have also shown that the proposed method results in more accuracy.

References

[1] L. Batista, D. Rivard, R. Sabourin, E. Granger, and P. Maupin, "State of the art in off-line signature verification," in *Pattern Recognition Technologies and Applications: Recent Advances*, 1st edition, B. Verma, M. Blumenstein Eds., IGI Global, Hershey, 2007.

[2] J. Coetzer, B. Herbst, and J. du Preez, "Off-line signature verification using the discrete radon transform and a hidden Markov model," *EURASIP Journal on Applied Signal Processing*, 2004.

[3] I. Guler and M. Meghdadi, "A different approach to off-line handwritten signature verification using the optimal dynamic time warping algorithm Digital Signal Processing," *Science Direct*, vol. 18, pp. 940–950, 2008.

[4] J. B. Fasquel and M. Bruynooghe, "A hybrid opto-electronic method for fast off-line handwritten signature verification," *International Journal on Document Analysis and Recognition*, vol. 7, issue 1, pp. 56-98, March 2004.

[5] S. M. S. Ahmad, A. Shakil, M. A. Faudzi, R. M. Anwar, and M. A. M. Balbed, "A hybrid statistical modeling, normalization and inferencing techniques of an off-line signature verification system," in *Proc. 2009*

[6] H. Baltzakis and N. Papamarkos, "A new signature verification technique based on a two-staged neural network classifier," *Engineering Applications of Artificial Intelligence*, vol. 14, pp. 95-103, 2001.

[7] A. Bansal, D. Garg, and A. Gupta, "A pattern matching classifier for offline signature verification," in *Proc. IEEE First International Conference on Emerging Trends in Engineering and Technology*, July 2008, pp. 1160-1163.

[8] B. Fang, Y. Y. Wang, C. H. Leung, Y. Y. Tang, P. C. K. Kwok, K. W. Tse, and Y. K. Wong, "A smoothness index based approach for off-line signature verification," in *Proc. the Fifth International Conference on Document Analysis and Recognition ICDAR"99*, 1999, pp. 785-787.

[9] A. C. Ramachandra, K. Pavithra, K. Yashasvini, K. B. Raja, K. R. Venugopal, and L. M. Patnaik, "Cross-validation for graph matching based offline signature verification," in *Proc. India Conference INDICON*, 2008.

[10] V. Nguyen, M. Blumenstein, and G. Leedham, "Global features for the off-line signature verification problem," in *Proc. IEEE 10th International Conference on Document Analysis and Recognition*, July 2009, pp. 1300-1304.

[11] M. Fakhlai and H. Pourreza, "Off-line signature recognition based on wavelet, curvelet and contourlet transforms," in *Proc. 8th WSEAS International Conference on Signal Processing and Computational Geometry and Artificial Vision (ISCGAV"08)*, Rhodes, Greece, August 20-22, 2008.

[12] E. Özgündüz, T. Şentürk, and E. M. Karslıgil, "Off-line signature verification and recognition by support vector machine," 4-8 September, 2005, Antalya, Turkey, pp. 113-116.

[13] M. A. Abdala and N. A. Yousif, "Off-line signature recognition and verification based on artificial neural network," *Eng & Tech Journal*, vol. 27, no. 7, 2009.

[14] V. A. Bharadi and H. B. Kekre, "Off-line signature recognition systems," *International Journal of Computer Applications*, vol. 1, no. 27, pp. 0975 – 8887, 2010.

[15] H. N. Prakash and D. S. Guru, "Offline signature verification - an approach based on score level fusion," *International Journal of Computer Applications*, vol. 1, no. 18, pp. 0975–8887, 2010.

[16] M. Hanmandlu, M. Hafizuddin, M. Yusof, and V. K. Madasu, "Off-line signature verification and forgery detection using fuzzy modeling," *Pattern Recognition*, vol. 38, issue 3, pp. 341–356, 2005.

[17] L. Ye and E. Keogh, "Time Series Shapelets: a novel technique that allows accurate, interpretable and fast classification," *Data Min. Knowl. Discov.*, vol. 22, no. 1-2, pp. 149-182, 2011.

[18] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Elsvier Publisher, 2nd edition.

[19] R. O. Duda and P. E. Hart, *Pattern Classification*, 2nd edition, Wiley India Private Limited, 2006.

[20] M. S. Arya and V. S. Inamdar, "A preliminary study on various off-line hand written signature verification approaches," *International Journal of Computer Applications*, vol. 1, no. 9, pp. 0975–8887, 2010.

[21] V. K. Madasu and B. C. Lovell, "An automatic off-line signature verification and forgery detection system," in *Pattern Recognition Technologies and Applications: Recent Advances*, B. Verma, M. Blumenstein, Eds., 1st ed. IGI Global, Hershey, 2007.

[22] P. C. Mahalanobis, "On the generalized distance in statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, 1936.

M. Arathi was born in Hyderabad, Andhra Pradesh, India on 8 October, 1979. She received her B.E.(CSE) from MVSREC, Hyderabad, Andhra Pradesh, India, in 2001, and M.Tech(CS), JNTUH, Hyderabad, Andhra Pradesh, India, in 2008. Her major field of study is data mining.

She has worked as an assistant professor in Sant Samarth Engineering College for 11 months. Now, she is working as an assistant professor in JNTUH, Hyderabad, Andhra Pradesh, India. It is more than 10 years since she has been with JNTUH. She has helped the University in finalizing syllabus for many subjects. She has 4 international journals, 6 international conference and 2 national publications.

Mrs. M. Arathi is an expert committee member for Institute for Innovations in Science and Technology. She has been a judge for many paper presentation contests in JNTUH. She has subject expert for QTP testing tool.

A. Govardhan was born in Nalgonda, Andhra Pradesh, India, on 10 March, 1970. He received his B.E.(CSE) from Osmania University, Hyderabad, Andhra Pradesh in 1992, M.Tech(CS) from JNU, New Delhi, India in 1994, and Ph.D(CS) from JNTU, Hyderabad, Andhra Pradesh in 2003. His areas of research include databases, data mining and information retrieval systems.

He is presently the director at SIT and Executive Council Member at Jawaharlal Nehru Technological University Hyderabad (JNTUH), India. He has 2 Monographs and has guided 125 M.Tech projects, 20 Ph.D thesis and has published 152 research papers at journals/conferences including IEEE, ACM, Springer, Elsevier and Inder Science. He has delivered more than 50 Keynote addresses. He held several positions including the director of Evaluation, principal, HOD and students' advisor.

Prof. A. Govardhan is a member on the editorial boards for eight international journals, and a member of several Advisory & Academic Boards & Professional Bodies and a Committee Member for several International and National Conferences. He is the chairman and a member on several boards of studies of various universities and the chairman of CSI Hyderabad Chapter. He is the recipient of 21 international and national awards.