



ISSN: 2319-5967

ISO 9001:2008 Certified

International Journal of Engineering Science and Innovative Technology (IJESIT)

Volume 2, Issue 4, July 2013

Offline Handwritten Signatures Classification Using Wavelets and Support Vector Machines

Poornima G Patil, Ravindra S Hegadi

Dayananda Sagar Institutions Bangalore, Solapur University Solapur

Abstract— Support Vector Machines is a statistically sound technique which has the capability of learning well separating hyper planes in high dimensional feature spaces. One of the goals of this technique is optimizing the generalization bounds. The various optimizations related to the parameters like the maximal margin, the margin distribution, the number of support vectors are well handled by this technique. In this paper, the handwritten signatures images from a standard database are preprocessed and are decomposed using wavelets. The wavelet approximation and detail coefficients in three directions are subjected to principal component analysis and are used to train the SVM classifier using a linear kernel and a non linear kernel which is Gaussian Radial Basis Function kernel. During the training of the SVM classifier, Sequential Minimal optimization technique is used which fastens up the process of optimization. The FAR and FRR rates of the linear kernel are 13% and 10% respectively. The FAR and FRR rates of Radial Basis Function kernel are 15% and 12% respectively. Further improvement in results can be achieved using higher levels of wavelet decomposition of the images, ranking of key features and selection of appropriate kernels for SVM.

Index Terms— Support vector Machines, Wavelets, linear kernel, non linear kernel .

I. INTRODUCTION

Handwritten signatures are an age old accepted means of a person's identification in government, legal and commercial transactions. As a result, the signatures are highly vulnerable and are often forged and misused. There are many situations where a small piece of information like handwritten signatures are used for identification. It is very essential to correctly classify whether a given signature is a genuine or a forgery. The data used for classification comes from either of the following two approaches.

(i) Offline Approach

(ii) Online Approach

Online approaches use a device or a digitizing surface to capture dynamic features like pressure, speed, direction etc which result in higher accuracies. Off-line verification deals with signatures that have been written on paper and digitized by scanning them. Here the dynamic information is missing which results in low accuracy results. But the offline approaches are needed in many situations. The forgeries related to handwritten signatures are classified into three types.

A. Skilled Forgery

This type of forgery is deliberately created by some professional forgers and they would have sufficiently practiced for a long time to forge other's signatures. These are very hard for detection and hence quite challenging.

B. Casual Forgery

The signer observes the signatures of others closely for a very brief moment and then puts them in his own style without any prior experience.

C. Random Forgery

The signer or the forger creates it by using the name of the victim in his own style to create a forgery known as the simple forgery or random forgery. The signatures are used for identification because it is a well accepted form of identification in the society and a non invasive method which does not annoy an individual being verified. However challenges are many. There are disadvantages like there are intra personal variations in an individual's signature and sometimes a greater variability can be observed in signatures according to age, time, habits or emotional state.



ISSN: 2319-5967

ISO 9001:2008 Certified

International Journal of Engineering Science and Innovative Technology (IJESIT)

Volume 2, Issue 4, July 2013

II. WAVELETS

Wavelet transforms are being extensively used in the domain of image processing since the wavelets have the capability for multiresolution analysis of any signal both stationary and non stationary. The most useful features of Wavelet transforms are the linear nature and orthogonality. Orthogonality eliminates redundancy because each Wavelet term is independent of the other. Since Wavelets can perform multiresolution analysis, they are best used for extraction of time-frequency wavelet coefficients from the signature image. Wavelet transform is especially suitable for any application where the information available is hardly represented by functions. Offline handwritten signatures classification is also one such application. The wavelets can match any signal by various versions of the mother wavelet with various translations and dilations. The various characteristics of wavelet transforms like multiresolution ability, linear nature, and orthogonality make it a desirable technique in many applications and hence we have used this method in our paper for feature extraction.

A. *Wavelets Theory*

Unlike sinusoids, a wavelet is a waveform of limited duration whose average value is zero. The sinusoids theoretically extend from minus infinity to plus infinity where as the Wavelets have a beginning and an end. The multiresolution feature is well supported by Wavelets and hence the representation and analysis of signals at more than one resolution is possible. The multi resolution analysis is capable of ensuring that the features which go undetected at one resolution will be detected at another. Wavelets can work with both stationary and non stationary signals which change frequency over time and can detect anomalies, pulses and events that exist within the signal being analyzed. This is possible because the Wavelet can be stretched or scaled to match the same frequency as the anomaly, pulse or event. Wavelets can also be shifted in time domain to align with the event. Having the knowledge of how much the wavelet was stretched or shifted to align or correlate with the event helps to know the frequency and time of the event. The stretching and shifting of wavelet calculates an index which is a measure of resemblance between the signal and the wavelet located at a particular position and a particular scale. The indexes are termed as Wavelet coefficients and larger the index higher is the resemblance, otherwise it is slight. The higher scales correspond to the most "stretched" wavelets and the higher the stretching of the wavelet, the longer is the portion of the signal with which it is being compared and thus the coarser the signal features being measured by the wavelet coefficients. There is a definite relationship between wavelet scales and frequency as revealed by wavelet. Low scale indicates a compressed wavelet which are the rapidly changing details or in other words high frequency components. High scale indicates a stretched wavelet which are slowly changing or coarse features or in other words low frequency. Wavelet analysis gives a time-scale view of the signal. There are two types of Wavelet transforms possible and they are Continuous Wavelet Transform and Discrete Wavelet Transform.

1) *Continuous Wavelet Transform*

The continuous wavelet transform is going to calculate wavelet coefficients at each possible scale which is indeed tedious. This transforms sums over the all time of the signal multiplied by scaled, shifted versions of the wavelet. This process produces wavelet coefficients that are a function of scale and position. The CWT can operate at many scales starting from that of the original signal up to some maximum scale that one may need. The CWT is also continuous in terms of shifting. The analyzing wavelet is shifted smoothly over the full domain of the signal being analyzed.

2) *Discrete Wavelet Transform*

Calculating wavelet coefficients at every possible scale is a tedious task and it generates a lot of data. In order to avoid this the Discrete Wavelet Transform the wavelet coefficients are calculated only at scales and positions which are chosen based on powers of two and they are called dyadic scales and positions, then the analysis will be much more efficient and just as accurate.

3) *Daubechies Wavelet Transform*

Daubechies wavelets are denoted as dbN where N is the order of the wavelet. This wavelet is both regular and orthogonal. It has compact support and both the continuous wavelet transform and discrete wavelet transforms are possible. Since the wavelets are regular, the regularity of a signal can be easily measured by analyzing the wavelet coefficients.



ISSN: 2319-5967

ISO 9001:2008 Certified

International Journal of Engineering Science and Innovative Technology (IJESIT)

Volume 2, Issue 4, July 2013

III. SUPPORT VECTOR MACHINES

The Support Vector Machines are popularly used in the area of machine learning. SVMs have been applied in various problem areas such as handwritten characters recognition, text classification, object detection and image classification etc. SVMs have proven to yield better results compared to many supervised learning methods. However the success of this method is highly dependent on the kernel selection and also setting the kernel parameters. We have worked with a standard database of 640 handwritten signature images consisting of both genuine signatures and skilled forgeries. The basic idea of support vector machines is to find an optimal hyper plane for linearly separable patterns using kernel function and also it is extensible to patterns that are not linearly separable by transformations of original data to map into new space using kernel functions which are non linear. One unique feature of SVM is that it can efficiently handle high dimensional feature spaces. SVMs maximize the margin around the separating hyper planes. The decision function is typically determined by a subset of training samples called as the support vectors. Support vectors are the data points that lie closest to the decision surface. They are the most difficult to classify. They have direct impact on the optimum location of the decision surface. If linear separability is possible then separation can be done by a line in two dimensions and for higher dimensions we need hyper planes. This has been shown in Fig.1.

IV. RELATED WORK

Many research papers have been published in the area of handwritten character recognition and also in handwritten signature images. Feed forward neural network has been used to verify the offline handwritten signature images which have resulted in good accuracy rates [1]. Text based directional signature recognition algorithm which verifies signatures from the standard database have reported higher accuracies [2]. English handwritten character and digit recognition using the structural micro features and multiclass SVM classifier have exhibited good generalization ability [3]. A hybrid technique of structural, statistical and correlation features used with a multiclass SVM classifier has produced greater accuracies compared to many other techniques for handwritten character recognition [4].

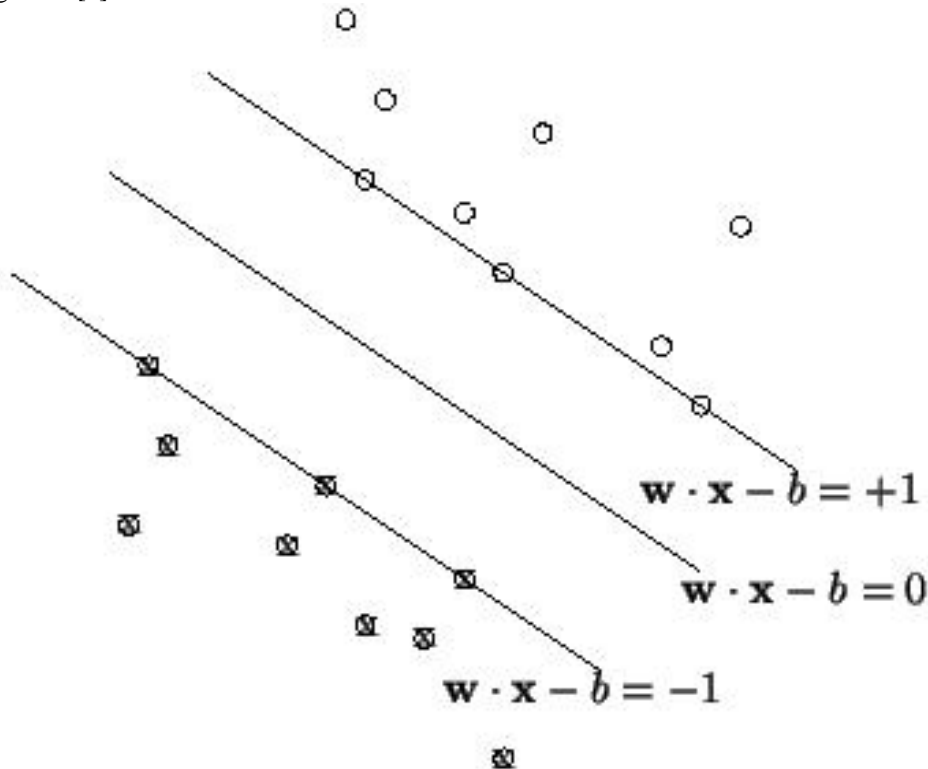


Fig.1 Hyper planes

Define the hyper plane H such that



ISSN: 2319-5967

ISO 9001:2008 Certified

International Journal of Engineering Science and Innovative Technology (IJESIT)

Volume 2, Issue 4, July 2013

$$x_i \cdot w + b \geq +1 \text{ when } y_i = +1$$

$$x_i \cdot w + b \geq -1 \text{ when } y_i = -1$$

$H1$ and $H2$ are the planes

$$H1: x_i \cdot w + b = +1$$

$$H1: x_i \cdot w + b = -1$$

The points on the planes $H1$ and $H2$ are the Support Vectors.

d_+ = the shortest distance to the closest positive point

d_- = the shortest distance to the closest negative point

The margin of a separating hyper plane is $d_+ + d_-$

The energy features of the segmented characters for Devanagari script are used for classification using a SVM classifier [5]. Clustering technique along with finding the correlation between the sample signature of a person and the test signature has been used for verification of offline signatures [6]. Offline signature verification system using discrete wavelet transform and wavelet neural network which uses feature of wavelet energy for classification have promised encouraging results [7]. Signature Verification using rotated complex wavelet filters (RCWF) and dual tree complex wavelet transform (DTCWT) have better identification rates compared to the Discrete Wavelet Transform (DWT) [8]. Offline Signature Verification using multiresolution Gabor Wavelet Coefficients for signatures in four different languages has low error rate in verification [9]. Multi scale discrete Fourier Transform with different wavelet families and distance measures applied for signature verification have shown low average error rate and best results were witnessed in case of Sym8 wavelet family and Manhattan distance [10]. Offline signature verification has been done using contourlet coefficients fed to a SVM classifier and the recognition rates are as high 98 percent [11]. The signatures in African and Persian languages have been recognized using Gabor wavelet coefficients fed to an ensemble of three classifiers with different thresholds applied results in low error rates [12]. A hybrid feature extraction combining wavelet analysis, central projection transformation and fractal theory computes information conserving micro-features and is said to produce higher recognition results [13]. Offline signature verification has been done based on grey level information using texture features [14].

V. PROPOSED SYSTEM

In this paper, we have designed an offline signature verification system which uses the Discrete Daubechies (db4) Wavelet Transform to extract approximation wavelet coefficients and detail coefficients in three directions namely horizontal, vertical, and diagonal. Two SVM classifiers one with a linear kernel and another with a non linear kernel i.e. Gaussian Radial Basis Function kernel have been used for training and classification. The signature images from Standard GPDS database have been preprocessed and then applied Daubechies wavelet (db4) transform to extract wavelet features. Principal Component Analysis of the approximation and detail wavelet coefficients is carried out and first ten principal components each belonging to approximation, horizontal, vertical and diagonal coefficients make the feature set. SVM classifiers are trained using the wavelet features of both genuine and forgery signatures and classification has been done.

A. Preprocessing

There are signatures belonging to 640 subjects where for each person there are 24 genuine signatures and 30 forgery signatures available in the database. From the database four genuine signatures and four forgery signatures of each person are considered for the study. All the signature images are first converted into binary images. Bounding rectangles are put over the signature images to cover only the signature area. Once the bounding rectangles are put, normalization is done in order to resize the signature images with the aspect ratio maintained of the original signature. Bilinear interpolation method has been used for resizing. After normalizing the size, the images have been thinned in order to eliminate the effect of using different types of pens.



ISSN: 2319-5967

ISO 9001:2008 Certified

International Journal of Engineering Science and Innovative Technology (IJESIT)

Volume 2, Issue 4, July 2013

B. Feature Extraction

The preprocessed images of the signatures after the steps of normalizing and thinning are decomposed by db4 wavelet transform. The resulting wavelet coefficients are large in number and in order to choose the important ones and hence reduce the dimensionality of feature space, the principal component analysis (PCA) is carried out for both approximation and detail coefficients. The principal component analysis is done using standardized variables which is based on correlations. As a result of principal component analysis, scores are produced. These scores are the data formed by transforming the original data into the space of the principal components. There are two more vectors namely latent and Hotelling's T^2 which are the output values of the principal component analysis. The variance of the columns of scores are the values of the vector latent and Hotelling's T^2 gives a measure of the multivariate distance of each observation from the center of the data set. The first ten values of latent and Hotelling's T^2 are chosen for the approximation and three detail coefficients namely horizontal, vertical and diagonal. These first ten values corresponding to four types of Wavelet coefficients are obtained as mentioned above for 4 genuine and 4 forgery signatures of each person and these form the feature vector for training the SVM classifier and classification is done. Each signature is represented by total 80 features (10 latent vector values and 10 Hotelling's T^2 values each for four types of wavelet coefficients approximation, horizontal, vertical and diagonal).

Sequential Minimum Optimization (SMO) technique is used in the SVM classifier for faster optimization. SVM classifiers have been designed using both a linear kernel and a non linear kernel namely Gaussian Radial Basis Function kernel. Matlab software has been used for implementation. In this paper, an attempt has been made to understand the performance of linear and non linear kernels over the same dataset with the same number of features.

C. Results

There are signatures belonging to a total of 640 people in the database. There are 24 genuine signatures and 30 forgeries for each person in the database. The database consists of 34560 signatures ($640 \times (24+30)$). This database is divided into training and testing datasets. The 4 genuine signatures and 4 forgery signatures of each person form the training set. They have been preprocessed and then applied the Daubechies wavelet transform in order to extract the features. The remaining signatures of each person form the testing set. The FAR and FRR in case linear kernel are 13% and 10% respectively. The FAR and FRR in case of the non linear kernel i.e. Gaussian Radial basis function are 15% and 12%. The results have been tabulated in Table I. and Table II. The comparison of our proposed method with other wavelet based research works has been presented in Table III.

I. CLASSIFICATION RESULTS OF LINEAR KERNEL FUNCTION KERNEL

Number of persons:640 Number of Genuine signatures per person:24 Number of Forgeries per person:30 Total No of Signatures=34560 (640 *(24+30))			
Number of genuine signatures used for training:4			
Number of forgeries used for training :4			
Number of genuine signatures used for testing	Number of forgeries used for testing	Number of genuine signatures classified correctly	Number of forgeries classified correctly
12800 (20*640)	16640 (26*640)	11520	14476
FAR	0.13%		
FRR	0.10%		

II. CLASSIFICATION RESULTS OF RADIAL BASIS

Number of persons:640 Number of Genuine signatures per person:24 Number of Forgeries per person:30 Total No of Signatures=34560 (640 *(24+30))			
Number of genuine signatures used for training:4			
Number of forgeries used for training :4			
Number of genuine signatures used for testing	Number of forgeries used for testing	Number of genuine signatures classified correctly	Number of forgeries classified correctly
12800 (20*640)	16640 (26*640)	11,264	14362
FAR	0.15%		
FRR	0.12%		



ISSN: 2319-5967

ISO 9001:2008 Certified

International Journal of Engineering Science and Innovative Technology (IJESIT)

Volume 2, Issue 4, July 2013

III. COMPARISON OF PROPOSED METHOD WITH OTHER WAVELET BASED SYSTEMS

Reference	FAR		FRR	
[7]	0.15%		0.12%	
[9]	0.15%		0.15%	
[12]	0.5%		0.35%	
Proposed method	Linear Kernel	Non Linear Kernel	Linear Kernel	Non Linear Kernel
	0.13%	0.15%	0.10%	0.12%

REFERENCES

[1] V. Pandey and S. Shantaiya, "Signature verification using morphological features based on artificial neural network," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, 2012.

[2] J. S. Viriri and J.-R. Tapamo, "Signature verification based on handwritten text recognition," Communications in Computer and Information Science, vol. 61, pp. 98–105, 2009.

[3] S. D. C and P. Hiremath, "Handwritten English character and digit recognition using multiclass svm classifier and using structural micro features," International Journal of Recent Trends in Engineering, vol. 2, pp. 193–200, 2009.

[4] M. N. Ayyaz, I. Javed, and W. Mahmood, "Handwritten character using multiclass svm classification with hybrid feature extraction," Pak. J. Engg. Appl. Sci, vol. 10, pp. 57–67, 2012.

[5] S. K. Shrivastava and P. Chaurasia, "Handwritten devanagari lipi using verification of signature image using clustering technique," International Journal of Smart Home, vol. 4, 2010.

[6] S. Biswas, T. hoon Kim, and D. Bhattacharyya, "Features extraction and studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Tarek, T. Hamza, and E. Radwan, "Off-line handwritten signature recognition using wavelet neural network," International Journal of Computer Science and Information Security, vol. 8, 2010.

[8] M. Shirdhonkar and M. Kokare, "Off-line handwritten signature identification using rotated complex wavelet filters," International Journal of Computer Science Issues, vol. 8, 2011.

[9] M. H. Sigari, M. R. Pourshahabi, and H. R. Pourreza, "Offline handwritten signature identification and verification using multi-resolution gabor wavelet," International Journal of Biometrics and Bioinformatics (IJBB), vol. 5, 2011.

[10] I. A. Ismail, M. A. Ramadan, T. S. E. danaf, and A. H. Samak, "Signature recognition using multi scale Fourier descriptor and wavelet transform," International Journal of Computer Science and Information Security, vol. 7, no. 3, 2010.

[11] M. Fakhrai, H. R. Pourreza, R. Moarefdost, and S. Shadroo, "Off line Signature recognition based on coutourlet transform," signature recognition based on contourlet transform," IPCSIT, vol. 3, 2011.

[12] M. H. Sigari, M. R. Pourshahabi, and H. R. Pourreza, "An ensemble classifier approach for static signature verification based on multi-resolution extracted features," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 5, no. 1, 2012.

[13] Y. Y. Tang, Y. Tao, and E. C. Lam, "New method for feature extraction based on fractal behavior," Pattern Recognition 35, pp. 1071-1081, 2002.

[14] M. J. F. Varga, C. M. Travieso, and J. B. Alonso, "Off-line signature verification based on grey level information using text features," Pattern Recognition, vol. 44, no. 2, pp. 375-385, 2011.

