IJIRST –International Journal for Innovative Research in Science & Technology| Volume 3 | Issue 01 | June 2016
ISSN (online): 2349-6010

Online Signature Verification System using DRT, DCT and K-NN Classifier

Kajal S. Khabiya
UG Student
Department of Information Technology
Pimpri Chinchwad College of Engineering

Lalit V.Sonawane
UG Student
Department of Information Technology
Pimpri Chinchwad College of Engineering

Sagar S. Sonawane
UG Student
Department of Information Technology
Pimpri Chinchwad College of Engineering

Tanuja S. Sali
Assistant Professor
Department of Information Technology
Pimpri Chinchwad College of Engineering

Abstract

The paper presents an effective method for "Online Signature Verification System". For feature extraction we use Discrete Radon Transform (DRT) and Discrete Cosine Transform (DCT). K-Nearest Neighbor Classifier(K-NN) is used as a classification technique which classifies the test signature as genuine of forged. Experimental results obtained on our signature database proves that Discrete Cosine Transform (DCT) works better than Discrete Radon Transform(DRT) and gives high accuracy level to the system.
Keywords: Pre-processing, Discrete Radon Transform(DRT), Discrete Cosine transform (DCT), K-Nearest Neighbor (K-NN) Classifier

I. INTRODUCTION

Biometrics are widely used methods for person identification or verifying the identity of a person based on a physiological or behavioral characteristic of that person such as fingerprints, voice, face, signature. As biometrics is very difficult to forge and cannot be forgotten or stolen, its authentication system offers an accurate, convenient, highly secure and irreplaceable alternative for an individual, which makes it has advantages over traditional authentication schemes. signature has been a distinguishing feature for identification of person.

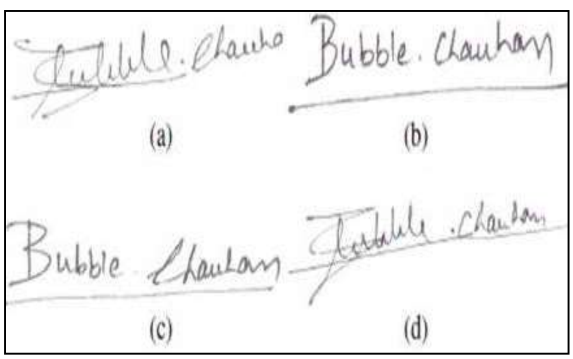

Fig. 1: (a) Original (b) Random (c) Simple (d) Skilled

Based on the ownership, the signature is classified as genuine and forge. Genuine is the original signature. A forged signature (forgery) is the imitation of the genuine signature without the concise of the signer, to the degree of its acceptance for authentication. Based on the knowledge of the forger about the signature and the signer in (see figure 1) forgery can be broadly classified into three types such as: skilled, random and simple forgery[1]. In simple forgery, the forger knows the name of the signer but does not have any idea about the genuine signature pattern and hence produces his/her own pattern of strokes. In random forgery, the forger neither knows the name of the signer nor the signature pattern and produces randomly some pattern of strokes. The skilled forger have the knowledge about the genuine signature sample pattern and also name of the signer, hence results into major threat for verification and authentication of a person through signature. Apart from these, Malik[2] introduced disguised forge, i.e. signature where the authentic signer deliberately tries to hide the identity with the purpose of denial at a later stage.

Signature verification has many applications such as financial transactions, providing electronic signatures for documents, and in providing additional security measures for authentication of computer systems. Unlike other biometrics techniques such as fingerprinting and iris scanning, signature verification is culturally accepted all over as it is less intrusive. These characteristics make signatures more easily accepted as a means of identity verification [3]. In addition to that, although the signature may change for individuals over time; a person signs his/her signature rather uniquely at any given time period and forgeries can be identified by human experts.

Signature verification is studied in two approaches on the basis of the acquisition of the data: On-line and Off-line. For offline signature verification, the images of signatures are found on bank checks and documents that used for verification and are useful in automatic verification of signatures. On the other hand for signatures that are captured by tablets are used in online signature verification. Online verification deals with both static features such as number of black pixels, height and length of signature, etc and some dynamic features as the speed of signing, pressure values, breakpoints, and time taken for generation of a signature, etc of the signature. The online verification extracts dynamic features as well as static features, while offline verification extracts only the static features[4].

II. Related work

There are various approaches that are used for offline signature verification. A review of online signature verification, presented by Gupta and McCabe in 1998, includes a summary of earlier work on the offline case [5].

Substantial research involving English signatures has been done in the field of signature verification. Pal and Blumenstein presented a survey of signature verification systems on non-English and non-Latin signatures to convey the state-of-the-art in the field to researchers. By discussing and comparing different existing approaches it is observed that the maximum work has been performed for Chinese language systems, among the literature of non-English signature verification research. For Arabic, Japanese and Persian only a few pieces of work have been done. From this survey, they conclude that there are still many challenges in this research area, despite of many works done in this research area [6].

Paper by Pippin proposed a technique for online signature verification system. It extracts global features of a signature and by using KNN classification these features are compared with stored signature templates. On the other hand it uses velocity based stroke segmentation to encode the signature in the form of series of strokes and then used to find the closest distance between test and template signatures uses dynamic time warping. The advantage of considering only global features of a signature is that it is simple to compute and addresses privacy concerns [7].

Paper by Kalera, Srihari, and Xu describe a novel approach of feature extraction for signature verification and identification in an offline pattern based on a quasi-multi-resolution technique. They involve structural, gradient, and concavity features, which uses the global as well as the local information of a signature. When these features are used at the word level, instead of the character level, they yield promising results with accuracies as high as 78% for verification and 93% for identification, respectively. Bayes classifier and the k-nearest neighbor classifier techniques are used to solve the verification and identification problem [8].

Nguyen and Bluemnstein [9] developed a global feature based approach, where the horizontal and vertical projection profiles of the signature are used as features. They also described a grid-based feature extraction technique that utilizes directional information extracted from the signature contour and apply 2*D* Gaussian filter on the feature matrices.

Malik, Liwicki and Dengel[2] proposed hybrid approach, where Gaussian mixture models for local features and *k*NN for global features are combined and experimented on ICDAR2009 and ICFHR2010 datasets. However, in their concluding remarks it is mentioned that local features are predominantly better.

Almazain, Fornes and Valveny [10] demonstrated the performance on GPDS-100 and GPDS-750 datasets by calculating the pixel density distribution resulting in non-rigid feature extraction of the signature, but posses better performance with random forgeries only.

In [11], a technique for a bi-script(English and Bengali) offline signature identification system is proposed considering the features such as under-sampled bitmaps, modified chain-code direction features and gradient features computed from both background and foreground components. The performance is evaluated on a database of 1554 English signatures and 1092 Bengali signatures through Support Vector Machine(SVM) and Nearest Neighbour(NN) techniques.

Alan et al.[12] presented a method using a neural network for verifying handwritten signatures. In the feature extraction phase for neural network training, static and dynamic features were extracted. After applying more than one network topologies 3.3% error rate was the achieved accuracy.

Edson et al.[13] presented Hidden Markov Model and a cross-validation for an offline signature verification system. Radon transform was used to detect stroke lines form signature image for the feature extraction phase and as a unique feature the Hough spaces from the signature skeleton were set. For the system evaluation performance neural network was used. For 70 test signatures from different persons recognition rate was 95.24%.

An offline signature verification system was developed by J.Coetzer et al.[14], using the Discrete Radon Transform (DRT) and a HMM. The Global features are extracted using DRT, and then using HHM in a ring the signature are modelled. The Euclidean distance was used for dissimilarity measure. They got 18% EER for skilled forgeries and 4.5% for casual. They suggested combining local features with their global features for future work.

III. Proposed method

The proposed method is based on major 3 phases Pre-processing, Feature Extraction using DRT and DCT, Classification using K-NN Classifier(see figure 2).

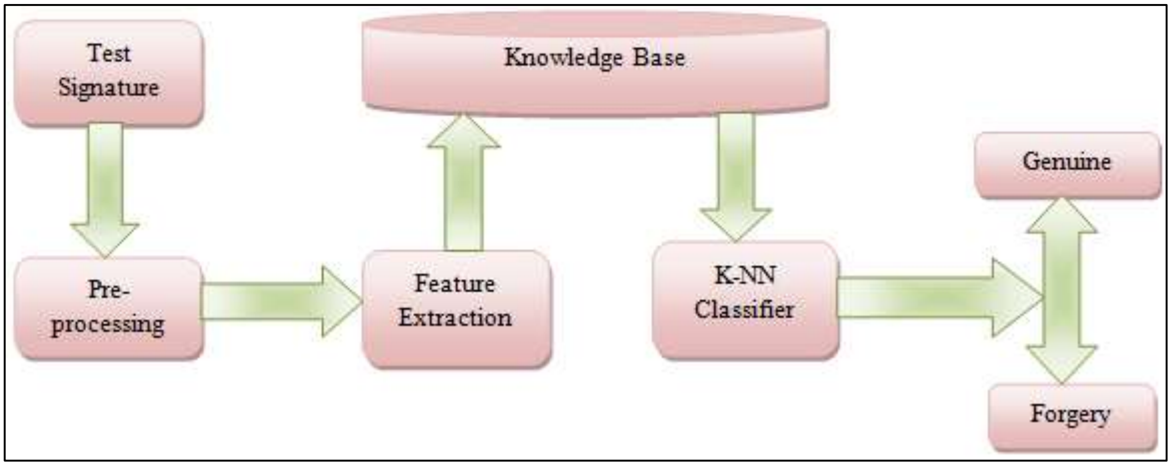

Fig. 2: System Architecture

A. Pre-processing
The pre-processing stage influences the accuracy and reduce the computational time[1]. The input signature image is first converted from RGB to Grey. Thinning is applied on the converted signature image. In thinning process the thickness of the signature is reduced which automatically removes the noise produced during the generation of the image. This signature image is then cropped to eliminate the extra boundary of the image into size 400*400. Finally this signature image is used for feature extraction further.

B. Feature Extraction
1) Discrete Radon Transform
The Discrete Radon Transform is function that computes projections of an image matrix or a shadow of the original image along specified directions. A set of line integrals is the projection of a two-dimensional function f(x,y). The radon function computes the line integrals in a certain direction from multiple sources along parallel paths, or beams. The spacing of the beam is 1 pixel unit apart. To represent an image from different angles, the radon function takes multiple, parallel-beam projections of the image by rotating the source around the center of the image(see figure 3).

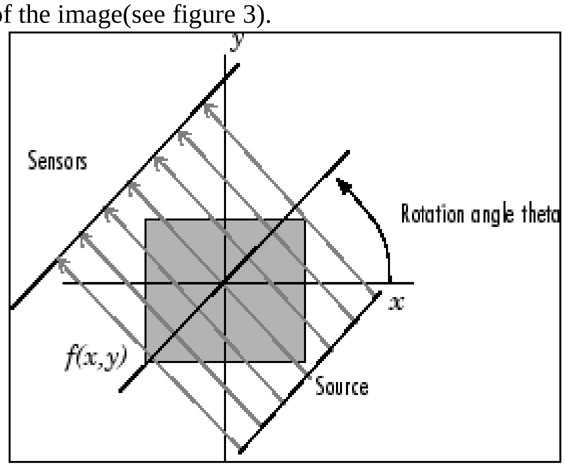

Fig. 3: Single projection at a specified rotation angle

DRT can be expressed as (equation 1)

$$R_j = \sum_{i=1}^{\Psi} w_{ij} I_j \; ; \quad j=1,2,......,N_{\varphi},N_{\theta} \qquad (1)$$

here ,
Rj= the cumulative intensity of the pixels that lie within the jth beam of the image.
Ψ= total number of pixels in an image.
wij= the contribution of the sum of ith pixel to the jth beam of the image.
Ii= the intensity of the ith pixel of the image.
Nφ= non-overlapping beams per angle in the image.
Nθ= number of total angles in the image.
DRT is not a shift invariant representation of a signature image, but shift invariance is ensured by the subsequent image processing. This shift invariant is done by decimating all the zero-valued components from each projection. Then these decimated vectors are shrunk through linear interpolation. Thus to ensure rotation invariance, the projections at angles that range from 180^o to 360^o are also been projected [15].

2) Discrete Cosine Transform

The DCT is a popular image processing and video compression technique which transforms the input signal of spatial domain into a frequency domain. DCT represents the image as a sum of varying frequencies and magnitudes. It is a separable linear transformation. We propose to use DCT-II in our work introduced by Wang [16]. The forward 2D-DCT of M × N block image (equation 2) and (equation 3), f is defined as

$$C(a,b) = \alpha(a)\,\alpha(b) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \times \cos\left[\frac{\pi(2x+1)a}{2M}\right] \cos\left[\frac{\pi(2y+1)b}{2N}\right] \quad (2)$$

The inverse 2D-DCT is

$$f(x,y)= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \alpha(a)\ \alpha(b)\ C(a,b) \times \cos\left[\frac{\pi(2x+1)a}{2M}\right] \cos\left[\frac{\pi(2y+1)b}{2N}\right] \quad (3)$$

where a,b =1,2,....,N-1.
x and y = spatial coordinates in the image block .
a and b = coordinates in transformed image.

It is known that the DCT coefficients are available in the top-left portion of the image of size, say M×N. The DCT coefficients which are located in the upper left corner of the image contains most of the important energy features and hence are considered for further processing. The size of the subset is selected in such a way that it sufficiently represents the input signature image and verify it properly. The top 10×10 =100 coefficients are enough to represent the signature image. [17].

C. *K-NN Classifier*

The main characteristics of any classifier is to classify the unseen data correctly which is not available in the training dataset. Any classifier separates data into training and testing set. K-Nearest Neighbour classifier is used for classification of the signatures[18]. K-NN has two learning techniques, instance-based and lazy learning techniques. K-nearest neighbour algorithm is simplest classification technique because of simple computations. The classification of objects is based on votes of its neighbours which represented by k. K-NN classifies the object to a particular class which has majority of votes. K-NN computes the distance between feature values of the test sample and the feature vector values of every training image. The class of majority among the k-nearest training samples is based on the Euclidian distance measures(4).

$$d(q,x) = \sqrt{\sum_{i=1}^{n}(x_i - q_i)^2} \quad (4)$$

The classification data set consists of few genuine signatures and same number of forged signatures, which is used to test the trained classifier.

In K−NN classification there are two stages:

1) Determination of the nearest neighbours
2) Determination of the class using those neighbours.

Given a set of training dataset D consists of (xi)_(i∈[1,|D|]) training samples. Each of the training example is labelled yj ∈ Y, with class labels. The goal of the classifier is to classify an unknown sample q. The distance between q and xi for each xj ∈ D is calculated as (5)

$$d(q, x_i) = \sum_{f\in F} |\omega_f\ \delta(q_f, x_{i\,f}) \quad (5)$$

F is the set of features and any numerical feature has been normalized to the range[0,1]. There are large range of possibilities for the distance matrices. For continuous and discrete attributes the basic version is (6).

$$\delta(q_f, x_{i\,f}) = \begin{cases} 0 & f \text{ discrete and } q_f = x_{i\,f} \\ 1 & f \text{ discrete and } q_f \neq x_{i\,f} \\ |q_f - x_{i\,f}| & f \text{ continuous} \end{cases} \quad (6)$$

The k-nearest neighbours are selected based on (6) distance matrices [19]

IV. EXPERIMENTAL RESULTS

Experiments are carried out on our dataset that contains 200 signatures from 20 writers. Each writer has 10 genuine signature . For each individual 7 genuine signatures are used as a reference set and the rest of the signatures are used as test set.

Based on the methods previously discussed we can obtain a performance of approximately (79%) when using K-NN as a classifier and DRT for feature extraction(see figure 4).

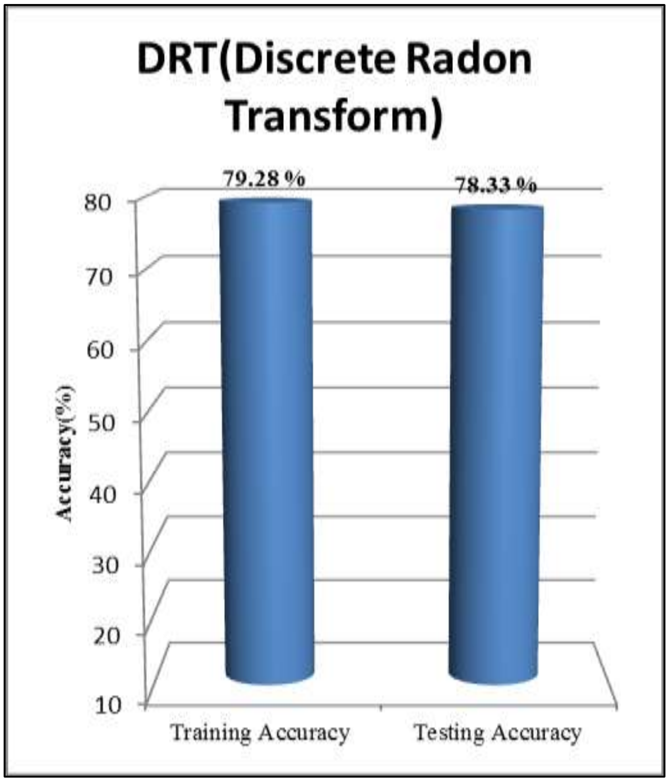

Fig. 4 :DRT based results

Based on the methods previously discussed we can obtain a performance of approximately (81%) when using K-NN as a classifier and DCT for feature extraction (see figure 5).

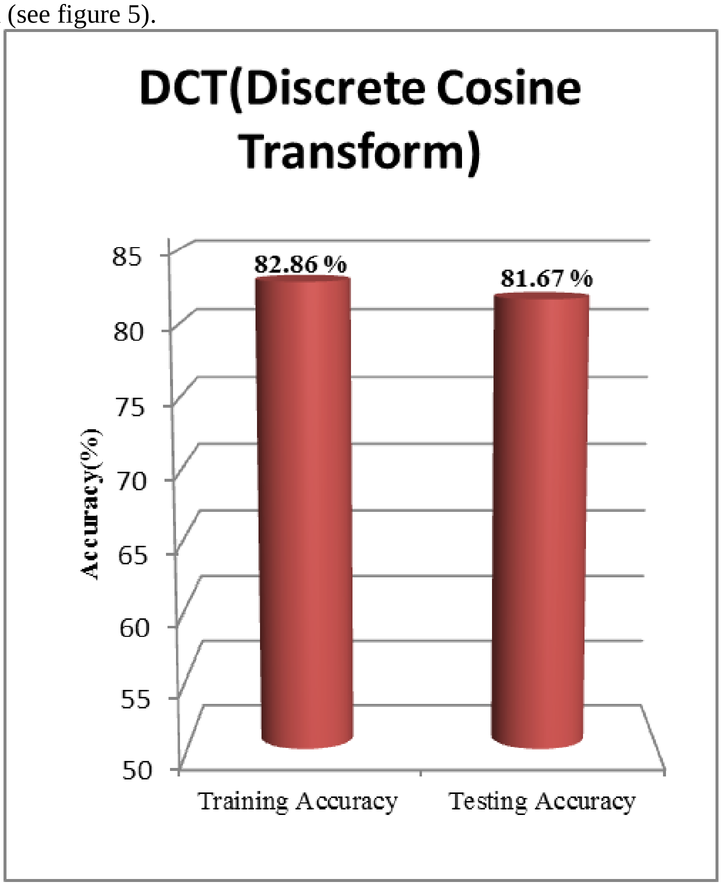

Fig. 5: DCT based Results

In Table 1 we can see the results obtained using K-NN inbuilt function of matlab

Table – 1
Results obtained from inbuilt K-NN function

Feature Extraction Technique	*Genuine Acceptance Rate (GAR)*	*False Acceptance Rate (FAR)*
DRT	*79.30%*	*20.70%*
DCT	*83.70%*	*16.30%*

In Table 2 we can see the results obtained from proper implementation of K-NN algorithm.

Table – 2
Results obtained from K-NN algorithm.

Feature Extraction Technique	*Genuine Acceptance Rate (GAR)*	*False Acceptance Rate (FAR)*
DRT	*79.28%*	*20.72%*
DCT	*83.47%*	*16.53%*

V. CONCLUSION

The experiments with DRT, DCT and K-NN classifier proves that DCT gives higher accuracy than DRT for a dataset of 200 signatures from 20 people. DRT can be proved more accurate and efficient if the dataset contains all the signature images in same alignment. The results obtained from DCT are approximately 81% and from DRT are approximately 78%.

REFERENCES

[1] Yazan M. Al-Omari, Siti Norul Huda Sheikh Abdhullah, and Khairuddin Omar, "State-of-the Art in Offline Signature Verification System," International Conference on Pattern Analysis and Intelligent Robotics, Putrajaya , Malayisa. IEEE, pp. 59-64, June 2011.
[2] M. I. Malik, M. Liwicki, and A. Dengel, "Evaluation of local and global features for offline signature verification," in Proceedings. International Workshop on Automated Forensic Handwriting Analysis, Beijing, China, pp. 26–30, 2011.
[3] C. E. Pippin, "Dynamic signature verification using local and global features", Georgia Institute of Technology, July 2004.
[4] Mayank Vasta, Richa Singh, Pabitra Mitra, and Afzel Noore, "Signature Verification Using Static and Dynamic Feature," Springer-Verlag Berlin Heidelberg, pp. 350-355, 2004.
[5] J. Gupta and A. McCabe, "A review of dynamic handwritten signature verification," Tech. Rep., James Cook University, Australia, 1997.
[6] S. Pal, M. Blumenstein, U. Pal, "Non-English and Non-Latin Signature Verification Systems: A Survey", Proceedings of the 1st International Workshop on Automated Forensic Handwriting Analysis (AFHA), 2011.
[7] Charles E. Pippin ,"Dynamic Signature Verification using Local and Global Features", Georgia Institute of Technology, July 2004.
[8] M. K. Kalera, s. Srihari, a. Xu, "Offline signature verification and identification using distance statistics" , International Journal of Pattern Recognition and Artificial Intelligence, vol. 18, No. 7, pp.1339-1360 , 2004.
[9] S. Khayam, "The discrete cosine transform (dct): theory and application," Michigan State University, 2003.
[10] J. Almazan, A. Fornes, and E. Valveny, "A non-rigid feature extraction method for shape recognition," in International Conference on Document Analysis and Recognition. IEEE, pp. 987–991, 2011.
[11] S. Pal, A. Alireza, U. Pal, and M. Blumenstein, "Off-line signature identification using background and foreground information," in International Conference on Digital Image Computing Techniques and Applications. IEEE, pp. 672–677, 2011.
[12] Alan McCabe, Jarrod Trevathan and Wayne Read, "Neural Network based Handwritten Signature Verification," Journal of computers, vol. 3, no. 8, August 2008.
[13] Edson J. R. Justino, Abdenaimel Yacoubi, Flaviob Ortolozzi and Roberts Abourin, "An Off-Line Signature Verification System Using Hidden Markov Model and Cross-Validation," IEEE Int. Workshop on Neural Networks for Signal Processing, pp. 859–868, 2000.
[14] Coetzer, Herbst and Preezp ,"Offline Signature Verification Using the Discrete Radon Transform and a Hidden Markov Model," EURASIP Journal on Applied Signal Processing , 2004.
[15] A.A. Abdalla Ali, V.F. Zhirkov, "Off line signature verification using radon transform and svm/knn classifiers", TSTU Trans, vol. 15, no 1. pp. 62-69, 2009.
[16] K. Franke and M. Kppen, "A computer-based system to support forensic studies on handwritten documents," International Journal on Document Analysis and Recognition, vol. 3, pp. 218–231, 2001.
[17] T.S. Sali, S. Gore, "Face Recognition Based on Feature Vector Generation using Centric Mean with Forward and Backward Diagonal Mean using DCT," IEEE, pp. 367-375,.
[18] U Ravi Babu, Dr. Y Venkateswarlu, Aneel Kumar Chintha, "Handwritten Digit Recognition Using K-Nearest Neighbour Classifier," World Congress on Computing and Communication Technologies, .IEEE, pp. 60-65, 2014.
[19] P. Cunningham, S. J. Delany, "K-Nearest Neighbour Classifiers," Technical Report, March 2007.