International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-2 Issue-2, December 2012

# An Empirical Study of Signature Recognition &Verification System Using Various Approaches

**Sachin A. Murab, Vaishali. M.Deshmukh**

***Abstract: -Signature used as a biometric is implemented in various systems as well as every signature signed by each person is distinct at the same time. So, it is very important to have a computerized signature verification system. In offline signature verification system dynamic features are not available obviously, but one can use a signature as an image and apply image processing techniques to make an effective offline signature verification system. In this paper, we present implementation of off-line signature recognition and verification system, which is based on moment invariant method, ANFIS, Pairwise distance (pdist) and Kmeans. The user introduces the scanned images into the computer, modifies their quality by image preprocessing followed by feature extraction, ANFIS training, pdist and kmeans.***

***Keywords-component: Image preprocessing, Feature extraction, Moment Invariant method, ANFIS training, pdist & kmeans.***

## I. INTRODUCTION

Signatures are composed of special characters and flourishes and therefore most of the time they can be unreadable. Also interpersonal variations and the differences make it necessary to analyze them as complete images and not as letters and words put together. As signatures are the primary mechanism both for authentication and authorization in legal transactions.A signature may be termed a behavioral biometric, as it can modify depending on many essentials such as: frame of mind, exhaustion, etc. The exigent aspects of automated signature recognition and verification have been, for a long time, a true impetus for researchers. Research into signature verification has been energetically pursued for a number of years [1] and is still being explored (especially in the off-line mode) [2]. Signature recognition and verification involves two separate but strongly related tasks: one of them is identification of the signature owner, and the other is the decision about whether the signature is genuine or forged. Also, depending on the need, signature recognition and verification problem is put into two major classes: (i) online signature recognition and verification systems (SRVS) and (ii) offline SRVS[3].

### A. *Types of Signature Verification*

Based on the definitions of signature, it can lead to two different approaches of signature verification.

#### 1) *Off-Line or Static Signature Verification Technique:*

This approach is based on static characteristics of the signature which are invariant. In this sense signature verification, becomes a typical pattern recognition task knowing that variations in signature pattern are inevitable; the task of signature authentication can be narrowed to drawing the threshold of the range of genuine variation.

In the offline signature verification techniques, images of the signatures written on a paper are obtained using a scanner or a camera.

#### 2) *On-line or Dynamic Signature Verification Technique:*

This is the second type of signature verification technique. This approach is based on dynamic characteristics of the process of signing. This verification uses signatures that are captured by pressure sensitive tablets that extract dynamic properties of a signature in addition to its shape. Dynamic features include the number of order of the strokes, the overall speed of the signature and the pen pressure at each point that make the signature more unique and more difficult to forge. Application areas of Online Signature Verification include protection of small personal devices (e.g. PDA, laptop), authorization of computer users for accessing sensitive data or programs and authentication of individuals [4].

### B. *B. Types of Forgeries*

1) *Random:* these signatures are not based on any Knowledge of the original signature
2) *Simple:* these signatures are based on an assumption of how the signature looks like by knowing the name of The signer.
3) *Skilled:* an imitation of the original signature, which means that the person knows exactly how the original Signature looks like [4].

## II. IMAGE PROCESSING

Image pre-processing represents a wide range of techniques that exist for the manipulation and modification of images. It is the first step in signature verification and recognition. A successful implementation of this step produces improved results and higher accuracy rates. After an image is acquired, it goes through different levels of processing before it is ready for the next step of feature extraction.

### A. *Converting Color image to gray scale image*

In present technology, almost all image capturing and scanning devices use color. Therefore, we also used a color scanning device to scan signature images. A color image consists of a coordinate matrix and three color matrices.

**Manuscript published on 30 December 2012.**
***** Correspondence Author (s)
**Mr. Sachin A. Murab**, Information Technology Department, P.R.M.I.T & R,Badnera, Sant Gadge Baba Amravati University.Amravati, India.
**Prof Vaishali M. Deshmukh, Assistant Professor**, CSE Department, P.R.M.I.T &R,Badnera, Sant Gadge Baba Amravati University.Amravti, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

Coordinate matrix contains x, y coordinate values of the mage. The color matrices are labeled as red (R), green (G), and blue (B). Techniques presented in this work are based on grey scale images, and therefore, scanned or captured color images are initially converted to grey scale.

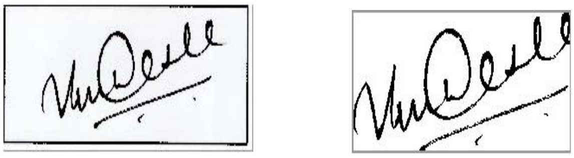
### *B. Background elimination and border clearing:*

Many image processing applications require the differentiation of objects from the image background. *Thresholding* is the most trivial and easily applicable method for this purpose. It is widely used in image segmentation [5, 6]. Thresholding is choosing a threshold value T and assigning 0 to the pixels with values smaller than or equal to T and 1 to those with values greater than T. We used thresholding technique for differentiating the signature pixels from the background pixels. Clearly, in this application, we are interested in dark objects on a light background, and therefore, a threshold value T, called the brightness threshold, is appropriately chosen and applied to image pixels f(x, y) as in the following[3]:

If f(x,y) ≥ T then
f(x,y) = Background
else f(x,y) = Object

### *C. Signature Normalization*

Signature dimensions may vary due to the irregularities in the image scanning and capturing process. Furthermore, height and width of signatures vary from person to person and, sometimes, even the same person may use different size signatures.



Original Signature Image        Normalized Signature

Figure. 1 Signature Normalization

## III. FEATURE EXTRACTION

Feature extraction is the second major step in signature recognition and verification. If we are to compare 2 sketches; there should be at least one measurement on which to base this comparison. The main function of this step is to generate features which can be used as comparison measurements. Since the issue of signature verification is a highly sensitive process, more than one feature/measurement has to be generated in order to enhance the accuracy of the results. Image Features

### *A. GLCM (Grey Level Co-occurrence Matrix)*

The GLCM is a tabulation of how often different combinations of pixel brightness values (grey levels) occur in an image. Statistics provide information about the texture of an image. The table I lists the statistics.

**Table: I**

| **Statistic** | **Description** |
|---|---|
| Contrast | Measures the local variations in the gray-level co-occurrence matrix. |
| Correlation | Measures the joint probability occurrence of the specified pixel pairs. |
| Energy | Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment. |
| Homogeneity | Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal |

### *B. Color Dominant*

Color is one of the most dominant and distinguishable low-level visual features in describing image.

Dominant color descriptor (DCD) :

DCD contains two main components: representative colors and the percentage of each color. DCD can provide anEffective, compact, and intuitive salient color representation, and describe the color distribution in an image or a region of Interesting[7].

Firstly the image is uniformly divided into 8 coarse partitions. If there are several colors located on the same partitioned block, they are assumed to be similar. After the above coarse partition, the centroid of each partition is selected as its quantized color. Let X=(XR, XG,XB) represent color components of a pixel with color components Red, Green, and Blue, and Ci be the quantized color for partition i. The average value of color distribution for each partition center can be calculated by

$$\bar{X}_i = \frac{\sum_{X \in C_i} X}{\sum_{X \in C_i} 1}$$

After the average values are obtained, each quantized color.

$$C_i = (\bar{X}_i^R, \bar{X}_i^G, \bar{X}_i^B) \ (1 \le i \le 8)$$

In this way, the dominant colors of an image will be obtained.

### *C. Histogram*

An image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image. The histogram plots the number of pixels in the image (vertical axis) with a Particular brightness value (horizontal axis), Image histogram function displays a histogram for the image given as a input above a grayscale colorbar. The number of bins in the histogram is specified by the image type. If the given image is a grayscale image, Image histogram function uses a default value of 256 bins. If given image is a binary image, function uses two bins.

## IV. MOMENT INVARIANT METHOD

Moment invariants are properties of connected regions in binary images that are invariant to translation, rotation and Scaling. They can be easily calculated from region properties and they are very useful in performing shape classification and part recognition. One of the techniques for generating invariants in terms of algebraic moment was originally proposed by Hu [8]. The algebraic moment of the characteristic function f(x,y) is defined to be:


Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

$m_{pq} = \int_x \int_y x^p y^q f(x,y) d_y d_x$

This can be approximated in discrete form by:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y)$$

A geometric figure can be uniquely determined by its algebraic moment. Therefore, instead of looking for invariants of moments, only invariants of low order moments are used in practical applications. Moment invariants are usually specified in terms of centralized moment. Here, the moment is measured with respect to the "center of mass", (x', y'). The central moment, μ, with respect to the centroid, and the normalized central moment,η, are calculated as:

$m_{pq} = \sum_x \sum_y (x - x')^p (y - y')^q a_{xy}$

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\lambda}$$

$$Where, \lambda = \frac{P+Q}{2} + 1, (P+Q) > 2$$

Extracted from the signature using various image processing techniques. This paper will be completed when the utility of signature verification is shown i.e. it helps in detecting the exact person and it provides more accuracy of verifying signatures for implementation of above, this paper The moment invariants used in our research are computed using the equations given in Following Table for all signatures at various angles.

*Global Properties:* seven global features are used for better results. These features are signature height-to width ratio, maximum vertical projection, maximum horizontal projection, image area, vertical center of signature, vertical projection peaks and horizontal projection peaks.

## V. ANFIS

The acronym ANFIS derives its name from adaptive neuro-fuzzy inference system. Using a given input/output data set, anfis function constructs a fuzzy inference system (FIS) whose membership function parameters are tuned (adjusted) ANFIS uses a hybrid learning algorithm for training to construct an input-output relationship [9]. The basic structure of the ANFIS under consideration will be described below.

ANFIS Training

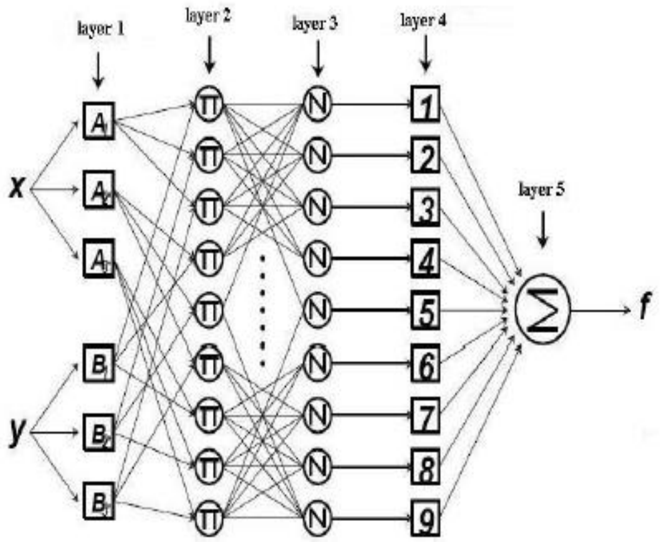



Figure. 2 Two-input one-output zero order ANFIS

A two-input one-output zero order ANFIS with three membership functions per input is shown in figure 1. It is composed of five layers. The square nodes in the first and fourth layer represent adaptive nodes, e.g. they depend on a set of parameters; the circle nodes are fixed. The outputs of the first layer are the membership values of the inputs that correspond to the membership functions. The values that are obtained from the first layer are multiplied in the second layer in order to obtain the firing strength or weight of the rules. In the third layer, the ratios of the firing strengths of the rules to the sum of all rules' firing strengths are calculated. In the fourth layer, these values are multiplied by the consequence parameters. Finally, the fifth layer adds all the incoming signals.

## VI. PAIRWISE DISTANCE (pdist)

Pairwise distance computes the Euclidean distance between pairs of objects in data matrix. Euclidean distance is th most common distance measure, is the geometric distance in multidimensional space.y = pdist(X) computes the Euclidean distance between pairs of objects in n-by-p data matrix X. Rows of X correspond to observations; columns correspond to variables. y is a row vector of length n(n–1)/2, corresponding to pairs of observations in X. The distances are arranged in the order (2,1), (3,1), ..., (n,1), (3,2), ..., (n,2), ..., (n,n–1)). y is commonly used as a dissimilarity matrix in clustering or multidimensional scaling.

### *A. Other Distance Measures :*

1) ***Squared Euclidean distance:***

   The squared Euclidean distance places greater emphasis on objects that are further apart.

2) ***City block distance:***

   Both city block distance and Euclidean distance are special cases of the Minkowski metric. Where the Euclidean distance corresponds to the length of the shortest path between two points, the city-block distance is the sum of distances along each dimension.

3) ***Cosine distance:***

   The cosine of the angle between two vectors of values

4) ***Pearson correlation distance:***

   The difference between 1 and the cosine coefficient of two observations. Cosine coefficient is the cosine of the angle between two vectors.

5) ***Jacquard distance:***

   Difference between 1 and the Jaccard coefficient of two observations. For binary data, Jaccard coefficient equals to the ratio of sizes of intersection and union of two observations

## VII.K-MEANS

Cluster analysis is one of the basic tools for exploring the underlying structure of a given data set and is being applied in a wide variety of engineering and scientific disciplines such as medicine, psychology, biology, sociology, pattern recognition, and image processing.


*Published By:*
*Blue Eyes Intelligence Engineering*
*and Sciences Publication (BEIESP)*
*© Copyright: All rights reserved.*

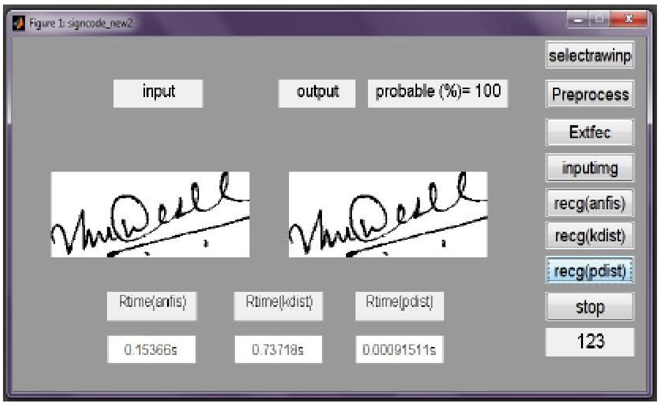
The primary objective of cluster analysis is to partition a given data set of multidimensional vectors (patterns) into so-called homogeneous clusters such that patterns within a cluster are more similar to each other than patterns belonging to different clusters [10].

The K-means algorithm assigns each point to the cluster whose center (also called centroid) is nearest. The center is the average of all the points in the cluster that is, its coordinates are the arithmetic mean for each dimension separately over all the points in the cluster [11]. K-means can be thought of as an algorithm relying on hard assignment of information to a given set of partitions. At every pass of the algorithm, each data value is assigned to the nearest partition based upon some similarity parameter such as Euclidean distance of intensity [12]. The partitions are then recalculated based on these hard assignments. With each successive pass, a data value can switch partitions, thus altering the values of the partitions at every pass. K-Means algorithms typically converge to a solution very quickly as opposed to other clustering algorithms.

The algorithm steps are:

1) Choose the number of clusters, *k*.
2) Randomly generate *k* clusters and determine the cluster centers, or directly generate *k* random points as cluster centers.
3) Assign each point to the nearest cluster center.
4) Recomputed the new cluster centers.
5) Repeat the two previous steps until some convergence criterion is met.

## VIII. IMPLEMENTATION AND TEST RESULT




In this work the signature samples were acquired from many individuals. Each individual was asked to sign non-overlapping signatures using black pen or any color pen on a white sheet of paper. A total10 signatures were collected from each person which are signed at different time to collect samples of intra-personal variations. All these sample signatures were scanned and stored as genuine signatures. After that stored signature image was preprocessed in the database, the preprocessed stage include border clearing, background removing and signature normalization. Next step is feature extraction in this step we have considered several features of an image such as GLCM, Color Dominant &Histogram for the extraction purpose, after feature extraction we used Moment Invariant method, In this method we have calculated results of seven invariant moments. These results and all the different extracted features are then given as input data for ANFIS training as well as input to the pairwise distance (pdist) & Kmeans approach. ANFIS Training, pdist and Kmeans works on the basis of extracted feature and output of seven moment invariant by moment invariant method. When we use all these methods for signature recognition, we found that the pdist method takes less time for recognition as compare to ANFIS and Kmeans, when we have small database, But if we increase the size of database it could take more time.

## IX. CONCLUSION

This paper is focused on Implementation of Offline signature recognition and verification. Signatures are verified based on parameters extracted from the signature using various image processing techniques. It helps in detecting the exact person and it provides more accuracy of verifying signatures. For implementation of this work we have used the ANFIS, pdist and Kmeans method which are worked on several different features and Moment Invariant method. Our recognition system exhibited a 100% success rate by identifying correctly all of the signatures that system was trained for. However, it exhibited poor performance when it was presented with signatures that it was not trained earlier.

## REFERENCES

[1]. K. Han, and I.K. Sethi, "Handwritten Signature Retrieval and Identification", Pattern Recognition 17, 1996, pp. 83-90.
[2]. S. Chen, and S. Srihari, "Use of Exterior Contour and Shape Features in Off-line Signature Verification", 8th International Conference on DocumentAnalysis and Recognition (ICDAR '05), 2005, pp. 1280-1284.
[3]. Cemil OZ, Fikret Ercal, Zafer Demir "Signature Recognition and Verification with ANN ",Skarya University Computer Eng. Department Sakarya,Turkey,UMR Computer Science Department Rolla,MO65401.
[4] Ms. Vibha Pandey Ms. Sanjivani Shantaiya. "Signature Verification Using Morphological Features Based on Artificial Neural Network", 2012 pp. 288-292
[5] Erdem, U.M., "2D Object Recognition In Manufacturing Environment Using Implicit Polynomials and Algebraic Invariants",Master Thesis, Bogazici University, 1997.
[6] Fu, K.S., Mui, J.K., "A survey On Image Segmentation", Pattern Recognition, Vol. 13, pp.3-16, Pergoman Press, 1981.
[7] M.Babu Rao, Dr.B.Prabhakara Rao and Dr.A.Govardhan," Content Based Image Retrieval using Dominant Color and Texture features", February 2011, pp. 118-123.
[8] MING-KUEI HU," Visual Pattern Recognition by Moment Invariants", February 4, 2010,pp. 179-187.
[9] Henry José Block Saldana And Carlos Silva Cardenas," Design and Implementation of an Adaptive Neuro- Fuzzy Inference System on an FPGA used forNonlinear Function Generation",IEEE 2010.
[10] Mu-Chun Su, Chien-Hsing Chou," A Modified Version of the K-Means Algorithm with a Distance Based on Cluster Symmetry", JUNE 2001,pp 674- 680.
[11] Yiu-Ming Cheung, k*-Means: "A new generalized k-means clustering algorithm", Pattern Recognition Letters 24, 2003.
[12] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, and Angela Y. Wu, "Efficient Algorithms for K Means Clustering".
[13]. Plamondon, "The Handwritten Signature as a Biometric Identifier: Psychophysical Model & System Design" IEE Conference Publications,R.1995,Issue CP408, 23-27.
[14] Ali Karouni Bassam Daya, Samia Bahlak," Offline signature recognition using neural networks approach", Procedia Computer Science 3,2010,pp.155- 161.
[15] E.J.R. Justino, F. Bortolozzi, and R. Sabourin. "A comparison of SVM and HMM classifiers in the off-line signature verification", Pattern Recognition Letters 26, 2005, pp. 1377-1385.
[16] Suhail M. Odeh, Manal Khalil ,"Off-line signature verification and recognition: Neural Network Approach",IEEE 2011,pp.34-38
[17]. Vu Nguyen, Yumiko Kawazoey, Tetsushi Wakabayashiy, Umapada Palz, and Michael Blumenstein, "Performance Analysis of the Gradient Feature and the Modified Direction Feature for Off-line Signature Verification", the IEEE 12th International Conference on Frontiers in Handwriting Recognition, 2010,pp. 303-307.


*Published By:*
*Blue Eyes Intelligence Engineering*
*and Sciences Publication (BEIESP)*
*© Copyright: All rights reserved.*

[18]. Bradley Schafer, Serestina Viriri, "An Off-Line Signature Verification System", IEEE International Conference on Signal and Image Processing Applications, 2009, pp. 95-100

[19]. Michael Blumenstein, Graham Leedham, Vu Nguyen,"Global Features for the Off-Line Signature Verification Problem", IEEE 10th International Conference on Document Analysis and Recognition, 2009,pp.1300-1304.

[20] Alan McCabe, Jarrod Trevathan and Wayne Read, school of mathematics, physics and Information Technology, James cook University, Australia." Neural Network-based handwritten signature verification",Journal of computers, vol.3, No. 8 August 2008, pp. 9-22.

[21] Minal Tomarand Pratibha Singh," A DirectionalFeature with Energy based Offline SignatureVerification Network" Vol.2, No.1, February 2011, PP48-57

**Mr. Sachin A. Murab**, Information Technology Dept, P.R.M.I.T & R,badnera,Sant Gadge Baba Amaravati University, Amravati.

**Prof. Vaishali M. Deshmukh**, Assistant Professoer at Computer Science & Engineering Dept, P.R.M.I.T. & R. Badnera, Amravati, Sant Gadge Baba Amaravati University, B.E (Computer Science & Engineering),M.E (Computer Science & Engineering).Professional Membership : ISTE, FIE, IETE.


© Copyright: All rights reserved.