Handwritten Signature Verification And Recognition Using ANN

Mohan Mandaogade
Department Of Electronics Engg.
P.R.E.C,LONI.

Saurabh
Department Of Electronics Engg.
P.R.E.C,LONI.

Vishal Mhaske
Department Of Electronics Engg.
P.R.E.C,LONI.

ABSTRACT
Automatic person identification is one of the major concerns in this era of automation. However, this is not a new problem and our society has adopted several different ways to authenticate the identity of a person such as signature and possessing a document. With the advent of electronic communication media (Internet), the interactions are becoming more and more automatic and thus the problem of identity theft has became even more severe. Even, the traditional modes of person authentication systems such as Possessions and Knowledge are not able to solve this problem. Possessions include physical possessions such as keys, passports, and smart cards. Due to inability of knowledge and possession based authentication methods to handle the security concerns, biometrics research have gained significant momentum in the last decade as the security concerns are increasing due to increasing automation of every field. Biometrics refers to authentication of a person using a physiological and behavioral trait of the individual that distinguish him from others. Biometric authentication has various advantages over knowledge and possession based identification methods including ease of use and non-repudiation. In this paper, we address the problem of handwriting biometrics and present a method for verifying handwritten signatures by using an ANN.

1. 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, the need for research in efficient automated solutions for signature recognition and verification has increased in recent years. Various methods have already been introduced in this field and application of Statistical models is one of them in this regard. Using statistical knowledge, we can easily perform the relation, deviation, etc between two or more data items. Strictly speaking, to find out the relation between some set of data items we generally follow the concept of Correlation Coefficients. In general statistical usage, correlation or correlation refers to the departure of two variables from independence, although correlation does not imply causation. Our approach is based on the above concept. To verify an entered signature with the help an average signature, which is obtained from the set of, previously collected signatures, we have followed the concept of correlation to find out the amount of divergence in between them using ANN.

1.1 Review of prior work:
Lot of previous work and techniques are there for signature identification and verification with their own pros and cons. In this section we look at nomenclature associated with signature verification and types of techniques along with some previous work in this field.

For person's identity various forms of biometric securities exists, such as finger print recognition, iris recognition, speech recognition, heart sound recognition and key stroke recognition. But the simplest method for verifying the person's identity is through the use of handwritten signature. The signature represents a personal style of human handwriting. Various projects have been carried out in the past on signature verification with varying degrees of success. Azriel Rosenfeld has done static signature work. In the Dynamic signature field a number of groups have taken several approaches. Especially the related work of Neural Network based handwritten signature verification by Alan McCabe, Jarrod Trevathan and Wayne Read School of mathematics, Physics and Information technology, and James cook university, Australia. Off-line signature verification using HMM for Random, Simple and Skilled forgeries by Edson J.R. Justino, Flavio Bortolozzi, Robert sabourin and Automatic on-line signature verification by Vishvjit S. Nalwa of Bell Laboratories, Holmdel, NJ. P.Drouhard, R.Sabourin and M. Godbout-A neural network approach to off-line signature verification using directional PDF, are under considerations.

Signature verification systems can be categorized into two main classes:

1. Off-line Signature Verification
2. Online Signature Verification.

2. IMAGE PROCESSING
The camera-captured or scanned real world images containing human signatures are processed using several image processing algorithms before the calculation of the moment invariants. These processes are given below.

2.1 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.Coordinate matrix contains x,y coordinate values of the image. The color matrices are labeled as red (R), green (G), and blue (B). The scanned or captured color images are initially coverted to grey scale using the following equation :

Gray color = 0.299*R + 0.5876*G+0.114*B

2.2 Noise Reduction
Noise reduction (also called "smoothing" or "noise filtering") is one of the most important processes in image processing. Images are often corrupted due to positive and negative impulses stemming from decoding errors or noisy channels. An image may also be degraded because of the undesirable effects due to illumination and other objects in the

environment. Median filter is widely used for smoothing and restoring images corrupted by noise. Different from linear filters such as the mean filter, median filter has attractive properties for suppressing impulse noise while preserving edges. Median Filter is used in this study due to its edge preserving feature .

2.3 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 . 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:

If $f(x,y) \geq T$ then

$f(x,y) =$ Background

else $f(x,y) =$ Object

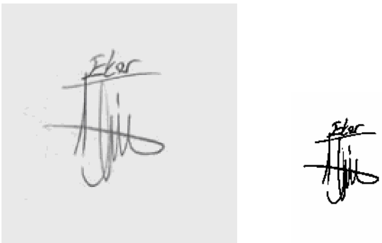


Figure 1. a) Captured signature b) signature image with background removed

Signature image which is located by separating it from complex background image is converted into binary image white background taking the pixel value of 1.Vertical and horizontal (histogram) projections are used for border clearing. For both direction, vertical and horizontal, we counted every row zeros and the resulting histogram is plotted sideways.

2.4 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. First, we need to eliminate the size differences and obtain a standard signature size for all signatures. After this normalization process, all signatures will have the same dimensions. In this study, we used a normalized size of 40x40 pixels for all signatures that will be processed further. During the normalization process, the aspect ratio between width and height of a signature is kept intact. Normalization process made use of the following equations:

$$x_i = \left\{ \frac{x_i - x_{min}}{x_{max} - x_{min}} \right\} * M$$

$$y_i = \left\{ \frac{y_i - y_{min}}{y_{max} - y_{min}} \right\} * M$$

In these equations:

x_i,,y_i : pixel coordinates for the normalized signature,

x'_i,,y'_i : pixel coordinates for the original signature,

M : one of the dimensions (width or height) for the normalized signature

The normalization process is demonstrated in the following figure.

a) Original signature image b) Normalized signature

Figure 2. Signature normalization

3. 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. The algebraic moment of the characteristic function f(x,y) is defined to be:

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

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 3"center of mass", (x', y'). The central moment, μ, with respect to the centroid, and the normalized central moment, □ , 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 $\lambda = (p+q)/2$, $(p+q) > 2$

The moment invariants used in our research are computed using the equations given in Table-1(b) for all signatures at various angles.

Table 1 : (a) Formulas used for specific central moments b) List of the derived invariant moments

Central Moments	**Derived Invariant Moments**
$\mu 00=m00$	$I1=\eta 20 + \eta\ 02$
$\mu 10=0$	$I2=(\eta\ 20 - \eta\ 02)^2+4\ \eta\ 11^2$
$\mu\ 01=0$	$I3=(\eta\ 30 - 3\ \eta\ 12)^2+ (3\ \eta\ 21- \eta\ 03)^2$
$\mu 20=m20-x'm00$	$I4=(\eta\ 30 + \eta\ 12)^2+ (\eta\ 21- \eta\ 03)^2$
$\mu 02= m02-y'm01$	$I5=(\eta\ 30 - 3\ \eta\ 12)\ (\eta\ 30 + \eta\ 12)^2(\ (\eta\ 30 + \eta\ 12)^2\ -3(\eta\ 21 +\eta\ 03)^2) + (3\ \eta\ 21 - \eta\ 03)\ (\eta\ 21 + \eta\ 03)(3\ (\eta\ 30 + \eta\ 12)^2\ -(\eta\ 21 + \eta\ 03)^2)$
$\mu 11= m11-y'm10$	$I6=(\eta\ 20 - \eta\ 02)\ ((\eta\ 30 + \eta\ 12)^2- (\eta\ 21 + \eta\ 03)^2)+ 4\ \eta\ 11(\eta\ 30 + \eta\ 21\ )(\ \eta\ 21+ \eta\ 03)$
$\mu 30= m30-3x'm20 +2x'm10$	$I7=(3\ \eta\ 12 - \eta\ 30)\ (\eta\ 30 + \eta\ 12)(\ (3\ \eta\ 30\ \eta\ 12) - 3(\eta\ 21 +\eta\ 03)^2) + (3\ \eta\ 21- \eta\ 03)\ (\eta\ 21 + \eta\ 03)((\ 3\ \eta\ 30\ \eta\ 12)^2\ -(\ \eta\ 21+ \eta\ 03)^2)$

4. DIGITIZATION OF SIGNATURES

4.1 Calculating moment invariants:

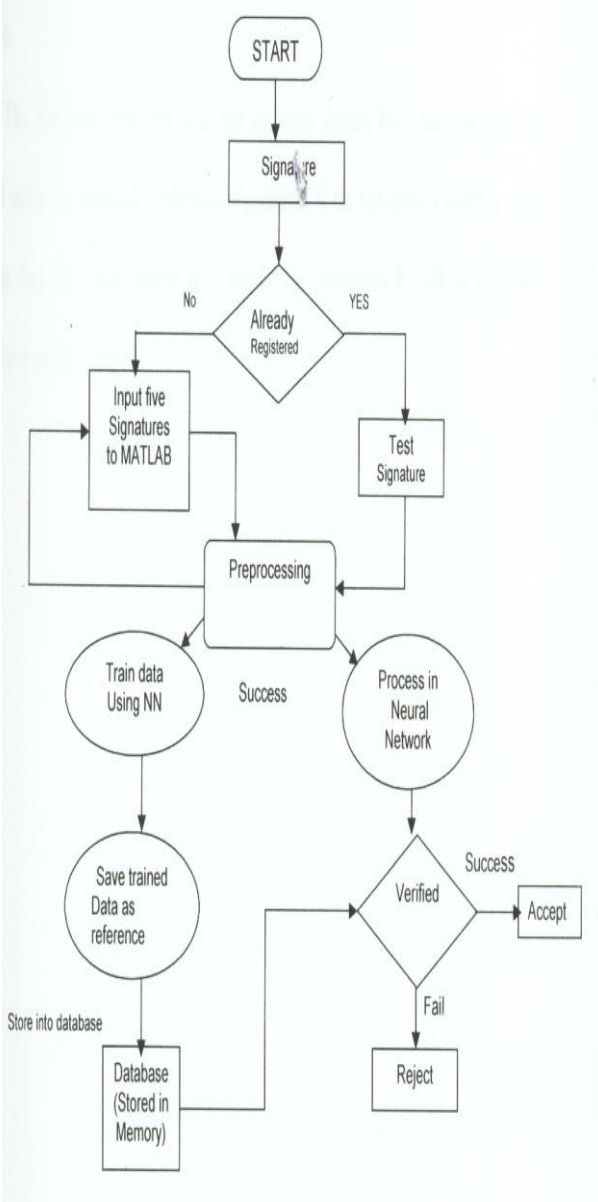
Feature vectors are generated using moment invariants. For this purpose, we use six different signature images which are sign different time. Then, we produced six different sets of feature vectors for every signature where each set consisted of seven moment invariant values listed in Table 1(b). A sample feature vector is shown in Figure 3.

5.18647933247972E-0001
1.72351384466328E-0001
1.32449449923652E-0001
2.13629276301316E-0001
1.45538602648840E-0002
8.86884674185101E-0002

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.

4.2 Proposed System:

The block diagram of proposed system is shown below:



5. ANN DESIGN FOR SIGNATURE RECOGNITION AND VERIFICATION

An Artificial Neural Network (ANN) is an information-processing paradigm that is inspired by the way biological nervous systems works, such as the brain, process information. The key element of this paradigm is the structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working tog.ether to solve specific problems. ANNs, its just like people, learn by example. An ANN is designed for a specific application, such as a data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well.

We designed a multilayer feed forward artificial neural network for recognition of off-line digitized signatures. The proposed ANN consists of 14 input variables, 18 hidden neurons, and 30 output variables and it is designed to recognize one signature at a time. Backpropagation algorithm is used for training.

5.1 Training for signature recognition

First, an input/output database is created manually for training and testing the ANN for six signature image which are belong to same person but signed different time. Each input vector consists of seven moment invariants obtained for a signature. As explained earlier in section 4, six different moment invariant vectors are produced for each signature. These six vectors are divided into two sets each containing three vectors. One of these sets (3 input vectors) is used in the training of ANN and the other set (remaining 3 input vectors) is used for testing. Additional, We also produce seven extra properties for a signature. The database contained a total of 30 different signature images which are used for both training and testing. Since 3 input vectors for each image is used for training purposes, there are a total of 90 (30*3) input vectors (data sets) in the training set. The remaining 90 data sets are used for testing. ANN contained 30 binary output values each corresponding to one signature being tested. Under normal (correct) operation of ANN, only one output is expected to take a value of "1" indicating the recognition of a signature represented by that particular output. The other output values must remain zero. In general, the number of outputs must be equal to the number signatures being considered.

5.2 Signature verification

In this part of the study, our purpose is to authenticate a signature i.e. to verify that the signature is not counterfeit and it really belongs to the person who is claimed to be the owner of the signature. The ANN used for this purpose is also a multilayer feed forward network which consists of 14 input variables, 10 hidden neurons, and 2 output variables indicating whether the signature is fake or true.

Backpropagation algorithm is used for training. The training data set is obtained from three original (authenticated) signatures provided by the real owner and three fake signatures. As it was done for the preparation of the training data for the ANN used in recognition, three invariant vectors per signature is used in the training set. Therefore, a total of 18 moment invariant vectors are used in the training set. A sample set of three signatures belonging to the same person is shown in Figure 4.

Figure 4. Three signatures which belong to the true owner of the signature

6. IMPLEMENTATION AND TEST RESULTS

6.1 Signature recognition

The program used a windows interface . This software allowed the signature images to be loaded one at a time and used in training and testing. First, the signature image is captured using a CCD camera or a scanner, then, through several image processing operations, it is converted to binary and normalized to a 40*40 image as explained earlier. Moment invariant and additional values are obtained from the normalized image which is then used as the input vector to the ANN. After the training of the ANN for signature recognition, the system is ready to recognize a given signature.

The signature recognition system is tested using 30 signatures chosen at random. The images were obtained using the following properties:

- Signatures were signed inside a special framed area.
- Images were taken with a simple CCD camera and they were shot from a fixed distance.

As explained in Section 5.1., 30 images in our database belonging to 30 different signatures are used for both training and testing. Since 3 (out of 6) input vectors for each image were used for training purposes, there are only 90 (30*3) input vectors (d ata sets) left to be used for the test set. Under normal (correct) operation of the ANN, only one output is expected to take a value of "1" indicating the recognition of a signature represented by that particular output. The other output values must remain zero. The output layer used a logic decoder which mapped neuron outputs between 0.5-1 to a binary value of 1. If the real value of an output is less than 0.5, it is represented by a "0" value. The ANN program recognized all of the 30 signatures correctly. This result translates into a 100% recognition rate. We also tested the system with 10 random signatures which are not contained in the original database. Only two of these signatures which are very similar to at least one of the 30 stored images resulted in "false positives" (output > 0.5) while the remaining 8 are recognized correctly as not belonging to the original set (the output value was <= 0.5). Since recognition step is always followed by the verification step, these kinds of false positives can be easily caught by our verification system. In other words, the verification step serves as a safeguard against "false positives" as well as "false negatives".

6.2 Testing the verification system

Training for verification is explained in Section 5.2.Signatures used for testing the verification system are obtained the same way as in the recognition system. We tested the verification software using 10 signatures; 5 imitations (counterfeit signatures) and 5 true signatures. The program detected (classified) 4 true signatures and 5 counterfeits correctly. In other words, all counterfeit signatures are detected correctly. Only one signature is classified as a counterfeit while it was not (i.e. a "false negative"). Obviously, a "false negative" should be more acceptable in comparison to a "false positive", because the person can always be given a second chance to prove that the signature is his/hers. On the other hand, a false positive in verification carries a lot of risk.

7. APPLICATION

This system can be applied in banking sector where 20 specimen signatures of each account holder is taken. Using these signatures neural network is trained and thus we can verify fake and frauds in banks.

8. FUTURE SCOPE

A reliable signature verification system is an important part of law enforcement, security control and many business processes. It can be used in many applications like cheques, certificates, contracts etc.

9. CONCLUSION

This paper presents a method for verifying handwritten signatures by using an ANN architecture. Various static and dynamic signature features are extracted and used to train the

NN. Several Network topologies are tested and their accuracy is compared. The most successful version of the NN based HSV system uses a single MLP with one hidden layer to model each user's signature. It is trained using hundred genuine signatures. Using this approach efficiency of detection increased.

Generally, the failure to recognize/verify a signature was due to poor image quality and high similarity between two signatures. Recognition and verification ability of the system can be increased by using additional features in the input data set.

10. REFERENCES

[1] Signature Recognition and Verification with ANN Cemil OZ,ozc@umr.edu, Sakarya University Computer Engg.,Department, Sakarya, Turkey

[2] Handwritten Signature Verification using Instance Based Learning Priya Metri , Ashwinder Kaur International Journal of Computer Trends and Technology- March to April Issue 2011

[3] A Preliminary Study on Various Off-line Hand Written Signature Verification Approaches Meenakshi S AryaJaypee ,University of Information Technology©2010 International Journal of Computer Applications (0975 – 8887)Volume 1 – No. 9

[4] Statistical Approach for Offline Handwritten Signature Verification, Debnath Bhattacharyya, Samir Kumar Bandyopadhyay, Poulami Das, Journal of Computer Science 4 (3): 181-185, 2008

[5] J. -P. Drouhard, R. Sabourin and M. Godbout, A neural network approach to off-line signature verification using directional PDF

[6] Simon Haykin, Neural Networks A Comprehensive Foundation

[7] Introduction to Digital Image Processing with MATLAB (Hardcover) by Alasdair McAndrew (shelved 1 time as image-processing) published 2004

[8] Fundamentals of Digital Image Processing (Prentice Hall Information and System Sciences Series) by Anil K. Jain (shelved 1 time as image-processing) published 1988