

Handwritten Numeral Recognition Using Wavelet Transform and Neural Networks

Bubakir Almuttardi, Abdulsalam Ambarek, Khadija Alshari,
University of Benghazi/Electrical and Electronic Department
Benghazi, Libya

abdulsalam_ambarek@yahoo.com bubakir_m@yahoo.com.uk k_alshari@yahoo.com

ABSTRACT:

Character recognition is one of the most popular practical applications of pattern recognition. The first step is to acquire a digital image. The next step is acquiring to improve it for the other processes. The next stage is segmentation, the pattern recognition problem can be represented into data acquisition, preprocessing, feature extraction, and classification. The wavelet transform is used in including image processing [1][2]. For a given image the wavelet transform produces a low frequency sub band image reflecting its basic shape and three sub band images that contain the high frequency components of the image at horizontal, vertical and diagonal directions. The artificial neural networks can be successfully used In this paper we will use wavelet coefficients as features in numeral recognition problem, these features are fed to MLP network which used as a classifier. We will train and test MLP networks by using different types of discrete wavelet transform which applied to a data base of 960 samples of

handwritten numerals. The goal of this paper is to demonstrate the strengths of wavelet transform as a tool of feature extraction combined with a neural network as a classifier in the field of numeral recognition.

keywords :

pattern recognition,
character recognition ,segmentation ,
feature extraction , classification ,
wavelet transform ,neural network

1. INTRODUCTION:

Recognizing characters is a problem that at first seems extremely simple but it's extremely difficult in practice to program a computer to do it. Figure 1 shows the different families of character recognition. Two families are included in the general term of character recognition:

- On-line character recognition
- Off-line character recognition

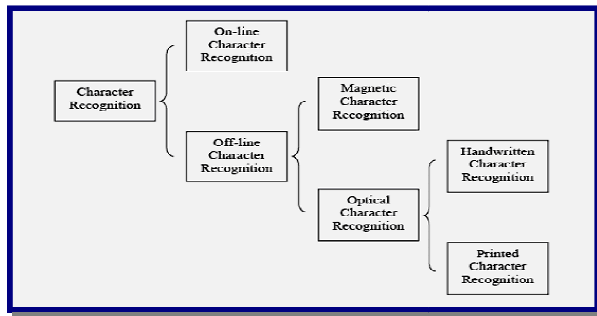


Figure 1 Different Families of Character Recognition

On-line character recognition deals with a data stream which comes from a transducer, while the user is writing. Off-line character recognition is performed after the writing is finished. The major difference between on-line and off-line character recognition is that on-line character recognition has time-sequence contextual information but off-line data does not. This difference generates a significant divergence in processing architectures and methods.

The off-line character recognition can be further grouped into:

- Magnetic Character Recognition (MCR)
- Optical Character Recognition (OCR)

In MCR, the characters are printed with magnetic ink. The reading device can recognize the characters according to the unique magnetic field of each character. The OCR deals with the recognition of characters acquired by optical means, typically a scanner or a camera. The

characters can be either printed or handwritten, of any size, shape, or orientation. The objective of an OCR system is to recognize alphabetic letters, numbers, or other characters, which are in the form of digital images, without any human intervention. This is accomplished by searching a match between the features extracted from the given character's image and the library of image models. At the same time, we also want the features to be robust enough so that they will not be affected by viewing transformations, noises, resolution variations and other factors. Figure 2 illustrates the basic processes of an OCR system.

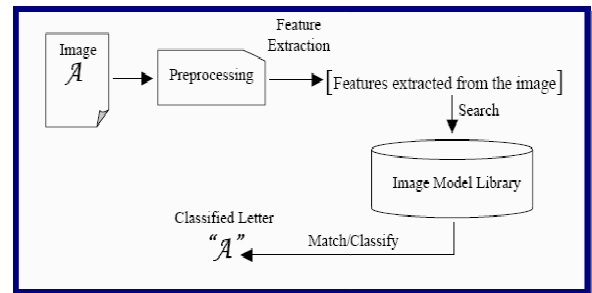


Figure 2 The basic processes of an OCR system.

The OCR can be subdivided into handwritten character recognition and printed character recognition. Handwritten character recognition is more difficult to implement than printed character recognition due to the diversified human handwriting styles and customs. [3]. The

design of character recognition system requires careful attention to the following issues: definition of pattern classes, sensing environment, pattern representation, feature extraction selection, classifier design learning, selection of training test samples, and performance evaluation. Figure 3 illustrates the main elements of pattern recognition system.

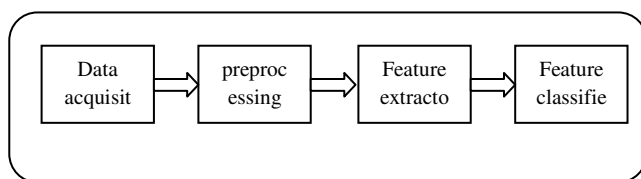


Fig 3 The main elements of pattern recognition system.

a) Data acquisition

The first stage in character recognition which is employed at the front end of the system to transform patterns from their original domain to a representation acceptable by the reminder of the system. The output of the transducer is generally a collection of digital information.

b) Preprocessing

The preprocessing is the next stage, digital filtering, image thresholding, normalization, and contour detection are some examples of preprocessing.

c) Feature extraction:

The two most important steps in pattern recognition systems are feature extraction and classifier design. Feature extraction can be defined as the conversion of a set of measurements in order to form a subset that can be useful in a classification module. Feature vector consisting of a number of features. These features can be represented as numerical quantities [4]. Feature extraction is the heart of a pattern recognition system. In the feature extraction phase, only the salient features necessary for the recognition process are retained such that the classification can be implemented on a vastly reduced feature space. Feature extraction can be viewed as a mapping which maps a pattern space into a feature space using wavelet transform.[5].

d) Classification:

The last stage is the classifier, which transfer the feature vectors into different pattern classes using neural networks. There are many types of classifiers; selecting classifier depends on many factors like the intended application, and feature extraction method used.

2. WAVELET TRANSFORM:

The wavelet transform decomposes a signal into a set of wavelet basis functions, or just “wavelets” for short, that are localized in time. Therefore signals with short bursts can be reconstructed with a much smaller set of wavelet basis functions. The wavelet transform has two great events:

- A class of wavelet bases is constructed, which is smooth, compactly supported and orthonormal. They are referred to as Daubechies bases that are applied to many fields as filter banks [6].
- Multiresolution analysis (MRA) is intrinsically consistent with sub-band coding in signal analysis. It is showed that wavelet series expansions could be implemented with filter banks.

In wavelet analysis the signals can be locally characterized in both time domain and frequency domain simultaneously and self-adaptively. the wavelet theory consists of two parts, namely the wavelet transform and the wavelet basis. The families of functions $\Psi_{a,b}$ Generated from one single function by the operation of dilation(a) (The variable a reflects the scale (width) of a particular basic (function), and translation (b)(specifies its translated position along

the X - axis, such families called “Wavelets”. $\Psi(x)$ is called mother wavelet or basic wavelet ,which is like the sine basis in the Fourier transform.

$$\Psi(x)=|a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \quad a, b \in \mathbf{R}, a \neq 0 \quad (1)$$

The function $\Psi(x)$ is a real function whose Fourier transform $\Psi(\omega)$ satisfies the admissibility criterion [6] [1].

$$C_{\Psi} = \int \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (2)$$

Where the Fourier transform $\Psi(\omega)$ is

$$\Psi(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-i\omega x} \Psi(x) dx \quad (3)$$

The “admissibility” of the previous equation implies, if Ψ has sufficient decay, always assumed in practice that Ψ has zero mean.

$$\int \Psi(x) dx = 0, \quad \Psi(0) = 0, \Psi(\infty) = 0 \quad (4)$$

Typically, the function Ψ will therefore have at least some oscillations [13]. Figure (4) illustrates some typical shapes of wavelet by different scales and translations.

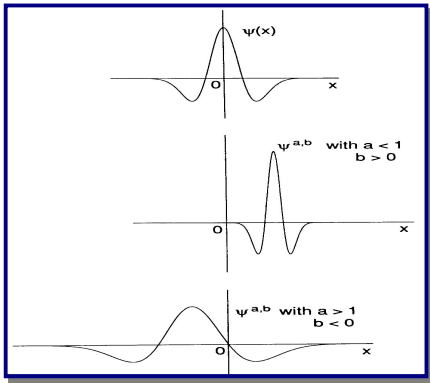


Figure (4) Typical shapes of wavelets.

3. IMAGE ANALYSIS USING TWO DIMENSIONAL WAVELET :

Mallat suggested the representation of one dimensional wavelet to two dimensional wavelet, where Mallat uses this extension for two dimensional wavelet transform. [7][8].

The two dimensions scaling function is separable defined as:

$$\phi(x, y) = \phi(x) \phi(y) \quad (5)$$

Where $\phi(x), \phi(y)$:are one dimensional scaling functions.

The three two-dimension wavelets functions are separable defined as:

$$\begin{aligned} \psi^h(x, y) &= \phi(x) \psi(y) \\ \psi^v(x, y) &= \psi(x) \phi(y) \\ \psi^d(x, y) &= \psi(x) \psi(y) \end{aligned} \quad (6)$$

Where: h: horizontal direction, v: vertical direction, d: diagonal directions and $\psi(x), \psi(y)$ are one dimensional wavelet functions.

4. NEURAL NETWORKS:

Neural networks can be trained to solve problems that are difficult for conventional computers or human beings. Neural networks have been trained to perform complex functions in various fields of applications including pattern recognition, identification, classification, speech, vision, and control systems. [9][10]. A neural network is a collection of interconnected identical nodes, or Processing Elements PE (neurons) each of which is relatively simple in operation. Each PE receives a number of inputs from several of the “upstream” PEs in the network, generates a scalar output, and sends it “downstream” to another group of PEs. Associated with each input is a weight, and the output (in most cases) is then a function of the weighted sum of inputs. The interconnection scheme, or network architecture, is one of the major design choices. The PEs are commonly organized into layers. The number of PEs in each layer is a design choice. In some networks each PE in one layer receives input from every PE in the previous layer

and sends its output to every PE in the subsequent layer. The first layer is called the input layer, the final one is called the output layer, and all other layers are termed hidden layers.[9][11].

The output of the neuron a is defined as :

$$\mathbf{a} = f(\mathbf{w}p + \mathbf{b}) \quad (7)$$

where p is the scalar input multiplied by the scalar weight w to form wp , the other input 1 is multiplied by a bias b , f is the transfer function of the output a . w and b are both adjustable scalar parameters of the neuron. The basic {central} idea of neural networks is that such parameters can be adjusted so the network exhibits some desired behavior. Thus, we can learn (train) the network to do a particular job (task) by adjusting these parameters.

A feed-forward neural network is a multilayer neural net with input layer, output layer, and hidden layers. Each neuron in each layer receives inputs from every neuron in the previous layer and sends its output to every neuron in the subsequent layer. [12][13].

The back propagation neural network “learns” a predefined set of input-output example pairs which is known as the training set. This is accomplished in a three

phases. The first phase is the forward propagation of the input that has been applied to the first layer of the network units. The output is then compared to the desired output; an error signal is compared for each output unit. In the second phase, the error signals are transmitted backwards from the output layer to each node in the intermediate layer such that each node receives only a portion of the total error. The weights are then updated in the third phase based on the error at each node. In this way the network will converge toward a state that allows all the training sets to give minimum error. [14].

5. EXPERIMENTAL WORK:

In this part a handwritten numeral recognition system was implemented, with Wavelet coefficients as features, and MLP Network (with Back-Propagation training) as a classifier. The system was trained and tested on a database of 960 samples of handwritten numerals of different handwritten styles. The recognition system was simulated on a personal computer by writing programs in the Matlab technical language. Figure 5 shows the main elements of the Numeral recognition system used in this project.

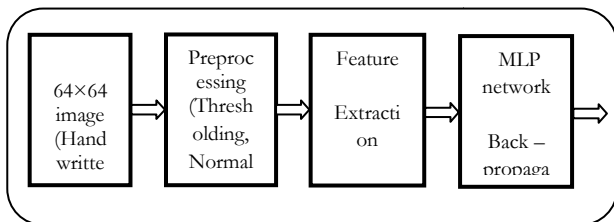


Figure 6 The main elements of Numeral recognition system used

5.1. Database Description

The used database consists of 960 handwritten numerals of various styles and pen thicknesses. These were written by people in different age groups including primary school children, several students and staff members in the Faculty of Engineering, Benghazi University [15]. There are 96 samples of each numeral, each sample is 64x64 gray level image. This database was divided into three sets, Training set which contains 500 samples to train the MLP; Validation set which contains 100 samples to stop the training when the validation error reaches its minimum; Test set which contains 360 samples to evaluate the system.

5.2. Feature Extraction(wavelet transform)

Wavelet decomposition algorithm is applied to the normalized image recursively for three times to obtain 8x8 smooth approximation of the original image as the final decomposition. These 64 coefficients are fed to the input layer of the MLP network. Figure 6 shows one level of

wavelet transform applied to the image of number four.

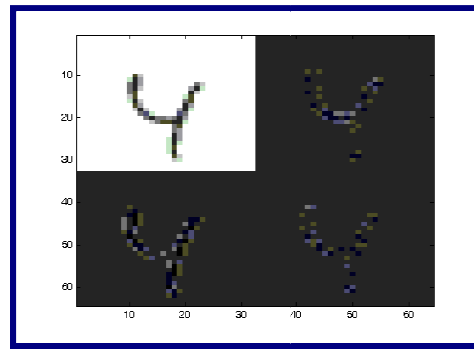


Figure 6 One level of wavelet transform applied to the image number.

A particular family of wavelets is specified by a particular set of numbers called Wavelet filter coefficients. We will use some of famous families of these wavelet transforms.

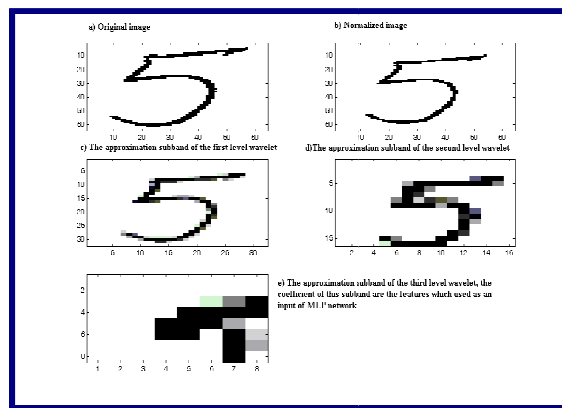


Figure 7 shows the stages that one image (from database) passed through the system.

Figure 7 the stages that one image (from database) passed through the system, a)Original image , b)Normalized image, c) the approximation sub band of the first level of wavelet transform, d)

approximation sub band of the second level of wavelet transform, e) approximation sub band of the third level of wavelet transform.

5.3. Feature Classification (Neural network)

The MLP network with the following architecture and parameters was used as a classifier:

- Input layer of 64 neurons (number of wavelet coefficients).
- One hidden layer of 64 neurons.
- Output layer of 10 neurons (number of classes or numerals).
- Sigmoid activation functions for all layers.
- Weights initialized between [-1, 1] using Nguyen-Widrow initialization.
- Back propagation learning algorithm.
- Termination of training: An effective strategy of judging training adequacy is the use of a validation set. With increased training, the recognition error on the validation set will decrease monotonically to a minimum value but then it starts to increase, even if the training error continues to decrease. For better network performance, training is terminated when the validation error reaches its minimum. In our simulations,

we considered about 10% of the data as the validation set.[16].

6. EXPERIMENTAL RESULTS:

6.1.Experiment 1

In this experiment we used one of the most famous wavelet transform families called Daubechies' wavelet; the Daubechies4 (db4) which has only four coefficients for low pass L (smoothing filter) and high pass filter H.

Low pass = [-0.1294, 0.2241, 0.8365, 0.4830], High pass = [-0.4830, 0.8365, -0.2241, -0.1294], After applying the wavelet filters to the training validation and test sets, we fed the MLP network by 64 coefficients of wavelet to train and test the system.

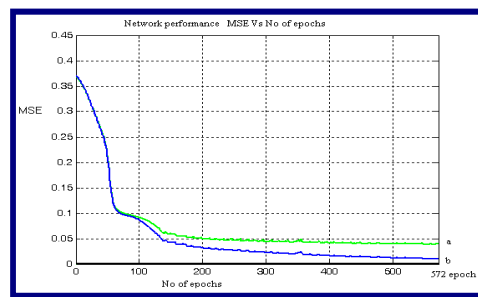


Figure 8 shows the MSE of both training and validation sets vs. No. of training epochs.

a) the MSE of the validation set, b) the MSE of the training set.

The results of the previous experiment can be seen in the confusion matrix in table (1), where the row represents the input class

and the column represents the recognition result.

Table (1) the Confusion matrix of experiment 1.

	0	1	2	3	4	5	6	7
0	100	0	0	0	0	0	0	0
1	0	100	0	0	0	0	0	0
2	0	0	97.22	0	0	0	0	2.78
3	0	0	0	100	0	0	0	0
4	0	0	0	0	94.44	0	2.78	2.78
5	0	0	0	0	0	100	0	0
6	2.78	0	0	0	0	0	94.4	0
7	0	0	2.78	0	2.78	0	0	94.44

The Recognition Rate = (sum of diagonal elements) / 10 = 97.22%.

6.2. Experiment 2:

In this experiment we used a biorthogonal wavelet transform filter, which has the following coefficients for low pass L (smoothing filter) and high pass filter H.

Low pass = [0, 0, 0, -0.0707, 0.3536, 0.8485, 0.3536, -0.0707, 0, 0], High pass = [0, -0.0152, 0.0758, 0.3687, -0.8586, 0.3687, 0.0758, -0.0152, 0, 0], After applying the wavelet filters to the training, validation and test sets, we fed the MLP network by 64 coefficients of wavelet to train and test the system. Figure 9 shows the MSE of both training and validation sets vs. No. of training epochs.

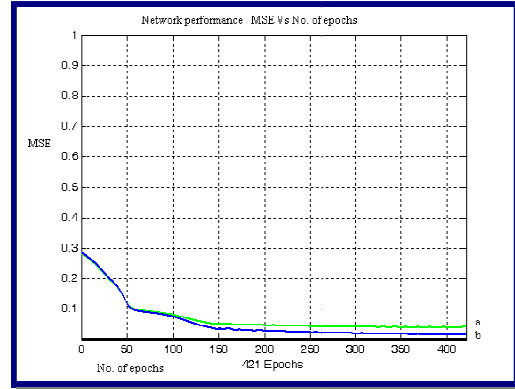


Figure 9 the MSE vs. No of training epochs,

a) MSE of the validation set b) MSE of the training set

Table (2) the Confusion matrix of experiment 2

	0	1	2	3	4	5	6	7
0	10 0	0	0	0	0	0	0	0
1	0	10 0	0	0	0	0	0	0
2	0	0	97.2 2	0	0	0	0	2.78
3	0	2.7 8	0	94.4 4	0	0	0	0
4	0	0	0	0	97.2 2	0	0	0
5	0	0	2.78	0	0	94.4 4	0	0
6	2.7 8	2.7 8	0	2.78	0	0	91.6 7	0
7	0	0	2.78	2.78	2.78	0	0	91.6 7

The Recognition Rate = (sum of diagonal elements) / 10 = 95.56%

6.3. Experiment 3:

In this experiment we used another kind of biorthogonal wavelet transform filter, which has the following coefficients for

low pass L (smoothing filter) and high pass filter H.

Low pass = [0, 0.0378, -0.0238, -0.1106, 0.3774, 0.8527, 0.3774, -0.1106, -0.0238, 0.0378],

High pass = [0, -0.0645, 0.0407, 0.4181, -0.7885, 0.4181, 0.0407, -0.0645, 0, 0],
 After applying the wavelet filters to the training, validation and test sets, we fed the MLP network by 64 coefficients of wavelet to train and test the system. Figure 10 shows the MSE of both training and validation sets vs. No of training epochs.

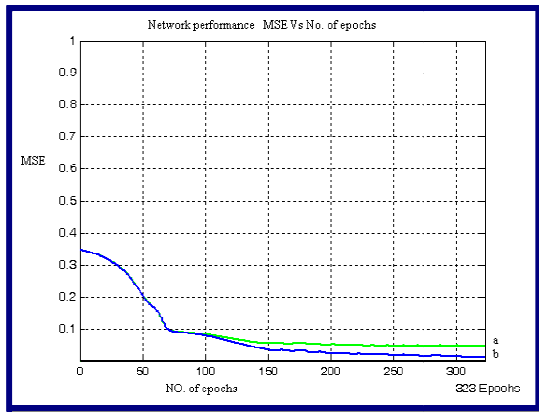


Figure 10 the MSE of both training and validation sets vs. No of training epochs, a) the MSE of the validation set b) the MSE of the training set

Table (3) the Confusion matrix of experiment 3

	0	1	2	3	4	5	6	7
0	10 0	0	0	0	0	0	0	0
1	0	10 0	0	0	0	0	0	0
2	0	0	91.6 7	0	2.7 8	0	0	2.78
3	0	0	0	97.2 2	0	0	0	0
4	0	0	0	0	94. 4	0	2.78	0
5	0	0	2.78	0	0	94.4 4	2.78	0
6	2.7 8	0	0	0	0	0	94.4 4	0
7	0	0	0	5.56	2.7 8	0	0	91.6 7

The Recognition Rate = (sum of diagonal elements) / 10= 95%.

6.4. Experiment 4:

In this experiment we used another kind of biorthogonal wavelet transform filter, which has the following coefficients for low pass L (smoothing filter) and high pass filter H.

Low pass = [0 0.0331, -0.0663, -0.1768, 0.4198, 0.9944, 0.4198, -0.1768, -0.0663, 0.0331],

High pass = [0, 0, 0, 0.3536, -0.7071, 0.3536, 0, 0, 0, 0],
 After applying the wavelet filters to the training, validation and test sets, we fed the MLP

Figure 11 shows the MSE of both training and validation sets vs. No of training epochs.

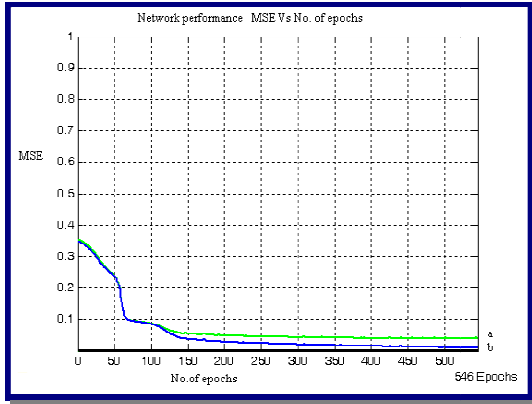


Figure 11 MSE vs. No of training epochs

a) MSE of the validation set , b) MSE of the training set

Table (4) the Confusion matrix of experiment 4.

	0	1	2	3	4	5	6	7
0	10 0	0	0	0	0	0	0	0
1	0	97.2 2	2.78	0	0	0	0	0
2	0	0	97.2 2	0	0	0	0	2.78
3	0	0	0	97.2 2	0	0	0	0
4	0	0	0	0	94.4 4	0	2.78	0
5	0	0	2.78	0	0	97.2 2	0	0
6	2.7 8	0	0	2.78	0	2.78	91.6 7	0
7	0	0	0	2.78	0	0	0	97.2 2

The Recognition Rate = (sum of diagonal elements) / 10= 95.28% .

7. CONCLUSION:

In a handwritten character recognition system one important work is to extract features, if we select the suitable feature, it will compress the useless information of the pattern and remain the meaningful information. Therefore, as the choice of fine features with stable ability to represent pattern features with little respect to shape

variation and writing style, as reaching the crucial point to improve the performance of the handwritten numerals recognition.

In this paper we proposed feature extraction method based on two-dimensional discrete wavelet transform for off-line recognition of unconstrained handwritten numerals using back propagation neural networks as a classifier. We used the wavelet transform because it gave us space and frequency information at the same time. This means most of the energy of the signal is well represented by a few expansion coefficients.

For a given image the wavelet transform produces a low frequency sub band image reflecting its basic shape and three sub band images that contain the high frequency components of the image at horizontal, vertical and diagonal directions. These components can be used to construct the feature vector in a recognition system. Two-dimensional one- level discrete wavelet transform (DWT) can be described in terms of filter banks. The calculation of the coefficients from the signal can be done efficiently. The experimental results showed that wavelet analysis is an efficient transform.

Several experiments on the wavelet filters were conducted to improve the efficiency of this system. The recognition rates of the test

sets are 97.22 %, 95.56%, 95%, 95.28% respectively. We conclude the following:

- The size of the database has a major effect in the learning performance. A large database means a bigger samples with much more variations that when captured will more likely to cover different writing styles and any actual real life outcome.
- In a practical situation handwritten numerals might come translated, scaled and/or rotated. In this work normalization has been used successfully to overcome the first two problems. Rotation remains to be a problem as can be seen from the results for (6, 9) numerals. In the worth mentioning that Daubechies4 wavelet seems to have a superior performance when it comes to this point, it means that its features represented the numerals much more accurately. However, this rotation invariance issue might not make a difference for some applications where the numerals will inherently be limited in their rotation. As an example, filling applications where numerals will be confined in boxes for instance.
- Although wavelet method might be difficult to learn as a mathematical tool, but its computation is a bit easy to

implement and it's not computationally intensive. In other words, wavelet as a feature-extraction tool, fits naturally with digital computer with its basis functions defined by just multiplication and addition operators. There are no derivatives or integrals.

- Note also the data visuality of transformed images will show only at a certain level of decomposition.
- The experimental results show that the proposed methods are simple and efficient representation for unconstrained handwritten numerals recognition using fewer image preprocessing.

REFERENCES:

- [1] Ingrid. Daubechies," Ten Lectures on Wavelets", CBMS-NSF Regional Conference Series In Applied Mathematics 1999 Sixth Printing.
- [2] Stephane.G.Mallat," Multifrequency Channel Decompositions of Images and Wavelet Models", IEEE PAMI Vol. 37,N0.12, December .1989.
- [3] Lasse Holmstr"om, Petri Koistinen, Jorma Laaksonen, and Erkki Oja, "Neural and Statistical Classifiers—Taxonomy and Two Case Studies" IEEE Transactions on Neural Networks, Vol. 8, No. 1, January 1997.

- [4] Yuki C. Yunes Saito "Algorithms for Image Classification Using Wavelets and Fusion Methods" A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering University of Puerto Rico Mayagüez Campus 2005.
- [5] Yuan.Yan.Tang,l,h, Yang ,and H.Ma," Wavelet Theory and ITS Application to Pattern Recognition", World Scientific 2000.
- [6] Ingrid.Daubechies," Orthonormal Bases of Compactly Supported Wavelets", Communication on pure and Applied Mathematics, Vol. XLI 909-996, 1988.
- [7] Stephane.G.Mallat," A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE PAMI vol. 11, NO. 7 July 1989.
- [8] Yves.Meyer," Wavelets and operators", Cambridge studies in advanced mathematics, 1990.
- [9] Millan S. VALCAV H. And ROGER B," Image Processing Analysis, and Machine Vision",1999 pws 2.^{ed}
- [10] Martin.Vetterli and Jelena.Kovacevic," Wavelets and Subband Coding", Prentice Hall 1995.
- [11] Kenneth.R. Castileman," Digital Image Processing", Prentice Hall 1996.
- [12] Síma Jirí, "Introduction to Neural Networks", Institute of Computer Science, Academy of Sciences of the Czech Republic, Technical report No. V-755, August 10, 1998.
- [13] Sinha Anshu, "An Improved Recognition Module for the Identification of Handwriting Digits," B.Sc in computer Science and Engineering and Master of Engineering Department of Electrical Engineering and computer Massachusetts Institute of Technology, 1999.
- [14] Laurene V. Fausett Fundamentals of Neural Network. Englewood Cliffs, N.J.: Prentice-Hall, Inc 1994.
- [15] Hatem M.R. Abou-zeid, Ammar A. Al-khatib, Supervised by Mr.Akrem Saad M.Elgazal "Computer Recognition Of Handwritten Numerals", Submitted For The Degree Of B.S.C In Electrical And Electronics Department, Garyounis University.
- [16] Khotanzad A., Hong Y.H. "Invariant image recognition by Zernike moments," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12(5), pp. 489-497, 1990.