Australian Journal of Basic and Applied Sciences, 6(7): 125-130, 2012
ISSN 1991-8178

Handwritten Numerals Recognition, New Approach

Mazin Al-Hadidi

Department of Computer Engineering, AlBalqa' Applied University,Faculty of Computer Engineering, Salt, Jordan Al-Salt, Jordan 19117, P.O Box 7181

Abstract: There are many important applications for recognizing the handwritten numerals, one of the most important and significant application is the automation of postal services especially in countries that use different languages at the post offices, such as India. The purpose of this study is to develop a recognition method for recognizing the handwritten numerals using neural networks and wavelets transform. This study is going to present a method that uses a directional continuous two dimensional wavelet transform besides the neural networks. The transform will be used as an extraction technique of the proper features that are required for the classification before using them as an input to the feed-forward neural network. The results are promising and significant improvements could be achieved to enhance the recognition performance. A recognition rate of 90.3% could be achieved.

Key wrods: Neural Networks, Character Recognition, Wavelet Transform.

INTRODUCTION

The main purpose of the handwritten numeral recognition is to find the possible applications areas for automations, for example postal automations, (Asthana *et al*., 2011; Pal *et al*., 2009), filling forms and bank automation (Chaudhari, 2007; Marinai and Fujisawa 2008)

Pattern recognition field has the challenge to automatically classify the handwritten digits and this area has many intricacies.

To assign an anonymous object to 10 or more predefined classes is a challenging problem to solve, because the objects from the same class could be highly different, on the other hand objects from different classes could be very similar.

Among the fields that receive special attention nowadays is the character recognition. There are many approaches; one of them is the extraction of invariant features to translation, rotation and scaling.

Wavelet transform is considered an efficient tool especially for image-processing applications. It has proved a good result for edge detection (Thomas, 2008) and texture identification (Liapis and Tziritas, 2004).

Before the processing of the digit recognition starts, a one-dimensional discrete wavelet which is dyadic orthogonal is applied to the extracted contour, which can be represented with 2 vectors x and y (Jangal *et al*., 2009; Chen *et al*., 2008; Asthana *et al*., 2011).

A multi-wavelet discrete one-dimensional transform has been applied to the contour in (Chen *et al*., 2003).

The Discrete Wavelet Transform (DWT) disassemble the image into pieces with different orientations and resolutions; (Jangal *et al*., 2009; Koko, 2009). It is mostly used for image compression (Skodras *et al*., 2001), but it is not a translation variant. However, the Continuous Wavelet Transform (CWT), is a translation variant, that gives a redundant representation of an image. CWT is mainly used for image analysis and it has been modified to construct directional wavelet transform (Du1 *et al*., 2006), that is achieved by giving one main orientation to the wavelet, by stretching one of its axes and by adding the rotational angle.

Wavelet transform:

The wavelet transform shows specific features such as scalability, orthogonality, separability and multiresolution capability.

The discrete wavelet transform, (eq.1) (Gonzalez and Woods, 2006) of an image g (a, b) of size X×Y is:

$$W\varphi\left(j_0, x, y\right) = \frac{1}{\sqrt{XY}} \sum_{a=0}^{X-1} \sum_{b=1}^{Y-1} f\left(a, b\right) \varphi_{j_{0,}\, x, y(a,b)} \quad (1)$$

Corresponding Author: Mazin Al-Hadidi, Department of Computer Engineering, AlBalqa' Applied University,Faculty of Computer Engineering, Salt, Jordan Al-Salt, Jordan 19117, P.O Box 7181
E-mail: trueamman@yahoo.com; Mobile :(+962)777131163

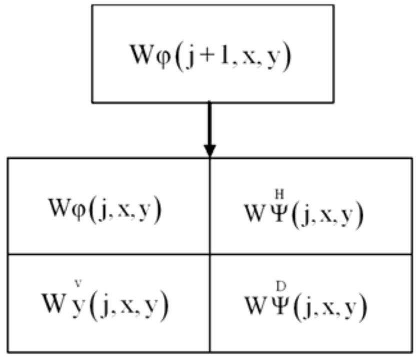

Fig. 1: Single level wavelet decomposition

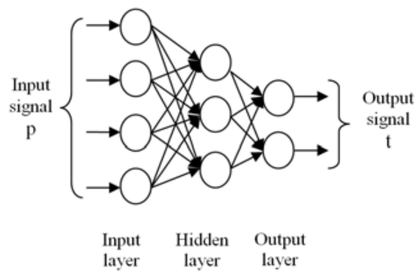

Fig. 2: Example of two layer FEED-Forward neural network architecture

$$w_{\Psi}^{i}\left(j_0,x,y\right)=\frac{1}{\sqrt{XY}}\sum_{a=0}^{X-1}\sum_{b=1}^{Y-1}f\left(a,b\right)\varphi^{i}_{j_{0,}x,y(a,b)} \tag{2}$$

Where:

$$\varphi_{j\,x,y(a,b)}=2^{j/2}\varphi(2^{j}a-x,2^{j}b-y,) \tag{3}$$

And:

$$\Psi^{\iota}_{j\,x,y(a,b)}=2^{j/2}\Psi^{\iota}(2^{j}a-x,2^{j}b-y,) \tag{4}$$

Are two dimensional functions, scaling and wavelet respectively, the index i represents the directional wavelet which takes the values, such as horizontal, vertical and diagonal details.

j_0 represents an arbitrary starting scale, while the W $\varphi(j_0,x,y)$ coefficients define an approximate value of g(a, b) at scale j_0. The W w^i (j_0,x,y) coefficients represents the horizontal, vertical and diagonal information for scales $j \geq j_0$. Ordinarily, $j_0 = 0$, X = Y = 2J so that j = 0, 1, 2…, J-1 and m, n = 0, 1, 2, …, 2j-1.

Down samplers and digital filters can be used to implement the discrete wavelet transform.

The image's information about the high frequency with vertical orient is characterized using the detail component. While the approximation component includes its low-frequency.

The sub images can be then filtered and sampled to produce four output images with equal sizes, Fig. 1 shows these sub images.

Neural networks:

One of the popular methods that are used in the optical character recognition is the artificial neural networks which are used as classifiers and feature extractors because of their potentials in learning and generalization.

Usually neural networks are used in character recognition applications, choosing from multiple architectures, the Multiple Layer Perception (MLP) is commonly used.

The MLP can be considered a completely connected network, consisting of different layers, each layer is consisted from multiple neurons, every neuron is connected to every neuron in the other layer by a weighted link, by this link the neuron state is transmitted.

This network represented by the MLP is composed of three layers, the input layer, the output layer and the hidden layer. Figure 2 shows such a network.

The input layer provides the input signal to the neurons in the hidden layer. It behaves the same way as a weight on a connection from a unit which has an activation value of 1 all the time. The hidden layer contains a nonlinear activation function for each neuron.

Connection weights are observed during the training phase. The weighted sum of the connected nodes in the previous layer forms the output.

This operation is explained in (Hagan *et al*., 2011).

$$x^{n+1} = f^{n+1}(w^{n+1}x^{n} + y^{n+1}) \tag{5}$$

For n = 0, 1,…, N-1, where N is the number of layers in the network. The neurons in the first layer receive external inputs:

$$x^{0} = p \tag{6}$$

Which provides the starting point for the network. The networks outputs are the outputs of the neurons in the last layer:

$$x^{N} = t \tag{7}$$

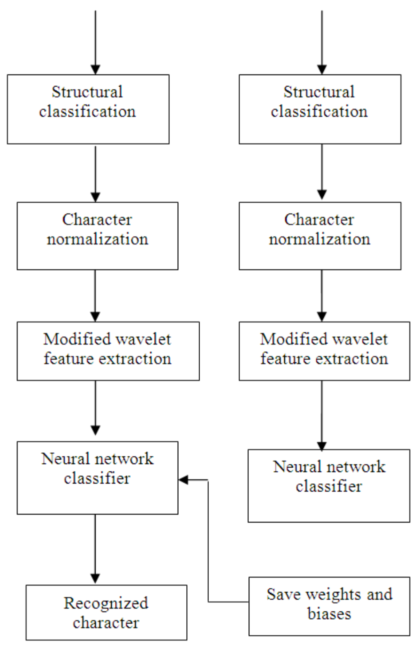

Fig. 3: Block diagram of the system used

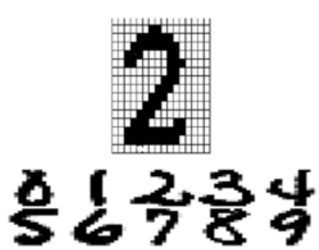

Fig. 4: Samples of handwritten numerals and grid

As soon as the networks weights are intialized, the traning phase is ready in the network. The extraction of meaningful features from the input vectors helps the hidden neurons to learn complex tasks.

The procedure for modifying weights of a network, is called a learning rule. The aim of the learning rule is to train the network to accomplish the task of pattern recognition.

Through the training process, the weights of the network are modified iteratively to minimize error. The training process requires a collection of examples for the proper network behavior, network inputs p and target outputs t. at the time the input is applied to the network, the output of the network is compared to the target.

The difference between the network output and the target output represents the error.

Minmizing the average sum of these errors is the purpose.

Each application requires different parametes for the optimal performance, these parameters could be, the activation function, number of neurons in the hidden layer and the network training function.

Recognition System:

The block diagram in Figure 3 shows the proposed recognition system of digits characters.

The data flow during recognition is on the left side, while the data flow during training is on the right side. The handwritten character patterns form the input to the system. When the system receives the inputs, the main coordinates top, bottom, left and right are calculated.

Figure 4 shows a digit captured in a grid, the character pixels are sensed in grid boxes, the digit is then digitized in a binary string, this string is applied for recognition and training.

A grid of size 16×16, (16 rows and 16 columns) was implemented in the experiments in a way that the aspect ratio for images is preserved.

After that, a directional two dimensional continuous wavelet transform is applied on each image(Figure 5). The recognition system is implemented using a feed-forward neural network.

A database containing 2000 unconstrained handwritten numerals is used as the source of our experiments; the database was collected from examination paper of the student of faculty of engineering at Al Balqa'a University.

Digits with different writing styles as well as different sizes and stroke widths were taken from our database. Some of the numerals are very complicated to recognize with human eyes.

The proposed system of handwriting numeral recognition, there are three steps for feature extraction: extreme coordinates measurement, holding characters into the grid and character digitization.

The handwritten character is grabbed by the extreme coordinates, from top-bottom and left-right and subdivided into rectangular grid that contains specific rows and columns(Figure6,7). The algorithm adjusts the size of the grid as necessary and adjusts its constituents, depending on the dimensions of the character.

After that it searches the existence of character pixels in each box of the grid.

Fig. 5:(a) An original sample digit 0. (b) Same after inverting color (c) Same after CWT- preprocessing

Fig. 6:(a) An original sample digit 2. (b) Same after inverting color (c) Same after CWT- preprocessing

Fig. 7:(a) An original sample digit 5. (b) Same after inverting color (c) Same after CWT- preprocessing

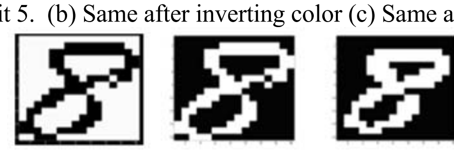

Fig. 8:(a) An original sample digit 8. (b) Sample after inverting color (c) Sample after CWT-preprocessing

Boxes found within character pixels are marked “on” and the rest are marked “off”.

To locate the “on” and “off” boxes, a binary string of every numeral is formed, in a process named numeral digitization, this binary string is presented to the neural network as an input for recognition and training purposes.

The number of binary inputs is represented by the total number of grid boxes.

Therefore a 16×16 grid generates 256 inputs to the recognition model. In another words a 16*16 grid provides a 256 dimensional input feature vector. The developed software displays this phenomenon by filling up the intersected squares. Figure 5 through 8 produces the effect.

The training phase includes the modified wavelet features which are extracted from a process of normalizing characters; these features are then used to fix the weights for each character, taking into account sufficient samples for each for each character(figure 8).

Similar features are extracted from the character in the testing phase. Structural classification and pre-processing of the character is done at the beginning. The features will be applied as inputs to the neural network. The neural network outputs the final recognition result.

Pre-processing:

RGB character images are converted to a gray image then these images are converted to binary ones by a point operator.

This operator segregates the pixels that have values within determined range, for example, the object from the background by choosing a threshold that segregates the object from the background.

After normalization, the threshold can be chosen using uniform thresholding, in which, pixels above the threshold are set to a specific color i.e. white and pixels below are set to black.

Knowledge of the gray levels is required in the uniform thresholding, without this knowledge the target features could be not selected or misclassified.

A testing was made for the handwritten characters for various threshold values before finalization.

Pre-processing is playing a significant role in the recognition of handwritten characters, as well as in any other pattern recognition task.

Handwritten characters might produce undesired effects, i.e. strokes, gaps, breaks, which happens during binarization.

A handwritten character might exhibit lesser width at curves than other parts of the character.

Most probably this point will break during binarization.

As mentioned before, because the image is inverted, character will have a white color, that's the reason the algorithm for character recognition is very simple.

It starts searching from the left to the right for white pixels starting from the left top corner of the area determined for writing.

The presence of a character can be inferred from a trace of white pixel.

Results:

To be able to evaluate the performance of our proposed system, many experiments are done on 6000 unconstrained handwritten numerals.

Table 1: Results of the test set

Digit	Correctly classified	Recognition (%)
0	184/200	92.0
1	193/200	96.5
2	183/200	91.5
3	169/200	84.5
4	190/200	95.0
5	170/200	85.0
6	182/200	91.0
7	181/200	90.5
8	167/200	83.5
9	187/200	93.5
Total	1806/2000	90.3

A bi-level format of the numerals is stored in the database. We use 2000 numeral for testing and 4000 for training.

This section shows results observed after the pre-processed digits in the neural networks in section 3 were classified.

Table 1 lists the results, the first column of this table represents the digit, the second column represents the fraction of classified samples, the last column represents the recognition percentage for the digit.

The last row in the table shows the total of the complete set.

The percentage of correctly classified patterns was 90.30% for the test set. This result is promising and it could be improved. The performance obtained in this study can be compared to other results reported in the literature for the same data set (Koko, 2009) (Asthana, 2011).

Conclusions:

In this study we presented a new digitized method for handwritten numeral character recognition, we used the directional wavelet transform besides and neural networks.

We have presented a pre-processing stage that is based on a two dimensional directional CWT, before the training of a multilayer feed-forward neural network for handwritten numeral classification.

We had an efficient wavelet descriptor for the handwritten numerals.

To decrease the training time, Levenberg-Marquardt (Alma'adeed *et al*., 2004) is used, which is a faster training algorithm.

The modified method of wavelet feature generation boosts the range of the signal to a significant amount, producing a new unique feature for every numeral, consequently increasing the recognition accuracy.

The results of the tested data can be compared with other proposed techniques (Koko, 2009) (Asthana, 2011), which are also based on a wavelet–transform pre-processing step and also adopts the training of a feed-forward neural network for pattern classification. However, these mentioned techniques employ a wavelet or multi-wavelet transform in one dimension; in addition it requires the identification of the contour of the digit, which is not mandatory in our case.

REFERENCES

Alma'adeed, S., C. Higgens and D. Elliman, 2004. Off-line recognition of handwritten Arabic words using multiple hidden Markov models. Knowledge-Based Systems, 17: 75-79.

Asthana, S., F. Haneef, R.K Bhujade, 2011. Handwritten multiscript numeral recognition using artificial neural networks. Int. J. Soft Comput. Eng. ISSN: 1: 2231-2307.

Chaudhari, B.B., 2007. Digital document processing-major directions and recent advances, Springer, London,

Chen, G.Y., T.D. Bui and A. Krzyzak, 2003. Contour-based handwritten numeral recognition using multiwavelets and neural networks. Patt. Recog., 7: 1597-1604.

Chen, G.Y., T.D. Bui and A. Krzy-Zak, 2008. Invariant pattern recognition using radon, dual-tree complex wavelet and Fourier transforms,

Directional continuous wavelet transform applied to handwritten numerals recognition using neural networks,

Du1, P., W.A. Kibbe1 and S.M. Lin, 2006. Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching,

Gonzalez, R. and R. Woods, 2006. Digital Image Processing. Sec. Edn. Pearson Education, India

Hagan, M., H. Demuth and M. Beale, 2011. Neural network design, Thomson learning, India, 2003. International J. Signal Processing, Image Processing and Pattern Recognition, 4: 1.

Jabril Ramdan; Khairuddin Omar, 2011. Comparative study of algorithms for voronoi diagram construction on segmentation of Arabic hand writing . Australian Journal of Basic and Applied Sciences, 5(11): 1653-1667.

Jangal, F., S. Saillant and M. Helier, 2009. Ionospheric clutter mitigation using one-dimensional or two-dimensional wavelet processing,

Koko, I.S., 2009. Member, IAENG and Herman Agustiawan, two-dimensional discrete wavelet transform memory architectures,

Liapis, S. and G. Tziritas, 2004. Color and texture image retrieval using chromaticity histograms and wavelet frames. IEEE Trans. Multimedia, 5: 676-686.

Marinai, S. and H. Fujisawa, 2008. Machine learning in document analysis and recognition, Springer-Verlag, berlin heidelberg,

Pal, U., R.K. Roy, K. Roy and F. Kimura, 2009. Indian multi script full pin-code string recognition for postal automation. Proceeding of the 10th International Conference Document Analysis and Recognition, Barcelona, Spain, pp: 456-460.

Shaikh, N.A., G.A. Mallah and Z.A. Shaikh, 2009. "Character Segmentation of Sindhi, an Arabic Style Scripting Language, using Height Profile Vector, "Australian Journal of Basic and Applied Sciences, 3(4): 4160-4169.

Skodras, A., C. Christopoulos and T. Ebrahimi, 2001. JPEG 2000: The upcoming still image compression standard. Elsevier Pattern Recognition Letters, 22: 1337-1345.

Structural decomposition and statistical description of Farsi/Arabic handwritten numeric characters. Proceeding of the Int'l Conference Document Analysis and Recognition, pp: 237-241.

Thomas, B., 2008. Moeslund image and video processing.