

HAND PRINTED CHARACTERS RECOGNITION USING WAVELET FEATURES AND NEURAL NETWORKS

التعرف على الحروف المطبوعة يدويا باستخدام سمات التحول المويجي

والشبكات العصبية

Dr. I.F. El-Nahry

Dept. of E. Eng., Suez Canal Univ., Port-Said, Egypt.

ملخص البحث

تبسيط لوغاريتمات الإدراك الآلي للحروف المطبوعة يدويا نال جهد هائل من الباحثين. نظم التعرف على الحروف تستطيع تحسين التفاعل بين الإنسان والآلة في كثير من التطبيقات، شاملة تطبيقات آلية المكتب، الأعمال ودخول البيانات. هذا البحث يقدم استخدام التحول المويجي wavelet ثنائي الأبعاد كمستخلص للسمات التي يتم تغذيتها الى الشبكات العصبية الأصبطناعية للتعرف على الحروف اللاتينية المطبوعة يدويا. التجربة لمراجعة كفاءة النظام تم أداءها. الطريقة المقترحة يمكن أن تقسم إلى ثلاث خطوات أساسية: الخطوة الأولى هي المعالجة الأولية التي عليها الصورة الأصلية حيث تحول إلى صورة رقمية باستخدام ماسح ضوئي بـ 300 نقطة في البوصة، ثانياً، استخلاص السمة باستخدام الموجات. وفي النهاية، الشبكة الذكية العصبية متعددة الطبقات الأصبطناعية استخدمت للتعرف على الحروف.

Abstract

Simplifying automatic recognition algorithms for hand-printed characters attracted immense research efforts [1-3]. Character recognition systems can improve the interaction between man and machine in many applications, including office automation, business and data entry applications. This paper introduces the use of bi-dimensional wavelet as features extractor that is feed to Artificial Neural Networks (ANNs) for recognition Latin hand-printed characters. An experiment to verify the efficiency of the system was performed. The proposed technique can be divided into three major steps: the first step is pre-processing in which the original image is transformed into a digitized image utilizing a 300 dpi scanner. Second, feature extraction using wavelets. Finally, multilayer artificial neural network is used for characters recognition.

Keywords : Pattern Recognition, Wavelet, Feature Extraction, Neural Network.

1- Introduction

Today, hand-printed characters recognition has been a very active area of research. The main object of hand-printed recognition is interpretation of data which describes hand-printed objects [1-3]. On-line hand-printed recognition deals with a data stream coming from a transducer while user is writing [4-5]. Off-line hand-printed recognition deals with data set which has been obtained from a scanned hand-printed document [6-7]. Useful reviews are found in [8-10]. Most researchers have adopted the pattern recognition approach in which image pre-processing is followed by feature extraction and classification. There have been many subtopics but these may be roughly divided into the following categories: feature extraction and selection, classification method, and combination of classifier.

The wavelet transform is a new tool that has been applied in many disciplines, including image processing [11-12]. Due to the multiresolution property, it decomposes the signal at different frequency scales. For a given image the wavelet transform produces a low frequency subband image reflecting its basic shape and three subband images that contains the high frequency components of the image at horizontal, vertical and diagonal directions. These components can be used to construct the feature vector.

This paper proposes a system for the recognition of hand-printed Latin characters using a multilayer network [13] trained with the back propagation [14] algorithm. This pilot study will be performed to analyse the usefulness of the wavelet feature selection technique combined with neural network classifier. The system was tested on a sample of 6760 hand-printed characters provided by 30 writers of different ages and with many different sizes and writing styles. The result obtained using 5200 characters for training the network and 1560 characters for testing, yielded a recognition rate of 89.2%.

The structure of the paper is as follows: section 2 introduce the image representation, section 3 describes the wavelet systems, section 4 represents the feature extraction, section 5 outlines the structure and training of neural network, section 6 presents of experiments performed

and finally, some concluding remarks are made in section 7.

2- Image Representation

A character image is scanned and digitized. A 300 dpi scanner is used to digitize the image. The result file (TIFF format) ready to preprocess. The image extracted is a 100x100 pixels, however, it has unpredictable size. Therefore, it is necessary to normalize (or scale) the image make sure the recognition is size invariant. The pattern is also centered because scaling keeps initial aspect ratio (this condition voids a deformation of the pattern). In this approach of image representation, the image size is fixed at 32x32 [15].

3- Wavelet Systems

Biorthogonal wavelet systems can be obtained using analysis filters for decomposition and synthesis filters for reconstruction [11]. Associated with the analysis filter (h for low-pass, g for high-pass) and the synthesis filter (h_0 for low-pass, g_0 for high-pass) are the scaling function $\Phi(x)$ and the dual $\bar{\Phi}(x)$, defined respectively as:

$$\Phi(x) = \sum_n h(n) \sqrt{2} \Phi(2x - n) \quad (1)$$

$$\bar{\Phi}(x) = \sum_n h_0(n) \sqrt{2} \Phi(2x - n) \quad (2)$$

For Φ and $\bar{\Phi}$ to exist the following criterion must hold

$$\sum_n h(n) = \sum_n h_0(n) = \sqrt{2} \quad (3)$$

The wavelet $\psi(x)$ and dual $\bar{\psi}(x)$ are defined as :

$$\psi(x) = \sum_n g(n) \sqrt{2} \Phi(2x - n) \quad (4)$$

$$\bar{\psi}(x) = \sum_n g_0(n) \sqrt{2} \bar{\Phi}(2x - n) \quad (5)$$

The wavelet system is said to be biorthogonal if :

$$\sum_n h(n) h_0(n+2k) = \delta(k) \quad (6)$$

$$\sum_n h(n) g_0(n+2k) = 0$$

$$\sum_n h_0(n) g(n+2k) = 0$$

$$\sum_n g(n) g_0(n+2k) = \delta(k)$$

The analysis and synthesis filters are related by

$$g(n) = (-1)^n h_0(1-n), g_0(n) = (-1)^n h(1-n) \quad (7)$$

In two dimensional wavelet decomposition the analysis scaling function can be written as :

$$\Phi(x, y) = \Phi(x) \Phi(y) \quad (8)$$

where $\Phi(x)$ is one dimensional scaling function

Let $\psi(x)$ the one dimensional wavelet associated with the scaling function $\Phi(x)$. The three two dimensional analysis wavelets (i.e. horizontal, vertical, diagonal direction respectively) are defined as:

$$\begin{aligned} \psi_{LH}(x, y) &= \Phi(x) \psi(y) \\ \psi_{HL}(x, y) &= \psi(x) \Phi(y) \\ \psi_{HH}(x, y) &= \psi(x) \psi(y) \end{aligned} \quad (9)$$

The image is split into an approximation and details images. The approximation image is then split itself into a second level of approximation and details. For an n-level decomposition, the signal is decomposed in the following way :

$$\begin{aligned} A_n &= [h_x * [h_y * A_{n-1}] \downarrow 2, I] \downarrow I, 2 \\ D_{n1} &= [h_x * [g_y * A_{n-1}] \downarrow 2, I] \downarrow I, 2 \\ D_{n2} &= [g_x * [h_y * A_{n-1}] \downarrow 2, I] \downarrow I, 2 \\ D_{n3} &= [g_x * [g_y * A_{n-1}] \downarrow 2, I] \downarrow I, 2 \end{aligned} \quad (10)$$

Where * denotes the convolution operator, $\downarrow 2, I$ sub-sampling along the rows (columns) and $A_0 = I(x, y)$ is the original image, A_n is obtained by low pass filtering and is the approximation image at scale n. the detailed images D_{ni} are obtained by band-pass filtering in a specific direction ($i=1, 2, 3$ for vertical, horizontal and diagonal direction respectively) and thus contain directional detailed information at scale n. The original

image I is thus represented by a set of subimages at several scales $\{A_n, D_{ni}\}$.

4- Feature Extraction

Feature extraction is an important step in achieving good performances for character recognizer. Extracted features must be invariant to the distortions and variations that can be expected in a specific application. The size of the feature set is also important in order to avoid phenomenon called the dimensionality problem. In The sequency, the wavelet decomposition is applied at one level of resolution. The normalized characters 32x32 pixels, yielding four subband images ($A_1, D_{11}, D_{12}, D_{13}$) which form the feature vector as shown in Figure (1) where

A_1 the subband image corresponds to the low frequencies (global characteristics),

D_{11} gives the vertical high frequencies (horizontal- details), D_{12} the horizontal high frequencies (vertical details) and D_{13} the high frequencies in both diagonal directions (diagonal details). Each subband containing 16x16 pixels.

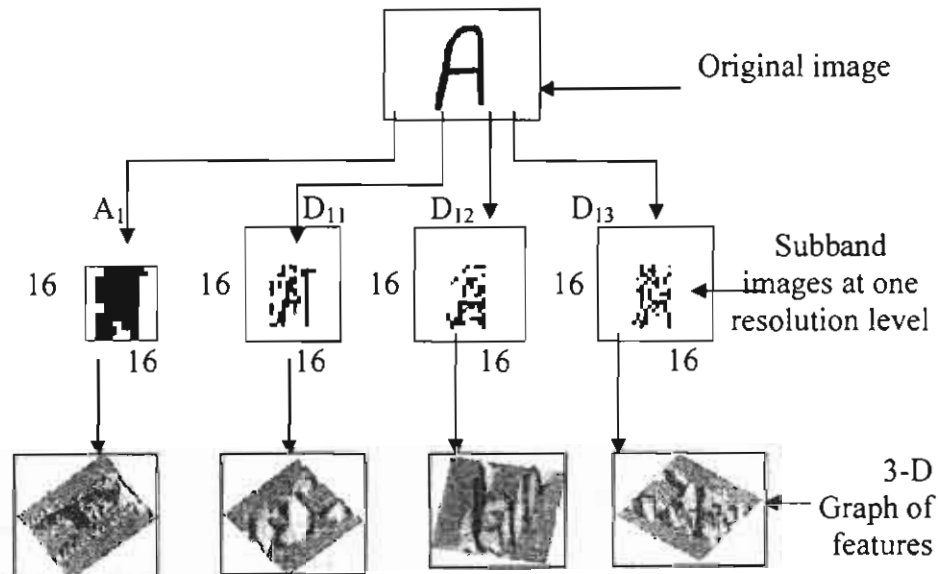


Figure (1): Feature extraction at one level 2-D wavelet decomposition using a bank of one-dimensional low-pass and high-pass analysis filter of character "A".

5- Neural Network Recognizer

Artificial neural network (ANN) are highly parallel information processing systems resembling that of human brain configured in regular architectures. The collective behavior demonstrates the ability to learn, recall and generalize from training patterns or data. A neural network learns about its environment through an iterative process of adjustments applied to its synaptic weights and thresholds. Ideally, the network becomes more knowledgeable about its environment after each iteration of the learning process. Learning can be defined as a process by which the free parameters of a neural network are adapted through a continuing process of simulation by the environment in which the network is embedded. Because of their learning and memorizing capability, a neural network recognizer is used for hand- printed character recognition.

5-1 Multi-layer Neural Net

The multilayer architecture consists of an input layer of information processing nodes, hidden layers with hidden nodes and finally an output layer, which consists of output nodes that usually equal the number of data classes [12],[14]. The nodes are also referred to as neurons. The input data is presented to the neural network through the input layer and this information is propagated through various network nodes under the control of the training algorithms. The main aim of a training algorithm is to modify the weights on individual network connections so that learning can take place. The training algorithm modifies network weights depending on training error which is defined as the difference between the actual and desired network output. At the end of the training process, which represents achieving a tolerable error, the network stores its knowledge as the final set of weights. The network structure used is shown in Figure (2). The input layer consists of four maps of 16x16 unit each. Each map represent the feature vector formed by the subband image. Each of these four input maps make up fully connected groups with their corresponding maps in the hidden layer. The output layer have 52 units, fully connected with all units in the hidden layer. So the architecture of the network yielding, the four independent subnets. The output represents the total number of classes.

The back propagation with momentum training scheme is followed. This method was selected because of its simplicity and because it has been previously used on a number of pattern recognition problems. The method works on the principle of gradient descent and has been described in the basic form in detail by Rumelhart et al. [14].

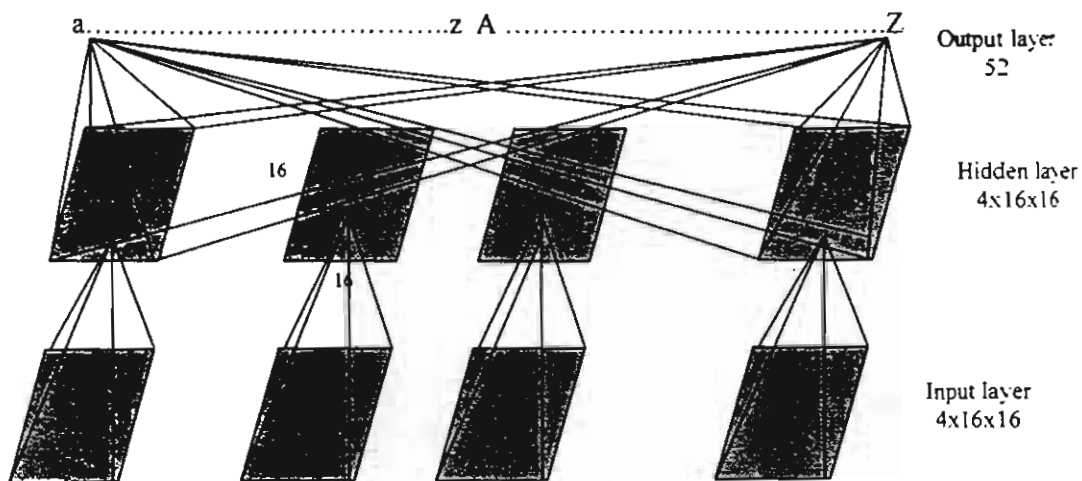


Figure (2): Network structure.

6- Experimental Results

Experiments have been performed using isolated character images. As preprocessing, the image considered only size normalization (32x32). No further preprocessing like tilt correction, smoothing etc. are considered. The wavelet decomposition is applied at one level of resolution resulting four 16x16 subband images. To improve the recognition by the neural network the values of the wavelet coefficients, are divided by their maximum value to be normalized to range [0 - 1].

The back propagation with momentum training scheme is followed. The method works in the principle of gradient descent and has been described in its basic form in detail by [14]. The algorithm uses two parameters, the learning rate and the momentum, these parameters allow the algorithm to converge more easily if they are properly set by experimenter[16]. In this paper the learning rate set as 0.6 and the momentum term as 0.9. The training error falls gracefully to an acceptable value ($e < 0.001$) at the end of the training over 1000 epochs, as shown in Figure (3).

Each output unit represents one class. Each class for character recognition represents a particular alphabets with the first 26 classes representing lower case letters (a-z) and the next 26 classes the alphabets in upper case (A-Z). when introducing a pattern that belong to class 'c' the trained output will be '1' for the c^{th} output unit and '0' for the others.

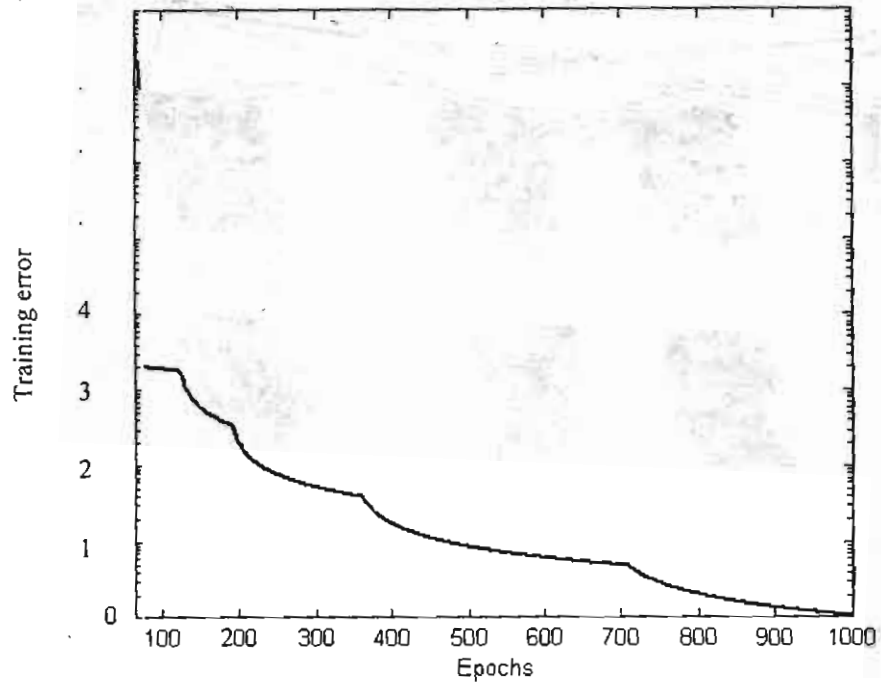


Figure (3): Change in the sum of squared error over 1000 epochs.

The performance rate of the recognition estimated using 'k.fold cross validation' by Fu [17]. The average recognition rate of 10-fold is 89.2% the recognition on the training data is extremely high, 99.77% which represents high quality training as shown in table (1).

Fold	Recognition rate (%)	
	Training	Testing
1	99.7	89.0
2	99.7	88.0
3	99.9	90.0
4	99.8	86.0
5	99.8	89.0
6	99.8	90.0
7	99.7	88.0
8	99.8	93.0
9	99.9	91.0
10	99.8	88.0
Average	99.77	89.2

Table (1): Neural network. Recognition rate performance using ten fold- cross validation on the training and test sets and the sum of the squared error at the end of training.

Conclusion

The system adopted in this paper for recognizing hand-printed Latin characters uses the neural network combined with bi-dimensional wavelet. The potential of the wavelet extracted features and the contribution of each subband image were investigated. Recognition rate of 89.2% is obtained using ten fold cross-validation. Thus, the wavelet proved to provide good feature extraction of hand-printed characters. Finally, the system study demonstrates that the bi-dimensional wavelet combined with neural network are well suited for the off-line recognition of hand-printed Latin characters. In order to improve system performance, other types of features may be considered in future study.

References

- [1] S.Mori, S.y.Suen and K.Yamamoto, "Historical review of OCR Research and development". Proceeding of The IEEE,80, pp.1029-1058,1992.
- [2] A. Wang, K. Fan, " Optimal recognition of handwritten character by hierarchical radical matching method". Pattern Recognition,34, pp. 15-35, 2001.
- [3] T. Wakahara, and Odaka," Adaptive Normalization of Handwritten Characters using Global/Local Affine Transformation". ICDAR, pp. 28-33, 1997.
- [4] D.J. Burr, " Designing a hand written reader". 5th Int. Conf. Pattern Recognition, PP 715-722, 1980.
- [5] J. Kim, and C.C. Tappert, "Hand written recognition accuracy versus tablet resolution and sampling rate". 7th Int. Conf. Pattern Recognition, PP 917-918, 1984.
- [6] M. Bosker, "Omnidocument Technologies". Proc. IEEE, 80(7): 1066- 1078, 1992.
- [7] S. Srihari, "From pixel to paragraphs the use of models Int. text recognition". Second Annual Symposium on Document Analysis and Information Retrieval, PP 47-64,1993.
- [8] R.O. Duda and P.E. Hart, "Pattern classification and scene analysis". John Wiley & Sons, Inc., New York, 1979.
- [9] R. Schalkoff, "Pattern recognition: Statistical, structural and neural approaches". John Wiley and Sons, Inc., New York, 1992.
- [10]J. Mantas. "An overview of character recognition methodologies". Pattern Recognition, 19(6): 425- 430, 1986.
- [11]I. Daubechies. "Ten lectures on wavelets". Prentice Hall, 1998.

- [12] I.S Oh and C.Y. Sven, " A class modular feed forward neural network for hand writing recognition". *Pattern Recognition*, (35):, pp. 229-244,2002.
- [13] S. When Lee, "Off-line recognition of totally unconstrained hand writing numerals using multilayer cluster neural network". *IEEE transactions on P.A.M.I.*, Vol. 18, Num. 6, PP. 648-652,1996.
- [14] D.E. Rumelhart ,R.Durbin , Golden , Chauvin, and DE Rumelhart (eds), "Back propagation: Theory, architectures and applications", Lawrence Erlbaum Hillsdale, NJ, PP 1-34,1995.
- [15] Wakahara, T. Odaka, "Adaptative normalization of hand written characters using global/ local Affine Transformation", *ICDAR-IEEE*, PP. 28-33,1997.
- [16] C.M. Bishop, "Neural Networks for Pattern Recognition", Clarendon Press, Oxford, 1995.
- [17] L. FU, "Neural Networks in Computer Intelligence". McGraw-Hill, Singapore, PP 331-348,1994.
- [18] Z. James, and Wang, "Unsupervised Multiresolution Segmentation for images with low depth of field", *IEEE Transaction of Pattern Analysis and Machine Intelligence*, Vol. 23, No.1,2001.