HANDWRITTEN NUMERICAL CHARACTERS RECOGNITION USING BIORTHOGONAL SPLINE WAVELETS

SUZETE E. N. CORREIA   JOÃO M. DE CARVALHO

Laboratório de Automação e Processamento de Sinais
Departamento de Engenharia Elétrica – Universidade Federal da Paraíba
Caixa Postal 10105, CEP 58.109-970 - Campina Grande, PB, Brasil
{suzete,carvalho}@dee.ufpb.br

ABSTRACT

In this paper, a novel approach for recognition of handwritten numerical characters using the biorthogonal spline wavelets Cohen-Daubechies-Feauveau (CDF) 3/7 is proposed. The main purpose is to optmize character recognition methods and develop a system which is able to efficiently absorb the variations presented in this type of data. The proposed scheme consists of three stages: scale normalization, feature extraction and classification. The classifier used is a multilayer cluster neural network. In order to verify the performance of the method, experiments were realized with the numerical database of Concordia University of Canada, obtaining an recognition rate of 94.7 %.

1. INTRODUCTION

The automatic recognition of handwritten numerical characters has been an interesting and useful research topic during the last two decades because both of its theoretical value in pattern recognition and its numerous possible applications, such as automatically processing postal ZIP codes from mail pieces and money amount in bankchecks [1]. However, it is regaderd as a difficult problem due to the large degree of variability the data may exhibit. There are changes and distortions not only from one writer to another, but even for the same writer [2]. In order to attain high recognition and reliability rates it is desired to develop a system which is able to efficiently extract features from handwritten numerals, despite all variations present. To achieve that, many methods have been proposed over the years [3], [4].

The wavelet transform is a new theory developed by scientists and mathematicians in the last few years and applied in many disciplines including signal processing and image processing [5], [6], [7]. Wavelet decomposition of a signal gives more information than other wave decompositions such as the classical Fourier transform [5]. Especially, due to the multiresolution property, it provides local information in both time and frequency domains. Therefore, this tool is suitable for representing the handwritten numerals in a recognition system, since the wavelet transform of a image reflects the basic shape of the numeral and its details along horizontal, vertical and diagonal directions [6]. Since wavelet studies are recent, only a few papers have been published on handwritten numerical character recognition [8], [9], [10]. In this paper, a novel apprach is proposed using the wavelet coefficients at one resolution level to construct the feature vector. The biorthogonal spline wavelets Cohen-Daubechies-Feauveau (CDF) 3/7 were choosen, based on experiments performed [10].

The authors would like to express their thanks to CNPq for the financial support to the research.

The proposed scheme consists of three stages: preprocessing, feature extraction and classification. Preprocessing deals with image normalization in order to reduce the class intravariability. Feature extraction represents the normalized images by wavelet coefficients obtained with the CDF 3/7. Classification performs the final decision according to extracted features and acquired knowledge. The classifier used is a mulitlayer cluster neural network trained with the backpropagation momentum algorithm [10].

The rest of this paper is organized as follows: Section 2 describes the preprocessing method employed. The design of a feature extractor is sketched in Section 3 and a multilayer cluster neural network is presented in Section 4. Experimental results are analysed and concluding remarks are given in Section 5 and 6, respectively.

2. PREPROCESSING

The handwritten numeral preprocessing considered in this work uses scale normalization to reduce the variations in size and produce characters images with the same dimension. The scale normalization algorithm selected is based on the method proposed by Veloso and Carvalho [11] which performs scaling with distinct factors along different axes. The maping function for the scaling invariant image is represented by:

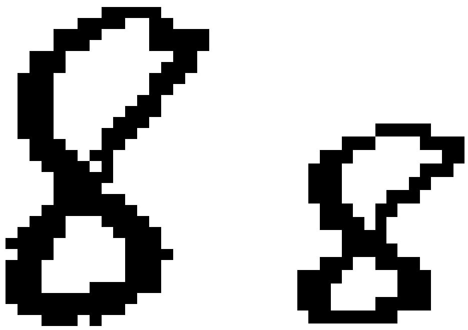
$$f(x_i, y_j) = f(S_x x_i, S_y y_j) \tag{1}$$

where $f(x,y)$ gives the pixel values of the original image at coordinates (x,y) and $f_s(x,y)$ gives the pixel values of the new image at coordinates (x,y). S_x and S_y are the scaling factors at the axes x and y, respectively, calculated by the equations:

$$S_x = \frac{N}{R_x}, \tag{2}$$

$$S_y = \frac{M}{R_y}. \tag{3}$$

The variables R_x and R_y are the desired dimensions of the normalized image and the variables N and M the dimensions of the original image. In equation 1, $S_x x_i$ and $S_y y_j$ may be fractional values. In this case, they can be converted to integer assuming that if the fractional value is > 0.5 rounding is done. Figure 1 gives an example of the preprocessing results. Here, the numerals are size normalized to 16×16 pixels.



(a) (b)

Fig. 1 - Normalized image of the numeral 8: (a) Original character with 29×18 pixels, (b) Nomalized character with 16×16 pixels.

3. FEATURE EXTRACTION

This section introduces the notation of a biorthogonal wavelet system, which is particularly useful for the application proposed. It is assumed that the reader is familiar with the principles of the wavelet transform. More general description can be found in the work by Burrus et al.[5], Mallat [6] and Daubechies [7].

3.1 Biorthogonal Wavelet System

Biorthogonal wavelet systems can be obtained using the analysis filters for decomposition and the synthesis filters for reconstruction. Associated with the analysis filters (h and g) and the synthesis filters (h_0 and g_0) are the scaling function $\phi(x)$ and the dual $\tilde{\phi}(x)$, respectively defined by

$$\phi(x) = \sum_n h(n)\sqrt{2}\phi(2x-n) \tag{4}$$

$$\tilde{\phi}(x) = \sum_n h_0(n)\sqrt{2}\tilde{\phi}(2x-n) \tag{5}$$

For ϕ and $\tilde{\phi}$ to exist, equation 6 must be satisfied

$$\sum_n h(n) = \sum_n h_0(n) = \sqrt{2} \tag{6}$$

The wavelet $\psi(x)$ and its dual $\tilde{\psi}(x)$ are defined as

$$\psi(x) = \sum_n g(n)\sqrt{2}\phi(2x-n) \tag{7}$$

$$\tilde{\psi}(x) = \sum_n g_0(n)\sqrt{2}\tilde{\phi}(2x-n) \tag{8}$$

The system is said to be biorthogonal if the following four conditions hold:

$$\begin{aligned} \sum_n h(n)h_0(n+2k) &= \delta(k) \\ \sum_n h(n)g_0(n+2k) &= 0 \\ \sum_n h_0(n)g(n+2k) &= 0 \\ \sum_n g(n)g_0(n+2k) &= \delta(k) \end{aligned} \tag{9}$$

The analysis and synthesis filters are related as

$$g(n) = (-1)^n h_0(1-n), \quad g_0(n) = (-1)^n h(1-n) \tag{10}$$

These filters not necessarally are identical, as they may be of unequal length. Table I shows the filters coefficients $h(n)$ and $h_0(n)$ for the bases CDF 3/7 [7].

TABLE I

FILTER COEFFICIENTS FOR THE ANALYSIS $h(n)$ AND SYNTHESIS $h_0(n)$ CDF 3/7 WAVELET BASES

$h(n)$	$h_0(n)$
0.1767766	-0.003021
0.530330	-0.009063
0.530330	-0.016831
0.1767766	0.074663
	0.031332
	-0.301159
	-0.026499
	0.951642
	0.951642
	-0.026499
	-0.301159
	0.031332
	0.074663
	-0.016831
	-0.009063
	-0.003021

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

$$\phi(x,y) = \phi(x)\phi(y) \tag{11}$$

where $\phi(x)$ is a one-dimensional scaling function.

Let $\psi(x)$ be the one-dimensional wavelet associated with the scaling function $\phi(x)$. Then, the three two-dimensional analysis wavelets are defined as

$$
\begin{aligned}
\psi_{LH}(x,y) &= \phi(x)\psi(y) \\
\psi_{HL}(x,y) &= \psi(x)\phi(y) \\
\psi_{HH}(x,y) &= \psi(x)\psi(y)
\end{aligned} \tag{12}
$$

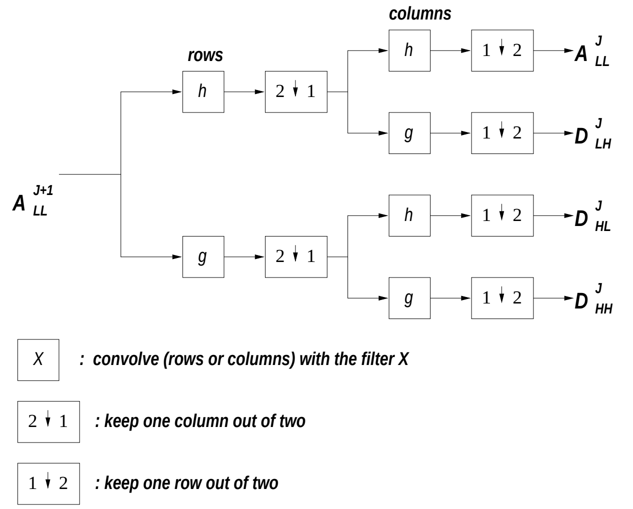


Fig. 2 - One level 2D wavelet decomposition using a bank of one-dimensional low- pass and high-pass analysis filters.

Figure 2 represents one level 2D wavelet decomposition. The rows and columns of the image are convolved with the one-dimensional analysis filters, where $h(n)$ and $g(n)$ act as low-pass and high-pass filters, respectively. This decomposition provides subband images corresponding to different resolution levels and orientations. The subband image A^{j}_{LL} corresponds to the lowest frequencies (global information), D^{j}_{LH} gives the vertical high frequencies (horizontal details), D^{j}_{HL} the horizontal high frequencies (vertical details) and D^{j}_{HH} the high frequencies in both diagonal directions (diagonal details).

3.2 Feature Vector

After the preprocessing stage, wavelet decomposition is applied at one level resolution, yielding four subband images $\{A^{-1}_{LL}, D^{-1}_{LH}, D^{-1}_{HL}, D^{-1}_{HH}\}$, each containing 8×8 pixels. Therefore, the feature vector consists of $4 \times 8 \times 8$ pixels from the subband images. For each subband image the values of the wavelet coeficients were normalized to the range [0,1]. This procedure improves recognition by the neural network. Figure 3 ilustrates the feature extraction steps.

4. CLASSIFICATION

Classification is performed using a three layer cluster neural network [9]. The input layer consists of four 8×8 clusters because the network inputs are the feature vector composed

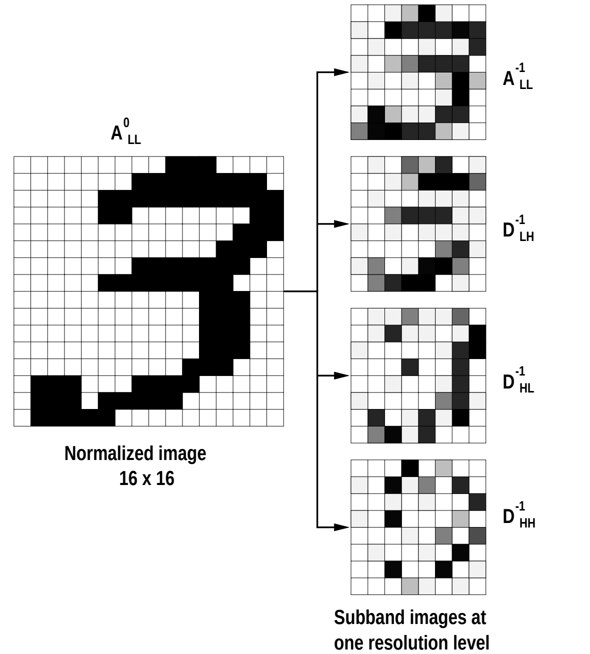


Fig. 3 - Feature Extraction at one resolution level with CDF 3/7.

by the four subband images at one resolution level. Each cluster in the input layer is fully indenpendently connected to a corresponding cluster in the hidden layer. However, the output layer is fully connected to all hidden units. The output consists of 10 units, one per class of numeral. All the weights are inicialized with random values between -1 and 1, and are up-dated using the backpropagation momentum algorithm [12].The network architecture is presented in Figure 4.

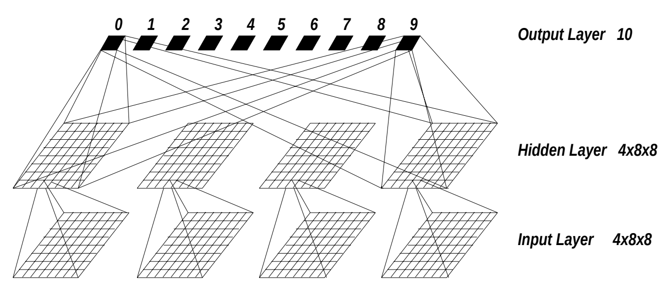


Fig. 4 - Architecture of the multilayer cluster neural network.

One intuitive advantage of this multilayer cluster neural network is that while one subnetwork may be confused by a given input, others may not because each subnetwork starts from different initial state and learn with different features [9].

TABLE II
Confusion Matrix with the CDF 3/7

Class	0	1	2	3	4	5	6	7	8	9	Rejected	Substituted	Corrected
0	200										0	0	200
1		194									6	0	194
2			195								5	0	195
3			4	190					2		4	6	190
4			1		186					3	10	4	186
5			2	2		176	2			1	17	7	176
6			2			1	190				7	3	190
7					2			189		5	4	7	189
8		1				1			184	2	12	4	184
9					2	1		2		190	5	5	190

5. EXPERIMENTAL RESULTS AND ANALYSIS

5.1 Database Used

The characters used in the experiments are from the database of Concordia University of Montreal, Canada, which consists of 17000 unconstrained handwritten numerals originally collected from dead letter envelopes of the U.S. Postal Service at different locations in the U.S. The numerals of this database were digitilized in bilivel on 64×224 grid of 0.153 mm square elements, giving a resolution of approximatelly 166 PPI [2]. Among the data, 4000 numerals were used for training the neural network and 2000 numerals for testing. Figures 5 and 6 shows some representative and confuse samples taken from the database, respectively. Many different writing styles could be observed, as well as numerals of different sizes and stroke widths.

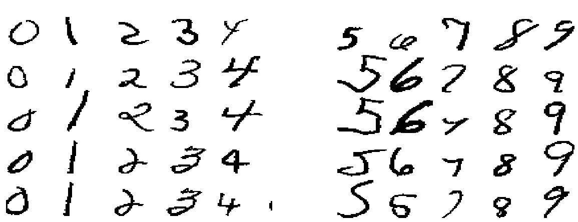


Fig. 5 - Representative samples of the database.

5.2 Experimental Results

To evaluate the performance of the presented method the parameters in the multilayer cluster neural network were set as 0.01 for the learning rate and 0.9 for the momentum term. Table II shows the confusion matrix with the results obtained and Table III the average rates.

A numeral is classified as belonging to class i, if the i-th output of the neural network produces the largest value and the difference between the two largest output values is at least 0.2. Otherwise the numeral is rejected. The thereshold

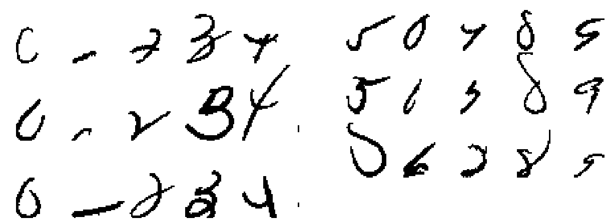


Fig. 6 - Confuse samples of the database.

TABLE III
Average Rates

CDF	Recognized	Substituted	Rejected	Reliability
3/7	94.7 %	1.8 %	3.5 %	98.13 %

value of 0.2 has been empirically determined through the experiments.

6. CONCLUSION REMARKS

In this paper, a novel approach for recognition of handwritten numerical characters using biorthogonal spline wavelets to extract features was presented. The wavelet decomposition enables an efficient representation of the characters, since global and directional features are provide. Results show that the recognition system obtained was very robust to variations of writing style and shape of the handwritten numerals.

7. REFERENCES

[1] S. Srihari High-Performance Reading Machines. *Proceedings of the IEEE, Vol. 80, No. 7*, pages 1120–1132, 1992.

[2] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai and L. Lam Computer Recognition of Unconstrained Handwritten Numerals. *Proceedings of the IEEE, Vol. 80, No. 7*, pages 1162–1180, 1992.

[3] C. Y. Suen, M. Berthod and S. Mori Automatic Recognition of Handprinted Characters - The State of the Art. *Proceedings of the IEEE, Vol. 68, No. 4*, pages 469–487, 1980.

[4] O. Trier, A. Jains and T. Taxt Feature Extraction Methods for Character Recognition - A Survey. *Pattern Recognition, Vol. 29, No. 4*, pages 641–662, 1986.

[5] C. S. Burrus, R. A. Gopinath and H. Guo Intoduction to Wavelets and Wavelet Transforms. *Prentice Hall*, 1998.

[6] S. G. Mallat A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7*, pages 674–6933, 1989.

[7] I. Daubechies Ten Lectures on Wavelets. *Prentice Hall*, 1998.

[8] G. S. Kapogiannopoulos and M. Papadakis Character Recognition using a Biorthogonal Discrete Wavelet Transform. *Wavelet Applications in Signal and Image Processing IV, SPIE, Vol. 2825*, pages 384–393, 1996.

[9] S.-W. Lee, C.-H. Kim and Y. Y. Tang Multiresolution Recognition of Unconstrained Handwritten Numerals with Wavelet Transform and Multilayer Cluster Neural Network. *Pattern Recognition, Vol. 29, No. 12*, pages 1953–1961, 1996.

[10] S. Correia and J. M. Carvalho Optimizing the Recognition Rates of Unconstrained Handwritten Numerals Using Biorthogonal Spline Wavelets. *ICPR 2000*, Barcelona, Espanha, September 2000.

[11] L. R. Veloso and J. M. Carvalho Neural versus Syntactic Recognition of Handwritten Numerals. *ICDAR'99*, Bangalore, India, 1999.

[12] Simon Haykin Neural Networks -A Comprehensive Foundation. *Prentice Hall*, 1996.