

*Keynote Address at 6<sup>th</sup> International Program on Wavelet Analysis and Active  
Media Technology*

## **Wavelet Feature Extraction for the Recognition and Verification of Handwritten Numerals**

P. ZHANG, T. D. BUI, C. Y. SUEN

*Centre for Pattern Recognition and Machine Intelligence  
Department of Computer Science and Software Engineering  
Concordia University*

*1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8 Canada*

Two kinds of wavelet features are proposed: (a) Kirsch edge enhancement based 2D wavelets and (b) 2D complex wavelets. The two sets of hybrid features are congregated by combining them with the geometrical features for the recognition of handwritten numerals. Experiments conducted on handwritten numeral recognition and verification show that the two hybrid feature sets can achieve high recognition and verification performance. In addition, the merits of the proposed wavelet feature extraction methods are discussed.

Key Words: Hybrid Feature Extraction, Wavelet Transform, Complex Wavelet Transform, OCR, Artificial Neural Networks.

### **1. Introduction**

Feature extraction is a vital step for pattern recognition, especially for Handwritten Optical Character Recognition (HOOCR) as there are varieties of handwritten styles depending on the subject's age, gender, educational and ethnical background, as well as his/her mood while writing. Although OCR is one of most successful applications in pattern recognition and machine intelligence, and great efforts have been made in terms of theoretical breakthrough and implementation [1, 2], there is still room for researchers to develop new theory and algorithm to seek higher recognition rate and reliability.

Many feature extraction methods have been reported such as various moment features, gradient and distance-based features, geometric features, transform domain features, etc. Wavelet transform has been widely used in image processing and signal processing for image/signal enhancement, denoising, texture segmentation [3] based on its properties of short support, orthogonality, symmetry, higher order of vanishing moments, and more importantly, its multiresolution decomposition analysis. Its application to pattern recognition, especially to OCR, is a relative new research field.

In this paper, we propose two kinds of discrete wavelets and successfully apply them for the recognition and verification of handwritten numerals. In Section II, a 2D discrete wavelet transform is applied to the Kirsch edge enhanced images to extract 2D wavelet features. The structure of 2D Complex

Wavelet Transform (CWT) feature extraction is described in section III, which has the following merits: directional selectivity and insensitivity to the translational shifts of the input images. Geometrical feature extraction is briefly introduced in Section IV. In Section V, the two sets of hybrid features are congregated by combining them with the geometrical features. The experimental results conducted on the recognition and verification of MNIST handwritten numerals are also given in this section. The merits of the two sets of wavelet transforms used for OCR feature extraction are discussed in Section VI and conclusions end our paper.

## 2. Wavelet Feature Extraction based on Kirsch Edge Enhancement

According to wavelet theory, a conventional two dimensional wavelet discrete transform (2D-DWT) can be regarded as equivalent to filtering the input image with a bank of filters, whose impulse responses are all approximately given by scaled versions of a mother wavelet. The output of each level consists of four sub-images: LL, LH, HL, and HH with 2:1 down-sampling. Mathematically, we can express this recursive algorithm in the following equation:

$$\begin{aligned}
 x_{LL, k_1, k_2}^{(n-1)} &= \sum_{l_1, l_2} h_{l_1-2 k_1} h_{l_2-2 k_2} x_{LL, l_1, l_2}^{(n)}, \\
 y_{LH, k_1, k_2}^{(n-1)} &= \sum_{l_1, l_2} h_{l_1-2 k_1} g_{l_2-2 k_2} x_{LL, l_1, l_2}^{(n)}, \\
 y_{HL, k_1, k_2}^{(n-1)} &= \sum_{l_1, l_2} g_{l_1-2 k_1} h_{l_2-2 k_2} x_{LL, l_1, l_2}^{(n)}, \\
 y_{HH, k_1, k_2}^{(n-1)} &= \sum_{l_1, l_2} g_{l_1-2 k_1} g_{l_2-2 k_2} x_{LL, l_1, l_2}^{(n)}
 \end{aligned} \quad \dots (1)$$

The gradient based feature extraction is implemented as follows: first, the Kirsch nonlinear edge enhancement algorithm is applied to an NxN character image to extract horizontal, vertical, right-diagonal and left-diagonal directional features and global features; then a 2-D wavelet transform is used to filter out high frequency components of each directional feature image and character image, respectively, and to convert the feature matrix into a 4x4 matrix. Suppose that we define the eight neighbors of pixel (i,j) as in Fig. 1.

$$\begin{array}{ccccc}
 A0 & A1 & A2 & & \\
 A7 & (i,j) & A3 & & \\
 A6 & A5 & A4 & & 
 \end{array}$$

Fig. 1 Definition of eight neighbors of pixel (i,j)

Kirsch defined a nonlinear edge enhancement algorithm as follows:

$$G(i, j) = \max \{1, \max_{k=0}^7 [|5S_k - 3T_k|]\} \quad \dots (2)$$

where

$$\begin{aligned} S_k &= A_k + A_{k+1} + A_{k+2} \\ \text{and} \\ T_k &= A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7} \end{aligned} \quad \dots (3)$$

For the detailed algorithm implementation, interested readers can refer to [4]. In addition, we develop four new templates to detect horizontal endpoints and vertical endpoints. The templates are shown below:

$$\begin{bmatrix} -5 & -1 & -1 \\ -5 & 5 & 5 \\ -5 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -5 \\ 5 & 5 & -5 \\ -1 & -1 & -5 \end{bmatrix} \quad \begin{bmatrix} -5 & -5 & -5 \\ -1 & 5 & -1 \\ -1 & 5 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 5 & -1 \\ -1 & 5 & -1 \\ -5 & -5 & -5 \end{bmatrix}$$

We apply Daubechies-4 wavelets to four directional feature matrices, two endpoint matrices and a character image, and only kept 4x4 low frequency components of each as features. Totally there are 4x4x7=112 Kirsch based Wavelet Features (KWF) for each character.

### 3. Complex Wavelet Transform and Feature Extraction

Complex Wavelet Transform adds some new merits such as approximate shift invariance, good directional selectivity for 2-D image. 2D-CWT functions have following form:

$$h(x, y) = a(x, y)e^{j(w_x x + w_y y)} \quad \dots (4)$$

where  $a(x,y)$  is a slowly varying Gaussian-like real window function centered at (0,0), and  $(w_x, w_y)$  is the center frequency of the corresponding subband. So the complex coefficients of the  $i$ th subband of the  $l$ th level can be written as:

$$c_i^l = u_i^l + jv_i^l \quad \dots (5)$$

The magnitude of each component of each subband is calculated as:

$$C_i^l = \sqrt{(u_i^l)^2 + (v_i^l)^2} \quad \dots (6)$$

Since  $a(x,y)$  is slowly varying, the magnitude is insensitive to the small image shift. Complex filters in two dimensions provide true directional selectivity. There are six subband images of complex coefficients at each level, which are strongly oriented at angles of  $\pm 15^\circ$ ,  $\pm 45^\circ$ ,  $\pm 75^\circ$ . These two properties are useful for pattern recognition.

2D-CWT can be implemented using a dual-tree structure. Fig. 2 shows a 2D-CWT feature extraction scheme for the recognition and verification of handwritten numerals. In addition, the up-sampling before filtering and the down-sampling after filtering are needed in order to maintain the approximate shift invariance. The outputs of the subband images are congregated into complex wavelet coefficients. Interested readers can refer to reference [5] for further details.

For example, if a  $32 \times 32$  character is decomposed into the third level, the final size of each subband image is  $4 \times 4$ . We keep only the amplitude coefficients for three high frequency components and both amplitude and phase information for the low frequency components. The number of Complex Wavelet Features (CWF) is 160.

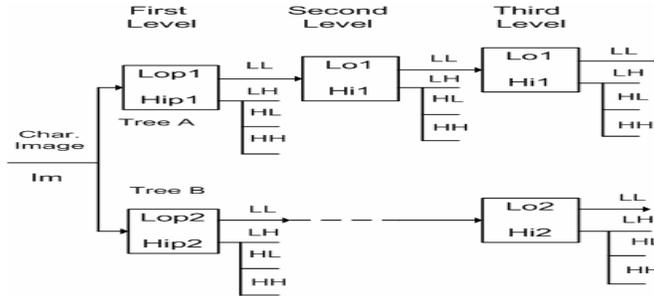


Fig. 2 The schematic diagram of 2D-CWT for character feature extraction

#### 4. Geometrical Feature Extraction

Character geometrical features such as No. of loops, No. of T-joints, No. of X-joints, No. of end points, concavity/convexity, middle line feature, and local segment features will be used and encoded as 20 geometrical features (GF).

## 5. Recognition and Verification Scheme for Handwritten Numerals

We propose an OCR recognition and verification system shown in Fig. 3. MNIST handwritten database is used in the recognition and verification experiments. The database includes 60,000 training samples and 10,000 testing samples.

An Artificial Neural Network (ANN) of three layers with Back Propagation (BP) algorithm is used as a classifier. In order to complementarily congregate two sets of statistical wavelet features with one set of structural geometrical features, two sets of new hybrid features are constructed. We conducted a series of recognition experiments based on the hybrid features. The recognition results are listed in Table I. Table II lists the verification rates conducted on some character pairs. The overall recognition rate is over 99%.

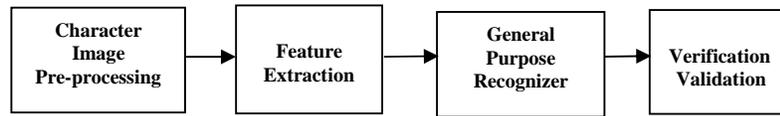


Fig. 3 A system architecture for OCR recognition and verification

Table I List of hybrid feature sets and recognition accuracies

Name of Feature Set	No. of Features	Recognition Accuracy (%)
KWF + GF	132	98.55
CWF + GF	180	98.30

Table II Verification rates (%) conducted on two wavelet feature sets

Char pairs	1-7	3-5	4-9	0-6	4-6	9-8	5-8	2-8	6-5
CWF	99.40	99.10	99.15	99.30	99.15	99.15	99.25	99.30	99.15
KWF	99.40	99.15	99.15	99.35	99.15	99.20	99.30	99.35	99.15

## 6. Discussion and conclusions

In this paper, we propose two kinds of wavelet feature extraction methods. Furthermore, the two sets of hybrid features are congregated by combining the respective statistical wavelet features and structural geometrical features for the recognition and verification of handwritten numerals. Experiments demonstrated that the proposed two sets of hybrid features can achieve high recognition performance. The overall recognition rate by combining the general purpose recognizer and the verifier can reach over 99%.

We use Kirsch edge enhancement operations before 2D wavelet feature extraction in order to make up for its deficiency of lacking of directional selectivity. The Kirsch directional operations produce four directional feature matrices, two endpoint matrices and a global character image matrix which incorporates with 20 geometrical features to form hybrid feature set I. As for the complex wavelet feature extraction, the merit of the directional selectivity and insensitivity to shifts of the input image is helpful for increasing its discriminant ability. We extract 160 complex wavelet features and 20 geometrical features to form hybrid feature set II. The experimental results show that hybrid feature set I has a higher recognition rate. The reason is that hybrid feature set I does not only extract directional information of the character image, but it also keeps character global and endpoints information. However, if we directly conduct 2D WT on a character image to extract the wavelet feature, the recognition rate will decrease substantially.

As the structure of 2D-CWT is similar to that of 2D-DWT, 2D-CWT has a dyadic fast algorithm. Experiments have shown that both proposed hybrid wavelet features make the ANN stable and converge quickly.

Theoretically, wavelet decomposition at a higher level will lose some information. In order to keep as much information as possible, one possible way is to decompose the character image into only the second level. As a result, there will be more features remaining. Our next research will focus on using the feature selection method to extract useful features from a large feature array and use them to recognize and verify handwritten numerals and characters.

### References

1. Ching. Y. Suen, Christine Nadal, Raymond Legault, Tuan A. Mai and Louisa Lam, Computer Recognition of Unconstrained Handwritten Numerals, Proc. IEEE 1992, Vol. 80, No. 7, pp. 1162-1180.
2. Cheng-Lin Liu, Kazuki Nakashima, Hiroshi Sako and Hiromichi Fujisawa, Handwritten Digit Recognition: Investigation of Normalization and Feature Extraction Techniques, Pattern Recognition, Volume 37, No. 2, pp. 265-279, 2004.
3. G. Y. Chen, T. D. Bui, A. Krzyzak, Image Denoising with Neighbour Dependency and Customized Wavelet and Threshold, Pattern Recognition, Vol. 38, No. 1, pp. 115-124, 2005.

4. P. Zhang, T. D. Bui, C. Y. Suen, Feature Dimensionality Reduction for the Verification of Handwritten Numerals, *Pattern Analysis and Applications*, Vol. 7, No. 3, pp. 296-307, 2004.

5. N. G. Kingsbury, Image Processing with Complex Wavelets, *Phil. Trans. R. Soc. Lond. A* 357, pp. 2543-2560, 1999.