Extraction of Hybrid Complex Wavelet Features for the Verification of Handwritten Numerals

P. Zhang, T. D. Bui, C. Y. Suen

Centre for Pattern Recognition and Machine Intelligence
Concordia University
1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8 Canada

Abstract

A novel hybrid feature extraction method is proposed for the verification of handwritten numerals. The hybrid features consist of one set of two dimensional complex wavelet transform (2D-CWT) coefficients and one set of geometrical features. As 2D-CWT does not only keep wavelet transform's properties of multiresolution decomposition analysis and perfect reconstruction, but also adds its new merits: its magnitudes being insensitive to the small image shifts and multiple directional selectivity, which are useful for handwritten numeral feature extraction. Experiments demonstrated that the features extracted by our proposed method can make the ANN classifier more reliable and convergence easily. A high verification performance has been observed in the series of experiments on handwritten numeral pairs and clusters.

Key Words: Feature Extraction, Wavelet Transform, Complex Wavelet Transform, Verification of Handwritten Numerals, Artificial Neural Networks.

1. Introduction

OCR is one of most successful applications of pattern recognition. Many methodologies have been published in the literature and various commercial OCR products have been in the markets in recent years [1, 2]. Current research on OCR is addressing more diversified and sophisticated problems, including the recognition of severely degraded, omni font mixed language texts and unconstrained handwritten texts. There are many ways to improve recognition performance.

One way is to improve the classifier's performance. There exist varieties of classifiers, for example, linear or nonlinear discriminating classifiers, decision tree classifiers, Neural Networks, HMM, SVMs and the combination of classifiers, etc.

Another way is to seek new feature extraction methods. Many feature extraction methods have been reported such as various moment features [3], transform features: Fourier and wavelet based features [4], gradient and distance-based features [5, 6], geometrical features [7], as well as hidden Markov model for unconstrained handwriting recognition [8].

The third way is the research on verification and validation of confusing characters in order to achieve higher recognition precision and better reliability, as well as its applications to the recognition and verification of handwritten numerals [9-13].

Wavelet has been widely used in the field of image processing for image enhancement, denoising, texture segmentation, etc. based on its properties of its multiresolution decomposition analysis and perfect reconstruction, etc. [14]. However, because Discrete Wavelet Transform (DWT) normally uses its decimated form with down-sampling, the coefficients of decomposed subband images will suffer from the following problems: the coefficients of DWT are very sensitive to a shift of the input image, and the subband images have poor directional selectivity. These problems limit its applications to pattern recognition.

Complex Wavelet Transform (CWT) has been developed in order to overcome DWT's deficiency. CWT adds some new merits such as approximate shift invariance, good directional selectivity for 2-D image, efficient order-N computation and limited redundancy. The computational complexity of CWT requires only twice that of DWT for 1-D (2^m times for m-D signal). The redundancy is independent of the number of scales: 2:1 for 1-D (2^m: 1 for m-D signal) [15]. These good properties have made CWT successfully applicable to image processing recently. However, the application of CWT to pattern recognition is almost a new research

field. CWT can be used for feature extraction for the recognition of handwritten characters as well as for the verification of handwritten characters. In this paper, we only discuss the feature extraction for the verification of handwritten numerals.

A brief review on methods of improving the performance on handwritten character recognition is given in the first section. Then the concept of DWT is analyzed and the structure of extracting CWT feature set is described in section II. Since only relatively few advanced studies have been conducted on verification and validation in the handwritten numerals, a verification scheme conducted on pair-wise and cluster characters by using our proposed feature extraction method is presented in Section III. In Section IV, some experimental results of handwritten numeral verification are listed. Finally, conclusions and future work will be proposed.

2. Complex Wavelet Transform and Feature Extraction

2.1 Wavelet Transform

For a continuous function $f(x)$, it is projected at each step j on the subset V_j, $(\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots)$. The scalar projection $c_{j,k}$ is defined by the dot product of $f(x)$ with the scaling function $\phi(x)$, which is dilated and translated:

$$c_{j,k} = <\ f(x), \phi_{j,k}(x) >,$$
$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k). \qquad \ldots (1)$$

where $\phi(x)$ is a scaling function. The difference between c_{j+1} and c_j is contained in the detailed component belonging to the space W_j, which is orthogonal to V_j.

$$W_j \oplus V_j = V_{j+1},$$
$$V_j \cap W_j = \{0\}, \ j \in Z. \qquad \ldots (2)$$

Suppose $\psi(x)$ is a wavelet function. The wavelet coefficients can be obtained by

$$w_{j,k} = <\ f(x), 2^{j/2} \psi(2^j x - k) >. \qquad \ldots (3)$$

Some relationships between $\phi(x)$ and $\psi(x)$ are listed below:

$$\frac{1}{2} \phi(\tfrac{x}{2}) = \sum_{i=0}^{n-1} h(i) \phi(x - i),$$
$$\frac{1}{2} \psi(\tfrac{x}{2}) = \sum_{i=0}^{n-1} g(i) \phi(x - i). \qquad \ldots (4)$$

where $h(i)$ and $g(i)$ represent unit impulse functions of lowpass and highpass filters, respectively, which are related to the scaling function $\phi(x)$ and the wavelet function $\psi(x)$; n is the length of the unit impulse functions. In other words, the low frequency components and high frequency components can be directly obtained by

$$c_{j-1,k} = \sum_{i=0}^{n-1} h(i - 2k) c_{j,i},$$
$$w_{j-1,k} = \sum_{i=0}^{n-1} g(i - 2k) c_{j,i}. \qquad \ldots (5)$$

In order to realize perfect reconstruction, the unit impulse functions $h(i)$ and $g(i)$ need to be carefully chosen to satisfy the following equation.

$$c_{j,k} = \sum_{i=0}^{n-1} h(k - 2i) c_{j-1,i} + g(k - 2i) w_{j-1,i}. \qquad \ldots (6)$$

According to the 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, HH with 2:1 down-sampling. Mathematically, we can express this recursive algorithm in the following equation.

$$x_{LL,k1,k2}^{(n-1)} = \sum_{l1,l2} h_{l1-2k1} h_{l2-2k2} x_{LL,l1,l2}^{(n)},$$
$$y_{LH,k1,k2}^{(n-1)} = \sum_{l1,l2} h_{l1-2k1} g_{l2-2k2} x_{LL,l1,l2}^{(n)},$$
$$y_{HL,k1,k2}^{(n-1)} = \sum_{l1,l2} g_{l1-2k1} h_{l2-2k2} x_{LL,l1,l2}^{(n)}, \qquad \ldots (7)$$
$$y_{HH,k1,k2}^{(n-1)} = \sum_{l1,l2} g_{l1-2k1} g_{l2-2k2} x_{LL,l1,l2}^{(n)}.$$

If the wavelet filters are real and we use Mallat's dyadic wavelet decomposition tree [14], which has a fast algorithm, the coefficients of decomposition will suffer from the following problems: lack of shift invariance and poor directional selectivity [15].

2.2 Complex Wavelet Transform for OCR Feature Extraction

In two dimensional complex wavelet transform (2D-CWT), we can set the basic functions to closely approximate complex Gabor-like functions, which exhibit strong characteristics of spatial locality and orientation selection, and are optimally localized in the space and frequency domains. Therefore, the 2D-CWT functions have following form:

$$h(x, y) = a(x, y)e^{j(w_x x + w_y y)}. \qquad \dots (8)$$

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

$$c_i^l = u_i^l + jv_i^l. \qquad \dots (9)$$

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

$$C_i^l = \sqrt{(u_i^l)^2 + (v_i^l)^2}. \qquad \dots (10)$$

Since $a(x,y)$ is slowly varying, the magnitude is insensitive to the small image shift.

The directional properties of the 2D-CWT arise from the fact that $h(x,y)$ has a constant phase along the lines such that $w_x x + w_y y$ is constant. 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 $\pm 15°$, $\pm 45°$, $\pm 75°$. These two properties are useful for pattern recognition.

2D-CWT can be implemented using a dual-tree structure. For each tree, its structure is similar to 2D-DWT, which has two decomposition operations on each level, namely row decomposition and column decomposition, except that the different filters are applied for perfect reconstruction and the outputs of subbands images are congregated into complex wavelet coefficients. Fig. 1 shows a 2D-CWT feature extraction scheme for the recognition and verification of handwritten numerals.

The dual-tree complex wavelet decomposition consists of two trees: Tree A and Tree B. These two trees have the same structure. In order to realize perfect reconstruction from decomposed subimages, a lowpass filter and a highpass filter at the first level need to be specially designed and denoted as Lop1 Hip1 for tree A; Lop2, Hip2 for tree B, which are called pre-filters. The other complex filters in the higher levels are set to Lo1 and Hi1 for Tree A, Lo2 and Hi2 for tree B. Interested readers can refer to reference [15] for further details.

A character image of size NxN is decomposed into four subband images: LL, LH, HL, HH at the first level of each tree and each of the subband images has a size of $\frac{N}{2} x \frac{N}{2}$. At the higher levels, the decompositions are based on LL subband image at the previous level. For example, if a 32x32 character is decomposed into the third level, the final size of each subband image is 4x4. Then we can extract two kinds of complex wavelet coefficients as features.
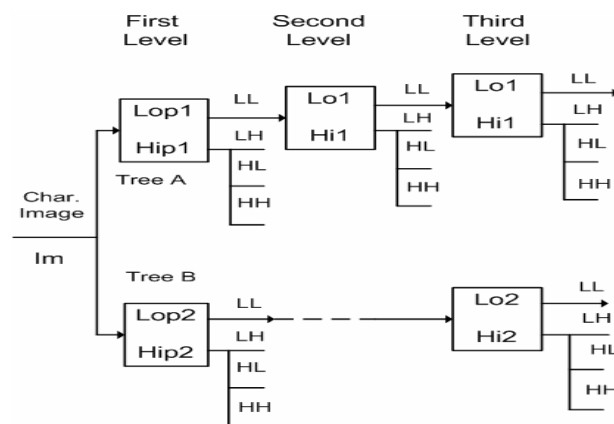

Fig. 1 The schematic diagram of 2D-CWT for character feature extraction

In the first method, the feature extraction is conducted at the third layer. We keep only amplitude coefficients for three high frequency components as well as both real part coefficients and imaginary part coefficients for low frequency component. The number of features = 4x4 (for each subband image) *3 (high frequency subband images for each tree) *2 (trees) +4x4 (for each subband image) *2 (trees)* 2(parts: real and imaginary) =160. As the real and imaginary coefficients of each LL subband image are extracted as features, the phase information is preserved as a good directional selectivity.

In the second method, we only use the magnitudes of each subimage at the third level as features. As a result, the phase information is lost. The No. of features =4x4 (for each subband image) * 4 (subimages for each tree) *2 (trees) =128.

2.3 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.

3. A Verification Scheme for Handwritten Numerals

It is common that the General Purpose Recognizer (GPR) may output two or more candidates with different confident values. Without verifier, the system will output the character with highest confidence value. In order to pursue a higher and reliable recognition rate, we need to build up a second level verification and validation engine (post-recognition) based on the GPR results, namely, to use our proposed verification engine to confirm a best candidate as the output. Fig. 2 shows a recognition and verification system.

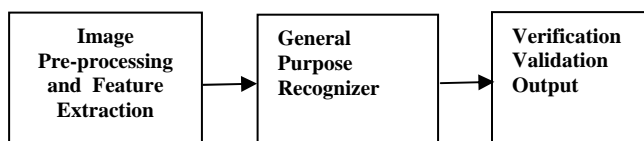| Image Pre-processing and Feature Extraction | → | General Purpose Recognizer | → | Verification Validation Output |

Fig. 2 A system architecture for OCR recognition and verification

There are four types of verifiers according to the number of classes [9]. Let Ω denote the working space of a verifier, and let $|\Omega|$ be the dimension of the space.

$|\Omega| = n$: General verifier, working on all classes in the problem.

$0 < |\Omega| < n$: Cluster verifier, verification in clustered categories (Is it a "4","6", or "9"?).

$|\Omega| = 2$: Pair-wise verifier, verification between two categories (Is it a "4" or "9"?).

$|\Omega| = 1$: Class-specific verifier, working on one candidate class (Is it a "1"?).

In this paper, we will only focus on pair-wise and cluster verification. Theoretically, for 10 numerals, there are 45 confusing digit pairs (10x9/2=45). From our observations based on GPR, some of the most confusing numeral pairs are {4,9}, {1,7}, {3,5}, {2,3}, {0,6}, {4,6}, {9,8}, {5,8}, {2,8}, {6,5}, {9,3}, {8,6}, etc.

For cluster verification, the clusters with three characters are chosen. Consequently, three verification engines for each cluster have been designed. For instance, for cluster {2, 3, 5}, the three verification

engines are designed for distinguishing (1): char "2" from chars "3" and "5"; (2): char "3" from chars "2" and "5"; (3): as well as char "5" from chars "2" and "3".

The MNIST handwritten numeral database is used to conduct the experiments. For pair-wise verification, the first 3000 samples of each character in the pair are used as training samples. The next 1000 samples of each character in the pair are extracted for verification, and the last 1000 samples of each character in the pair are used as testing samples. Some character images are shown in Fig. 3.

For cluster verification experiments, the same database is applied. However, for those sub cluster set with two chars, we will use fewer training samples in order to balance the overall training samples. For example, for distinguishing character "2" from chars "3" and "5". We will use 1500 training samples of chars "3" and "5" each.
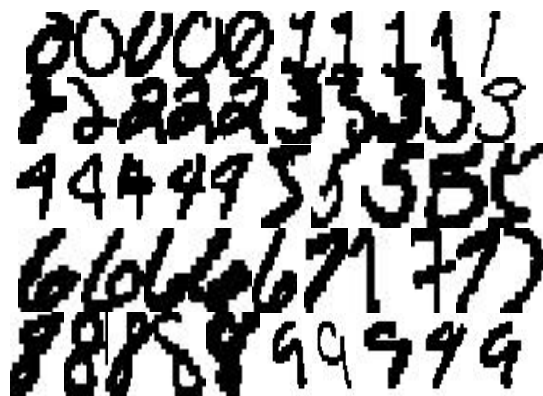
Fig. 3 Numeral images extracted from MNIST

Actually, the verification of character pairs is implemented by a two-classification classifier. A three-layer ANN with Back Propagation algorithm (BP) is employed as the verifier. The configuration of ANN is set as follows: No. of nodes in the input layer is the same as No. of features. No. of nodes in the hidden-layer is set at 20. No. of nodes in the output layer is set at No. of verification number. The termination condition of training the ANN depends on the conditions that either the No. of iterations reaches predefined limit or the verification rate conducted on verification data samples while training achieves a predefined accuracy.

4. Experiments

We conducted two kinds of experiments based on complex wavelet features, which are described in Section 2.1. The first feature vector (Feature Set I) has 160 complex wavelet features and 20 geometrical features;

whereas in the second feature vector (Feature Set II), there are 128 complex wavelet features and 20 geometrical features.

4.1 Verification Experiments Conducted on Numeral Pairs

Handwritten characters "4" and "9" can form a very confusing numeral pair among 10 handwritten numerals. We conducted verification experiments by using two sets of features. The verification results conducted on training samples, verifying and testing samples are listed in Table I. The verification(two-class recognition) rate conducted on training samples, the ANN error, and the verification rates conducted on numerals "4" and "9" with ANN training iterations are shown in Fig. 4. Table II lists the verification results of other character pairs conducted on the testing set by using feature set I.

Table I Verification results using two feature Sets

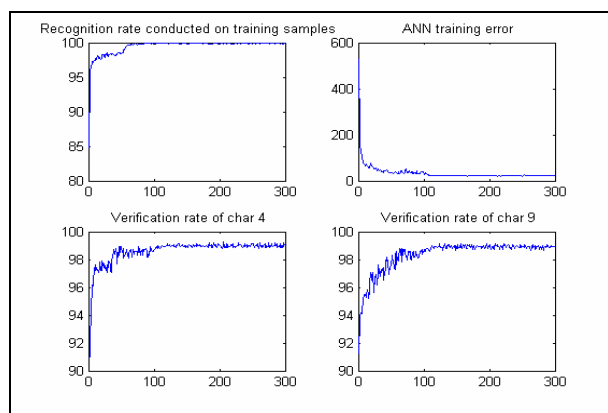| Feature Set I (180 features)	Correct Verification No.	Verification Rate
6000 training samples	5990	99.83%
2000 verifying samples	1983	99.15%
2000 testing samples	1982	99.10%
Feature Set II (148 features)		
6000 training samples	5985	99.75%
2000 verifying samples	1975	98.75%
2000 testing samples	1972	98.60%

Fig. 4 ANN training, verifying performance conducted on chars "4" and "9"

Table II Verification results obtained from char pairs

Character Pairs	Verification Results
1-7	99.40
3-5	99.10
2-3	99.15
0-6	99.30
4-6	99.25
9-8	99.20
5-8	99.25
2-8	99.30
6-5	99.15
8-6	99.15

From Tables I and II, it can be concluded that: (1): complex wavelet coefficients used as features for the verification of handwritten numerals are reliable and stable. (2): The classifier ANN trained by our proposed features can achieve a good recognition rate. These results compare favorably with those obtained by other researchers in the field. (3): The phase information of the complex wavelet transform helps to increase the verification rate provided that the input images have been de-skewed and normalized in the preprocessing procedure.

4.2 Verification Experiments Conducted on Clusters

It is possible that GPR output three or more candidates with similar confident values as these characters have similar features used in GPR. In order to investigate how our proposed complex wavelet features and geometrical features can be used to cluster verification, we conducted verification experiments on the some easily confusing clusters with three characters in each cluster, such as {2, 3, 5}, {1, 7, 9}, {4, 6, 9,},{0, 6, 9}, etc. As an example, the recognition rate conducted on the training set, the verification rates conducted on the verifying set and the testing set of cluster {2, 3, 5} are shown in Table III.

Table III   Recognition rates conducted on the training, verifying, and testing data for cluster {2, 3, 5}

Char pairs	2 - {3,5}	3 - {2,5}	5 - {2,3}
Training set	99. 65	99.55	99.60
Verifying set	99.10	99.10	99.15
Testing set	99. 05	99.10	99.10

Similar verification experiments are conducted on other three clusters {1, 7, 9}, {4, 6, 9} and {0, 6, 9}. Table IV lists the overall verification results on the testing sets for three clusters.

Table IV Verification rates conducted on testing sets of three clusters

Clusters	Overall verification rate (%)
{1,7,9}	99.20
{4,6,9}	99.10
{0,6,9}	99.15

5. Conclusions

Two CWT features are proposed in this paper. One is the complex wavelet magnitude features which are invariant to the shift of input character images and another set of features which keeps the phase information. Experiments demonstrated that the feature set with phase information can achieve a higher verification rates when tested on the normalized handwritten numerals. It means the directional selectively property of 2D-CWT is useful for the verification of normalized handwritten numerals.

As the structure of 2D-CWT is similar to that of 2D-DWT, 2D-CWT has a dyadic fast algorithm. Experiments have shown that our proposed hybrid complex 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. Cheng-Lin Liu, Kazuki Nakashima, Hiroshi Sako and Hiromichi Fujisawa, Handwritten Digit Recognition: Investigation of Normalization and Feature Extraction Techniques, Pattern Recognition, Vol. 37, No. 2, 2004, pp. 265-279.

2. 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.

3. S. O. Belkasim, M. Shridhar and A. Ahmadi, Pattern Recognition with Moment Invariants: A Comparative Study and New Results, Pattern Recognition, Vol. 24, No. 12, Dec. 1991, pp. 1117-1138.

4. Guangyi Chen and Tien D. Bui, Invariant Fourier-Wavelet Descriptor for Pattern Recognition, Pattern Recognition, Vol. 32, No. 7, July 1999, pp. 1083-1088.

5. Meng Shi, Yoshiharu Fujisawa, Tetsushi Wakabayashi, Fumitaka Kimura, Handwritten Numeral Recognition Using Gradient and Curvature of Gary Scale Image, Pattern Recognition, Vol. 35, No. 10, 2002, pp. 2051-2059.

6. II-Seok Oh, C. Y. Suen, Distance Features for Neural Network-based Recognition of Handwritten Characters, International Journal on Document Analysis and Recognition, 1998, pp. 73-88.

7. Qivind Due Trier, Anil K. Jain and Torfinn Taxt, Feature Extractions for Character Recognition A-Survey, Pattern Recognition, Vol. 29, No. 4, April 1996, pp. 641-662.

8. Se-Chang Oh, Jin-Young Ha and Jin-H. Kim, Context Dependent Search in Interconnected Hidden Markov Model for Unconstrained Handwriting Recognition, Pattern Recognition, Vol. 28, No. 11, Nov. 1995, pp. 1693-1704.

9. Jie Zhou, Adam Krzyzak, Ching Y. Suen, Verification- A Method of Enhancing the Recognizers of Isolated and Touching Handwritten Numerals, Pattern Recognition, Vol. 32, No. 5, 2002, pp. 1179-1189.

10. L. S. Oliveira, R. Sabourin, F. Bortolozzi, C. Y. Suen, Impacts of Verification on a Numeral String Recognition System, Pattern recognition Letters, Vol. 24, No. 7, July 2002, pp. 1023-1031.

11. Xiangyun Ye, Mohamed Cheriet, Ching Y. Suen, StrCombo: Combination of String Recognizers, Pattern Recognition Letters, Vol. 23, No. 4, April 2002, pp. 381-394.

12. Luiz S. Oliveira, Robert Sabourin, Flavio Bortolozzi, Ching Y. Suen, Automatic Recognition of Handwritten Numerical Strings: A Recognition and Verification Strategy, IEEE Transaction on PAMI, Vol. 24, No. 11, Nov. 2002, pp. 1438-1454.

13. Zhixin Shi, V. Govindaraju, Segmentation and Recognition of Connected Handwritten Numeral Strings, Pattern Recognition, Vol. 30, No. 9, Sept. 1997, pp. 1501-1504.

14. Mallat, S. G. A Thory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Trans. PAMI, Vol. 11, 1989, pp. 674-693.

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