PERGAMON

Pattern Recognition 32 (1999) 1083–1088

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

# Invariant Fourier-wavelet descriptor for pattern recognition

Guangyi Chen, Tien D. Bui*

*Department of Computer Science, Concordia University, 1455 De Maisonneuve West, Montreal, Que, Canada H3G 1M8*

Received 10 April 1997; accepted 16 October 1998

**Abstract**

We present a novel set of descriptors for recognizing complex patterns such as roadsigns, keys, aircrafts, characters, etc. Given a pattern, we first transform it to polar coordinate $(r, \theta)$ using the centre of mass of the pattern as origin. We then apply the Fourier transform along the axis of polar angle $\theta$ and the wavelet transform along the axis of radius $r$. The features thus obtained are invariant to translation, rotation, and scaling. As an example, we apply the method to a database of 85 printed Chinese characters. The result shows that the Fourier-wavelet descriptor is an efficient representation which can provide for reliable recognition. © 1999 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

*Keywords*: Feature extraction; Fourier transform; Invariant descriptor; Multiresolution analysis; Pattern recognition; Wavelet transform

## 1. Introduction

Feature extraction is a crucial processing step for pattern recognition [1]. Some authors [2–5] extract 1-D features from 2-D patterns. The advantage of this approach is that we can save space for the database and reduce the time for matching through the whole database. The apparent drawback is that the recognition rate may not be very high because less information from the original pattern is retained. In this paper we use 2-D features for pattern recognition in order to achieve higher recognition rate.

Fourier descriptor has been a powerful tool for pattern recognition [6–11]. It has many useful properties, one of which is that shifts in the time domain do not affect the spectrum in the frequency domain, i.e. Fourier transform is translation invariant with respect to the spectrum. However, the frequency information obtained from the Fourier descriptor is global, and, therefore, local variation of the shape can affect all the Fourier coefficients. In addition, the Fourier descriptor does not have a multiresolution representation. Therefore, we seek to develop descriptors that have better properties.

In the past few years, wavelet basis functions have become popular for localized frequency analysis, because they have short-time resolution for high frequencies and long-time resolution for low frequencies. Although wavelet descriptors have many advantages, they are not translation invariant. A small shift of the original signal will cause totally different wavelet coefficients. This is the reason why wavelet transforms are not widely used in pattern recognition community.

Because both the Fourier and wavelet descriptors have good properties and drawbacks, we are going to combine them in order to compensate each other to obtain a descriptor which is not only invariant under translation,

*Corresponding author. Tel.: 001 514 848 3014; Fax: 001 514 848 2830; E-mail: bui@cs.concordia.ca


PII: S0031-3203(98)00148-4

rotation and scaling, but also has a multiresolution matching ability. It should be mentioned that both Fourier transform and wavelet transform used in this paper are discrete transforms.

The paper is organized as follows. Section 2 derives the algorithm and provides a brief overview of its connection with the ring-projection approach. Section 3 introduces a set of wavelet filters and shows the multiresolution technique of wavelet transform. And finally, as an example, Section 4 gives experimental results for recognizing printed Chinese characters.

## 2. Fourier-wavelet descriptor for pattern recognition

Given an $N \times N$ pattern image $f(x, y)$ which may consist of several disconnected parts such as roadsigns or oriental characters, we are going to derive invariant features from it. The translation invariance can be achieved by translating the origin of the coordinate system to the centre of mass of the pattern, denoted by $(x_0, y_0)$.

The scale invariance can be obtained by transforming the pattern image into polar coordinate system. Let $d = \max_{f(x,y)\neq 0} \sqrt{(x-x_0)^2 + (y-y_0)^2}$ be the longest distance from $(x_0, y_0)$ to a point $(x, y)$ on the pattern. We draw $N$ concentric circles centered at $(x_0, y_0)$ with radius $(d \times i)/N$, $i = 1, 2, \ldots, N$. Also, we form $N$ angularly equispaced radial vectors $\theta_i$ departing from $(x_0, y_0)$ with angular step $2\pi/N$. For any small region

$$S_{ij} = \{(r, \theta) | r_i < r \leqslant r_{i+1}, \theta_j < \theta \leqslant \theta_{j+1}\}$$

we calculate the average value of $f(x, y)$ over this region, and assign the average value to $g(r, \theta)$ in the polar coordinate system. The feature $g(r, \theta)$ obtained in this way is also invariant to scaling, but the rows may be circularly shifted if we use different orientation.

With regard to rotational invariance, we can apply 1-D Fourier transform along the axis of the polar angle $\theta$ of $g(r, \theta)$ to obtain its spectrum. Since the spectra of Fourier transform of circularly shifted signals are the same, we obtain a feature which is also rotation invariant. Wavelet coefficients represent pattern features at different resolution levels, and, therefore, we apply wavelet transform along the axis of radius of the resulting $G(r, \phi)$ so that we can query the pattern feature database from coarse scales to fine scales.

The pattern feature database contains all scales of wavelet coefficients for each pattern in the training data set. For the coarsest scale, the target feature is matched against all possible patterns in the database. Because the number of coefficients in the coarse scale is quite small, the matching process can be carried out quickly even though the number of patterns in the database may be very large. During each scale we have three decisions to make: (1) If only one valid target identification is found, we terminate the matching process and mark the target to be unumbiguously identified; (2) If all patterns have to be rejected, then we stop the querying process and mark the target as an unknown target; (3) If we have more than one valid target identifications, we move on to the next finer scale and follow the same procedure as above. Because at finer scales we only need to consider those patterns marked to be refined by the last step, we can fulfill the querying process quickly even though we have more coefficients at finer scales. The matching process continues in this way until we identify the target or reject it.

The steps of the algorithm called PFW can be summerized as follows:

1. Find the centroid of the pattern $f(x, y)$ and transform $f(x, y)$ into polar coordinate system to obtain $g(r, \theta)$.
2. Conduct 1-D Fourier transform on $g(r, \theta)$ along the axis of polar angle $\theta$ and obtain its spectrum:

   $$G(r, \phi) = |FT_\theta(g(r, \theta))|.$$

3. Apply 1-D wavelet transform on $G(r, \phi)$ along the axis of radius $r$:

   $$WF(r, \phi) = WT_r(G(r, \phi)).$$

4. Use the wavelet coefficients to query the pattern feature database at different resolution levels.

Fig. 1 is the block diagram of the PFW algorithm. Fig. 2 depicts how a printed Chinese character is transformed after each step of the PFW algorithm. Fig. 2a is the character in $(x, y)$ -coordinate system. Fig. 2b is the polarized character $g(r, \theta)$ in polar coordinate system

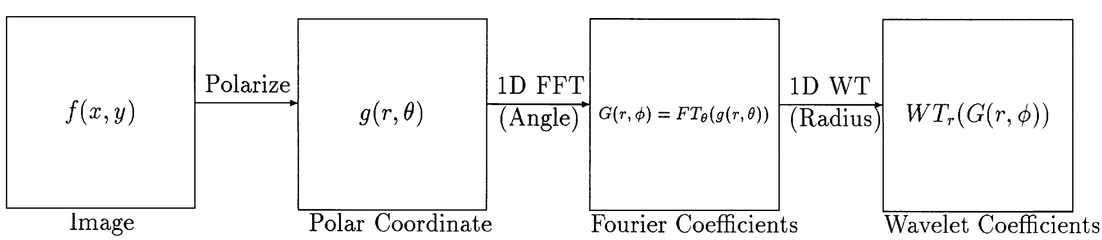


Fig. 1. Block diagram of the PFW algorithm.

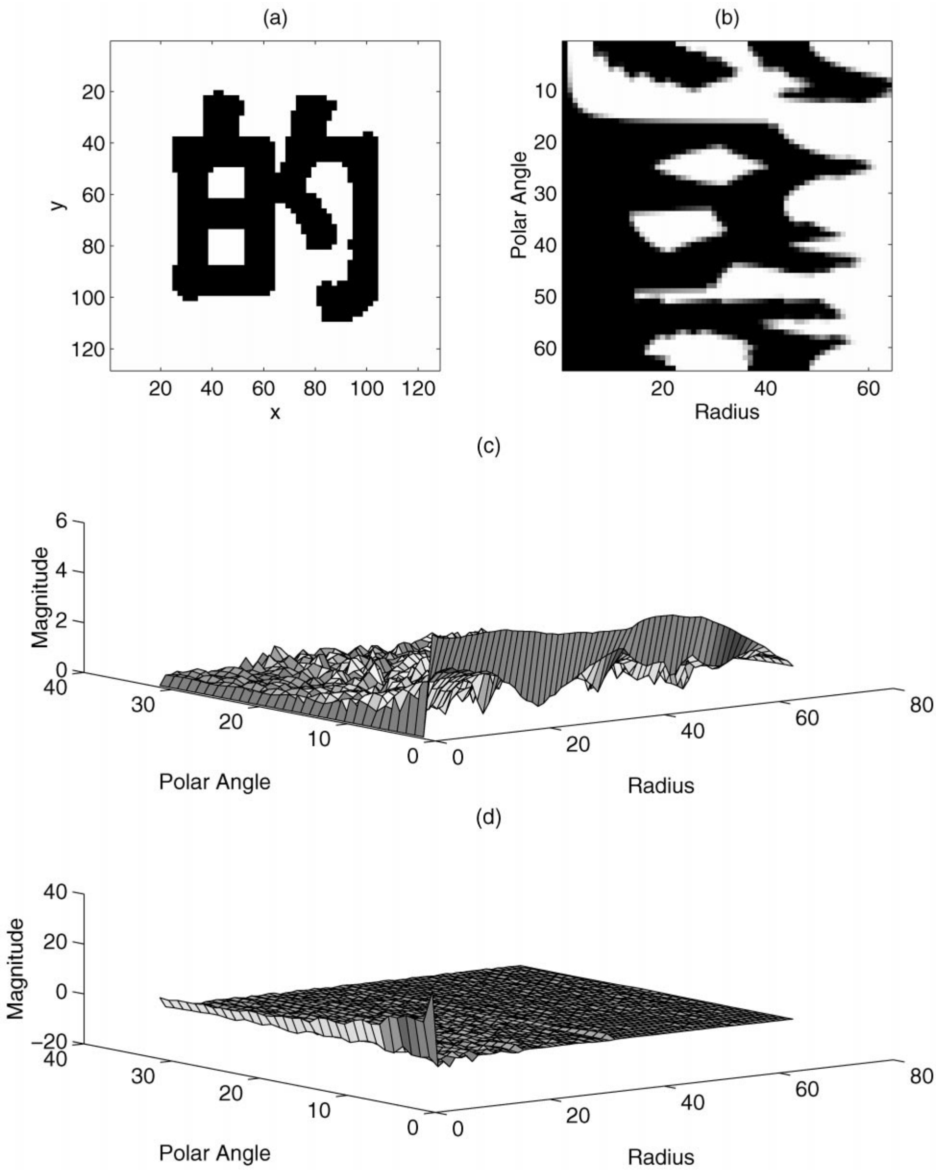


Fig. 2. An illustration of how a printed Chinese character is transformed after each step of the PFW algorithm. (a) The original printed Chinese character in Cartesian coordinates. (b) The polarized character in polar coordinates where each unit in the axis of the polar angle represents 6°. (c) The Fourier spectrum of the polarized character. (d) The wavelet coefficients based on the Fourier spectrum.

where each unit in the axis of the polar angle represents 6°, Fig. 2c shows the spectrum density of the Fourier transform $G(r, \phi) = |FT_\theta(g(r, \theta))|$, and Fig. 2d shows the wavelet coefficients $WF(r, \phi) = WT_r(G(r, \phi))$.

It is noted that the features extracted by the PFW algorithm are a superset of that of the ring-projection approach. Tang et al. [4] introduces the ring-projection of a pattern by

$$P(r) = \int_0^{2\pi} f(r\cos\theta, r\sin\theta)\,\mathrm{d}\theta,$$

where $r$ is the radius of the ring. It is shown that $P(r)$ is equal to the pattern mass distributed along circular rings. From Fourier transform we have

$$G(r, \phi) = \frac{1}{N}\sum_{\theta=0}^{N-1} g(r, \theta)\mathrm{e}^{-i2\pi\theta\phi/N}.$$

When $\phi = 0$, we get the average value along the axis of radius $r$:

$$G(r, 0) = \frac{1}{N}\sum_{\theta=0}^{N-1} g(r, \theta),$$

i.e.

$$G(r, 0) = \frac{1}{N} P(r).$$

The PFW algorithm extracts more features from the pattern than the ring-projection approach does. Therefore, we can expect that the PFW algorithm gives higher recognition rate.

## 3. Wavelet and multiresolution analysis

The wavelet transform [12–14] is well suited for localized frequency analysis, because the wavelet basis functions have short-time resolution for high frequencies and long-time resolution for low frequencies. In addition, wavelet representation provides a coarse-to-fine strategy, called *multiresolution matching* [15]. The matching starts from the coarsest scale and moves on to the finer scales. The costs for different levels are quite different. Since the coarsest scale has only a small number of coefficients, the cost at this scale is much less than for finer scales. In practice, the majority of patterns can be unambiguously identified during the coarse scale matching, while only few patterns will need information at finer scales to be identified. Therefore, the process of multiresolution matching will be faster compared to the conventional matching techniques.

The basic equation of the multiresolution analysis theory is the dilation (or scaling) equation

$$\phi(x) = \sqrt{2} \sum_k h_k \phi(2x - k),$$

which defines the cascade of the multiresolution approximation space using the wavelet family with the scaling function $\phi$. It has been shown that except for the Haar wavelet, all other wavelets with desirable properties can only be expressed by implicit equations. Nevertheless, once the coefficients $h_k$'s are known, all other properties of this family are completely determined.

Associated with the scaling function $\phi$, we can define the wavelet function $\psi$ by

$$\psi(x) = \sqrt{2} \sum_k g_k \phi(2x - k),$$

where $g_k = (-1)^k h_{1-k}$.

In order to compare the performance of different wavelet families, we use the following wavelet filters (see Fig. 3). These wavelet filters are reproduced from WAVELAB developed by D.L. Donoho.

The Haar filter is discontinuous, and can be considered a Daubechies-2. Its scaling filter is

$$h = (1/\sqrt{2}, 1/\sqrt{2}).$$

The Daubechies-4 filter has its advantage on its most compact support of 4 and its orthonormality. The size

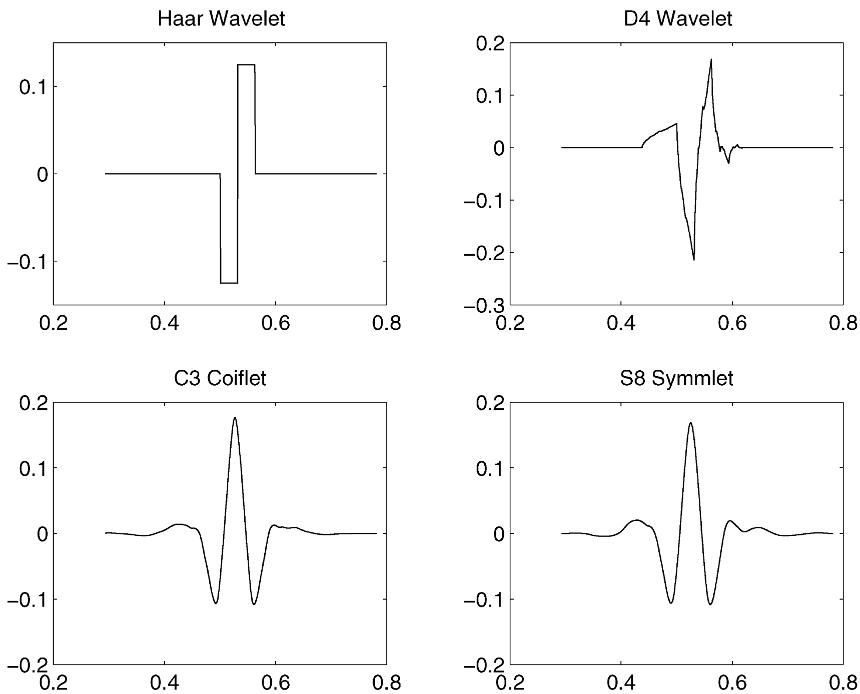


Fig. 3. The wavelet families used in our experiment.

4 is indeed shortest even span in which the second derivatives are computable. Its scaling filter is

$$h = (0.482962913145, 0.836516303738, 0.224143868042, -0.129409522551).$$

The Coiflet filters are designed to give both the mother and father wavelets 2, 4, 6, 8, or 10 vanishing moment (see Daubechies for the definition of vanishing moments and their usefulness). Here we only test the 2 vanishing moment case. Its scaling filter is

$$h = (0.038580777748, -0.126969125396, -0.077161555496, 0.607491641386, 0.745687558934, 0.226584265197).$$

The Symmlet-8 is the least asymmetric compactly supported wavelets with 8 vanishing moments. Its scaling filter is

$$h = (-0.107148901418, -0.041910965125, 0.703739068656, 1.136658243408, 0.421234534204, -0.140317624179, -0.017824701442, 0.045570345896).$$

## 4. Experimental results

In order to test the efficiency of the PFW algorithm, we use a set of 85 printed Chinese characters in our experiment. In Zhang et al. [11], Fourier descriptors together with a new associative memory classifier for recognition were developed and tested on the same set of Chinese characters. The original Chinese character is represented by $64 \times 64$ pixels, and so is the polarized character. Since the spectrum of 1-D Fourier transform is symmetric, we only keep half of the Fourier coefficients. Therefore, the size of $G(r, \phi)$ is $64 \times 32$ and so is the size of the wavelet coefficients $WF(r, \phi)$.

Because translation will not change the relative position of the centre of mass of the character, our major concern is the system's performance on rotation and scaling. For each character, we tested six rotation angles and six scaling factors. The six rotation angles are 30, 60, 90, 120, 180 and 270°, and the six different scaling factors are 0.5, 0.6, 0.7, 0.8, 0.9 and 1.2. We get 100% recognition rate for all the rotation angles. The recognition results for different scaling factors are given in Table 1, while the recognition results for a combination of rotation and scaling are shown in Table 2. These results demonstrate the effectiveness of this feature extraction algorithm against geometric distortion.

Table 1
Recognition result for different scaling factors

| Percentage (%) | Scaling factor | | | | | |
|---|---|---|---|---|---|---|
| | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.2 |
| Recognition rate | 98.82 | 100 | 100 | 100 | 100 | 100 |
| Error rate | 0 | 0 | 0 | 0 | 0 | 0 |
| Rejection rate | 1.18 | 0 | 0 | 0 | 0 | 0 |

Table 2
Recognition rate for different rotation angles and scaling factors

| Scaling factor | Rotation angle | | | | | |
|---|---|---|---|---|---|---|
| | 30° (%) | 60° (%) | 90° (%) | 120° (%) | 180° (%) | 270° (%) |
| 0.5 | 97.65 | 96.47 | 92.94 | 90.59 | 92.94 | 95.29 |
| 0.8 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1.2 | 100 | 100 | 100 | 100 | 100 | 100 |

The above experimental results are obtained from the Haar wavelet. We also test the Daubechies-4, Coiflet-3, and Symmlet-8 wavelets. The experimental results are nearly identical for all wavelet families considered. That is, we obtain a 100% correct classification for most test cases. Because the wavelet coefficients of a signal have multiresolution representation of the original signal, we use a coarse-to-fine matching strategy. The coarse scale wavelet coefficients normally represent the global shape of the signal, while the fine scale coefficients represent the details of the signal. Due to noise introduced in the original image and the errors accumulated in the process of polarization, the detail coefficients are becoming less important than the coarse scale coefficients. However, for characters with similiar shapes the coefficients at the finer scales must be used for efficient discrimination among these shapes.

## 5. Conclusion

The PFW algorithm proposed by this paper is a computational reliable tool for pattern recognition. The algorithm is invariant to translation, rotation, and scaling. We achieve very high recognition rate for all different rotation angles and scaling factors by using different wavelets. We employ a multiresolution matching technique in our algorithm so that the matching process can be accomplished cheaply. One should note that although our experiments are done on a set of printed Chinese characters, our method is equally applicable to other pattern recognition problems such as airplanes, key sets, or roadsigns. Future work can also be done for recognizing more deformed and noisy patterns by incorporating neural network into the PFW algorithm.

## Acknowledgements

This work was supported by research grants from the Natural Sciences and Engineering Research Council of Canada and by the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche of Quebec.

## References

[1] O.D. Trier, A.K. Jain, T. Taxt, Feature extraction methods for character recognition – a survey, Pattern Recognition 29 (1996) 641–662.

[2] G.C.-H. Chang, C.-C. Jay Kuo, Wavelet descriptor of planar curves: theory and applications. IEEE Trans. Image Process. 5 (1996) 56–70.

[3] P. Wunsch, A.F. Laine, Wavelet descriptors for multiresolution recognition of handprinted characters, Pattern Recognition 28 (1995) 1237–1249.

[4] Y.Y. Tang, B.F. Li, H. Ma, J. Liu, C.H. Leung, C.Y. Suen, A novel approach to optical character recognition based on ring-projection-wavelet-fractal signatures, ICPR'96, vol II, 1996, pp. 325–329.

[5] S.-S. Wang, P.-C. Chen, W.-G. Lin, Invariant pattern recognition by moment Fourier descriptor, Pattern Recognition 27 (1994) 1735–1742.

[6] G.H. Granlund, Fourier processing for handwritten character recognition, IEEE Trans. Comput. 21 (1992) 195–201.

[7] C.T. Zahn, R.Z. Roskies, Fourier descriptors for plane closed curves, IEEE Trans. Comput. C-21 (1972) 269–281.

[8] E. Persoon, K.-S. Fu, Shape discrimination using Fourier descriptors, IEEE Trans. Systems Man Cybernet. SMC-7 (1977) 170–179.

[9] A.E. Grace, M. Spann, A comparison between Fourier–Mellin descriptors and moment-based features for invariant recognition using neural networks, Pattern Recognition Letters 12 (1991) 635–643.

[10] M. Shrihar, A. Badreldin, High accuracy character recognition algorithm using Fourier and topological descriptors, Pattern Recognition 17 (1984) 515–524.

[11] M. Zhang, C.Y. Suen, T.D. Bui, Feature extraction in character recognition with associative memory classifier, Int. J. Pattern Recognition Artif. Intell. 10 (1996) 325–348.

[12] C.K. Chui, An introduction to wavelets, Academic Press, Boston, 1992.

[13] I. Daubechies, Ten lectures on wavelets, SIAM, Philadelpha, 1992.

[14] I. Daubechies, Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math. 41 (1988) 909–996.

[15] S. Mallat, Multiresulotion approximations and wavelet orthonormal bases of L2(R), Trans. Amer. Math. Soc. 315 (1989) 69–87.

**About the Author**—GUANGYI CHEN received the B.S. in Applied Mathematics from Dalian University of Technology in 1986, and the M.S. in Computing Mathematics from the Computing Center of the Chinese Academy of Sciences in 1989. He was a research associate at Shenyang Institute of Computing Technology of the Chinese Academy of Sciences during 1989–1994. He is now a graduate student in the Computer Science Department of Concordia University, Canada. His research interests include numerical methods, pattern recognition, image processing, and neural networks.

**About the Author**—TIEN D. BUI is Full Professor in the Department of Computer Science, Concordia University, Montreal, Canada. He was Chair of the Department from 1985 to 1990, and Associate Vice-Rector Research at the same University from 1992 to 1996. Dr. Bui has published more than 120 papers in many different areas in scientific journals and conference proceedings. He was an invited professor at the Instituto per le Applicazioni del Calcolo in Rome under the auspices of the National Research Council of Italy (1978–1979), and a visiting professor at the Department of Mechanical Engineering, the University of California at Berkeley (1983–1984). He has received many research grants and contracts from governments and industries. His current research interests are in the area of mathematical modeling and image processing. He is a Fellow of the British Physical Society, a Senior Member of the Society for Computer Simulation, and a Member of the IEEE. He is an Associate Editor of the journal Simulation and listed in Who's Who in the World and American Men and Women of Science.