J. of Basic Sciences, Cheju Nat'l Univ.
11(2), 25~34, 1998

기초과학연구 제주대학교
11(2), 25~34, 1998

A Solution to translation of the two dimensional wavelet transform using a second-order neural network

Cheol-Min Kim and Bong-Kyu Lee

1. Introduction

Image processings using the wavelet transform in two dimensions (2D-WT) have been a very active research area in recent years because the 2D-WT possess many good properties. For example, the 2D-WT provides a simple hierarchical framework for interpretating the image information and enable us to have a scale-invariant interpretation of a image. Also, important information of a image, such as edges, is better characterized by 2D-WT due to its good locality in both time and frequency domains[1,2,3]. In addition, the computation of the 2D-WT of a given image is very fast and efficient with multirate filterbank.

Daubechies[4] studied important properties of a wavelet transform and its discretization. Meyer and Stromberg found a very important class of discrete wavelet transform called orthogonal wavelets[1,2,3]. Mallat[5] showed that the computation of the wavelet transform can be achieved with a pyramidal algorithm based on convolutions with quardrature mirror filters. These algorithms are a good tool for image processings.

However, the discrete 2D-WT can not be used for pattern recognition directly. In pattern recognition, patterns might be located anywhere in the image. Hence, it is necessary to build an image representation which is invariant to translation. The discrete 2D-WT does not have this property[5,6]. Its results are given in a translated representation including wrap-around.

There have been three different approaches to solve the translation problem of the discrete wavelets. The first approach focuses on finding wavelet bases that are amicable to translation while keeping the methods of calculating wavelet transform. These methods are efficient in terms of computation complexity. However, in many cases, it is difficult to find wavelet bases for the specific domain. The second approach is to use a large redundant library of wavelet bases that correspond to different translations[5,6,7]. Mallet represented a signal using the zero crossing representation in the wavelet domain. Though the second approach offers a lot of flexibility, the computation is usually intensive and the orthogonality is not preserved. The third approach is to use variable sampling rates or adaptive sampling rates[8,9]. However this approach considerably increases the redundancy of the representation. Recently, Liang and Parks[9] proposed a new method based on the third approach. They described a translation invariant 2D-WT using symmetric extension and tree search algorithm. This method satisfies translation invariant, size limited(= the output coefficients are of the size $N \times N$ for a $N \times N$ input image) and has reduced edge effects. But, requirements, such as additional preprocessing step(=symmetric extension of a input signal) and a cost function for a tree search algorithm, make it hard to understand and implement.

In this paper, we propose a new method to solve the translation problem of the discrete 2D-WT The proposed method uses a second-order neural network which has translation invariance including wrap around. The method is simpler and can be implemented easily than the existing methods. 80 images (20 images per one airplane) were used for the experiments. Each set of 20 images has the same but translated airplane pattern. Our method recognizes 80 images correctly regardless of the positions of the pattern. These results indicates the efficiency of the proposed system.

2. The discrete two dimensional Wavelet Transform (2D-WT)

Morlet[10] defined the wavelet transform by decomposing the signal into a family of functions which are the translation and dilation of a unique function $\psi(x)$ called a wavelet.

The two dimensional wavelet transform of a function $f(x,y) \in L^2(R^2)$ at the scale s and a point *(u, v)* is defined by [11]

$$Wf(s,(u,v)) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)s\Psi(s(x-u),s(y-v))\,dxdy \qquad (1)$$

where $\Psi(x,y) \in L^2(R^2)$ is a function whose Fourier transform $\psi(\omega_x, \omega_y)$ satisfies

$$\forall\ (\omega_x,\omega_y) \in R^2 \int_0^{+\infty} \frac{\psi(s\omega_x, s\omega_y)^2}{s}\, ds = C_\Psi < +\infty \tag{2}$$

Let $\Psi_s(x,y) = s\Psi(sx, sy)$ and $\widetilde{\Psi}_s(x,y) = s\Psi(-x,-y)$. The wavelet transform of $f(x,y)$ at the scale s and a point (u, v) can be rewritten as a convolution product

$$Wf(s,(u,v)) = f * \widetilde{\Psi}_s(u,v) \tag{3}$$

The two dimensional wavelet transform can be discretized by sampling both the scale and the translation parameters [12]. A sequence of scales $(a^j)_{j\in Z}$ is selected, where a is the elementary dilation step. For each scale a_{j}, the translation vector (u,v) is uniformly sampled. Let βa^{-j} be the sampling rate at the scale a^j. The discrete 2D-WT is defined by

$$W_d f(j,(n,m)) = Wf(a^j, (\frac{n\beta}{a^j}, \frac{m\beta}{a^j})) = f * \widetilde{\Psi}_{a^j}(\frac{n\beta}{a^j}, \frac{m\beta}{a^j}) \tag{4}$$

A discrete wavelet transform for $a=2$ and $\beta=1$ corresponds to a wavelet orhonormal basis. Mallet[5] showed that the wavelet transform of a function in such a wavelet orthonormal basis can be computed with a quadrature mirror filter bank. Therefore, the discrete 2D-WT can be computed with a pyramidal algorithm based on convolutions with quadrature mirror filters[2,5,11,12,13,14,15]. This algorithm is illustrated by the Figure 1. The filters G and H are the quadarture mirror filters.

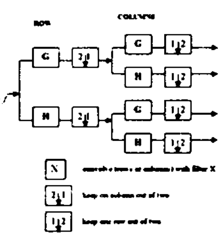

Figure 1. Decomposition of an image $A^d_{2^{j-1}}f$ into $A^f_{2^j}$ and $D^1_{2^j}f$, $D^2_{2^j}f$ and $D^3_{2^j}f$

Although the discrete 2D-WT has good capabilities for the image processing, it cannot be used for pattern recognition application. In order to use the discrete 2D-WT for pattern recognition applications, we must be able to build models of patterns within the wavelet representation. The patterns can be located anywhere in the image. Hence, the models must be independent from the pattern location. When a pattern is translated, its model should only be translated but not modified. Unfortunately, the discrete 2D-WT does not verify this translation invariant property[5,6,7,8]. Figure 2 shows the two original images composed of the same but translated airplane pattern and results of their one layer decompositions. In the figure 2, it can be seen that the decomposed images are varied according to the positions of the pattern in images. They are translated with possible wrap around. This problem must be solved for the effective use of the discrete 2D-WT in pattern recognitions.

(a) original images (b) transformed images

Figure 2. Various positions of a pattern in images and their transformed images

3. Wrap-translation invariant second-order neural networks

Since the discrete 2D-WT maps input images into translated subimages with possible wrap around, it cannot be used for pattern recognition directly. In this section, we present a second-order neural network which has translation and/or wrap-around(wrap-translation) invariant property.

The net input h_i for a node i in a second-order neural network is given as follows[17].

$$h_i = \sum_j W_{i(jj)} x_j^2 + \sum_j \sum_k W_{i(jk)} x_j x_k + \sum_j W_{ij} x_j + W_i \tag{5}$$

$x_j's$: the pixel values of the input nodes

In equation (5), we need to take only the two dimensional correlation terms into account to obtain wrap-translation invariance. Second-order neural networks can be trained for invariances by adding constraints to weights. To obtain constraints for wrap-translation invariance, types of distance between $x_j's$ must be divided into two categories, the inner distance and the outer distance. The inner distance $|k-j|$ is the euclidean distance between two points. The outer distance is defined as $|n-k+j|$, where n is the size of an input. For wrap-translation invariant learning, weights corresponding to wrap-translated positions should be equal as follows[18].

$$W_{i(jk)} = W_{i|k-j|} = W_{i|n-k+j|} \tag{6}$$

Thus, the new learning rule for wrap-translation invariance is given by (7).

$$h_i = \sum_j \sum_k (W_{i|j-k|} x_j x_k + W_{i|n-j+k|} x_j x_k) \tag{7}$$

Since all weights in wrap-translated positions have the same values, all the combinations of two pixels having the same inner and outer distances are added beforehand and the cumulative value can be represented by a single input node of the second-order neural network. This feature, summation of all the combinations for given relative position, is called SOP (Summation Of Products) or second-order feature[18,19,20]. Using second order features, we can obtain the wrap-translation invariant representations of the wavelet transformed subimages. The second order neural network using second-order features was used to solve the translation problem of the discrete 2D-WT.

4. Experimental Results

Experiments were performed using the images which have a pattern with various positions. Patterns within images composed of four kinds airplanes. 20 images per each airplane were used for experiments. Figure 3 shows the airplanes within images. The images were digitized (8 bits grey-level) using a scanner and normalized to 128×128.

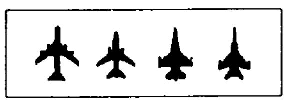

Figure 3. Airplanes used in the experiments

The proposed method consists of two stages, the decomposition stage and the recognition stage. Figure 4 shows the block diagram of the proposed system.

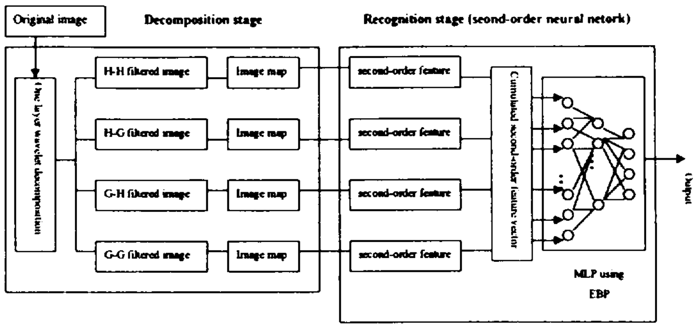

Figure 4. block diagram of the proposed system

In the decomposition stage, original 128×128 image is decomposed with the discrete 2D-WT into four 64×64 subimages. The 12 coefficients of the impulse response $h(n)$ of the filter H is given by Table 1. The impulse response of the G is obtained by equation (8).

<Table 1> Filter coefficients used in the wavelet transform

n	$h(n)$	n	$h(n)$
0	0.542	6	0.012
1	0.307	7	-0.013
2	-0.035	8	0.006
3	-0.078	9	0.006
4	0.023	10	-0.003
5	-0.030	11	-0.002

The four subimages of each image can be translated with possible wrap around according to the positions of an airplane within an image and are also grey level images. Pixel connectivity is not critical in current experiments. Thus, grey level subimages are converted into binary images using 3×3 gradient operator[21]. Four binary images called the edge map are obtained after the operation. But they still have translation variation. The four edge maps should be feed into the second-order neural network to solve the translation with possible wrap around.

The recognition stage consists of the second-order feature extractor and the Multilayer Perceptron (MLP) trained by EBP(Error Backpropagation). Generally, second-order neural networks are a form of preprocessing for the standard backpropagation neural networks that use geometrically motivated nonlinear combination of image pixels to achieve invariant pattern recognition feature spaces[20]. Thus, the proposed MLP using second-order features defined by equation (7) is also a second-order neural network. Table 2 shows the numbers of nodes of the proposed second-order neural network.

<Table 2> The number of nodes of the proposed second-order neural network

input	hidden	output
8192	10	4
size of the cumulated vector	Randomly selected	number of output categories

In the recognition stage, the second-order features of the four edge maps are extracted respectively. A $2 \times 64 \times 64$ feature vector per each edge map is generated. Then, four feature vectors are cumulated into the one vector by adding the same dimensions. The cumulated vector is the actual inputs of the neural network. The processing examples are shown in Figure 5.

The recognition results in Table 3 show that the proposed second-order neural network recognizes all 80 images correctly. The results showed that the second-order features are a good invariant representation of the wavelet transform. Thus, the proposed method makes it possible to use the wavelet transforms to the pattern recognition directly.

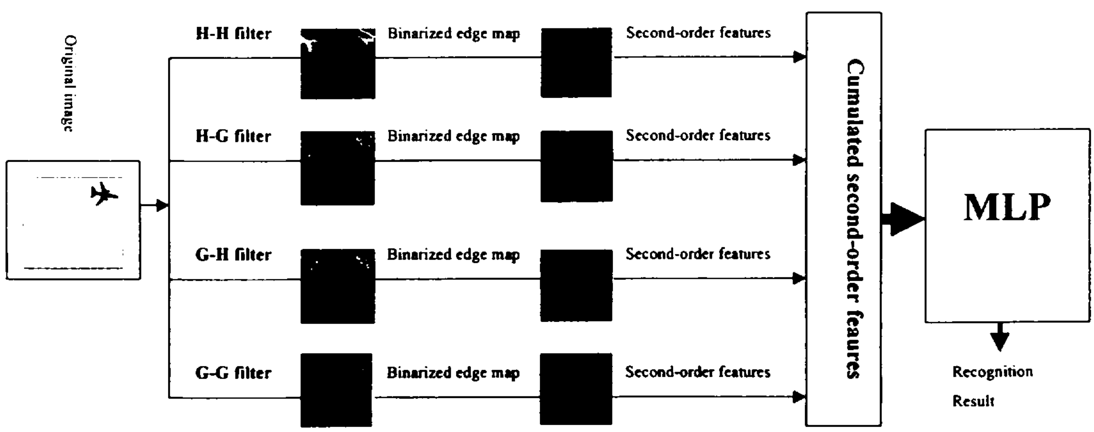

Figure 5. An Example of processing steps

<Table 3> The results of the experiments

	airplane 1	airplane 2	airplane 3	airplane 4
number of correct recognitions	20/20	20/20	20/20	20/20
recognition rates	100%	100%	100%	100%

5. Conclusion

The discrete 2D-WT can be widely used for image processing application due to its capability. However, the discrete 2D-WT transform can not be used for pattern recognition directly because it is not invariant to translations of patterns. This problem is the major obstacle to use the wavelet transform to pattern recognitions.

In this paper, a new method was proposed to solve the translation problem of the discrete 2D-WT. The proposed system uses a second-order neural network without additional preprocessing. Our system has simpler architecture and can be implemented easily than the existing methods and can be applied to pattern recognitions regardless of sizes and types of images. The experimental results shows the method based on the second-order neural network is a good solution for the translation problem of the discrete 2D-WT.

<References>

[1] M. Vetterli, "A theory of multirate filter banks," IEEE Trans. on ASSP, vol. 35, no. 3, Mar., 1987.

[2] S. G. Mallat, "A theory for multiresolution signal decomposition : The wavelet representation," IEEE Trans. on PAMI, vol. 41, no. 7, July, 1989.

[3] S. G. Mallat, S. Zhong, "Characterisation of signals from multiscale edges,", IEEE Trans. on PAMI, vol, 14, no. 7, pp. 710 - 732, 1992.

[4] I. Daubechies, "Orthnormal bases of compactly supported wavelets," Commun. Pure Appl. Math., vol. 41, pp. 909-996, Nov. , 1988.

[5] S. G. Mallat, "Multifrequency Channel decompositions of images and wavelet models ,"

IEEE Trans. on ASSP, vol. 37, no. 12, Dec., 1989.

[6] Q. M. Tieng, W.W. Boles, "Wavelet based affine invariant representation: A tool for recognizing planar objects in 3D space," IEEE Trans. on PAMI, vol, 19, no. 8, pp. 846 - 857, 1997.

[7] Q. M. Tieng, W.W. Boles, "Recognition of 2D objects contours using the wavelet transform zero-crossing representation," IEEE Trans. on PAMI, vol, 19, no. 8, pp. 910 - 916, 1997.

[8] I. Cohen, S. Raz, D. Malah, "Shift invariant wavelet packet bases," Proceeding of ICASSP, pp. 1081 - 1084, 1995.

[9] J. Liang and T. W. Parks, "Image Coding Using Translation Invariant Wavelet Transforms with Symmetric Extensions," IEEE Trans. on Image processing, vol. 7, no. 5, pp. 762 - 768, 1998.

[10] A. Grossmann, J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," SIAM J. Math., vol. 15, pp. 723 - 736, 1984.

[11] R. K. Young, Wavelet theory and its application, Kluwer Academic Publishers, 1993.

[12] M. J. Shensa, " The discrete wavelet transform : Wedding the a' Trous and Mallat algorithm," IEEE on Signal Processing, vol. 40, pp. 2464 - 2482. Oct., 1992. R. K. Young, Wavelet Theory

[13] M. Aotonini, M. Barland, P. Mathien, I. Daubechies, " Image coding using wavelet transform," IEEE trans. Image Processing, vol. 1, no. 2, Apr. 1992.

[14] R. H. Bamberger, S. L. Eddins, V. Nui, "Generalized symmetric extensions for size limited multirate filter banks," IEEE trans. Image Processing, vol. 3, no. 1, pp. 82 - 87, Jan. 1994.

[15] G. Pirani, V. Zingarelli, "Analytical formula for design of quadrature mirror filters,", IEEE trans. acoustic, speech, signal processing, vol. 32, pp. 645-648, June, 1984

[16] J. W. Woods, S. D. O'Neil, "Sunband coding of images," IEEE trans. acoustic, speech, signal processing, vol. 34, Oct., 1984.

[17] C. L. Giles, T. Maxwell, "Learning,invariance, and generalization in high-order neural

networks," Applied Optics, vol. 26, pp. 4972-4978, 1987.

[18] H. Y. Kwon, B. C. Kim, H. Y. Hwang, "Rotation and scale invariant pattern recognition using complex log mapping and augmented second-order neural network," IEE Electronics Letters, vol 29, no. 7, pp. 620 - 621, 1992.

[19] B. K. Lee, H.H Lee, D. K. Kim, Y. K. Cho, H. Y. Hwang, " Reduction of input nodes for shift invariant second-order neural networks using PCA," ICANN'94, vol. 2, pp. 1029 - 1032, 1994.

[20] W. A. C. Schmidt, J. P. Davis, "Pattern Recognition Properties of various feature spaces for higher order Neural Networks," IEEE Transaction on PAMI, vol. 15, no. pp. 795 - 801, 1993.

[21] Anil K. Jain, Fundamentals of digital image processing, Englewood Cliffs, N.J.: Prentice Hall, 1989.