

Texture Segmentation Based on Wavelet and Kohonen Network for Remotely Sensed Images

Z. Chen*, T.J Feng*, Z. Houkes**

*Department of Electrical Engineering,
Ocean University of Qingdao,
Qingdao 266003, P. R. China

E-mail: chenzhe@mail.ouqd.edu.cn tjfeng@mail.ouqd.edu.cn

**Department of Electrical Engineering,
University of Twente,

P.O.Box 217, 7500 AE Enschede, The Netherlands

E-mail: Z.Houkes@el.utwente.nl

ABSTRACT

In this paper, an approach based on wavelet decomposition and Kohonen's self-organizing map is developed for image segmentation. After making the 2-D wavelet transform of image, some features are extracted for texture segmentation, and the Kohonen neural network is used to accomplish feature clustering. The experimental results demonstrated the satisfactory effect of proposed approach both for simulated textured image and multi-spectral remotely sensed images.

Keywords: Texture segmentation, wavelet decomposition, feature extraction, Kohonen network, clustering

1. INTRODUCTION

Texture is interpreted in the literature as a set of statistical measures of spatial distribution of gray levels in an image, the texture feature is a value which quantifies some characteristic of the gray-level variation with the object. The texture of image contains not only the gray statistical information, but also spatial distribution and structural information, which are all of importance in image processing study. There is a rich history of research for texture segmentation of image, among the traditional methods have auto-correlation function, Fourier transform, gray statistical method and so on. In recent years, There have been many new approaches in this field, such as Markov random field [1], Gibbs model [2], fractal [3], wavelet [4-6].

Image segmentation is one of important applications in remotely sensed images, which is the foundation of the classification, recognition processing. The quality of sea surface temperature (SST) measurement derived from spaceborne instruments depends largely on the accuracy of the cloud removal techniques for eliminating cloud-

contaminated pixels. Therefore, the accurate detection of clouds in satellite imagery, also referred to as cloud screening or cloud segmentation, is fundamental to the analysis of remotely sensed data. Clouds are characterized by high albedoes and low temperature, traditional cloud detection methods are often based on its multi-spectral characteristic. Unfortunately, the data of multi-spectral remotely sensed images are rather complicated and these methods often fail or have poor precision in the case of complex cloud variation due to their simplicity. Since mere spectral information is not enough to segment the image, it is essential that the cloud segmentation algorithm can adapt dynamically to image specific conditions.

Some scholars introduced the techniques of pattern recognition to cloud detection to finish cloud detection, segmentation and classification using texture information in different channels. Gray level co-occurrence matrix (GLCM) method based on second-order statistics is a point of case. Spatial-frequency analysis approach is a useful way to describe texture. Among which the wavelet transform seems to be the most valuable. In this paper, according to the characteristics of remotely sensed image, we developed a kind of approach based on feature extraction with wavelet decomposition and feature clustering with Kohonen network to perform image segmentation. Kohonen self-organized map (SOM) of unsupervised clustering method is particularly suitable for the complicated remotely sensed image processing due to its no need of *a priori* knowledge.

This paper is organized as follows: Section 2 provides a brief background about the wavelet decomposition and multi-resolution analysis. Section 3 describes the feature extraction we employed in the experiment based on wavelet decomposition. In section 4, the Kohonen SOM network for feature clustering is introduced. The experimental results and conclusion are given in section 5 and section 6, respectively.

2. WAVELET DECOMPOSITION AND MULTIREOLUTION ANALYSIS

Wavelet analysis, derived from Fourier analysis, has made great progress both in mathematics and engineering since it is founded in 1980s. Owing to the excellent properties of temporal-frequency localization and adaptive multi-resolution, wavelet has been widely used as a powerful tool in signal processing showing a promising prospect such as image processing, computer vision, data compression and so on [7]. In the filter bank interpretation, wavelet can be viewed as the impulse response of a band-pass filter, wavelet transform, can be explained as that the original signal is filtered by a series of band-pass filters in different scales. Therefore, we can decompose the signal to a series of frequency bands for analyzing and processing. Wavelet theory provides a united framework for the space scale analysis of image. Two-dimension wavelet transform and multi-resolution analysis (MRA) are powerful tools in image processing. In the digital image processing, we usually introduce dyadic wavelet by discrete binarization to make multi-channel multi-resolution analysis to image.

Denote $\{V_j\}$ as a closed subspace of $L^2(R)$ (vector space of square-integrable, one-dimensional function), thus $L^2(R^2)$ can be decomposed as a tensor product of two identical subspace of $L^2(R)$, i.e., $V_j = V_j \otimes V_j$, which can form a MRA in $L^2(R^2)$. Given 1-D scaling function ϕ and wavelet function ψ , the 2-D scaling function and 2-D separable wavelets can be defined as follows:

$$\begin{cases} \Phi(x, y) = \phi(x)\phi(y) \\ \Psi^1(x, y) = \phi(x)\psi(y) \\ \Psi^2(x, y) = \psi(x)\phi(y) \\ \Psi^3(x, y) = \psi(x)\psi(y) \end{cases} \quad (1)$$

Thus, an image $f(x, y)$ can be represented as a hierarchy as follows:

$$\begin{aligned} & \{A_{2^{-j}}^d f, D_{2^j}^1 f, D_{2^j}^2 f, D_{2^j}^3 f, -J \leq j \leq 1\} \\ & A_{2^j}^d f = [f(x, y) * \Phi_{2^j}(-x, -y)] \cdot (2^{-j} m, 2^{-j} n) \\ & D_{2^j}^{1-3} f = [f(x, y) * \psi^{1-3}_{2^j}(-x, -y)] \cdot (2^{-j} m, 2^{-j} n) \end{aligned} \quad (2)$$

This set of images is called an orthogonal wavelet representation in two dimensions. The wavelet decomposition can be interpreted as a signal or image decomposition in a set of independent, spatially oriented frequency channels. The schematic of 2-level wavelet decomposition for an original image is shown in Fig.1. We can see that the original image is decomposed to a pyramid structure composed of smoothing image in the low frequency at the resolution 2^J and the high frequency detail information of J levels images. $D_{2^j}^{1-3} f$ represent high frequency details in the horizontal, vertical and diagonal

direction at the resolution 2^j , which are orthogonal each other.

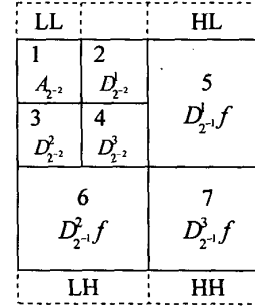


Fig.1 Wavelet decomposition of image

Mallat [8,9] had studied the relation between the MRA and wavelet transform, he used a set of quadratic mirror filters (QMF) to perform fast wavelet decomposition and reconstruction of signal by making convolution product. The importance of Mallat fast algorithm in the wavelet analysis is same as that of FFT in Fourier analysis.

Reference [10] used non-orthogonal Morlet wavelet and Mexico hat wavelet to perform the classification of texture. However, the advantage of choosing of orthogonal and compactly supported wavelet is that since the representation features at each scale are obtained by decomposing an image onto an orthonormal basis, correlation between scales is avoided and hence texture segmentation is more effective. We use the simplest Haar wavelet, Daubechies wavelet (D2,D4) and biorthogonal wavelet [11] in our experiment to test their effects, the experimental results demonstrate the biorthogonal filters are more efficient. In such way, wavelet decomposition provide us a different perspective on various scales and MRA representation provide a hierarchical framework for interpreting the image information.

3. FEATURE EXTRACTION

Image is usually comprised of smoothing region and textured region. The experimental studies have shown that the textured region has more high-frequency component and the smoothing region has more low-frequency components, both of which have conspicuous energy disparity, and different types of textures also show energy difference among them. The energy contrast of wavelet-decomposed sub-images of channels is illustrated in Fig.2. Additional experiment about Fourier energy spectrum analysis to the remotely sensed image also demonstrate that most of energy concentrates on the low-frequency region, while the smoothed images in the high-frequency embody the textural information of the image, where horizontal, vertical and diagonal sub-images comprise of the directional information of texture. Mallat [9] suggested that

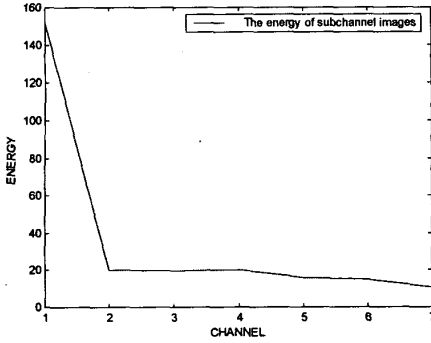


Fig.2 The energy contrast of wavelet-decomposed channels

by using wavelet decomposition, statistics based on first-order distribution of gray levels might be sufficient for the perception of textural difference. In this paper, we define five features to describe the texture of image as below.

1. *Energy ratio feature between low-frequency against high-frequency*

$$F1 = \frac{e_1}{e_2 + e_3 + e_4 + e_5 + e_6 + e_7} \quad (3)$$

where e_n is energy of n^{th} channel of image defined as:

$$e_n = \frac{1}{M^2} \sum \sum |x(i, j)| \quad (4)$$

2. *Energy feature of 5~6 channels*

$$F2 = \frac{1}{255^2} \sum_{i=0}^{255} i^2 p(i) \quad (5)$$

3. *Entropy feature of 5~6 channels*

$$F3 = \frac{-1}{\log(255)} \sum_{i=0}^{255} p(i) \log(p(i)) \quad (6)$$

where $p(i)$ is the probability of number of gray i against the number of all pixels of image.

4. *Variance feature of 5~6 channels*

$$F4 = \frac{1}{M} \sum (x(i, j) - \bar{m})^2 \quad (7)$$

where \bar{m} is the mean of all pixels in the selected image.

5. *Mean feature of 1 channel*

Eventually, we will get eight feature vectors as $F = \{F1, F2_{5-6}, F3_{5-6}, F4_{5-6}, F5\}$. Among which, the feature 1 represents the energy ratio of smoothing sub-image against

sub-textured images. Feature 2~4 are only processed in the high frequency channels, representing information of energy, entropy and gray statistics in those channel images. Feature 5 represents the mean feature of smoothing image in low frequency. In our experiment, an overlap 3×3 neighborhood window is scanned over each pixel of the whole image to compute the features 2~5 of the local neighborhood and replace the central pixel gray value with the computed feature values. The procedure is repeated and seven feature images are produced. For the feature 1, an overlap 4×4 window is scanned over the whole image in which the 2-level wavelet decomposition is made and the energy ratio value is computed to replace the origin at the leftmost and upmost position in the neighborhood. The experiments below will demonstrate these feature images can well distinguish different textured regions.

4. KOHONEN SOM NETWORK FOR FEATURE CLUSTERING

There are many feature-clustering methods in pattern recognition technique, such as K-mean clustering, Bayesian estimation, Maximum likelihood. But the traditional clustering methods are also in demand of *a priori* knowledge and probability distribution, and therefore usually hard to produce the complex hyper-plane due to computational complexity. Neural network is a highly nonlinear dynamic system with strong fault-tolerance, good robustness and adaptability, which is thus widely applied in pattern recognition. Kohonen network, proposed by T.Kohonen as a self-organized feature map, is a kind of efficient unsupervised clustering method, which can be used as classifier without the need of empirical knowledge about pattern distribution. It employs the self organizing learning algorithm without specifying the desired output, the weights will be organized such that topologically close nodes are sensitive to inputs that are physically similar. Kohonen's algorithm creates a vector quantizer by adjusting weights to make every output node correspond to some kind of the input vector. After enough input vectors have been presented, weights will specify cluster or vector centers that sample the input space such that the point density function of vector centers tends to approximate the probability density function of the input vectors. Kohonen network is a competitive two-layer network, whose input layer is composed of some input neurons and its competitive layer (output layer) is a two-dimensional array of output nodes. The learning algorithm for SOM networks [12] is described as below:

1. Initialize weights from N inputs to the M output nodes to the $[0,1]$ interval randomly, set the initial radius of neighborhood.
2. Present the input pattern to network.
3. Compute Distance d_j between the input and each output node j using

$$d_j = \sum_{i=0}^{N-1} (x_i(t) - W_{ij}(t))^2 \quad (j=1,2, \dots, M) \quad (8)$$

4. Select node j^* as output node with minimum d_j .
5. Update the weights of node j^* and all nodes among the neighborhood $N_g(t)$.

$$W_{ij}(t+1) = W_{ij}(t) + \eta(t)(x_i(t) - W_{ij}(t))$$

$$0 < \eta(t) < 1, \quad j \in N_g(t), \quad (i=1,2, \dots, N) \quad (9)$$

6. Sent the next learning pattern to the input layer, repeat step 3 until all learning patterns are sent once.
7. Update the learning rate $\eta(t)$ and the neighborhood radius $N_g(t)$.

$$\eta(t) = \eta_0(1 - t/T) \quad (10)$$

8. Make $t=t+1$, return step 2 until $t=T$.

We combined the feature vectors of all pixels of image to make a set, whose elements are actually the pixels of the feature image. In this way, eight sets are produced, which contain some feature vectors or some types of textures of image to be segmented. A random subset whose size is about amount to the one tenth of total number is picked out in order to improve the learning speed. The elements of subset are chosen as the input vectors of SOM network for learning. At the beginning of learning, every neuron of output layer is given a counter whose initial value is set as zero. When the input feature vectors respond to some neuron, the counter of which add one. At the end of learning, the same type of input vectors should concentrate the same region in the output layer, while different types of input vectors should respond different regions in the output layer of network. With excellent characteristic of topological clustering among the input vectors, Kohonen network's output eventually classify the output layer as several fields, each of which represents a kind of input vector or a kind of texture. The number of field represents the types of texture, the neuron with largest counter represents the center of the clustering.

5. EXPERIMENTAL RESULTS

The block diagram of our experiment procedure is shown in Fig.3. Histogram equalization is to enhance image in order to guarantee higher contrast degree. In the need of our experiment for keeping the channel images the same size, we make convolution product with discrete-time filters without sub-sampling by 2 when performing discrete wavelet transform so that the seven channel sub-images

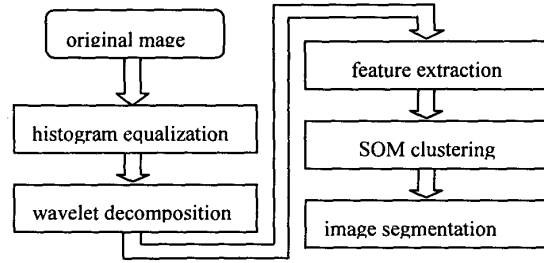


Figure 3: Block diagram of the proposed approach

with same size as the original image are produced. After feature images are computed after 2-level wavelet decomposition and feature extraction, the texture features of images or the elements of the feature images are sent sequentially to Kohonen network whose initial radius is set as two. For each feature vector of input, there must exist a node response in a region of output layer, which can be used to classify the type of feature vectors. After confirming the type of all pixels, we can represent different types of textures or regions with different gray values to finish image segmentation.

Experiment 1: Figure 4.a is a simulated textured image, which contains four kinds of texture information. Feature set 1 is only employed to make texture segmentation for this image. The input layer has 25 nodes that represent the pixels in the local neighborhood of this feature image. The corresponding feature image and the final segmented image are shown in Fig. 4b and Fig.4c respectively.

Experiment 2: In addition, we choose a real multi-spectral remotely image. The satellite image data we used in the experiment is an ATSR (Along Track Scanning Radiometer) scene over the Southern Sea of China on February 7th, 1993 derived from the ERS-1 satellite. The scene of ATSR channel near-infrared $11 \mu m$ have a gray level range of 0~255, composing of 512×512 pixels, shown in Fig.5a. The 2-level wavelet decomposition of this image is given in Fig.5b. Fig.5c~5f show the feature images computed from feature 1~4. (feature 2~4 only give the feature image of channel 5). It may be seen that the feature image can distinguish the region of cloud and sea very well. The network architecture whose input nodes are eight and output layer are two-dimensional array of 10×10 nodes are employed in this experiment. The last result of texture segmentation is given in Fig.5g. The method of histogram-threshold segmentation to the same image is also given in Fig.6 for comparison. In contrast, the proposed approach is very suitable for the segmentation of cloud-sea field, especially capable of detecting the thin cloud region and is obviously superior to the histogram-threshold image segmentation method. Fig. 7 is the standard cloud-sea binary image for reference. As shown from the Fig.5g and Fig.7, we can see that the proposed method is very effective for cloud-sea segmentation.

6. CONCLUSION

An approach based on wavelet decomposition, texture feature extraction and SOM feature clustering is proposed in this paper to investigate the effect of image segmentation both for simulated textured image and multi-spectral remotely sensed image. The experiments demonstrate that the given approach can segment the image with different textures very well, especially suitable for the data-complicated and non-anterior knowledge remotely sensed image. In the application in cloud detection for a satellite image, the proposed method can classify the cloud and sea field very well and can detect the region of thin cloud. An open problem still under further investigation is the effect for image texture segmentation when selecting different wavelet basis and considering 2-D *non-separable wavelet*.

7. ACKNOWLEDGEMENTS

We would thank Prof. He Ming-Xia of ORSI of Ocean University of Qingdao for providing the remotely sensed image in the experiment. We also appreciate the helpful discussion with Professor Qin Qian-qing of Department of Mathematics of Ocean University of Qingdao. This work is supported by NSF of China (No.69675005)

8. REFERENCES

- [1] B. S. Manjunath and R. Chellappa, "Unsupervised Textures Segmentation Using Markov Random Field Model", *IEEE Trans. Pattern Analysis and Machine Intelligence*. Vol. 13, No. 5, 1991, pp. 478-482.
- [2] H. Derin and H. Elliot, "Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Field", *IEEE Trans. Pattern Analysis and Machine Intelligence*. Vol. 9, No. 1, 1987, pp. 39-55.
- [3] F. Dubuisson, "Efficacy of Fractal Features in Segmentating Images of Natural Textures", *Pattern Recognition Letters*. Vol. 15, No. 4, 1994, pp. 419-431.
- [4] P. Robert and C. Nishan, "A Robust Automatic Clustering Scheme for Images Segmentation Using Wavelets", *IEEE Trans. Image Processing*. Vol. 5, No. 4, 1996, pp. 662-665.
- [5] L. Andrew and F. Jian, "Texture Classification by Wavelet Packet Signatures", *IEEE Trans. Pattern Analysis and Machine Intelligence*. Vol. 15, No. 11, 1993, pp. 1186-1191.
- [6] L. C. Jia and K. Amlan, "Rotation and Gray Scale Transform Invariant Texture Identification Using Wavelet Decomposition and Hidden Markov Model", *IEEE Trans. Pattern Analysis and Machine Intelligence*. Vol. 16, No. 2, 1994, pp. 208-214.
- [7] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM Press, 1992.
- [8] S. Mallat, "Multifrequency Channel Decomposition of Images and Wavelet Models", *IEEE Trans. Acoustics, Speech, and Signal Processing*. Vol. 37, No. 12, 1989, pp. 2091-2110.
- [9] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", *IEEE Trans. Pattern Analysis and Machine Intelligence*. Vol. 11, No. 7, 1989, pp. 674-693.
- [10] P. H. Carter, "Texture Discriminationo Using Wavelets", *SPIE applications of digital image processing XIV*. Vol. 1567, 1991, pp.432-438.
- [11] J. M. Bruneau and P. Mathieu, "Image Restoration Using Biorthogonal Wavelet Transform", *Optical Engineering*. Vol. 33, No. 7, 1994, pp. 2378-2384.
- [12] R. P. Rippmann, "An Introduction to Computing with Neural Networks", *IEEE ASSP Magazine*. Vol. 4, No. 2, 1987, pp. 4-22.

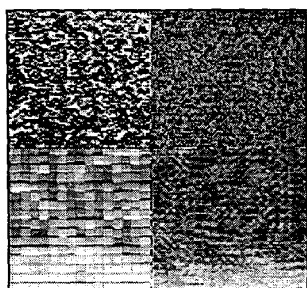


Fig 4.a Simulated textured image

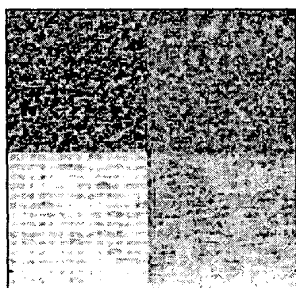


Fig 4.b Energy ratio feature image of 4.a

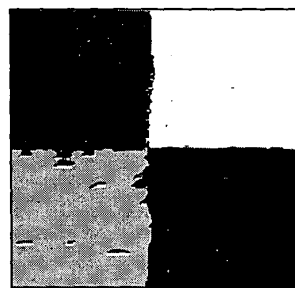


Fig 4.c Segmented result of 4.a



Fig 5.a Remotely sensed image

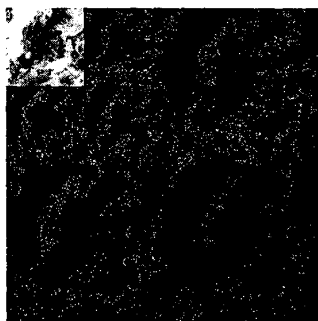


Fig 5.b 2-level wavelet decomposition of 5.a

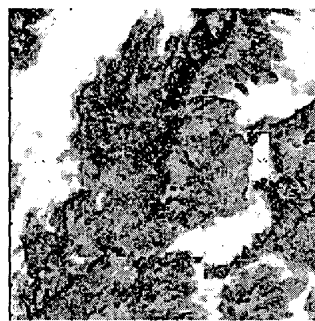


Fig 5.c Energy ratio feature image of 5.a

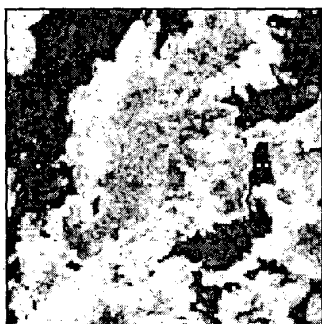


Fig 5.d Energy feature image

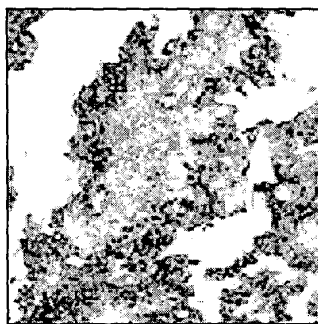


Fig 5.e Entrophy feature image

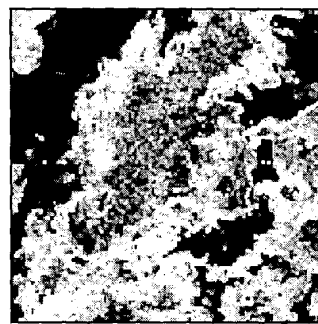


Fig 5.f Standard deviation feature image

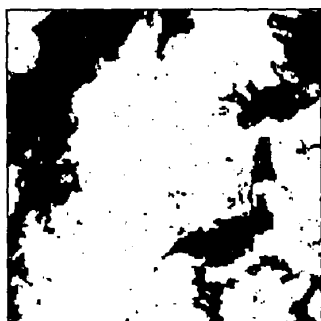


Fig 5.g Segmented result of 5.a

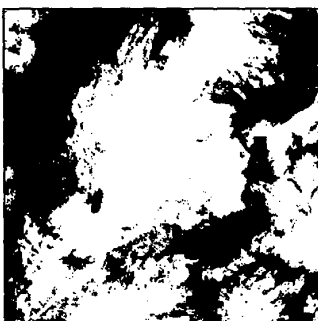


Fig.6 Histogram-thresholding segmented result

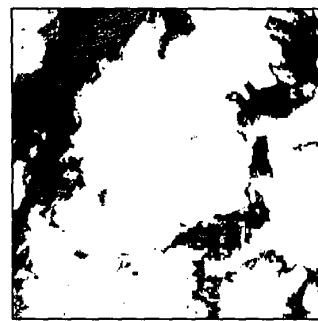


Fig. 7 Standard segmented image