IEICE TRANS. COMMUN., VOL.E78-B, NO.12 DECEMBER 1995

LETTER *Special Issue on Satellite Remote Sensing*

JERS-1 SAR Image Analysis by Wavelet Transform

Yoshio YAMAGUCHI†, *Member*, **Takeshi NAGAI**†, *Student Member*, *and* **Hiroyoshi YAMADA**†, *Member*

SUMMARY The wavelet transform provides information both in the spatial domain and in the frequency domain because of its inherent nature of space-frequency analysis. This paper presents a classification result of synthetic aperture radar image obtained by *JERS-1* based on the discrete wavelet transform. This paper points out that the wavelet analysis has yielded a fine result in texture classification compared to a conventional method with less computation time.

key words: *radio applications, SAR image, classification, remote sensing*

1. Introduction

The Japanese Earth Resources Satellite-1 *(JERS-1)* has provided Synthetic Aperture Radar (SAR) Image operative at the frequency of 1.275 GHz since its launch in Feb. 11, 1992 [1]. The resolution in the terrain is 12.5 m by 18 m. This microwave remote sensing has advantages because it can provide terrain images regardless of weather and regardless of operation time. The image is a result of microwave backscattering and is expected to provide important but different information on terrain which cannot be obtained by any optical sensors.

The authors have been engaged in the SAR data analysis, especially classification of terrain targets based on the wavelet transform. The wavelet transform has been recently attracting attention in diverse areas such as signal processing [2], numerical analysis and reconstruction problems [3], especially for non-stationary signal analysis applications [4]. Since SAR images consist of thousands of pixels pertaining to backscattering coefficient from natural terrain, it displays non-stationary signal characteristics. Using this specific characteristics of SAR images and the characteristics of wavelet transform, *JERS-1* radar scenes are examined to verify the performance of wavelet transform in target classification. In the following, a brief introduction of wavelet transform and multiple resolution representation scheme are given in Sects.2 and 3. Sections 4 and 5 discuss the texture classification procedure and the results comparing with other method indicating the superiority of wavelet transform.

Manuscript received May 30, 1995.
Manuscript reviced August 11, 1995.

† The authors are with the Faculty of Engineering, Niigata University, Niigata-shi, 950-21 Japan.

2. Wavelet Transform

The wavelet transform is given [2]-[5] by

$$Wf(b,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}\psi^{*}\left(\frac{t-b}{a}\right)f(t)\,dt \qquad (1)$$

where $f(t)$ is the input signal, $\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$ is called a wavelet which is a localized function of space and frequency and is derived from the analyzing wavelet function $\psi(t)$. The symbol * denotes complex conjugation. The analyzing wavelet function can be defined by an orthogonal scaling function $\phi(x)$ which is a cubic spline function [5]. Equation (1) is of convolutional integral form given by the input signal and the wavelet functions. The parameter a is a scaling factor, while the parameter b is a shifting factor in the space domain. It is possible to choose any desired parameters, which leads to provide any resolution (under the uncertainty restriction) both in the space and in the frequency domain. Therefore, the wavelet transform is suitable for non-stationary signal analysis. If the analyzing wavelet functions are orthogonal, the transform is said to be orthogonal. The orthogonal wavelet transforms has specific features such that

1) Signal is well reconstructed by the inverse transform,
2) The orthogonality enables easy treatment of signal components.

However, it is difficult to obtain explicit orthogonal analyzing wavelet functions from the definition of the wavelet transform. Mallat [5] has proposed a method to derive analyzing wavelet functions based on the multi-resolution representation. The wavelet transform can be roughly classified into two types, i.e., the continuous wavelet transforms and the discrete wavelet transforms. In this paper, we adopt discrete transform because the image pixels are discrete signals. In the discrete wavelet transform, we have chosen the scaling factor a and the shifting factor b as $(b,a)=(2^{-n}\cdot m, 2^{-n})$ with m, n being integer [3].

The discrete wavelet transform corresponds to multi-resolution approximation expressions [3]. This method utilizes the fact that the complementary space belonging to the 2^{-n}-th resolution representation but not belonging to the 2^{-n-1}-th resolution representation becomes the component of the orthogonal wavelet transform of the n-th order. The transform scheme is depicted in Fig.1. In the actual

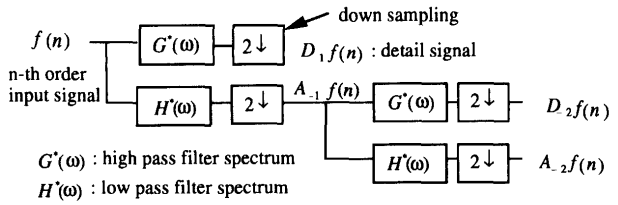


Fig.1 Transformation scheme for wavelet transform.

processing, a recursive signal execution is employed to obtain the discrete wavelet transform as follows:

1) Divide the bandwidth of the input signal into two parts using high-pass and low-pass filters whose frequency spectrum are given by $G(\omega)$ and $H(\omega)$ which are Fourier transform of $g(n)$ and $h(n)$ (see Eqs.(4), (5)), respectively,
2) Sample these divided parts at a half rate (1/2-down sampling) of the previous data,
3) Use the lower bandwidth signal for the next 1/2-down sampling.

By repeating this procedure, it is possible to obtain wavelet transform of any order. This 1/2 sampling procedure does not change the scaling parameter a=1/2 throughout successive wavelet transforms so that it benefits for a simple computer implementation. Using this procedure, the input signal is represented by multi-resolution approximation expression.

3. Multi-Resolution Representation

This multi-resolution representation technique stems from image processing area and has the following advantages: 1) similarity of data structure with respect to the resolution, i.e., the higher resolution image includes the lower resolution images, and 2) decomposition of the image is available at any level where an appropriate signal processing can be carried out. These advantages benefit especially to remotely sensed data analysis for which one may wish to see macroscopic characteristics of the scene at first stage using a low resolution level and then analyze an area of interest in detail using an appropriate higher resolution level. For computer implementation, it is desirable to use 2 to the n-th power in the resolution representation.

Now, let the original image be the highest resolution image and the original pixel values represent the input data function f. Define an operator A_j which approximates f at the resolution stage 2^{-j}, j being an integer. Then the data sets $A_{-1}f(n)$, $A_{-2}f(n)$, $\cdots, A_{-j}f(n)$ represent discrete values at the stages 2^{-1}, 2^{-2}, $\cdots, 2^{-j}$, respectively. The $A_{-j}f(n)$ is related to $A_{-j+1}f(n)$ by the detail signal $D_{-j}f(n)$ such that

$$A_{-j}f(n) = \sum_k h(k-2n)\, A_{-j+1}f(k), \tag{2}$$

$$D_{-j}f(n) = \sum_k g(k-2n)\, A_{-j+1}f(k), \tag{3}$$

where $h(k-2n)$ and $g(k-2n)$ are identical with the low-pass and the high-pass filter, respectively. The coefficients

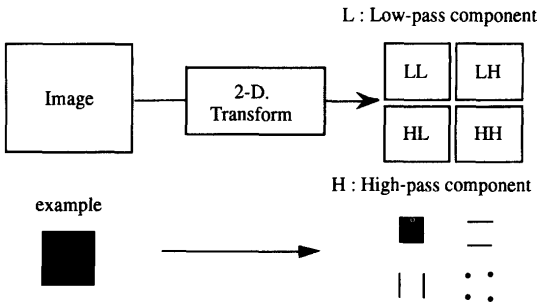


Fig.2 Two-dimensional transform.

of these filters are given in terms of the scaling functions $\phi(x)$ as

$$h(n) = \int \phi(x)\, \phi(2x-n)\, dx, \tag{4}$$

$$g(n) = (-1)^{1-n} h(1-n). \tag{5}$$

This formulation can be easily expanded to the 2-dimensional (x and y-directions) case using

$$h(n_x, n_y) = h(n_x)\, h(n_y). \tag{6}$$

The detail function in 2-D case may be found from

$$g^1(n_x, n_y) = g(n_x)\, h(n_y), \tag{7}$$

$$g^2(n_x, n_y) = h(n_x)\, g(n_y), \tag{8}$$

$$g^3(n_x, n_y) = g(n_x)\, g(n_y). \tag{9}$$

Therefore, the wavelet transform yields 4 sub-images of the original one as shown in Fig.2. The number of pixel in each sub-image reduces to 1/4 of the original number.

4. Texture Classification

An image consists of thousands of pixels. In general, pixel value (backscattering coefficient) itself does not provide any specific information on a target because it is a function of many parameters. Texture is a specific uniform pattern representing a target consisting of some pixels, in other words, it is a spatial variation of backscattering coefficient from one region to another in an image. Texture can be roughly classified into statistical and structural textures. Texture analysis is frequently employed for target classification, which accounts for correlation between the adjacent pixels or neighboring pixels. Since the terrain texture obtained by remotely sensed data has statistical properties, we used a statistical method to classify the *JERS-1* data.

The classification algorithm here is based on the maximum likelihood method [7] using training data sets (supervised classification). The maximum likelihood method determines the degree to which pixel is the suitable to a group (cluster) among a certain categorized set. If the correlation exceeds over a certain threshold level, the pixel is categorized into the group. In the signal processing, pixel is represented by a characteristic vector as,

$$X = \left[x_1, x_2, x_3, \cdots, x_m \right]^t \tag{10}$$

where m denotes the dimension of the vector and superscript t represents transpose. The component of the vector, x_k, is the pixel value of a sub-image in the wavelet transform. Thus, the dimension of X becomes 4 provided that the wavelet transform is applied to the original image once, and becomes 7 for two-times wavelet transforms. Then a cluster in an image is chosen so that the average characteristic vector and the covariance matrix pertaining to the cluster are determined. A similar procedure to obtain different clusters together with the mean characteristic vectors and the covariance matrices which may represent specific targets. The discriminant criterion for each cluster category is defined as

$$G_i = -\log \left| V_i \right| - (X - \overline{X_i})^t V_i^{-1} (X - \overline{X_i}) \tag{11}$$

where the subscript i indicates the number of category to be classified, V_i is the covariance matrix and $\overline{X_i}$ is the mean characteristic vector of the category i, respectively. Any pixel which maximizes the discriminant is classified into the category i. The dimension of the characteristic vector X increases with the number of wavelet transform execution. Since the component of X is the component of wavelet transform of a specific target, it represents the characteristics in sub-frequency bands, which also represents the characteristics of the pixel. Since the component of X contains information on neighboring pixels, a repeated wavelet transform to a specific image may serve to classification of targets efficiently. This point is an advantage compared to the conventional method which requires extensive computation time on correlation between adjacent pixels.

5. The Classification Results

The classification using the wavelet transform is applied to the *JERS-1* SAR data (data no.C0490A03). The data was acquired on Aug.20, 1993. The SAR data contains the following specific regions : the Sea of Japan, Niigata city and surrounding residential area, Toyano lake, and rice fields.

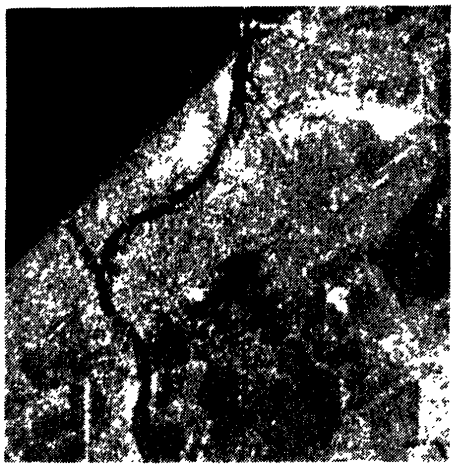
The categories classified are densely built area (downtown), surrounding residential area, rice fields, natural terrain (such as pine tree forests), and water region (sea/river/lake). Figure 3 shows the original *JERS-1* image, and Fig.4 displays the results of the wavelet transform (low-resolution images). Five quasi-gray scale is used to cover these regions with black indicating small backscattering coefficient in sea/river. White color corresponds to large backscattering in urban area. For an efficient classification, we have chosen an average characteristic vector in each region so that it may represent specific features of the region referring to a detailed geological city map. After determining the average vectors, we have carried out classification procedure for the image. The classified results are shown in Tables 1-3. These tables show the percentage of SAR pixels in a test area classified to actual 5 categories. The vertical row corresponds to actual categories, whereas the horizontal column corresponds to categories classified.

It is seen that the image quality degrades as the wavelet transforms proceeds. On the other hand, the degree to the classification performance increases. This is due to the property of wavelet transform, i.e., the lower resolution images are the low-pass filtered version of the higher resolution images bearing low-frequency information on neighboring pixels. This characteristics contributed to an efficient classification performance. However, an excess of transformation would cause the lack of the high frequency information in the classification procedure. Therefore there would be an appropriate number of wavelet transform execution which might be dependent on the degree to which resolution is desirable.

It is also seen that the discrimination between natural terrain and rice field is difficult. This seems to be caused by a similarity in the backscattering mechanism. It is anticipated that rice crops are close to natural terrain including grasses and trees. If the averaged characteristic vectors are resemble each other, or if the covariance matrices are similar, it becomes difficult to discriminate these two targets.

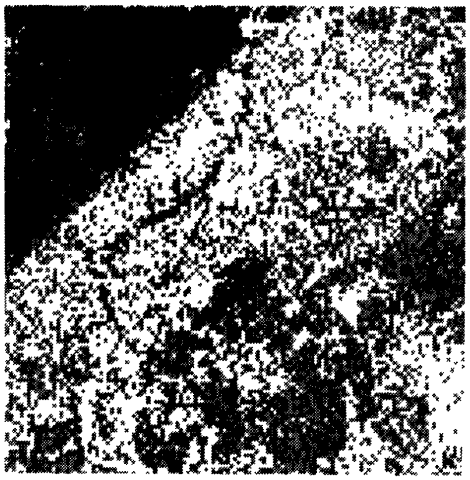
The backscattering coefficient is a function of many parameters such as incidence angle onto a target, target size and orientation, material constant of target, frequency, polarization, etc. Since the frequency and polarization are fixed in the entire *JERS-1* SAR, the incidence angle seems to be constant within a SAR image. The condition of same incidence angle, suggests a possibility to use the same training data in Niigata-city to other neighboring regions. We tried to classify other portions of the scene, resulting in a similar classification performance [8]. The reason why the classification yielded a satisfactory result is due to the

Fig.3 The original *JERS-1* image (1024 × 1024 pixels) of Niigata-city on Aug.20, 1993.



(a) 1/2 resolution (512 × 512 pixels)

(b) 1/4 resolution (256 × 256 pixels)



(c) 1/8 resolution (128 × 128 pixels)

Fig.4 Wavelet transformed image.

similarity in the training data and in the backscattering coefficients for targets. For a different scene of SAR data which includes different types of target, it is necessary to obtain the corresponding training data.

Table 1 Classification result at 1/2 resolution (%).

	densely built area	residence area	rice field	natural terrain	sea/river lake
densely built area	85.9	9.4	0.0	4.7	0.0
residence area	7.0	68.5	17.6	5.7	1.2
rice field	0.0	3.5	63.7	8.6	24.2
natural terrain	0.0	10.9	40.6	40.7	7.8
sea/river lake	1.6	1.5	31.2	1.6	64.1

The vertical row corresponds to actual categories, and the horizontal column corresponds to categories classified.

Table 2 Classification result at 1/4 resolution (%).

	densely built area	residence area	rice field	natural terrain	sea/river lake
densely built area	87.3	8.3	0.7	3.7	0.0
residence area	6.9	80.3	5.6	6.5	0.7
rice field	0.0	0.0	79.2	12.5	8.3
natural terrain	0.0	14.6	30.5	54.9	0.0
sea/river lake	4.2	0.0	18.4	1.4	76.0

Table 3 Classification result at 1/8 resolution (%).

	densely built area	residence area	rice field	natural terrain	sea/river lake
densely built area	89.4	4.4	0.0	6.2	0.0
residence area	6.2	84.6	0.0	9.2	0.0
rice field	0.0	8.1	81.5	9.6	0.8
natural terrain	0.0	20.8	9.4	69.8	0.0
sea/river lake	0.0	0.5	3.8	0.7	95.0

Comparison to other classification method

The density histogram analysis [7] is one of the most frequently used methods in texture classification. For the sake of comparison, we made up a characteristic vector $\boldsymbol{X}$ whose components are produced by a density histogram [7], [9] (instead of wavelet transform). The density histogram yields four statistical characteristic quantities, i.e., average, variance, skewness, and kurtosis within a selected area. Using this specific characteristic vector with four components, we carried out the classification procedure for the same scene. The final classification result is shown in Table 4.

It is seen that the classification result using this method

Table 4 Classification result by histogram method.

	densely built area	residence area	rice field	natural terrain	sea/river lake
densely built area	76.9	11.6	0.0	11.5	0.0
residence area	2.8	5.9	27.4	63.9	0.0
rice field	30.2	0.0	5.6	0.0	64.2
natural terrain	19.6	0.0	45.5	30.7	4.2
sea/river lake	7.2	2.8	0.0	0.2	89.8

is comparable in densely built area and water regions, however, the classification result in rice field, natural terrain and residential area is much poor compared with that of the wavelet transform. The performance may be similar in the specific areas, however, the wavelet transform can proceed to the next resolution stage where the performance of classification increases further (Table 3). Since there is no other resolution stage in this histogram method, it is impossible to compare the result in a strict sense. In additon, the computation time necessary to classify one scene is constant in this method. For the wavelet transform case, the computation time becomes shorter because the number of pixels reduces to 1/4 as the transform proceeds. The overall CPU time ratio was found to be 2:1 for the present analysis indicating the wavelet transform scheme efficient.

6. Conclusion

The discrete wavelet transform provides multi-resolution representation of a SAR image and the information on texture in detail. The information here is assigned as the component of the characteristic vector of a pixel. The classification using the wavelet transform together with the maximum likelihood method was successfully applied to a *JERS-1* SAR image, yielding a fine classification result compared to the histogram method. Once the training data (both averaged characteristic vector and the covariance matrix pertaining to a desired class) are obtained, it is possible to use them to the other scene provided the targets are close to the previous region. For further efficient classification, it will be needed to classify various radar scene together with an consideration of backscattering coefficients of targets and polarimetric information.

Acknowledgment

Authors are grateful to National Space Development Agency of Japan for providing *JERS-1* SAR data. MITI/NASDA retains ownership of the data. Authors also thank M. Koda and K. Misawa of Niigata University for their contributions to the data analysis.

References

[1] National Space Development Agency of Japan, Final Report of JERS-1/ERS-1 System Verification Program, vols.1 and 2, MITI, March 1995.
[2] M. Vetterli and O. Rioul, "Wavelet and signal processing," IEEE Signal Processing Magazine, vol.8, no.4, pp.14-38, 1991.
[3] "Mathematical Sciences," no.354, Science Publisher, Dec. 1992.
[4] H. Kikuchi, "Pictorial introduction to wavelets," IEICE Trans., vol.75, no.7, pp.781-785, 1992.
[5] S. Mallat, "The theory for multiresolution signal decomposition : the wavelet representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.11, no.7, pp.654-693, 1989.
[6] L. -J. Du, J. -S. Lee, K. Hoppel, and S. A. Mango, "Segmentation of SAR images using the wavelet transform," International Journal of Imaging Systems and Technology, vol.4, no.4, pp.319-326, 1992.
[7] Takagi, Shimoda ed., "Handbook of Image Analysis," pp.661-663 and 517-518, Tokyo-Daigaku-Shuppan-Kai, 1991.
[8] M. Koda, Y. Yamaguchi, H. Takahashi, and H. Yamada, "A trial to wavelet analysis on JERS-1 SAR data," IEICE Technical Report, SANE94-40, Aug. 1994.
[9] M. Ogami et al, "Image Processing Handbook," pp.296-297, Syoko-Dou, 1987.
[10] F. T. Ulaby and M. C. Dobson, "Handbook of Radar Scattering Statistics for Terrain," Artech House, 1989.