Korean Journal of Remote Sensing, Vol.23, No.5, 2007, pp.355~363

Application of the 3D Discrete Wavelet Transformation Scheme to Remotely Sensed Image Classification

Hee-Young Yoo*, Kiwon Lee†, and Byung-Doo Kwon***

*Department of Earth Science Education, Seoul National University, Shillim-dong, Kwanak-gu, 151-742, Seoul, Korea

**Department of Information System Engineering, Hansung University, Samsun-dong, Sungbuk-gu, 136-792, Seoul, Korea

Abstract : The 3D DWT (The Three Dimensional Discrete Wavelet Transform) scheme is potentially regarded as useful one on analyzing both spatial and spectral information. Nevertheless, few researchers have attempted to process or classified remotely sensed images using the 3D DWT. This study aims to apply the 3D DWT to the land cover classification of optical and SAR (Synthetic Aperture Radar) images. Then, their results are evaluated quantitatively and compared with the results of traditional classification technique. As the experimental results, the 3D DWT shows superior classification results to conventional techniques, especially dealing with the high-resolution imagery and SAR imagery. It is thought that the 3D DWT scheme can be extended to multi-temporal or multi-sensor image classification.

Key Words : 3D DWT, Land cover classification, High-resolution imagery, SAR imagery.

1. Introduction

The discrete wavelet transformation has been widely used in 2D image processing techniques: compressing large-sized image sets, reducing the noise of imagery, fusing and analyzing multi-scale images. The 2D DWT focuses on the spatial arrangement of pixel in one band rather than analyzing color components or multi-spectral features. The 3D DWT expands the 2D DWT to 3D for multi-channel data or volumetric data. Recently, the 3D DWT has been used in several engineering and scientific domains. For instances, it was applied to image compression and noise reduction in the medical images like MRI and CT (Ghugre *et al*, 2003; Chen and Ning, 2004). This scheme also was used in video image compression (Klock *et al*, 1997; Luo *et al*, 2004; Xiong *et al*, 2005). In other fields, seismic data processing using the 3D DWT was suggested for determining local orientations of three dimensional reflectors and extracting the local scale in migrated seismic data (Spendonck *et al*, 2001). Regarding remote sensing, space-borne images are composed of several bands in the frequency order so that it can be considered as a kind of 3D data. However, there are rare application of the 3D wavelet scheme to the analysis and the processing of space-borne imagery. Few studies have been just carried out

Received 13 September 2007; Accepted 17 October 2007.

† Corresponding Author: Kiwon Lee (kilee@hansung.ac.kr)

on hyper-spectral imagery compression (Cheung *et al*, 2006) till now. As the 3D wavelet scheme uses both spectral and spatial information simultaneously, the 3D DWT is available to classify road, bare soil, forest and building class, which are mixed with various materials. In this study, the 3D DWT in classifying satellite images is applied, and its results are compared with their results quantitatively with that of conventional classification.

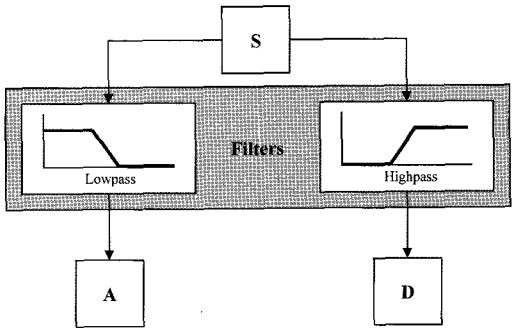


Fig. 1. Two-band filter banks for discrete wavelet transform.

2. Methodology

1) Wavelet Transform

The wavelet transform is a method of converting a function into another form which either makes certain features of the original signal more amenable to investigate or enables the original data set to be described more concisely (Addison, 2002). The basic idea of the wavelet transform is to represent signal as a superposition of wavelets. Wavelets are functions that satisfy certain requirements of being integrated to zero, especially the wave above and below the x-axis and insure quick and easy calculation of the direct and inverse wavelet transform (Antonini *et al.*, 1992). The signals are represented by translations and dilations of the wavelet, $\psi(t)$ as follows.

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

The scale parameter 'a' is defined as |1/frequency| and corresponds to the frequency information. The translation parameter 'b' relates to the location of the wavelet function as it is shifted through the signal. Thus, it corresponds to the time information in the wavelet transform. This is the generic case of the continuous wavelet transform (CWT). The CWT computation may consume huge amount of time and resources depending on types of used data sets and target resolution. The discrete wavelet transform (DWT), which is based on sub-band coding, is to yield a fast wavelet transformation computation. As well, it is rather easy to implement and reduce the computation time and resources required. Sub-bands of the digital signal are obtained by passing digital filters. The target signal (S) is passed through the filters with different cutoff frequencies at different scales (Fig. 1). After passing the filters, two decomposed signals, which are approximation (A) and detail (D) or L band and H band, are produced.

$$y_{high} = \sum_n x[n]g[2k-n]$$
$$y_{low} = \sum_n x[n]h[2k-n] \qquad (2)$$

y_{high} and y_{low} represent the output of the highpass(g) and lowpass(h) filters, respectively after subsampling by a factor 2 (Tzanetakis *et al.*, 2001).

Similarly, the 2D DWT generates results by two step processing: the initial filtering on the data in row direction and the next filtering with respect to the initial output in column direction by the same filters. This leads at each level to 4 different sub-bands HH, HL, LH and LL obtained by a combination of high and low filtering (Mallat, 1989).

2) 3D discrete wavelet transformation

The 2D DWT decomposes data in row and column direction. Meanwhile, the 3D DWT decomposes data not only in row and column direction but also in

depth direction. In terms of the wavelet space decomposition, the 3D DWT can be constructed by a tensor product by

$$\begin{aligned} V^3 &= (L^x \oplus H^x) \otimes (L^y \oplus H^y) \otimes (L^z \oplus H^z) \\ &= L^x \otimes L^y \otimes L^z \oplus L^x \otimes L^y \otimes H^z \\ &\quad \oplus L^x \otimes H^y \otimes L^z \oplus H^x \otimes L^y \otimes L^z \\ &\quad \oplus L^x \otimes H^y \otimes H^z \oplus H^x \otimes L^y \otimes H^z \\ &\quad \oplus H^x \otimes H^y \otimes L^z \oplus H^x \otimes H^y \otimes H^z \end{aligned} \quad (3)$$

where $\oplus$ and $\otimes$ denote space direct sum and direct multiplication, respectively. H^λ and L^λ represent the high and low-pass directional filters along the direction of λ-axis, where λ={x, y, z} (Chen and Ning, 2004). After the 2D DWT is applied in one level, 4 sub-bands are created. Meanwhile, 8 sub-bands are formed from the 3D DWT. Each sub-band obtained through two frequency filters is named by directions, frequency filter passed in: LLL, LLH, LHL, HLL, LHH, HLH, HHL, and HHH. Fig. 2 is the diagram of 3D wavelet decomposition in one level.

Fig. 3 shows the result of the 3D DWT using a simple test model such as a cube having computed numeric values differing from those of the neighborhood. This test model is composed of 64*64*64 pixels. The size of the cube in the middle of test model is 54*54*54 pixels. In this model, LLL band, the result passed through the low frequency in all direction refers the approximation of a given model. Sub-bands which are passed high frequency filter in only one direction show the each marginal plane in that direction. Sub-bands by high frequency filter in two directions delineate the boundary lines. HHH band passed through high frequency in all directions shows vertexes of angle.

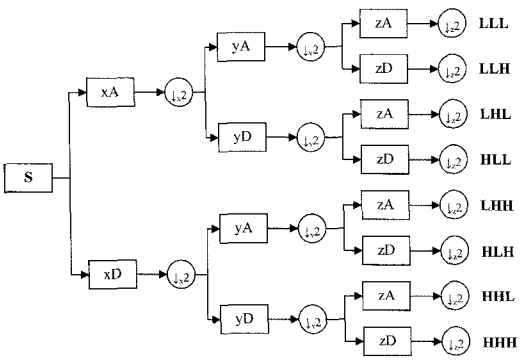


Fig. 2. The diagram of one level decomposition for the 3D DWT.

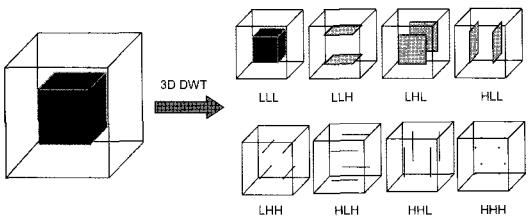


Fig. 3. The 3D DWT result simple model. The size of simple model is 64 by 64 by 64 and 54 by 54 by 54 cube is in the middle of model.

3) The procedure of the 3D DWT based classification

The work flow in this study is shown in Fig. 4. Firstly, bands are arranged in the order of frequency. And then, the number of data in row, column and depth directions should be a power of two to implement the 3D DWT using multi-spectral imagery. In case of row and column directions, we can adjust the size into a power of two when we make a subset image. Whereas, in case of z direction,

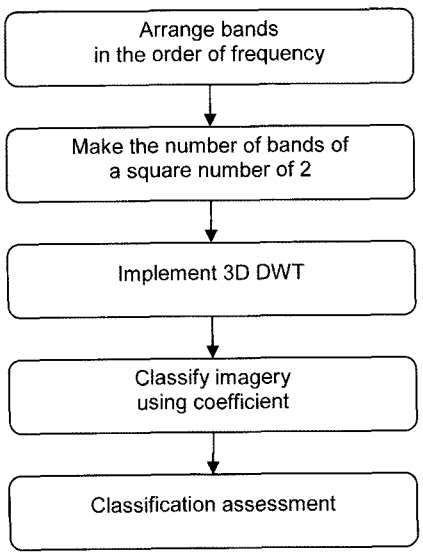
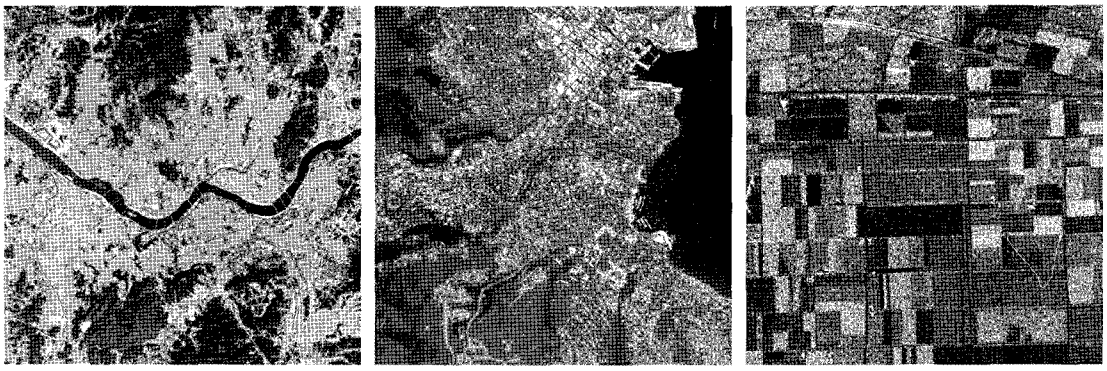


Fig. 4. Flowchart of the 3D DWT based classification procedure.

that is a band direction, we can make band numbers a power of two by adding last band not to add arbitrary detail information. The next step is to implement the 3D DWT. The used basis function is haar wavelet which is advantageous in urban images containing buildings and road boundaries where pixel value changes dramatically and multi-spectral image has not have many bands. Lastly, we classify the image using original pixel and 3D wavelet coefficients and then evaluate their results. The maximum likelihood classifier is used for these image data sets. In previous study, we tested classification using 3D DWT coefficients and evaluate their results quantitatively with those of traditional classification techniques including the 2D wavelet scheme (Yoo *et al*, 2007). In former study, we confirmed that the classification technique of the 3D DWT is more effective than 2D DWT based scheme as well as traditional classification technique. However, all sub bands were used in that study and there was the problem about uncertain reference data. In this study, we focus on comparing 3D DWT based classification method with pixel based classification method. Among the 3D DWT sub-bands, only LLL and LLH band are used for classification because too much complicated information in row and column direction interfere with making land cover map. As only LLL and LLH bands are used for classification, the higher accuracy can be obtained than the case of using all sub-bands. Training data sampled at random and reference data offered by National Geography Information Clearinghouse of Korea (NGIC) are newly used for this study.

3. Classification using the 3D DWT coefficient

Three images were used for this study. Three data images among them are optical images with different resolutions: Landsat7 ETM+ and Ikonos images (Fig. 5(a), (b)). The other is the synthetic aperture radar image, AirSAR image of Flevoland, Netherlands. GRSS DFC provided that image for public uses, which is composed of C, L, and P bands and full polarization; HH, HV and VV (Fig.5 (c)). Landsat 7 ETM+ imagery is covered with Seoul, Korea and has 8 bands with spatial resolution of 15m (panchromatic band) 30m (multi-spectral bands) and 60m (thermal infrared band), but 7 bands were used for the case studies in this study except for the panchromatic band. The Ikonos imagery in Hobart, Australia is one of ISPRS dataset collections (http://www.isprs.org/data/index.html) and is composed of red, green, blue,



(a) Landsat7 ETM+ (b) Ikonos (c) AirSAR

Fig. 5. The original images: (a) Landsat7 ETM+ (30m), (b) Ikonos (1m), (c) AirSAR (10m)

near infrared bands (4m).

Fig. 6 shows the training data sets and reference data sets. The reference data of Landsat7 ETM+ (Fig. 6(b)) is the land cover map provided by NGIC and this map contains 6 classes: forest, water, grass, residential or commercial area, bare soil and field. The map can not be subdivided into individual building because of low resolution. The reference data of Ikonos image is divided into residential area, forest, water, commercial area and grass (Fig. 6(d)). The number of reference pixels by class is 142277 in residential area, 42546 in forest, 84391 in water, 29224 in commercial area and 1892 in grass, respectively. This reference set is collected from image by visual check because classes in high resolution image can be divided through eyes. The reference data set of SAR image is obtained by visual check from ground truth data provided by GRSS

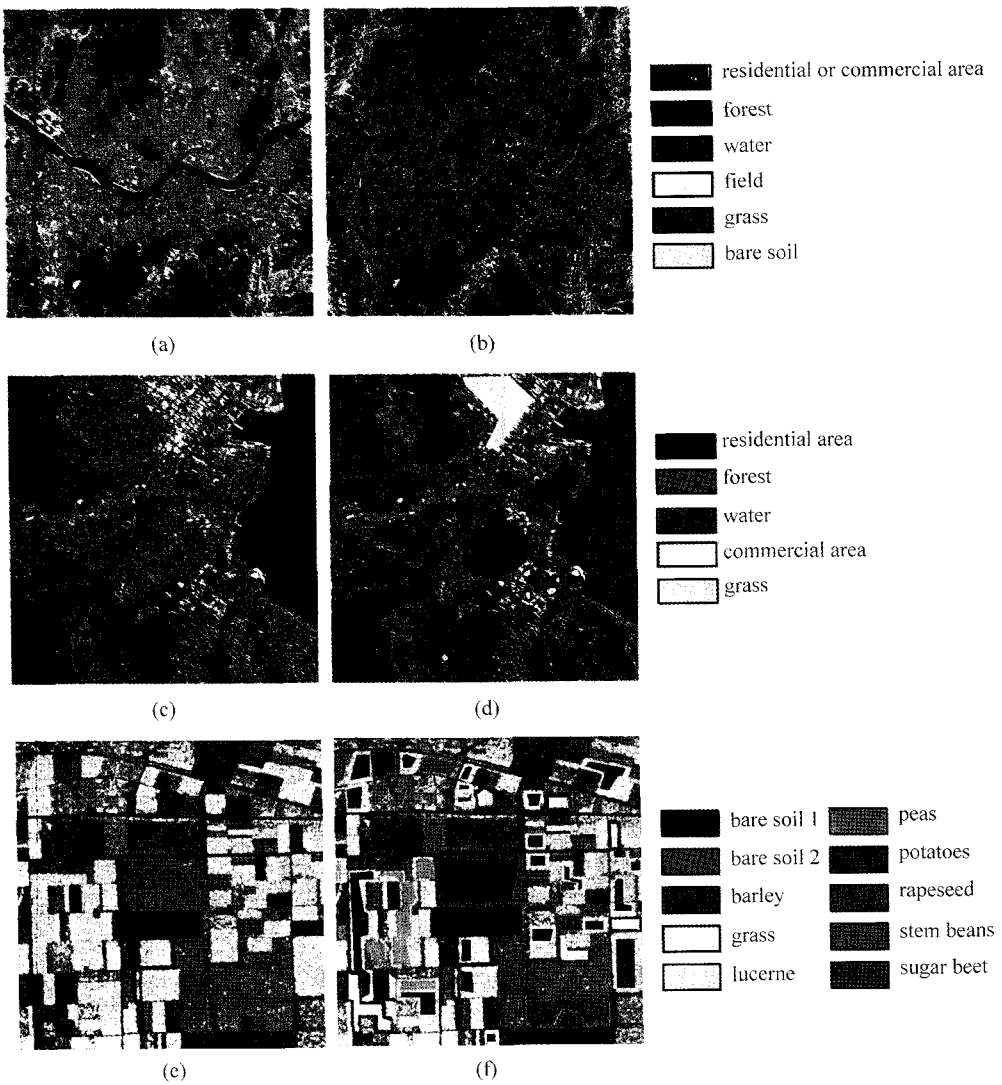


Fig. 6. The training and reference data sets for supervised classification with class indexes: (a) The training data set of Landsat7 ETM+ image. (b) The reference data set of Landsat7 ETM+ image. (c) and (d) show the training data and reference data of Ikonos image. The training data of AirSAR image is (e) and the reference data is (f).

Data Fusion Community test dataset. SAR image is classified by 10 classes: bare soil 1 (3919 pixels), bare soil 2 (3469 pixels), barley (3121 pixels), grass (5610 pixels), lucerne (921 pixels), peas (2144 pixels), potatoes (17702 pixels), rapeseed (10579 pixels), stem beans (3571 pixels) and sugar beet (5895 pixels). The training data sets of three images are obtained 10% random sampling of each reference data set (Fig. 6(a), (c), and (e)). The explanation between the class and the color is given in the right part of the reference data set.

Fig. 7 shows the classification results according to image and method. In case of Landsat image, we cannot check differences between the classification result of original pixel value and the 3D DWT coefficient in eyes (Fig. 7(a) and (d)). In the Ikonos image, the result of classification using the 3D DWT coefficient differs from the result of original method (Fig. 7(b) and (e)). Especially residential area is extended remarkably. Forest, grass and water classes are composed of simple features respectively. Meanwhile, residential area is composed of various features such as house, grass, wood, and road. Accordingly, the result of 3D DWT based classification is shown extended residential area because of reducing spatial variations.

In Fig. 7(c) and (f), the 3D DWT coefficient is more effective than the original pixel in SAR imagery. The speckle noises of SAR image are reduced by enhanced Lee filter before the classification. However, Fig. 7(c) has too much speckles yet. Meanwhile, the 3D DWT method

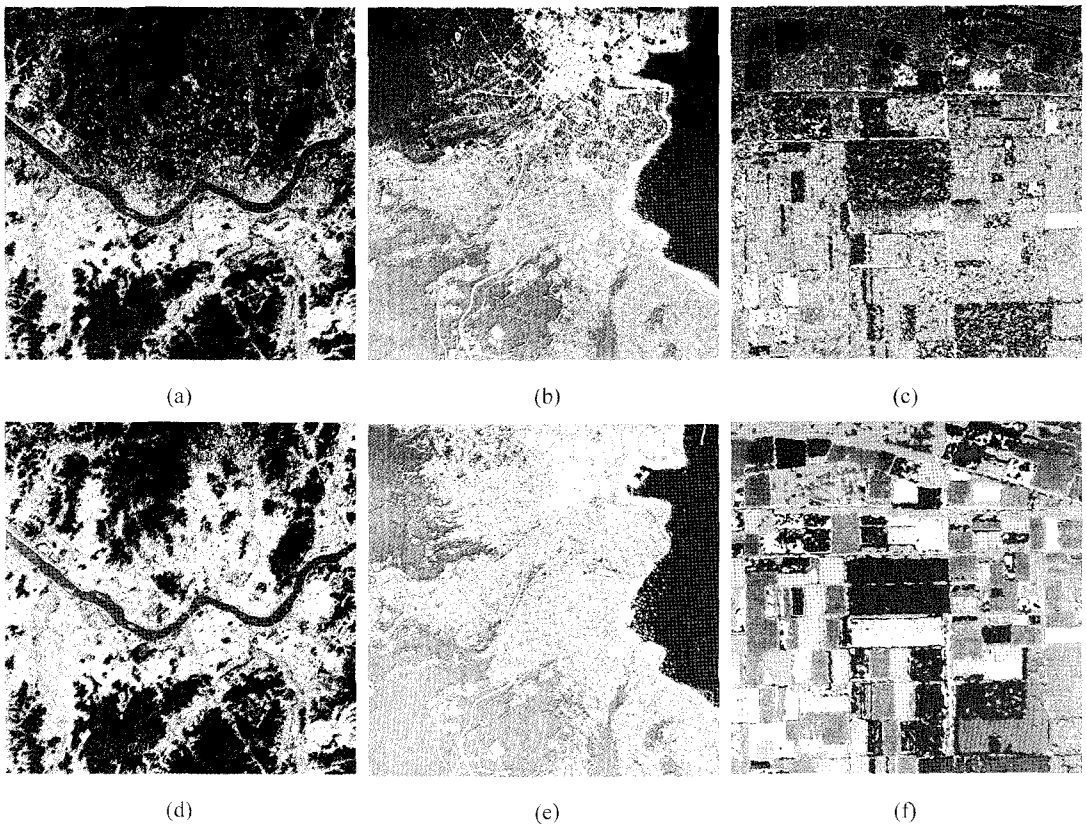


Fig. 7. The maximum likelihood classification results. Original pixel values are used for classification in (a), (b) and (c) and the 3D DWT coefficients are used in (d), (e) and (f).

makes dramatic effect in SAR imagery as removing H bands in spatial direction among 3D DWT sub-bands makes speckle noises reduce. These are only visual assessments. For the accuracy assessment to evaluate quantitatively, the following section shows quantitative accuracy assessments for these results.

4. Accuracy Assessment

To evaluate the ability of classifying image linked with the 3D DWT scheme, the classification confusion matrixes are presented. Table 1 shows the overall accuracy and kappa coefficient for all applied images. As shown in the Table 1, the classification accuracy using the wavelet coefficients is higher than that of conventional methods on all cases (Landsat7 ETM+, 68.6293%; Ikonos, 86.1267%; AirSAR, 97.6832%). These results are thought to cause that traditional methods have used only spectral information, whereas the 3D wavelet method can use both spectral and spatial information in classification processes. The accuracy improvement of AirSAR images is the most conspicuous among three test images from 60.2203% to 97.6832%. These results seem to be shown due to the elimination of the speckle noise effect caused by 3D DWT. The overall accuracies of Landsat7 ETM+ image are the lowest among three images. This result is referable to reference data covered with whole area. The overall accuracies of other images are higher than that of Landsat image, as we used not whole area but selected representative area by class for accuracy assessment.

To check the accuracies by class, user's accuracy by each image are presented in Table 2, 3 and 4 respectively. User's accuracy is computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified that category. This is a measure of commission error and indicates the probability that a pixel classified into a given category actually represents that category on the ground.

Table 2 shows the user's accuracy of Landsat imagery according to classes. The water class has the highest accuracy on using the 3D DWT coefficients (83.53%). In forest, water, residential or commercial area and field classes, the cases of using the 3D DWT coefficient have higher accuracy than the cases of using original pixel values. The accuracy of forest class increases from 79.64% to 82.11%. The accuracy of water class improves from 75.61% to 83.53%. The accuracy's improvements of residential or commercial area and field classes are from 78.72% to 81.69% and from 24.61% to 27.82%, respectively. In grass and bare soil class, there are only small differences between original pixel and the 3D DWT coefficient but these classes are more accurate in case of using original pixel values. Grass, bare soil and field classes have relatively low accuracy (grass (25.95%), bare soil (18.91%) and field (27.82%)). These classes are the cause of low overall accuracy.

Table 3 is the user's accuracy of Ikonos image.

Table 1. The overall accuracy and kappa coefficient of classification results (in percent). Kappa coefficients are given in parentheses after overall accuracy. The case of using original pixel values is marked as 'original'. '3D DWT' means using the 3D DWT coefficients for classification.

	Landsat7 ETM+	Ikonos	AirSAR
original	68.2593 (0.5330)	77.0351 (0.6914)	60.2203 (0.5409)
3D DWT	68.6293 (0.5554)	86.1267 (0.8043)	97.6832 (0.9722)

Table 2. The user's accuracy of each class in Landsat7 ETM+ image (in percent). 'R or C' means residential or commercial area class.

	forest	water	grass	R or C	bare soil	field
original	79.64	75.61	26.74	78.72	19.14	24.61
2D DWT	81.47	82.23	27.22	80.78	19.01	27.67
3D DWT	82.11	83.53	25.95	81.69	18.91	27.82

The 3D DWT coefficient is more effective in residential area, forest and water classes for Ikonos image. Especially, appreciable improvement is shown in residential area class (from 54.72% to 73.64%). However, the commercial area and grass classes have lower accuracy in case of using the 3D DWT coefficients.

The accuracy statistics for each class in AirSAR image is shown in Table 4. On using SAR, the accuracy using the 3D DWT is the highest in all classes. Especially, the accuracy of sugar beet class increases from 30.17% in case of using original pixel values to 97.50% in case of using the 3D DWT coefficient. The 3D DWT coefficient makes a great advance in bare soil 2, barley, lucerne, peas stem beans and sugar beet classes. The user's accuracies in almost classes are greater than 90% in case of using the 3D DWT coefficient. These all results are attributed that removing H bands of the 3D DWT scheme reduce speckle noises.

Judging from this accuracy assessment, the proposed 3D DWT coefficient shows good results on the high resolution imagery and SAR image containing speckle noises.

Table 3. The user's accuracy of each class in Ikonos image (in percent). Residential area and commercial area are signified by 'residential' and 'commercial', respectively.

	residential	forest	water	commercial	grass
original	95.03	77.69	99.99	38.15	14.97
3D DWT	96.33	89.35	100.00	45.77	55.80

Table 4. The user's accuracy of each class in AirSAR image (in percent).

	bare soil1	bare soil 2	barley	grass	lucerne
original	100	40.75	50.89	93.88	25.95
3D DWT	100	95.41	99.21	99.41	90.26
	peas	potatoes	rapeseed	stem beans	sugar beet
original	23.12	82.45	74.66	63.34	30.17
3D DWT	81.07	99	99.25	95.58	97.50

5. Concluding Remarks

In this study, it is carried out the classification using the 3D DWT coefficients and the quantitative accuracy evaluation concerned. The suggested methods are more effective than the pixel based scheme. The 3D DWT coefficient can increase classification accuracy without adding new information because the 3D wavelet scheme uses both spectral and spatial information simultaneously. The 3D DWT is more effective to classify road, bare soil, forest and building class, which are mixed with various materials, than water and grass class that are classified rather correctly even though using traditional method. The 3D DWT based classification method shows especially higher accuracy in SAR imagery contained speckle noises. Also, the 3D DWT value can be expected to apply for simplifying classification result and generating land-cover thematic map.

Only one image by sensor is used for this study and some of area is used for accuracy assessment. However, this result shows that possibility of applying the 3D DWT scheme to the classification of multi-spectral and SAR image. We just applied the 3D DWT based scheme to multi-spectral image classification, but it can be extended to multi-temporal image classification or other image processing later on.

Acknowledgement

This work is financially supported by Korean Research Foundation (KRF) and Korean Aerospace Research Institutes (KARI). The author (Yoo, H.Y.) is supported by the Brain Korea 21 project in 2007.

References

Addison, P. S., 2002. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, IOP Publishing, Bristol, Philadelphia, 6p.

Antonini, M., M. Barlaud, P. Mathieu, and I. Daubechies, 1992. Image coding using wavelet transform, *IEEE Transaction on Image Processing*, 1(2): 205-220.

Chen, Z. and R. Ning, 2004. Breast volume denoising and noise characterization by 3D wavelet transform, *Computerized Medical Imaging and Graphics*, 28(5): 235-246.

Cheung, N., C. Tang, A. Ortega, and C. S. Raghavendra, 2006. Efficient wavelet-based predictive Slepian-.Wolf coding for hyperspectral imagery, *Signal Processing*, 86(11): 3180-3195.

Ghugre, N. R., M. Martin, M. Scadeng, S. Ruffins, T. Hiltner, R. Pautler, C. Waters, C. Readhead, R. Jacobs, and J. C. Wood, 2003. Superiority of 3D wavelet-packet denoising in MR microscopy, *Magnetic Resonance Imaging*, 21(8): 913-921.

Klock, H., A. Polzer, and J. M. Buhmann, 1997. Region-Based Motion Compensated 3D-Wavelet Transform Coding of Video, *Proceedings of the 1997 IEEE International Conference on Image Processing*, Vol.2, 26-29 October 1997, Santa Barbara, California (IEEE Signal Processing Society, Piscataway, New Jersey) , pp. 776-779.

Luo, L., F. Wu, S. Li, Z. Xiong, and Z. Zhuang, 2004. Advanced motion threading for 3D wavelet video coding, *Signal Processing: Image Communication*, 19(7): 601-616.

Mallat, S. G., 1989. A theory of multi-resolution signal decomposition: The wavelet representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11(7): 674-693.

Spaendonck, R. L. C. V., T. P. H. Steeghs, I. Overeem, F. C. A. Fernandes, and J. T. Fokkema, 2001. Wavelet based volume attributes for seismic interpretation, Proceedings of EAGE 63rd Conference & Technical Exhibition, 11 - 15 June, Amsterdam, Netherlands (European Association of Geoscientists & Engineers, Houten, The Netherlands), unpaginated CD-ROM.

Tzanetakis, G., G. Essl, and P. Cook, 2001. Audio Analysis using the Discrete Wavelet Transform, *Proceedings of WSES International Conference Acoustics and Music: Theory and Applications* (AMTA 2001), 26-30 September, Skiathos, Greece (World Scientific and Engineering Academy and Society, Wisconsin, USA), unpaginated CD-ROM.

Xiong, R., J. Xu, F. Wu, and S. Li, 2005. Studies on Spatial Scalable Frameworks for Motion Aligned 3D Wavelet Video Coding, *Proceedings of SPIE*: Visual Communications and Image Processing 2005, Vol. 5960, 31 July 2005, Bellingham, Washington (SPIE, Bellingham, Washington), pp. 189-200.

Yoo, H. Y., K. Lee, and B. D. Kwon, 2007, Implementation of 3D Discrete Wavelet Scheme for Space-borne Imagery Classification and Its Application, *Proceedings of IEEE International Geoscience and Remote Sensing Symposium 2007*, 25 July 2007, Barcelona, Spain, unpaginated CD-ROM.