

The Impacts of Decomposition Levels in Wavelet Transform on Anomaly Detection from Hyperspectral Imagery

Hee Young Yoo*[†] and No-Wook Park**

*Geoinformatic Engineering Research Institute, Inha University

**Department of Geoinformatic Engineering, Inha University

Abstract : In this paper, we analyzed the effect of wavelet decomposition levels in feature extraction for anomaly detection from hyperspectral imagery. After wavelet analysis, anomaly detection was experimentally performed using the RX detector algorithm to analyze the detecting capabilities. From the experiment for anomaly detection using CASI imagery, the characteristics of extracted features and the changes of their patterns showed that radiance curves were simplified as wavelet transform progresses and H bands did not show significant differences between target anomaly and background in the previous levels. The results of anomaly detection and their ROC curves showed the best performance when using the appropriate sub-band decided from the visual interpretation of wavelet analysis which was L band at the decomposition level where the overall shape of profile was preserved. The results of this study would be used as fundamental information or guidelines when applying wavelet transform to feature extraction and selection from hyperspectral imagery. However, further researches for various anomaly targets and the quantitative selection of optimal decomposition levels are needed for generalization.

Key Words : Hyperspectral image, Discrete wavelet transform, Dimension reduction, Anomaly detection

1. Introduction

Hyperspectral imagery consists of hundreds narrow spectral bands with very fine spectral resolution and full spectral curves at each pixel location can be obtained from it. Hence, it can provide much more detailed information than multispectral imagery and objects of interest can be effectively detected and identified using this detailed

information. Thus, hyperspectral imagery has been applied to various fields such as mining, geology, ecology, vegetation and water quality analysis (Cochrane, 2000; Igamberdiev *et al.*, 2011; Kurz *et al.*, 2012). However, hyperspectral imagery has too many bands which include highly correlated redundant spectral information. These may make it hard to analyze data in detail and require time-consuming processing steps. Thus, one of main issues

Received October 21, 2012; Revised November 23, 2012; Accepted November 26, 2012.

[†] Corresponding Author: Hee Young Yoo (yhy1003@gmail.com)

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

in hyperspectral image analysis is to summarize and extract significant information effectively.

This issue is very important for anomaly detection among various application fields of hyperspectral remote sensing, which aims at extracting anomalous pixels from hyperspectral imagery. The anomaly detection procedure for hyperspectral imagery is mainly divided into two stages: feature extraction and anomaly detection. (Chang and Chiang, 2002; Kwon and Nasrabadi, 2005; Banerjee *et al.*, 2006; Gu *et al.*, 2008). Among two main stages, this study focuses on feature extraction stage using wavelet transform to define the characteristic of representative normal and anomalous regions and to divide effectively each other. For the feature extraction stage, PCA (principal component analysis) and kernel PCA have been mainly used for both dimension reduction and extraction of effective features for anomaly detection. When using PCA for dimensionality reduction or feature extraction, the higher ranked PCs are usually selected and used for further analysis. The first principal component tends to have the largest possible variance of original data. Lower ranked components have high frequency information or noise. By the definition of anomaly, however, anomaly may occur relatively infrequently. So, the higher ranked PCs may be inappropriate for anomaly detection and it is not easy to distinguish differences between noise and anomaly from low ranked PCs. Meanwhile, wavelet transform decomposes data into both frequency and time or location information and it is useful for the analysis of finite and non-stationary datasets. Wavelet transform has been used for many applications in image processing including feature extraction and dimensionality reduction. Especially, 3D DWT (3 dimensional discrete wavelet transform) developed for volumetric data has been effectively applied to hyperspectral image analysis for dimension reduction (Bruce *et al.*, 2002; Cheung *et al.*, 2006).

Furthermore, Park *et al.* (2012) verified that features extracted from wavelet transform outperformed for anomaly detection. Although 3D DWT has proven to be useful in hyperspectral imagery analysis, the study on how the information content changes according to the change of a wavelet decomposition level has not been fully investigated.

In this study, the effects of features extracted from wavelet decomposition are investigated as pre-processing for anomaly detection from hyperspectral imagery. After applying wavelet decomposition, some points are sampled at both anomaly and background areas. At these sampled points, the characteristics of extracted features are examined for every decomposition level. Then, features extracted by wavelet transform are experimentally applied to anomaly detection to evaluate the detecting capability.

2. Theory

1) Anomaly detection

An anomaly can be defined as a deviation from the normal or common order or form or rule. In science fields, an anomaly is to be out of uniformity in physical properties or deviation from predictable measurements. Researchers in remote sensing field have used anomalies as pixels having very high or low. Anomaly detection attempts to locate anything that looks different spatially or spectrally from its surroundings. It is often confused with target detection. Anomaly detection is to search something different from neighborhood without any prior information. Meanwhile, target detection aims to search the pixels for the presence of a specific material and in this case, it is assumed that the characteristics of target pixels have been known

(Camps-Valls and Bruzzone, 2009). Therefore anomaly detection is more broad and comprehensive approach than target detection and the result can be used as input for further target detection analysis.

2) Wavelet transform

Wavelet transform is a numerical analysis to capture not only frequency but also location information unlike Fourier transform. Discrete wavelet transform is a sub-band coding for discretely sampled data. The wavelet transform decomposes data into L (low) bands or approximation components and H (high) bands or detail components using band-pass filtering. 2D DWT results are obtained by first filtering the signal in row direction and then re-filtering the output in column direction by the same filters. 3D DWT is the expansion of the wavelet theory with respect to one-dimensional signal into three dimensions (Muraki, 1993). After data is

decomposed by the 3D DWT, 8 sub-bands are generated such as LLL, LLH, LHL, HLL, LHH, HLH, HHL, and HHH (Yoo *et al.*, 2007). The size of dataset is halved after the decomposition because of downsampling stage in the wavelet transform. The decomposition process can be iterated with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components. Fig. 1 shows the 2-level decomposition procedure of wavelet transform for 1, 2 and 3D dataset respectively.

When applying 3D DWT to hyperspectral imagery, the input imagery can be summarized and analyzed in multi-resolution through downsampling process. Spatial and spectral resolutions become lower in this process simultaneously as shown Fig. 1. Thus, the integrated process of 3D DWT and 1D DWT is applied to test imagery to detect anomalies by preserving spatial resolution in this study. As a

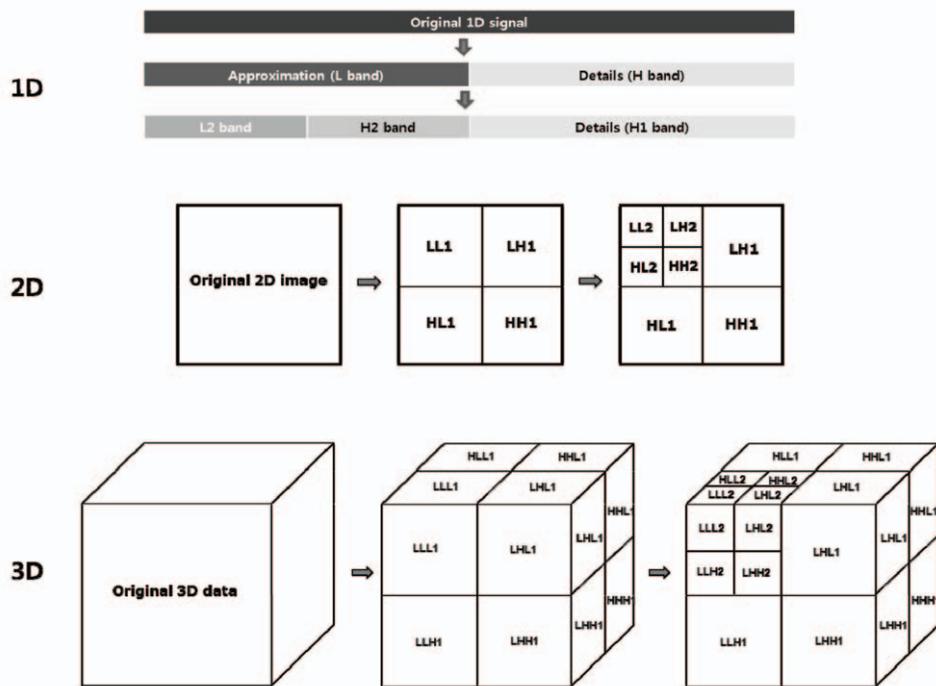


Fig. 1. The wavelet decomposition procedure for 1, 2, and 3D dataset.

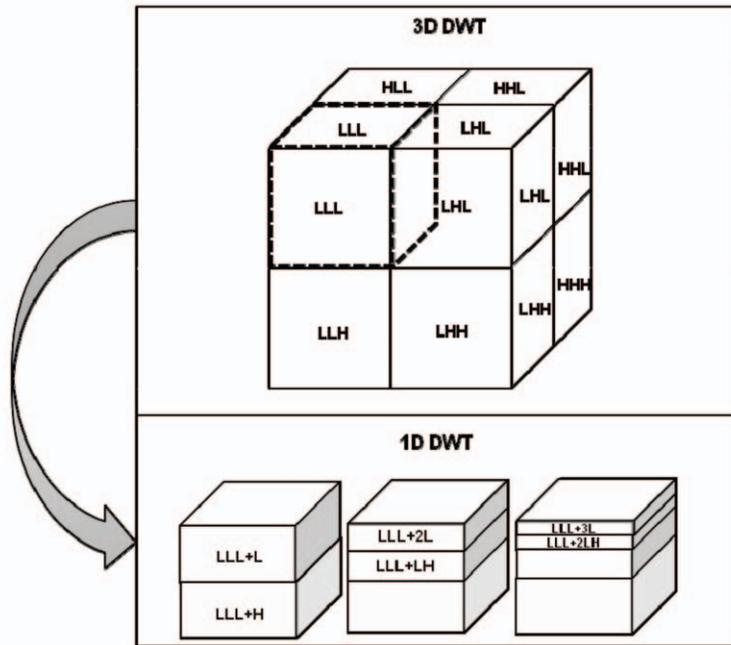


Fig. 2. The integrated process of 3D wavelet decomposition in all directions and 1D wavelet decompositions in the spectral direction.

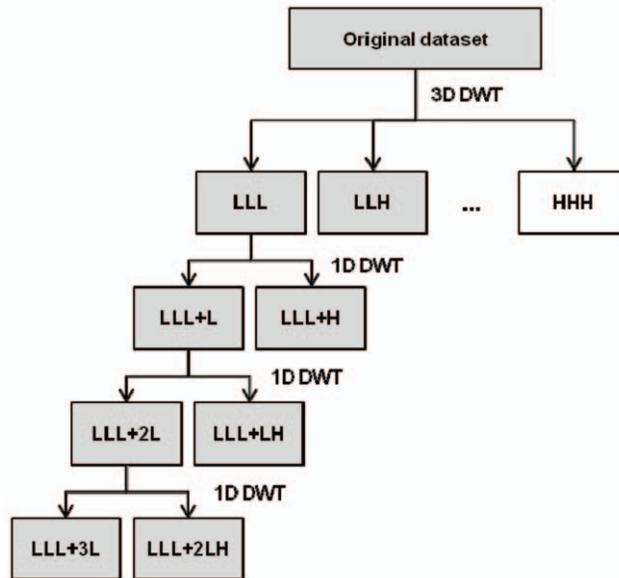


Fig. 3. The overview of the wavelet transform process and the dataset for detailed analysis (Gray boxes).

result of application of 3D DWT, the original data are decomposed both spatially and spectrally in a simultaneous way. Meanwhile, 1D DWT is used for dimensionality reduction and spectral information analysis. At first, full hyper spectral imagery is

transformed by 3D DWT. And then, 1D DWT is iterated three times for only LLL band which is obtained by 3D DWT. 1D DWT is carried out in only spectral direction. For the next 1D transform, only L band is used. Details of the process used in this study

are presented in Fig. 2 and 3. Fig. 2 shows the wavelet decomposition process of 3D hyperspectral cube in detail. The repetition of 1D DWT is marked after plus sign. For example, LLL+L means that 1D DWT is once applied to LLL band after 3D DWT. LLL+2LH stands for H band obtained after three 1D DWT procedures. Among a lot of basis functions, Db4 basis function is used for the wavelet transform. Db4 constructed by Daubechies(1988) is one of the widely used basis functions in remote sensing field. Finally, 9 datasets including original data are used for the experiment and these dataset are presented in gray boxes in Fig. 3.

3. Data

The airborne CASI-1500 imagery is used for this experiment (Fig. 4(a)). The detailed specification of CASI-1500 is described in Table 1. The area is simply classified into target and background areas as shown in Fig. 4(b). The targets are power line towers appearing very brightly in the study area. Forests are dominant in non-target background areas.

At first, the integrated process of 3D DWT and 1D DWT is performed on the airborne test imagery. After wavelet decomposition, some samples representing target and non-target area are extracted for detailed analysis. 10 sample points are displayed with CASI-1500 imagery in Fig. 4(a). Black asterisks are selected for target areas and white cross signs are for non-target areas. The sampled points for non-target area are mainly extracted from forest areas for the comparison with bright target areas. The characteristics of extracted features and the changes of their patterns are examined at every decomposition level. Then, to verify features extracted by 3D DWT, anomalies are detected experimentally using the Reed-Xiaoli (RX) detector algorithm which is traditionally used for anomaly detection (Reed and Yu, 1990). The RX detector algorithm extracts targets that are spectrally distinct from the image background and expression for the conventional RX detector is given by

$$d(x, N(x)) = (x - \mu)^T \Sigma^{-1} (x - \mu) \quad (1)$$

where $N(x)$ is N by N neighbor matrix, μ is the mean of $N(x)$, and Σ is the covariance matrix of $N(x)$.

Consequently, the pixel having higher possibilities

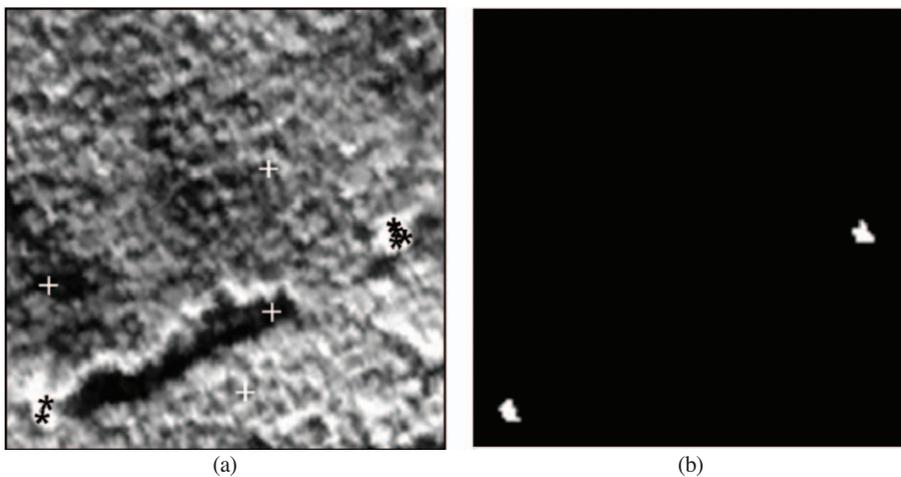


Fig. 4. The test dataset: (a) CASI-1500 imagery with sampling points; black asterisk means sample location for target area and white cross sign is for non-target area, (b) The binary image classified into target (white) and non-target areas (black).

Table 1. The specification of experimental data set

	CASI-1500
Acquisition date	Oct. 26 2010
Acquisition height	2,500m
Spatial resolution	1.25m
Wavelength range	365-1,050nm
The number of bands	96
Band width	7nm
Radiometric resolution	14bit

to be anomaly is displayed brighter. If the value is greater than a defined threshold, then pixel x is an anomaly. Finally, the detecting capabilities are analyzed using ROC curves. ROC curves are provided to evaluate the results of anomaly detection according to wavelet decomposition level and frequency.

4. Results

1) Wavelet analysis

As aforementioned approach, the integrated wavelet transform is performed on the test dataset in 4 decomposition levels. Then, some samples representing target and non-target area are extracted and radiance curves are plotted at each sampling location and every decomposition level. As a result, the every sample data in single class tend to show very similar radiance curves. Thus, the average values are calculated from all sample points for target and non-target classes to display the results simply. The average of radiance curves are presented in Fig. 5. Solid line is the radiance curve of target area and dash line is that of non-target area. The target curve is generally higher than the non-target curve. Especially, it is much higher in short wavelength bands. Some breaks are appeared simultaneously in specific spectral bands in both target and non-target

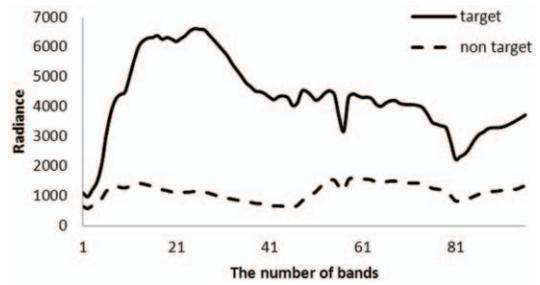


Fig. 5. The average of radiance curves of sampled points as target and non-target class, respectively.

areas.

Fig. 6 shows full decomposition curves. Each curve at top shows the original curve for each class. Below original curves, the left panel shows approximation components, L bands. Detail components, H bands are displayed in the right panel. It is expected that LLH band have a lot of information about changes between bands. However, these curves show similar patterns regardless of classes although the magnitude is different from each other. H band at the first decomposition level is interpreted to reflect not relative changes but absolute displacements. In case of LLL band, there are few changes. These results indicate that LLL bands can almost explain fully the dataset with a half bands of original dataset. LLL+L bands show more simplified patterns of original dataset but the overall shapes are preserved. The curves of LLL+H bands still show similar patterns to each other. After the second 1D DWT, overall shapes of L bands start to change but overall patterns are still preserved. In LLL+LH, there are some unique patterns for each class in short wavelength bands. In LLL+3L bands, the whole shapes are not retained anymore and LLL+2LH bands look very similar to each other. Similar patterns between target and non-target classes are observed in all H bands, regardless of the decomposition level except H band in the third level. It means that it is difficult to distinguish target and non-target classes

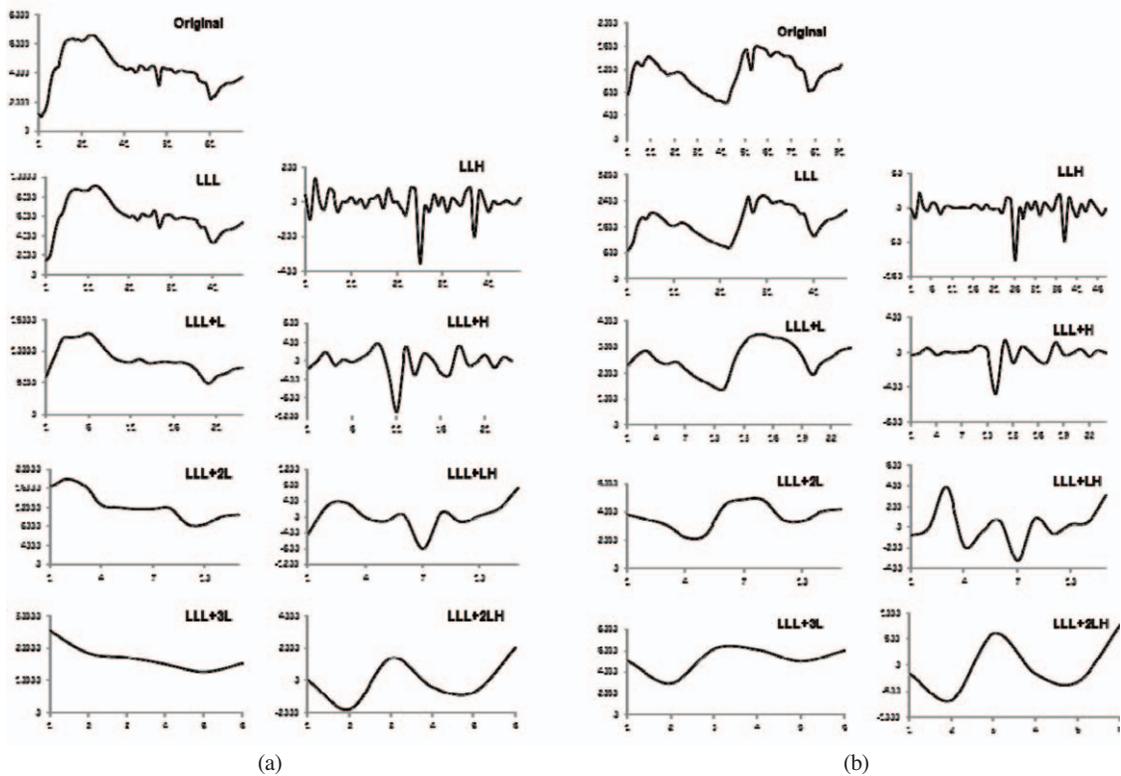


Fig. 6. Multi-stage analyses for hyperspectral imagery belonging to target and non-target areas using wavelet transform: (a) target area and (b) non-target area. X-axes of all graphs mean band number, Y axis of the graph for original is radiance and Y axes for wavelet sub-bands mean wavelet coefficients of radiances.

using very high frequency component. To sum up all wavelet analysis, these results show the appropriate wavelet decomposition level might be the second 1D DWT level for the test dataset and L bands can be more useful than H bands.

2) Anomaly detection

To evaluate the applicability of wavelet analysis results for anomaly detection, anomalies are experimentally detected using the RX detector algorithm. The results are given in Fig. 7. Under the algorithm in section 3, the brighter pixels have higher possibilities to be anomaly. All cases detect two main target areas and some non-target areas are also displayed brightly. However, it is difficult to distinguish the capability differences visually.

Therefore, to visualize the performance this result, the ROC curve is used. For the ROC curve, the probability of detection is plotted in function of the probability of false alarm for different thresholds. The closer the ROC curve is to the upper left corner, the higher the overall accuracy of the test. Fig. 8 is comparing original dataset, LLL band and LLH band. As mentioned wavelet analysis part, H band in early level cannot explain the difference between target and non-target area. Therefore L bands can be more effective than H bands for anomaly detection. Fig. 9 shows the ROC curves of only L bands according to decomposition level. The higher level of wavelet decomposition shows better performance, until LLL+2L band. However, using LLL+3L band, the performance is not improved anymore. So, for this

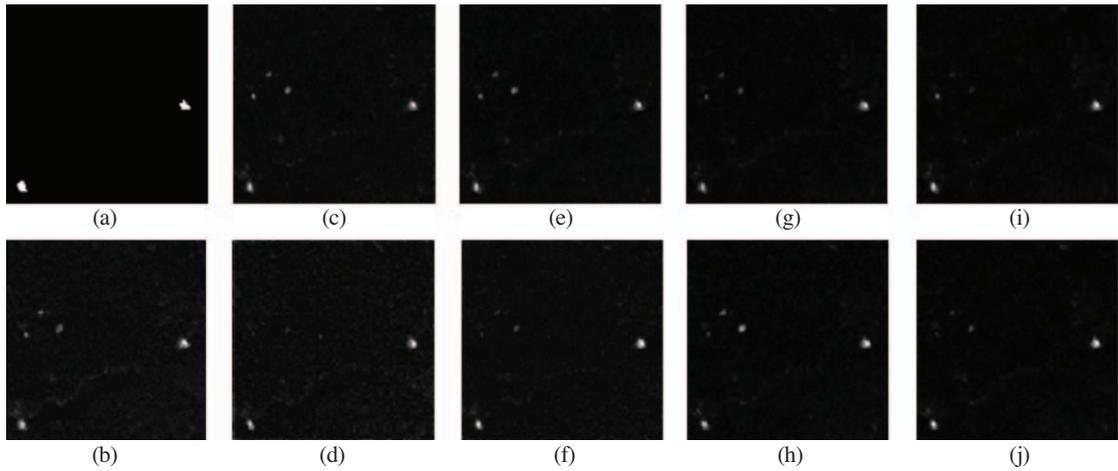


Fig. 7. Ground truth for anomaly detection (a) and the results of anomaly detection using the RX detector algorithm at every decomposition level (b~j); (b) original, (c) LLL, (d) LLH, (e) LLL+L, (f) LLL+H, (g) LLL+2L, (h) LLL+LH, (i) LLL+3L and (j) LLL+2LH.

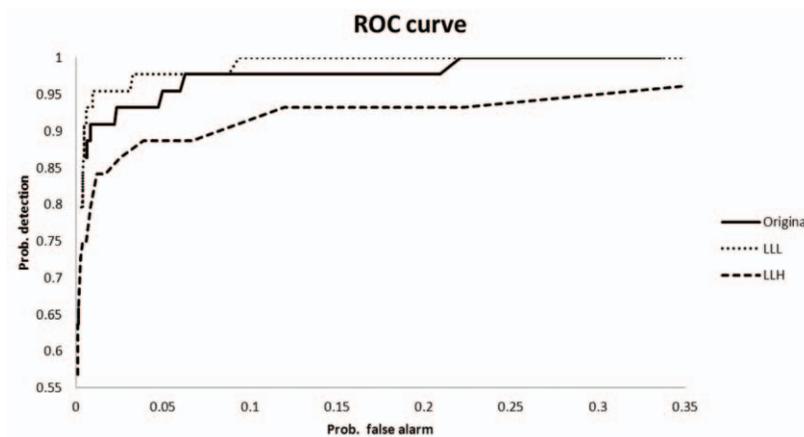


Fig. 8. ROC curves for comparing the detecting capability of original dataset, LLL band and LLH band.

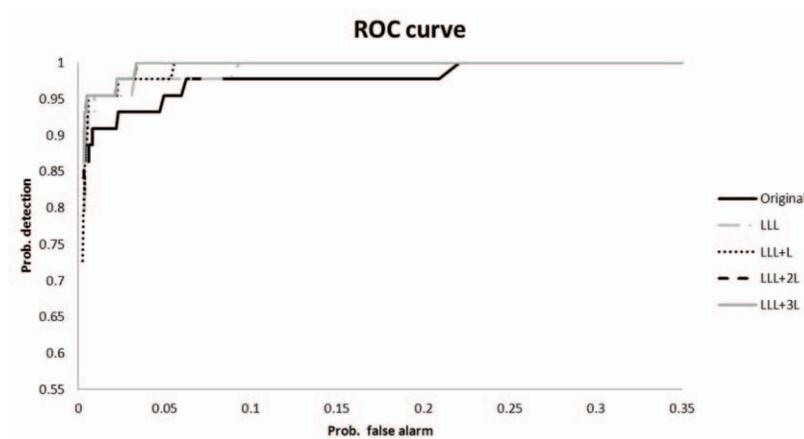


Fig. 9. ROC curves for L bands at every decomposition level.

dataset, the appropriate decomposition level is the 2nd level of 1D DWT. These all results of anomaly detection and their ROC curves are support the finding from wavelet analysis that is the appropriate decomposition level could be the 2nd level of 1D DWT and L band is better than H band for anomaly detection. Thus, wavelet analysis could be reference information to extract features for anomaly detection.

5. Conclusions

This study has investigated how DWT features can affect the anomaly detection results. A special focus is on the aspects of wavelet based features according to the change of decomposition level. From the experiment using CASI imagery, it was observed that L bands were appreciably summarized and simplified according to the decomposition level. Meanwhile, H bands did not show any unique characteristics irrespective of target or non-target areas. Thus, L bands can be more effective than H bands for anomaly detection, especially for spatially clustered anomalies like test dataset. The appropriate decomposition level could be the one where the overall shape of profile is preserved. The results of experimental anomaly detection were consistent with the findings on visual wavelet analysis for the decision of appropriate decomposition level. Thus, the results of this study would be used as fundamental information or guidelines when applying wavelet transform to feature extraction and selection for hyperspectral imagery. For major findings of this study to be general ones, extensive experiments on various anomaly targets should be carried out and additional researches for selecting wavelet decomposition level quantitatively are needed.

Acknowledgements

This research was supported by the Defense Acquisition Program Administration and Agency for Defense Development, Korea, through the Image Information Research Center at Korea Advanced Institute of Science & Technology under the contract UD100006CD. Participation of No-Wook Park was partially supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology(2012R1A1A1005024). The authors thank Prof. Kyu-Sung Lee and Dr. Jeong-Il Shin at Inha University for providing the preprocessed CASI imagery.

References

- Banerjee, A., P. Burlina, and C. Diehl, 2006. A support vector method for anomaly detection in hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 44(8): 2282-2291.
- Bruce, L.M., C.H. Koger, and J. Li, 2002. Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction, *IEEE Transactions on Geoscience and Remote Sensing*, 40(10): 2331-2338.
- Camps-Valls, G. and L. Bruzzone, 2009. *Kernel Methods for Remote Sensing Data Analysis*, John Wiley & Sons Ltd, Chichester, UK.
- Chang, C.-I. and S.-S. Chiang, 2002. Anomaly detection and classification for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 40(6): 1314-1325.
- 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.
- Cochrane, M.A., 2000. Using vegetation reflectance variability for species level classification of hyperspectral data, *International Journal of Remote Sensing*, 21(10): 2075-2087.
- Daubechies, I., 1998. Orthonormal Bases of Compactly Supported Wavelet, *Communications on Pure and Applied Mathematics*, 41(7): 906-966.
- Fauvel, M., J. Chanussot, and J.A. Benediktsson, 2009. Kernel principal component analysis for the classification of hyperspectral remote-sensing data over urban areas, *EURASIP Journal on Advances in Signal Processing*, 2009: 1-14.
- Gu, Y., Y. Liu, and Y. Zhangm, 2008. A selective KPCA algorithm based on high-order statistics for anomaly detection in hyperspectral imagery, *IEEE Geoscience and Remote Sensing Letters*, 5(1): 43-47.
- Igamberdiev, R.M., G. Renzdoerffer, R. Bill, H. Schubert, M. Bachmann, and B. Lennartz, 2011. Determination of chlorophyll content of small water bodies (kettle holes) using hyperspectral airborne data, *International Journal of Applied Earth Observation and Geoinformation*, 13: 912-921.
- Kurz, T.H., J. Dewit, S.J. Buckley, J.B. Thurmond, and D.W. Hunt, 2012. Hyperspectral image analysis of different carbonate lithologies (limestone, karst and hydrothermal dolomites): the Pozalagua Quarry case study (Cantabria, North-west Spain), *Sedimentology*, 59(2): 623-645.
- Kwon, H. and N.M. Nasrabadi, 2005. Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 43(2): 388-397.
- Muraki, S., 1993. Volume data and wavelet transforms, *IEEE Computer Graphics and Applications*, 13(4): 50-56.
- Park, N.-W., H.Y., Yoo, J-I., Shin, and K.-S., Lee, 2012. Anomaly Detection from Hyperspectral Imagery using Transform-based Feature Selection and Local Spatial Auto-correlation Index, *Korean Journal of Remote Sensing*, 28(4): 357-367.
- Reed, I.S. and X. Yu, 1990. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 38(10): 1760-1770.
- Rodarmel, C. and J. Shan, 2002. Principal Component Analysis for Hyperspectral Image Classification, *Surveying and Land Information Systems*, 62(2): 115-123.
- Yoo, H.Y., K. Lee, and B.D. Kwon, 2007. Application of the 3D Discrete Wavelet Transformation Scheme to Remotely Sensed Image Classification, *Korean Journal of Remote Sensing*, 23(5): 355-363.