MULTI-TEMPORAL SAR AND OPTICAL DATA FUSION FOR LAND COVER CLASSIFICATION BASED ON THE BAYESIAN THEORY

Kamolratn Chureesampant[1] and Junichi Susaki*[2]
[1]Graduate student, Department of Urban and Environmental Engineering, Kyoto University
C1-1-209, Kyoto-Daigaku Katsura, Nishikyo-ku, Kyoto, 615-8530 Japan; Tel: + 81-7-53833302;
E-mail: kamolratn.c@ft5.ecs.kyoto-u.ac.jp

[2]Associate Professor, Department of Civil and Earth Resources Engineering, Kyoto University
C1-1-206, Kyoto-Daigaku Katsura, Nishikyo-ku, Kyoto, 615-8530, Japan; Tel: + 81-7-53833300;
E-mail: susaki.junichi.3r@kyoto-u.ac.jp

KEY WORDS: Land Cover Classification, Multi-temporal Synthetic Aperture Radar (SAR) Images, Texture, Data Fusion, Bayesian Classification Theory

ABSTRACT: This paper addresses the land cover classification capabilities of multi-temporal synthetic aperture radar (SAR) data and optical data fusion based on Bayesian approach. Multi-temporal SAR data were used to extract average backscattering coefficient, backscatter temporal variability and long-term coherence while the reflectance values were calculated using the optical data. Grey Level Co-occurrence Matrix (GLCM) based texture measure including mean, standard deviation, correlation, contrast, homogeneity, dissimilarity, and entropy was used to parameterize texture in the image. These features are integrated in the Bayesian approach. Three processing steps for the classification were used in this study: 1) Information fission by feature extraction. 2) Supervised classification with grayscale value of information fission. Then, the maximum a posteriori (MAP) estimation can be used to label each class. 3) The combination of logical operators was applied to compute the final combined Bayesian membership value function. Finally, the classification results were generated taking Osaka city of Japan as the study area. In the experiment, fourteen ALOS/PALSAR level 1.1, single-polarization data, and ALOS/AVNIR-2 level 1B2G data were used. The major classes were selected to be built-up areas, fields, forests, and water bodies. It was found in SAR and optical images that mean images produced the best result among the texture measures because of the smoothing effect for image. Moreover, the correlation results of texture measurements with mean texture showed that using highly correlated textures can have lower accuracies. In addition to this, a slight increase in accuracy was found when using multi-temporal SAR and optical data fusion combining all texture images. The study shows that combination of various textures can improve over single-set texture, but the low correlation between textures should be considered.

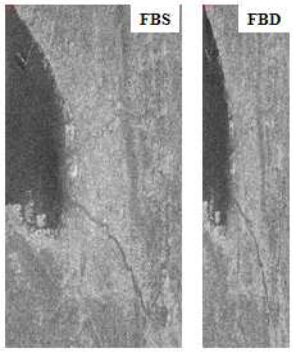
1. INTRODUCTION

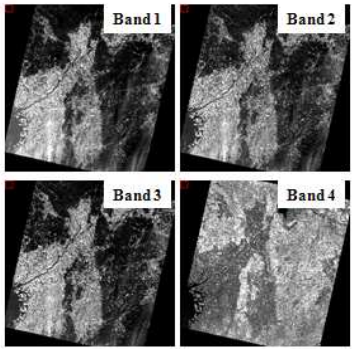
Earth-observing satellites currently operate in the visible, infrared, and microwave regions of the spectrum. Hence, the remote sensed imagery was produced by the active microwave and passive optical sensors. In term of classification, an optical image is easy for interpretation, but on the other hand SAR images with the added advantageous of its all weather, day and night operational capabilities and the penetration possibilities over cloud, haze and rainfall, is difficult to interpret. Their significant is based on the properties of signals. SAR signals contains specific information such as the sensitivity of the backscattering coefficient to target geometry and permittivity; and the coherence of the electromagnetic pulse that permits interferometry. The spectral reflectance is the specific information contained in optical sensor signal. However, Bruzzone *et al*. [1] stated that, there are some limitations of using SAR and optical data for classification such as the presence of speckle noise and geometric features (i.e., foreshortening, layover and radar shadow) while the optical sensors fails due to the unavailability of cloud-free data. With the above causes, it is difficult to obtain high classification accuracy if used only a single source image or a single polarization SAR image. To mitigate these problems, the typical operational frameworks are considered as follows: 1) use of fully polarimetric images; 2) data fusion between SAR data and multispectral images; and 3) using series of multi-temporal SAR images acquired on the same geographic area. Though the fully polarimetric data may increase the separability of land-cover classes, but there is limitation of remote sensing system to provide for all the region of interest area. The accuracy of the SAR and optical image fusion depends on the complementary information provided by active microwave and passive optical sensors. In recent years, the use of multi-temporal SAR data has increased for the production of land-cover maps by analyzing the temporal behavior of the backscattering coefficient. For these reasons, the integration of multi-temporal SAR images and optical data can be considered to become an operational tool for land cover classification. In this framework texture can be considered as a significant parameter to identify land cover classes in the image.

Tso *et al.* [4] stated that data fusion is a formal framework for the bringing together of imagery originating from different sources, viewing the same scene. In term of classification, the assumption is made that the classification accuracy should improve if additional features to be incorporated. Mainly because the greater the amount of relevant information that is included, the higher the probability that interclass confusion will be reduced. Thus, it can be foreseen that the development of multisource classification methodologies will become increasingly important. Generally, standard features associated with the single-date SAR or optical signals are commonly exploited by using maximum-likelihood classifiers. But when nonlinear features extracted from multi-temporal images or fused images from different sources are involved, Bayesian decision theory is considered as the classification methodology because it is the fundamental statistical approach to the problem of pattern recognition. The main objective of this paper is to investigate the land cover classification capabilities of individual multi-temporal synthetic aperture radar (SAR) and optical images, and there combination with the GLCM based texture measures in a Bayesian framework. The paper is organized in seven sections. Section 2 introduces the characteristics of the study areas. Section 3 and 4 describes the pre-processing and processing techniques. The classification method based on the Bayesian theory is presented in Section 5. The experimental results are reported and discussed in Section 6. The paper is concluded in Section 7.

2. STUDY AREA AND DATA DESCRIPTION

The study area was selected in Osaka city, Japan. Osaka city area is completely surrounded by more than ten cities. With the development of this proposed framework was focused, using four major land cover classes consisting human settlements, forest, fields and water. The multi-temporal ALOS/PALSAR fine beam HH-polarization (L-band) data and an optical ALOS/AVNIR-2 data with four spectral bands were used as shown in Figure 1(a) and (b), respectively.





(a) Original HH polarization ALOS/PALSAR images. (b) Original four bands of ALOS/AVNIR-2 images.

Figure 1. ALOS satellite images of Osaka city, Japan.

The fourteen PALSAR images at level 1.1 with ascending orbit of observation and 34.3º off-nadir angles, acquired on 8 October 2006 to 16 July 2009 were used. The pixels size are 9,344 x 18,432 (FBS) and 4,640 x 18,432 (FBD). The range and azimuth pixel size (m) are 4.68 x 6.35 (FBS) and 9.37 x 12.7 (FBD). An optical AVNIR-2 image, at level 1B2G (geometrically corrected data) with descending orbit of observation image was acquired on 18 May 2009. The size is 8,491 x 8,390 pixels with 10m spatial resolution. The sun angle elevation and azimuth were 69.39º and 133.58º, respectively. Moreover, the DEM 90m resolution was also used to georeference the SAR data.

3. PRE-PROCESSING TECHNIQUES

The PALSAR images were orthorectified by using a digital elevation model to correct foreshortening. Before performing fusion activities, all images have to be preprocessed to obtain the consistent image size and geolocation. In this paper, SAR images were resampled at the AVNIR-2 image pixel size (10m.), and were coregistered to UTM projection, zone 53 north, and WGS-84 geodetic datum. Besides, all images were clipped to 5,493 x 6,292 pixels.

4. PROCESSING TECHNIQUES

The used features were extracted based on the scattering properties of multi-temporal SAR data from the study of Bruzzone *et al.* [1]. The spectral reflectance [2] was used as the optical signal. In this paper, the Gray Level Co-occurrence Matrix (GLCM) is applied to measure texture. GLCM is a tabulation of how often different combinations of pixel gray levels occur in an image. The GLCM texture measures including mean, standard deviation, correlation, contrast, homogeneity, dissimilarity, and entropy were analyzed. The feature extraction approaches of multi-temporal SAR images are described as follows:

4.1 Average Backscattering Coefficient

The backscattering coefficient can be derived from the complex SAR data. With the multi-temporal images used, an appropriate noise reduction processing should be considered. Therefore, the multi-temporal filtering approach of *Quegan et al.* [3] was applied. Then, the filtered average backscattering coefficient is defined as:

$$\sigma^0_{fil,ave} = \sum_{i=1}^{M} \frac{\sigma^0_{fil,i}}{M} \quad \text{where } \sigma^0_{fil,i} = \frac{\sigma_{forst,i}}{M} \sum_{i=1}^{M} \frac{I_i}{\sigma_i} \tag{1}$$

where $\sigma_{fil,i}$ is the filtered output for the i^{th} input image. σ_i is the backscatter value. I_i is the intensity values. M is the number of images. σ_{forst} is the forst adaptive filtering of backscattering coefficient image i.

4.2 Backscattering Temporal Variability

In this experiment, the "Standard Deviation" was used because of its capabilities in identifying the spread of the backscattering coefficient. It is effective and its values are more meaningful and easier to understand. Therefore, the filtered standard deviation backscattering temporal variabiity is defined as:

$$stdev = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \sigma_i^2 - \sigma_{ave}^2} \ . \tag{2}$$

4.3 Long-term Coherence

Coherence is defined as the absolute value of the normalized complex correlation coefficient. The long-term L-band coherence is derived from tested multi-temporal images and can be computed as follows. 1) Select several pairs of base and slave images; 2) Generate the coherence images; and 3) Average the images.

5. CLASSIFICATION METHOD BASED ON THE BAYESIAN THEORY

5.1 Data Fusion Classification

There are three processing steps of data fusion classification method based on Bayesian theory as shown in Figure 2. The data sources used are the multi-temporal SAR data and an optical data. It is assumed that these data are preprocessed as described in Section 3. The data fusion classification method is delineated as follows.

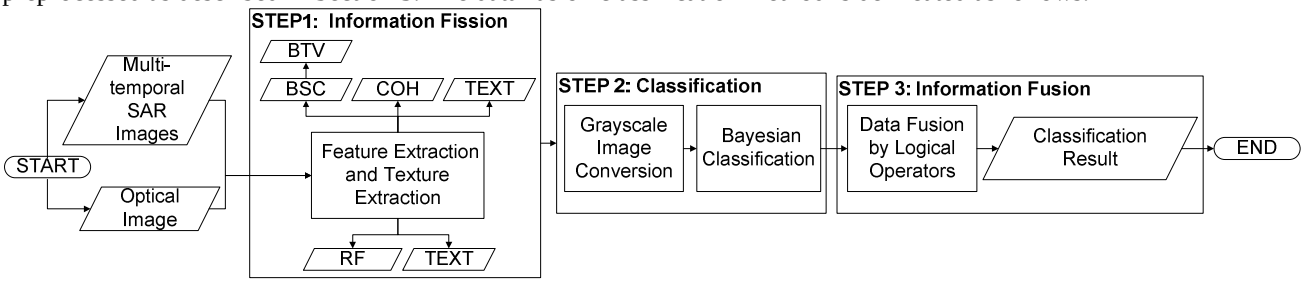


Figure 2. Classification method of multi-temporal SAR and an optical data fusion based on Bayesian theory.

5.1.1 Information Fission

The information fission performed by feature extraction from each source or set of sources as described in Section 4. For multi-temporal SAR data, average backscattering coefficient (BSC), backscatter temporal variability (BTV) and long-term coherence (COH) are extracted. Spectral reflectance (RF) is calculated from optical data. Moreover, eight GLCM texture measures (TEXT) are also used. These features can describe various properties of the observed scene.

5.1.2 Classification

The second step is supervised classification with grayscale value of information fission using Bayesian theory. Each extracted feature is converted to the grayscale to reduce the complexity of the data before implementing with Bayesian classification. The user interactively provides the training set of each class with accordance to the rule described in Section 5.2. Then, Bayesian theory is performed. This theory allows analyzing uniformly the uncertainties in the data, which were acquired from heterogeneous sources and non-commensurable features, also it easily incorporate additional prior information. In the general multisource case, a set of observations from n, $n>1$ different classes is used to implement the relations between classes. Let x_i, $i \in [1,n]$ denote the measure of a specific

pixel from measurement class *i*. These relations are described by the probabilities. As a result, the posteriori probabilities are obtained for each of *c* information class ω_j, $j \in [1,c]$. These probabilities are used further in the classification that is assigning the label with the highest probability to each pixel of the data. The Bayesian probability theory states that:

$$P(\omega_j|X) = \frac{P(\omega_j) \times P(X|\omega_j)}{P(X)}, X = x_1, x_2, ..., x_n \text{ where } P(X) = \sum_{j=1}^{n} P(\omega_j) \times P(X|\omega_j) \quad (3)$$

$$P(\omega_j|x_1, x_2, ..., x_n) = \frac{P(\omega_j) \times P(x_1|\omega_j) \times P(x_2|\omega_j) \times ... \times P(x_n|\omega_j)}{P(x_1, x_2, ..., x_n)} \quad (4)$$

where $P(\omega_j|x_1,x_2,\ldots x_n)$ is known as the conditional or posterior probability that ω_j is the correct class, given the observed data vector $(x_1,x_2,\ldots x_n)$. $P(x_1,x_2,\ldots x_n|\omega_j)$ is the likelihood probability. The prior probably of class ω_j is denoted by $P(\omega_j)$.

5.1.3 Information Fusion

In the third step, the AND and OR logical operators are applied to compute the final combined Bayesian membership value function. In case of optical fusion (or only reflectance combination of four bands); or multi-temporal SAR data fusion (or only three SAR signals combination), OR operator is used. Since the texture measures are fused, AND operator is used together with OR operator to join with original reflectance combination of four bands or combination of three SAR signals. Finally, the classification result is generated.

5.2 Selection of Training Set and Test Set

By analyzing the data for each class, training set was selected based on proportionate sampling method in order to train the Bayesian network, for larger classes more sample pixels were selected. To access the classification accuracy, the test set was selected from different set of regions of interest (ROIs) of the samples of training set. In this study, the selection of training and test sets were based on SPOT-5 satellite imagery as shown in Table 1.

Table 1. Number of training and test set used in the experiments.

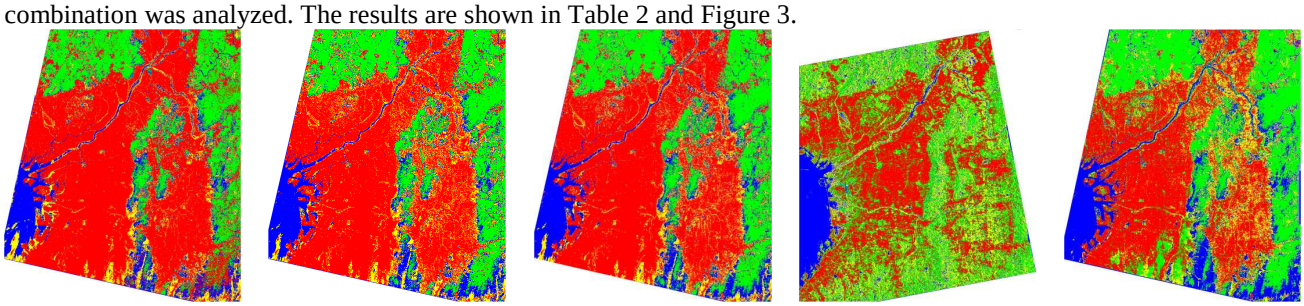
Land cover class	Water bodies(Blue)	Forests(Green)	Fields(Yellow)	Human Settlements(Red)	Total Pixels
Training set (pixels)	2,893	3,779	3,325	6,259	16,256
Test set (pixels)	28,108	37,181	32,857	60,660	158,806

5.3 Post Classification Smoothing and Accuracy Analysis

Since the raw classified image appears noisy due to the isolated pixels spread in more homogeneous regions. Such pixels can be removed by making use of majority filter. In this study, median filter (5x5 window size) was applied to the classification results. To assess the classification performance, an overall accuracy and kappa coefficient were analyzed.

6. EXPERIMENTAL RESULTS AND DISCUSSION

The data fusion experiments were classified into three major cases consist optical data fusion (case 1); multi-temporal SAR data fusion (case 2); and multi-temporal SAR data and optical data fusion (case 3). In each case, the textures combination was analyzed. The results are shown in Table 2 and Figure 3.



(a) Reflectance. (b) Reflectance + Mean. (c) Reflectance + All Texture. (d) SAR signals + All Texture. (e) Reflectance and SAR signals + All Texture.

Figure 3. Classification results of optical data fusion [(a), (b), and (c)]; multi-temporal SAR data fusion (d); and multi-temporal SAR data and optical data fusion; (e) with texture measures.

Table 2. Overall classification accuracies and kappa coefficient of accuracy exhibited on the test set.

Data Fusion Cases		Number of files	Before post classification			After post classification	
			Overall		Kappa	Overall	Kappa
1. Optical	1.1	Reflectance	4	69.43%	0.5545	70.16%	0.5646
	1.2	Mean	4	75.47%	0.6465	75.77%	0.6465
	1.3	Reflectance + Mean	8	76.17%	0.6567	76.56%	0.6609
	1.3.1	Reflectance + Mean + Variance	12	74.58%	0.6327	75.70%	0.6482
	1.3.2	Reflectance + Mean + Correlation	12	73.86%	0.6207	73.81%	0.6192
	1.3.3	Reflectance + Mean + Homogeneity	12	76.12%	0.6548	75.98%	0.6521
	1.3.4	Reflectance + Mean + Contrast	12	75.72%	0.6495	76.09%	0.6538
	1.3.5	Reflectance + Mean + Dissimilarity	12	76.07%	0.6541	76.05%	0.6533
	1.3.6	Reflectance + Mean + Entropy	12	74.55%	0.6307	74.52%	0.6295
	1.4	Reflectance + Statistic Group	16	76.09%	0.6544	76.10%	0.6538
	1.5	Reflectance + Statistic + Contrast Groups	28	76.19%	0.6565	76.29%	0.6568
	1.6	Reflectance + All Texture	32	76.26%	0.6576	76.42%	0.6588
2. Multi-temporal SAR	2.1	SAR signals	3	75.87%	0.666	82.03%	0.7516
	2.2	SAR signals + Mean	6	77.20%	0.6824	82.54%	0.7575
	2.3	SAR signals + All Texture	24	77.24%	0.6829	82.63%	0.7587
3. Multi-temporal SAR and Optical	3.1	Reflectance and SAR signals	7	79.34%	0.7175	84.92%	0.7924
	3.2	Reflectance + Mean and SAR signals + Mean	14	81.73%	0.7578	86.29%	0.81
	3.3	Reflectance + All Texture and SAR signals + All Texture	56	82.75%	0.7632	86.83%	0.8149

Table 3. The correlation coefficient between mean texture and other texture images of optical data

Correlation coefficient	Band1	Band2	Band3	Band4	Average of 4 bands	Overall accuracy
Mean + Variance	0.92	0.88	0.96	0.88	0.91	73.96%
Mean + Correlation	0.33	0.33	0.34	0.39	0.35	72.92%
Mean + Homogeneity	0.22	0.21	0.23	0.20	0.22	75.92%
Mean + Contrast	0.89	0.87	0.94	0.89	0.90	73.89%
Mean + Dissimilarity	0.84	0.89	0.88	0.87	0.87	73.38%
Mean + Entropy	0.41	0.38	0.39	0.37	0.39	74.00%

Consider the results obtained from optical data fusion in Table 2 (case1) with Figure 3 (a, b, and c). Figure 3(a) presented the classification of reflectance images for all four bands combination. From visual interpretation of the result reveal that the forest class was misclassified as the water class in the bottom-center part of the image due to cloud cover from optical data. It also showed the lowest overall accuracy (69.43%) as shown in case1.1. In addition, the texture parameters were analyzed as well. Particularly, the mean texture of reflectance combining all four bands in case1.2 showed 8% improvement over original reflectance combination. This improved classification result is present in Figure 3(b). The classified forest class was corrected from misclassified urban class in the bottom-right of the image. With this result, it was found that mean texture is a significant parameter for classification. Therefore, the combination of the original reflectance and mean texture was also investigated as shown in case1.3. However, the accuracy increased slightly in the range of 0.92%.

In addition to this, other seven textures measures were analyzed. The data fusion between reflectance and mean texture with all the other textures measures of all four bands were experimented respectively, from 1.3.1 to 1.3.6 in Table 2. In all of these cases, lower accuracies were generated. It can be seen that the reliability or uncertainty of each data source should be considered before the classification of multisource data is applied. In this study, the correlation coefficient is used as the indicator for measuring the fusion quality. From the result in Table 3, the lowest correlation between the mean and homogeneity textures generated the highest overall accuracy of classification. It was found that the correlation coefficient closer to zero was preferred because the highly correlated textures show no variation between the classes of interest to be classified. This could be a drawback to the achievement of higher classification accuracy. Among texture measures, mean texture performed better than others because of the smooth effect of the image. Moreover, the data fusion between reflectance and texture statistic group (including mean, variance and correlation) in Table 2 case1.4; the data fusion between reflectance with texture statistic group and contrast group (including homogeneity, contrast and dissimilarity) in case1.5; and the data fusion between reflectance and all textures in case 1.6 with Figure 3(c) were tested and found that their accuracies were slight increased.

The classification results obtained from multi-temporal SAR data fusion in Table 2, case 2 and Figure 3(d) are discussed as follows. Table 2, case 2.1 shows classification result of the combination of SAR signals including (average backscattering coefficient, backscattering temporal variability and long-term coherence). The overall accuracy was 75.87%. From this result it can be found that the integration of a reliable feature extraction approaches

based on the physics of multi-temporal SAR signals can solve a multi class problem with multi-temporal SAR data. According to the experiment of SAR signals with mean texture of each signal combination as shown in case 2.2, the results show 1.72% improvement than the case2.1. With this result, mean texture can be used to improve the classification accuracy for SAR data, mainly the mean texture can reduce speckle noise most effectively. From the comparison of the improvement percentage of mean texture combination between optical and multi-temporal SAR data, it was found that the changing range for multi-temporal SAR data is shorter. The reason is the use of multi-temporal filtering as described in Section 4.1, the noise of SAR data can be reduced. In case2.3 with Figure 3(d), the data fusion of SAR signals and all texture of each signals, outperformed with the 0.05% increased accuracy (overall accuracy: 77.24%).

The classification results obtained from multi-temporal SAR data and optical data fusion in Table 2, case 3 and Figure 3(e) are discussed as follows. Table 2, case 3.1 shows classification result of the combination of SAR signals and reflectance of four bands. The overall accuracy was 79.34%. The experiment of data fusion with mean texture of each signal combination is shown in case 3.2. It represented 81.73%% overall accuracy, which was 2.92% improvement from case3.1. Finally in case3.3 with Figure 3(e), the data fusion of SAR signals and reflectance with all texture of each signals, outperformed the case3.2 with 1.23% increased accuracy (overall accuracy: 82.75%). The post classification filtering can improve the capabilities of classified results and accuracies. In Table 2, the results of all the cases after using post-classification smoothing yielded the better accuracies, increasing approximately 0.6% for optical data fusion, 6.8% for multi-temporal SAR data fusion and 5.5% for multi-temporal SAR data and optical data fusion respectively. In addition, it should be noted that the amount of smoothing depends on the size of the window.

7. CONCLUSION

The classification method for multi-temporal SAR data and optical data fusion based on Bayesian theory has been proposed based on three processing steps including: 1) information fission by extracting the features on analyzing the physical properties of the multi-temporal SAR signals, reflectance data from optical data, and GLCM texture measures; 2) supervised classification with grayscale value of information fission based on Bayesian theory. This classifier learns the relationships between the classes and the maximum a posteriori (MAP) estimation assign the classes; 3) final combined Bayesian membership value function for each class was computed using combination of logical operators. Thus, the classified results were generated. The water, forest, fields and human settlements classes were studied. From the analysis of all the experimental results, it can be concluded that the feature integration of SAR signals including the average backscattering coefficient, backscattering temporal variability and long-term coherence are effective in modeling the series of multi-temporal SAR data. However, the classification using single source is still insufficient to carry out the high accuracy and good quality, due to the sensor based drawbacks such as the cloud cover of optical data; or the relief displacement and the speckle noise from SAR data. These effects should be accounted prior to the multisource classification procedures. GLCM texture measures were also analyzed. The results demonstrated that, using mean texture combination significantly improved the classification accuracies due to the influence of the smoothening effect. In addition, the study of correlation between mean and other textures shows that the use of highly correlated textures can lower the accuracy. However, the combination of various texture measures showed better results over single-set texture measure, because of their different and complementary information. The result suggests that the multi-temporal SAR and optical data fusion combining all texture images was slightly accurate than in the case of combining only with the mean texture. Post-classification filtering process can improve classification result considerably. Eventually, the classification using multisource data fusion outperformed single source fusion. This is mainly because appropriate extracted features can increased the capacities of the results and can reduce the complexity of functions in classification method. For the future development of this work, more different class combinations can be incorporated to see their discrimination in the classification. As a final remark, this classification method worth to be an operational tool for classification problems involving the land covers classes.

ACKNOWLEDGEMENT

This research was supported by a program of the 3rd ALOS Research Announcement.

REFERENCES

[1] Bruzzone, L., Marconcini, M., Wegmuller, U., and Wiesmann, A., 2004. An advanced system for the automatic classification of multi-temporal SAR images. IEEE Transactions on Geoscience and Remote Sensing, 42 (6), pp. 1321-1334.

[2] Chander, G., and Markham, B., 2003. Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. IEEE Transactions on Geoscience and Remote Sensing, 41 (11), pp. 2674-2677.

[3] Quegan, S., and Yu, J. J., 2001. Filtering of multichannel SAR images. IEEE Transactions on Geoscience and Remote Sensing, 39 (11), pp. 2373-2379.

[4] Tso, B., and Mather, P. M., 2001. Classification Methods for Remotely Sensed Data. 2nd Ed Taylor and Francis Group, London, pp. 290-297.