Urban Area Extraction Using X-band Fully Polarimetric SAR Imagery

Junichi Susaki, *Member, IEEE*, and Masaaki Kishimoto, *Non-Member, IEEE*

***Abstract*—Here we present a method for extracting urban areas from X-band fully polarimetric synthetic aperture radar (SAR) data by reducing the effects of polarization orientation angle (POA). The proposed classifier performs two classifications utilizing characteristics of X-band scattering in land cover. One classification uses total power of scattering and volume scattering derived by using four-component decomposition methods with correction for the POA effect. The other classification uses polarimetric coherence between S_{HH} and S_{VV}. The two results are intersected and final urban areas are extracted after post-classification processing. We applied the proposed method to airborne X-band fully polarimetric SAR data of the Polarimetric and Interferometric Airborne Synthetic Aperture Radar System, developed by the National Institute of Information and Communications Technology, Japan. Validation of the results for three Japanese urban areas shows that the proposed method provides an acceptable overall accuracy of approximately 80–90% at a 100-m spatial scale. It is also shown that texture-based classifiers using single polarimetric data have accuracy limitations when applied to extracted urban areas where the POA of objects is not uniform.**

***Index Terms*— Urban areas, X-band fully polarimetric SAR, Polarization orientation angle effect, Deorientation.**

I. Introduction

Urban monitoring is an important application of satellite and airborne remote sensing. Both urban area mapping for urban planning and continual map updates are required. Whereas optical sensors are capable of delineating urban areas, optical images are vulnerable to cloudy conditions. Synthetic aperture radar (SAR) has potential for monitoring urban areas, because man-made structures return relatively strong backscattering that distinguishes them from other objects. SAR is also much less sensitive to cloud contamination than optical sensors. However, backscattering detected by SAR is sensitive to the orientation angle (OA) of man-made structures. This OA effect is critical in analyzing urban areas from SAR images. Land cover classification without considering OA effects may lead to parts of urban areas being omitted due to fluctuating backscattering. Reduction of OA effects through "deorientation" has thus been investigated over the past decade.

Xu and Jin [1] proposed deorientation theory, in which the parameters for the deoriented state are derived by rotating target scattering vectors and minimizing cross-polarization. This idea has been applied in other research, such as Kimura [2], An et al. [3] and Yamaguchi et al. [4]. Hereafter, we call the OA effect the polarization orientation angle (POA). As An et al. [3] pointed out, Yamaguchi decomposition [5,6] suffers from the problem that volume scattering may be overestimated due to POA effects, soon after which Yamaguchi et al. [4] presented an improved four-component decomposition method. The methods presented by both parties essentially follow the work of Xu and Jin [1] from the viewpoint that the coherence matrix should be rotated. Lee and Ainsworth [7] reported the POA effect on the coherency matrix and that neither the Freeman–Durden three-component decomposition [8] nor the Yamaguchi decomposition utilizes the full polarimetric information.

Following these works, several improvements were achieved via analysis. Chen et al. [9] showed that the POA effect in scattering for large POA may not be corrected even after deorientation. This is because model-based decomposition assumes that only volume scattering contributes to the cross-polarization term. We also reported the results of experiments performed in an anechoic room, which showed that the total powers of backscattering and the four components derived from fully polarimetric scattering are highly dependent on the orientation angles [10]. With these findings, we have already reported a method for extracting urban areas by using Advanced Land Observing Satellite (ALOS) / Phased Array type L-band Synthetic Aperture Rader (PALSAR) imagery [10] and another method for estimating urban densities by using a single fully polarimetric SAR (PolSAR) image [11,12].

For detailed urban monitoring, very high resolution (VHR) SAR has been examined because it has potential for mapping urban areas on a per-district or per-building basis. Esch et al. [13] proposed a method for extracting human settlements by using backscatter amplitude and speckle divergence obtained from single-polarization TerraSAR-X and TanDEM-X data. Ferro et al. [14] studied the relationship between the double-bounce effect of buildings and the orientation angle in SAR images. Most of these studies assume single polarimetric SAR images, and pay little attention to the above-mentioned POA effects. Therefore, these methods are not guaranteed to effectively extract urban areas, even those with heterogeneous POA distributions. Therefore, in this paper we propose an urban-area classifier that is applicable to X-band PolSAR images by extending our previous work [10].

The remainder of this paper is organized as follows: Section II explains the indices used in the proposed algorithm. The location and data collected from the site are described in Section III. Section IV explains the proposed method and reports the experimental results. The implications of these results and the validity of the algorithm are then discussed in Section V. Finally, Section VI concludes the paper.

Junichi Susaki, *Member, IEEE*, and Masaaki Kishimoto, *Non-Member, IEEE*

II. INDICES USED

A. Coherency Matrix

PolSAR data consist of complex scattering values, which can be represented by the 2 × 2 scattering matrix shown in (1):

$$s = \begin{pmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{pmatrix} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}. \quad (1)$$

For mono-static radar imaging of a reciprocal medium, we have $S_{HV} = S_{VH}$. In a PolSAR image, each pixel is represented by a 3 × 3 coherency matrix

$$T = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} |a+b|^2 & (a+b)(a-b)^* & 2(a+b)c^* \\ (a-b)(a+b)^* & |a-b|^2 & 2(a-b)c^* \\ 2c(a+b)^* & 2c(a-b)^* & 4|c|^2 \end{pmatrix}. \quad (2)$$

where * denotes the complex conjugate operation.

B. Four-component Decomposition with Rotation of Coherency Matrix

From the coherency matrix, four-component decomposition decomposes the observed backscattering into the surface scattering power (P_s), the double-bounce scattering power (P_d), the volume scattering power (P_v), and the helix scattering power (P_c) [6]. The coherency matrix is decomposed into four components as follows:

$$\langle T \rangle = f_s \langle T_s \rangle + f_d \langle T_d \rangle + f_v \langle T_v \rangle + f_c \langle T_c \rangle. \quad (3)$$

Here, f_s, f_d, f_v, and f_c respectively denote coefficients of the surface, double-bounce, volume, and helix scatterings, and T_s, T_d, T_v, and T_c respectively denote coherency matrices of the surface, double-bounce, volume, and helix scatterings. The four-component decomposition method calculates the four coefficients, and thus the contribution of each scattering is determined.

Yamaguchi et al. [7] proposed an algorithm that rotates the coherency matrix by the polarization orientation angle (POA) to reduce the dependence of the components on the relative azimuth. The rotated coherency matrix $T(\theta)$ is derived by rotating the original coherency matrix in (2) by POA θ:

$$T(\theta) = \begin{pmatrix} T_{11}(\theta) & T_{12}(\theta) & T_{13}(\theta) \\ T_{21}(\theta) & T_{22}(\theta) & T_{23}(\theta) \\ T_{31}(\theta) & T_{32}(\theta) & T_{33}(\theta) \end{pmatrix} = [R_p(\theta)][T][R_p(\theta)]^{\dagger}. \quad (4)$$

Here, † denotes complex conjugation and transposition, and $R_p(\theta)$ is the rotation matrix given by

$$[R_p(\theta)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{pmatrix}. \quad (5)$$

The POA is reported to estimate the orientation angle of the target [1,2]. θ is obtained by minimizing $T_{33}(\theta)$ in (3) and estimated as

$$\theta = \frac{1}{4}\tan^{-1}\frac{2\,\mathrm{Re}(T_{23})}{T_{22} - T_{33}}, \left(-\frac{\pi}{4} \le \theta \le \frac{\pi}{4}\right). \quad (6)$$

The θ is determined by requiring that the wall normal be parallel to the sensor's ground range direction.

C. Magnitude balance of HH and VV

In four-component decomposition, Yamaguchi et al. [5] employed the magnitude balance of $\langle |S_{HH}|^2 \rangle$ and $\langle |S_{VV}|^2 \rangle$ for classifying volume scattering type:

$$10 \log\left[\frac{\langle |S_{VV}|^2 \rangle}{\langle |S_{HH}|^2 \rangle}\right]. \quad (7)$$

The larger (7) is, the more vertical dipoles are dominant in the pixel. After the coherency rotation, (7) is modified as (8) [4]:

$$10 \log\left[\frac{T_{11}(\theta) + T_{22}(\theta) - 2\,\mathrm{Re}(T_{12}(\theta))}{T_{11}(\theta) + T_{22}(\theta) + 2\,\mathrm{Re}(T_{12}(\theta))}\right]. \quad (8)$$

D. Polarimetric Coherence

Polarimetric coherence between two polarimetric data is a coefficient of scattering values of the two polarimetric data in a certain window. We can calculate polarimetric coherence of any combination of different polarimetric data. For example, the polarimetric coherence between S_{HH} and S_{VV} can be expressed in (9) as

$$\gamma_{HHVV} = \frac{|\langle S_{HH} S_{VV}^* \rangle|}{\sqrt{\langle S_{HH} S_{HH}^* \rangle \langle S_{VV} S_{VV}^* \rangle}}. \quad (9)$$

III. DATA USED AND STUDY AREA

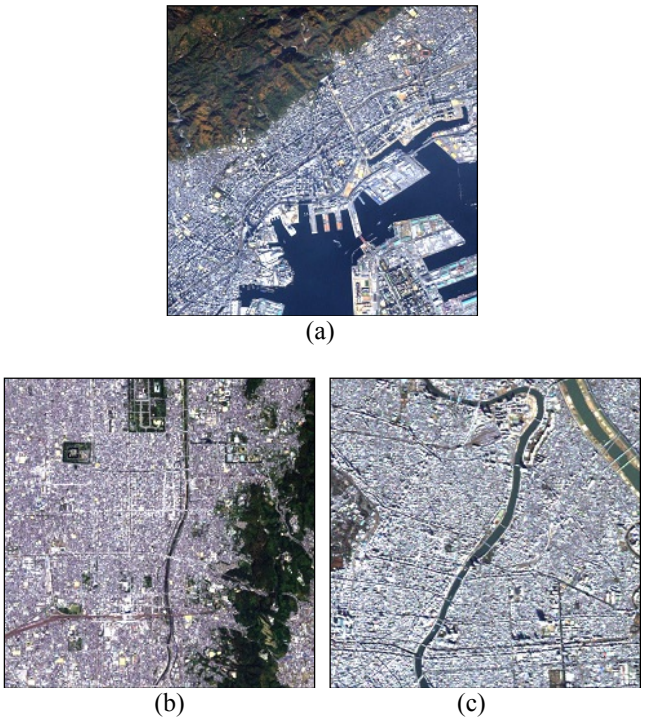
In this research, we used airborne X-band fully polarimetric SAR images. The National Institute of Information and Communications Technology (NICT) developed the X-band fully polarimetric Polarimetric and Interferometric Airborne Synthetic Aperture Radar System (Pi-SAR2). The slant range resolution has been improved to 0.3 m. The specification of Pi-SAR2 is listed in Table I [15]. The system has been used to monitor damage caused by disasters such as earthquakes [16] and volcanic eruptions [17].

We selected three observed cities as study areas—Kobe

(study area 1), Kyoto (study areas 2), and Tokyo (study area 3)—because they have several categories of land cover, including urban, mountainous, and aquatic. Fig. 1 shows Advanced Land Observing Satellite (ALOS)/Advanced Visible and Near Infrared Radiometer type 2 (AVNIR-2) optical images of the three study areas. All three study areas have a 6 × 6 km area. Pi-SAR2 images were projected to the WGS 1984 coordinate system with a universal transverse Mercator projection. Quegan's polarimetric calibration procedure [18] was used for polarimetric calibration.

TABLE I. Specification of Pi-SAR2 [15].

Center frequency	9.55 / 9.65 GHz
Bandwidth	500 / 300 / 150 MHz
Slant-range resolution	0.3 / 0.5 / 1.0 m
Azimuth resolution	0.3 (1 look) / 0.6 m (2 look)
Swath width (grand range)	> 10 km
NESZ	< -23 / -27 / -30 dB
Data rate	200 MB/s × 4 ch



(a) (b) (c)

Fig. 1. Study areas in the research. (a), (b), and (c) respectively show ALOS/AVNIR-2 images of study areas 1, 2, and 3. Each panel has an area of 6 km × 6 km.

IV. Existing Algorithms

A. Texture-based Classifier for Single Polarimetric Data

Here, we briefly introduce the work by Esch et al. [13], a texture-based classifier. First, local image heterogeneity H is calculated for each pixel as

$$H = \sigma_A / \mu_A. \tag{10}$$

Here, μ_A and σ_A are the mean and standard deviation of the amplitude image A in 9 × 9 local pixels. They assume that the relationship in (11) exists between the image heterogeneity H, the true image T, and the fading texture F:

$$H^2 = T^2F^2 + T^2 + F^2. \tag{11}$$

T^2, defined as the speckle divergence S, is derived from (11):

$$T^2 = S = (H^2 - F^2)(1 + F^2)^{-1}. \tag{12}$$

F can be calculated as $\hat{F} = 0.5233 \cdot N^{-0.5}$, where N denotes the number of looks of a SAR image.

Then, classification is implemented. They empirically set backscattering amplitude for non-urban areas $Th^A = 100$. Changing the threshold for S, the urban and non-urban candidates are generated by thresholding. The Jensen–Shannon divergence is calculated to measure the distance of two probability distributions in terms of amplitude. The optimal threshold with minimum divergence is selected.

The most serious concern is that the method is not designed to compensate for POA effects. The probability distributions of amplitude directly suffer from the amplitude decrease caused by the POA effect. Because the Jensen–Shannon divergence uses the probability distributions without any compensation, districts having heterogeneous POAs may be omitted. We examine the performance of the method in Section V.

B. Kajimoto–Susaki Classifier for Fully Polarimetric Data

The classifier proposed by Kajimoto and Susaki [10] is based on experimental results using ALOS/PALSAR images. They showed that the POA dependence of four components and total power (*TP*) decreases as a result of POA correction. After the correction, P_v is almost constant regardless of POA, whereas the other components still depend on POA. This fact was also supported by experiments conducted in an anechoic radio wave chamber in 2011 [11]. In those experiments, they measured X-band (10 GHz) fully polarimetric backscattering reflected by concrete rectangular blocks arranged at several spacings and orientation angles. As a result, P_v had the smallest difference in terms of POA effect.

The classifier performs a two-stage classification. The first classification uses POA-corrected P_v and *TP* data, because classification using only P_v cannot extract urban areas. *TP* was the power second-least affected by POA among the four components and *TP* [10]. Principal components of P_v and *TP* are generated, and pixels are categorized into five POA categories. A different classification threshold is set for each category to reduce POA effects.

In the first stage, urban areas can be discriminated from other land covers except mountains. Because P_v – *TP* scattergrams of urban and mountain areas mostly overlap, these two land

covers cannot be clearly classified by setting a threshold. Therefore, in the second classification stage POA randomness is calculated. The majority of buildings in an urban area are aligned in the same direction, so POAs in a given district will be nearly homogenous. In contrast, mountain areas are covered by natural vegetation, so POAs randomly differ, providing a good indicator for discriminating urban areas from mountains.

Note that the classifier is a kind of semi-supervised classifier requiring training data for urban areas, mountains, and farmland in the study area. The areas of training data are manually selected from an image. When classifying multiple images, optimal thresholds used in first-stage classification of the study area are automatically updated by referring to the $P_v - TP$ scattergram of the study area. The threshold used in the second-stage classification is fixed.

We confirmed that the first-stage classifier using $P_v - TP$ principal components can perform for X-band data even when we change the image pixel spacing. However, the second-stage classifier may not work for X-band data. Fig. 2 shows the POA randomness in L-band and X-band SAR images. Whereas L-band results have a clear contrast of POA randomness between urban areas and mountains, X-band results have no significant contrast. Therefore, the second-stage classifier proposed by Kajimoto and Susaki [10] should be replaced by another classifier that can effectively discriminate mountains.

C. Indices Extracting Vegetation by Using Fully Polarimetric Data

We examined the performance of extracting vegetation by using (8) to exclude mountains remaining after the classifier described in the previous subsection. We calculated (8) and set the threshold to –21.419 dB. We took samples of urban areas and forests and calculated the weighted mean of the two means, where the standard deviations were used as weights. The results show that higher values of (8) correspond to vegetation in general. However, it was found that simple thresholding of the value derived by (8) overestimates vegetation. For example, urban areas near the boundaries between urban areas and mountains, and urban areas having heterogeneous POA distributions were misclassified as vegetation. The results are shown in Section V. This finding is consistent with Chen et al. [9], which showed that model-based decomposition may overestimate volume scattering.

We next focused on (9) for extracting vegetation. The results, shown in Fig. 3, indicate that vegetation can be accurately extracted. As a result, we determined that (9) is a good indicator for extracting vegetation.

V. Method and Results

A. Proposed Method

From the previous results shown in Section IV, we propose a classification method for extracting urban areas. Fig. 4 shows a flowchart of the proposed method. We set five land cover categories: urban, forest, other vegetation, water body and

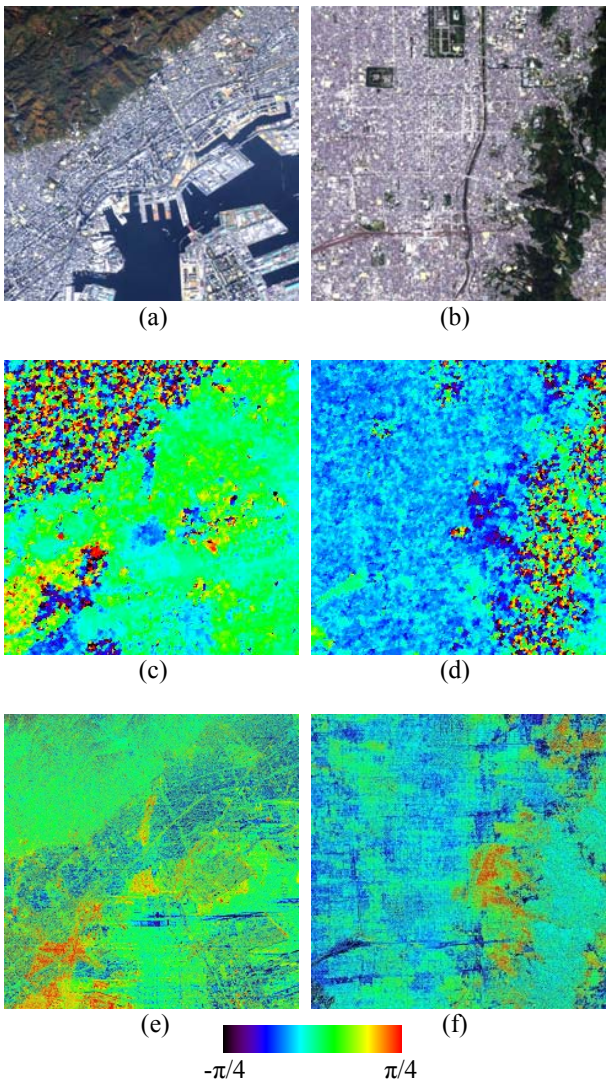


Fig. 2. Difference in POA randomness between L-band and X-band images. (a)(c)(e) and (b)(d)(f) are related to study areas 1 and 2, respectively. (a)(b) ALOS/AVNIR-2 images, (c)(d) POA derived from ALOS/PALSAR images, and (e)(f) POA derived from Pi-SAR2 images.

others. As preprocessing, multilooking is first implemented and the four components are generated with correction of the POA effect. The multilooked images had 2.5-m pixel spacing. Next, the first classification using P_v and TP principal components is implemented. As described in Subsection IV *B*, the method proposed by Kajimoto and Susaki [10] set different thresholds based on POA space. However, the dependency of Pi-SAR2 data on POA was not significant. That is, scattergrams of P_v and TP in different POA subspaces are almost same. Therefore, we applied the method without dividing POA space into subspaces.

In this experiment, the areas of interest (AOIs) of urban areas and forests were manually selected from study area 1. Each AOI has an area of 750 m × 750 m, equivalent to 300 pixel × 300 pixel. In [10], training data for urban areas and

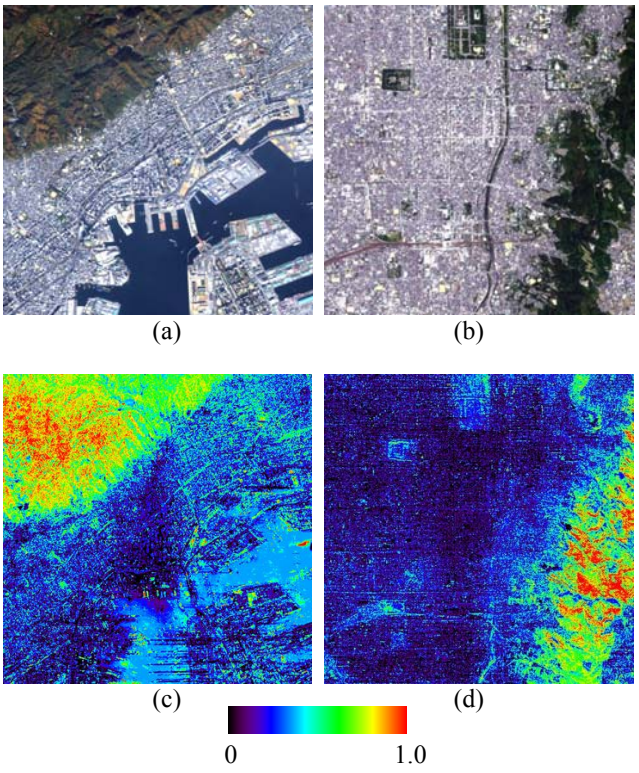


Fig. 3. Indices for extracting vegetation from X-band PolSAR images. (a)(c) and (b)(d) are related to study areas 1 and 2, respectively. (a)(b) show ALOS/AVNIR-2 images, and (c)(d) polarimetric coherence between S_{HH} and S_{VV}, expressed in (9).

farmlands were selected, but we could not find homogeneous large farmlands in these study areas. This was also the case in many other study areas. Therefore, we propose that the method requires training data for urban areas and forests. First, we generate a line for discriminating urban areas from forests in the plane of TP and P_v principal components using training data. The difference between the means of the first principal components of urban areas and forests is divided in proportion to the standard deviations, and a break point is obtained. As a result, a line is determined that is parallel to the second component axis and passes through the break point. The value of the first principal component of this break point can be regarded as a threshold for discriminating urban areas from forests.

This line is then applied to another study area to generate the line that is optimal to the new study area without training data. Tentative urban areas in the new study area are obtained by applying the line of the initial study area to the plane of TP and P_v of the new study area. The gravity point of the tentative urban areas is determined by using the values of TP and P_v. The difference is calculated between the first principal components of the urban gravity points in the new study area and in the initial study area. The threshold and line are shifted by this difference. The updated line is again applied to the new study area, and the gravity point difference is calculated. Iteration of this process is terminated when the change in the threshold is within a predefined limit. We set the limit to 0.01 dB. In short, the line is updated so that the gravity point of the urban areas in the new study area should be accordance with that of the initial study area. As a result, the line of the new study area is determined.

In this research, three lines for discriminating urban areas from forests were obtained. They can be expressed by using original TP and P_v instead of principal components of TP and P_v:

$$\begin{aligned} &\text{Study area 1: } TP = -0.902\,P_v - 69.89 \\ &\text{Study area 2: } TP = -0.902\,P_v - 42.10 \\ &\text{Study area 3: } TP = -0.902\,P_v - 54.68 \end{aligned} \tag{12}$$

Units for TP and P_v are dB. They have the same slope.

The result generated at this stage includes urban areas and part of the forests. The forests included in the result correspond to ridges that return strong backscatter. In addition, the extracted urban areas mainly correspond to isolated buildings. To generate an urban area map, but not a building distribution map, we applied closing processing (dilation and erosion). The results of this stage are shown in Figs. 5(b), 6(b), and 7(b). In the closing, a 3×3 window was applied for dilation twice, and then, the same size of window was applied for erosion twice.

The second classification uses (9) for extracting vegetation. In the experiment, the ensemble average was calculated within a 5×5 window, and the threshold for (9) was set to 0.80. This value was determined as follows: we calculated classification accuracy as a function of polarimetric coherence by using the AOIs of urban areas and forests. The accuracy was defined as $(tp - fp) / (tp + fp + tn + fn)$. Here, tp, fp, tn and fn denote true positive (pixels those were correctly classified as forests), false positive (pixels those were falsely classified as forests), true negative (pixels those were correctly classified as urban areas) and false negative (pixels those were falsely classified as urban areas), respectively. The increment of polarimetric coherence was set to 0.01. Finally, we determined the best threshold that gave the best accuracy among the interval of [0, 1].

In the same manner as classification using TP and P_v, we applied closing processing. The results are shown in Figs. 5(c), 6(c), and 7(c). After both results are intersected, urban area candidates are extracted. These intersected images, shown in Figs. 5(d), 6(d), and 7(d), show that urban areas are well extracted, but roads and radar shadow are not extracted. Therefore, we applied filtering, closing, and labeling to the urban candidate image. Filtering examined the percentage of candidate urban pixels within a window. We set the window size to 5×5, and the minimum percentage for urban areas to 20%. Labeling is implemented to remove small areas. The minimum urban area was set to 2,500 pixels, equivalent to 15,625 m^2. Figures 5(e), 6(e), and 7(e) show the urban areas finally extracted by the proposed method in the three study areas.

We applied the methods of Esch et al. [13] and Kajimoto and Susaki [10] as references. Figs. 5(f), 6(f), and 7(f) show the urban areas extracted by using the method in [13]. Figs. 5(g),

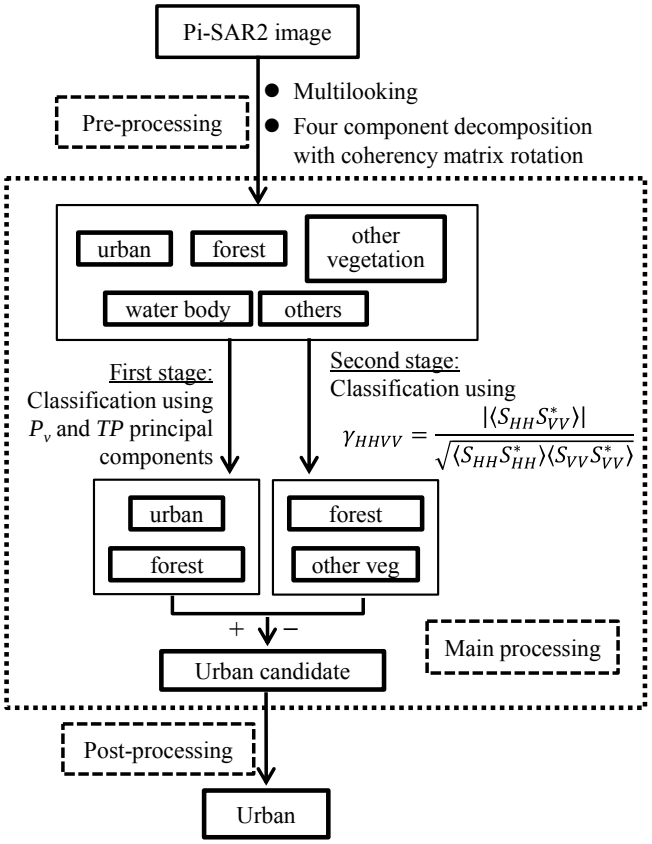


Fig. 4. Flowchart for extracting urban areas using X-band PolSAR images.

6(g), and 7(g) show the urban areas extracted by using the method in [10]. In addition, we extracted urban areas by replacing the second-stage classifier of the proposed method with (8). The results are shown in Figs. 5(h), 6(h), and 7(h).

B. Accuracy Assessment

Urban area data were obtained from Zmap-TOWN II (ZENRIN) data, which are residential maps of Japan [19]. First, building polygon data were intersected by a mesh. Buildings lying across the mesh border were divided into pieces by borderlines. The mesh size was set to 100 m × 100 m. Then, the mesh was classified into urban or non-urban categories. The minimum percentage of buildings within a mesh for classifying into urban category was set to 20%, thus generating a reference map for the urban area. Finally, the urban areas extracted from Pi-SAR2 images were compared with the reference map. Table II shows accuracy assessment results for the three study areas when using the method of Esch et al. [13] and the proposed method.

VI. Discussion

A. Comparison with Existing Method

Table II shows that accuracy of the proposed method is higher than that of Esch et al. [13] and that of Kajimoto and Susaki [10]. As shown in Fig. 5(e), (f) and (g), in study area 1, the method proposed by Esch et al. [13] and that of Kajimoto and Susaki [10] cannot remove forests in areas of layover to the extent of the proposed method. Because of off-nadir angle differences, airborne SAR often suffers from layover more than spaceborne SAR does. The method of Esch et al. [13] thus omitted hilly urban areas in study area 2, whereas the proposed method extracted them.

In terms of discriminating forests from mixed urban and forest areas, we examined the feasibility of using (8). The intent of (8) is to classify a pixel of interest into one of three scattering types: scattering where vertical dipoles are dominant, scattering where horizontal dipoles are dominant, or scattering where random dipoles are available. We found that classification using (8) can work in study area 1, where the boundaries between forests and urban areas are clear—in other words, the case where there are few areas of mixed urban and mountainous areas. In study area 2, which has mixed areas in the skirts of forests, the method overestimated vegetation areas. Thus, the final urban areas shown in Fig. 6(h) were underestimated.

In addition, Fig. 7(h) shows that the method using (8) omitted urban areas in the right part of study area 3. An optical image of that area (Fig. 8(d)) reveals that the omitted areas have small districts with different POA distributions. This defect is similar to the results when using the method of Esch et al. [13]. More districts in urban areas (Fig. 8(b)) are omitted than in the results obtained by using (8) (Fig. 8(c)). As described in Subsection IV.A, the method of Esch et al. [13] is found to be sensitive to POA effects. As a result, the proposed method is more effective than the method of Esch et al. [13] for extracting urban areas by using X-band fully polarimetric SAR data.

B. Post-classification and Spatial Resolution

In terms of mapping urban areas, classification using VHR X-band SAR images requires multiple processing steps. Figs. 5(b), 6(b), and 7(b) show clear contrast between buildings and roads. This indicates that VHR X-band images can extract buildings, but that mapping of urban areas should include other objects in urban areas. To deal with this issue, we took the approach that tentative urban areas should be extended by applying filtering and closing. Such approaches can compensate for areas where radar shadow occurs. We suggest that extraction of urban areas using VHR X-band SAR images should be accompanied by such algorithms to extract widely homogeneous areas.

The original Pi-SAR2 images have very high slant-range (0.3 m) and azimuth (0.3 m) resolutions. In this research, we applied multilooking process to the images to reduce the speckle noise and save computation time. The multilooked images had 2.5-m pixel spacing, so to examine the pixel spacing effect we classified 0.5-m pixel spacing images. We confirmed that the extracted urban areas had no significant difference with those using 2.5-m pixel spacing images. However, urban areas extracted by using 25-m pixel spacing images had worse accuracy than those extracted using 2.5-m pixel spacing images. This was partly because the polarimetric coherence of urban areas increases with pixel spacing, so thresholding falsely

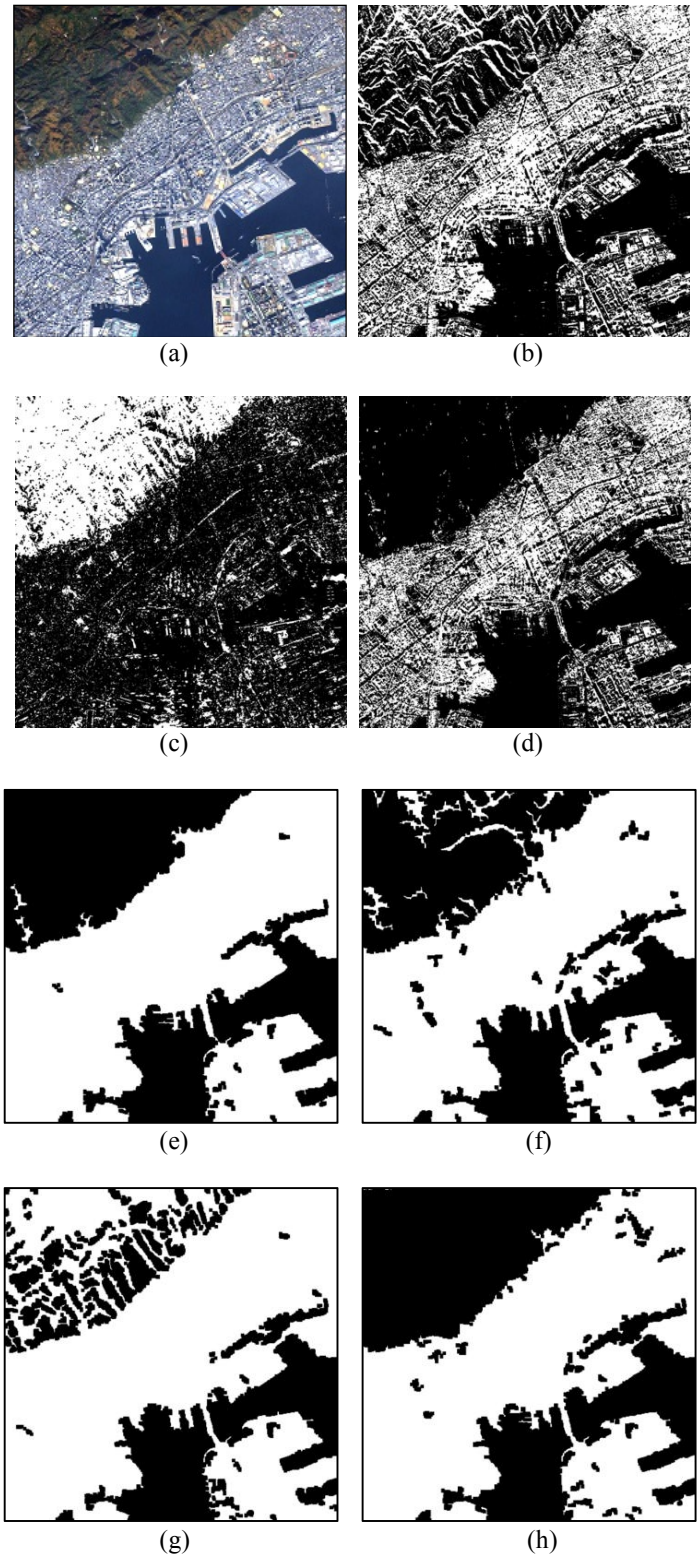


Fig. 5. Urban areas extracted from Pi-SAR2 images in study area 1. (a) ALOS/AVNIR-2 image; (b) result of the first-stage classification of the proposed method; (c) result of the second-stage classification of the proposed method; (d) intersected image of (b) and (c); (e) urban areas finally extracted by using the proposed method by applying post-processing to (d); (f) urban areas extracted by using the method in [13]; (g) urban areas extracted by using the method in [10]; (h) urban areas obtained by applying the method that replaces the second-stage classifier of the proposed method with (8). Each panel has an area of 6 km × 6 km.

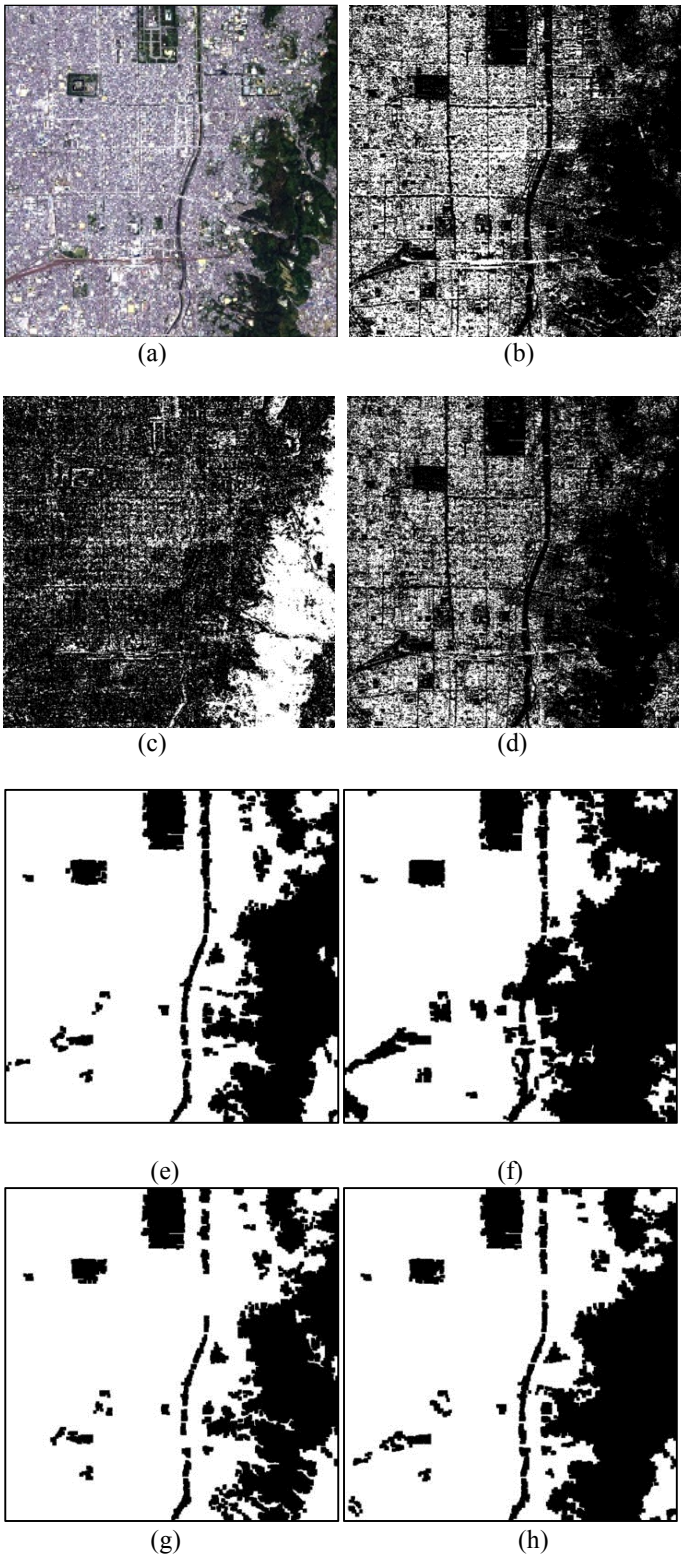


Fig. 6. Urban areas extracted from Pi-SAR2 images in study area 2. See Fig. 5 for a description of each panel.

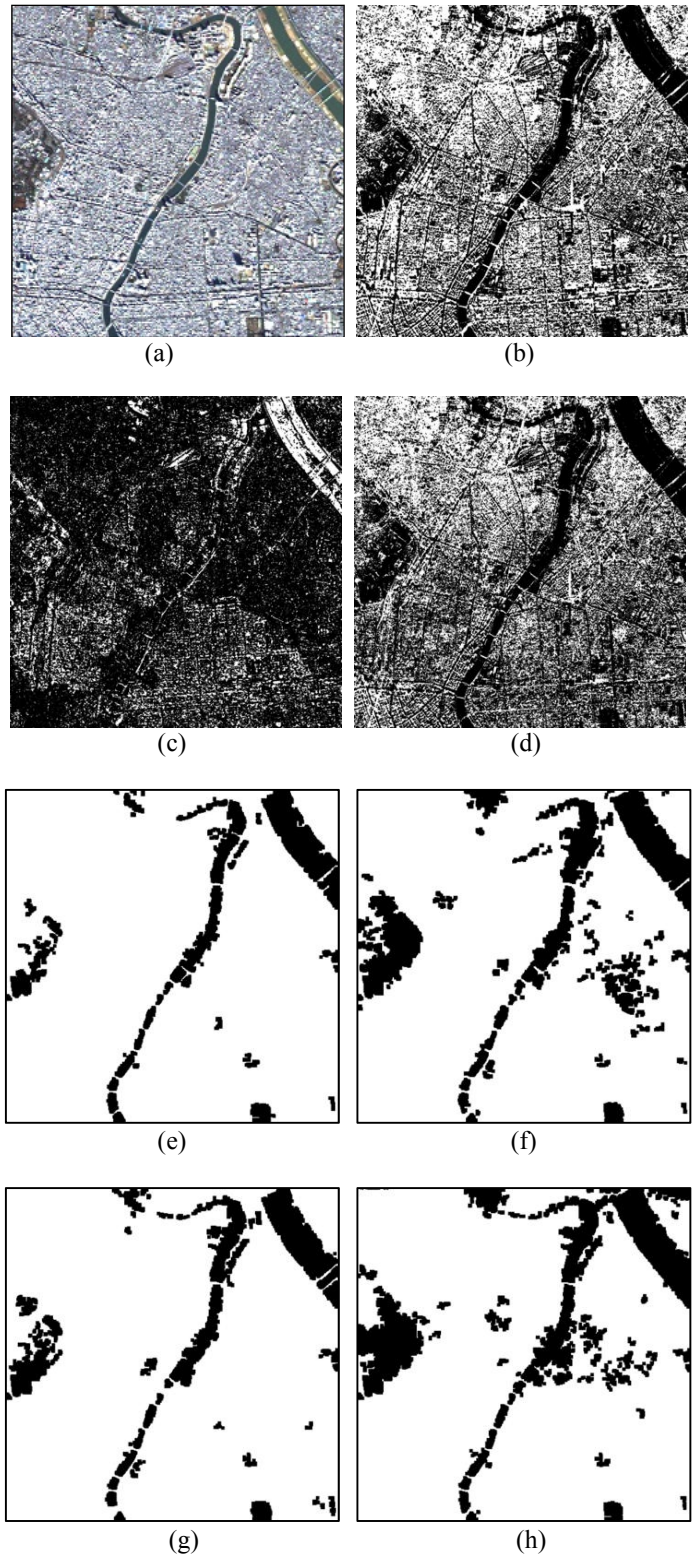


Fig. 7. Urban areas extracted from Pi-SAR2 images in study area 3. See Fig. 5 for a description of each panel.

TABLE II. Accuracy of extracted urban areas in study area 1, area 2, and area 3. "Proposed", "Esch (2013)" and "Kajimoto (2013)" denote the results obtained by using the proposed method, the method in [13] and the method in [10], respectively. "Classified using (8)" denotes the results obtained by replacing the second-stage classifier of the proposed method with (8).

Study area 1 (Kobe)			Actual (pixels)			User's accuracy
			Urban	Non-urban	Total	
Estimated (pixels)	Urban	Proposed	1520	753	2273	66.9%
		Esch (2013)	1514	847	2361	64.1%
		Kajimoto (2013)	1518	1278	2796	54.3%
		Classified using (8)	1509	698	2207	68.4%
	Non-urban	Proposed	9	1318	1327	99.3%
		Esch (2013)	15	1224	1239	98.8%
		Kajimoto (2013)	11	793	804	98.6%
		Classified using (8)	20	1373	1393	98.6%
	Total		1529	2071	3600	
	Producer's accuracy	Proposed	99.4%	63.6%	Overall accuracy	78.8%
		Esch (2013)	99.0%	59.1%		76.1%
		Kajimoto (2013)	98.6%	38.3%		64.2%
		Classified using (8)	98.7%	66.3%		80.1%

Study area 2 (Kyoto)			Actual			User's accuracy
			Urban	Non-urban	Total	
Estimated	Urban	Proposed	2495	231	2726	91.5%
		Esch (2013)	2186	140	2326	94.0%
		Kajimoto (2013)	2531	297	2828	89.5%
		Classified using (8)	2439	221	2660	91.7%
	Non-urban	Proposed	175	699	874	80.0%
		Esch (2013)	484	790	1274	62.0%
		Kajimoto (2013)	139	633	772	82.0%
		Classified using (8)	231	709	940	75.4%
	Total		2670	930	3600	
	Producer's accuracy	Proposed	93.4%	75.2%	Overall accuracy	88.7%
		Esch (2013)	81.9%	84.9%		82.7%
		Kajimoto (2013)	94.8%	68.1%		87.9%
		Classified using (8)	91.3%	76.2%		87.4%

Study area 3 (Tokyo)			Actual			User's accuracy
			Urban	Non-urban	Total	
Estimated	Urban	Proposed	2861	490	3351	85.4%
		Esch (2013)	2764	433	3199	86.5%
		Kajimoto (2013)	2843	116	2957	85.8%
		Classified using (8)	2737	422	3159	86.6%
	Non-urban	Proposed	98	151	249	60.6%
		Esch (2013)	195	208	403	51.6%
		Kajimoto (2013)	116	172	288	59.7%
		Classified using (8)	222	219	441	49.7%
	Total		2959	641	3600	
	Producer's accuracy	Proposed	96.7%	23.6%	Overall accuracy	83.7%
		Esch (2013)	93.4%	32.4%		82.6%
		Kajimoto (2013)	96.1%	26.8%		83.8%
		Classified using (8)	92.5%	34.2%		82.1%

extracts urban areas as vegetation. Therefore, we found that over-multilooking lowers the accuracy of urban area classification.

C. Sensitivity Analysis

Finally, we discuss the sensitivity of the parameters used in the proposed method. The threshold for the first-stage classification using P_v and TP principal components is optimized. Once an a priori threshold for a study area is available, the optimal threshold for another study area is automatically generated without obtaining training data. In this sense, we regard the proposed method as a semi-supervised classifier. The classification results may depend on selection of the training data obtained in the first study area.

The threshold for the second-stage classification was found to affect the final classification results more than did the threshold for the first-stage classification. In this research, we determined the threshold for $\gamma_{HH,VV}$ as 0.8 by taking samples from Pi-SAR2 images. This value may be dependent on the study area and the data, and thus users should determine the optimal threshold by taking samples. While (8) is corrected in terms of POA effects, (9) is not. Therefore, after thresholding some districts in urban areas—those with high-amplitude HV scattering—were classified as vegetation. By intersection, such districts were tentatively classified as non-urban areas. However, as mentioned in the previous subsection, the tentative urban areas are extended by filtering and closing. As a result, the omitted urban areas are reclassified as urban areas if they are close to areas tentatively classified as urban areas. Therefore, due to its image processing technique, the proposed method is designed so that threshold errors have less effect in the second-stage classification.

VII. Conclusions

We presented a method for extracting urban areas from X-band PolSAR data. The proposed method is based on two classifications: classification using a combination of TP and P_v, and classification using polarimetric coherence of S_{HH} and S_{VV}. The two results are intersected and final urban areas are extracted after post-classification processing. We applied the proposed method to airborne X-band Pi-SAR2 images of Kobe, Kyoto, and Tokyo. Validation showed that the results of the proposed method are acceptable, with an overall accuracy of approximately 80 to 90% at a 100-m spatial scale. Whereas sensitivity analysis revealed that the most sensitive threshold for extracting urban areas was that for polarimetric coherence of S_{HH} and S_{VV}, we confirmed that the extracted results were stable in most cases. We conclude that the proposed method can be applied to X-band PolSAR images for robustly extracting urban areas.

Acknowledgment

This research was supported by a JSPS Grant-in-Aid for Scientific Research (KAKENHI) for Young Scientists (B) (24760412) and by a Research Announcement of Pi-SAR2, NICT, Japan. Zmap-Town II was provided by the Center for Spatial Information Science, The University of Tokyo.

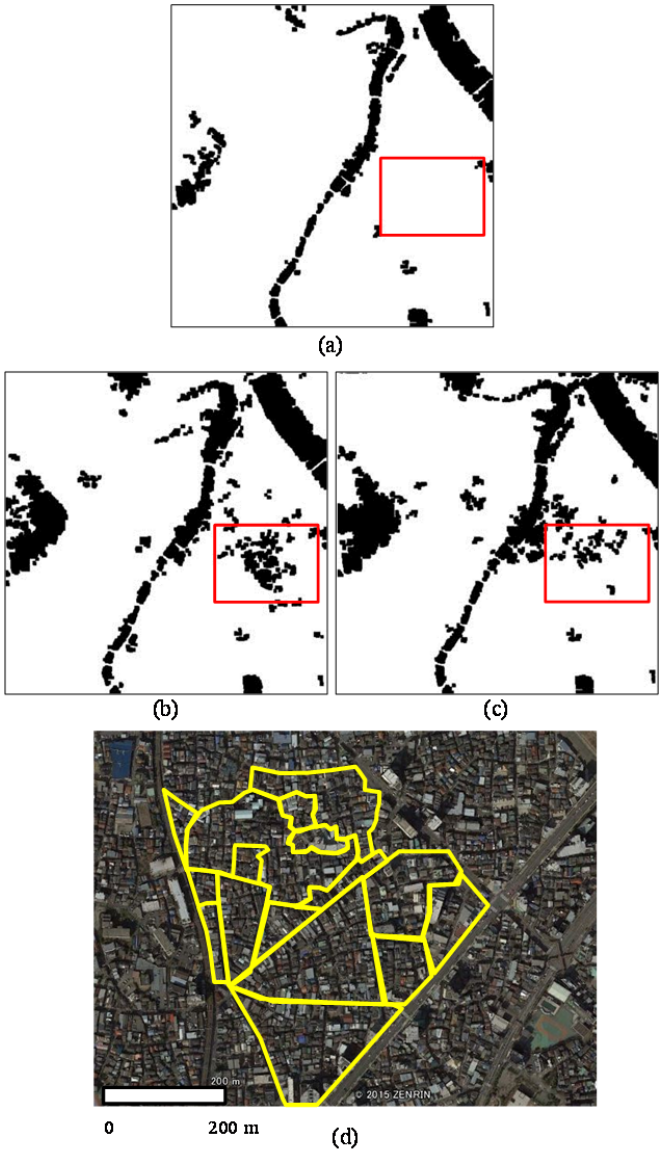


Fig. 8. Urban areas omitted via classification. (a) Result obtained by using the proposed method, (b) result obtained by using the method in [13], (c) result obtained by using the classification method with (8), and (d) optical image from Google Earth. In (d), the districts surrounded by yellow solid lines are areas having homogeneous POA.

References

[1] F. Xu and Y. Q. Jin, "Deorientation theory of polarimetric scattering targets and application to terrain surface classification", *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2351-2364, 2005.

[2] H. Kimura, "Radar polarization orientation shifts in built-up areas", *IEEE Geosci. Remote Sens. Letters*, vol. 5, pp. 217-221, 2008.

[3] W. An, C. Xie, X. Yuan, Y. Cui and J. Yang, "Four-component decomposition of polarimetric SAR images with deorientation", *IEEE Geosci. Remote Sens. Letters*, vol. 8, no. 6, pp. 1090-1094, 2011.

[4] Y. Yamaguchi, A. Sato, W. M. Boerner, R. Sato and H. Yamada, "Four-component scattering power decomposition with rotation of

coherency matrix", *IEEE Trans. Geosci. Remote Sens.*, vo. 49, no. 6, pp. 2251-2258, 2011.

[5] Y. Yamaguchi, T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric SAR image decomposition", *IEEE Trans. Geosci. Remote Sens.*, vol. 43, pp. 1699-1706, 2005.

[6] Y. Yamaguchi, Y. Yajima, and H. Yamada, H., "A four-component decomposition of POLSAR images based on the coherency matrix", *IEEE Geosci. Remote Sens. Letters*, vol. 3, pp. 292-296, 2006.

[7] J. S. Lee, and T. L. Ainsworth, "The effect of orientation angle compensation on coherency matrix and polarimetric target decompositions", vol. 49, no. 1, pp. 53-64, 2011.

[8] A. Freeman and S. L. Durden, "A three-component scattering model for polarimetric SAR data", *IEEE Trans. Geosci. Remote Sens.,* vol. 36, no. 3, pp. 963-973, 1998.

[9] S. W. Chen, M. Ohki, M. Shimada and M. Sato, "Deorientation effect investigation for model-based decomposition over oriented built-up areas", *IEEE Geosci. Remote Sens. Letters*, vol. 10, no. 2, pp. 273-277, 2013.

[10] M. Kajimoto and J. Susaki, "Urban-area extraction from polarimetric SAR images using polarization orientation angle", *IEEE Geosci. Remote Sens. Letters*, vol. 10, no. 2, pp. 337-341, 2013.

[11] M. Kajimoto and J. Susaki, "Urban density estimation from polarimetric SAR images based on a POA correction method", *IEEE J. Selected Topics Applied Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1418-1429, 2013.

[12] J. Susaki, M. Kajimoto and M. Kishimoto, "Urban density mapping of global megacities from polarimetric SAR images", *Remote Sens. Env.*, vol. 155, pp. 334-348, 2014.

[13] T. Esch, M. Marconcini, A. Felbier, A. Roth, W. Heldens, M. Huber, M. Schwinger, H. Taubenböck, A. Müller and S. Dech, "Urban footprint processor – fully automated processing chain generating settlement masks from global data of the TanDEM-X mission", *IEEE Geosci. Remote Sens. Letters*, vol. 10, pp. 1617–1621, 2013.

[14] A. Ferro, D. Brunner, L. Bruzzone and G. Lemoine, "On the relationship between double bounce and the orientation of buildings in VHR SAR images", *IEEE Geosci. Remote Sens. Letters*, vol. 8, pp. 612-616, 2011.

[15] J. Uemoto, S. Uratsuka, T. Umehara, S.-I. Yamamoto, S. Taira, M. Satake, S. Kojima, T. Kobayashi, M. Satoh, K. Kawasaki, T. Matsuoka, A. Nadai and R. Suzuki, Development of the onboard processor for Pi-SAR2, *Proceedings of IGARSS 2011*, pp. 906–909, 2011.

[16] M. Satake, T. Kobayashi, J. Uemoto, T. Umehara, S. Kojima, T. Matsuoka, A. Nadai, and S. Uratsuka, "Damage estimation of the Great East Japan earthquake with airborne SAR (PI-SAR2) data", *Proceedings of IGARSS 2012*, pp. 1190-1191, 2012.

[17] T. Kobayashi, T. Umehara, J. Uemoto, M. Satake, S. Kojima, T. Matsuoka, A. Nadai, and S. Uratsuka, "Volcanic monitoring by polarimetric and interferometric airborne SAR (Pi-SAR2), *Proceedings of APSAR 2013*, pp. 487-490, 2013.

[18] S. Quegan, "A unified algorithm for phase and cross-talk calibration of polarimetric data-theory and observations", *IEEE Trans. Geosci. Remote Sens.*, vol. 32, pp. 89-99, 1994.

[19] Zenrin, Co. Ltd, 2014. Zmap Town II. Available at http://www.zenrin.co.jp/product/gis/zmap/zmaptown.html (in Japanese: accessed May 29, 2015).