APPLICATION OF RAJSKI DISTANCE TO LAND-COVER CLASSIFICATION USING POLARIMETRIC SAR AMPLITUDE IMAGE DATA

T. Yamada[a*], T. Hoshi [b]

[a] Dept. of Electrical Engineering,, Fukushima National College of Technology, 30 Kamiarakawa Taira Iwaki, Fukushima, 970-8034 Japan - yamada@fukushima-nct.ac.jp
[b] Dept. of Computer and Information Science, Ibaraki University, 4-12-1 Nakanarusawa Hitachi, Ibaraki, 316-8511 Japan - hoshi@cis.ibaraki.ac.jp

Poster Session, Working Group VII/1

KEY WORDS: Remote Sensing, Land Cover, Classification, Polarization, Algorithms, SAR

ABSTRACT:

Polarimetric SAR can observe scattering matrix for each resolution cell and provide amplitude images that have gray level in proportion to amplitude of the conjunction matrix elements. However, these amplitude image data have been used for pseudo color synthesize and construction of feature vector for land-cover classification as the vector elements mainly, discussion about features derived from amplitude image data was scarcely. In this paper, we consider pattern difference in different polarization amplitude image of polarimetric SAR as probability, and discuss about contribution of expanding dimension of feature vector by introducing Rajski distance as a features. To calculate Rajski distance, gray level co-occurrence matrix (GLCM) method that has been often used for texture analysis was used. In the proposed method, gray levels of the pixel that was located at the same coordinates in two different transmit and receive polarization amplitude images are adapted to line and column of GLCM, then co-occurrence probability are calculated. From this matrix, joint entropy and conditional entropy are derived and Rajski distance is found. To inspect the proposed theory, we investigate the effect of expanding dimension of feature vector for land-cover classification by Euclid minimum distance method and maximum likelihood method as well known supervised classification method generally using SIR-C polarimetric data obtained in two kinds of scenes including different land-cover objects. As the results, improvement of accuracy in the point of classification score and ambiguity scale, so that the effectiveness of introduction of Rajski distance to expanding dimension of feature vector for land-cover classification is demonstrated.

1. INTRODUCTION

The various land-cover classification algorithms using the data of the polarimetric SAR have been proposed until now, and the validity is reported. For example, they are supervised classification methods based on Bayes theory and using neural network, and unsupervised classification methods based on scattering types and scattering entropy. Each of these corresponds to full-polarimetric data that used both amplitude information and phase information for each element of scattering matrix obtained by polarimetric SAR (Yueh, H. A., et al., 1988; Ito, Y. and Omatsu, S., 1997; Zyl, J. J., 1989; Cloude, S. R. and Pottier, E., 1997). However, in the case where whole information of scattering matrix is acquired, and the situation that only amplitude information is acquired, some of these techniques cannot be apply or deteriorate classification accuracy. In order to analyze to image data only for amplitude information, the classification technique as a general digital image are applied. It could judge visually that the difference has arisen in the amplitude image data for the different polarization used for observation by polarimetric SAR, when performing such a classification. Therefore, it is necessary to extract quantitatively the difference in the feature by the difference kind of polarization that is latent in the amplitude image data of the polarimetric SAR, and to raise classification accuracy with the application of it for dimension extension of the feature vector of land-cover classification for amplitude image. In this paper, application of the Rajski distance based on information theory is proposed as the parameter. Although entropy is conventionally used for the texture analysis of a digital image, Rajski distance is used as a distance measure that indicates the similarity of two probability phenomenon systems using entropy. In this paper, Rajski distance is calculated for each pixel by making into a probability phenomenon system the pixel value distribution of the amplitude image data from that the combination of two different polarizations of polarimetric SAR data: SIR-C (Shuttle Imaging Radar-C) and airborne Pi-SAR (Polarimetric and Interferometric SAR), the distances for some area are characterized, and the application effect of Rajski distance to land-cover classification is reported.

2. CALCULATION OF RAJSKI DISTANCE FROM POLARIMETRIC SAR IMAGE DATA

2.1 Overview of Rajski Distance

In this section, overview of Rajski distance is described (Isomichi, Y., 1980). Two probability phenomenon systems are set to X and Y, respectively, and each phenomenon is set to a_i (i = 0, …, m–1) and b_j (j = 0, …, n–1), respectively. The probable relevance of both phenomenon systems is determined by the simultaneous occurrence probability $p(a_i, b_j)$. Joint entropy $H(XY)$ for X and Y is given using simultaneous occurrence probability as follows.

$$H(XY) = -\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} p(a_i,b_j)\log p(a_i,b_j) \quad (1)$$

The conditional information content $H(X|Y)$ and $H(Y|X)$ are given as follows

$$\left.\begin{aligned} H(X \mid Y) &= -\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} p(b_j)\,p(a_i \mid b_j)\log\, p(a_i \mid b_j) \\ H(Y \mid X) &= -\sum_{j=0}^{n-1}\sum_{i=0}^{m-1} p(a_i)\,p(b_j \mid a_i)\log\, p(b_j \mid a_i) \end{aligned}\right\} \quad (2)$$

where, $p(a_i|b_j)$ and $p(b_j|a_i)$ are conditional probability, respectively. Using these information content, Rajski distance $\rho(X, Y)$ is defined by the following formula.

$$\rho(X,Y) = \frac{H(X \mid Y) + H(Y \mid X)}{H(XY)} \quad (3)$$

In addition, the range of this Rajski distance is 1 or less and 0 or more, it takes 0 when two probability phenomenon systems are in agreement, and it takes 1 when these are independent.

2.2 Calculation Technique of Rajski Distance from Polarimetric SAR Image Data

As the technique of calculating Rajski distance using polarimetric SAR amplitude image data, the method by co-occurrence matrix used in texture analysis is applied. This matrix constructs a square matrix with the size of N x N, when the gray level in a partial image is set to N. In this matrix, when a pixel value is taken in the position with row and column, the element is appearance probability of pixel with the combination of the pixel value. This matrix is called gray level co-occurrence matrix (GLCM). When GLCM is used for texture analysis, the pixel value of the circumference pixel which takes the pixel value of an attention pixel in the row direction of a procession, and is about an attention pixel and its contiguity pixel in the same picture at the place which only a specific distance separated from the attention pixel with the specific angle in the direction of a sequence is taken. Generally the distance between attention pixel and contiguity pixel is one, and an angle is taken at 0, 45, 90, and 135 degrees (Haralick, R. M., et al., 1973).

In this paper, it considers constituting GLCM from combining two kinds of polarization from the amplitude image to HH, HV, and VV polarization used in many cases for visualizing of polarimetric SAR data. When the gray level of one of amplitude image is set i, and that of another amplitude image is set j, the frequency of appearance of the combination of the gray level is expressed with $P(i, j, q)$. Where, q expresses the combination of transmit and receive polarization. The element of GLCM $p(i, j, q)$ can be expressed as follows.

$$p(i,j,q) = \frac{P(i,j,q)}{\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} P(i,j,q)} \quad (4)$$

In this formula, N is the number of gradation. GLCM $[\boldsymbol{G}(q)]$ for combination of polarization q is expressed as follows.

$$[G(q)] = \begin{bmatrix} p(0,0,q) & \cdots & p(0,N-1,q) \\ \vdots & \ddots & \vdots \\ p(N-1,0,q) & \cdots & p(N-1,N-1,q) \end{bmatrix} \quad (5)$$

In the case of SIR-C, combination of polarization q = 3; HH-VV, HH-HV, and HV-VV.

Next, the technique of calculating co-occurrence matrix to Rajski distance is described. In the two kinds of transmit and receive polarization amplitude images, the probability that the pixel value will appear in one of images is corresponded to the probability phenomenon system X, and the probability that the pixel value will appear in another image is corresponded to the probability phenomenon system Y, respectively. Using the elements of co-occurrence matrix, joint entropy for the combinations of transmit and receive polarization q is obtained as follows.

$$H_q(XY) = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j,q)\log\, p(i,j,q) \quad (6)$$

And the conditional information contents are also obtained as follows.

$$\left.\begin{aligned} H_q(X \mid Y) &= -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} p(i,j,q)\log\frac{p(i,j,q)}{p_q(j)} \\ H_q(Y \mid X) &= -\sum_{j=0}^{N-1}\sum_{i=0}^{N-1} p(i,j,q)\log\frac{p(i,j,q)}{p_q(i)} \end{aligned}\right\} \quad (7)$$

Using these information contents, Rajski distance $\rho(q)$ for the combinations of transmit and receive polarization q is obtained the same as equation (3).

2.3 Creation of Rajski Distance Image

The algorithm that creates Rajski distance image constructed by assigning Rajski distance obtained from two kinds of single polarization amplitude image to pixel value is described. Main processing that creates Rajski distance image is as follows.

2.3.1 Extraction of Sub Area: To construct co-occurrence matrix, sub areas are extracted with M x M pixels. Note the size of sub areas because reliability of co-occurrence probability will decline with too small size and influenced area of tiny noise that appears in sub area will be spread with too large size.

2.3.2 Construction of Co-occurrence Matrix: In the extracted sub area, the appearance frequency for the combinations each pixel value of two amplitude images are calculated and co-occurrence frequency matrix is constructed. Each element in the frequency matrix is transformed to probability using the size of sub area.

2.3.3 Calculation of Information Contents: Joint entropy and conditional information contents are calculated using the constructed co-occurrence matrix.

2.3.4 Calculation of Rajski Distance: Rajski distance is calculated for attention pixel using obtained joint entropy and conditional information contents.

2.3.5 Quantization of Rajski Distance to Pixel Value: Calculated Rajski distance is quantized to pixel value that set a limit from 0 to 255. The range of Rajski distance corresponds to one gray scale of pixel value is 1/256.

Reiterating a series of these processing for whole pixels with scanning to x and y direction in two images, Rajski distance image is created. The outline of this process is illustrated in Figure 1 (Yamada, T. and Hoshi, T., 2002a).

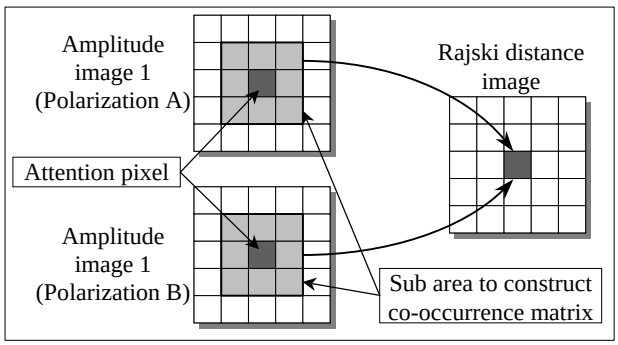


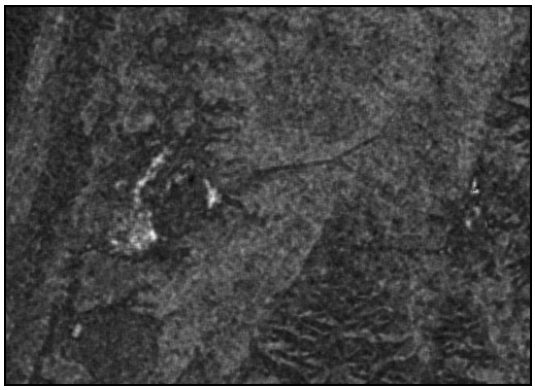
Figure 1. Creation process of Rajski distance image

3. APPLICATION OF POLARIMETRIC SAR DATA

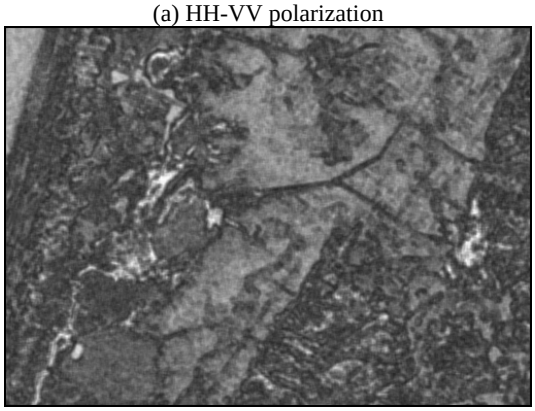
The proposed technique in this paper was applied to land-cover classification using two types of polarimetric SAR data: SIR-C and Pi-SAR. L-band and C-band data observed by SIR-C at Sarobetsu site and Kashima site in 1994 were applied, and X-band data observed by Pi-SAR at Hitachi site in 1998 and 2002 were applied. Examples of Rajski distance image created by each observation site in Figure 2, 3, and 4. The created images for SIR-C Sarobetsu site data are shown in Figure 2, for SIR-C Kashima site data are shown in Figure 3, and for Pi-SAR Hitachi site data are shown in Figure 4, respectively. In this experiments, sub areas were extracted with the size of 9 x 9 pixels when Rajski distance was calculated.

As the common property, the tendencies that Rajski distance is higher in water area and lower in vegetation are shown in all images. In water area, pixel values that exist in sub area are different slightly although these values are quite similar. So that it can consider that Rajski distance is higher because mutual information contents was lower due to narrow deviation of pixel value. On the other hand, the property of Rajski distance in vegetation can be described relating with microwave scattering model. It is known widely that notable diffusion scattering appears in this area. In this case, helix components are included largely according to basic scattering model. Rajski distance is lower due to similar pixel value increases equally in each amplitude image because the elements in scattering matrix for helix have same value. This reason has been confirmed by simulation.

To introduce Rajski distance to land-cover classification, feature vector is constructed with pixel value of Rajski distance images and amplitude images.



(a) HH-VV polarization



(b) HH-HV polarization

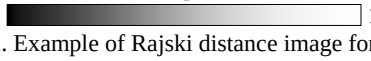
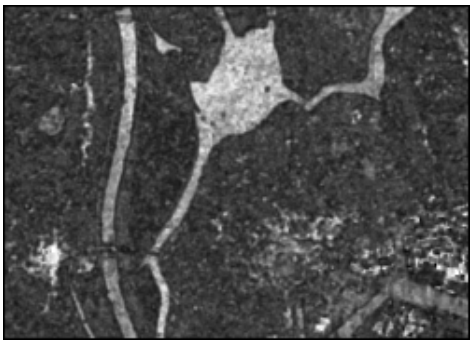
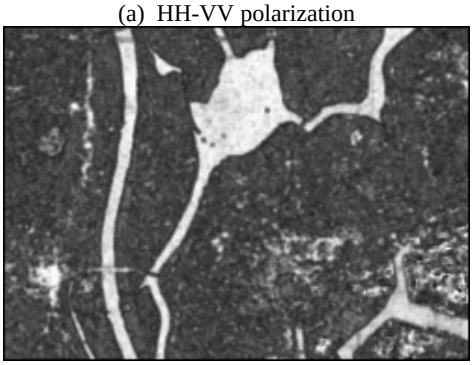
0  1

Figure 2. Example of Rajski distance image for SIR-C Sarobetsu site C-band data (1994)



(a) HH-VV polarization



(b) HH-HV polarization

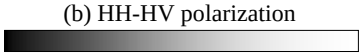
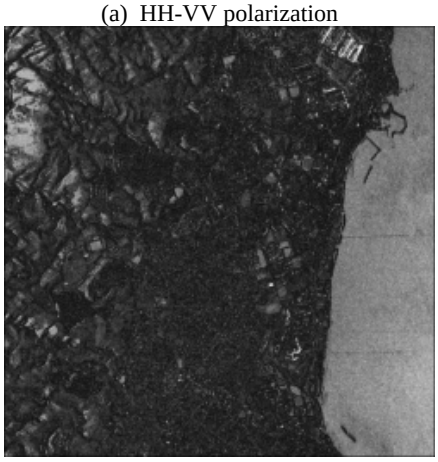
0  1

Figure 3. Example of Rajski distance image for SIR-C Kashima site C-band data (1994)

(a) HH-VV polarization



(b) HH-HV polarization

0  1

Figure 4. Example of Rajski distance image for Pi-SAR Hitachi site X-band data (2002)

4. CLASSIFICATION RESULTS

4.1 Classification Results for SIR-C Data

Using SIR-C Data, three amplitude images for HH, HV, and VV polarization can be obtained for two wavelength band: L-band and C-band. In this section, supervised classification results are shown for two different feature vectors. One of them has six elements that are pixel values of six amplitude images, three kinds of polarization for two kinds of wavelength band. Another of them has twelve elements that are six elements from amplitude images and other elements are obtained Rajski distance images. As the classification technique, Euclid distance method and maximum likelihood method.

The classification score matrix for Sarobetsu site is shown in Table 1. And the matrix for Kashima site is shown in Table 2. In these table, table (a) and (b) are the case of six elements, table (c) and (d) are the case of twelve elements. And table (a) and (c) are the results of Euclid distance method, table (b) and (d) are the results of maximum likelihood method. There numbers shows the score after otherwise class was excluded from test areas.

In these tables, category names are defined under the tables. Category names of column direction, $C1$, $C2$, and so on, mean the categories before classification and ones the names of row direction, $\overline{C}1$, $\overline{C}2$, and so on, mean the categories after classification. And the numbers in ( ) mean pixel numbers classified to the categories. The sum of scores to row direction will be 100%.

In Table 1, average classification accuracy was 70.92% for (a), 75.31% for (b), 78.86% for (c), and 88.28% for (d), respectively.

Table 1. Classification score matrix for Sarobetsu site

(Unit: %)

(a) Euclid distance method (six elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$
$C1$	74.6 (546)	4.1 (30)	21.0 (154)	0.1 (1)	0 (0)	0.1 (1)
$C2$	0 (0)	86.4 (6798)	0.9 (70)	0.2 (18)	11.7 (925)	0.8 (61)
$C3$	8.1 (430)	42.7 (2261)	43.5 (2305)	2.7 (145)	1.5 (82)	1.5 (77)
$C4$	0 (0)	38.1 (444)	0.1 (1)	58.3 (680)	0 (0)	3.5 (41)
$C5$	0 (0)	0 (0)	0 (0)	0 (0)	100.0 (144)	0 (0)
$C6$	0 (0)	34.3 (472)	0 (0)	0.3 (4)	0.1 (1)	58.4 (804)

(b) Maximum likelihood method (six elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$
$C1$	76.8 (562)	0 (0)	23.2 (170)	0 (0)	0 (0)	0 (0)
$C2$	0.1 (4)	76.4 (6018)	16.0 (1258)	0.1 (8)	0 (0)	7.4 (584)
$C3$	13.1 (694)	14.1 (748)	70.0 (3709)	0 (0)	0 (0)	2.8 (149)
$C4$	0 (0)	4.9 (57)	0.7 (8)	94.1 (1097)	0 (0)	0.3 (4)
$C5$	0 (0)	47.2 (68)	0 (0)	0 (0)	52.8 (76)	0 (0)
$C6$	0 (0)	14.7 (203)	2.2 (30)	0 (0)	0 (0)	76.1 (1048)

(c) Euclid distance method (twelve elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$
$C1$	70.7 (442)	1.0 (6)	28.3 (177)	0 (0)	0 (0)	0 (0)
$C2$	0 (0)	97.4 (6486)	1.8 (118)	0 (0)	0 (0)	0.8 (55)
$C3$	2.1 (108)	32.1 (1640)	64.4 (3285)	0.3 (13)	0 (0)	1.1 (54)
$C4$	0 (0)	10.1 (115)	0 (0)	89.9 (1025)	0 (0)	0 (0)
$C5$	0 (0)	14.3 (20)	0 (0)	0 (0)	85.7 (120)	0 (0)
$C6$	0 (0)	34.8 (410)	0 (0)	0.2 (2)	0 (0)	65.0 (766)

(d) Maximum likelihood method (twelve elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$
$C1$	82.1 (513)	0 (0)	17.9 (112)	0 (0)	0 (0)	0 (0)
$C2$	0 (0)	90.3 (6012)	5.8 (387)	0 (0)	0 (0)	3.9 (260)
$C3$	1.5 (76)	12.7 (647)	85.5 (4363)	0 (0)	0 (0)	0.3 (14)
$C4$	0 (0)	1.4 (16)	0 (0)	98.6 (1124)	0 (0)	0 (0)
$C5$	0 (0)	15.0 (21)	0 (0)	0 (0)	85.0 (119)	0 (0)
$C6$	0 (0)	11.3 (133)	0.5 (6)	0 (0)	0 (0)	88.2 (1039)

$C1$: urban, $C2$: grass, $C3$: forest, $C4$: sea,
$C5$: opened lake, $C6$: frozen lake

In Table 2, average classification accuracy was 73.91% for (a), 66.79% for (b), 87.57% for (c), and 84.47% for (d), respectively. In the both sites, classification accuracy was improved by increasing the elements of feature vector.

Table 2. Classification score matrix for Kashima site

(Unit: %)

(a) Euclid distance method (six elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$	$\overline{C}7$
$C1$	55.7 (374)	0.3 (2)	12.4 (83)	0.1 (1)	31.5 (212)	0 (0)	0 (0)
$C2$	0 (0)	78.5 (439)	0.2 (1)	21.3 (119)	0 (0)	0 (0)	0 (0)
$C3$	7.5 (290)	1.1 (41)	87.9 (3421)	2.8 (108)	0.8 (30)	0 (0)	0 (0)
$C4$	0.2 (4)	4.5 (109)	2.8 (68)	87.0 (2116)	0.6 (14)	0 (0)	5.0 (122)
$C5$	11.0 (54)	5.7 (28)	4.9 (24)	0 (0)	78.4 (385)	0 (0)	0 (0)
$C6$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	66.4 (3255)	33.6 (1645)
$C7$	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	36.6 (256)	63.4 (444)

(b) Maximum likelihood method (six elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$	$\overline{C}7$
$C1$	32.0	0	41.2	0	26.8	0	0
	(215)	(0)	(277)	(0)	(180)	(0)	(0)
$C2$	0	91.1	3.0	5.9	0	0	0
	(0)	(509)	(17)	(33)	(0)	(0)	(0)
$C3$	0.5	0.2	97.5	1.8	0	0	0
	(20)	(7)	(3792)	(70)	(0)	(0)	(0)
$C4$	0.1	12.3	6.6	81.0	0	0	0
	(2)	(299)	(161)	(1970)	(1)	(0)	(0)
$C5$	7.9	2.0	21.2	0.2	68.6	0	0
	(39)	(10)	(104)	(1)	(337)	(0)	(0)
$C6$	0	0	0	13.2	0	15.5	71.3
	(0)	(0)	(0)	(646)	(0)	(761)	(3493)
$C7$	0	1.6	0	14.3	0	2.3	81.9
	(0)	(11)	(0)	(100)	(0)	(16)	(573)

(c) Euclid distance method (twelve elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$	$\overline{C}7$
$C1$	76.8	0	0	0	23.2	0	0
	(384)	(0)	(0)	(0)	(116)	(0)	(0)
$C2$	0	83.5	0	16.5	0	0	0
	(0)	(449)	(0)	(89)	(0)	(0)	(0)
$C3$	0	1.8	95.2	2.6	0.4	0	0
	(0)	(71)	(3681)	(100)	(15)	(0)	(0)
$C4$	0	7.2	1.9	90.8	0.1	0	0
	(0)	(171)	(45)	(2146)	(2)	(0)	(0)
$C5$	7.7	1.0	1.2	0	90.1	0	0
	(32)	(4)	(5)	(0)	(375)	(0)	(0)
$C6$	0	0	0	0	0	84.0	16.0
	(0)	(0)	(0)	(0)	(0)	(3858)	(737)
$C7$	0	0	0	1.5	0	5.8	92.6
	(0)	(0)	(0)	(10)	(0)	(38)	(605)

(d) Maximum likelihood method (twelve elements)

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$	$\overline{C}6$	$\overline{C}7$
$C1$	63.2	0	0	0	36.8	0	0
	(316)	(0)	(0)	(0)	(184)	(0)	(0)
$C2$	0	90.1	2.6	7.2	0	0	0
	(0)	(485)	(14)	(39)	(0)	(0)	(0)
$C3$	0	0	99.9	0	0.1	0	0
	(0)	(0)	(3866)	(0)	(1)	(0)	(0)
$C4$	0	5.7	6.9	87.3	0.1	0	0
	(0)	(136)	(162)	(2064)	(2)	(0)	(0)
$C5$	1.0	0	1.7	0	97.4	0	0
	(4)	(0)	(7)	(0)	(405)	(0)	(0)
$C6$	0	0	0	0	0	59.1	40.9
	(0)	(0)	(0)	(0)	(0)	(2717)	(1878)
$C7$	0	0.2	0	0.3	0	5.4	94.2
	(0)	(1)	(0)	(2)	(0)	(35)	(615)

*C*1: urban, *C*2: grass, *C*3: paddy 1, *C*4: paddy 2, *C*5: factory, *C*6: lake, *C*7: river

4.2 Classification Results for Pi-SAR Data

Pi-SAR is an airborne polarimetric SAR that can observe with multi-frequency, L-band and X-band, and multi-polarization. However, the observation in 1998 for Hitachi site could obtain only two kinds of polarization in four polarizations. For that data, polarization synthesize and pseudo color synthesize have been difficult. Finding Rajski distance proposed in this paper for two obtained amplitude images, influence for land-cover classification results using feature vector that have increased elements was estimated (Yamada, T. and Hoshi, T., 2002b). For this data, land-cover classification score matrices were calculated by most likelihood method.

The results are shown in Table 3. In the table, (a) is the case that feature vector has two elements, and (b) is the case of three elements. In these tables, category names are defined under the tables.

In table 3, average classification accuracy was 74.0% for (a) and 85.28% for (b), respectively. As the same as the case of SIR-C, classification accuracy was improved by increasing the elements of feature vector.

Table 3. Classification score matrix for Hitachi site

(Unit: %)

(a) Two elements

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$
$C1$	93.1	4.8	2.0	0	0.1
	(1489)	(77)	(32)	(0)	(2)
$C2$	9.8	49.4	37.2	0	3.6
	(33)	(169)	(126)	(0)	(12)
$C3$	1.5	17.0	64.8	0	16.7
	(14)	(153)	(583)	(0)	(150)
$C4$	0	0	0	87.7	12.3
	(0)	(0)	(0)	(3157)	(443)
$C5$	1.7	3.3	6.7	13.3	75.0
	(1)	(2)	(4)	(8)	(45)

(b) Three elements

	$\overline{C}1$	$\overline{C}2$	$\overline{C}3$	$\overline{C}4$	$\overline{C}5$
$C1$	99.3	0.2	0.1	0	0.4
	(1589)	(3)	(2)	(0)	(6)
$C2$	5.1	59.8	28.7	0	6.4
	(17)	(203)	(98)	(0)	(22)
$C3$	1.9	13.4	79.1	0	5.7
	(17)	(121)	(711)	(0)	(51)
$C4$	0	0	0.1	99.9	0
	(0)	(0)	(4)	(3596)	(0)
$C5$	1.7	5.0	5.0	0	88.3
	(1)	(3)	(3)	(0)	(53)

*C*1: urban, *C*2: grass, *C*3: forest, *C*4: sea, *C*5: shadow

5. DISCUSSIONS

By increasing elements of feature vector for land-cover classification, improvement of average classification accuracy was confirmed generally for Table 1, 2, and 3. Although classification accuracy was improved for most of category by increasing elements of feature vector, that was lower for category C2 and C5 in Sarobetsu site by Euclid distance method. Sub area extracted to calculate Rajski distance also realizes filter effect. So misclassifications are happened in boundary of areas. Spatial ranges for both categories that classification accuracy was lower are narrow, so that the influences caused by misclassification in boundary parts are higher than other areas. This problem will be avoided by correcting the size to construct GLCM and selecting targets distributing widely. Actually, classification accuracy was increased in urban pattern of Kashima site distributing widely.

For Hitachi site Pi-SAR data that suffers a loss of some polarization data, classification accuracies for some categories were improved remarkably by increasing elements of feature vector although peculiar elements were scanty. Feature vector

will be constructed by more elements for the data acquired at the same area in 2002 because whole polarization data were acquired for the observation.
From the mentioned discussions, Rajski distance proposed in this paper is considered to be suitable as the element of feature vector for land-cover classification.

6. CONCLUSIONS

In this paper, Calculating Rajski distance from polarimetric SAR amplitude image data and introduce of it to land-cover classification as element of feature vector were proposed. To obtain Rajski distance, gray level co-occurrence matrix was constructed using two amplitude image data and joint entropy and conditional information contents were calculated using the matrix. Actual polarimetric SAR data, SIR-C and Pi-SAR, were applied to proposed algorithm. To introduce Rajski distance as elements of feature vector, the distance was quantized to gray scale image like other amplitude images. In Rajski distance images, some characteristic properties were appeared for water area and vegetation. Average classification accuracies for whole data were improved when extended feature vectors were applied actually to land-cover classification. The size of sub area extracted to construct GLCM for the area that has narrow spatial range needed correction. Optimize of size of sub area and introduction of other parameter obtained from GLCM will be needed to advance.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Masaharu Fujita, Professor of Tokyo Metropolitan Institute of Technology, Dr. Seiho Uratsuka, Group Leader of Environment Information Technology Group, Applied Research and Standard Dept., National Institute of Information and Communications Technology, offered the data of SIR-C and Pi-SAR.

REFERENCES

Yueh, H. A., et al., 1988. Bayes Classification of Terrain Cover Using Normalized Polarimetric Data. *J. Geophysical Research*, 93(B12), pp.15261-15267.

Ito, Y. and Omatsu, S., 1997. Polarimetric SAR Data Classification Using Neural Networks. *J. Japan Society of Photogrammetry and Remote Sensing*, 36(3), pp.13-22.

Zyl, J. J., 1989. Unsupervised Classification of Scattering Behavior Using Radar Polarimetry Data. *IEEE Trans. Geoscience and Remote Sensing*, 35(1), pp.68-78.

Cloude, S. R. and Pottier, E., 1997. An Entropy Based Classification Scheme for Land Applications of Polarimetric SAR. *IEEE Trans. Geoscience and Remote Sensing*, 35(1), pp.68-78.

Isomichi, Y., 1980. *Information Theory*. Corona publishing co., ltd., Tokyo Japan, pp.6-34

Haralick, R. M., et al., 1973. Textural Features for Image Classification. *IEEE Trans. Systems, Man, and Cybernetics*, SMC-3(6), pp.610-621.

Yamada, T. and Hoshi, T., 2002a. Expansion of Feature Vector Introduced Rajski Distance for Land-cover Classification using Polarimetric SAR Image Data. *J. Japan Society of Photogrammetry and Remote Sensing*, 41(6), pp.14-26.

Yamada, T. and Hoshi, T., 2002b. Application of Rajski distance to pseudocolor synthesis for recognition of dual polarization SAR image data. *Proc. 35th Conference of The Remote Sensing Society of Japan*, pp.127-128.