Commission No. VII

Takao Yoshida, H. Kurioka,
Y. Matsuo, S. Kasahara

Kajima Corporation,
Information Processing Center
2-7, Motoakasaka, 1-chome,
Minato-ku, Tokyo, Japan, 107

Post Processing for Classified Land Use

Abstract:

When remote sensing data obtained by a low flight is used to grasp the pattern of land use, the classified results must be generally corrected because housing areas, fields, roads and woods are so intermingled.

In this research, a qualitative smoothing method for the classified results will be proposed by using a probability based majority-rule.

The method aims to realize ① Clear specification of zone boundary from the available point data, ② Improvement of the accuracy in the data by using the surrounding element information, and ③ Removal of noise contained in the data.

The application of the method to a rural area will be given, varying various parameters required in the present proposed method.

1. Introduction

Remote sensing (R/S) techniques have become an inevitable tool for various fields such as environmental assessment and feasibility studies on land development.

However, when it is going to be applied in Japan, there arises some difficulty. The LANDSAT multi-spectral scanner (M.S.S.) data taken by artificial satellites can not be used directly. This is because forests, fields, rivers, roads and houses are so intermingled in Japan that the satellite M.S.S. data are not enough in their accuracy when used for actual purposes. Therefore, we have to use the air borne M.S.S. data taken by a low flight. Those data can give sufficiently detailed information on the present land use with dividing a desired area into many small unit areas.

The quality of information can be improved considerably when aeroplanes are used to gather data. However, since many and too unnecessarily detailed information such as high frequency noise are also included, we can not use the data directly in this case too. Special smoothing technique will be introduced here to eliminate such unnecessary informations. Some procedures have already been proposed for noise elimination and clear boundaries enhancement (1, 2). However, a qualitative smoothing method will be given in this paper, extracting representative surface characteristic as exactly as possible. A few examples are also given by using the proposed smoothing method.

2. Algorithm of the Qualitative Smoothing

The classified results of R/S data are expressed with the norminal scale variables that can indicate what class is good for each point information (i.e., pixel information). Broadly speaking, there are two ways in data smoothing. One is the already-established method for quantitative variables (3). The other is the present method for qualitative variables.

The qualitative smoothing proposed here is based on the information levels in a pixel and the surrounding ones. So-called "majority rule" is applied to uniquely determine the content in each pixel. In order to estimate the reasonable content in each pixel, the contents in the surrounding pixels are first considered. If considerable amount of surrounding pixels are observed to show the same information, then the information in the center pixel is altered with the surrounding one. The actual process will be described in the following.

Majority Rule

Let us take an example to explain the majority rule. Fig. 1(a) shows a pixel and its surrounding region (called mask). The size of surrounding region is given by a parameter a in that figure. The first block picture in Fig. 1 (a) shows the case when a=1, i.e., eight adjacent elements exist around a pixel. On the other hand, the second one is the case when a=2. Since we can define the surrounding region, the majority rule can be stated as follows.

$$Kij = \begin{cases} K & \text{if } M < Mo \\ Ko & \text{otherwise} \end{cases} \qquad (1)$$

where the symbol Kij expresses the class that the pixel at (i, j) point must take. The constants K and Ko denote the classes of the pixel in question and its adjacent pixels respectively. The constant M is the total number of pixels in the mask whose class are the same. The constant Mo is a certain threshold for the number M. In a word, eq. (1) indicates that if more than Mo elements are found such that whose class are all Ko, the class in the (i, j) pixel is replaced with the surrounding one. The threshold Mo considered in this paper is given by

$$Mo = 2a^2 + 2a + 1 \qquad (2)$$

Note that the threshold defined by eq. (2) depends on the mask size a.

Fig. 1 (b) shows a schematic example of the present majority rule eq. (1). Suppose we are given raw class data as shown in the left picture of Fig. 1 (b). Then we observe that there are five pixels whose class are A. On the other hand, the threshold number Mo in this case is given by Mo=5 since the unit mask size (a=1) case is considered. Therefore, according to the present majority rule eq. (1), the class of the pixel at the center is changed to A from B. The obtained result is shown in the right picture of Fig. 1 (b)

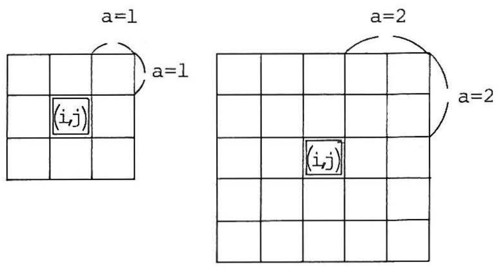
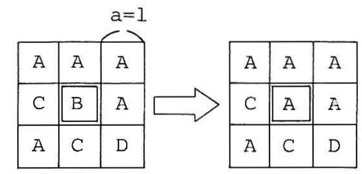


Fig. 1 (a) Neighbourhood of Pixel (i,j) defined by the Mask size a.



Pixel counts of Class A = 5 a=1
" B = 1
" C = 2 $2a^2+2a+1=5$
" D = 1

Fig. 1 (b) Schematic example of the "Majority rule" Smoothing

However, since all raw data can be considered as probabilistic quantities, we have to investigate the reliability of given data before applying the afore-mentioned majority rule. The reliability of data contained in each pixel is studied based on its posterior probability Px. The probability Px can be defined by

$$Px = \frac{P_J}{\sum_{i=1}^{k} Pi} \qquad (3)$$

where the notation Pi is the probability that a pixel takes the i^{th} class and the number k expresses the number of all conceivable classes. The notation P_J, on the other hand, is the maximum probability among all possible P_i (i=1...k) for the pixel considered.

The probability Pi for the i^{th} class can be calculated by assuming that the probability distribution is Gaussian. It is given by the following formula:

$$Pi=(2\pi)^{-\ell/2} \left| \mathbb{K}i \right|^{-1/2} \exp\left[-\frac{1}{2} (\mathbf{x} - \mathbb{M}i)^{T} \mathbb{K}i^{-1} (\mathbf{x} - \mathbb{M}i) \right] \qquad (4)$$

where $\mathbb{K}i$ = covariance matric

$\mathbb{M}i$ = average value vector

ℓ = channel number

The larger value of the posterior probability Px implies that the present data is more reliable than the ones with less Px. Therefore, we can suggest an elaborate and probability-based majority rule (a devised majority rule) as follows.

(1) If the calculated posterior probability for a pixel is larger than a certain threshold probability Pc, then the information contained in the pixel can be judged probable so that no smoothing is required with referring surrounding data.
(2) However, on the contrary, if the condition $Px \leqq Pc$ occurs, the majority rule already described must be applied.
Such devised majority rule will be examined for actual problems in the next section.

3. Application

Parametric studies on the qualitative smoothing method in the previous section will be performed to see its representative natures. The following four points will be investigated in the present parametric studies.

(a) How the number of classes affects the smoothing results.
(b) How the repetition of smoothing influences the results.
(c) What threshold value should be taken for the posterior probability in each pixel and
(d) Whether the mask size taken influences the final results.

The R/S data used are for a representative rural area in Japan. The data were taken by an aircraft flight above 3550 ft from the ground. The covered area (test area) consists of 240x240 pixels, in which one pixel occupies about 4mx4m actual area of ground. The test area contains mild slopes of ground and coastal lines.

In table 1, the characteristics of wave length contained in the original M.S.S. data are shown. The twenty six different representative groups are considered for the test area according to various land use such as roads, fields and etc. (see Table 2). In order to clearly express the computed results, the five larger groups are prepared after combining some of the 26 groups (see the details in Table 2). In the Table 3, the results of supervised-classification are shown, after combining some of the categories in Table 2.

Table 1 Spectral bands of M.S.S.

Channel No	Test channel	Wave - Length (μm)
0		0.30 - 0.35
1		0.35 - 0.40
2		0.40 - 0.45
3	O	0.45 - 0.50
4		0.50 - 0.55
5	O	0.55 - 0.60
6		0.60 - 0.65
7	O	0.65 - 0.70
8		0.70 - 0.80
9	O	0.80 - 0.90
10	O	0.90 - 1.10
11	O	10.5 - 12.50
12		4.30 - 5.50
13		4.50 - 4.90

Table 3 Classification of Test Area

Class	Category	Land Use Results
1	Road	6.47 (%)
2	Bare Soil	20.70
3	Vegetation	65.30
4	Building	1.35
5	Sea	6.18

Table 2 Legend of Land Use Category

5 Classes	26 Classes	Category
1	1	Road (Bare)
	4	Road (Asphalt)
2	2	Bare Soil
	3	"
	6	"
3	5	Crop Field
	7	Paddy Field
	8	Radish
	9	Woods
	10	Orchard
	12	Woods in Shadow
	13	Orchard in Shadow
	14	Grass in Shadow
	15	Woods in Shadow
	16	Orchard in Shadow
	17	Orchard
	18	Grass
	19	Weeds
	20	Orchard
	21	Woods
	22	Woods in Shadow
	23	Orchard in Shadow
	24	Woods in Shadow
4	11	Artifical Building
	26	Artifical Building
5	25	Sea

Four different cases shown in Table 4 were considered varying the values of parameters such as the total number of classes, the iteration number, the threshold value Pc as well as the mask size.

Table 4 Case Study

Case No.	Objective	Total Classes	Iteration N	Threshold Pc	Mask Size a
1	Effects of Iteration & Total Classes	26 5	1~10	1.0	1
2	Effects of Mask Size & Total Classes	26 5	2.3	1.0	1~3
3	Influence of Threshold	26	1	0.5 0.793	1~3
4	Influence of Threshold & Iteration	5	1~5	0.9 1.0	1

Case 1 example was used to investigate the effect of smoothing when the numbers of classes and repetition are varied. See the actual parameters employed for the Case 1 in Table 4. The following observations could be made in the present case.

(a) When larger number of classes are used, remarkable smoothing effect could not be observed. This might come from the fact that since there are too many classes in the mask, the majority rule can not work.

(b) Repetition of smoothing widened the area for vegetation. On the other hand, the point or line informations like buildings and roads tended to be disregarded. The required number of repetition to convergence was almost three.

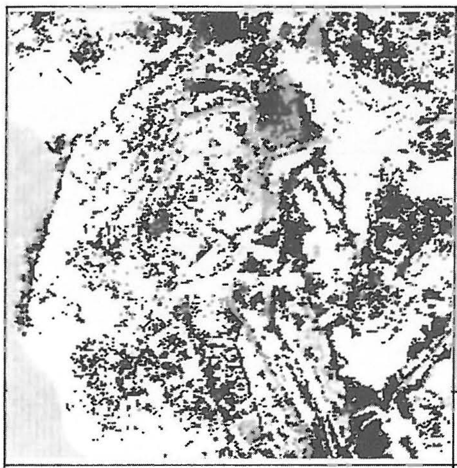
The above tendency can be seen from Tables 5 and 6 as well as from Fig. 2.

Table 5 Effects of Iteration N Classes: 26 a=1 Pc=1.0

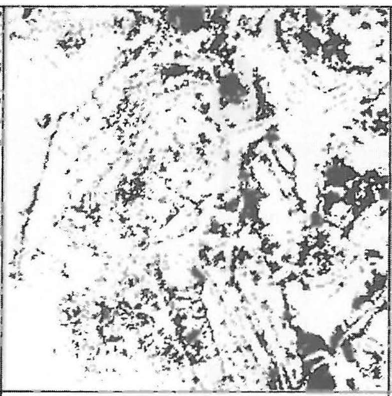
Class	Category \ Iteration	0	1	2	3	4	5	6	7	8	9	10
1	Road	6.47(%)	6.68	6.75	6.75	6.77	6.78	6.80	6.80	6.80	6.81	6.81
2	Bare Soil	20.70	20.79	20.82	20.74	20.69	20.66	20.65	20.63	20.64	20.63	20.63
3	Vegetation	65.30	65.29	65.27	65.39	65.41	65.49	65.45	65.48	65.46	65.47	65.47
4	Building	1.35	1.02	0.92	0.89	0.87	0.85	0.85	0.84	0.84	0.84	0.84
5	Sea	6.18	6.23	6.24	6.24	6.24	6.24	6.24	6.24	6.24	6.24	6.24

Table 6　Effects of Iteration N　Classes: 5 a=1 Pc=1.0

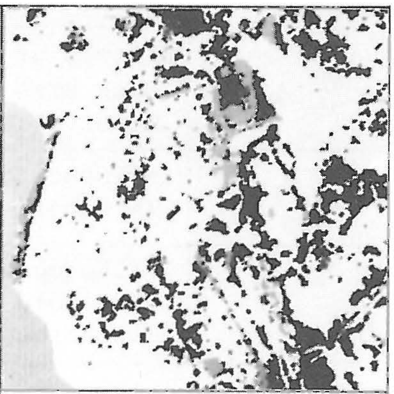
Class	Category \ Iteration	0	1	2	3	4	5	6	7	8	9	10
1	Road	6.47	6.10	5.88	5.76	5.69	5.64	5.60	5.58	5.56	5.56	5.56
2	Bare Soil	20.70	16.62	15.42	14.72	14.25	13.90	13.70	13.54	13.44	13.35	13.29
3	Vegetation	65.30	70.46	72.02	72.90	73.48	73.90	74.15	74.34	74.46	74.56	74.63
4	Building	1.35	0.59	0.46	0.39	0.36	0.33	0.32	0.31	0.31	0.31	0.30
5	Sea	6.18	6.22	6.23	6.23	6.23	6.23	6.23	·6.23	6.23	6.23	6.23



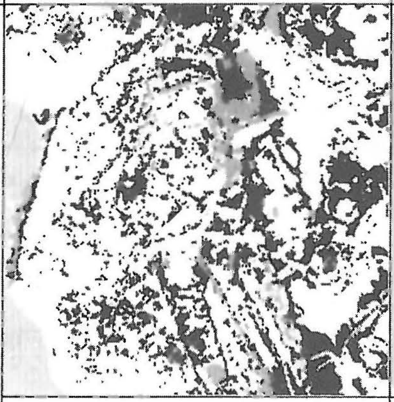
Unsmoothed (n=0)



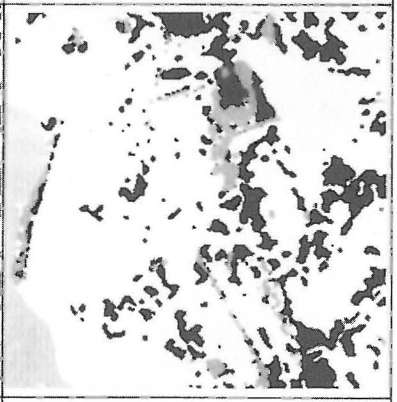
n=1　Classes: 26



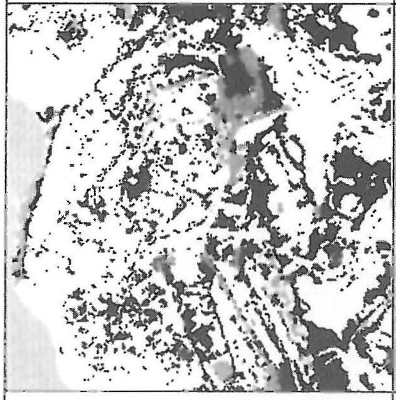
n=1　Classes: 5



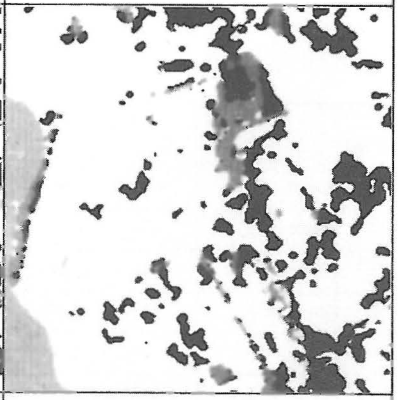
n=2　Classes: 26



n=3　Classes: 5



n=3　Classes: 26



n=5　Classes: 5

Fig. 2　Effects of Iteration No.

Mask size　a=1

Threshold　P_C=1.0

The case 2 was used to see the effect of the mask size a and repetition number in smoothing. The obtained results are listed in Table 7 and shown in Fig. 3. As can be seen from Fig. 3, larger vegetation areas was obtained as the mask size became large. When the 26 supervised groups were used, no alteration was observed in the smoothed results even when the mask size was changed. However, when the five groups were employed, the change of the mask size considerably influences the final output. Note that when unnecessarily large mask size is taken, it was found that there will be a slight danger of distorting the final results.

Table 7 Effects of Mask Size & Interation of Post Processing Pc = 1.0

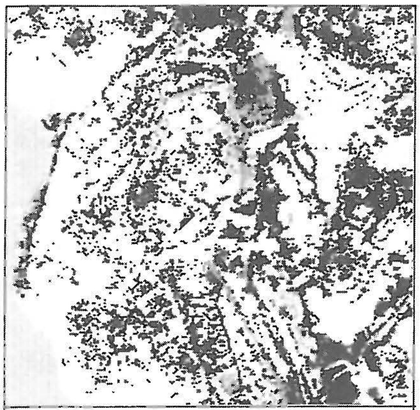
		26							26 → 5					
	a		2			3			2			3		
Class	N / Category	0	1	2	3	1	2	3	1	2	3	1	2	3
1	Road	6.47(%)	6.68	6.79	6.81	6.68	6.78	6.83	5.64	5.33	5.11	5.33	4.83	4.71
2	Bare Soil	20.70	20.58	20.43	20.34	20.58	20.54	20.50	14.04	12.53	11.73	12.42	10.80	10.32
3	Vegetation	65.30	65.43	65.54	65.65	65.35	65.34	65.42	73.57	75.59	76.70	75.49	77.83	78.54
4	Building	1.35	1.09	0.98	0.90	1.16	1.05	0.96	0.51	0.32	0.21	0.52	0.28	0.17
5	Sea	6.18	6.24	6.26	6.26	6.26	6.28	6.30	6.23	6.24	6.24	6.24	6.26	6.26

The case 3 problem was investigated to see the effects of varying the threshold value Pc and the mask size a. Two threshold values were considered; Pc=0.5 and 0.793. The latter values is the average value of posterior probabilities of all pixels in the test area. The obtained result is shown in Table 8. The smoothing was not influenced for the change of the threshold values (see Table 8).

Table 8 Effects of Threshold (Pc) & Mask Size (a)

Threshold Pc			0.5			0.793		
Iteration N		0	1			1		
Class	Category / a	0	1	2	3	1	2	3
1	Road	6.47(%)	6.47	6.49	6.61	6.60	6.60	6.61
2	Bare Soil	20.70	20.75	20.72	20.66	20.76	20.66	20.66
3	Vegetation	65.30	65.30	65.30	65.28	65.27	65.29	65.28
4	Building	1.35	1.30	1.32	1.26	1.17	1.24	1.26
5	Sea	6.18	6.18	6.18	6.21	6.20	6.20	6.21

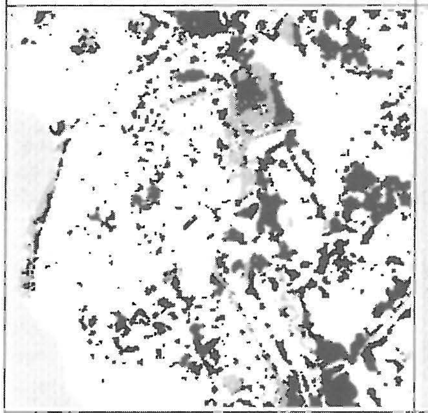
Total Classes: 26



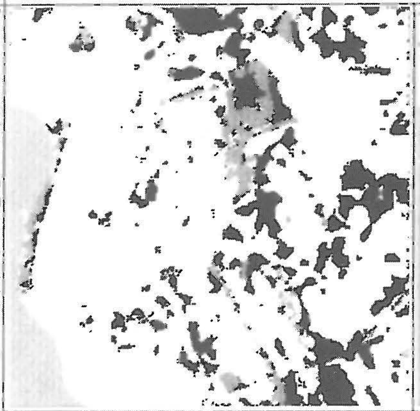
Unsmoothed

Fig. 3 Effects of Mask Size a & Iteration n

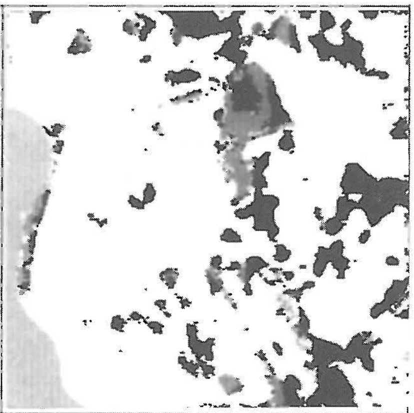
(Classes: 5, Threshold Pc=1.0)



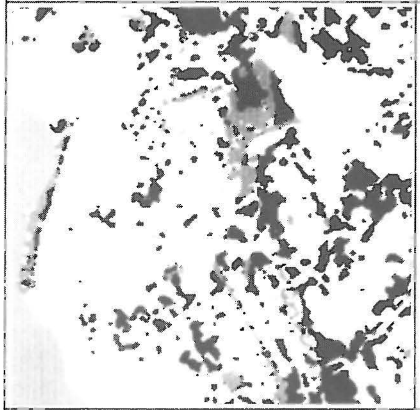
a=1, n=1



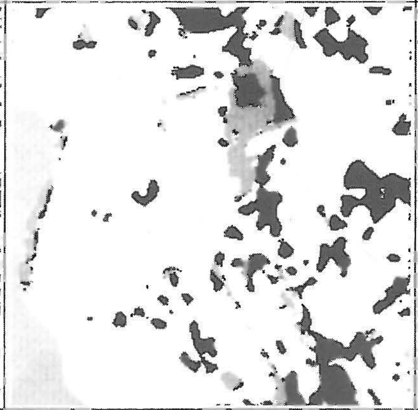
a=2, n=1



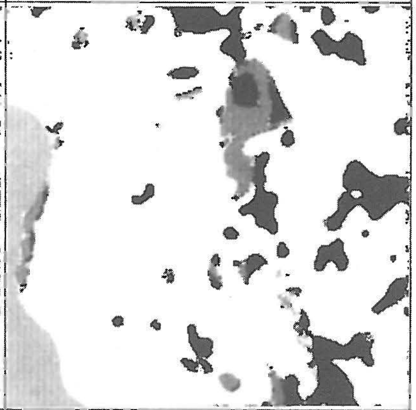
a=3, n=1



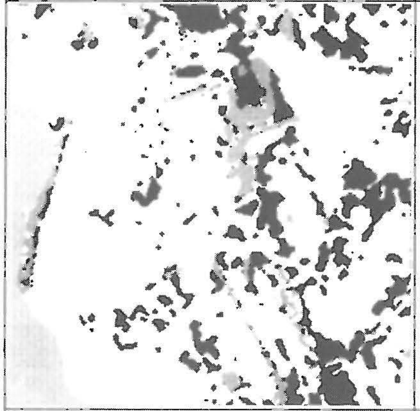
a=1, n=2



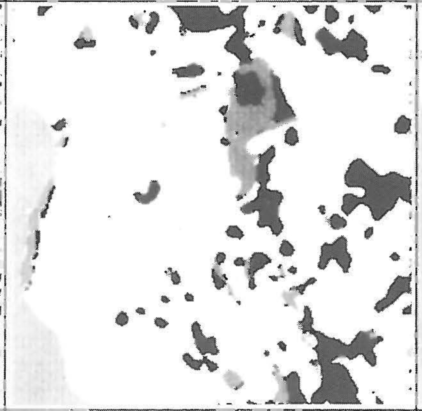
a=2, n=2



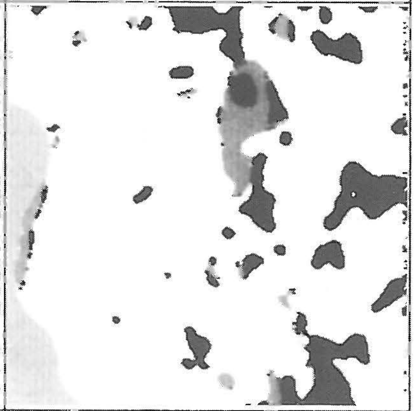
a=3, n=2



a=1, n=3



a=2, n=3

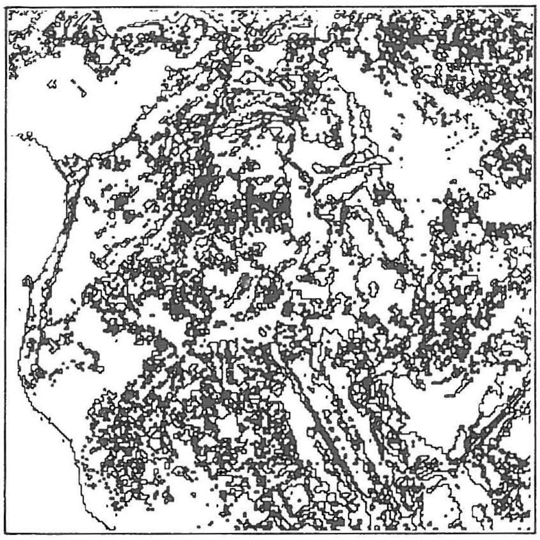


a=3, n=3

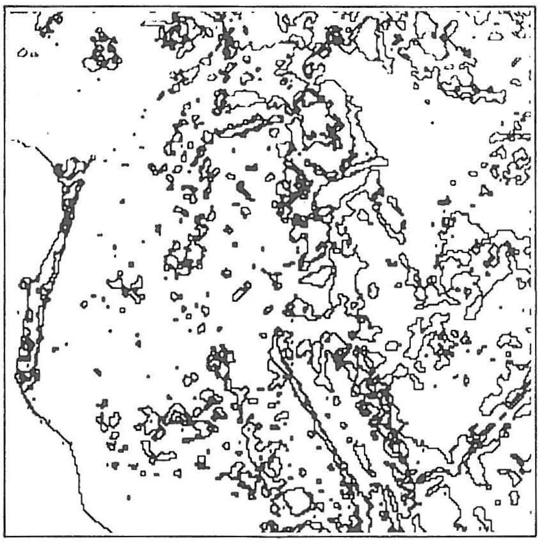
Fig. 4 Influence of Threshold Pc

Classes: 5, a=1

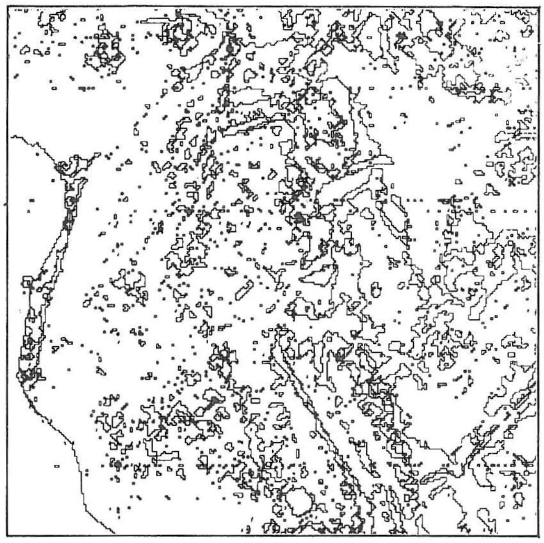
All figures presented here are drawn by X-Y plotter to clearly show complex boundaries. Note that in original figures, the boundaries were clearly shown with using different colors.



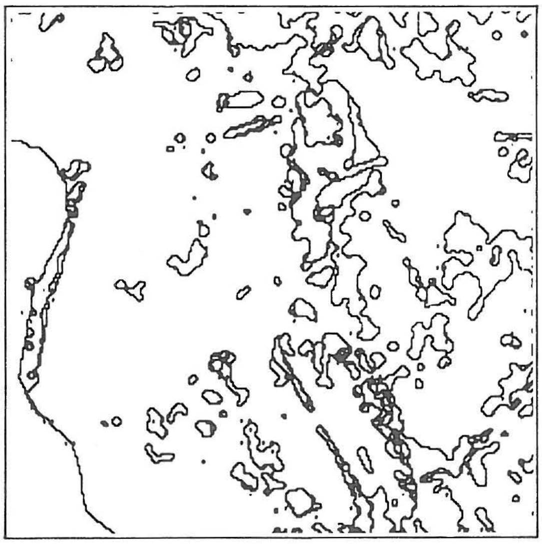
(a) Unsmoothed



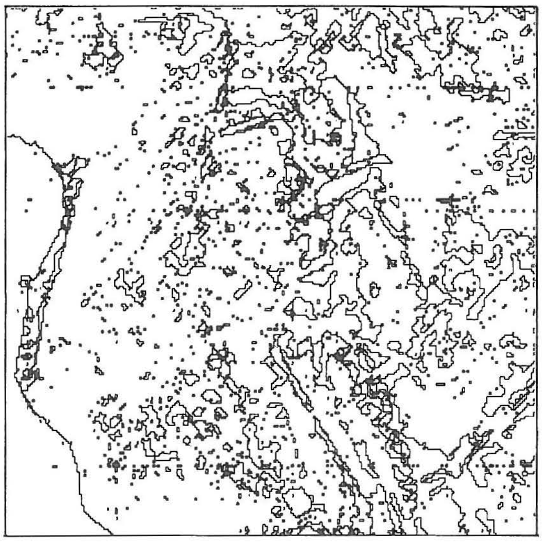
(b) Pc=1.0 n=1



(d) Pc=0.9 n=1



(c) Pc=1.0 n=5



(e) Pc=0.9 n=5

The case 4 uses the parameters: a=1, P_C=0.9 and n=1 to 5. Fig. 4 (a) shows the land use before smoothing. Figs. 4 (b) and (c) shows the results for n=1 and n=5 cases, respectively. Both results were obtained with P_C= 1.0. The corresponding results are shown in Figs. 4 (d) and (e) for threshold value 0.9. Original land use in Fig. 4 (a) does not exhibit clear distinction between bare land and vegetation area. However, when P_C=1.0 and n=1 are taken like in Fig. 4 (b), bare land and vegetation area could be clearly identified, while leaving the detailed information on roads and buildings. In Fig. 4 (b), it can be observed that the scanning noise could also be removed. Fig. 4 (c) gave the result that can assist our gross understanding on the land use, although some detailed information are eliminated. Figs. 4 (d) and (e) show that the iteration of smoothing is not a major factor in improving the quality of final outputs. In any event, it might be said that all obtained results Figs. 4 (b) to (e) are better than the original one Fig. 4 (a) in the sence that we can easily grasp the land use by using them.

4. Conclusions

The following conclusions could be drawn from the present research.

(a) When too many classes exist in a test area, significant smoothing effect can not be expected.
(b) However, when smaller number of classes is used for a test area, the present smoothing technique can exhibit its effectiveness.
(c) Only one iteration is sufficient to remove the scanning and random noises.
(d) Of course, actual required iteration number depends on the demand for final output quality. However, at most 3 to 5 iterations are enough.
(e) Similarly, the appropriate mask size depends on how the final output is used. But actually, only a=1 case has better be considered.
(f) If a smaller value than the average posterior probability is used as the threshold, any significant smoothing effect can not be anticipated. Further researches would be necessary for determining its optimum value.

As a final comment, it might be concluded that the present method will be particularly useful in obtaining practical results.

Acknowledgement

The authors would like to express their sincere thanks to those who kindly assist them to finish this research. Notable persons are Prof. S. Murai in University of Tokyo, and Mr. M. Shoji and Dr. K. Imai in Kajima Corporation. They also owe to the staffs in the Japan Foundation for Shipbuilding Advancement.

References

1. M. Nagao and T. Matsuyama: "Edge Perserving Smoothing" 4IJCPR, 1978
2. A.L. Steven, W. Zucker, A. Rosenfeld: "Iterative Enhancement of Noisy Images", IEEE. Trans. vol. SMC-7 No.6, 1977
3. F. Tomita, S. Tsuji, "Extraction of Multiple Regions by smoothing in Selected Neighborhoods" IEEE. Trans. vol. SCM-7, Feb. 1977