

MAXIMUM LIKELIHOOD ESTIMATION OF CATEGORY PROPORTIONS AMONG MIXELS

Masao MATSUMOTO:

Faculty of Engineering, Nagasaki University, Japan.

Yasunori TERAYAMA & Kohei ARAI:

Faculty of Science and Engineering, Saga University, Japan.

Abstract

To make an accurate retrieval of the proportion of each category among mixed pixels (Mixel's) of a remotely sensed imagery, a maximum likelihood estimation method of category proportion is proposed. In this method, the observed multispectral vector is considered as probability variables along with the approximation that the supervised data of each category can be characterized by normal distribution. The results show that this method can retrieve more accurate proportion of each category among Mixel's than the conventional generalized inversion method. And a index that can estimate the degree of error in each category is proposed.

Key words: Mixel, Category proportion, Maximum likelihood estimation, Non-linear optimization.

1. INTRODUCTION

Generally, each pixel of a remotely sensed imagery contains the multiple information from the different categories, such pixel is called Mixel (mixed pixel). Some methods to estimate the category proportions among Mixel's are proposed (Hallum, 1972), (Inamura, 1987), (Ito and Fujimura, 1987). Such methods used generalized inversion method which is based on least square criterion, and supervised data of each category is characterized by only the average or mode of multispectral vector of the category. In a general case, even the supervised data fluctuated because of sampling effect, terrain effect, shadow, selection of physically undefinable category such as urban area, and so on. Such fluctuation makes the error of the estimated category proportion based on previous methods (Inamura, 1987). Moreover, in the non-determine case (the number of categories is greater than the number of multispectral bands), the estimation error from the previous methods become larger than the result from over-determine case because of lack of the consideration of the observation error.

In this study, To make more accurate estimation of category proportion among Mixel's, maximum likelihood estimation of category proportion is proposed. This method can explicitly contain the effect of observation error. And in this method, the supervised data of each category can be represented by the average and variance, so this thing gives more freedom to select the categories to the users.

In this paper, at first the maximum likelihood estimation methods is described. Next, the verification of the method based on the simulated Mixel data and the comparison between the result from the proposed method and the generalized inversion methods are made. And at last, the index to show the estimation error is proposed.

2. MAXIMUM LIKELIHOOD ESTIMATION

The multispectral vector (number of band is M) of the Mixel which contain the information from N categories is shown in Eq. (1),

$$\begin{aligned} \mathbf{I} &= \mathbf{A} \cdot \mathbf{B} + \boldsymbol{\varepsilon} \\ \mathbf{I} &= [I_1, I_2, \dots, I_M]^t, \\ \mathbf{B} &= [B_1, B_2, \dots, B_N]^t, \\ \mathbf{A} &= \begin{bmatrix} A_{11} & \dots & A_{1N} \\ \vdots & \vdots & \vdots \\ A_{M1} & \dots & A_{MN} \end{bmatrix}, \\ \boldsymbol{\varepsilon} &= [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M]^t, \end{aligned} \quad (1)$$

where I_i , B_j and A_{ij} are observed data in i -th band, category proportion of j -th category and supervised data of j -th category in i -th band, respectively and superscript t means transpose. In this study, on the basis of following approximations, the observed multispectral vector can be considered as a probability variable.

Approximations:

- (1) A_{ij} is expressed by normal distribution that average is \bar{A}_{ij} and variable is σ_{ij}^2 .
- (2) The supervised data of each category is independent.
- (3) ε_i is independent from the proportions and it is expressed by the normal distribution that average is 0 and variance is $\sigma_{\varepsilon_i}^2$.
- (4) The observation error of each band is independent.

The correlation matrix R of each spectral band is expressed as follows,

$$\begin{aligned} \mathbf{R} &= \begin{bmatrix} 1 & \dots & \rho_{1M} \\ \vdots & \vdots & \vdots \\ \rho_{1M} & \dots & 1 \end{bmatrix}, \\ \rho_{kl} &= \frac{COV_{kl}}{\sqrt{VAR_k \cdot VAR_l}}, \quad (k, l = 1, \dots, M) \end{aligned} \quad (2)$$

where ρ_{kl} is correlation coefficient between band k and l . And VAR and COV mean integrated variance shown in Eq. (3) and integrated covariance, respectively.

$$\begin{aligned} VAR_k &= \mathbf{B}^t \cdot S_k \cdot \mathbf{B} \\ S_k &= \text{diag}(\sigma_{k1}^2, \dots, \sigma_{kM}^2) \end{aligned} \quad (3)$$

From above consideration, the probability $P(\mathbf{I}; \mathbf{B})$ when the proportion vector \mathbf{B} is occurred and multispectral vector \mathbf{I} is observed is expressed as Eq. (4),

$$P(\mathbf{I}; \mathbf{B}) = \frac{1}{(2\pi)^{M/2} \sqrt{\det(\mathbf{Z}')}} \exp \left[-\frac{1}{2} (\mathbf{I} - \bar{\mathbf{A}}\mathbf{B})' \mathbf{Z}'^{-1} (\mathbf{I} - \bar{\mathbf{A}}\mathbf{B}) \right]$$

$$\mathbf{Z}' = \begin{bmatrix} \text{VAR}_1 & \dots & \text{COV}_{1M} \\ \vdots & \ddots & \vdots \\ \text{COV}_{1M} & \dots & \text{VAR}_M \end{bmatrix} + \text{diag}[\sigma_{\varepsilon 1}^2, \dots, \sigma_{\varepsilon M}^2] \quad (4)$$

$$\bar{\mathbf{A}} = \begin{bmatrix} \bar{A}_{11} & \dots & \bar{A}_{1N} \\ \vdots & \ddots & \vdots \\ \bar{A}_{MI} & \dots & \bar{A}_{MN} \end{bmatrix}$$

The category proportion \mathbf{B} has following constraints,

$$0 \leq B_j \leq 1, \quad (j=1, \dots, N),$$

$$\sum_{j=1}^N B_j = 1. \quad (5)$$

From the concept of maximum likelihood estimation (Rodgers, 1976), the proportion which maximize $P(\mathbf{I}; \mathbf{B})$ under the above constraints is the solution. This problem can be solved by constrained non-linear optimization technique (Konno and Yamashita 1978).

3. VERIFICATION

To verify the maximum likelihood estimation, the comparison between the result from the proposed method and conventional method (generalized inversion method) based on the simulated Mixel data.

3.1 Simulated Mixel Data

3.1.1 Classified Categories and the Supervised Data In this study, the classified categories are selected from sub-scene (380 pixels and 310 lines) of LANDSAT-5 TM data of HAKONE area of Japan which acquired at 6. June '87. As the classified categories, 5 categories of residential area, bare soil, grass land, broad leaf tree and needle leaf tree are selected from vegetation map and land use map (Amano, Furuya and Ishikawa, 1990). Average and variance of each category are calculated and shown in Table 1 and 2, respectively. These data are use as the supervised data for the estimation. The calculation of $P(\mathbf{I}; \mathbf{B})$ of Eq. (4) is numerically sensitive to the correlation matrix R of each band, Table 3 shows the correlation matrix R calculated from whole image data except band 6 (thermal IR band). In this study, because of the high correlation of each band, the first and second principal components (stretched into 8 bits) of the imagery are used as the multispectral data. Table 4 and 5 show that eigenvalue and eigenvector of each principal component and the average and variance of the supervised data in the principal components, respectively.

3.1.2 Generation of Simulated Mixel Data The 100 simulated Mixel data are generated in the following manner.

(1) Category Proportion

By using uniform random number from 0 to 1, proportion of each category is generated as in Eq. (6),

$$B_k = \frac{r_k}{\sum_{i=1}^N r_i}, \quad (k=1, \dots, N). \quad (6)$$

where r_k is uniform random number from 0 to 1.

(2) Supervised data

In this method, the supervised data of each category (each component of matrix A of Eq. (1)) is allowed to fluctuate, so by using random number in multivariate normal distribution (Takane, 1980) which average and variance are those of each category and band.

(3) Observation Error

Also the observation error of each band is generated by normal distribution which average is 0 and variance is 1/12 [count²]. This variance correspond to quantization error.

(4) Mixel data

After above simulated data from (1-3) and Eq. (1), the 100 Mixel data is generated.

3.2 Generalized Inversion Method

As a conventional category proportion method, the generalized inversion method (Inamura, 1987) is used in this study to compare with the proposed method. In this case the number of category (5 categories) is greater than the number of band (2 bands), so this problem is non-determine. And this method, the non-negative constraint is not considered.

4. RESULT

Eq. (4) is solved by the grid search method with 1/64 of step width (Kanno and Yamashita, 1978). From the results of two methods, two kinds of root means square error (RMSE) are calculated, one is total RMSE: $RMSE_T$ and the other is maximum RMSE: $RMSE_M$,

$$RMSE_T = \sqrt{\frac{1}{N \cdot S} \sum_{l=1}^S \sum_{j=1}^N [B_j^{(e)} - B_j^{(a)}]^2}, \quad (7)$$

$$RMSE_M = \max_{l=1}^S \sqrt{\frac{1}{N} \sum_{j=1}^N [B_j^{(e)} - B_j^{(a)}]^2},$$

where $B_j^{(e)}$, $B_j^{(a)}$ and S are estimated proportion of j-th category, actual proportion of j-th category and number of the Mixel data (100), respectively. Table 6 shows these RMSE's from the two methods. the result from maximum likelihood estimation shows the better result correspond to that from generalized inversion method. This is because that in maximum likelihood estimated, supervised data is represented by 2 parameters (average and variance), while in generalized inversion method it is represented by only one parameter (average).

5. MAXIMUM ERROR OCCURRENCE PROBABILITY

To predict the degree of estimation error in each category, the index named maximum error occurrence probability between category k an category l: $P_e(\mathbf{I}; k, l)$ is proposed as follows.

$$P_e(I; k, l) = P[I_{\text{pure}}(k); B_{\text{pure}}(l)], \quad (8)$$

where $B_{\text{pure}}(l)$ and $I_{\text{pure}}(k)$ are the proportion which k -th category is 1 and rest are 0 and average multispectral vector of k -th category, respectively. The maximum error occurrence probability show the maximum probability which occurs in the case of composition of 2 categories. So if this value is larger, it can be assumed to be occurrence of larger estimation error. In Table 7, the maximum error occurrence probabilities calculated from previous 5 categories are shown. In the combination of grass land and bare soil and the combination of grass land and broad leaf tree, the probabilities are relatively large, so it can be predicted that relative large estimation error occur in the estimation of the proportion of grass land. This is because of the relatively large variance of this category. Fig. 1 shows the histogram of maximum RMSE. As it can be expected, in the estimation of the proportion of grass land, the frequency of maximum RMSE is the largest. To make more accurate estimation, it has to reduce the maximum error occurrence probabilities among each category.

6. CONCLUSIONS

The previous results lead to the following conclusion. The observed multispectral vector is considered as probability variables along with the approximation that the supervised data of each category can be characterized by normal distribution. The results from simulated Mixel data show that this method can retrieve more accurate proportion of each category among Mixel's than the conventional generalized inversion method. And the maximum error occurrence probabilities can be a good index which can predict the estimation error in each category.

7. REFERENCES

- *Amano, M., K. Furuya and T. Ishikawa. 1990. Classification of forest type based on satellite data. Proc. 1990 Autumn Conf. of JSPRS, Nagasaki Japan, 73-78.
- Hallum, C. R. 1972. On a model for optimal proportions estimates for category mixture, Proc. 8th. Intl. Symp. on Remote Sens. of Env., ERIM, Vol. 2, pp. 951-958.
- *Inamura, M. 1987. Analysis of remotely sensed image data by means of category decomposition. J. of IECSE., J70-C(2), 241-250.
- *Ito, T, and S. Fujimura, 1987. Estimation of cover area of each category in a pixel by pixel decomposition into categories. Trans. of SICE., 23(8), 20-25.
- *Konno, H. and H. Yamashita, 1978. Non-linear programming. Nikka-Giren Pub. Co. pp. 354.
- Rodgers, C. D. 1976. Retrieval of atmospheric temperature and composition from remote measurement of thermal radiation. Rev. of Geophys. and Space Phys., 14(4), 609-623.
- *Takane, Y. 1980. Multidimensional scaling. Tokyo-Daigaku-Shuppan-Kai Pub. Co. pp.332.

(*: Original text written in Japanese.)

Table 1 Average of each category

Band	Resid.	Bare	Grass	Broad	Neddle
1	122.60	143.98	109.33	98.52	100.40
2	50.15	69.79	47.69	38.89	41.38
3	50.27	80.74	43.68	32.44	33.86
4	68.25	87.73	119.52	81.73	137.15
5	75.19	107.64	113.34	60.23	95.32
7	39.76	52.41	40.69	19.86	28.22

Table 2 Variance of each category

Band	Resid.	Bare	Grass	Broad	Neddle
1	64.24	222.61	17.07	10.30	18.29
2	14.73	96.21	60.45	1.19	2.22
3	29.00	186.40	15.14	1.46	2.08
4	75.49	136.85	193.23	121.76	228.44
5	61.95	246.88	104.45	59.84	110.08
7	40.94	49.05	29.53	53.45	9.04

Table 3 Correlation of each band

Band	1	2	3	4	5	7
1	1.00e+0	7.61e-1	8.43e-1	-5.29e-2	2.21e-1	5.20e-1
2	7.61e-1	1.00e+0	9.22e-1	3.92e-1	6.64e-1	8.50e-1
3	8.43e-1	9.22e-1	1.00e+0	1.34e-1	4.64e-1	7.56e-1
4	-5.29e-2	3.92e-1	1.34e-1	1.00e+0	8.91e-1	6.12e-1
5	2.21e-1	6.64e-1	4.64e-1	8.91e-1	1.00e+0	8.73e-1
7	5.20e-1	8.50e-1	7.56e-1	6.12e-1	8.73e-1	1.00e+0

Table 4 Eigenvalue and eigenvector of the principal components

	PC1	PC2	PC3	PC4	PC5	PC6
	5.47	0.49	0.03	0.01	0.00	0.00
BAND1	0.41	-0.33	0.00	-0.83	0.13	-0.08
BAND2	0.42	-0.26	-0.27	0.32	-0.35	-0.68
BAND3	0.42	-0.29	-0.12	0.19	-0.42	0.72
BAND4	0.36	0.75	-0.53	-0.12	0.07	0.06
BAND5	0.41	0.39	0.79	0.02	-0.23	-0.09
BAND7	0.42	-0.16	0.09	0.39	0.79	0.06
Cont.	0.91	0.08	0.01	0.00	0.00	0.00

Table 5 Average and variance in PC's

	Resid.	Bare	Grass	Broad	Needle
A PC1	97.8	162.4	127.3	60.9	107.8
V PC2	62.2	135.1	162.0	100.9	187.7
V PC1	160.4	841.1	185.7	94.0	178.2
R PC2	309.9	681.3	430.4	329.3	586.2

Table 6 Root mean square error

	RMSE _T	RMSE _M
Maximum likelihood	0.102	0.288
Generalized inversion	0.125	0.384

Table 7 Maximum error occurrence probability

	Resid.	Bare	Grass	Broad	Needle
Resid.	7.14e-04	3.58e-07	5.10e-10	6.49e-08	5.44e-10
Bare	3.10e-13	2.10e-04	8.80e-06	2.40e-28	1.09e-08
Grass	4.96e-12	5.94e-05	5.63e-04	2.02e-16	9.68e-05
Broad	9.02e-07	1.96e-07	5.13e-11	9.05e-04	1.66e-09
Needle	4.78e-15	4.70e-06	9.42e-05	7.98e-14	4.92e-04

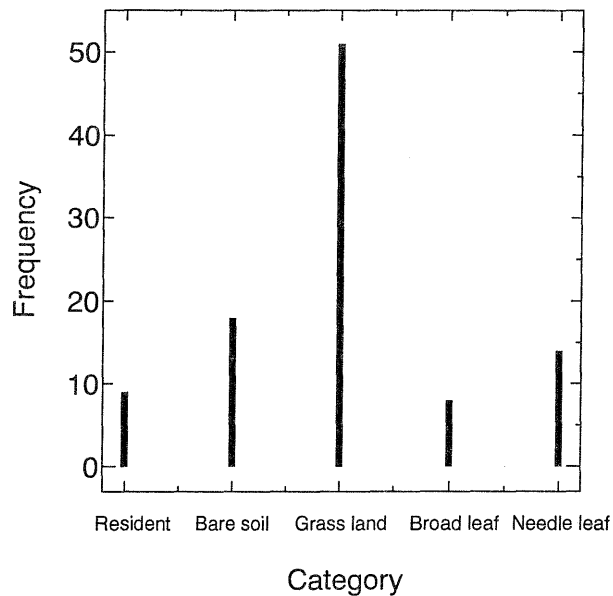


Fig. 1 Histogram of Maximum RMSE