A study on evaluating the spatial distribution of satellite image classification error

Yong-Il Kim*, Byoung-Kil Lee and Myung-Ki Chae*****

Division of Urban Eng., Seoul National University
Assistant professor, 880-7364, E-mail yik@plaza.snu.ac.kr *
Ph. D. Course, 880-7371, E-mail basil@plaza2.snu.ac.kr **
Master***

Abstract

This study overviews existing evaluation methods of classification accuracy using confusion matrix proposed by Cohen in 1960's, and proposes ISDd(Index of Spatial Distribution by distance) and ISDs(Index of Spatial Distribution by scatteredness) for the evaluation of spatial distribution of satellite image classification errors, which has not been tried yet. Index of spatial distribution offers the basis of decision on adoption/rejection of classification results at sub-image level by evaluation of distribution, such as status of local aggregation of misclassified pixels. So, users can understand the spatial distribution of misclassified pixels and, can have the basis of judgement of suitability and reliability of classification results.

1. Introduction

Since 1972, satellite images have become commercially available. Thus experts focus their research on extraction of accurate information from satellite data.[1)]

Remote Sensing can be divided into two categories. They are, quantitative field to build DEM using inclined observing satellite images and qualitative field to obtain spatial information using statistical analysis of DN(digital number) of pixels.

Until now, one of the main subjects of researches is the improvement of accuracy of satellite data analysis. Consequently, it became possible to analyze complicate spatial distribution in urban area.[2)] Especially, improvements on classification accuracy through the development of classification methods are very important to the practical use of remote sensing technology.

By this reason, researches in objective and accurate evaluating the classification results are widely performed.

Congalton and Mead(1983) quantitatively evaluated similarities between interpreters or variables of interpretation using Kappa coefficient and its normalized Z-statistic.[4)]

Rosenfield and Fitzpatrick-Lins(1986) evaluated both overall accuracy and class accuracy by introducing Kappa coefficient to evaluation of classification accuracy.[5)]

Story and Congalton(1986) evaluated overall accuracy and user's/producer's accuracy using main diagonal elements only from confusion matrix, and explained their meanings.

This study overviews existing evaluation methods of classification accuracy using confusion matrix, and proposes ISDd(Index of Spatial Distribution by distance) and ISDs(Index of Spatial Distribution by Scatteredness) for the evaluation of spatial distribution of satellite image classification errors, which has not been tried yet. Classification accuracy by item is the ratio of correctly classified pixels, and does not contain any information about the spatial distribution of misclassified pixels. So, presenting the index for spatial distribution of misclassified pixels can improve user's reliability on classification results.

To accomplish this, firstly, the evaluation methods of classification accuracy by item are studied. Secondly, evaluation methods of spatial distribution of classification errors are researched. And finally, the estimators of classification accuracy by item and index of spatial distribution of classification errors are proposed and tested.

2. Study on Evaluation Methods of Classification Accuracy by Item

2.1 Overall and User's/Producer's Accuracy

As shown in eq.(2.1), the proportion of correctly classified pixels to the entire pixels is overall accuracy. User's/producer's accuracy could be evaluated for each class using eq.(2.2) and eq.(2.3).

$$\text{Overall Accuracy} = \frac{\sum_{i=1}^{r} X_{ii}}{N} \quad (2.1)$$

$$\text{User's Accuracy} = \frac{X_{ii}}{X_{i+}} \quad (2.2)$$

$$\text{Producer's Accuracy} = \frac{X_{ii}}{X_{+i}} \quad (2.3)$$

here, r : No. of row/column of Confusion matrix

X_{ii} : No. of correctly classified pixels

$X_{i\cdot}$: Sum of the ith row of Confusion matrix

$X_{\cdot i}$: Sum of the ith column of Confusion matrix

N : No. of entire pixels

2.2 Kappa Coefficient [4),5)]

Kappa coefficient is a measure of agreement of a contingency table. And, Kappa statistics do not directly include the effects of off-diagonal entries on the accuracies of individual classification categories and overall classification. [7)]

Eq.(2.4) shows Kappa coefficient for overall accuracy, $\sum_{i=1}^{r} X_{ii}$ means the number of all accurately classified pixels and $\frac{1}{N}\sum_{i=1}^{r}(X_{i+} \times X_{+i})$ means the number of all pixels that are classified accurately by accident..

$$\tilde{k} = \frac{\sum_{i=1}^{r} X_{ii} - \frac{1}{N}\sum_{i=1}^{r}(X_{i+} \times X_{+i})}{N - \frac{1}{N}\sum_{i=1}^{r}(X_{i+} \times X_{+i})} \quad (2.4)$$

3. Research on Evaluating the Spatial Distribution of Classification Error

3.1 Measuring the Spatial Distribution of Classification Error

To measure the spatial distribution of classification error, ISDd(Index of Spatial Distribution by distance) and ISDs(Index of Spatial Distribution by scatteredness) are used. ISDd is the indexed mean distance of misclassified pixels and ISDs is the statistical indicator of the scatteredness of misclassified pixels.

3.1.1 Index of Spatial Distribution by distance

In ISDd, the mean distance is calculated by dividing the sum of distances between misclassified pixels by the number of combination of distance (eq.(3.1)).

$$\text{Mean Distance} = \frac{l_{12}+l_{13}+\cdots+l_{1n}+l_{23}+\cdots+l_{n-1n}}{{}_{n}C_{2}} \quad (3.1)$$

here,

$$l_{12} = \overline{P_1P_2}$$

$$\vdots \qquad \ddots$$

$$l_{1n} = \overline{P_1P_n} \qquad l_{2n} = \overline{P_2P_n} \qquad \cdots \qquad l_{n-1n} = \overline{P_{n-1}P_n}$$

The mean distance of spatial distribution of satellite image classification error varies with image size. Because mean distance of large image is greater than that of smaller one, comparing and evaluating two or more images are difficult. Therefore mean distance should be indexed. As shown in eq.(3.2). mean distance is divided by the representative distance of image. On this study, the representative distance is a mean value of width and height of image.

$$\text{ISDd} = \frac{\text{Mean Distance}}{(V+H)/2} \quad (3.2)$$

In evaluating ISDd for P misclassified pixels in n x n image, the number of misclassified pixels, image size and the distribution pattern of misclassified pixels affect ISDd.

A. Number of Misclassified Pixels

Relative size of misclassified pixels is more important than absolute size. If P/(n x n) closes to 0, i.e. P can be neglected, ISDd varies 0 to $ISDd_{max}$. This means that the misclassified pixels can be neglected. If, P/(n x n) closes to 1, i.e. P ≒ n x n, ISDd equals $ISDd_{even}$(≒0.52). Then, classification results should be rejected. Therefore, in order to explain the spatial distribution pattern using ISDd, P/(n x n) should be between 0 and 0.3, i.e. overall accuracy should be over 0.7.

Theoretical maximum of ISDd occurs, when we allow the duplication of misclassified pixels. In the case that the number of pixels is 4n, ISDd is formulated as eq.(3.3).

$$ISDd = \frac{(\sqrt{2}+2)n^2}{4n^2 - n} \quad (3.3)$$

As shown above, ISDd decreases as the number of pixel increases. The limit of ISDd is 0.85355.

For evenly distributed n x n misclassified pixels, calculation of $ISDd_{even}$ by numerical analysis converges on 0.5215 at 39 x 39(1,521 pixels).

B. Image Size

In real world, satellite images do not allow any duplicating pixels. So the number of misclassified pixels cannot override the number of entire pixels.

That means there is no convergent value of $ISDd_{max}$.

Theoretically, distance between two points can have maximum of width and height. In satellite images, the distance between two pixels is calculated using center points of pixel. And, the maximum value of the size of image is the multiplication of the number of pixels by the size of a pixel. Because of these, image size affects the determining ISDd.

Eq.(3.4) shows ISDnp; ISDd of 4 misclassified pixels, which are located at 4 corner pixels in n x n image.

$$ISD_{n4} = \frac{1}{{}_4C_2}(1 - 2 \cdot \frac{1}{2} \cdot \frac{1}{n})(4 + 2\sqrt{2}) \qquad (3.4)$$

ISDd varies with 1-1/n. ISDd increases as the number of pixel increases. That is, image resolution increases for same area and same error. But when n is greater than 150, then marginal ISDd is less than 0.01, and effect of image size can be neglected. Or, representative distance can be calibrated by eq.(3.5) to minimize the effect of image size.

$$(X' + Y')/2 \qquad (3.5)$$

here, $X' = X - 1/n$, $Y' = Y - 1/n$

C. Distribution of Misclassified Pixels

The pattern of distribution of misclassified pixels can be categorized as follows;

1. 0 < ISDd < 0.3
   - Densely aggregated
2. 0.3 < ISDd < 0.7
   - Evenly distributed
   - Randomly distributed
   - Slightly aggregated pixels distributed neighborly
3. 0.7 < ISDd
   - Sporadically distributed

3.1.2 Index of Spatial Distribution by scatteredness

To evaluate the scatteredness of misclassified pixels, following assumptions are required.

If misclassification is initiated by random error, the probability of the misclassified pixel at any point is fixed. Dividing image by several grids, probability of λ misclassified pixels in each grid has Poisson distribution with parameter $\lambda = P / k$ (k = No. of grids, P = No. of misclassified pixels).

An important property of Poisson distribution is that mean is equal to variance. Consequently, a variance(υ) of difference between number of misclassified pixels in each grid and a mean(λ) offers the index of scatteredness.

$$ISDs = \upsilon / \lambda \qquad (3.3)$$

If misclassified pixels distribute randomly, ISDs equals 1. If misclassified pixels distribute evenly, the number of misclassified pixels in each grid equals λ and ISDs equals 0. And, if misclassified pixels distribute aggregately, the number of misclassified pixels differs from λ and ISDs is greater than 1.

However, ISDs gives good explanation of the scatteredness for intra-grid, but poor explanation for inter-grids.

3.2 Determining the Spatial Distribution Pattern

ISDd and ISDs compensate mutually. So, using two indices at the same time, the spatial distribution pattern can be determined as in fig.3.1

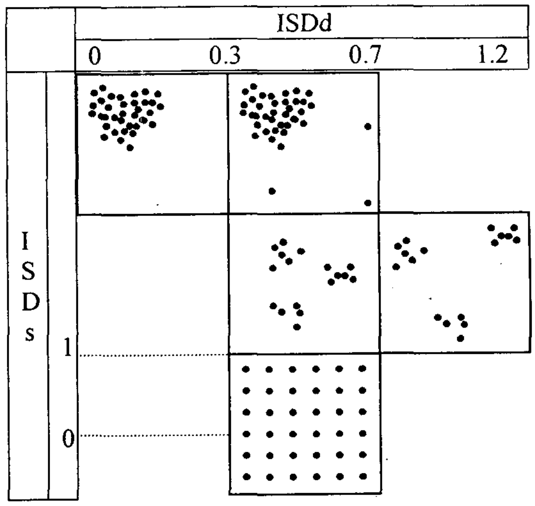

Fig. 3.1 Determining the Spatial Distribution Pattern by ISDd/ISDs

4. Estimating and Analyzing the Results

4.1 Study Area and Data

Gong-Ju and Chung-Yang area from 1:50,000 topographic maps have appropriate terrain features for remote sensing study. Satellite data sets are LANDSAT MSS, TM and SPOT XS. Reference data is extracted from 1:50,000 topographic maps. 1,200 samples are selected systematically at every 1,000m grid and for classes that have few samples, 42 samples are added randomly.

Considering terrain features of study area, 6 classes are selected (table 4.1).

The analysis contains classifying using MLC, and estimating overall accuracy, normalized Kappa coefficient and ISDd/ISDs for each image.

Borland C++ Builder(Borland International, Inc.) have been used to compiled every program needed in this study and ER Mapper(Earth Resource Mapping Pty Ltd) for classifications.

Table 4.1 Classes and Reference Data

Class Number	Class	Reference Data
1	Field	262
2	Paddy	73
3	Residential	76
4	Lake	10
5	Stream	28
6	Forest	793
Total		1,242

4.2 Results and Analysis

Table 4.2 presents overall accuracy, normalized Kappa coefficient and ISDd/ISDs of classification results.

Table 4.2 Estimators of Accuracy

Item \ Image	MSS	TM	SPOT
Overall Accuracy	0.603	0.622	0.639
Kappa Coefficient	0.542	0.713	0.470
ISDd	0.419	0.428	0.380
ISDs	0.975	0.879	1.165

Table 4.2 shows that the classification accuracy varies with scenes and that the ISDd has stabilized values. This is because that the ISDd suggests spatial distribution of misclassified pixels and misclassification probability, that is originated in spatial characteristics of study area.

However, ISDs shows slightly larger value for SPOT and smaller value for MSS, TM. The MSS, TM seems to be affected by the systematic sampling. SPOT has some aggregations in upper and lower part of the image, although the number of misclassified pixels is relatively less than that of MSS and TM. In this case it should be reconsidered whether the region, which contains more misclassified pixels, could be accepted or not.

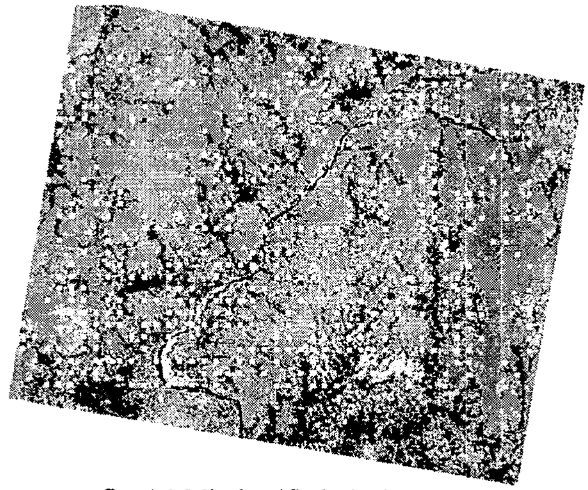

fig. 4.1 Misclassified Pixels on MSS

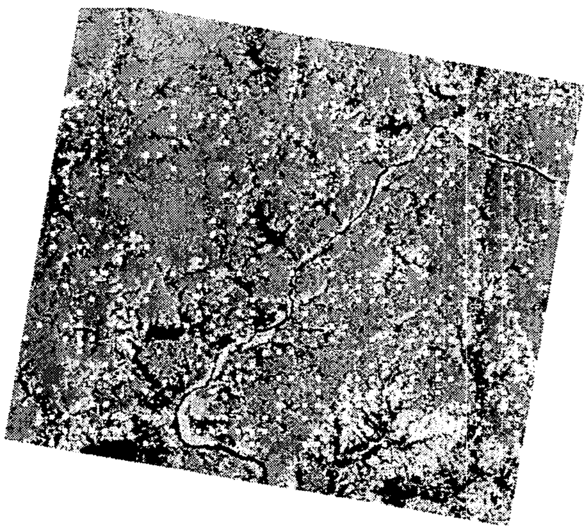

fig. 4.2 Misclassified Pixels on TM

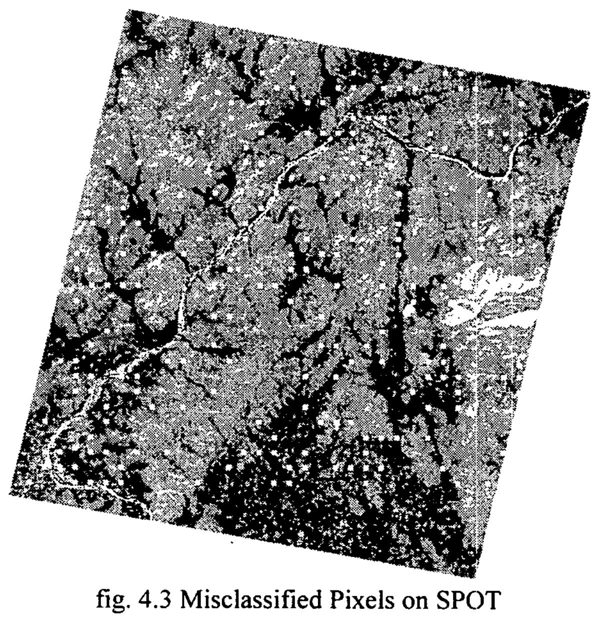

fig. 4.3 Misclassified Pixels on SPOT

5. Conclusion

The results of this study are as follows;

First, ISDd/ISDs are proposed for estimating the satellite image classification error.

Second, it became possible that the adoption or the rejection could be determined at sub-image level using ISD.

Third, ISD can offer index of the local aggregation of misclassified pixels which is originated in topography and land-cover.

The automatic extraction of rejected regions at sub-image level and the estimation of mixel and noise effects, which forms linear error, are needed for further study.

Acknowledgements

Special thanks to ADD(Agency for Defense Development) for help with satellite data.

< References >

1. Kim, Y.I., "Improving Correctness in the Satellite Remote Sensing Data Analysis –Laying Stresses on the Topographic Spectral Reflectance Correction and the Statistical Classification Techniques-", Seoul National University Ph. D. Paper, 1991.
2. Park, M.H., "A New Method for Classification of and Image Data Using Canonical Correlation Analysis", Seoul National University Ph. D. Paper, 1996.
3. Lee, H.Y., *Geographic Statistics*, Peop-Mun-Sa, 1991, p659
4. R. G. Congalton and R. A. Mead, 1983, "A Quantitative Method to Test for Consistency and Correctness in Photointerpretation", Photogrammetric Engineering & Remote Sensing, 49(1) : pp.69 - 74.
5. Resenfield, G. H. and K. Fitzpatrick-Lins, 1986, "A Coefficient of Agreement as a Measure of Thematic Classification Accuracy", Photogrammetric Engineering & Remote Sensing, 52(2) : pp.223 - 227.
6. Michael Story and Russell G. Congalton, 1986, "Accuracy Assessment: A User's Perspective", Photogrammetric Engineering & Remote Sensing, 52(3) : pp.397 - 399.
7. Xin Zhuang, B. A. Engel, Xiaoping Xiong and C. J. Johannsen, 1995, "Analysis of Classification Results of Remotely Sensed Data and Evaluation of Classificatio Algorithms", Photogrammetric Engineering & Remote Sensing, 61(4) : pp.427 - 433.
8. Thomas M. Lillesand and Ralph W. Kiefer, *Remote Sensing and Image Interpretation*, Wiley, 1994.