Optimum Spatial Resolution of Satellite-based Optical Sensors for Maximizing Classification Performance

Kohei Arai
Faculty of Science and Engineering
Saga University, Saga City
Japan

Abstract—Optimum spatial resolution of satellite-based optical sensors for maximizing classification performance is investigated. Also, classification performance assessment method considering spatial resolution of satellite-based optical imagers is proposed. Optimum spatial resolution which makes the highest classification accuracy is determined from spatial frequency components, spectral features of objects and classification method. First, in this paper, based on the relationship between variance of pixels and classification accuracy, classification accuracy for Landsat Multiple Spectral Scanner: MSS images with various Instantaneous Field of View (IFOV) will be shown. In their connection, variance of pixel values for images with various IFOV will be clarified. Second, assuming the shape of boundary line between adjacent categories is circle, relationship among IFOV, ratio of Mixels and classification accuracy will be cleared under the supposition that the number of Mixels equals to that of misclassified pixels. Finally, it will be also shown that aforementioned relationships and optimum spatial resolution have been confirmed by using airborne based MSS data of Sayama district in Japan.

Keywords—Spectral information; spatial information; maximum likelihood decision rule; satellite image; image classification; mixed pixel (Mixels); optimum spatial resolution; classification performance; spatial and spectral features

I. INTRODUCTION

Optimum spatial resolution of satellite-based optical sensors for maximizing classification performance is investigated. From the point of view of classification performance, there must exists an optimum spatial resolution of the spaceborne onboard optical sensors because the variances of the class categories are getting large in accordance with the spatial resolution. The classification performance is getting down because the overlapped areas among the class categories are getting large which results in confusion probabilities are getting large in accordance with spatial resolution.

Since variance of pixels correspond to that in the feature space increases in accordance with improvement of spatial resolution, classification accuracy will be gotten worse in accordance with improvement of spatial resolution under the limitations of variety of objects and class categories. On the other hand, classification accuracy gets better in accordance with improvement of spatial resolution because of decreasing of a ratio of "Mixels" which are pixels composing with plural class categories. Since aforementioned two effects contribute to classification accuracy multiplicatively, it seems that there exists an optimum spatial resolution.

The Instantaneous Field of View: IFOV of the multispectral radiometer mounted on the earth observation satellite with the highest classification accuracy, that is, the optimal spatial resolution, is generally determined by the spatial frequency component, spectral characteristics, classification method, and the like of the observation target [1]-[9]. Classification accuracy is defined as the discrimination efficiency (diagonal element of the confusion matrix) in maximum likelihood classification, and when parameters such as the object to be observed and the number of classes are limited, increasing the spatial resolution generally increases the variance in the feature space [10].

The classification accuracy becomes worse. On the other hand, the ratio of Mixels (mixed pixels) of different classes becomes smaller as the spatial resolution is improved, so that the classification accuracy is improved [11],[12]. Since the above effects synergistically contribute to the classification accuracy, it is considered that there is an optimal spatial resolution.

The motivation of this research study is to clarify relations between spatial resolution of optical sensors onboard satellites and classification performance and then find out optimum spatial resolution for maximizing classification performance.

This paper first clarifies the relationship between the instantaneous visual field and the classification accuracy [13] by deriving the case variance based on the relationship between the variance of pixel values and the classification accuracy. Next, assuming that the shape of the boundary of the same class region is an arc, the relationship between the instantaneous visual field and the ratio of the Mixels is obtained, and further, the relationship between the instantaneous visual field and the classification accuracy is obtained assuming that the ratio of the Mixels is a false recognition rate. Both relationships show that there is an optimal spatial resolution, and this is confirmed by using MSS data of the aircraft that observed Sayama Hills, in Japan.

In the following section, related research works and research background including motivation of the research are
II. RELATED RESEARCH WORKS

Classification by re-estimating statistical parameters based on auto-regressive model is proposed for purification of training samples. Meanwhile, multi-temporal texture analysis in Landsat Thematic Mapper: TM classification is proposed for high spatial resolution of optical sensor images [15]. On the other hand, Maximum Likelihood (ML) TM classification taking into account pixel-to-pixel correlation is proposed [16].

Supervised TM classification with a purification of training samples is proposed [17] together with TM classification using local spectral variability is proposed [18]. A classification method with spatial spectral variability is also proposed [19] together with TM classification using local spectral variability [20].

Application of inversion theory for image analysis and classification is proposed [21]. Meanwhile, polarimetric SAR image classification with maximum curvature of the trajectory in eigen space domain on the polarization signature is proposed [22]. On the other hand, A hybrid supervised classification method for multi-dimensional images using color and textural features is proposed [23].

Polarimetric SAR image classification with high frequency component derived from wavelet multi resolution analysis: MRA is proposed [24]. Comparative study of polarimetric SAR classification methods including proposed method with maximum curvature of trajectory of backscattering cross section in ellipticity and orientation angle space is conducted and well reported [25].

Comparative study on discrimination methods for identifying dangerous red tide species based on wavelet utilized classification methods is conducted [26]. On the other hand, multi spectral image classification method with selection of independent spectral features through correlation analysis is proposed [27]. Image retrieval and classification method based on Euclidian distance between normalized features including wavelet descriptor is proposed [28].

Gender classification method based on gait energy motion derived from silhouettes through wavelet analysis of human gait moving pictures is proposed [29]. Also, human gait skeleton model acquired with single side video camera and its application and implementation for gender classification is proposed [30] together with human gait skeleton model acquired with single side video camera and its application and implementation for gender classification [31]. On the other hand, gender classification method based on gait energy motion derived from silhouette through wavelet analysis of human gait moving pictures is proposed [32] together with human gait gender classification using 3D discrete wavelet transformation feature extraction [33].

Image classification considering probability density function based on simplified beta distribution is proposed [34]. Meanwhile, Maximum Likelihood Classification: MLH based on classified result of boundary mixed pixels for high spatial resolution of satellite images is proposed [35]. On the other hand, context classification based on mixing ratio estimation by means of inversion theory is proposed [36].

III. RESEARCH BACKGROUND

A. Relationship between Instantaneous Visual Field and Variance of Pixel Values

According to Friedman et al. [13], the classification accuracy P is determined by the variance $\sigma^2$ of the pixel values and the size $\alpha$ in the feature space of each class, which is determined by the classification method.

$$ p = 10^{-0.4/\beta} \beta = \alpha/\sigma \quad (1) $$

where, $\alpha$ is defined as a range of a discrimination threshold in an arbitrary dimension of an arbitrary class (this threshold is determined by a classification method that regards the range of this threshold as the same class). The unit is the digital count value, which is the same as the unit of the standard deviation ($\sigma$) of the pixel value. $\sigma^2$ changes depending on the instantaneous field of view. If the time series $x(t)$ of the original pixel values is now sampled at the time interval $l$ and the obtained series is $g(\xi)$, $\sigma^2(\alpha)$ is shown in Appendix 1. Equation (2) can be expressed as follows.

$$ \sigma^2(\alpha) = 2\sqrt{\left(\frac{1}{l/\alpha\sigma^2}\right)} \Gamma\left(\frac{1}{l/\alpha\sigma^2}\right) - \frac{a^2}{2\sqrt{\pi}} \int_0^\frac{1}{2\sqrt{\pi}} x^2 d\xi + \ldots \right) \quad (2) $$

That is, it can be seen that $\sigma^2(\alpha)$ can be represented by only the even-order terms of $a$, and can be represented by an equation that is negative from the second term. Therefore, when $x(t)$ is a continuous function, it can be seen that $\sigma^2(\alpha)$ becomes a constant value when $a$ approaches $0$, and decreases at first as an objective line as it increases. Since $\sigma^2(\alpha)$ in the range of the instantaneous visual field near the optimal spatial resolution is considered to decrease parabolically from a constant value, here, Eq. (2) is approximated by the second term.

B. Classification Method in Concern

According to Crapper, the Mixel ratio $F$ is expressed by the following equation.

$$ F = K_1 \sqrt{\pi} \left( \frac{a^2}{L} \right)^{\frac{3}{2}} \left( \frac{a^2}{L} \right)^\frac{1}{2} \frac{1}{L} \sum_j L_j/2\sqrt{\pi} \sum_j \sqrt{A_j} \right) $$

where $K_1 = \sum_j L_j/(2\sqrt{\pi} \sum_j \sqrt{A_j})$, $L_j$ denotes the perimeter of the class $j$ area, $A_j$: the area of the class $j$ area, $a$: the instantaneous visual field (pixel size), $L$: the average length when the boundary of the class area that crosses the pixel is approximated by a straight line.

Fig. 1 shows the definition of $L$. In the figure, the hatched area is defined as class $i$ and the other area is defined as class $j$. The boundary line length is defined as $L_{ij}$ and the length connecting the straight lines (broken line length) is defined as $L$. The average of $L$ is defined as $L$. 

www.ijacsa.thesai.org
The parameters other than \( a \) are obtained by the Crapper [11] as experimental values using 1605 types of sample data as follows. In the process of deriving Eq. (3),

\[
L=0.7935a, K_1=1.82, \sqrt{A} = 1483(m), A=3718600(m^2)
\]

Crapper makes the following assumptions.

1) There are at most two classes in a pixel.
2) The length of the line segment that intersects the edge of the pixel with the boundary between different classes in the image space is not different from the line segment length when it is approximated by a straight-line.

Both of the above assumptions hold if \( a \) is sufficiently small compared to the size of the class area. However, when discussing the optimal spatial resolution, it is generally considered that the size of the class area is equivalent to the pixel size. Must be made and this assumption does not hold. In particular, the assumption of (2) often does not hold. Here, it is assumed instead that "the category boundary line in a pixel is an arc." Under this assumption, the ratio \( q \) between the average length \( L_q \) of the class area boundary line and the average length \( L \) obtained by linearly approximating it is the radius \( r \) and the pixel size \( a \) when the boundary line of the class area is assumed to be circular. It becomes a function and the relationship shown in Fig. 2. Considering this ratio and adopting the experimental values of Crapper as other parameters, \( F \) can be expressed by the following equation.

\[
F = 1.62 \times \frac{10^{-3}a}{q(\pi/7)}, q(\frac{a}{r}) = L/L_{ij}
\]

Assuming that the classification accuracy of Mixels is 0, the classification accuracy of pixels (pure pixels) composed of a single class is 1, and their combined classification accuracy is equal to (1-\( F \)).

C. Overall Classification Accuracy

The total classification accuracy \( P_t \) is defined by the following equation in consideration of the influence of the variance of pixel values and the ratio of the Mixels on the classification accuracy.

\[
P_t = P \times (1-F)
\]

The pixel size at which this \( P_t \) is the highest is defined as the optimal spatial resolution.

---

**Fig. 1.** Definition of Notations.

**Fig. 2.** Ratio of the Line Length of Boundaries between Adjacent Class Categories \( L_{ij} \) to its Approximated Straight-Line Length \( L \).

---

IV. EXPERIMENT WITH AIRCRAFT MSS DATA OBSERVING SAYAMA HILLS: EXAMPLE OF OPTIMAL SPATIAL RESOLUTION

A. Geographical Characteristics of the Intensive Study Area

The analysis area is around Sayama as shown in Fig. 3. The south side consists of Sayama hills at about 150m above sea level, and the north side consists of flat land about 100m above sea level. The Sayama hill has Sayama lake, broadleaf / coniferous forest, grassland, etc. The flat land consists of development land including bare land, urban area, residential area, paddy field, upland field, tea plantation, etc.

B. Generation and Classification of MSS Data of Instantaneous Field of View: IFOV

On December 12, 1981, data of an instantaneous field of view of 1.25 m was collected using an MSS; DS-1250 for airborne use at an altitude of 500 m. Apply \( 8 \times 8, 12 \times 12, 16 \times 16, 24 \times 4, 48 \times 48 \) pixel window size smoothing filter to this to generate instantaneous visual field data of 10, 15, 20, 30, 60m.

Then, the maximum likelihood classification was tried using each band data in the wavelength range of 0.55 to 0.60, 0.65 to 0.69, 0.8 to 0.89, 8.0 to 14.0 \( \mu m \), and the discrimination efficiency was evaluated.  

Figure 4 shows examples of images of each instantaneous visual field. Table 1 shows the discrimination efficiency \( P_i \) for each class by maximum likelihood classification using the data of each instantaneous visual field. \( P_i; i \) is the class type, the area ratio of each class.

The discrimination efficiency \( P_a \) at each instantaneous visual field represented by the formula is shown.

\[
P_a = \Sigma R_i P_i
\]

\( R_i \) is the average of the area ratio of each class obtained by classifying the simulation images of each instantaneous visual field by the maximum likelihood method. From the table, it can be seen that the instantaneous visual field showing the maximum discrimination efficiency differs for each class, and that coniferous forests, fields, development areas, urban areas, etc. have a relatively large number of composite pixels of different classes, and the dimensions of the components are small.
As shown in Fig. 5, it can be seen that the variance of the pixel values is relatively low in the classification accuracy. It was also found that the average classification accuracy in each instantaneous visual field, weighted by the area ratio of each class, was highest when the instantaneous visual field was 30m.

### TABLE I. CLASSIFICATION ACCURACY FOR THE SIMULATION DATA OF SEVERAL INSTANTANEOUS FIELD OF VIEW (IFOV)

<table>
<thead>
<tr>
<th>a</th>
<th>1.25</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>30</th>
<th>60</th>
<th>Ri</th>
</tr>
</thead>
<tbody>
<tr>
<td>Needle Leaf Tree</td>
<td>33.1</td>
<td>58.5</td>
<td>87.8</td>
<td>69</td>
<td>80.2</td>
<td>86.2</td>
<td>22.5</td>
</tr>
<tr>
<td>Deciduous Tree</td>
<td>76.9</td>
<td>79.4</td>
<td>63.2</td>
<td>91.9</td>
<td>92.4</td>
<td>86.5</td>
<td>17.5</td>
</tr>
<tr>
<td>Paddy Field</td>
<td>83.5</td>
<td>84.5</td>
<td>73.4</td>
<td>83</td>
<td>95.5</td>
<td>87</td>
<td>6.4</td>
</tr>
<tr>
<td>Perm Area</td>
<td>57.9</td>
<td>91.5</td>
<td>81.8</td>
<td>87</td>
<td>82</td>
<td>45.5</td>
<td>9</td>
</tr>
<tr>
<td>Tea Garden</td>
<td>75.6</td>
<td>75.1</td>
<td>88.4</td>
<td>67.9</td>
<td>98.3</td>
<td>93.7</td>
<td>6.9</td>
</tr>
<tr>
<td>Grass Field</td>
<td>87</td>
<td>90.9</td>
<td>95.6</td>
<td>80.1</td>
<td>100</td>
<td>88.6</td>
<td>3.2</td>
</tr>
<tr>
<td>Developmental Area</td>
<td>38.2</td>
<td>89.5</td>
<td>63.1</td>
<td>98.1</td>
<td>92.8</td>
<td>100</td>
<td>3.6</td>
</tr>
<tr>
<td>Water Body</td>
<td>99.5</td>
<td>98.9</td>
<td>99.8</td>
<td>100</td>
<td>99.9</td>
<td>99.8</td>
<td>13.5</td>
</tr>
<tr>
<td>Mulberry Field</td>
<td>61.8</td>
<td>86</td>
<td>86.5</td>
<td>92.9</td>
<td>74.6</td>
<td>79.9</td>
<td>7.4</td>
</tr>
<tr>
<td>Urban Area</td>
<td>48.8</td>
<td>66.2</td>
<td>88.2</td>
<td>88.6</td>
<td>83.6</td>
<td>70.6</td>
<td>10</td>
</tr>
<tr>
<td>Pa</td>
<td>63.7</td>
<td>78.3</td>
<td>83</td>
<td>88.9</td>
<td>88.9</td>
<td>83.5</td>
<td></td>
</tr>
</tbody>
</table>

C. Relationship between IFOV and Variance

As shown in Fig. 6, the variance differs depending on the channel and class; where, in order to clarify the average relationship between the IFOV and the variance, the function was approximated by the least square method.

As a result, Eq. (7) was obtained. The approximation error normalized by the variance value at this time was 0.71.

$$\sigma^2(\alpha) = 498.32 - 0.49a^2 \tag{7}$$

By substituting Eq. (7) into Eq. (1) and evaluating the classification accuracy $p$ using $\alpha$ as a parameter, the result is as shown in Fig. 7. In the figure, the solid line is the calculated value of Eq. (7), and the broken line is the estimated value.

When the aircraft data used in the analysis were classified by the maximum likelihood method, the range $\alpha$ of each class was about 10 to 120. Therefore, Fig. 7 requires the classification accuracy when $\alpha$ is changed in 20 steps from 30 to 90.

D. Relationship between IFOV and Mixing Ratio

Estimated from the land cover classification map of the Sayama area used in the analysis, the average of the radius $r$ was determined to be about 20 m when the boundary between each class was assumed to be an arc. Therefore, here, $r$ was changed to 10 to 30 m, and Eq. (4) was opened to evaluate the effect of the change of the Mixels by the instantaneous visual field on the classification accuracy.

Fig. 8 shows the results. The figure also shows the calculated values of Crapper [11] and the experimental values of Jackson [12].

Since the calculated value of Crapper is based on the assumption that the instantaneous visual field is sufficiently smaller than the class size, the calculated value of the model proposed in this paper asymptotically approaches the region where the instantaneous visual field is small. Also, Jackson's experimental values correspond to those where $r$ in the calculated values of this model is about 25 m.
Comparing the experimental values (circles in Fig. 8) for calculating the ratio of the Mixels for the images of each instantaneous visual field with the calculated values of this model, it can be seen that r is about 20 m. This experimental value is based on the assumption that the result of classifying aircraft MSS data with an instantaneous field of view of 1.25 m into 10 classes by maximum likelihood classification is 0 (%) for the Mixels, and this classification result is the MSS image of each instantaneous field of view. In each pixel, the ratio of the number of categories that are composed of a plurality of categories is calculated as the number of Mixels.

This indicates that, when the analysis target area is classified into 10 classes in Table I, the average of the class sizes can be considered to be equivalent to a circle with a semicircle of about 20 m.

E. Relationship between IFOV and Classification Accuracy

The total classification accuracy \( p_t \) considering the variance and the Mixels was calculated by Eq. (5) and shown in Fig. 9. At this time, 20m was selected as the parameter r, and 50, 70, and 90 were selected for \( \alpha \). In addition, Fig. 9 also shows the classification accuracy \( p_t \) obtained by actually performing the maximum likelihood classification, and (Table 1). The two values are almost the same, especially when \( \alpha = 90 \), which indicates that the method proposed in this paper for finding the optimal spatial resolution is appropriate.

[Diagram showing the relationship between IFOV and classification accuracy]

Fig. 9. Comparison of Theoretical Classification Accuracy with Experimental one for Various IFOV.

V. CONCLUSION

Optimum Spatial Resolution of Satellite Based Optical Sensors for Maximizing Classification Performance is investigated. Also, classification performance assessment method considering spatial resolution of satellite based optical imagers is proposed. Optimum spatial resolution which makes the highest classification accuracy is determined from spatial frequency components, spectral features of objects and classification method.

The validity of the method of obtaining the ratio of the Mixels and the variance of the pixel values and the method of deriving the classification accuracy considering them were confirmed by the analysis example of the Sayama area of Japan scene of Landsat satellite imagery data. According to this method, the classification accuracy can be estimated by giving the class size in the feature space and the radius when the class region in the image space is approximated by an arc. It is also possible to find the optimum spatial resolution that maximizes the accuracy.

VI. FUTURE RESEARCH WORKS

The proposed method is adopted in the real earth observation satellite imagery data, and it is a future subject to realize a more usable classification method. Also, optimum spatial resolution would be better to be realized with actual remote sensing satellite based optical sensors.

ACKNOWLEDGMENT

The author would like to express our sincere thanks to Dr. Crapper of CSIRO, Prof and Strahler of Hunter College, former Professor Tsuchiya of Chiba University and others for their valuable advices and discussions. The author, also, would like to thank Professor Dr. Hiroshi Okumura and Professor Dr. Osamu Fukuda for their valuable discussions.

REFERENCES


www.ijacsa.thesai.org


AUTHOR’S PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor for the Remote Sensing Society of Japan for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is a Science Council of Japan Special Member since 2012. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission “A” of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 55 books and published 620 journal papers as well as 450 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. http://teagis.ip.is.saga-u.ac.jp/index.html

www.ijacsa.thesai.org