# Object-oriented remote sensed image classification accuracy assessment

**Zhaocong Wu, Lina Yi,** and **Guifeng Zhang**
School of remote sensing and information engineering
Wuhan university
Wuhan, China
lina1986350@163.com

*Abstract*—**This paper proposes a new accuracy assessment scheme for object-oriented remote sensing imagery classification. It measures both the geometrical and thematic accuracy of objects. The geometrical accuracy are measured from two aspects of the area and boundary accuracy. The thematic accuracy is measured associated with objects. The classification results of a Quickbird image with different object location are evaluated by the proposed method. Experiments show the method can provide more information about the classification accuracy and is potential to solve the problem of the traditional statistical accuracy assessment measures.**

***Keywords:*** ***object-oriented; classification accuracy assessment; segmentation quality***

## I. INTRODUCTION

In recent years, object-oriented remote sensing image classification is widely used in various application fields. This technique provides a means to produce object-based thematic maps. It is always required to evaluate the classification accuracy before using the map for further application. The common used accuracy assessment measures, such as confusion matrix, can only characterize the correctly classified objects or pixels in a statistical way without considering the geometrical accuracy. It suffers a problem that two classification results, in which the objects location precisions are different, may statistically have the same accuracy measure.

To comprehensively evaluate the classification accuracy, it is needed to evaluate both the geometrical and thematic accuracy in an objective way. This paper proposes a new classification accuracy assessment scheme for object-based classification (Fig. 1). In this supervised scheme, objects are the base elements to be evaluated. The classification result and the reference data are firstly matched, and then both the area-based and boundary-based discrepancy measures are calculated to assess the object geometrical accuracy. Moreover, the thematic accuracy can be measured by counting how well the classification object is correctly identified at the pixel or object level. Here, we use two representative measures of precision, recall (Crevier, 2008) and FOM (Pratt et al., 1978) to evaluate the geometrical accuracy. Then, the thematic accuracy is evaluated by assign the correctly classified objects 1 and incorrectly classified objects 0. This scheme can provide a means to retain uncertainty information and distributing it to the users at the object level. The overall classification accuracy can also be measured by the statistical average of the object-based accuracy measures.

The research was supported by the National Natural Science Foundation of China (No.40201039 and No. 40771157), the National High Technology Research and Development Program of China (No.2007AA12Z143) and the PhD candidates self-research program of Wuhan Unversity (NO. 20082130201000048).

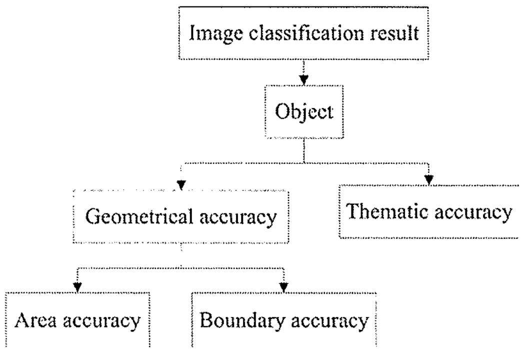


Figure 1. Classification accuracy assessment based on objects

The paper is organized as follows: firstly, we depict the best matching criteria and the measures for object accuracy assessment; Then, the geometrical accuracy measures, are verified by geometrical accuracy comparison experiments for different types of land cover regions; Finally, a Quickbird image classification is evaluated by the proposed method to prove its application potential.

## II. GEOMETRICAL ACCURACY ASSESSMENT MEASURES

To evaluate the geometrical accuracy of the classification result, firstly the matching relationships between the classification result and reference data are identified, then the accuracy measures can be calculated.

### *A. The Best Matching Criteria*

Here, we used the best matching criteria proposed by (Crevier, 2008). It identify the best matching reference object $S^h_{l(i)}$ of the object by a similarity function $l(i)$ in (1), where $S^m_i$ and $S^h_j$ are the grouping of pixels in the object and the reference object respectively, $i$ and $j$ are the sequential number of objects. The criteria means that the reference object with the largest ratio value of the intersection area and union area is the matching object of the current $i$ th segmented object.

$$l(i) = \arg\max_j \frac{\left|S_i^m \cap S_j^h\right|}{\left|S_i^m \cup S_j^h\right|} \quad (1)$$

*B. Area Accuracy Measure*

Crevier (2008) suggested that the *precision* and *recall* are two effective measures to evaluate the region discrepancy. Consider a certain object and a reference object on the same image, let the pixels in the reference object be *positives*, and those in the object be *detections*. Then the *precision* of the object is simply the area of the intersection of two objects divided by the area of the object. Likewise the *recall* associated with the object is the area of the intersection divided by the area of the reference object. For the entire image, we can define the overall precision or recall as the average precision $\bar{p}$ or recall $\bar{r}$ for all objects, weighted by their areas as shown in (2) and (3), where $n_m$ is the number of objects, | X | represents the number of pixels of set X, A is the total area of the matching objects in accuracy assessment. Weight parameter $w_i = \frac{S_i^m}{A}$.

$$\bar{p} = \sum_{i=1}^{n_m} w_i \frac{\left|S_i^m \cap S_{l(i)}^h\right|}{\left|S_i^m\right|} = \frac{1}{A}\sum_{i=1}^{n_m}\left|S_i^m \cap S_{l(i)}^h\right| \quad (2)$$

$$\bar{r} = \sum_{i=1}^{n_m} w_i \frac{\left|S_i^m \cap S_{l(i)}^h\right|}{\left|S_{l(i)}^h\right|} = \frac{1}{A}\sum_{i=1}^{n_m}\left|S_i^m\right| \frac{\left|S_i^m \cap S_{l(i)}^h\right|}{\left|S_{l(i)}^h\right|} \quad (3)$$

*Precision* measures how well the objects are contained in the matched reference object; *recall* measures how well the objects cover the matched reference object. An object that perfectly matches reference object will have both *precision* and *recall* of 1. These quantities will be smaller than 1 for all other conditions.

*C. Boundary Accuracy Measure*

Figure of Merit (FOM) (Pratt et al., 1978) is an effective boundary-based accuracy assessment measure. It can be used to measure the boundary accuracy of objects by measuring the empirical distance between the objects' boundary pixels. For each reference pixel on a object, a measure $M_i$ is defined in (4) to measure the matching degree of the compared object's boundary pixels. Where $d_i$ is the minima distance of boundary pixels to the reference boundary pixel $i$, $d_l$ and $d_h$ define the distance tolerances for matching. If $d(i) \le d_l$, the corresponding object pixel is ideally matched, $M_i$ is 1. If $d(i) \ge d_h$, no pixel on the object is matched with the reference object boundary pixel, the $M_i$ equals to 0. If $d_l \le d(i) \le d_h$, $M_i$ ranges from (0, 1).

$$M_i = \begin{cases} \dfrac{1}{1+(d(i)-d_l)^2} & d_l \le d(i) \le d_h \\ 0 & d(i) \ge d_h \\ 1 & d(i) \le d_l \end{cases} \quad (4)$$

After calculating all the measure $M_i$ of the reference object boundary pixels, the object FOM measure can be calculated with (5) by averaging the $M_i$ by $N$, where $N$ is the number of boundary pixels of the reference object.

$$FOM_{obj} = \frac{1}{N}\sum_{i=1}^{N} M_i \quad (5)$$

Moreover, the average measure $\overline{FOM}$ of all the involved objects can be used to assess the overall geometrical accuracy of classification results.

*D. Thematic Accuracy Measure*

For each classified object, the corresponding thematic accuracy $P_i$ can be calculated by comparing its thematic attribute to the reference attribute. If the matching objects are of the same class attribute, $P_i$ is 1, otherwise $P_i$ is 0, *i* is the sequential number of segmented object. The overall thematic accuracy of the classification results can then be measured by the weighted average of $P_i$, the weight for each object *i* is $\frac{1}{A}\left|S_i^m \cap S_{l(i)}^h\right|$. When the involved objects are correctly assigned the thematic attributes, and are ideally matched with, contain or inside the corresponding reference object, the thematic measure is approaching 1. This measure is derived from the geometrical accuracy measure $\bar{p}$. Like the $\bar{p}$, it favors the over-segmented results.

## III. Geometrical Accuracy Evaluation

To validate the geometrical accuracy evaluation measures, four different types of land cover regions, including water, forest, farmland and residential area are considered for the accuracy assessment experiments. For each image, the reference objects are extracted by human segmentation; the geometrical accuracy of three different results (Fig. 2) is evaluated respectively. The overall geometrical accuracy of each result measured by the different discrepancy measures is shown in Fig.3. It shows that the precision measure is approaching 1 when the image is normal or over segmented; the recall measure is approaching 1 when the image is normal or under segmented; $FOM_{obj}$ is approaching 1 when the image is normal segmented, the over or under segmentation results measured by $FOM_{obj}$ are lower than the normal segmentation. $FOM_{obj}$ is a good measure to represent how accurate the object boundary is. However, it may be high at the situation that the image is over-segmented (water) or under segmented (farmland). When all of these three measures are approaching 1, the segmentation result is mostly similar to the reference data and is of high accuracy.

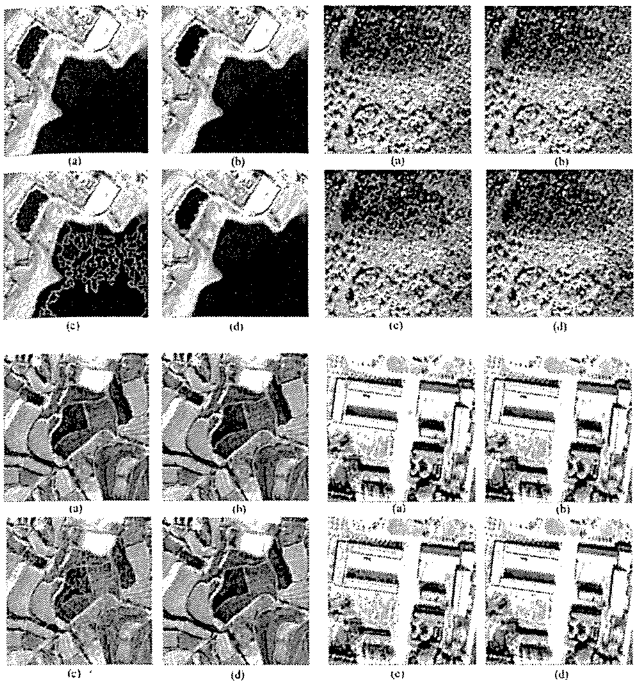

Figure 2. Segmentation results of different land cover class regions: (a)human segementation;(b)normal segmentation;(c)over-segmentation;(d)under-segmentation

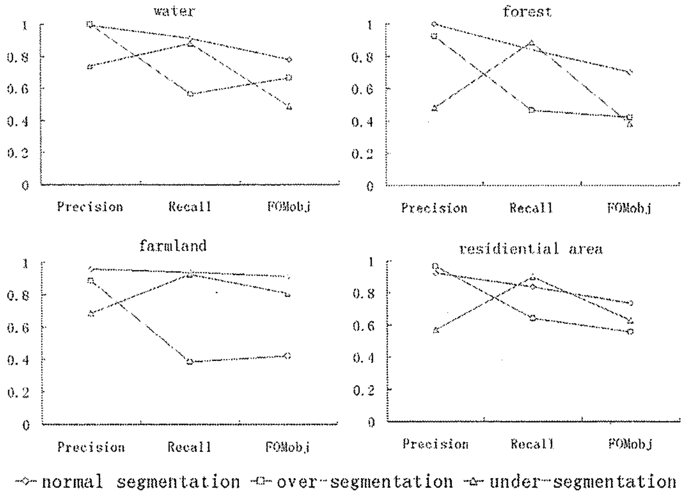


Figure 3. Geometrical accuracy evaluation results

## IV. CLASSIFICATION ACCURACY EVALUATIION EXPERIMENT

To validate the method, we used a 2050*2060 Quickbird image (Fig. 4a) covering wuhan suburb in the classification evaluation experiment. The reference data is derived by

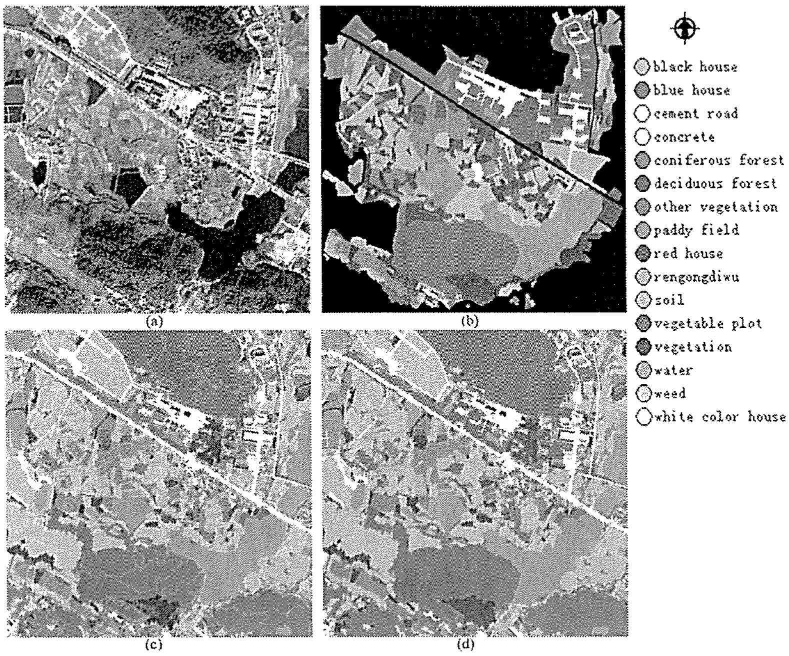


Figure 4. Quickbird image and the classification results: (a) Quickbird image; (b) reference data; (c)classification result based on over-segmented objects; (d)classification result by merging neighboring objects on result 1

human segmentation (Fig. 4b), two classification results are evaluated (Fig. 4c, Fig. 4d). The first classification result is achieved based on the over-segmented objects. The second classification result is achieved by merging the neighboring objects that have the same thematic attributes on result 1. Therefore, these two results are with the same thematic accuracy if they are assessed by the pixel-based statistical accuracy measures, such as confusion matrix. However, they are different in segmentation accuracy. Compared with the reference data, the first result is over-segmented and the second one is under-segmented. The overall geometrical accuracy of these two classification results measured by the three discrepancy measures are shown in table I.

The overall thematic accuracy of the classification results 1 and result 2 calculated by the proposed method is 52.5% and 43.8% respectively.

TABLE I. OVERALL GEOMETRICAL ACCURACY OF THE CLASSIFICATION RESULTS

| Classification result | Discrepancy measures | | |
|---|---|---|---|
| | $\overline{p}$ | $\overline{r}$ | $\overline{FOM_{obj}}$ |
| Result 1 | 0.844881 | 0.592786 | 0.467658 |
| Result 2 | 0.644129 | 0.786967 | 0.522793 |

## V. CONCLUSION

In the accuracy assessment of object-based classification, both the correctness of the geometrical and thematic attributes can be measured. The proposed method can solve the problem of the traditional pixel based statistical measures that two classification results, in which the objects location precisions are different, statistically have the same accuracy measure. Further research is needed to find more objective accuracy measures for application.

## ACKNOWLEDGEMENTS

The authors would like to thank the supports from the National Natural Science Foundation (No.40201039 and No. 40771157), the National High Technology Research and Development Program of China (No.2007AA12Z143) and the PhD candidates self-research (including 1+4) program of Wuhan Unversity in 2008 (NO. 20082130201000048).

## REFERENCES

Crevier, D. (2008). Image segmentation algorithm development using ground truth image data sets. *Computer Vision and Image Understanding*. 112, 143–159.

Carleer, A. P., Debeir, O., Wolff, E. (2005). Assessment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering & Remote Sensing*. 71, 1285-1294.

Pratt W., Faugeras O.D., Gagalowicz A. (1978). Visual discrimination of stochastic texturefields . *IEEE Trans.on Systems, Man and Cynernetics*, 11, 796–804.