SCALE SELECTION BASED ON MORAN'S I FOR SEGMENTATION OF HIGH RESOLUTION REMOTELY SENSED IMAGES

Yan Meng[1], Chao Lin[2], Weihong Cui[1] and Jian Yao[1]

[1]School of Remote Sensing and Information Engineering, Wuhan University, China
[2]Beijing GEOWAY Software Co., Ltd, Beijing, China
Email: mengyan@whu.edu.cn

ABSTRACT

Image segmentation is a prerequisite for object-based image analysis (OBIA). However, selecting an optimal segmentation scale is often time consuming and needs trial-and-error. This paper presents an unsupervised scale selection method based on the rate of change of a spatial autocorrelation indicator – the global Moran's I for segmentation of high resolution remotely sensed images. It was compared with other two scale selection methods and its effectiveness is validated through both visual analysis and by referencing to multiple manual segmentations. Experimental results on our own data and statistical data from an external reference showed that the optimal scale could be easily selected through the proposed method.

Index Terms— Local Variance, Spatial Autocorrelation, Segmentation Scale, High Resolution, Moran's I

1. INTRODUCTION

This paper reports an unsupervised scale selection method for segmentation of high resolution remotely sensed images. It belongs to the category of methods based on the rate of change of certain indexes (ROC methods), which utilize the dynamics of segmentation, i.e. the change from the segmentation result at one scale to the next. The dynamics may be different using different indexes. However, in a bottom-up segmentation scheme, it is expected the rate of change of the used index will be great until most of the objects in the image are formed and afterwards the rate of change will keep small for a certain range of scales. Kim and Madden [1] and Kim et al. [2] proposed the rate of change of the local variance with respect to the segmentation scale for scale selection. Drăguţ et al. [3] developed the ESP (Estimation of Scale Parameter) tool to estimate multiple suitable scales, and an effort in its automation was followed by [4-5]. Most existing ROC methods utilize un-weighted local variance (UW-LV), which is calculated as Formula (1):

$$UW\text{-}LV(s)=\frac{1}{R}\sum_{i=1}^{R}LV_i(s), \qquad (1)$$

where R is the number of segments at the scale s, and $LV_i(s)$ is the local variance of the i-th segment at scale s. Note that [3] used local standard deviation for calculation.

Most existing ROC methods assume that UW-LV increases greatly when the segments merge fast. However, if some segments of high local variances are merged, the UW-LV may drop because the number of the segments of high variances reduces. On the other hand, combining segments of moderate local variances may exhibit insignificant fluctuation in UW-LV. Note whether a segment is of high local variance is decided by its ranking among all segments at the same scale. Moreover, the rate of change of local variance has not been fully utilized in existing ROC methods, leading to unsatisfactory segmentation results. For example, [4-5] rely on too strict conditions that may be difficult to meet. To address these issues, this paper proposes to use the rate of change of a spatial autocorrelation indicator – the global Moran's I [6] to select the best segmentation scale.

2. METHODOLOGY

The principle underpinning the proposed method is as follows. When raising the segmentation scale, many sub-regions with similar features are merged and the spatial autocorrelation of the segments changes greatly until the segments match the actual objects. Beyond that the spatial autocorrelation will stay with insignificant difference. The above principle is similar to that of existing ROC methods, however, it utilizes a spatial autocorrelation indicator – the global Moran's I to consider the spatial distribution of segments:

$$MI=\frac{N\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\omega_{ij}(y_i-\bar{y})(y_j-\bar{y})\right)}{\left(\sum_{i\neq j}\sum\omega_{ij}\right)\left(\sum_{i=1}^{N}(y_i-\bar{y})^2\right)}, \qquad (2)$$

where N is the total number of segments, y_i is the mean spectral value of region R_i, and $\bar{y}$ is the mean spectral value of the image. ω_{ij} measures the spatial proximity, and if region R_i and R_j are neighbors, ω_{ij}=1, otherwise, ω_{ij}=0. In our method, MI is first normalized to the range [0, 1] :

$$Norm\text{-}MI(s) = \frac{MI(s) - MI_{\min}}{MI_{\max} - MI_{\min}}, \quad (3)$$

where $MI_{\max}$ and $MI_{\min}$ are the maximum and minimum values of MI, respectively. The rate of change of Moran's I is calculated as Formula (4):

$$ROC(s) = \frac{1}{\mathrm{M}} \sum_{i=1}^{M} \left| Norm\text{-}MI_i(s+1) - Norm\text{-}MI_i(s) \right|, \quad (4)$$

where $Norm\text{-}MI_i(s)$ denotes the normalized MI at scale s of the i-th band, and M is the number of spectral bands.

Since considerably under-segmented results may also lead to a small rate of change of MI, the GS (Global Score) method [7] is used to exclude seriously under-segmented scales. According to the GS method, the minimal GS value corresponds to the optimal segmentation scale. If we denote the minimal GS value as $GS_{\min}$ and the maximal GS value $GS_{\max}$, only the scales that meet the following conditions will be kept for further evaluation:

$$GS(s) \le GS_{\min} + C_1 \times (GS_{\max} - GS_{\min}), \quad (5)$$

where C_1 is the coefficient to control the scale range that the proposed method searches (the selected scale range), which is suggested to be 0.6, which is a technical number great enough to include all optimal scales in the selected scale range. $GS_{\max}$ is experimentally set to 1 because its numerical value is almost identical to 1 in theory when the heterogeneity is not too high for most segments.

If we denote the mean of $ROC(s)$ values as ROC_{mean}, then a scale can be regarded as a very prominent scale (a scale corresponds to a very prominent valley) if it satisfies the condition described in Formula (6):

$$ROC(s) \le ROC_{\min} + C_2 \times (ROC_{mean} - ROC_{\min}), \quad (6)$$

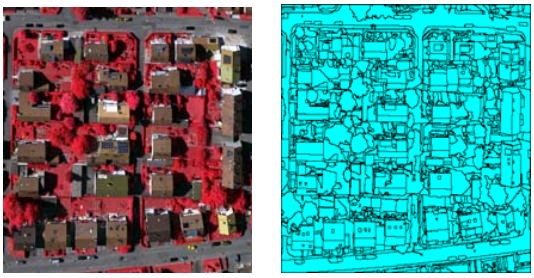
where C_2 is used to control the threshold of defining prominent valleys (PV) and is suggested to set to about 0.15 and 0.45 to define PVs at two different levels, based on the experimental results over the dataset used in the present paper. Note that if several PVs are serial, only the first will be adopted. Then according to the two values of C_2, we can get two selected scales, one of which is just the optimal scale. Thus the scale selection issue becomes selection from two instead from multiple (usually more than 10). Note that we will refer to the proposed method as ROC-MI below.

3. EVALUATION

The proposed method was tested on a dataset composed of 8 images at a spatial resolution of 0.08 m/pixel over Vaihingen in Germany [8], and 8 images at a spatial resolution of 0.15 m/pixel over Toronto in Canada. All the 16 images were originally released for urban classification and 3D building reconstruction by ISPRS. For each image in the dataset three reference segmentations were manually delineated for this study.

The PR (Probabilistic Rand) index [9], which calculates the expectation of Rand indexes [10] between an automatic segmentation and a set of reference segmentations, is used for objective evaluation. The Rand index between two segmentations is the sum of the number of pixel pairs that have the same labels and those that have different labels in both segmentations, divided by the total number of pixel pairs. The larger the PR index value, the higher quality the corresponding segmentation. A series of segmentations were performed with eCognition at 41 different scales, ranging from 10 to 410 with an increment of 10, so that both over-segmented and under-segmented cases were included. The weights of all bands, and the weight of smoothness and compactness were at their default values. The weight for shape and color was set to 0.6 because small ones lead to significant jag for objects with smooth boundaries.

Two images were taken as example images (see Fig. 1(a) and Fig. 2(a)). In the experiments we took the scales that yielded five highest PR index values as optimal (acceptable). For the Vaihingen example image the optimal scales were from 240 to 280, and from 30 to 70 for the Toronto example image. From the two displayed manual segmentations (Fig. 1(b) and Fig. 2(b)), it was seen that almost all meaningful objects were manually delineated, based on the belief that a good segmentation should keep all the necessary details for subsequent high-level tasks.



(a) (b)

Fig. 1. Vaihingen example image (a) and one of its three manual segmentations (b). The image is with size of 1683×1626 pixels and displayed in green, red and NIR.

For clarity, we will refer to [3] as the ESP method and [4-5] as the Auto-ESP method below. To search for the optimal scale, we need to inspect all the prominent peaks for the ESP method [3] until the optimal scale is found. From Fig. 3, however, we could see that the first prominent peak appeared at scale 150 for the Vaihingen image and 70 for the Toronto image. For the Vaihingen example image, though one prominent peak (at scale 250) appeared at the optimal scale, there were several other similarly or more prominent peaks to distinguish from, which made the selection of optimal scale difficult. No scale was selected by the Auto-ESP method [4-5], because no scale existed whose

local variance was less than or equal to its previous scale as required by the method. For the Toronto example image, the ESP method functioned well, selecting scale 70 as the optimal one. But the Auto-ESP method selected scale 230, yielding an under-segmented result.

(a) (b)

Fig. 2. Toronto example image (a) and one of its three manual segmentations (b). The image is with size of 777×739 pixels and displayed in blue, green and red.

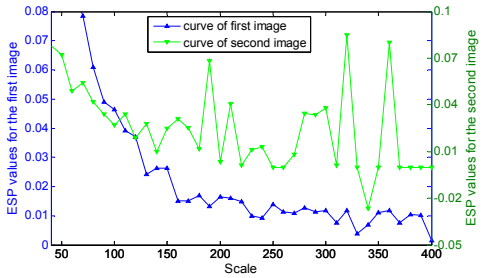


Fig. 3: ESP values at different scales. The curve of the Vaihingen (first) test image is marked blue and the Toronto (second) test image by green.

Table 1. The comparison of selected scales by different methods.

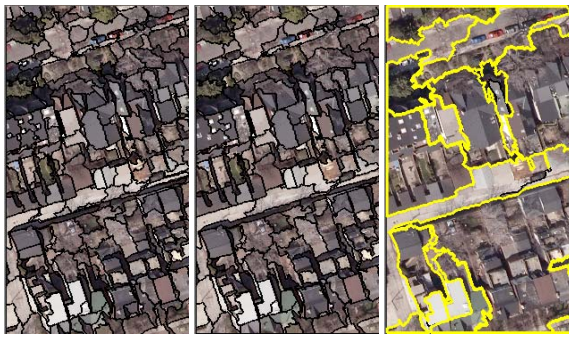
image / method	Vaihingen	Toronto
PR index	[240, 280]	[30, 70]
ESP	150	70
Auto-ESP	Failed	230
ROC-MI	180, 250	60, 100

As for the proposed method, the scales selected for the Vaihingen example image were 180 and 250, and scale 60 and 100 for the Toronto example image. In these two selected scales, the optimal scales were always captured, yielding satisfactory segmentation outcome. The comparison of selected scales by different methods was presented in Table 1.

Visual analysis was also made to further compare the three methods. For the Vaihingen example image, segmentations at scale 150 and scale 250 were presented in Fig. 4, from which we could conclude scale 150 led to over-segmentation and the outcome at scale 250 was more acceptable. Similarly, segmentations at scale 60, scale 70 and scale 230 were presented in Fig. 5 for the Toronto test image. Scale 60 and 70 balanced over-segmentation and under-segmentation while scale 230 yielded under-segmentation. Therefore, visual comparisons on the results of the two example images reconfirmed that the proposed method functioned best.

(a) (b)

Fig. 4. Segmentations at scale 150 and 250 for the left part of the Vaihingen example image. (a) for scale 150 and (b) for 250.



(a) (b) (c)

Fig. 5. Segmentations at scale 60, 70 and 230 for the left part of the Toronto example image. Among the three results, under-segmentation which only existed at scale 230 was marked by yellow boundaries. (a) for scale 60, (b) for scale 70 and (c) for scale 230.

As for the entire dataset, the ESP method had prominent peaks for none of the images over Vaihingen while 6 for the 8 images over Toronto in Canada. Note that it is a semi-automatic method, and we had to interpret prominent peaks visually from the ESP curve. The Auto-ESP method got accurate results for 6 images in all with a success rate of 37.5%. In contrast, there were prominent valleys corresponding to the optimal scales in the ROC-MI curves for all the 16 images, 13 of which could be covered by the two scales of the proposed method, yielding a recall of

81.25%. Note that $C_1 = 0.6, C_2 = 0.15 \text{ or } 0.45$ for all images and if we only set C_2 to 0.45, we could get a success rate of 62.5%, meaning the optimal scales of 10 images could be got automatically.

Note that the proposed method also got satisfying result when using the statistical data in [7] for validation. However, they used Moran's I as a metric to measure the inter-segment disparity while we used its change rate to model the dynamics of segmentations.

This section compared the proposed ROC-MI with the ESP method and the Auto-ESP method, and the proposed ROC-MI method got satisfying results comparing to the other two methods. Maybe the rate of change of UW-LV will be very small at the optimal scale, but whether it will be zero depends on the increment of segmentation scales. A very small increment may ensure a UW-LV value of zero at the optimal scale, but this may be heavily affected by noise and needs many segmentation computations. Otherwise, for example in the present paper the increment was set to 10, the Auto-ESP method may miss the optimal scale just because of the too strict condition even when the rate of change of UW-LV can model the dynamics of segmentations well. When the segmentation becomes under-segmented, the results also tend to be un-changed and may be accompanied with smaller rate of change of local variance than that at the optimal scale, which make the under-segmented scales more inclined to satisfy the condition required by the Auto-ESP method and explains the under-segmented result selected for the Toronto example image. The proposed method utilizes a more suitable indicator to model the dynamics of segmentations and small rate of change is searched instead of zero values, so the proposed method is more robust with the scale increment, both of which make the proposed method tend to get accurate result.

5. CONCLUSIONS

An unsupervised segmentation scale selection method based on the rate of change of Moran's I was put forward in this paper for high resolution remotely sensed imagery segmentation. Experiments based on our own data and statistical data from an external reference showed that the proposed method is useful for segmentation scale selection. The rate of change of Moran's I can model the dynamics of segmentations well, but how to find the optimal scale from the curve of ROC-MI still needs more research. In this paper, we find two scales to ensure high recall. How to achieve the same recall by finding only one scale is obviously necessary and will be our future effort.

6. ACKNOWLEDGEMENTS

The study was partially supported by the National Natural Science Foundation of China under Grant No. 41101410 and Natural Science Foundation of Hubei Province of China under Grant No. 2011CDB273. The Vaihingen dataset was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) [8]: http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html. The authors would like to acknowledge the provision of the downtown Toronto dataset by Optech Inc., First Base Solutions Inc., GeoICT Lab at York University, and ISPRS WG III/4. We would also appreciate Ying Wang, Meng Dai, Zheng Zhang, Yatong Xia and Wei Hu in Wuhan University for their help in producing the manual segmentations.

REFERENCES

[1] M. Kim and M. Madden, "Determination of optimal scale parameters for alliance-level forest classification of multispectral IKONOS images," in *Proceedings of the 1st International Conference on Object-based Image Analysis (OBIA 2006)*, 2006.

[2] M. Kim, M. Madden, and T. Warner, "Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery," in *Object-based image analysis*, ed: Springer, 2008, pp. 291-307.

[3] L. Drăguţ, D. Tiede, and S. R. Levick, "ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data," *International Journal of Geographical Information Science,* vol. 24, no. 6, pp. 859-871, 2010.

[4] L. Drăguţ, O. Csillik, C. Eisank, and D. Tiede, "Automated parameterisation for multi-scale image segmentation on multiple layers," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 88, pp. 119-127, 2014.

[5] L. Drăguţ and C. Eisank, "Automated object-based classification of topography from SRTM data," *Geomorphology,* vol. 81, no. 3-4, pp. 21-33, 2012.

[6] P. A. Moran, "Notes on continuous stochastic phenomena," *Biometrika,* vol. 37, no. 1-2, pp. 17-23, 1950.

[7] B. Johnson and Z. Xie, "Unsupervised image segmentation evaluation and refinement using a multi-scale approach," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 66, no. 6, pp. 473-483, 2011.

[8] M. Cramer, "The DGPF-test on digital airborne camera evaluation overview and test design," *Photogrammetrie-Fernerkundung-Geoinformation,* vol. 2010, no. 2, pp. 73-82, 2010.

[9] R. Unnikrishnan and M. Hebert, "Measures of similarity," in *Proceedings of Seventh IEEE Workshop on Applications of Computer Vision, WACV 2005*.

[10] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical Association,* vol. 66, no. 336, pp. 846-850, 1971.