# A QUANTITATIVE EVALUATION OF IMAGE SEGMENTATION QUALITY

**Honglei Zhu,** Senior Research Associate
**Hao Chen,** GIS Research Associate
Clark Labs, Clark University
950 Main Street, Worcester, MA 01610
hzhu@clarku.edu
hchen@clarku.edu

## ABSTRACT

Segmentation is often a procedure to group spatially adjacent image pixels into segments. Spatially adjacent pixels form one image segment if they meet some criteria, such as spectral similarity. A segmentation result may vary, depending on a given generalization level and other constraints such as compactness. In practice, the quality of a segmentation result is often assessed visually by the analyst, and that lacks a quantitative support and the quality of the result relies on the experience of an analyst. This research gives a quantitative estimate of a segmentation result using indices such as: a) a summed standard deviation of the input images within each image segment; b) a summed absolute difference within each image segment; and c) a summed difference of a segment to its adjacent segments. A series of segmentation results were studied from two perspectives. One focused on a series of segmentation results with different generalization levels, the other compared two series of segmentation results obtained using two different methodologies. The findings of this study can be used a) to guide an algorithm optimization for image segmentation, b) to assist the evaluation of different software programs to be used for a certain application and, c) to assist in identifying an optimal segmentation result for a given analysis.

## INTRODUCTION

Along with the advances of remote sensing sensor technology and improvements in image processing techniques, object oriented image analysis has been one of the most active research topics in satellite image processing. Image segmentation is often the first procedure employed in such an analysis and its result may then be utilized for further analysis. Therefore the quality of image segmentation is fundamental to a high quality object oriented image analysis. There are different segmentation algorithms for forming segments within an image (Blaschke *et al.*, 2006; Harlick and Shariro, 1985; Ton *et al.*, 1991; and Chen *et al.*, 2004). Such approaches fall into different categories, such as spatial-aware and aspatial approaches, pixel-based, edge-based, and region-based methods. Following a different approach, or using different constraints for the same approach, the segmentation results may vary. In practice, visual interpretation is often used to assess the quality of segmentation. Such practice lacks a quantitative measurement that leads to an optimal result, and the quality of an analysis relies on the experience of an analyst.

In an effort to address such an issue, this research proposes three indices for a quantitative estimate: a) a sum of size-weighted standard deviation within each segment, derived from the input images, b) a sum of absolute differences within each image segment and, c) a sum of sized-weighted difference of a segment to its adjacent segments. A series of segmentation results were studied from two perspectives. One focused on a series of segmentation results with different generalization levels, and the other compared two series of segmentation results obtained using two different segmentation algorithms.

## SEGMENTATION METHODS, DATA SET, AND ANALYSIS

To facilitate this study, two commonly used segmentation approaches were chosen to derive the series of segmentation. One is a watershed-based segmentation and the other, a region-grow approach. The detailed description about the region-grow approach can be found from Baatz and Schape, 2000, and a watershed-based method is given as follows.

## Watershed-based Segmentation

There are three steps involved in watershed-based image segmentation. 1) Derive surface image: a variance image is derived from each image layer. Centered at every pixel, a 3 by 3 moving window is used to derive its variance for that pixel. The surface image for watershed delineation is a weighted average of all variance images from all image layers. Equal weight is assumed in this study. 2) Delineate watersheds. From the surface image, pixels within a homogeneous region form a watershed. 3) Merge segments. Adjacent watershed may be merged to form a new segment with larger size according to their spectral similarity and a given generalization level.

## Data Set

QuickBird multispectral satellite imagery was used. The data set was collected on October 29, 2005. The area was around (42.4N, 83.8W), in the southeast of Michigan, U.S.A.. The image consisted of four bands, at the wavelengths of blue, green, red, and near infra-red. A subset with 1024 by 1024 rows and columns was utilized for this study. With a spatial resolution of 2.6 meters, the data set covers about 7.1 square km (see Figure 1).

## Statistical Analysis

A series of segmentation for both watershed-based and region-grow approaches were derived. And each series contained 20 generalization levels, from as fine as 39595 segments, to as coarse as one segment for the whole imagery. Two segmentation results were given in Figures 2 and 3. They were at the intermediate generalization level. The break down of the 20 levels can be seen from the x-axis in Figures 4-7. The three indices of measuring the quality of segmentation are given in Figures 4-6. In order to describe the degree of compactness, a measure of sum of perimeters of segments was adopted, and the result is given in Figure 7.

**Figure 1.** Study area: false color composite of the study area.

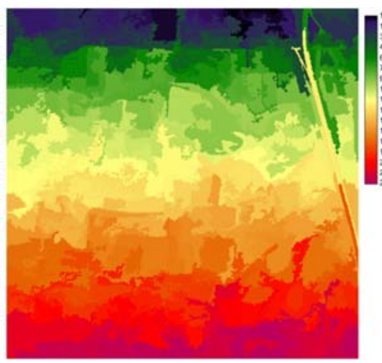

**Figure 2.** Watershed-based segmentation with 240 segments.

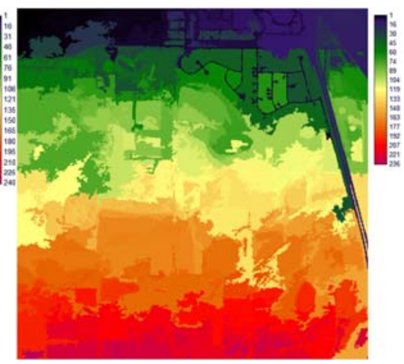

**Figure 3.** Region-grow segmentation with 236 segments.

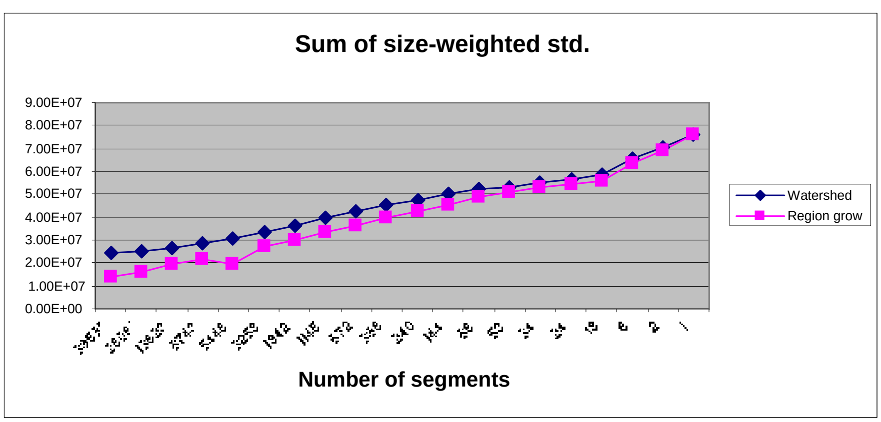


**Figure 4.** Sum of size-weighted standard deviation. The x-axis is the number of segments in one segmentation. The series of segmentation were derived from 20 different generalization levels for both watershed-based and region-grow approaches.

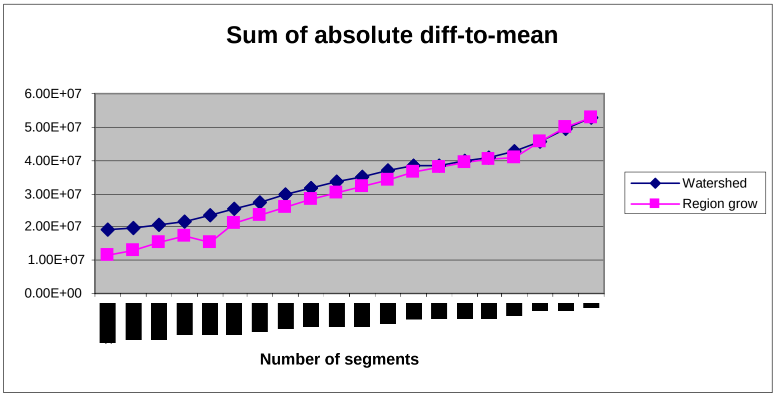


**Figure 5.** Sum of the absolute difference-to-mean.

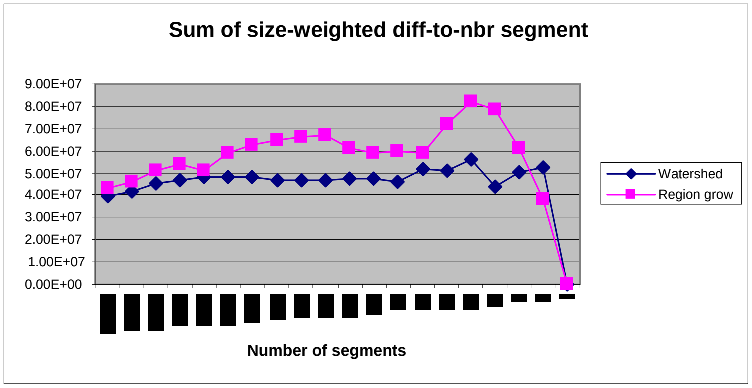


**Figure 6.** Sum of the size-weighted difference-to-neighboring-segments. The average difference to all neighbors was used to represent each segment, and its size was taken as its weight.

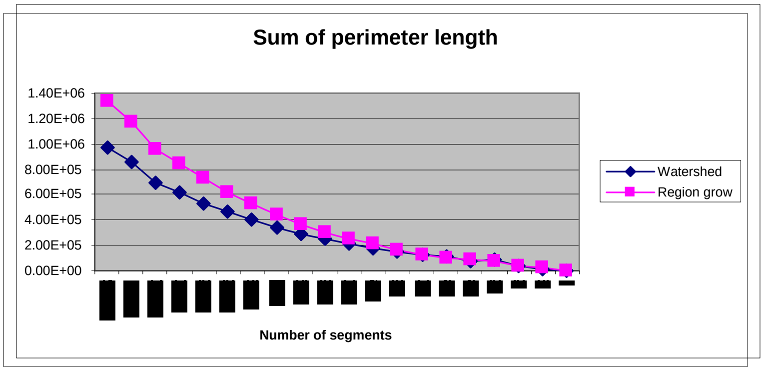


**Figure 7.** Sum of the perimeter length of all segments.

## CONCLUSION AND DISCUSSIONS

This study gave a quantitative estimate about the statistical characteristics of a segmentation result. Three indices, sum of size-weighted standard deviation, sum of absolute differences to mean, and sum of size-weighted differences to neighboring segments, were introduced for a quantitative estimate from two perspectives. The first perspective was the responses of the indices to the generalization levels. From Figures 4 and 5, we noticed that the first two indices are monotonically increasing when segmentation changed from finer to coarser. We also noticed that the third index, demonstrated in Figure 6 by two curves, representing watershed-based and region-grow methods respectively, behaved differently. But fluctuation was one feature they were in common. Such a behavior could be a response to a turning point when generalization levels went across a scale of natural landscape. The second perspective was quality comparison of the results that came from different segmentation algorithms. We noticed that statistically, the region-grow approach yielded more optimal results than that of the watershed-based approach. It was demonstrated by the curves in Figures 4 and 5, where the region-grow curves were consistently lower than that of the watershed-based curves. And in Figure 6, the region-grow curve was higher, which was better, because it indicated a larger difference of a segment to its neighbors.

The proposed indices in this study can be used for software development and for applications as well. For software development, it can be used to assist an algorithm optimization. For image analysis applications, it can assist in choosing the best parameter set to be used in getting an optimal result. It also might be helpful in identifying a proper generalization level for a given analysis.

## FUTURE WORK

There are two aspects need to be further looked into. One is the response of the indices to a change of spatial resolution, and the other is their responses to different landscapes. It is expected that Figure 6 can be better interpreted when the indices become available using more data sets.

## ACKNOWLEDGEMENTS

The authors would like to thank Ronald Eastman for his valuable comments and helpful discussions. The authors also feel grateful for the support from James Toledano during the development of this research. The QuickBird imagery data set was a subset of a larger image from a USDA funded research project at Clark Labs.

## REFERENCES

Baatz, M., and A. Schape, 2000, Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation, *XII Angewandte Geographische Informationsverarbeitung*, Wichmann-Verlag, Heidelberg: 12-23.

Blaschke, T., C. Burnett, and A. Pekkarinen, 2006, Image segmentation methods for object-based analysis and classification, *Remote Sensing Image Analysis: Including the Spatial Domain, Springer,* edited by S.M. de Jong and F.D. van der Meer, pp. 211-236.

Chen, Q., C, Zhou, J. Luo, and D. Ming, 2004, Fast segmentation of high-resolution satellite images using watershed transform combined with an efficient region merging approach, *IWCIA 2004*: 621-630.

Harlick, R.M., and L.G. Shariro, 1985, Image segmentation techniques, *Computer Vision, Graphics, and Image Processing* 29:110-132.

Ton, J., J. Sticklen, and A.K. Jain, 1991, Knowledge-based segmentation of Landsat images, *IEEE Transactions on Geoscience and Remote Sensing*, 29 (2): 222-232.