

ASSESSING EDGE AND AREA METRICS FOR IMAGE SEGMENTATION PARAMETER TUNING AND EVALUATION

H. P. Meyer^a, A. Van Niekerk^{a,*}

^a Dept. of Geography & Environmental Studies, Stellenbosch University, South Africa Private Bag X1, Matieland, Stellenbosch, 7602, South Africa

KEY WORDS: Segmentation, Edge metric, Area metric, Boundary delineation, Accuracy, Object-based image analysis

ABSTRACT:

Image segmentation algorithms allow the creation of variables that influence the attributes of the resulting objects. There is currently no set way of identifying the best segmentation parameters for a specific application. It would consequently be useful to calculate the accuracy of a particular segmentation for determining how well it can delineate an object of interest. This article assesses the use of both an area and edge metric to evaluate and quantify the ability of a segmentation algorithm to delineate objects of interest based on manually collected reference data. SPOT 5 imagery was used for the segmentation of two 2x3km study areas in the Eastern-Cape of South Africa. The aims of this study were to use the area metric to identify a scale parameter for which the ratio between oversegmentation (OS) and undersegmentation (US) yields the most favourable result and, use the edge metric to determine the accuracy to which the boundary of objects of interest is delineated. The usefulness of estimating scale parameter (ESP) and segmentation parameter tuner (SPT) are also investigated to identify the suitable scale parameter. The root mean square error (RMSE) between OS and US, was used in conjunction with the edge metric results to produce an overall accuracy value between 0-1. The ESP tool was found useful as a guide, but failed to identify the scale parameter yielding the most accurate results, whereas, the SPT tool succeeded in both cases to identify this scale parameter. It was concluded that the edge and area metric can be used in conjunction to, not only determine the ideal scale parameter, but also calculate the overall accuracy of the segmentation algorithm.

1. INTRODUCTION

The recent increase in availability of high resolution satellite imagery has led to a growing interest in object-based image analysis (OBIA). Fundamental to this relatively new paradigm in image analysis is the grouping of pixels with similar spectral or spatial attributes to form image objects (Blaschke, 2010; Blaschke et al., 2014; Lucieer & Stein, 2002; Moller et al., 2007; Dragut et al., 2010). This process, called image segmentation, is not trivial as poor segmentations have a negative effect on the accuracy of any further classification or processing that follows (Baatz et al., 2008; Navulur, 2006). Image segmentation is a spatial clustering technique that divides an image into non-overlapping regions or segments (Moller et al., 2007). Image segmentation can be categorized into edge detection methods and region-growing techniques. Edge detection segmentation regionalizes an image by locating feature boundary elements, while region-growing algorithms make use of seed pixels that are grown into objects, based on a region's homogeneity (Fortin et al., 2000; Munoz et al., 2003; Blaschke et al., 2014). Many factors affect the quality of segmentation results (Moller et al., 2007), and much research has been done on improving image object delineation by considering different segmentation algorithms (Meinel & Neubert, 2004), tuning input parameters (Achanccaray et al., 2014; Achanccaray et al., 2015) and studying the effect that the spatial and spectral characteristics of the objects of interest have on the output (Lucieer & Stein, 2002; Burnett & Blaschke, 2003; Tian & Chen, 2007; Blaschke, 2010; Petropoulos et al., 2012). Effective validation techniques that quantify a segmentation result's ability to delineate objects of interest

remain elusive (Lucieer & Stein, 2002; Benz et al., 2004; Clinton et al., 2010). The accuracy of image objects can be validated by studying their topological (area) and/or geometrical (edge) relationships (Molenaar, 1998; Zhan et al., 2005). Area metrics generally determine the extent of US and OS, where US is quantified by calculating the area of segments intersecting the boundaries of the reference object (Moller et al., 2007) and OS is determined by calculating the area of segments falling inside of the reference object. When using reference polygons for segmentation validation, the ideal result must adhere to two conditions, namely: 1) The number of segments within the reference polygon must be at its lowest; and 2) The area of intersecting segments falling outside of the reference polygon must be at a minimum (Benz et al., 2004). The first condition refers to OS, while the second refers to US.

In order for a segmentation to produce meaningful and accurate objects based on a reference dataset, it is important that the segmentation parameters are properly tuned. This can be a time consuming process of trials-and-errors. Segmentation parameter tuning (SPT) is a free software tool that automatically tests certain parameter values and compares the level of agreement between the segmented objects and a manually collected reference objects. The level of agreement between the segmentation output and reference objects are calculated for each parameter combination and referred to as segmentation goodness (Achanccaray et al., 2014; Achanccaray et al., 2015). The main purpose of the experiments was to determine to what extent edge and area metrics can be used to assess the output of segmentation algorithms and for selecting optimal input features and parameters. The estimated scale parameter (ESP) tool (Dragut et al., 2010) is also employed for comparison purposes.

* Corresponding author

2. STUDY AREA

For the purpose of this study, two 2x3km study sites, representing a river and a dam respectively, were selected for carrying out the experiments. Study site 1 includes a section of the Kouga River and is situated in a mountainous region. This site was chosen as it includes a surface water body that is particularly challenging to delineate using remote sensing. The river course is relatively narrow due to the mountainous morphology, the intermediate channel gradient and stability, and the wandering channels associated with this type of landscape (Hogan & Luzi, 2010).

Study site 2 covers the Glen Melville dam just north of Grahamstown which was built as part of the Orange-Great Fish Interbasin transfer scheme, and has a capacity of 6 million m³ (Lorraine & Mullins, 2011). The dam is dominated by mountainous landscape to the north and relatively flat hills to the south.

The variation in reference object size and shape between the two study areas was taken into account and will consequently contribute to testing the robustness of the analysis.

3. IMAGE DATA

The data for the study consist of 10 m resolution multispectral SPOT5 images as well as 2.5 m resolution panchromatic images of the eastern-cape. The images used for study areas 1 and 2 were captured in March 2012 and March 2013 respectively. Images were ordered from SANSa (South African National Space Agency) Earth Observation. The 2.5m panchromatic and 10m multispectral SPOT5 imagery were orthorectified using 0.5m ground sample distance orthorectified aerial photography, provided by the National Geo-spatial Information (NGI), as basis and geographical reference. The images were then atmospherically corrected (ATCOR 2) and pansharpned to 2.5m using the PANSHARP tool in PCI Geomatica. The data used as reference were collected manually by digitizing the water bodies in both study sites to minimize error.

4. METHODS

4.1 Area overlay

The optimal segmentation for an intensive background feature (e.g. trees, buildings) would result in the reference polygon to be covered by a single object. This is however not expected to be the case for a spatially extensive feature (e.g. roads, rivers) which would likely contain multiple meaningful objects (Tian & Chen, 2007). All the segmented objects intersecting the reference shape are divided into two classes namely, partially overlaying and overlaying. Therefore, a slightly modified equation from the one proposed by Clinton et al. (2010) is used for both the calculation of OS as well as US. With an increase in scale, OS drops due to the decrease in objects that fall completely within the borders of the reference object, and US increases. The area of intersecting objects falling outside of the reference object are used to calculate under segmentation, while all the resulting objects falling completely within the reference object are used to calculate over segmentation. The RMSE between US and OS is calculated and used as a measure of accuracy for the area metric.

4.2 Distance between edge pixels

In this study, the purpose of using the edge metric is, not to judge the goodness of fit between the objects, but the accuracy of boundary delineation which in order to obtain, we used the algorithm proposed by Delves et al. (1992) calculating the average distance from the reference shape to the nearest overlaying segment boundary.

The edge metric is used to calculate the average distance error between the boundaries of the resulting objects that intersect the reference object and the boundaries of the reference object. This is a measure of accuracy for the boundary delineation for the segmentation.

4.3 Estimating scale parameter

ESP is used to evaluate rate of change of local variation to identify suitable scales where the segmentation would have the greatest capacity to differentiate between spectral classes. ESP is a method presented by Dragut et al. (2010) that estimates scale parameter for the segmentation of remotely sensed data in eCognition. ESP is an iterative algorithm that builds on the principle of local variance (LV). It generates image objects on multiple scale levels to test the heterogeneity between objects. A measure called rate of change is used as a test for the LV variation between objects at certain scales

4.4 Segmentation parameter tuner

SPT is an automatic method that rely on optimization algorithms, to search the parameter space for an optimal set of values in order to produce the most accurate segmentation, and in that manner minimize human intervention (Achanccaray et al., 2015). The goodness of these segmentations are measured by a metric that expresses numerically the level of disagreement between the segmentation output and a set of reference polygons. In this study, differential evolution (Storn & Price, 1997) is used as optimization algorithm for parameter tuning, while the multiresolution region growing algorithm (Happ et al., 2013) is used for the segmentation. The area-fit index is used to express segmentation quality, by comparing the segmentation results with the reference polygon and therefor helping to guide the optimization process (Achanccaray et al., 2015). The lower the value of the metric, the more similar the resulting segmentation is to the reference polygon. For this study, the updated version called SPT 3.1 has been used as a guide and verification to the scale, shape and compactness parameters chosen.

5. RESULTS AND DISCUSSION

The initial segmentation was run using the standard optical bands as input, and varying scale from 1-100. A range was identified, for both study areas, where the optimal scale parameter might lie. With this range restricted to 5-30, the processing time was reduced substantially and consequently the creation of redundant data prevented. In Table 1 below, are the parameters that produced the most accurate segmentation results according to both metrics, as well as the results obtained from ESP and SPT. Various shape and compactness parameters were tested, whereas only the most accurate is stated below.

Results		
	Study area 1	Study area 2
Shape	1	0.5

Compactness	0.5	0.5
US	0.366	0.390
OS	0.659	0.955
RMS	0.533	0.729
Edge metric	0.044	0.030
Equilibrium (Scale)	21	-
Scale parameter	12	25
ESP output (Scale)	8	14
SPT output (Scale)	12	24
Overall accuracy	0.024	0.022

Table 1. Combined overall results

The segmentation results of the two study sites vary substantially. This is attributed to the size difference between the reference objects of both study areas as well as land cover characteristics. It therefore made it difficult to use one segmentation method for both study areas. The expected trend was that with a low starting scale, OS is high and US low. The RMSE is used in this case to observe the general accuracy of the area metric for the various scale parameters, as well as combined with the edge metric result, producing an overall accuracy value between 0-1, with higher accuracy the closer the value is to 0.

The trend for the edge metric however starts off with a relatively high error at a low scale, this error decreases as the scale increases until a lowest point (highest accuracy) and then increases again.

The smaller the scale, the higher the OS ratio, but because of each object being smaller, the error in boundary delineation is also lower.

5.1 Study area 1

In this segmentation based on optical bands, parameters shape and compactness were chosen as 1 and 0.5 respectively. OS and US was found to intersect at scale 19, at which the accuracy is 0.079 pixels. However, when the scale is lowered to 11, a highest accuracy is reached of 0.042 pixels error. The expected nature of results obtained from the edge metric, would be accurate at a smaller scales and increase in error as scale increases. As the size of objects increase, so does the distance between the outline of the reference shape and the overlaying segmented objects. This is where the area metric helps to identify the scale parameter range where OS and US are equal.

Although the results produced by the ESP tool aren't very accurate, the values suggested does fall within a small range from the value identified to produce the most accurate results. The SPT suggested the use of scale parameter as 12 and shape and compactness as 0.49 and 0.62 respectively. Although the scale parameter identified were correct, more accurate results were obtained from using different shape and compactness parameters.

At scale parameter 12, the RMS error of 0.533 was multiplied by the edge metric result of 0.044, which resulted in a lowest overall accuracy value of 0.024.

5.2 Study area 2

Study area 2 contains a large object for which the assumption is not true, that the perfect segmentation will yield one object per reference object. Basing the accuracy on the OS/US ratio alone would therefore be problematic for smaller objects in the same study area, essentially resulting in the misclassification thereof. OS is based on the total area of objects falling inside the reference object, because this is expected to be high in the case of large objects, OS on its own is consequently neglected. The

focus however shifts to US, which calculates the area of boundary objects falling outside of the reference object, and the RMSE between the two. Using the optical bands for the segmentation weighting with shape and compactness parameters both 0.5, scale parameter 25 was identified as where the boundary delineation accuracy is at a high with an error of 0.03 pixels. The US ratio produced by the area metric at scale 25 is 0.389, which is another example that would fit the norm of the previous study area and the use of optical bands to base segmentation on.

Again, the results suggested by ESP were not very useful. While SPT also yielded quite similar results with the shape and compactness parameters suggested not being accurate, but with scale parameter being out by one, at 24. Again the RMS error was calculated and multiplied with the edge metric result for scale 25. The resulting value of 0.022 was found as the highest overall segmentation accuracy for study area 2.

From the results obtained, it is possible for both instances to identify the parameters that yield the most accurate results. Determining the segmentation accuracy was achieved with both metrics, and a useful way of combining the results obtained to give an overall accuracy was proposed.

6. CONCLUSION

Both metrics have their advantages and weaknesses. Using them in conjunction with one another makes it possible to take into account the different facets of accuracy, whether it is area of overlap or boundary delineation. Favourable parameters changed from study area 1 being shape 1 and compactness 0.5, to shape 0.5 and compactness 0.5 in the second area. This means that the importance of spectral and shape characteristics taken into account for region merging shifted due to the landscape change. ESP was found to be the least useful of the methods investigated, with only identifying scale parameters fairly close to the most accurate range. SPT in terms of scale parameter tuning was found quite useful, identifying the correct parameter in study area 1 and being out by 1 in study area 2. With regard to parameters shape and compactness, SPT did not yield very accurate results. In both study sites, the edge metric succeeded in determining the accuracy and ability of a segmentation algorithm to delineate the boundary of a reference object. The area metric however can be divided into OS and US. US is useful to determine the percentage area of objects overlaying the reference object, which falls either within or outside the reference object. OS however is only useful for objects under which the assumption, that a perfect segmentation will yield one object for every reference object, is true. This is not expected for large reference objects and were therefore not of use in study area 2. The RMSE were calculated between US and OS, and used in conjunction with the edge metric result, to consequently yield a ratio value between the two metrics. The lower this number is, the more accurate the overall result of the segmentation. In both cases, the scale parameter for which this value was the smallest, were identified.

Being able to accurately map features of interest using an object oriented approach depend primarily on the segmentation algorithm's ability to delineate these features. It is confirmed that using three of the four investigated methods, that it is indeed possible, to ultimately quantify the accuracy to which this is achieved.

REFERENCES

- Achancarray, P., Ayma, VA., Jimenex, LI., Bernabe, S., Garcia, S., Happ, PN., Feitosa, RQ. & Plaza, A., 2014. A free software tool for automatic tuning of segmentation parameters. *South-Eastern European Journal of Earth Observation and Geomatics* 3(2):707-711.
- Achancarray, P., Ayma, VA., Jimenex, LI., Bernabe, S., Happ, PN., Costa, GA., Feitosa, RQ. & Plaza, A., 2015. SPT 3.1: A free software for automatic tuning of segmentation parameters in optical, hyperspectral and SAR images. *IEEE International Geomatics and Remote Sensing Symposium*. Milan, 26-31 July 2015. 4332-4335.
- Baatz, M., Hoffmann, C. & Willhauk, G., 2008. Progressing from object-based to object-oriented image analysis. In Blaschke T, Lang S & Hay GJ (ed) *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Springer, Berlin.
- Benz, UC., Hofmann, P., Willhauk, G., Lingenfelder, I., & Heynen M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing* 58, pp. 239-258.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, pp. 2-16.
- Blaschke, T., Hay, GJ., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, RQ., Van der Meer, F., Van der Werff, H., Van Coillie, F., and Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS journal of Photogrammetry and Remote Sensing* 87, pp. 180-191.
- Burnett, C., & Blaschke, T., 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. *Ecological Modelling* 168, pp. 233-249.
- Clinton, N., Holt, A., Scarborough, J., Yan, L., & Gong, P., 2010. Accuracy Assessment Measures for Object-based Image Segmentation Goodness. *Photogrammetric Engineering & Remote Sensing* 76(3), pp. 289-299.
- Delves, LM., Wilkinson, R., Oliver, CJ., & White, RG., 1992. Comparing the performance of SAR image segmentation algorithms. *International Journal of Remote Sensing* 13(11), pp. 2121-2149.
- Dragut, L., Tiede, D., & Levick, S., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24(6), pp. 859-871.
- Fortin, MJ., Olson, R., Ferson, S., Iverson, L., Hunsaker, C., Edwards, G., Levine, D., Butera, K., & Klemas, V., 2000. Issues related to the detection of boundaries. *Landscape Ecology* 15, pp. 453-466.
- Happ, P., Feitosa, R., Bentes, C., & Farias, R., 2013. A region growing segmentation algorithm for GPUs. *IEEE Geoscience and Remote Sensing Letters* 10(6), pp. 1612-1616.
- Hogan, DL., and Luzi, DS., 2010. Channel geomorphology: Fluvial forms, processes, and forest management effects. In: Pike RG, Redding TE, Moore RD, Winkler RD, and Bladon KD (Eds.), *Compendium of Forest Hydrology and Geomorphology in British Columbia*. BC Ministry of Forests and Range, Forest Science Program, Victoria, BC, and FORREX Forum for Research and Extension in Natural Resources, Kamloops, BC Land Management Handbook No (Vol. 66, pp. 331-371).
- Lorraine, R., & Mullins, G., 2011. Grahamstown's Water Supply: a brief history from 1812 to 2008. *Annals of the Eastern Cape Museums* 7, pp. 1-64.
- Lucieer, A., & Stein, A., 2002. Existential uncertainty of spatial objects segmented from satellite sensor imagery. *IEEE Transactions on Geoscience and Remote Sensing* 40 (11), pp. 2518-2521.
- Meinel, G., & Neubert, M., 2004. A comparison of segmentation programs for high resolution remote sensing data. Leibniz Institute of Ecological and Regional Development (IOER). Dresden, Germany.
- Molenaar, M., 1998. *An Introduction to The Theory of Spatial Object Modelling*. Research Monographs in Geographical Information Systems, Taylor & Francis.
- Moller, M., Lymburner, L., & Volk, M., 2007. The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation* 9, pp. 311-321.
- Munoz, X., Freixenet, J., Cufi, X., & Marti, J., 2003. Strategies for image segmentation combining region and boundary information. *Pattern Recognition Lett.* 24, pp. 375-392.
- Navulur, K., 2006. *Multispectral Image Analysis Using the Object-Oriented Paradigm*, CRC Press, New York.
- Petropoulos, GP., Kalaitzidis, C., Prasad Vadrevu, K., 2012. Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Comput Geosci.* 41, pp. 99-107.
- Storn, R., & Price, K., 1997. Differential Evolution – A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* 11, pp. 341-359.
- Tian, J., & Chen, DM., 2007. Optimization in multi-scale segmentation of high resolution satellite images for artificial feature recognition. *International Journal of Remote Sensing* 28(20), pp. 4625-4644.
- Zhan, Q., Molenaar, M., Tempfli, K., & Shi, W., 2005. Quality assessment for geo-spatial objects derived from remotely sensed data. *International Journal of Remote Sensing* 26(14), pp. 2953-2974.