

IMPROVING THE SPEED OF MULTIREOLUTION SEGMENTATION USING SLIC SUPERPIXELS

O. Csillik^{a,*}, S. Lang^a

^a Department of Geoinformatics – Z_GIS, University of Salzburg, 5020, Salzburg, Austria - (ovidiu.csillik, stefan.lang)@sbg.ac.at

KEY WORDS: pixels, superpixels, OBIA, oversegmentation, runtime, eCognition

ABSTRACT:

The popular multiresolution segmentation (MRS) algorithm is time and memory consuming when dealing with large images because it uses the pixel-grid for the initial object representation. In this study, we have tested a new workflow for image segmentation of remote sensing data, starting the MRS (using the ESP2 tool) from the superpixel level (using SLIC superpixels) and aiming at dramatically reducing the amount of time and computational resources needed to automatically partition relatively large datasets of very high resolution (VHR) remote sensing images. Tests were done on Quickbird and WorldView-2 data and the results show that the proposed workflow outperforms the traditional approach (MRS starting from pixels). The computational time was reduced in all cases, the biggest improvement being from 5h 35min to 13 min, for a WorldView-2 scene with 8 bands and an extent of 12.2 million pixels. This also comes with a slight improvement of the geometric accuracy of the extracted objects. This approach has the potential to enhance the automation of big remote sensing data analysis and processing, especially when time is an important constraint.

1. INTRODUCTION

The aim of image segmentation is to partition an image into relatively homogeneous objects, non-overlapped and spatially adjacent (Blaschke, 2010). The pixel-grid is used as the underlying representation for many image segmentation algorithms, including the popular multiresolution segmentation (MRS) (Baatz and Schäpe, 2000). However, pixels are not natural entities and are unlikely to match the content of the space represented (Fisher, 1997). It would be more natural and efficient to work with perceptually meaningful entities that are derived from pixels using a low-level grouping process (Ren and Malik, 2003; Neubert and Protzel, 2012). To achieve this, we can partition an image into superpixels, which are the result of perceptual grouping of pixels based on similar characteristics (e.g. color) (Neubert and Protzel, 2012).

In computer vision, using superpixels (Achanta et al., 2012) to speed up later-stage processing are becoming increasingly popular in many applications (Achanta et al., 2012; Neubert and Protzel, 2012; Van den Bergh et al., 2012). In remote sensing few studies have used superpixels (Thompson et al., 2010; Guanyun et al., 2015; Ortiz Toro et al., 2015; Vargas et al., 2015). Even so, they did not tackle the usage of superpixels in order to improve the computational efficiency of segmenting remote sensing data.

We have tested an alternative workflow for image segmentation of remote sensing data, starting the MRS from the superpixels level and aiming at dramatically reducing the amount of time and computational resources needed to partition relatively large datasets of VHR remote sensing data. This approach is compared with the traditional one, starting the MRS from the pixel level, regarding the geometric accuracy of the objects and the computational time.

The following section (Section 2) describes the dataset used, the theoretical background of the superpixels, the workflow of generating objects and the accuracy measures. Section 3 compares the results of the two approaches in terms of accuracy and computational time. Section 4 drafts the major implications of the results, while Section 5 concludes the main findings of our study.

2. METHODS

2.1 Datasets

Tests were conducted on very high resolution remote sensing data (T1, T2 – Quickbird and T3 – WorldView-2). The spatial resolution ranges from 0.5 m (T3) to 0.6 m (T1 and T2), respectively, as well as the number of bands, between 4 and 8. The extent (in pixels) ranges between approx. 4 mil. pixels (T1), 12.2 mil. pixels (T3) and 12.3 mil. pixels (T2). Test area T1 covers a dense residential and services area in the city of Salzburg, Austria. The T2 test area is comprised of dense clustered residential and commercial buildings with large green spaces in between. Test area T3 represents a sensitive riparian habitat with forests, agricultural fields and water bodies. We have used different extents, landscapes and number of bands to assess how the computational time is affected by these factors.

2.2 Simple Linear Iterative Clustering (SLIC) superpixels

Simple Linear Iterative Clustering (SLIC) (Achanta et al., 2012; Achanta et al., 2010) have been proven to outperform other state-of-the-art superpixel methods, because of its simplicity, adherence to boundaries, computational speed and more memory efficiency (Achanta et al., 2012). SLIC has only one parameter, which is the desired number of equally sized superpixels to be generated.

* Corresponding author

SLIC is an adapted k -means clustering, but what it makes it so fast and computationally efficient is that SLIC does not compare each pixel with all pixels in the scene. For a region of approximate size $S \times S$, the distance D (which combines distance of color proximity and spatial proximity) is computed in a region $2S \times 2S$ around the superpixel center, reducing dramatically the number of D calculations. Achanta et al. (2012) found out that 10 iterations for superpixels derivation are sufficient for most images and, therefore, we are using this threshold for SLIC superpixels generation.

The same authors proposed a parameter-free SLIC version (SLICO), which generates regular shaped superpixels across the scene, regardless of textured or non-textured regions in the image, while SLIC is influenced by the texture, generating smooth regular-sized superpixels in the smooth regions and highly irregular superpixels in the textured regions (Achanta et al., 2012).

For generating superpixels, we have used a freely available GDAL implementation (GDAL-segment), available on <https://github.com/cbalint13/gdal-segment>. Superpixels were generated using SLIC and SLICO algorithms (Achanta et al., 2012), as implemented in the tool mentioned, with an initial size of clustering of 10×10 pixels. We have used this value after trial-and-error, in order to avoid an extreme oversegmentation or too large superpixels (Figure 1).

2.3 Multiresolution segmentation: pixels vs. superpixels

One of the biggest issues of MRS is the selection of parameters, of which the most important one is the scale parameter (SP). Drăguț et al. (2010) developed Estimation of Scale Parameter tool (ESP) to detect optimal scales based on local variance graph, using a single layer. Drăguț et al. (2014) extended this approach into an automated tool (ESP2) for multiple layers. The ESP2 tool is a fully automated methodology for the selection of scale parameters to extract three distinct scales using MRS, implemented in the eCognition Developer software (Trimble Geospatial).

For reasons of objectivity, we have used the ESP2 tool starting from the pixel level, as usual, and starting from the superpixel level, using a hierarchical bottom-up region merging approach (i.e. starting from an initial level, the next level is generated based on the previous one) to derive only the finest level of objects (Level 1 of the hierarchy approach of ESP2).

2.4 Segmentation accuracy of the results

We evaluated the segmentations results by comparing the geometries of resulted objects with 50 manually digitized reference objects for each test area. We have used 5 measures of accuracy: Area fit index (AFI) (Lucier and Stein, 2002), Over-segmentation (OS), Under-segmentation (US), Root mean square (D) (Clinton et al., 2010) and Quality rate (QR) (Winter, 2000), with a minimum percent overlap of 50%. In the case of a perfect match between the geometries of objects, AFI, OS, US and D would be 0 and QR would be 1. The measures are implemented into an eCognition tool by Eisank et al. (2014).

Besides geometry comparison, we measured the computational time needed for both approaches. For the ESP2 starting from pixel level we retain only the time needed to run the tool, while for ESP2 starting from superpixel level, we added to the time

ESP2 needs to run the computational time of generating superpixels.



Figure 1. Visual comparison of superpixels produced by SLIC and SLICO methods, on a subset of T2. The average superpixel size in the upper left of the image is 5×5 pixels, in the middle of the image is 10×10 pixels and 15×15 pixels in the lower right.

Test	Segmentation results			Segmentation accuracy metrics					Time	
		Number of	SP	Number of objects	AFI	OS	US	D		QR
T1	Pixels	1,403,574	69	3,109	0.499	0.560	0.121	0.405	0.414	1min 29s
	SLIC	13,835	81	2,017	0.388	0.463	0.122	0.338	0.499	27s
	SLICO	13,906	61	2,757	0.447	0.515	0.122	0.374	0.454	24s
T2	Pixels	12,696,684	172	4,670	0.174	0.229	0.067	0.169	0.729	2h 42min 40s
	SLIC	123,153	173	4,204	0.088	0.161	0.079	0.127	0.782	13min 02s
	SLICO	125,842	148	5,354	0.335	0.386	0.075	0.278	0.584	9min 49s
T3	Pixels	12,217,001	220	1,632	0.100	0.148	0.052	0.111	0.813	5h 35min 24s
	SLIC	131,415	212	1,702	0.062	0.124	0.066	0.099	0.823	13min 03s
	SLICO	121,525	172	2,338	0.223	0.275	0.066	0.200	0.688	10min 46s

Table 2. Comparison of segmentation accuracy and computational time for the three test areas (T1 to T3), using pixels, SLIC and SLICO superpixels, respectively.

3. RESULTS

All the tests were conducted on a computer station with Intel Core i5-4590 CPU (3.30 GHz) processor with 8 GB RAM, using a 64-bit Windows 7 operating system.

Generating SLIC and SLICO superpixels is a very fast and memory efficient procedure. In the case of SLIC superpixels, the scored runtime was 2s for T1, 18s for T2 and 26s for T3, while for SLICO was 3s for T1, 34s for T2 and 37s for T3. It is obvious that computing SLICO superpixels is a slightly slower procedure and this could be mainly because SLICO generates regular shaped superpixels across the scene, thus having supplementary computational constraints (i.e. compactness) (Figure 1).

The complexity of the three test areas was significantly reduced by aggregating similar pixels into superpixels. Due to this fact, SLIC and SLICO superpixels are making a big difference in segmentation, regarding total computational time. Larger the scene and bigger the number of bands, bigger the time needed for ESP2 to run starting from a pixel-grid (from 1min 29s to 5h 35min 24s) (Table 2). Runtime of ESP2 starting from SLIC superpixels is higher than that of ESP2 starting from SLICO superpixels, the difference ranging between 3s for T1 and 3min 13s for T2. This is mainly because the scale parameter detected for the latter case is smaller and therefore the processing of ESP2 stops earlier. This is explained by the fact that SLICO superpixels have compactness constraints to follow a regular lattice and, as a consequence, it can omit meaningful image boundaries, increasing the internal heterogeneity of the superpixels. SLIC superpixels have reduced the runtime by a magnitude of 229% for T1, 1146% for T2 and 2476% for T3, respectively.

The scale parameters for pixel-grid and SLIC superpixels are similar. The most evident case is for T3 (220 and 212, respectively), where approximately the same number of objects were extracted in the end (1632 and 1702, respectively).

In all cases, using SLIC superpixels outperformed the other two approaches regarding the geometric accuracy of the final objects (Table 2). Compared to the pixel approach, SLIC has better values of QR for T1 (0.499 compared to 0.414), for T2 (0.782 compared to 0.729) and for T3 (0.823 compared to 0.813). The only indicator where SLIC had worse values than the pixel approach is the US, the difference ranging from 0.001 for T1 and 0.014 for T3, but this could be considered a negligible difference when aiming at considerably improving the runtime. The oversegmentation of the scene using SLICO superpixels

negatively impacts the accuracy measures and, therefore, they are better than the pixel approach only in the case of T1 (QR of 0.454 compared to 0.414), where the number of objects is slightly smaller. In test area T1, even if there is a big difference in number of objects extracted between pixel approach and SLIC superpixels (3109 and 2017, respectively), the US has the same value, while OS decrease by 0.1. This could mean that SLIC superpixels are better adhering to the boundaries and, therefore, creating more meaningful objects than those generated by starting from a pixel-grid.

4. DISCUSSION

Using the adapted workflow, we are able to increase the maximum extents on which ESP2 for MRS runs. Supplementary tests (not shown here) proven that for scenes of tens of millions of pixels ESP2 is successfully fast and memory efficient when starting from superpixels, while in the case of starting from pixels it crashes due to the immense amount of resources needed to compute the statistics at the pixel-level. The size of the superpixels should carefully be chosen, not to worsen the computational efficiency (by generating too small superpixels) or to contain more than one class inside a superpixel (by having too coarse superpixels).

We have used SLIC and SLICO superpixels because they were shown that outperform the other state-of-the-art superpixel algorithm (Achanta et al., 2012). In our study, applying MRS on SLIC superpixels has proven to be the most efficient in terms of geometric accuracy of final objects. One minor drawback of using SLIC and SLICO superpixels in MRS is that we introduce further parameterization in the process, but as implemented in the GDAL-segment tool, the only parameters you have to set is the desired size of the generated superpixels, that means how finer the generated superpixels should be.

Using VHR images with different characteristics gave us an overview of their effects in generating superpixels. For the same size of desired superpixels, the larger the extent of the scene and higher the number of layers, the higher the computational time. But the most important influence over the runtime is the extent of the generated superpixels: smaller superpixels will require longer time than larger superpixels.

5. CONCLUSION

In this paper, we have tested a workflow to efficiently partition an image into objects by using SLIC superpixels as the starting point for MRS. When compared to the traditional approach (starting from pixel-grid), our approach outperformed both in terms of geometric accuracy of the extracted objects and computational time. This approach has the potential to enhance the automation of big remote sensing data analysis and processing, especially when time is an important constraint.

ACKNOWLEDGEMENTS

This work was supported by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W1237-N23). WorldView-2 imagery was provided through the FP7 Project MS.MONINA (Multi-scale Service for Monitoring NATURA 2000 Habitats of European Community Interest), Grant agreement No. 263479 and the INTERREG Project EuLE (EuRegional Spatial Analysis). Quickbird imagery was kindly provided by the Department of Geoinformatics - Z_GIS, Salzburg, Austria.

REFERENCES

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34, pp. 2274-2282.
- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2010. Slic superpixels. École Polytechnique Fédérale de Lausanne (EPFL), Technical Report 149300.
- Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation-an optimization approach for high quality multi-scale image segmentation, In: *Angewandte Geographische Informationsverarbeitung*. Strobl, J., Blaschke, T., Griesebner, G. (Eds.), Wichmann-Verlag, Heidelberg, pp. 12-23.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, pp. 2-16.
- Clinton, N., Holt, A., Scarborough, J., Yan, L.I., Gong, P., 2010. Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering and Remote Sensing*, 76, pp. 289-299.
- Drăguț, L., Csillik, O., Eisank, C., Tiede, D., 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, pp. 119-127.
- Drăguț, L., Tiede, D., Levick, S., 2010. ESP: a tool to estimate scale parameters for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24, pp. 859-871.
- Eisank, C., Smith, M., Hillier, J., 2014. Assessment of multiresolution segmentation for delimiting drumlins in digital elevation models. *Geomorphology*, 214, pp. 452-464.
- Fisher, P., 1997. The pixel: a snare and a delusion. *International Journal of Remote Sensing*, 18, pp. 679-685.
- Guangyun, Z., Xiuping, J., Jiankun, H., 2015. Superpixel-Based Graphical Model for Remote Sensing Image Mapping. *Geoscience and Remote Sensing, IEEE Transactions on*, 53, pp. 5861-5871.
- Lucieer, A., Stein, A., 2002. Existential uncertainty of spatial objects segmented from satellite sensor imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 40, pp. 2518-2521.
- Neubert, P., Protzel, P., 2012. Superpixel benchmark and comparison, *Proc. Forum Bildverarbeitung*, pp. 1-12.
- Ortiz Toro, C., Gonzalo Martín, C., García Pedrero, Á., Menasalvas Ruiz, E., 2015. Superpixel-Based Roughness Measure for Multispectral Satellite Image Segmentation. *Remote Sensing*, 7, pp. 14620-14645.
- Ren, X., Malik, J., 2003. Learning a classification model for segmentation, In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. IEEE*, pp. 10-17.
- Thompson, D.R., Mandrake, L., Gilmore, M.S., Castano, R., 2010. Superpixel Endmember Detection. *Geoscience and Remote Sensing, IEEE Transactions on*, 48, pp. 4023-4033.
- Van den Bergh, M., Boix, X., Roig, G., de Capitani, B., Van Gool, L., 2012. Seeds: Superpixels extracted via energy-driven sampling, In: *Computer Vision—ECCV 2012. Springer*, pp. 13-26.
- Vargas, J., Falcao, A., dos Santos, J., Esquerdo, J., Coutinho, A., Antunes, J., 2015. Contextual superpixel description for remote sensing image classification, In: *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, pp. 1132-1135.
- Winter, S., 2000. Location similarity of regions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55, pp. 189-200.

Revised July 2016