GI_Forum 2017, Issue 1
Page: 244 - 252
Short Paper
Corresponding Author:
stefan.lang@sbg.ac.at
DOI: 10.1553/giscience2017_01_s244

ETRS Grid-constrained Superpixel Generation in Urban Areas Using Multi-Sensor very High Resolution Imagery

Stefan Lang and Ovidiu Csillik
[1]University of Salzburg, Austria

Abstract

Massive amounts of geospatial data require better techniques to analyse and display them. A common practice in applying statistical analysis and reporting units is represented by the regional statistical unit grids, which are used for national and European statistics and reporting. In line with the EU's INSPIRE directive, Austria provides statistics constructed on the European Terrestrial Reference System, which uses the Lambert Azimuthal Equal Area projection (ETRS-LAEA) for spatial analysis and display. Although these units are fixed and allow comparability and replicability, there is no relation with the underlying phenomenon. In this study, we evaluate the suitability of SLICO superpixels to replace the artificial ETRF grid squares when delineating an image for further analysis and using reporting units, or for display purposes. In this approach, we minimize any further parametrization, which is introduced by many other available segmentation algorithms, and aim to replicate the ETRF grid by imposing a size constraint and using the square's centres as seeds for superpixel generation.

Keywords:

SLICO superpixels, cell aggregation, multi-scale, image segmentation

1 Introduction

Cell aggregation ('resampling') is one approach to upscale detailed geospatial information (Hay et al., 2001). Similar to these scaled representations, regional statistical unit grids are used for national and European-wide statistics and reporting, including for pan-European geospatial information layers – as for example those produced by the Copernicus Land Monitoring (CLM) service. The CLM provides high-resolution layers for various land cover parameters such as forest, sealing degree or grassland distribution, aggregated to 20x20m cells. Many geospatial datasets are available at aggregated levels based on a regular grid cell and, depending on the reporting units being used, may introduce a bias into the analysis since these are artificial delineations of geographic reality (Hagenlocher et al., 2014; Hengl, 2006).

Suitable for grouping are those units that are reasonably small (Woodcock and Strahler, 1987), more precisely some 3–4 times smaller than the object of interest. Since here the size of the target objects depends on the policy scale and the level of detail, the minimum size of the grid cell needs to be estimated, but it cannot be derived deterministically. Census authorities in countries such as Finland, Sweden or Austria maintain socio-economic and other relevant data based on address location. Statistik Austria provides socio-economic indicators in standardized gridded data sets with a cell size in multiples of 125m, alongside the standard products for hierarchical administrative levels (Wonka and Strobl, 2006). Such data hosted and managed according to the European INSPIRE directive can neatly be integrated into the existing transnational terrestrial reference frames (e.g. ETRS-LAEA) (Annoni and Smits, 2003). Even more importantly, they allow for any (re-)aggregation and upscaling, independently of given administrative units. A regular grid cell has no particular meaning in terms of the studied phenomena on the ground, but enumeration units are rarely significant either (Byrne, 1998). Overall, grid cells have a neutral geometry and imply a predictable MAUP (Openshaw, 1984) effect. Thus, aggregating grid cells into new zones such as qualitatively established neighbourhoods may be preferable over aggregating enumeration areas.

Using resampling techniques, any continuous grid can be transferred into another one with larger (or smaller) cell size. VHR imagery with 0.5m resolution can be resampled into 10x10m cells by averaging the pixel values and assigning the resulting mean value to the (new) overlying cell. Categorical data can be resampled by using the majority of classes that occur. In terms of the studied phenomenon, the aggregation grid is deliberately 'neutral', e.g. the European transnational reference grid and national grids from it, such as the Statistik Austria reference grid. The adverse effect of the ecological fallacy, i.e. treating all group members within a zone the same way yet with arbitrary zone boundaries, is compensated for only by the fact that the units being used are fixed and therefore comparability is guaranteed in repeated studies.

Object-based image analysis (OBIA), or more general spatial image analysis, builds on hierarchically structured image objects (Blaschke et al., 2014; Lang, 2008). When working with image objects instead of pixels, we face the problem referred to above: the reference grid boundaries do no usually match the object boundaries, so that image objects are arbitrarily cut. It would be a coincidence to reach a 1:n relation between the aggregation cell and the underlying image objects. Studies assessing functional land cover classes (Lang et al., 2014) or spatial vulnerability units (Kienberger et al., 2009) used a grid resampling as an intermediate step in the construction of geons as units of uniform behaviour in terms of the underlying phenomenon (Lang et al., 2014). This step of preparing regular grid cells could be optimized: rather than using regular grid cells, methods of superpixel generation could be used to create the intermediate level, which matches more closely the actual spectral characteristics of the underlying image. The seed pixels for computing the superpixels would then match the centroids of the grid.

In this study, we discuss a novel approach: constraining image segmentation by borrowing size and position parameters from the reference grid. The idea is to use the spacing of the ETRF reference grid (and its sub-divisions) to control both the average size and the centroids for superpixel segmentation. This would result in a 1:1 relation between generated objects (superpixels) and reference units, and a 1:n relation between objects and multiples of reference units. The question remains: would superpixels still represent 'reasonable' objects? The hypothesis is that in urban settings, where there is a continuous fabric of built-up structures, the grid-constraint superpixels would lead to a satisfactory provision of image objects, which would be used as building blocks for further aggregation.

2 Material and Methods

Study area and datasets

The study area is located in the southern part of the city of Salzburg, Austria. The river Salzach flows from south to north, and the urban area comprises mainly dense residential and commercial areas separated by large green spaces with economic or recreational functions. The very high resolution datasets were acquired at equal time intervals in 2005, 2010 and 2015, although using different sensors (Table 1). The number of spectral bands varied between four (QuickBird and Pleiades) and eight (Worldview-2), and all images had red and near-infrared bands, which are important in differentiating the vegetation from other urban components.

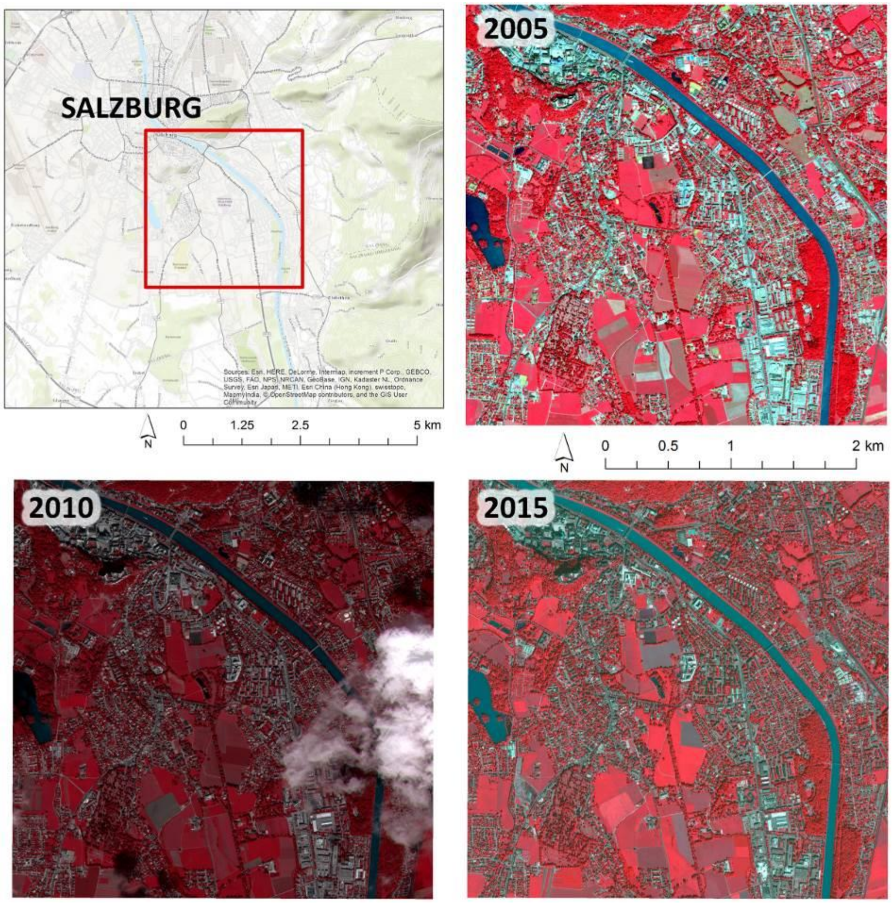

Figure 1: Location of the study area within the city of Salzburg, Austria, and the three temporal datasets used: 2005 (QuickBird), 2010 (WorldView-2) and 2015 (Pleiades).

Table 1 shows the characteristics of the VHR imagery used and the extent of the respective subsets (in pixels).

Table 1: Dataset characteristics.

Imagery	Spatial resolution (m)	Number of bands	Extent (pixels)	Date of acquisition
QuickBird	0.6	4	5583×5500	June 25, 2005
WorldView-2	0.5	8	10460×10444	Sept. 10, 2010
Pleiades	0.5	4	9963×9941	Oct. 1, 2015

SLIC superpixel-generation based on ETRF grid

In line with the EU's INSPIRE directive, Austria provides statistics constructed on a pan-European grid based on the European Terrestrial Reference System, which uses the Lambert Azimuthal Equal Area projection (ETRS-LAEA) for spatial analysis and display. Using a regional statistical grid which covers the entire territory of a country can assure transferability and uniformity of statistics, but the statistics are not related to the administrative boundaries or to the existing discontinuities in the landscape. An alternative to these rigid grids is to aggregate pixels into similar-sized image objects as the initial cell size. The resultant objects, superpixels (Shi and Malik, 2000), would better represent the information within a satellite image.

Superpixels have a number of advantages when compared to pixels or, in our case, to the rigid ETRF grid: they (1) are perceptually meaningful regions (Achanta et al., 2012); (2) have a low computational complexity (Ren and Malik, 2003); (3) preserve the structure in an image (Ren and Malik, 2003), and (4) have better adherence to the natural boundaries of features (Neubert and Protzel, 2012). Many superpixel algorithms are available in the literature (Neubert and Protzel, 2012), but the Simple Linear Iterative Clustering (SLIC) algorithm has been shown to outperform other state-of-the-art methods (Achanta et al., 2012; Csillik, 2016). The SLIC algorithm is an adaptation of k-means clustering, starting from a regular grid and having the initial seeds as the centroids of the cells (Figure 2). At this stage, we have a representation that is similar to the official ETRF grid – a rigid structure made of equally distributed squares, for which we can synchronize the seeds to match the centres of the squares. Furthermore, each of the seeds is shifted in a 3×3 window, in order to reduce its susceptibility to be being placed on an outlier. An iterative procedure is performed, assigning each pixel to the nearest cluster centre based on a distance measure (Eq. 1). This distance measure combines a distance of colour proximity (Eq. 2) and a distance of spatial proximity (Eq. 3):

$$\mathrm{D} = \sqrt{\left(\frac{\mathrm{d_c}}{\mathrm{m}}\right)^2 + \left(\frac{\mathrm{d_s}}{\mathrm{S}}\right)^2}; \tag{1}$$

$$d_c = \sqrt{\sum_{s_p \in B} \left(I(x_i, y_i, s_p) - I(x_j, y_j, s_p) \right)^2}; \quad (2)$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \quad (3)$$

where d_c and d_s are the colour and spatial distances between pixels $I(x_i, y_i, s_p)$ and $I(x_j, y_j, s_p)$ in the spectral band s_p, B represents the number of spectral bands used, S is the sampling interval of the seeds, and m dictates the compactness of the superpixels (Ortiz Toro et al., 2015). In order to produce regular and equal-sized SLIC superpixels, a compactness constraint needs to be added (optimized SLIC, known as SLICO), the resulting superpixels being less sensitive to the texture of the image. We used SLICO superpixels in order better to maintain a similarity with the ETRF grid.

Figure 2: The overlap between ETRF grid and the SLICO superpixels generated. The centre of each ETRF grid cell can be designated as the starting seed in generating superpixels. In this way, the superpixels generated coincide closely with the ETRF grid, but with a better adherence to the actual image boundaries.

3 Results and outlook

The grid-constrained SLICO method produces similar-sized and similarly-positioned superpixels for different input sensor data (see Figure 3), despite the different sensor resolutions and image acquisition dates: QuickBird (0.6m, 2005), WorldView-2 (0.5m, 2010), and Pléiades (0.5m 2015). However, the power of the superpixels to represent reasonable building blocks may be limited when linear features with a width of less than double the cell size, such as narrow rivers or roads, need to be represented. The bottom line of Figure 3 shows the superpixels generated with a slightly smaller ('adaptive') average size of 50 x 50m in order to better capture linear features such as bridges.

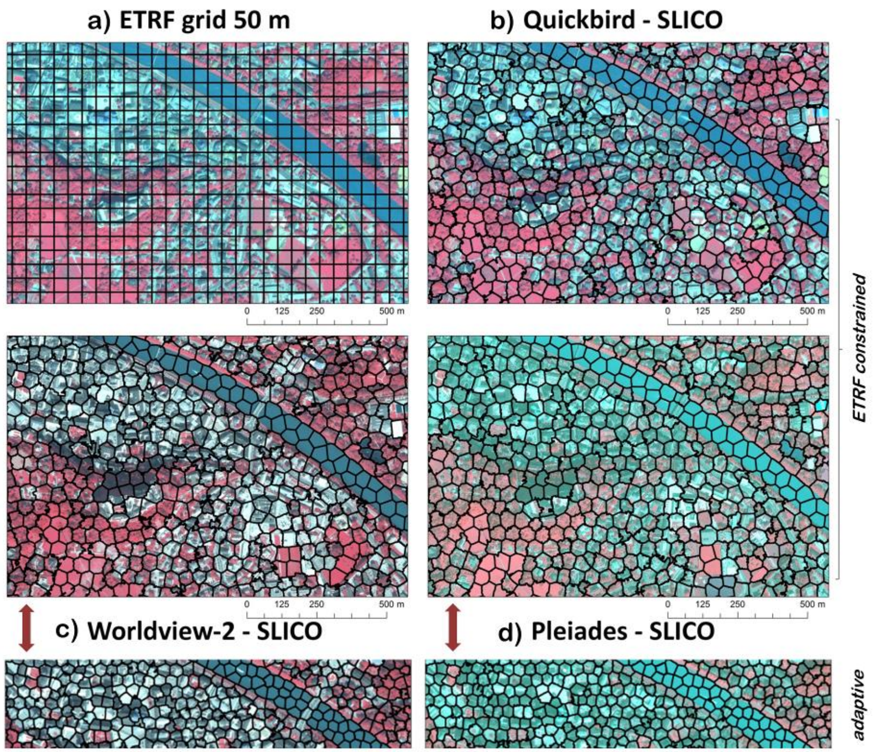

Figure 3: Results of the superpixel generation for the three VHR sensor types and dates.

A qualitative assessment based on the visual comparison between the ETRF rigid grid and the SLICO superpixels generated shows that the latter better adhere to the image-feature boundaries (Figure 4). Starting from the centroids of the ETRF grid and using the same cell size in constructing the superpixels does not ensure that both will overlap perfectly. In the case of QuickBird image, we had 4,422 reference grid squares and 4,273 superpixels. This difference is due to the fact that the SLICO algorithm iteratively refines the boundaries of the superpixels until it reaches a stable delineation of the image.

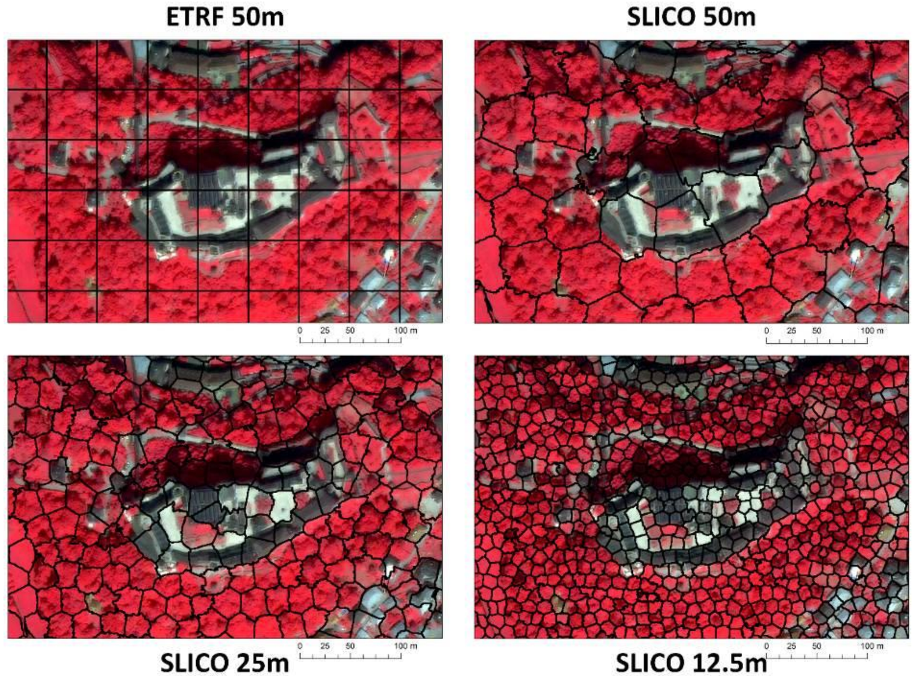

Figure 4: A comparison between a 50m-square grid and multi-scale representations of the same scene (the fortress in Salzburg, surrounded by green areas) using SLICO superpixels of 50m, 25m and 12.5m. In all four images, the grid or pixels are laid over an image produced by the QuickBird 432 band combination.

The standard deviation inside the ETRF grid ranges between 2.29 and 444, while for the SLICO superpixels it is lower, between 2.5 and 397. This is seen in Figure 5, where the SLICO segmentation better conserves the spatial structure within the image, while the artificial grid overlaps the image features randomly.

In this study, we evaluated the suitability of SLICO superpixels to replace the artificial ETRF grid squares in delineating an image for further analysis based on reporting units, or for display. In this approach, we control parametrization, which is introduced by many other available segmentation algorithms, and aim to replicate the ETRF grid by using the cell centres as seeds for the generation of superpixels. In future studies, this approach could be used as a methodological contribution to monitoring studies to assess for example the status and dynamics of urban greenness. As a means to such ends, the generation of geons as units of uniform green impression could prove to be more suitable if it is based on superpixels rather than on grid cells. A scale-adaptive strategy for 'over-'segmentation and for reducing the complexity of an image by grouping similar pixels, superpixels form the building blocks for further aggregation. A challenge remains with the representation of linear features that need to be captured by an adaptive spacing of the initial sampling grid.

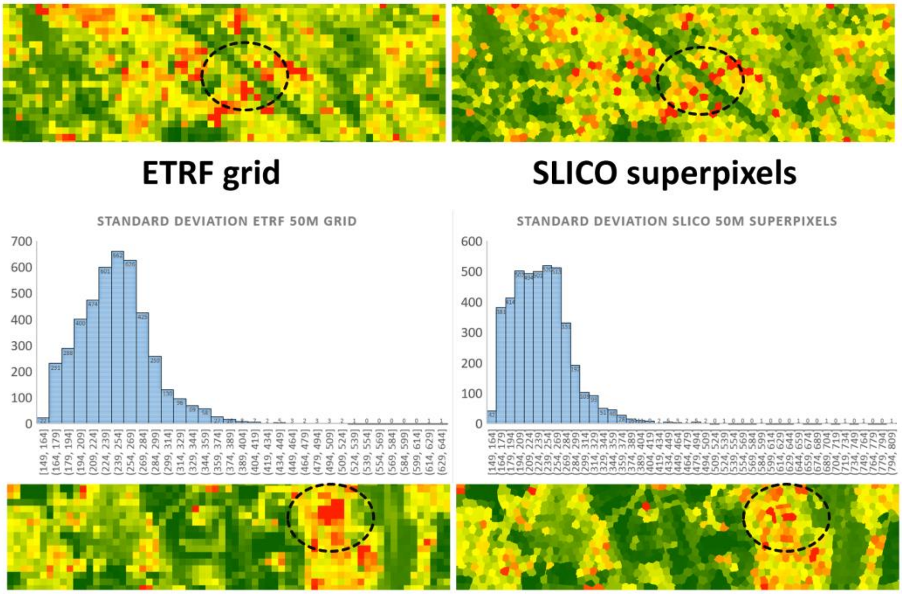

Figure 5: Standard deviation for the ETRF grid and similar-sized generated SLICO superpixels. The circles indicate corresponding areas where superpixels show a more adaptive behaviour than the grid cells.

Acknowledgements

This work was supported by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W1237-N23). Satellite imagery was kindly provided by the Department of Geoinformatics – Z_GIS, Salzburg, Austria. We acknowledge time and effort of the anonymous reviewers and proofreaders.

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(11) 2274-2282.

Annoni, A., Smits, P., 2003. Main problems in building European environmental spatial data. International Journal of Remote Sensing 24(20) 3887-3902.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS Journal of Photogrammetry and Remote Sensing 87(0) 180-191.

Byrne, D.S., 1998. Complexity theory and the social sciences: an introduction. Psychology Press.

Csillik, O., 2016. Superpixels: the end of pixels in OBIA. A comparison of state-of-the-art superpixel methods for remote sensing data, In: Kerle, N., Gerke, M., Lefevre, S. (Ed.), GEOBIA 2016 : Solutions and Synergies. University of Twente Faculty of Geo-Information and Earth Observation (ITC).

Hagenlocher, M., Lang, S., Holbling, D., Tiede, D., Kienberger, S., 2014. Modeling Hotspots of Climate Change in the Sahel Using Object-Based Regionalization of Multidimensional Gridded Datasets. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 7(1) 229-234.

Hay, G.J., Marceau, D.J., Dubé, P., Bouchard, A., 2001. A multiscale framework for landscape analysis: Object-specific analysis and upscaling. Landscape Ecology 16(6) 471-490.

Hengl, T., 2006. Finding the right pixel size. Computers & Geosciences 32(9) 1283-1298.

Kienberger, S., Lang, S., Zeil, P., 2009. Spatial vulnerability units-Expert-based spatial modelling of socio-economic vulnerability in the Salzach catchment, Austria. Natural Hazards and Earth System Sciences 9 767-778.

Lang, S., 2008. Object-based image analysis for remote sensing applications: modeling reality–dealing with complexity, Object-based image analysis. Springer, pp. 3-27.

Lang, S., Kienberger, S., Tiede, D., Hagenlocher, M., Pernkopf, L., 2014. Geons – domain-specific regionalization of space. Cartography and Geographic Information Science 41(3) 214-226.

Neubert, P., Protzel, P., 2012. Superpixel benchmark and comparison, Proc. Forum Bildverarbeitung, pp. 1-12.

Openshaw, S., Openshaw, S., 1984. The modifiable areal unit problem. Geo Abstracts University of East Anglia.

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(11) 14620.

Ren, X., Malik, J., 2003. Learning a classification model for segmentation, Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. IEEE, pp. 10-17.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on 22(8) 888-905.

Wonka, E., Strobl, J., 2006. Regionalstatistik in Österreich: von der Tabelle zu räumlicher Analyse und Visualisierung. Fachbereich Geographie und Geologie der Universität Salzburg.

Woodcock, C.E., Strahler, A.H., 1987. The factor of scale in remote-sensing. Remote Sensing of Environment 21(3) 311-332.