

SUPERPIXELS: THE END OF PIXELS IN OBIA. A COMPARISON OF STATE-OF-THE-ART SUPERPIXEL METHODS FOR REMOTE SENSING DATA

O. Csillik *

Department of Geoinformatics – Z_GIS, University of Salzburg, 5020, Salzburg, Austria - ovidiu.csillik@sbg.ac.at

KEY WORDS: SLIC, SLICO, SEEDS, LSC, segmentation, computer vision

ABSTRACT:

In computer vision, using superpixels or perceptually meaningful atomic regions to speed up later-stage processing are becoming increasingly popular in many applications. Superpixels are used as a pre-processing stage to organize an image into a low-level grouping process through oversegmentation, thus simplifying the computation in later stages. However, in remote sensing domain few studies use superpixels. Even so, there is no comparison between superpixel methods and their suitability for remote sensing images. In this study, we compare four state-of-the-art superpixel methods: Simple Linear Iterative Clustering (SLIC and SLICO), Superpixels Extracted via Energy-Driven Sampling (SEEDS) and Linear Spectral Clustering (LSC). We applied them to very high resolution remote sensing data of different characteristics (extent, spatial resolution and landscape complexity) in order to see how superpixels are affected by these factors. The four algorithms were compared regarding their computational time, ability to adhere to image boundaries and the accuracy of the resulted superpixels. Furthermore, we discuss the individual strengths and weaknesses of each algorithm and draw further applications of superpixels in OBIA.

1. INTRODUCTION

Nowadays, remote sensing community have to deal with the big volume, variety and velocity of the acquired geospatial data (Chen et al., 2015). Besides many advantages, this also comes with shortcomings, like the limited computational capabilities and a lack of powerful tools to handle the complexity of the data, especially when time is an important constrain in delivering of high quality geospatial information (e.g. Tiede et al., 2011).

What makes the computation so intensive is that many image segmentation algorithms and spatial analysis have as the underlying representation the pixel-grid. In the case of very high resolution (VHR) data the local spatial autocorrelation between the pixels is high, thus an object will be composed by many pixels with the same characteristics (Chen et al., 2012). To overcome this, it would be more natural and efficient to work with superpixels, which are an oversegmentation of the image, a low-level grouping of similar pixels in agreement to some desired homogeneity criterions (e.g. color) (Ren and Malik, 2003; Neubert and Protzel, 2012). Moving to superpixels allows us to measure feature statistics on a naturally adaptive domain rather than on a fixed window (Fulkerson et al., 2009). Furthermore, creating the final objects that matches the reality is a simpler task, by finding the superpixels which are part of the object (Fulkerson et al., 2009).

Using superpixels instead of pixels in the segmentation process have certain advantages: (1) superpixels carry more information than pixels and adhere better to the natural image boundaries (Neubert and Protzel, 2012; Guangyun et al., 2015); (2) superpixels are perceptually meaningful objects, having the scale between the pixel level and the object level (Achanta et al., 2012; Neubert and Protzel, 2012); (3) superpixels are of low computational complexity and can speed-up the subsequent

image processing (Ren and Malik, 2003; Li and Chen, 2015); (4) superpixels reduce the susceptibility to noise and outliers and capture image redundancy (Shi and Wang, 2014) and (5) because superpixels are results of an oversegmentation, most structures in the image are conserved (Ren and Malik, 2003).

In computer vision, superpixels are used in many applications. However, in remote sensing superpixels are not used to their full capacity, even if they can improve or speed-up later processing. The main objective of this study is to give a first comparison of four state-of-the-art superpixel methods, applied on remote sensing data.

The following section (Section 2) describes the datasets used, briefly explains the superpixel algorithms and the evaluation methodology. Section 3 compares the results, while Section 4 explains and discusses the main findings and further directions.

2. METHODS

2.1 Datasets

We have used very high resolution remote sensing data: a Quickbird scene of 0.6 m spatial resolution and WorldView-2 scene of 0.5 m spatial resolution, as well as a LiDAR derived DSM, of 1 m spatial resolution. All three datasets have an extent of 4 million pixels. In the rest of the article, we will refer to the datasets as QB (Quickbird), WV2 (WorldView-2) and DSM (Digital Surface Model). QB and WV2 datasets are located in the city of Salzburg, with residential, industrial and urban green areas, while DSM covers a rural area, with sparse housing and forest patches. In order to run the superpixel algorithms, all three datasets were used in a RGB combination with an 8-bit color depth, for QB and WV2 the red, green and blue bands being used.

* Corresponding author

2.2 Superpixel algorithms

Many superpixel algorithms exist, each with its advantages and limitations (Neubert and Protzel, 2012; Achanta et al., 2012). For this study, we compared 4 state-of-the-art superpixel algorithms, namely Simple Linear Iterative Clustering (SLIC and SLICO – parameter free) (Achanta et al., 2012), Superpixels Extracted via Energy-Driven Sampling (SEEDS) (Van den Bergh et al., 2012) and Linear Spectral Clustering (LSC) (Li and Chen, 2015). In computer vision, the 4 algorithms were found to be very efficient and accurate, outperforming many existing algorithms (Achanta et al., 2012; Van den Bergh et al., 2012; Li and Chen, 2015). For derivation of superpixels, we have used the open-source GDAL implementation, available on <https://github.com/cbalint13/gdal-segment>. In the following paragraphs we are briefly describing the methods and, for more details, the reader is referred to the sources mentioned for each algorithm.

SLIC is a gradient-ascent based algorithm, which starts from a rough initial clustering of pixels and iteratively refine the clusters until some criterions are met to form the superpixels (Achanta et al., 2012). SLIC is an adapted k -means clustering and, by default, the only parameter of the algorithm is k , the desired number of approximately equally sized superpixels (Achanta et al., 2012). What makes SLIC fast and computational efficient is that, when clustering the pixels, it does not compare each pixel with all pixels in the scene. For a region of size $S \times S$, a distance D (distance of color proximity and spatial proximity) is computed around the superpixel center, minimizing the number of D calculations, and, therefore, improving the speed over conventional k -means clustering, where each pixel must be compared with all cluster centers (Achanta et al., 2012).

The same authors (Achanta et al., 2012) proposed a parameter-free version of SLIC (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).

SEEDS algorithm is a simple hill-climbing optimization which starts from an initial superpixel partitioning and continuously refines the superpixels by modifying the boundaries (Van den Bergh et al., 2012). The algorithm is based on a robust and fast to evaluate energy function, based on enforcing color similarity between the boundaries and the superpixel color histogram (Van den Bergh et al., 2012).

In LSC, each image pixel is mapped to a point in a ten dimensional feature space where weighted k -means is applied for segmentation (Li and Chen, 2015). Non-local information is implicitly preserved due to the equivalence between the weighted k -means clustering in this ten dimensional feature space and normalized cuts in the original pixel space (Li and Chen, 2015).

For a better overview of the algorithms, for each test area we have derived superpixels starting from initial sizes of 5×5 , 10×10 , 15×15 and 20×20 , respectively. A number of 10 iterations for superpixels clustering and refinement was used for each method.

2.3 Evaluation

Comparison of the algorithms was done qualitatively and quantitatively. Each superpixel segmentation results were visually inspected in order to draw conclusions about the quality of boundary adherence of each algorithm. Since time is an important issue of an algorithm, we also retained the time needed for each algorithm to derive the superpixels. As a measure of internal homogeneity of the superpixels, we compared the overall standard deviation (SD) computed for the derived superpixels.

3. RESULTS

3.1 Runtime

The fastest algorithm is SEEDS, while the slowest is LSC (Figure 1). When increasing the size of the superpixels, the differences in time between the algorithms tend to significantly decrease, LSC and SLIC having similar values. This is due to the fact that a small number of superpixels are derived (20×20 pixels, approx. 10.000 superpixels) and the computational effort is less significant compared to the derivation of the finest superpixel segmentation (5×5 pixels, approx. 160.000 superpixels). LSC and SLICO are slower compared to SLIC and SEEDS and this can be explained by the fact that the former two are having additional compactness constraints, deriving a more regular lattice of superpixels than the latter two (e.g. Figure 2).

Because for all three test areas we have used three bands as input, there are no significant differences between datasets, for the same method.

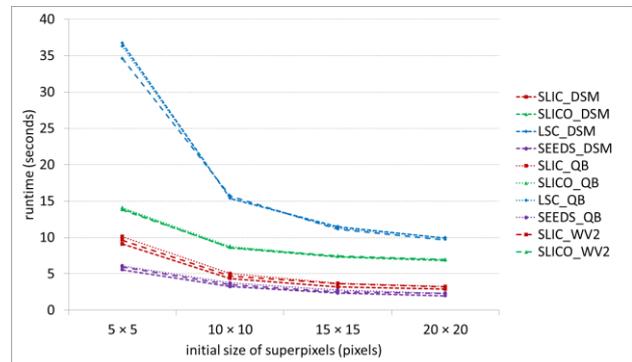


Figure 1. Runtime (in seconds) for each algorithm and for each test area, starting from initial size of the superpixels of 5×5 , 10×10 , 15×15 and 20×20 , respectively.

3.2 Visual evaluation

In the case of an initial size of superpixels of 5×5 pixels, the oversegmentation is high and the superpixels boundaries tend to follow the boundaries of natural features within the test areas (Figure 2, 3 and 4). Since the approx. size of superpixels is 25 pixels, there are low chances that a superpixel contains information from more than one class. For DSM test area, due to the fine transition between elevation of low buildings and their surroundings, LSC algorithm can omit some of these boundaries.

When increasing the size of superpixels (10×10 pixels), patterns starts to be visible in the scenes. Even if a superpixel is now approx. 100 pixels in size, there is a good adherence of the superpixels boundaries to the features in the images and, therefore, low changes that a superpixel contains information of more than one class. In non-textured regions (e.g. industrial buildings in Figure 3), all the algorithms produce superpixels that come closer to a regular lattice.

The initial superpixel size of 15×15 pixels is appropriate for detection of buildings in DSM test area (Figure 4), for SLIC, LSC and partially SLICO algorithms. In the same case, SEEDS have the tendency to create false superpixels where there is

a smooth transitions of boundaries (Figure 4). At this size of superpixels, individual trees and road segments can accurately be detected by all four algorithms. For 15×15 pixels size for superpixels there is a higher chance that a superpixel contains information from more than one class.

The last tested size of superpixels, 20×20 pixels, is reaching the capabilities of some algorithms to follow the correct boundaries, because many objects inside the scene have the size smaller than the size of superpixels. However, objects with similar or larger size are having the boundaries well detected (e.g. road segments, grass fields, rooftop parts, trees and forest patches in Figure 2 and 3). In the case of DSM, only SLIC and partially



Figure 2. Results for the QB test area representing the LSC, SEEDS, SLIC and SLICO segmentation, starting from initial size of the superpixels of 5×5 , 10×10 , 15×15 and 20×20 , respectively.

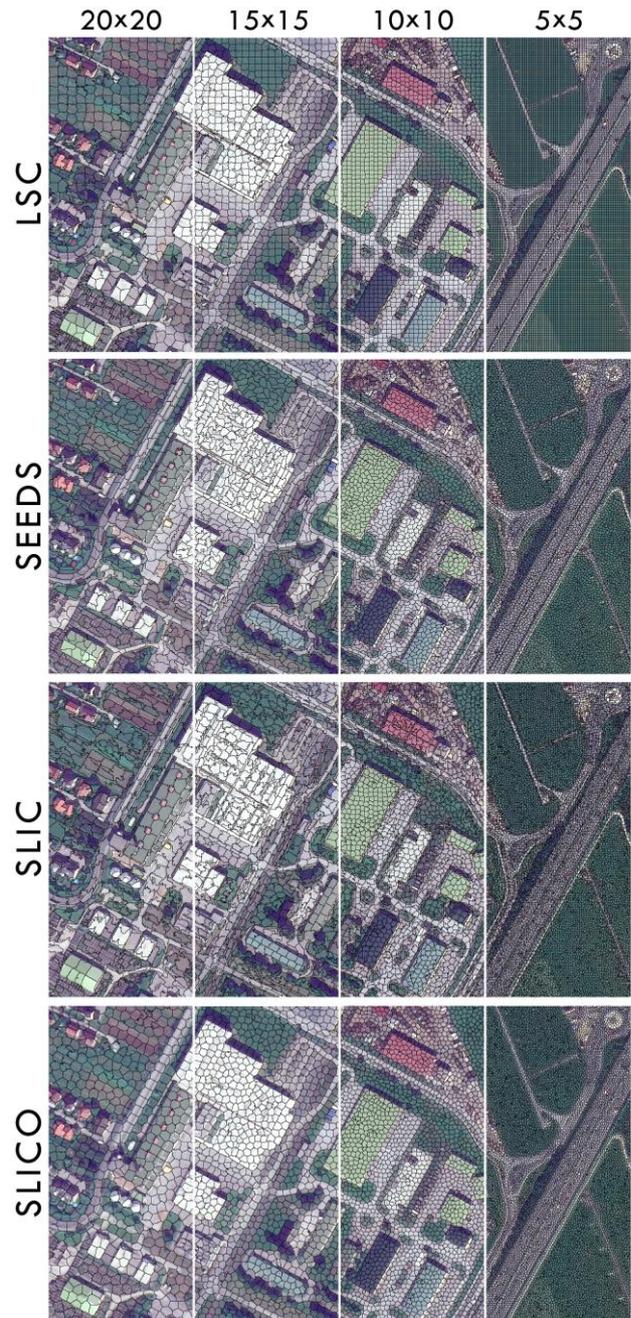


Figure 3. Results for the WV2 test area representing the LSC, SEEDS, SLIC and SLICO segmentation, starting from initial size of the superpixels of 5×5 , 10×10 , 15×15 and 20×20 , respectively.

SLICO superpixels adhere to the boundaries of buildings, while LSC and SEEDS fails (Figure 4). In the same test area, LSC performs better at delineating the boundaries of the forest patch (Figure 4).

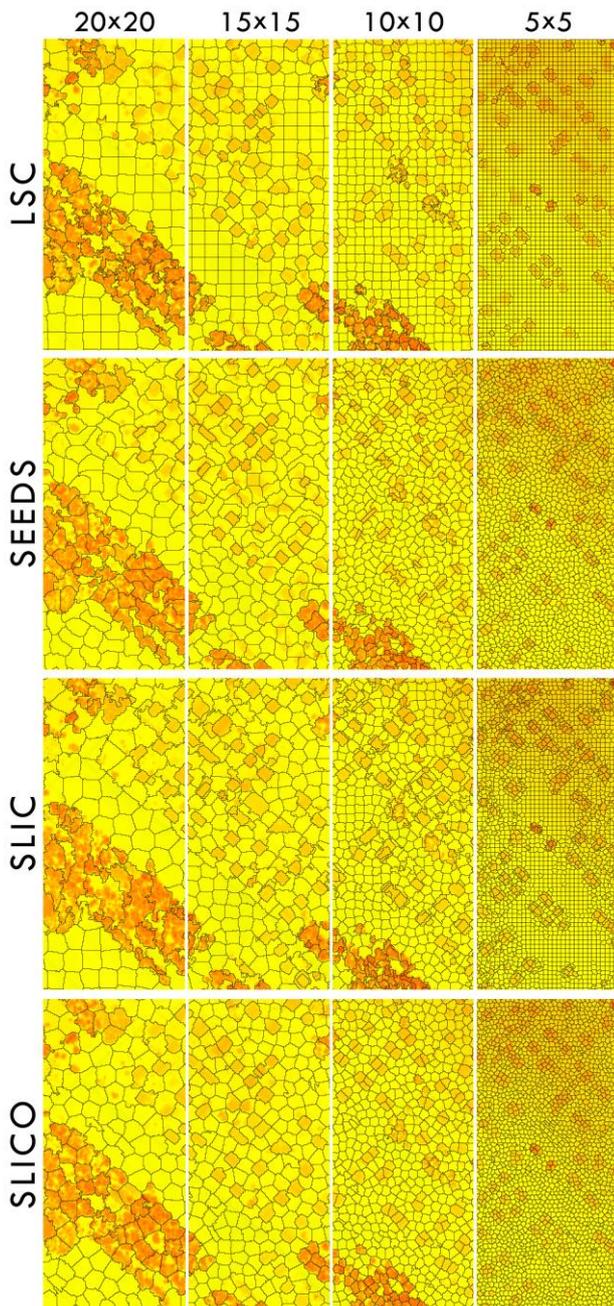


Figure 4. Results for the DSM test area representing the LSC, SEEDS, SLIC and SLICO segmentation, starting from initial size of the superpixels of 5x5, 10x10, 15x15 and 20x20, respectively.

3.3 Superpixels homogeneity

In the case of QB test area there are no big differences regarding the SD of the final superpixels (Figure 5). However, small differences occur between the algorithms. At the finest level, SLIC has the most internal homogeneous superpixels while SEEDS is on the opposite side. At the coarser level, LSC, SLIC

and SEEDS have similar values of SD of approx. 15, while SLICO superpixels have a SD value of approx. 17.

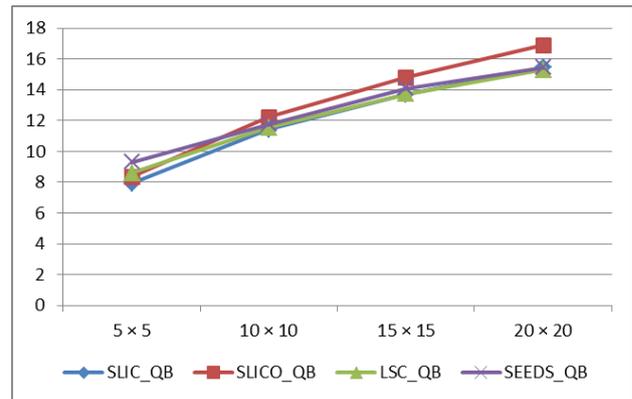


Figure 5. SD values (vertical axis) of the superpixels from QB test area, starting the algorithms from initial size of the superpixels of 5x5, 10x10, 15x15 and 20x20, respectively.

In the case of WV2 the differences regarding the SD of the final superpixels are more obvious than in the previous test area (Figure 6). SLIC algorithm outperform the other at all sizes of the generated superpixels. As in the previous study area, SLICO superpixels have good values for smaller sizes of the superpixels and tend to get worse at the coarser scale, because of the compactness constraints. LSC and SEEDS have similar values for all four sizes of superpixels.

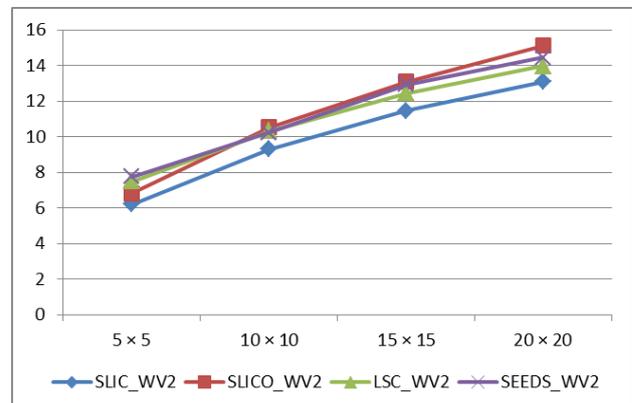


Figure 6. SD values (vertical axis) of the superpixels from WV2 test area, starting the algorithms from initial size of the superpixels of 5x5, 10x10, 15x15 and 20x20, respectively.

In the last test area (DSM), SLIC superpixels are having the best values of SD at all tested sizes (Figure 7). The SD of LSC superpixels is significantly getting worse as the size of the superpixels increases. This can be explained by the fact that LSC mixes the buildings with their surroundings at the size of superpixels of approx. 20x20 pixels, while the others don't (Figure 4). The SEEDS superpixels have the worse SD for the finest size of superpixels, but tends to reduce the difference to SLIC as the size of the superpixels increases.

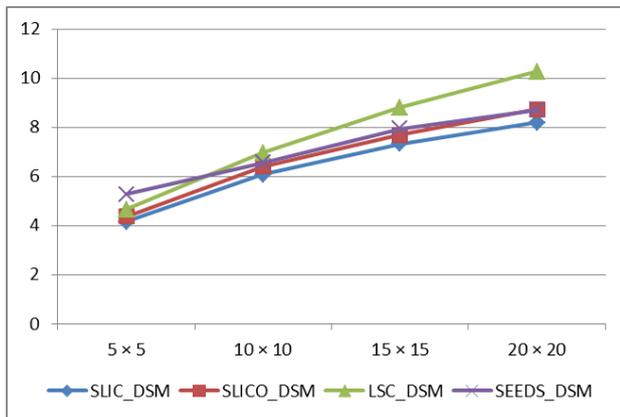


Figure 7. SD values (vertical axis) of the superpixels from DSM test area, starting the algorithms from initial size of the superpixels of 5×5, 10×10, 15×15 and 20×20, respectively.

4. DISCUSSIONS AND CONCLUSION

This study offers an initial overview of comparing four state-of-the-art superpixel methods: SLIC, SLICO, SEEDS and LSC. The fastest one was SEEDS, closely followed up by SLIC. The desired size of initial superpixels influences the speed of the computation. Therefore, the size of the superpixels should be carefully chosen: a smaller size is increasing the runtime, while a larger size is mixing up more than one class inside a superpixel. We suggest that an initial size of 10×10 pixels for the superpixels is a good compromise between the speed and accuracy of the final superpixels.

In this study, we have used 10 iterations for clustering and refinement of superpixels, while this value was found to be sufficient (Achanta et al., 2012). However, a higher number of iterations can lead to a better adherence of the superpixels to the boundaries of natural features in the scene but affecting the runtime negatively.

Regarding internal homogeneity of resulted superpixels, SLIC overcomes the other algorithms. This is mainly because SLIC generates superpixels that don't have a compactness constraint. Therefore, SLIC superpixels better follow even the most irregular shapes in the image.

Overall, after comparing runtime and homogeneity of the resulted superpixels for all three test areas, we can come to the conclusion that SLIC superpixels are the best choice when taking into account the two aspects: speed and accuracy. We suggest that superpixels should be preferred against the rigid structure of pixels, having in mind all the advantages described in this study. Further studies are needed in order to have a more detailed comparison amongst many other existing superpixel algorithms and their possible usage in OBIA applications.

ACKNOWLEDGEMENTS (OPTIONAL)

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).

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.
- Chen, J., Dowman, I., Li, S., Li, Z., Madden, M., Mills, J., Paparoditis, N., Rottensteiner, F., Sester, M., Toth, C., Trinder, J., 2016. Information from imagery: ISPRS scientific vision and research agenda. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, pp. 3-21.
- Chen, Y.X., Qin, K., Liu, Y., Gan, S.Z., Zhan, Y., 2012. Feature modelling of high resolution remote sensing images considering spatial autocorrelation. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, pp. 467-472.
- Fulkerson, B., Vedaldi, A., Soatto, S., 2009. Class segmentation and object localization with superpixel neighborhoods, In: *ICCV*, Vol. 9, pp. 670-677.
- 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.
- Li, Z., Chen, J., 2015. Superpixel segmentation using linear spectral clustering, In: *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pp 1356-1363.
- Neubert, P., Protzel, P., 2012. Superpixel benchmark and comparison, *Proc. Forum Bildverarbeitung*, pp. 1-12.
- 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.
- Shi, C., Wang, L., 2014. Incorporating spatial information in spectral unmixing: A review. *Remote Sensing of Environment*, 149, pp. 70-87.
- Tiede, D., Lang, S., Füreder, P., Hölbling, D., Hoffmann, C., Zeil, P., 2011. Automated damage indication for rapid geospatial reporting. *Photogrammetric Engineering & Remote Sensing*, 77, pp. 933-942.
- 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.

Revised July 2016