

ADVANCED EXTRACTION OF SPATIAL INFORMATION FROM HIGH RESOLUTION SATELLITE DATA

T. Pour^{a,*}, J. Burian^a, J. Miřijovský^a

^a Palacký University in Olomouc, Faculty of Science, Department of Geoinformatics, Olomouc, Czech Republic - (tomas.pour02, jaroslav.burian, jakub.mirijovsky)@upol.cz

ICWG III/VII

KEY WORDS: Image Analysis, Seath, Obia, Satellite Imagery, Urban Areas

ABSTRACT:

In this paper authors processed five satellite image of five different Middle-European cities taken by five different sensors. The aim of the paper was to find methods and approaches leading to evaluation and spatial data extraction from areas of interest. For this reason, data were firstly pre-processed using image fusion, mosaicking and segmentation processes. Results going into the next step were two polygon layers; first one representing single objects and the second one representing city blocks. In the second step, polygon layers were classified and exported into Esri shapefile format. Classification was partly hierarchical expert based and partly based on the tool SEaTH used for separability distinction and thresholding. Final results along with visual previews were attached to the original thesis. Results are evaluated visually and statistically in the last part of the paper. In the discussion author described difficulties of working with data of large size, taken by different sensors and different also thematically.

1. INTRODUCTION

Analysis of the urban structure is very important task that helps us understand the fast growing environment of large cities. This phenomenon was described multiple times using either GIS technologies (Burian, 2014) or remote sensing methods (Du, 2014; Blaschke et al., 2014, Aleksandrowicz et al., 2014). Remote sensing analysis remains a challenge even in the 21st century. Most of the articles including new remote sensing methods are based on data on which the method works perfectly (Nussbaum, 2006). Sadly, when the method is used for different data it hardly results in a similar quality and needs revision (Gao, 2011). Two things that increase this error and adds difficulties to the data processing are different sensors and different conditions of data acquisition (Novack, 2011). In this work the challenge consisted in processing different data from different sensors acquired at different date and time of the day using single set of methods. The main objective of this paper was to determine if versatile mathematical analysis of the remote sensing data can provide comparable results for different input data.

2. DATA AND SOFTWARE

2.1 Input data

Dataset including five European cities were used as the input dataset. These data were bought and are owned by Palacký University in Olomouc, Department of Geoinformatics. All of those cities, one per country, are middle-sized settlements and are located in post-communist countries. The cities are Ostrava (Czech Republic), Katowice (Poland), Košice (Slovakia), Szekesfehervar (Hungary) and Leipzig (Germany). The data were collected by three commercial satellites including WorldView2, GeoEye-1 and QuickBird. All sensors provide comparable data with sub-meter resolution in panchromatic

band and about 2 m resolution in multispectral bands. Images were taken in the summer of 2011 within 3 months from 27th June to 27th September. Time of the day was between 9:46 and 10:11. Size of the dataset was about 10 GB before any processing. The image of Katowice and Ostrava were cut into two pieces. The reason was the limited size of TIFF format and in the case of Katowice, it was also merge of two adjacent images. The image of Katowice was cut in half horizontally which resulted in difficulties later on. Another issue was with the image of Leipzig where we couldn't buy image of the whole city and at the image of Katowice there was low level of cloudiness. Because of these reasons, the image of Katowice was the worst of this dataset. In general, we can say that this dataset has high quality because of the time difference and resolution.

2.2 Software used

In cooperation with Technical University in Ostrava, we used older version of Definiens Developer (7.0, 2007). This software was used for segmentation, hierarchical classification and export to shapefile. For pre-processing and mosaicking the software ERDAS Imagine 2015 has been used. ESRI ArcGIS was used for visualisation and final corrections. Project R and R Studio were used for programming function for calculating SEaTH method for different classes. For minor pre-processing tasks, sharpening and conversion to NITF the software Exelis ENVI was used.

3. PRE-PROCESSING

3.1 Mosaicking

The first part of pre-processing was mosaicking images of Katowice and Ostrava and their conversion into NITF format. Both parts were geo-referenced so the only difficulty was to

* Corresponding author.

create seam line. Software ERDAS Imagine offers automatic seam line creation but the created seam line doesn't seem to respect any borders despite the fact that it has been set up to do so. Custom seam line was created in MosaicPRO Editor so that it will not cut objects into parts.



Figure 1. Part of the seamline in the city of Katowice

3.2 Image fusion

Image fusion was the second part of the pre-processing. The method, also known as pan-sharpening or resolution merge, is well known and adjusted. The original idea was to use subtraction as a general method for all images. This method has been successfully used before on QuickBird images but it doesn't work that good for other sensor. It actually changes the tone of the image completely making it very difficult to interpret. The method that mostly solved this issue was Gram-Schmidt pan-sharpening in Exelis ENVI. The method showed superior results on some images and very decent on others. Results had to be exported as NITF format which allows file size more than 4 GB which TIFF does not allow. The downside of NITF is that it removes geographic location of the image. Since the analysis was focused on comparison, this wasn't an issue.

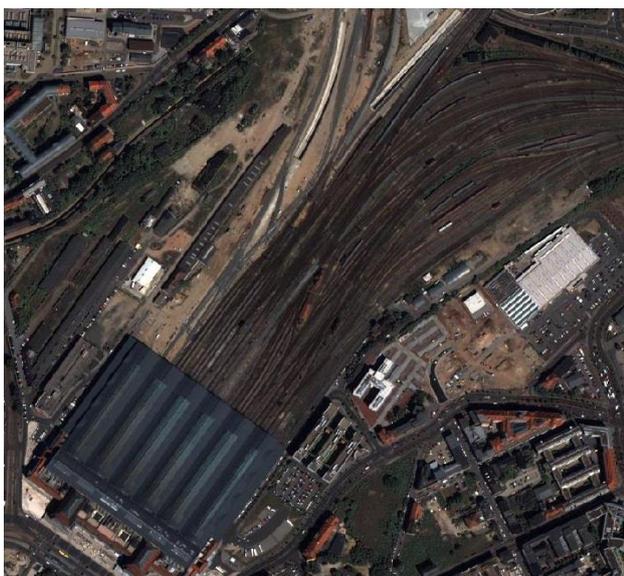


Figure 2. Results of Gram-Schmidt Pan-sharpening, the city of Leipzig

3.3 Segmentation

The third and the last part of the pre-processing was segmentation. The original idea was to use two levels of segmentation. One of small scale created by the algorithm and the other imported from governmental data – city blocks. In the end due to bad support of older version of Definiens Developer and issues getting the data, both levels were generated by the software, one of small scale and the other of larger scale. Segmentation is one of the most important parts of the object based classification. Since there are different algorithms and most of them, like the one in Definiens Developer, are secret. The setup was scale 200 and other parameters default for small scale and scale 1000, compactness 0.8 and shape 0.8 for large level.

4. CLASSIFICATION

4.1 Hierarchical Classification

Hierarchical classification was chosen to be the method for information extraction. Basic classes easily recognisable in multispectral images were chosen and formed hierarchical tree. In the first step we distinguished vegetation and non-vegetation. Second step divided vegetation to high and low and non-vegetation to tar materials and other. The last step was to divide other into red roofs and other materials. At the non-vegetation level, dark objects were removed from the classification.

4.2 SEaTH

SeaTH method (Nussbaum, 2006) was coded in R Project to distinguish the feature and the threshold for decision junctions in classification. The method firstly measures separability for a pair of classes and then counts two thresholds for each pair – upper and lower ones. The image of Košice was used as the training area for choosing the feature with highest separability between two classes. There was a number of features tested such as brightness, means in various wavelength bands, density, texture, ndvi, elliptical and rectangular similarity or maximum difference within the object. Features used in the hierarchical classification were ndvi, brightness, standard deviation and mean in blue band.

4.3 Large scale level

Small scale level of the classification was calculated and based on it also the second larger level. The classes on the larger level was assigned based on relatively dominant classes from smaller level. Four classes were created. These were dominant vegetation, dominant red roofs, dominant tar areas and mixed. In the case of city of Košice, which showed the best results, two residential areas were distinguished along with industrial zones and green areas in suburbs.

5. INTERPRETATION

5.1 Small scale level

Statistically the differences between land cover areas were very large. The city of Košice has less than 25 percent of tar land cover while other cities have between 40 and 50 percent. Detection of red roofs turn out to be the most difficult with same method in different images. Szekesfehervar has less than 1 percent of red roofs which is error, but other cities range from 5 to 10 percent, which reflects the actual structure of the city. Vegetation is the largest in the city of Košice and is more than

50 percent of the image. Ostrava and Leipzig are larger cities with man developed and built-up areas. This also reflects in the amount of vegetation which is the lowest among those cities. The remaining two cities – Katowice and Szekesfehervar – have vegetation coverage about 40 percent.



Figure 3. Small scale level classification, the city of Košice, light and dark green represent low and high vegetation respectively, grey represents tar areas, red areas are red roofs, black is for dark objects and yellow represents other surfaces

5.2 Large scale level

On the larger scale level, we can distinguish spatial distribution and concentration of different areas. Košice have the clearest city structure as mentioned before. Leipzig is mainly built-up settlement with green areas in form of parks but lacking other inner-urban vegetation. Similar situation is in Ostrava where, on the other hand, we can see vegetation lanes between industrial zones. Szekesfehervar has lots of vegetation spread in across the city but has only a little concentrated vegetation. In Katowice we can identify old city centre with residential red roofs. Besides this, the inner urban structure is scattered.

6. CONCLUSIONS

This work has proven that it is possible to use universal methods to a dataset consisting of images from different sensors. Cities of Ostrava and Leipzig were identified as industrial cities with low relative vegetation and heavy industry and tar areas. Katowice and Szekesfehervar show a lot mixed areas and spread different types of land cover. Košice on the other hand shows clear structure from the centre of the city changing older red rooftops into mixed into industrial into vegetation areas.

This paper showed good results in terms of analysing city structure and its comparison. Despite these results we would advise a revision of the whole process when applied on different data set and also improving the process technically. There were issues including data management, the secrecy of algorithms and even differences between results from different sensors.

ACKNOWLEDGEMENT

The paper was supported by Internal Grant Agency of Palacký University, Olomouc (project IGA_PrF_2016_008 - Advanced monitoring, spatial analysis and visualization of urban landscape).

REFERENCES

- Burian, J., 2014. Possibilities of The Definition of City Boundaries in Gis: The Case Study of a Medium-Sized City. In: *14th SGEM GeoConference on Informatics, Geoinformatics and Remote Sensing*. <http://sgem.org/sgemlib/spip.php?article4297>
- 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. and Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. In: *ISPRS Journal of Photogrammetry and Remote Sensing* Vol. 87, p. 180-191.
- Novack, T., Esch, T., Kux, H. and Stilla, U., 2011. Machine Learning Comparison between WorldView-2 and QuickBird-2-Simulated Imagery Regarding Object-Based Urban Land Cover Classification. In: *Remote Sensing* Vol. 3, no. 12, p. 2263-2282.
- Nussbaum, S., Niemeyer, I. And Canty, M.J., 2006, Seath - A New Tool for Automated Feature Extraction in The Context of Object-Based Image Analysis. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol. XXXVI-4/C42.
- Aleksandrowicz, S., Turlej, K., Lewiński, S. and Bochenek, Z., 2014, Change Detection Algorithm for the Production of Land Cover Change Maps over the European Union Countries. In: *Remote Sensing*. 2014. Vol. 6, no. 7, p. 5976-5994
- Gao, Y., Mas, J., Niemeyer, I., Marpu, P. and Palacio, J., 2007. Object-Based Image Analysis for Mapping Land-Cover in a Forest Area. *International Symposium for Spatial Data Quality* Enschede, Netherlands
- Du, P., Liu, P., Xia, J., Feng, L., Liu, S., Tan, K. and Cheng, L., 2014. Remote Sensing Image Interpretation for Urban Environment Analysis: Methods, System and Examples. In: *Remote Sensing*. 2014. Vol. 6, no. 10, p. 9458-9474