Use of Automatic Image Classification for Analysis of the Landscape, Case Study of Staré Jesenčany, the Czech Republic

Eva Trojovská, Pavel Sedlák, Jitka Komárková, Ivana Čermáková

University of Pardubice, Faculty of Economics and Administration, Czech Republic
eva.trojovska@gmail.com, {pavel.sedlak, jitka.komarkova}@upce.cz, ivana.cermakova1@student.upce.cz

Abstract. The landscape is changing, mainly due to man. The changes in the landscape are increasingly important, so it is important to monitor these changes, mainly due to the use of resources in the country in the future. The article deals with the use of automatic classification of data from remote sensing to analyse the development of the landscape on the case study in the village Staré Jesenčany, the Czech Republic.

Keywords: Image processing · Automatic classification · Comparison of classification · Landscape changes

1 Introduction

Studying changes of landscape is a very important issue in the modern society. It means that the society is interested in new findings, which can help to understand how nature and its particular elements exist and behave. Understanding functions, relationships and rules can support landscape management and further sustainable development e.g. prevention of devastating impacts of floods or drought. For example, studies [1, 2] are focused on landscape changes.

Remote sensing is very often used as a source of data for observation of landscape and terrain. There are many issues why this method is more reasonable than in situ observations, sampling and measurements and land surveying. The costs are lower (namely in a case when bigger areas are monitored), accuracy and spatial resolution is adequate to aims of studies and finally, data measured in various parts of electromagnetic spectrum are available. The last advantage is important for researches based on thermal imagery, various indices, etc. Satellites or aerial imagery is available for monitoring larger areas. Unmanned aerial vehicles (UAV) are increasingly used to monitor small areas. UAVs can often provide results faster and usually with higher spatial resolution.

Remotely acquired image material is either analogue or digital, and therefore the form determines also the possibility of further processing. Each image contains information about topographic (geometric) properties of objects (e.g. size, the distance of

adfa, p. 1, 2011.

objects and their relative position) and thematic information (e.g. type of surface soil moisture content). Processing of the digital image information, which is stored in digital form, together with the development of computer technology presents a series of processes to automate and accelerate some, particularly computational methods (transformation, creation of orthophoto, etc.). Other significant advantages of objectivity, accuracy and in a sense even lower processing costs. Besides these advantages the digital processing of remote sensing materials brings some disadvantages, and particularly in the area of automatic classification must also use knowledge from analogue interpretation. [3] Structure and sequence of methods of digital image processing are not completely uniform and closed.

2 Previous researches

In the most general sense, the classification is a process in that each pixel is assigned a corresponding information. Its aim is to replace the value of the radiometric characteristics of the original image by values called informational class. The result of classification is then presented for example in the form of thematic maps. [3]

The landscape as a term was emerged in the early nineties of the 20th century as one of the key words this time. The landscape is heterogeneous portion of the earth's surface, consisting of a set of interacting ecosystems, which is repeated in similar forms in that part of the surface. [4, 5]

Monitoring changes in the landscape over time is generally based on monitoring changes in individual landscape components, their surface representation, dynamics and spatial configurations. To change the landscape type occurs when a different type of landscape elements becomes a landscape matrix, which either landscape component grows or fades. [6] Quantification of the structure of land cover always encouraged environmentalists to creating the most diverse indices or their acceptance of other disciplines. [7, 8]

At present, the theme of monitoring changes in the landscape and image classification goal of many projects. These themes are attractive for students in the processing of theses and become a commercial product of software companies in the field of geoinformatics also.

Detection of landscape changes is a very important part of landscape management. Remote sensing is used to collect up-to-date data in various parts of electromagnetic spectrum. Data are processed by many different methods.

Qindong and Shengyan [2] analysed landscape pattern changes in monitored period (1990-2013) and identified driving sources during that period. They used various landscape indices to calculate landscape changes.

Bortoleto et al. [1] focused on index of restoration in landscapes. They proposed a mathematical index named SIR that describes suitability of individual habitat patches for restoration within a landscape. A model based on the SIR used a map of distance classes among fragments and a map of habitat quality established according to each land cover category. SIR was obtained as a result of calculation.

At present, there are many works focusing on the automatic classification of images with high spatial resolution for the monitoring of change of landscape. They primarily use aerial photographs or images from the UAV. Automatic classification of images with high spatial resolution are solved, for example, in works [10, 11, 12, 13, 14]. However, the use of images from the archives of satellite images is often cheaper, a lot of works use this source [15, 16, 17].

3 Case Study: Classification of image and quantification of areas

Methods of digital image processing with focus on automatic classification and monitoring changes in the landscape are applied to the analysis of the municipality Staré Jesenčany. The village is located near to city Pardubice in the Czech Republic. It is a small cadastral area of approximately 3.71 km^2. Typical for this landscape is a higher percentage of arable land, grass surface and construction of houses. There were analysed aerial photographs from firm Geodis, which were captured between 2003 and 2008. The classification was performed on the Landsat 7 satellite images from 2000 and 2010. The work was processed in software ArcGIS for Desktop from firm Esri.

A sequence of individual steps corresponds with the steps that are caught in a process map (Fig. 1), which was also created by the authors of this article [18].

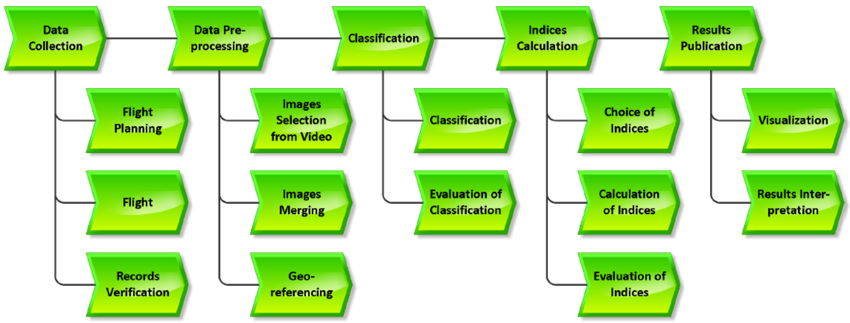

Fig. 1. The sequence of steps of image processing [18]

Algorithm ISODATA was applied on aerial photographs as the first method. At the beginning of this algorithm, number of clusters and the number of iterations is defined. Cluster, which becomes heterogeneous (based on the standard deviation values) is divided into two new clusters. Clusters, of that centers are closer than a predetermined value, are combined into one cluster. Clusters that contain fewer pixels than the pre-given number are cancelled and the pixels are assigned to the nearest clusters [3].

The results of this classification were particularly useful for obtaining pre-conditions of what we can expect from another classifier. For example, we can say that there is a great similarity between the areas of arable land and grass surface. The results were not used as input during the further processing.

The surface of the features creek or railway consists of a typical pixel structure and supervised classification will achieve better results here.

Maximum likelihood classification method was used as the basic method of aerial photographs. For each DN value, probability of its occurrence can be calculated. When it is presented it in the correlation field, each cluster with the same probability of a pixel can be connected to a certain value [3].

A training area for the following categories was created: building area, arable land, grass, communications, forest and pond. The big problem was to classify forested areas due to great similarities with the class grass, so they were only classified areas with shadows cast by trees. It is a major simplification, but it was assumed that there will be some compensation for the actual area of forest cover and the shadows of buildings. We can take this simplification into account if it is to monitor changes in the landscape percent per unit of the monitored area. The result of classification was full of isolated pixels and therefore postclassification adjustment was necessary. The difference between the original classification and adjusted image illustrates Fig. 2.

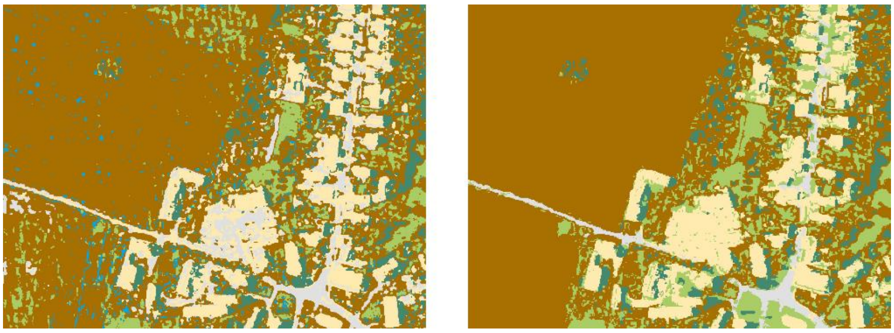

Fig. 2. The classification of image 2003 before and after adjustment (source: authors)

The classification results for each time horizons were quantified in hectares and percentage of different classes in the study area.

An objective way to determine the accuracy of classification is to build the error matrix in which they are compared randomly selected reference points and classified points. Values of both error matrix were achieved through Raster Calculator tool, and function Combine in ArcGIS. From those matrices, it was determined that the average accuracy for 2003 is 0.56 and for 2008 is 0.63.

Another way to determine the accuracy of the classification was to compare the results of maximum likelihood classifier with the results of visual interpretation, which were available through the work of [19].

We can say that the automatic classification of aerial photographs gave generally satisfactory results. Since the classification of satellite images was to be expected that due to the small size of the territory and low resolution images of Landsat 7 (resolution 30 m), the results are less conclusive. Application of ISODATA classifier to satellite images from 2000 and 2010 provided very rough classification results that lead into three and four classes. Either classification using maximum likelihood algorithm did not provide a new information, it only provided already acquired knowledge about the area. We can see the general structure of the territory from this classification, but the details of it could not be obtained. Frequent erroneous classification of the pixels due to great similarities between some classes can be seen like in the case of classification of aerial photographs, see Fig. 3.

The visual interpretation was used to obtain the most accurate results of the classification of satellite images also. Although in this process, there is rather more emphasis put on interpretive skills of man than on manner of calculation program; and it is necessary to take into account the data itself. As already mentioned, the images are of a low resolution. Elements of territory in the image from 2010 are more recognizable than in the picture from 2000, where for example it is not possible to identify area of the pond. Visual interpretation of these images from the Landsat 7 shows Fig. 4.

Fig. 3. Satellite images from 2000 and 2010, classified by maximum likelihood algorithm (source: authors)

For example, the very purpose of the study area and location of the pond, it is preferable to use a combination of bands of multispectral images in the order: red = 4, green = 5, blue = 3, which reveals the frontier land and water. The classification follows the displayed image by maximum likelihood classifier and visual interpretations. By means of Calculate Area, it was found that the area of a polygon obtained by visual interpretation is equal to 37 508 m^2, while the area resulting from supervised classification is 29 820 m^2, both in 2008. The area calculated by means of automatic classification is approximately 10 000 m2 smaller. The surface of the pond during the years 2000-2010 is constant, so there is no change in the case of the second image.

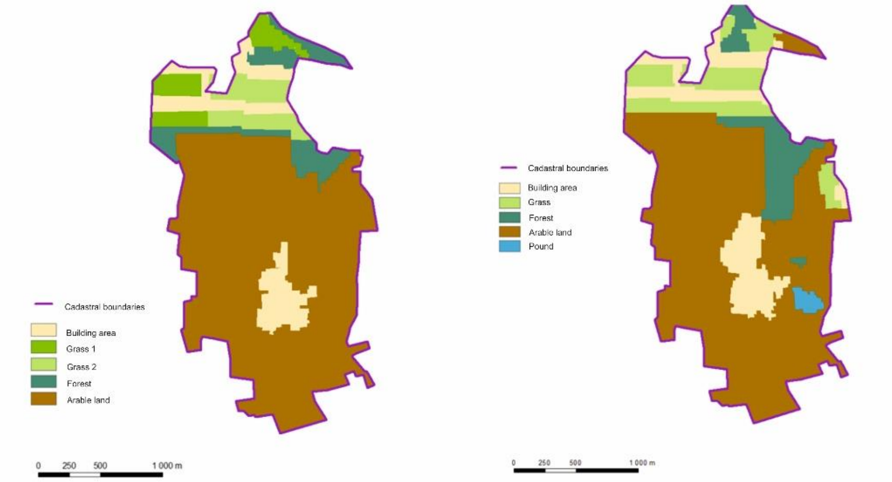

Fig. 4. Visual interpretation of satellite images from 2000 and 2010 (source: authors)

4 Interpretation of changes of Staré Jesenčany municipality

In the case of the classification of aerial photographs, the difference between these two time horizons can be neatly summed up – see Fig. 5 and the following Table 1, where are listed quantitative differences from the visual interpretation of study area also.

Table 1. Quantitative capture of changes in each category between years 2008 and 2003, based on aerial photographs (source: modified according [20])

Information class	**Changes in individual categories [ha]**		**Proportional representation [%]**	
	Automatic classification	Visual interpretation	Automatic classification	Visual interpretation
Building area	0.54	1.24	0.15	0.33
Arable land	-30.73	-5.21	-8.24	-1.40
Grass	32.87	3.39	8.90	0.91
Forest	-3.67	0.55	-0.99	0.15
Pond	0.65	0.20	0.17	0.06
Communication	0.04	0.02	0.01	0.01
Others	-	0.66	-	-0.05

Building area has increased, especially in the northwest part of village. Visual interpretation, however, reveals this increase more than twice higher than the automatic classification. The decline is typical for the area of arable land. In the case of forest monitoring for the result reflected a big difference between the visual interpretation and automatic classification of 2008. Although in 2008 were felled many trees and shrubs around the pond, on the whole territory, there is identifies a small increase in forest category (see visual interpretation results). The automatic classification gives bad results for this category. The calculation of result of the forest area in 2003 from automatic classification on the contrary is very close to reality. Category pond is a feature with an almost constant area. Growth of areas of communication category is small.

The graphic visualisation is a very good way to present the results of changes in the landscape, based on an analysis of satellite images – see Fig.4. The horizontal axis represents particular years 2000, 2003, 2008, 2010 and 2012. The vertical axis represents areas in hectares. In a line graph, there are showed the basic categories of building area, arable land and grass, with values of the results from visual interpretation, therefore, the most accurate values. The values of year 2012 were obtained from Czech Office for Surveying, Mapping and Cadastre [21].

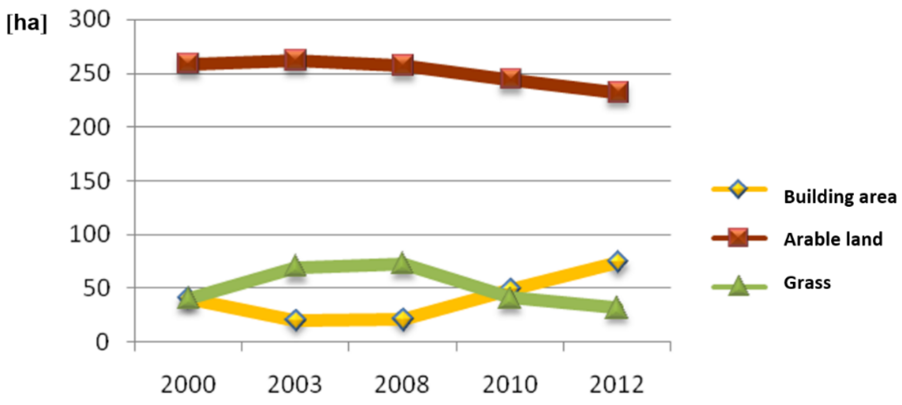

Fig. 4. The changes of basic categories in landscape (source: authors)

5 Conclusion

The automatic classification of aerial and satellite images were used to determine changes in the landscape in the municipality Staré Jesenčany. Software ArcGIS for Desktop 10 was used for data processing. The following data were used: aerial images from the company Geodis from years 2003 and 2008, satellite images from Landsat 7 ETM +, years 2000 and 2010. The supervised classification (Maximum Likelihood classifier) and unsupervised classification (method ISODATA) were used. The classification of satellite images was easier, but it did not provide good results due to the small size of the area and the low spatial resolution Landsat 7 images. Spectral image enhancement (composition change in RGB and vegetation indices) belongs to the significant benefits of multispectral images.

The classification results show that the proportion of arable land decreased due to new buildings. Also, there was a slight increase in area of forest and scrubland. The small increase can be observed in the category of communications as a result of new developments.

As far as spatial data are very important for the smart cities and regions sustainable development, suitable automatic classification, which is described in the paper, can be used as a part of smart technologies to support the development.

Acknowledgements. The paper has been completed with the kind support of SGS_2017_17 project of Faculty of economics and Administration, University of Pardubice.

References

1. Bortoleto L. A., 2016. Suitability index for restoration in landscapes: An alternative proposal for restoration projects. Ecological Indicators 1, pp. 724-735 (2016)
2. Qindong, F. and Shengyan, D., 2016. Landscape pattern changes at a county scale: A case study in Fengqiu, Henan Province, China from 1990 to 2013. Catena 2(2016), pp. 152-160

3. Lillesand, T. M., Kiefer, R. W.; Chipman, Jonathan W. Remote Sensing and Image Interpretation. 6th ed. Hoboken: Wiley, 756 ps. (2008)
4. Forman, R. T. T. Krajinná ekologie. 1. vyd. Praha: Academia, 583 ps. (1993)
5. Hradecký, J., Buzek, L. Nauka o krajině: Učební texty Ostravské univerzity. Vyd. 1. Ostrava: Ostravská univerzita, 2001, 215 ps. https://katedry.osu.cz/kfg/user/hradecky/download/skripta.rar
6. Lipský, Z. Sledování změn v kulturní krajině: učební text pro cvičení z předmětu Krajinná ekologie. Kostelec nad Černými lesy: Lesnická práce, 71 ps. (2000)
7. Guth, J., Kučera, T. Monitorování z měn krajinného pokryvu s využitím DPZ a GIS. Příroda [online]. 1997(10), 107-124 [cit. 2012-01-30]. https://www.usbe.cas.cz/people/kucera/LE/TEXTY/landcov.pdf
8. Herzog, F. et al. Landscape metrics for assessment of Landscape destruction and rehabilitation. Springer-Verlag: Environmental Management, 2001, Vol. 27., No. 1.
9. Pechanec, V., Brus, J., et al.: Decision Support Tool for the Evaluation of Landscapes. In: Ecological Informatics, vol. 30, pp. 305-308 (2015)
10. Gonzalez, A., Rafael, G., et al.: Robus Segmentation of Aerial Image Data Recorded for Landscape Ecology Studies. In: Huang, F., Sugimoto, A. (eds.), LNCS, vol. 9555, pp. 61-72 (2016)
11. Movia, A., Benait, A., Crosilla F.: Shadow Detection and Removal in RGB VHR Images for Land Use Unsupervised Classification. In: ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING, vol. 119, pp. 485-495 (2016)
12. Zhang, L., Li, A., et al.: Global and Local Saliency Analysis for the Extraction of Residential Areas in High-Spatial-Resolution Remote Sensing Image. In: IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, vol. 58, pp. 3750-3763 (2016)
13. Caccetta, P., Collins, S., et al.: Monitoring Land Surface and Cover in Urban and Peri-urban Environments Using Digital Aerial Photography. In: INTERNATIONAL JOURNAL OF DIGITAL EARTH, vol. 9 (2016)
14. Hung, C., Xu, Z., Sukkarieh, S.: Feature Learning Based Approach for Weed Classification Using High Resolution Aerial Images from a Digital Camera Mounted on a UAV. In: REMOTE SENSING, vol. 6, pp. 12037-12054 (2014)
15. Landsat Data Access. USGS. Landsat Missions [online]. 2015. In: http://landsat.usgs.gov/Landsat_Search_and_Download.php Access Date: 20 September, 2015
16. Brunclík, T., Danquah, K. A. B., "Model of chlorophyll-a concentrations in inland water bodies based on Landsat data," Advances in Remote Sensing, Finite Differences and Information Security, pp. 215-218 (2012)
17. Pásler, M., Komárková, J., Sedlák, P.. Comparison of Possibilities of UAV and Landsat in Observation of Small Inland Water Bodies. In International Conference on Information Society: Proceedings. Londýn: Infonomic Society, pp. 45-49 (2015)
18. Čermáková, I., Komárková, J. Modelling a Process of UAV Data Collection and Processing. In International Conference on Information Society: proceedings. Londýn: Infonomics Society, pp. 161-164 (2016)
19. Hromádko, T. Využití leteckých snímků pro analýzu vývoje krajiny v okolí obce Staré Jesenčany. Pardubice. Bachelor thesis. University of Pardubice (2009)
20. Trojovská, E. Využití automatické klasifikace obrazu pro analýzu vývoje krajiny v obci Staré Jesenčany. Diploma thesis. University of Pardubice (2012)
21. Czech Office for Surveying, Mapping and Cadastre (Český úřad zeměměřičský a katastrální) K. ú.: 754412 - Staré Jesenčany - podrobné informace [online]. 2012 [cit. 2012-07-30].

http://www.cuzk.cz/Dokument.aspx?PRARESKOD=10&MENUID=10016&AKCE=META:SESTAVA:MDR002_XSLT:WEBCUZK_ID:754412

After revision

1. The paper was formatted according to the guidelines.
2. More literature sources were added.
3. English was corrected.