ANALYSIS OF THE CONNECTION BETWEEN URBAN LAND COVER AND CENSUS DISTRICTS USING GEOINFORMATICAL METHODS

HENITS LÁSZLÓ[1] – MUCSI LÁSZLÓ[2]

[1,2]University of Szeged, Department of Physical Geography and Geoinformatics, Hungary 6722 Szeged, Egyetem str. 2-6. Email: [1]henits@geo.u-szeged.hu, [2]mucsi@geo.u-szeged.hu

Received 3 October 2012; accepted in revised form 15 December 2012

Abstract
Remote sensing resources are usually used in research to better understand urban built-up density, spatial structure and the processes of change. Based on results of image segmentation, landscape metrics indexes, texture and pattern may be analyzed beside the spatial changes in urban reflectance. Social processes within the settlement can be analyzed efficiently, although the census data may also be connected to the urban land cover data through geoinformation systems. On the research project different parameters of urban segments, i.e. patch number, mean patch area, total patch area, total patch perimeter, patch density and edge density, formations that make up the urban pattern were analyzed. Urban functional districts of different built-up density were separated using appropriate indexes, and extending the database with spectral content made it possible to review district boundaries and to mark new boundaries due to these changes.

Keywords: urban land cover, remote sensing, census, image segmentation, landscape metrics

1. Introduction

An adequate data source of information on urban landscape is the database of the census conducted every 10 years (Kátainé et al., 2012). The spatial data are suitable for analysis in geoinformatical systems but for researchers of geography the primarily available data are the so-called combined area data published by KSH (Hungarian Central Statistical Office) rather than the raw data regarding the smallest area units, the enumeration districts (KSH, 2003). The establishment of enumeration districts is an important element in the preparation of a census, thus, ensuring that the whole area of the country is covered and classifying the addresses known by KSH into districts. The basis for creating districts is the census address database; the addresses within it are from the address register of KSH, which is usually collected during the previous census.

So-called functional units are determined within the enumeration districts with regard to built-up density, to land use and to the settlement's historical quarters. The census returns, which mainly consist of demographic data, are determined for these units. Besides, data that can be well used in urban ecology studies, e.g. regarding housing supply, also appear.

Based on the census database, social processes within the settlement can be analyzed efficiently, although the data may also be connected to the urban land cover data through geoinformatical systems. The basic sources of urban land cover data are medium and high resolution satellite and aerial images (Weng, 2012).

The goals of our study were: (1) to analyze the connection between land cover and statistical functional units and (2) to investigate by using statistical methods if functional units can be characterized by landscape metrics parameters derived from urban land cover pattern image segmentation.

Through the spatial and time series analysis of the connection between land cover and functional units, important spatial information can be obtained to assess the quality of urban ecosystem services. Urban environmental conflicts (environmental load, allocation of green areas, etc.) can also be analyzed through the changes in built-up density and through demographic processes.

When analyzing land use, districts need to be viewed as homogeneous in terms of land use and built-up density (entity). When creating an urban land use map (inner functional units), the first step is to identify the nomenclature categories and to classify the enumeration districts in these categories. Then a larger merged district is created from adjacent districts that are included in the same category. This is usually done visually, by using large-scale maps and/or high resolution satellite, aerial images, and historical maps. Thus from a geoinformatical point of view, the smallest spatial unit, i.e. the object is the patch-like enumeration district bordered by polygons. However, in public reports, data are not given for enumeration districts, but rather for inner functional units merged from these districts.

When analyzing long term urban development, the change of inner functional unit areas and land use, together with the homogeneity of these units regarding the specific function are important aspects. The homogeneity of functional units may also be examined from a different point of view, e.g. land cover, texture.

Remote sensing resources are usually used in research to better understand urban built-up density, spatial structure and the processes of change. Based on these data, used together with landscape metrics indexes, texture and pattern may be analyzed beside the spatial changes in urban reflectance. The use of remote sensing resources for modelling urban processes has been known for a long time, but the use of remote sensing resources together with landscape metrics indexes may bring the understanding of urban area growth and changes to a higher level (Herold et al., 2005).

Urban remote sensing studies require intra-urban discrimination of urban land cover and land use types. According to several studies (Small, 2003), spatial sensor resolution of at least 5 m is necessary to accurately acquire the land cover object, but medium resolution satellite images may also be used if the pattern allows it.

The result of pixel or sub-pixel classification, rather than urban land use information, is naturally suitable for the analysis of urban land cover. The VIS model was the first attempt at modelling the connection between land cover and land use (bibliography), but no better or usable model has been created since then. Mainly visual methods remain for solving this problem. In these methods, after the use of modern devices and having completed training, the researcher makes the decisions through recognizing land pattern and cover (Urban Atlas).

2. Location of the area, data and methods

The investigated area was city of Szeged, and its environment. Szeged is the third most populated city in Hungary (cca. 170 thousand habitants), which is located on the southern part of Hungary, not so far from the Hungarian-Serbian border.

The basis of urban land cover mapping was a multispectral RapidEye satellite image. The recording was done on March 24 and 25, 2011, of Szeged and its surroundings, of an approximately 4000 km^2 area (Fig. 1). The constellation consisting of 5 satellites is able to make recordings of the same area daily, producing 5-band (blue, green, red, red edge, near IR), 5 meter spatial resolution images. Due to the cloudy and humid weather characteristic of the beginning of 2011, an atmospheric correction was completed on the images using the ATCOR2 add-on in ERDAS IMAGINE 2011. During calibration, DN values were converted to reflectance values, and haze reduction was also successfully completed on the images.

The database was further extended with a vector file (shp file) containing boundaries of functional districts (43 districts) which were established based on the 2001 census data (KSH, 2003). This vector file was the basis of district statistics in this study (Fig. 2).

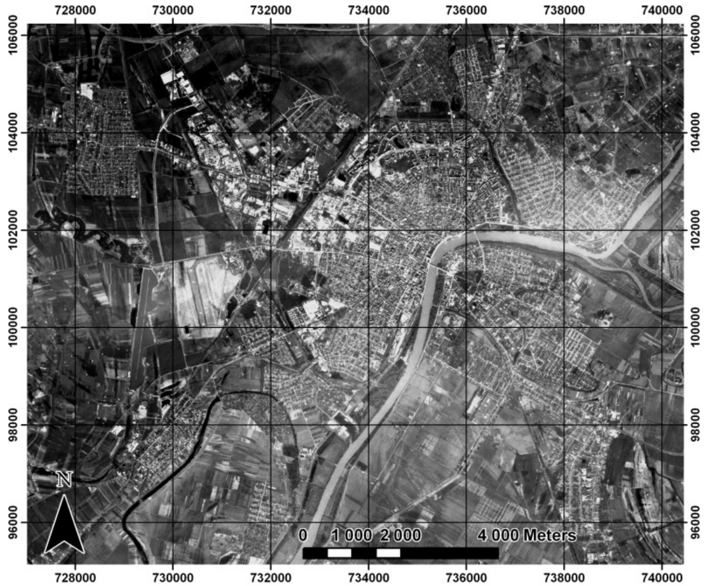


Fig. 1. RapidEye satellite image in the range of visible light (RGB 321), recorded March 24, 2011.

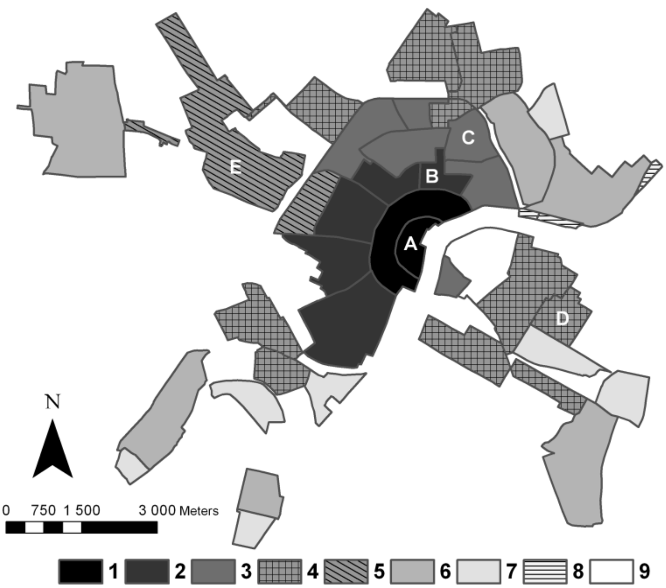


Fig. 2 Inner functional division of Szeged (KSH, 2003)
1 - Downtown, 2 - Inner residential area, 3 - Housing estates, 4 - Suburban residential area,
5 - Industrial area, 6 - Rural residential area, 7 - Weekend house area, 8 - Inner resort area,
9 - Garden suburb A, B, C, D, E are five selected districts, which were analyzed in detail

An important preliminary step in the remote sensing analysis of urban land cover is the spatial analysis of land cover reflectance, the determination of spatial scale of urban reflectance (Small, 2003). The segmentation method was used to determine the spatial scale of reflectance in each district. In the segmentation module of ERDAS IMAGINE 2011, homogeneous land cover patches and pixel groups representing those were selected for the whole area of the city, using a region growing algorithm. The first step of segmentation is edge detection, which can be carried out by setting a threshold. Based on the boundaries set by the edges, the process creates the segments with the help of minimal value difference. The result of the segmentation is a thematic layer where the pixel values denote the identifiers of individual objects. The raster layer containing the segments was vectorized without smoothing, so the RapidEye 5 x 5 m pixel size was retained in case of vector objects, too. The resulting vector file was intersected with the KSH file, which contained the districts. Then all the segments smaller than 1 pixel were removed. The main landscape metrics indexes (patch area and perimeter, shape index) were determined for the 50871 polygons using the ArcGIS V-LATE (Vector-based Landscape Analysis Tool Extension). Then zone statistics were completed for each district (number of patches, mean area of patches, total area of patches, total perimeter of patches, patch density, edge density) on classification level. On the basis of these data, mean patch sizes were obtained for KSH-districts as classes. Patch density can be calculated as the ratio of the number of patches and the total area of the class (district). Therefore this ratio expresses the number of patches in unit area (patches/ha). Edge density can be expressed in a similar way, it is the measure of edge length in unit area (m/ha) (Szabó, 2009). Maximum likelihood classification was used to map five land cover classes (trees, buildings, roads, grassland and shadow). Following the classification, the ratio of impervious surfaces was calculated from the land cover map.

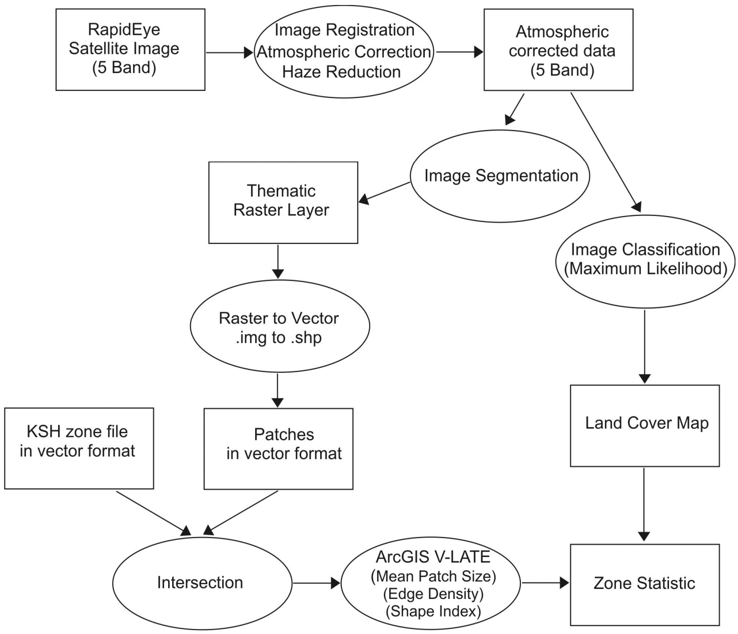


Fig. 3. Flowchart about segmentation and calculating the parameters of landscape metrics

3. Results

3.1. Patch size and edge density

One of the main parameters of urban land cover is the size of patches determined by spectrally similar image elements within the functional district. Roof surfaces of buildings, impervious land cover surfaces (pavement), open soil surfaces, water surfaces and plant-covered surfaces make up patches whose sizes can be determined by segmentation. On this list, impervious surfaces remain unchanged for a long time, but the size of reflecting natural surfaces may change even within a year. In an urban landscape, patch size indicates the spatial differences in land use and technological, architectural features of building up previously open surfaces changing with time. By spatial statistical analysis it could be determined that the mean patch size was the smallest (740-830 m^2) in the downtown area (Fig. 4a) and the traditional inner residential area.

Due to the dense, close and block-like build-up (Fig. 4b), homogeneous patches of relatively small area, consisting of impervious materials can be found in these districts. Patches of larger size consist of pixel groups that indicate boulevards, avenues and parks. Inside the boulevards the grid-like avenue structure does not allow for creating a corridor of plant-covered patches, which would be important from an ecological standpoint.

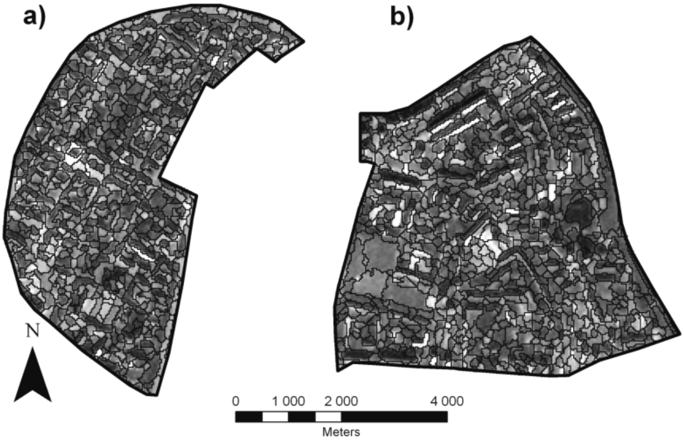


Fig. 4. A (a) city centre (downtown) and a (b) housing estate cut with segment boundaries

It was also observed that moving outwards from the city centre, the mean patch area increases. From the 1960s the building of housing estates was restricted to the area between the former city limits and the round dam (Körtöltés), which was built after the great flood of 1879. As opposed to the 2- or 3-story buildings of the city centre, in this part of the city there are 4-, 10-story houses of brick and panel blocks, with larger ground-space than the buildings in the centre. Between the blocks of houses it was possible to create larger green areas, and since the end of the 1990s most of these areas have been occupied by shopping malls with large parking lots. In the northern housing estate zones the mean average patch size is 900-1000 m^2 due to the blocks of houses and the open green spaces between them. (Fig. 5a) In the industrial area in the north-western part of the city the mean segment area size is 950-1100 m^2 because of the warehouses, factory buildings and open unbuilt surfaces. In the suburban residential area the mean size of pixel groups is 800-1000 m^2, in this zone there are family houses of smaller surface with larger yards and open green spaces.

During the reconstruction following the great flood, standard plans and building plots of 600-800 m^2 were offered to the former proprietors (Lechner, 1891). These plots had 400-500 m^2 backyards behind the building built on the front line. Thus inside the blocks, the collective green area of the adjoining backyards make up 800-1000 m^2 patches. The patch boundaries are the back walls of the buildings, and the fences, walls and hedges border the plots on the sides.

Based on edge density (Fig. 5b), which indicates landscape fragmentation in landscape ecology, the centre of the city displays high values: 1800-2100 m/ha. This is naturally related to the fact that smaller patches have relatively longer edges. These areas are fragmented in this degree, owning to the streets between the buildings, the avenues and the boulevards in the district. In the case of housing estates and industrial establishments this index is reduced, since the more compact and larger open green areas decrease its value. In the suburban, weekend house areas this value is higher due to the many small streets, the adjacent buildings and the small gardens.

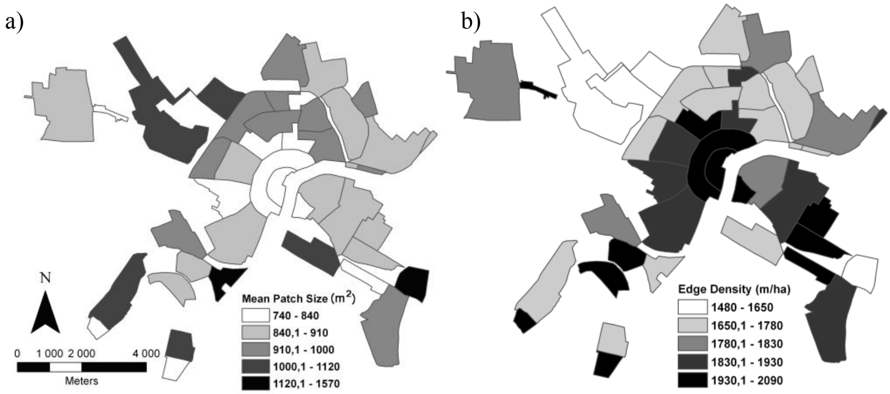


Fig. 5. Average values of mean patch size and edge density for KSH 2003 districts

The box-plot diagram illustrating the segment areas for each district serves with additional information (Fig. 6). It can be concluded that the elements falling in the same zone have boxes that are of similar size and contain the medium 50 % of the data. In the city centre segments that fall between the first and the third percent constitute 575 m^2, in the inner residential area this value is 525-575 m^2, but in the housing estate zone it ranges from 600 to 675 m^2. The pixel groups of the suburban residential area display variety, they range from 550 to 650 m^2. The size of the box comprising the medium 50 % is the largest in the case of the industrial zone, which has a 650 and 725 m^2 band. It can be concluded that out of all functional units, there is a significant difference in the patch size of individual districts of resort areas and weekend house areas, which is a consequence of different construction regulation conditions.

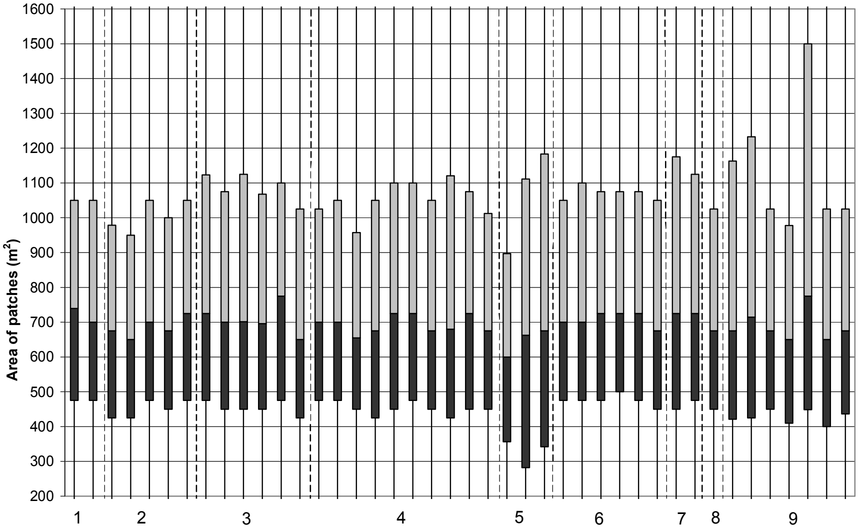
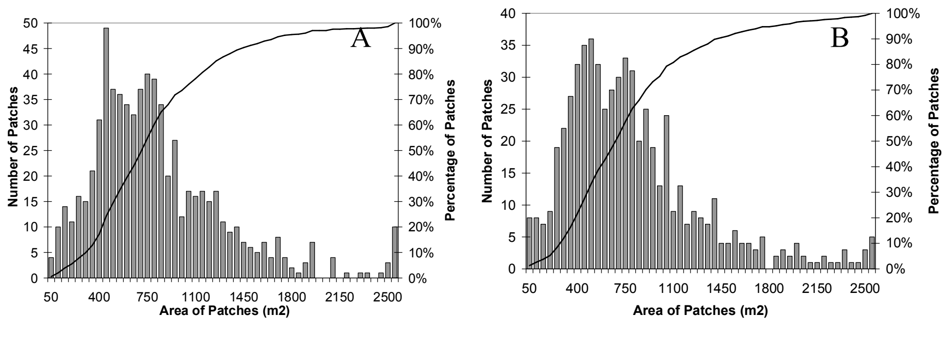


Fig. 6. Box-plots of KSH 2003 districts based on their patch sizes
1 - Downtown, 2 - Inner residential area, 3 - Housing estates, 4 - Suburban residential area, 5 - Resort area, 6 - Rural residential area, 7 - Industrial area, 8 - Garden suburb, 9 - Weekend house area

Next, one representative district was selected from the downtown area, the inner residential area, the housing estates, the suburban residential zone and the industrial zone each (Fig. 7), which was analyzed in detail. When preparing histograms, segments were sorted in 50 m^2 groups up to 2500 m^2, i.e. up to 100-pixel size, and one class that incorporates more than 100 pixels was also created.



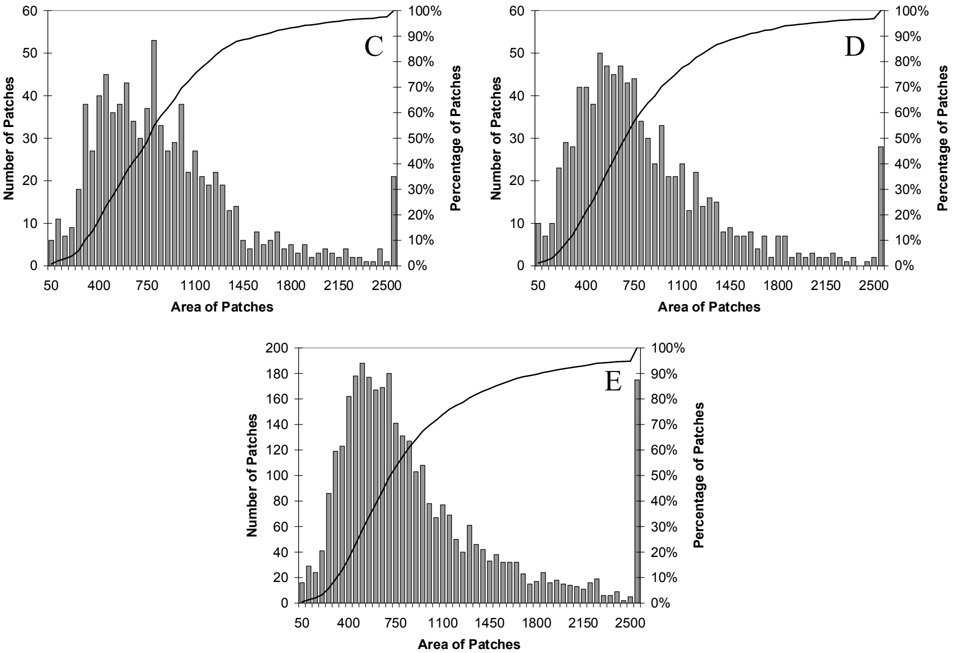


Fig. 7. Histograms of the following segment areas: A - Downtown, B - Inner residential area, C - Housing estates, D - Suburban residential area, E - Industrial area (A, B, C, D, E see Fig. 2)

Based on the histograms and the box-plots it can be determined that 50% of all segments fall in the following ranges: 475-1050 m^2 in the downtown area, 425-978 m^2 in the inner residential area, 475-1100 m^2 in housing estates, 450-1050 m^2 in the suburban residential area and 475-1125 m^2 in the industrial zone. It was also concluded that in the downtown area there are two peaks at 450 and 750 m^2; 36.9 % of the segments are between these two values. In the inner residential area there is no such peak, in this area the segments are distributed more evenly in similar value ranges. 49.6 % of the segments may be considered small, ranging from 350 to 800 m^2. The ratio of larger segments with areas of more than 1500 m^2 is 9.5 % in the inner residential area, and 9.6 % in the downtown area due to the green areas between the buildings and the image elements of avenues. As it has already been established, in the case of housing estates the 50 % frequency box is wide, which is also shown on the histograms: from 300 to 1100 m^2 in each class the frequency is more than 2.5 %. These are parts of grass-covered areas between buildings and smaller apartment blocks. Homogeneous pixel groups larger than 1500 m^2 are usually wide green surfaces, blocks and multi-story buildings or shaded areas. Their ratio is 11.4 %, which is higher compared to the two previously mentioned districts. The histogram of the suburban residential area is very similar to that of the inner residential area, most of the segments are small, the ratio of segments sized 200-800 m^2 is 56.1 %, which can be explained with houses and their yards or gardens that fragment the area. The larger patches are the yet open spaces, which constitute 11.4 % of this area. Segments larger than 1500 m^2 are mostly represented in the industrial zone; they take up 15.9 % of the whole area. The fact that this value is so high is mainly due to large industrial establishments, warehouses and large adjacent grass-covered surfaces and un-built areas.

3.2. Shape index

The shape index is calculated by dividing the patch perimeter by the minimum perimeter, which is expressed as the ratio of the perimeter of a compact patch to the same area. This index is not influenced by the size of the polygons, and it is considered the best shape marker (Szabó, 2009). It can be noted that by using this index, areas of different functionality can be separated well (Fig. 8). The downtown area is an individual unit, moreover, the inner residential area and the housing estates fall in the same category (1.58-1.6). Districts of the industrial zone have values of 1.6-1.624, while zones that are located outside the round dam (Körtöltés) and are not closely connected to the city's core area display high values (1.7-1.91).

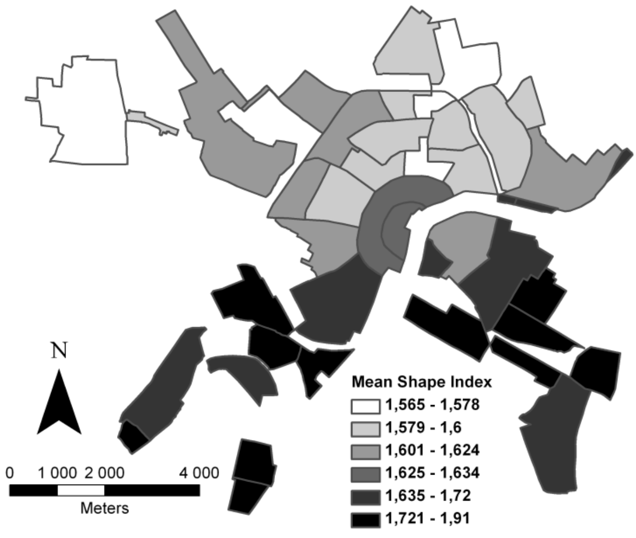


Fig. 8. Mean shape indexes for KSH 2003 districts

Districts of different functions are displayed on a plane spanned by mean patch size and mean Shape index. Fig. 9 illustrates downtown, inner residential and housing estate zones by these two parameters. The downtown districts have nearly equal mean patch sizes and mean shape indexes, since these areas are independent units regarding how much they are built up, and they are characterized by dense, close built-up and low mean patch size. In case of the inner residential area, a grouping may be determined. These districts have the same values as the downtown area, because the dense, block-like build-up continues even after the Nagykörút boulevard bordering the Downtown. The exception is district 23 (Alsóváros), the northern part of which is rather similar to the Downtown, while its southern part resembles the suburban residential area with wider spaces, yards and gardens connected to the houses. The housing estates are also definitely separated in this diagram. The closely positioned blocks in Tarján (3, 21) and Makkosháza (19) were built approximately at the same time, the green spaces and parks were already established between the blocks, while in the Rókus housing estate (10) the apartment blocks are on wide grass-covered and forest areas or are dissected by lakes. In addition, this last district borders the industrial area, so buildings of industrial function appear in its peripheral parts. In this category, there is also a particular value that stands out: Odessza, the first housing estate built in Szeged (29), which is densely built up with multi-story brick and panel blocks, allowing little space for parks and grass-covered areas.

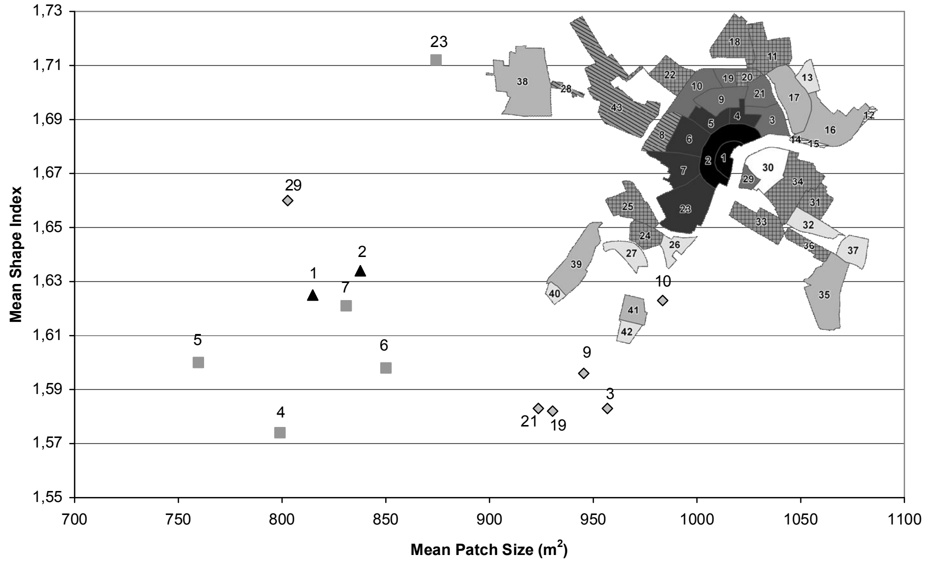


Fig. 9. Downtown, inner residential and housing estate districts in the plane spanned by mean patch size and mean shape index

Fig. 10 illustrates rural, suburban residential and weekend house areas**.** The points displayed here show higher dispersion, the groups are more heterogeneous based on these two parameters, so the groups are more difficult to determine. The rural residential zone constitutes two classes, the adjacent Petőfitelep (16) and Tápé (17), together with Kiskundorozsma (38) are in the first one – these are fragmented by parallel roads that are perpendicular to each other, creating patches of nearly 900 m^2. The other three districts located in the southern part of the city have larger open, yet unbuilt surfaces of 950-1050 m^2 mean area. Suburban residential zones display similar values as weekend house areas, but they do not form sharp-cut groups. Most of them are characterized with large mean patch size and shape index, but the dispersion of these values is high.

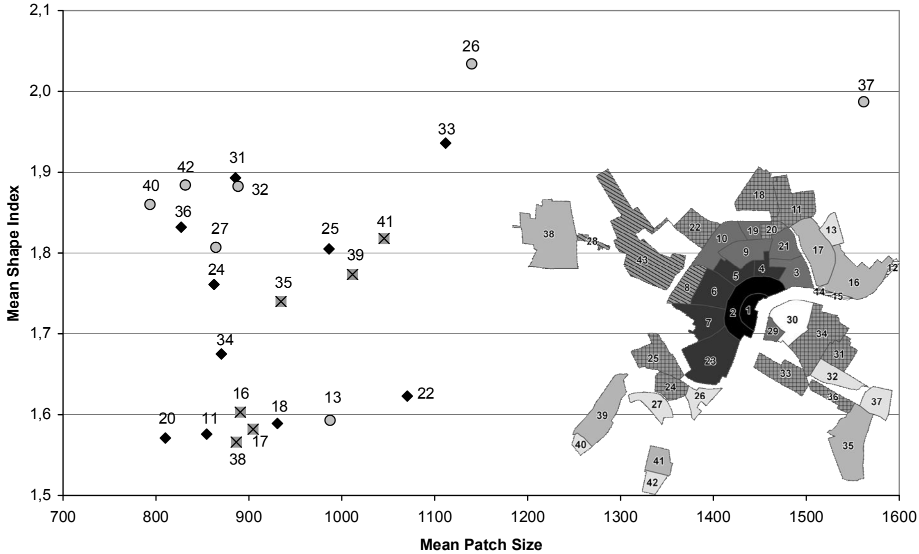


Fig. 10. Rural, suburban residential zones and weekend house areas in the plane spanned by the mean patch size and mean Shape index

3.3. Study of district boundaries

The results of the 2011 census are not known yet, but it would be important for the comparability of functional units that the demographic data refer to the same districts as the districts determined after the 2001 census. Although in this case the changes in function within a district cannot be displayed. It would be important to analyze if areas of different functions have been created within the functional units in the last 10 years, or if there are areas of identical functions within the 2001 district boundaries.

After studying 43 districts, it can be established that because district boundaries are often marked along boulevards, avenues and streets, there are some less homogeneous districts. That is why marked zones do not always adhere to the various land use categories. Such an example is a district that belongs to the inner residential area (Fig. 11), and is located between a downtown and an industrial zone. The part marked "B" is characterized by apartment houses, apartment blocks and the grass-covered areas connected to them, although a shopping mall and a bus station is also located here. The other part, marked "A", is more similar to the adjacent industrial zone. For this reason, the district was separated into two parts, and zones "A" and "B" were processed separately. The previously mentioned landscape metrics were recalculated for these two zones.

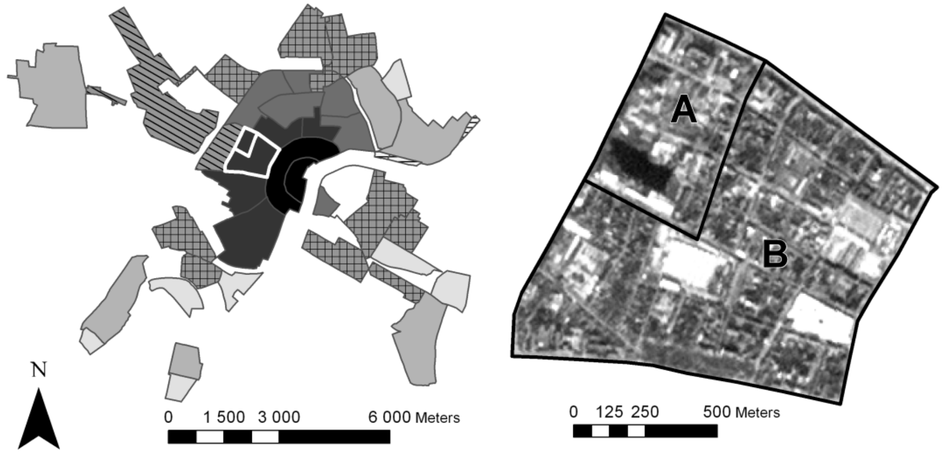


Fig. 11. The division of the inner residential zone between the downtown and industrial area separated into zone A - similar to the industrial zone, and B - similar to the inner residential zone

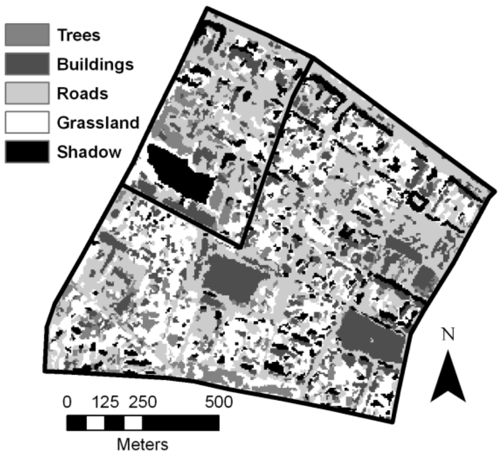


Fig. 12. The land cover image of the inner residential zone created by using supervised maximum likelihood classification

The mean segment area for the whole district is 850.2 m^2, but after separating the segments, this value was 889.1 m^2 for zone "A" and 816.9 m^2 for zone "B". This way zone "B" displays more similarity in mean patch size with the downtown area and the inner residential area, while "A" is more similar to the industrial zones.

The different landscape metrics parameters of the segments, such as mean area, are not necessarily enough to separate the districts; therefore spectral content may be introduced to the investigation as an additional parameter. A supervised maximum likelihood classification may be carried out within the selected district (Fig. 12), where 5 classes were defined (Trees, Grassland, Buildings, Roads and Shadow). Following the classification, the ratio of impervious surfaces in the whole district without Shadows is 60.9%, for zone "A" this ratio is 64.7%, for "B" it is 59.9%, where no big difference may be noted.

4. Conclusion

Using the different parameters of segments, i.e. patch number, mean patch area, total patch area, total patch perimeter, patch density and edge density, formations that make up the urban pattern can be analyzed. Urban districts of different built-up density may be separated using appropriate indexes, and extending the database with spectral content may make it possible to review district boundaries and to mark new boundaries due to these changes.

The RapidEye image of 5 m spatial resolution was usable for evaluation of urban land cover based on segmentation, because the spatial scale of urban reflectance is bigger in this city (10-60 m), then the spatial resolution of the satellite image.

According to our results the functional districts of KSH were clearly characterized by parameters of segments. Among the functional zones, especially the districts of downtown, the inner residential area and housing estate can be described by similar parameters (mean patch size and mean shape index).

The differences among the calculated parameters of distinguished part of the functional zones mark the long term temporal change in urban land cover, notably in the case of housing estate districts.

Acknowledgement

The Rapideye image was ordered within the frame of the Inland Waters Information System project (GOP–1.1.1–08/1-2008–0025).

References

ERDAS Field Guide (2010) - ERDAS, Inc. Norcross – GA 30092-2500 USA p. 812 http://geospatial.intergraph.com/products/erdasimagine/ERDASIMAGINE/ProductLiterature.aspx (downloaded in 2012)

Herold M. – Couclelis H. – Clarke K.C. (2005): The role of spatial metrics in the analysis and modeling of urban land use change. *Computers, Environment and Urban Systems* **29**: pp. 369–399

Kátainé Csincsák É. – Borbély A. – Erdei V. – Lakatos G. M. (2011) A 2011. évi magyarországi népszámlálás területi előkészítése, *Területi statisztika* **14.** (51.) Vol. 5. pp. 469-482.

KSH (2003) A nagyvárosok belső tagozódása, KSH Szeged 2003

Lechner L. (1891): *Szeged újjáépítése* - Szeged, 2000. (hasonmás kiadás)

Small C. (2003): High spatial resolution spectral mixture analysis of urban reflectance *Remote Sensing of Environment* **88** pp.170–186.

Szabó Sz. (2009) Tájmetriai mérőszámok alkalmazási lehetőségeinek vizsgálata a tájanalízisben 107 p. (Habilitation)

Weng Q. (2012): Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends *Remote Sensing of Environment* **117**, pp. 34–49.