THE AUTOMATIC CLASSIFICATION OF B&W AERIAL PHOTOS

L. Halounova

Remote Sensing Laboratory, Faculty of Civil Engineering, Czech Technical University Prague
Thakurova 7, 166 29 Prague 6, Czech Republic, halounov@fsv.cvut.cz

KEY WORDS: Land Cover, Photography, Texture, Fuzzy Logic, Segmentation, Aerial, Multiresolution

ABSTRACT:

The paper shows possibility to classify B&W aerial orthophotographs and other monochromatic remote sensing data using image enhancement phase and object-oriented analysis phase for the automatic classification.
The first phase enlarges the spectral signature space by channels calculated by image filtering (median filter and Gauss filter) and by texture measures. The combination of various filter sizes (texture measures) and kernel sizes (filters) enlarges the signature space allowing the following image segmentation and classification in two or three scale levels. The at least two level classification simplifies thematically complex aerial orthophotographs by dividing the photo into thematically more homogenous areas in the higher level. The lower level brings the final sought classes, which can be slightly corrected in the third (lowest) level. The eCognition software was used for the image segmentation and for the automatic classification. It is the first method of automatic classification of land cover of monochromatic remote sensing data bringing accuracy better than 80 per cent.

1. INTRODUCTION

The paper author was a responsible person for analyzing one part of a project of the Czech Ministry of Agriculture. The aim of the project was to find solutions for automatic information extraction from B&W aerial orthophotographs. The scale of these aerial photographs was 1: 23 000. Each othophoto was a result of mosaicking. The classical way of automatic classification is based on close spectral signatures or other signatures of pixels representing the same classes. This assumption is not valid in case of B&W photographs where different areas are formed by pixels with the same values and on the other side one class is formed by wide range digital values. The signature space for individual classes overlapped and could not have been used for their distinguishing in this state. No references were found to show solution for the similar monochromatic data type. Known image enhancement methods were applied to be obtained more separable class signature space for individual classes. The automatic pixel-by-pixel classification was excluded from the analysis and replaced by the object-oriented classification performed for segmented image data.
Segmentation has been used by several specialists who applied various interpretation keys (Borisov et al., 1987, Jagtap et al., 1994, Naesset, 1996, Žihlavník, Palaga, 1995). The segmentation used in this project was the Fractal Net Evolution Approach (FNEA) commercially introduced by Baatz and Schäpe (1999) incorporated in commercial software eCognition.
The proposed method can be applied for large number of with relatively high level of automation. That was the project goal. The result of the project should be applied for the whole country and therefore their image processing operability was necessary.

2. METHODOLOGY

There were two main tasks in the image processing. To enlarge signature space of individual classes to be separable was the first task. To perform the automatic classification formed the second task.

2.1 Signature space enlargement

Signature space enlargement was done by using two ways of new channel calculation. One way was the image filtering by low-pass filters where Median filter and Gauss filter were applied. They supressed local image unhomogeneities. They showed high correlation with original image data. That was the reason why they did not assure sufficient class separability. Channels calculated from two kernel sizes and repeatable filtering were used within the project. Channels filtered by Gauss filter were calculated for standard deviation equal to 2, 3, and 4.
Another tool was necessary for the successful solution. Haralick functions were the tool as functions characterizing textures. There are more Haralick functions used in image processing. Different numbers of them were used for different level of classification detail. Haralick functions were tested for several window sizes. Window sizes depended on individual class member sizes. Smaller window sizes were useful for small resulting class members sizes. The tests showed that there were no prevailing trends in a certain direction and that was the reason why all directions where relations between pixels are determined, were taken into account. All directions meant that differences between a pixel and a reference pixel were calculated for directions 0°, 45°, 90°, and 135°. These differences were used for GLCM (Grey Level Co-occurrence Matrix) and GLDV (Grey Level Difference Vector) calculations from Haralick functions. Mean, standard deviation and dissimilarity functions were chosen for the orthophoto classification defined in following expressions. Their window size were 5x5 pixels, 11x11 pixels, and 21 x 21 pixels.

$$Mean\_i = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} P_{i,j}.i$$

$$St\_deviation = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\left[P_{i,j}\left(i - Mean\_i\right)^2\right]^{1/2}$$

$$Dissimilarity = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} P_{i,j}.\left|i - j\right|$$

where $P_{i,j}$ is the normalized grey level co-occurrence matrix (GLCM) with n x n size, where original n = 256 for 8bit data is pre-scaled to n=16 such that

$$\sum_{i,j=0}^{n-1} P(i,j) = 1 \; .$$

2.2 Object-oriented analysis

Problem with wide range of digital values representing one thematic class and overlapping values for individual classes were partly solved by the previous step – new channel calculations. This step has not solved all problems with overlapping values of individual classes. That was image segmentation what helped to divide image data into large regions representing important parts of land use. This segmentation allowed separating of urban parts from agricultural parts and from forest areas.

The segmentation performed by object-oriented analysis defined in eCognition software simplified thematically complicated image data content. The object-oriented analysis comprises two parts. The first one prepares image data by creating segments from them and the second one allows their classification.

The segmentation is based on heterogeneity evaluation. The heterogeneity is characterized either by spectral heterogeneity, or spatial heterogeneity and their mutual combination. Higher influence of spectral heterogeneity is accomplished by lower spatial heterogeneity while their sum is equal to 1. The spatial heterogeneity compares either the compactness taking into account segment length and its number of pixels, or the smoothness expressing relation between segment length and the shortest segment length defined by a rectangle circumscribing the segment.

2.3 Class definition

Two-level class definition was the result of the presented methodology. The higher-level segmentation served to classification into basic regions:

- old forest,
- young forest
- agricultural area,
- urban area.

The lower-level classification comprised higher number of classes. Each of them belonged into the only higher-level class. The following list shows classes for the lower-level (more detailed classification). Class names are created from three parts. The first part part (F = forest, NF = nonforest, A= agricultural area and U urban area) defines the real situation of segments derived from the lower-level classification. The second part determines belonging to higher-level classification where OF means old forest, YF represents young forest, AA is an abbreviation for agricultural area and UA urban areas. The third part of class names indicates the lower-level class. The complete lower-level class names are:

- F_OF_coniferous forest,
- F_OF_deciduous forest,
- F_OF forest older than 7 years,
- NF_OF forest up to 7 years,
- F_YF coniferous forest,
- F_YF deciduous forest,
- F_YF forest older than 7 years,
- NF_YF forest up to 7 years
- U_YF road,
- F_AA tree,
- A_AA field,
- U_AA house
- U_AA road,
- F_UA tree,
- NF_UA green area,
- U_UA light house,
- U_UA dark house,
- U_UA house,
- U_UA road.

The first part of names shows possible regrouping of certain classes, which were originally classified into thematically wrong higher-level classes. This process of regrouping brings new improvements into the analysis. The forest class is formed by three classes from old forest, three classes from young forest and one class from agricultural areas and one class from urban areas. Class definitions and their list were adapted to real situation in the image data and can change from to region to region.

2.4 Segmentation

The segmentation called multiresolution segmentation allows segmentation in more levels. The higher-level segmentation for image data division into thematically simple regions was calculated for high scale value (250) in the first image-processing phase. The original orthophoto and the channel calculated by median filter with 5x5 kernel size formed the input image data for the segmentation processes in the higher-level segmentation.

The lower-level segmentation being the second image-processing phase was a repetition of the first one for lower scale value (35 – 50 according to image data). The segmentation process used the same channels.

The influence of spectral heterogeneity varied from higher value for the higher segmentation level to lower value for more detailed classification where spatial heterogeneity played more important role.

2.5 Classification

After the segmentation, segments can be classified into classes. The segmented image data classification was done by the nearest neighbor classifier. The classification signature space used mean segment values of orthophotograph, channels calculated from median filter and Gauss filter with three different standard deviation values (equal to 2, 3, and 4). The last part of signature space was created by channels calculated as texture measures. Texture measures – mean, dissimilarity and standard deviation using three Haralick functions were calculated for three window sizes – 5x5 pixels, 11x11 pixels and 21x21 pixels. Classifications in both levels used the same signature space.

3. RESULTS

The higher-level classification had to be corrected in several cases (several segments). It is a relatively quick part of processing being performed during visual control and being done manually. Fig. 1 shows the original image data and Fig. 2 shows result of the higher-level segmentation. The second-level classification result is on Fig. 3.
The accuracy of classification result was controlled in random sample areas. The accuracy was calculated for producer´s accuracy PA(class_i) defined by

$$PA(class\_i) = \frac{a_{ii}}{\sum_{k=1}^{N} a_{ki}}$$

and for user´s accuracy UA(class_i)

$$UA(class\_i) = \frac{a_{ii}}{\sum_{k=1}^{N} a_{ik}}.$$

The producer´s accuracy estimates the probability that a pixel, which is of class_i of the reference classification, is correctly classified. The total number of pixels of *class_i* in the reference classification is obtained as the sum of column *i*. The user´s accuracy estimates the probability that a pixel classified as *class_i* is actually of *class_i*. It compares the correctly classified number of pixels of *class_i* with the total number of pixels classified as *class_i* . The total number of pixels classified as *class_i* is in the row *i*.

Classified classes	Reference classes							
	house	tree	road	field	forest up to 7y	deciduous	coniferous	Σ of pixels
house	8354	0	680	0	0	0	0	9034
tree	358	1609	199	288	0	0	0	2454
road	1831	0	4681	0	0	0	0	6512
field	792	275	264	297253	0	0	0	298584
forest up to 7y	0	0	0	0	49813	0	0	49813
deciduous	29	855	0	0	32792	24173	580	58429
coniferous	0	415	0	0	9	964	81086	82474
Σ of pixels	11364	3154	5824	297541	82614	25137	81666	507300
	Accuracies for individual classes							
Producer´s PA	0.74	0.51	0.80	1.00	0.60	0.96	0.99	
User´s UA	0.92	0.66	0.72	1.00	1.00	0.41	0.98	

Table 1. Results of classification accuracy. The best results are in yellow and the worst in gray

The overall accuracy was 0.9 with kappa coefficient equal to 0.87.

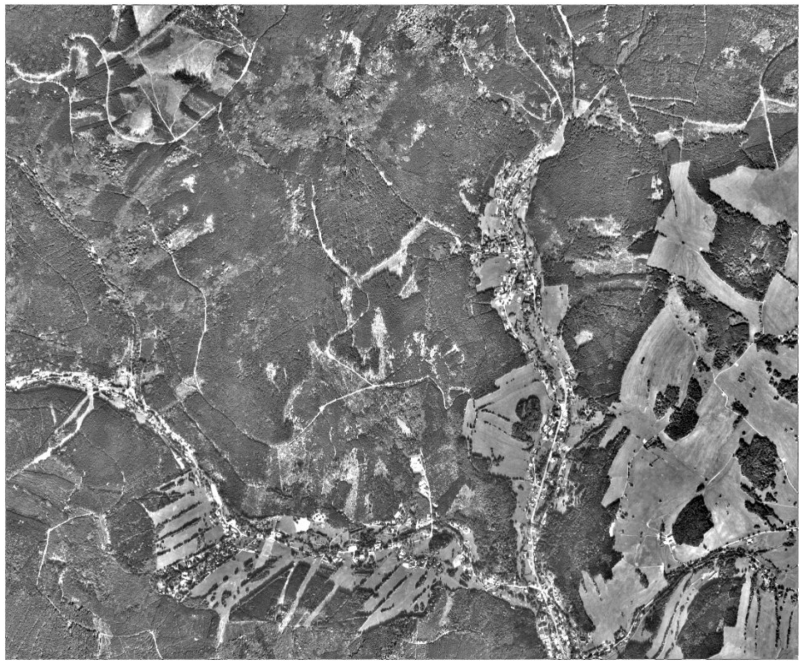

Figure 1. The image data (scanned B&W aerial orthophoto)

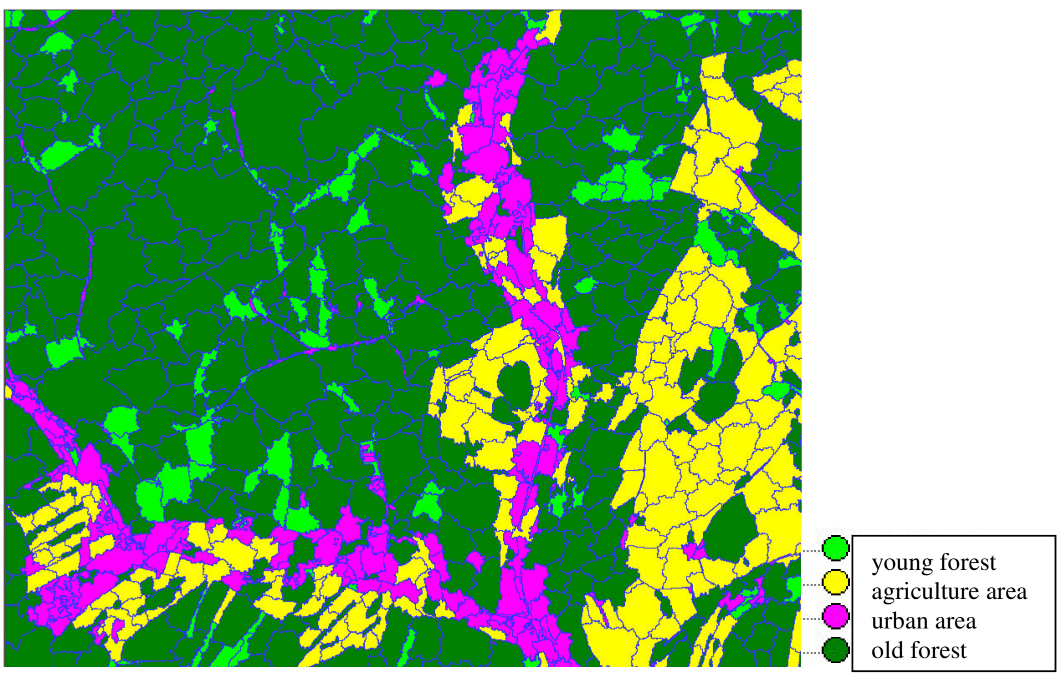

Figure 2. Higher-level classification

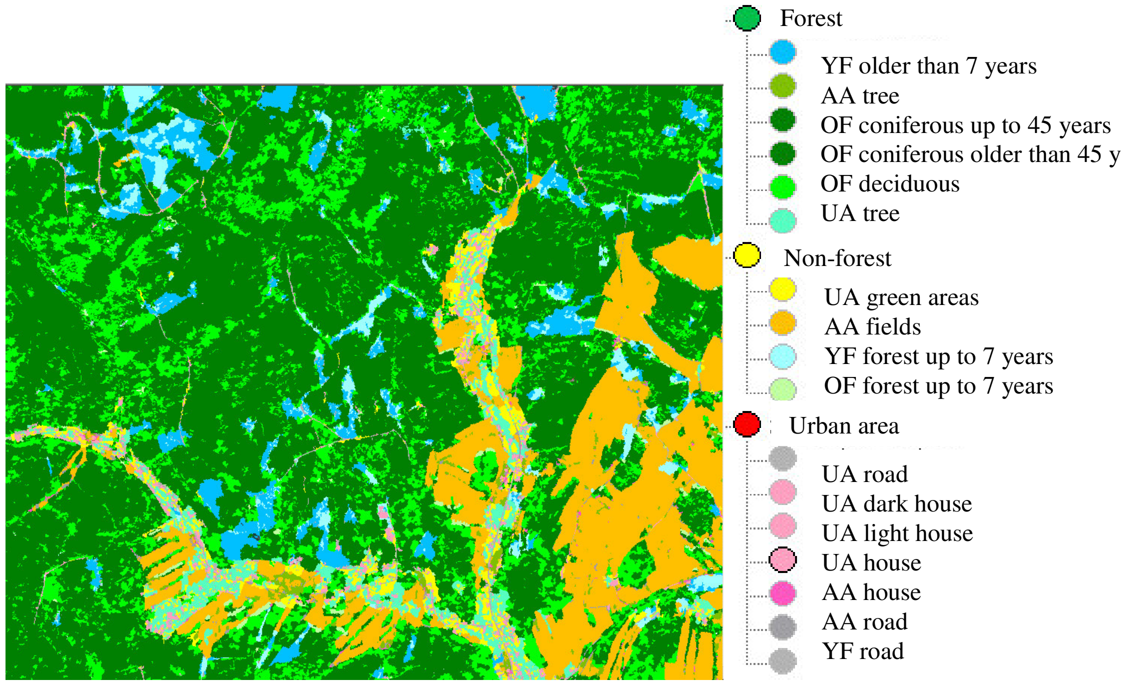

Figure 3. Lower-level classification

Results shows that certain parts of roads were classified into class houses. Pixels from class houses were classified into trees, roads, and fields. Classified class called tree comprised pixels from houses, roads, and fields, class roads was a class having also pixels from houses. Forest classes showed very good results for the coniferous forest. The forest younger than 7 years old was classified partly into the deciduous forest. The best results were obtained for fields. Results reliability is higher between urban areas and forests or urban areas and fields. Smaller reliability was found between urban classes – houses and roads, etc.

4. CONCLUSION

The B&W photograph classification can be performed with a relatively high accuracy. There are three necessary conditions offering such good results. New channels have to be calculated from original data (photograph) – channels calculated by filtering where median and Gauss filters are used, and channels using Haralick functions. Careful testing of kernel sizes for filtering and careful choice of filter window sizes for Haralick functions compared to final classified object sizes can sufficiently improve classification accuracy. The presented project used kernel sizes smaller than the smallest classified objects. Filter window sizes were smaller, equal and a little bit larger than these objects. The higher number of Haralick functions improves classification accuracy.
Object-oriented analysis using image segmentation followed by segment classification is the second necessary condition for good results of classifications.
Two level segmentation in the reverse order of segmentation (from higher to lower) is the third condition. The higher-level segmentation ensures fragmentation of image data with overlapping pixel values for different classes into thematically closer and smaller image parts whose further segmentation and classification offers good results. The lower level classification is performed individually for already fragmented parts of image data.
The confusion in resulting classification can be found in urban areas. The overall accuracy was about 90 per cent. However, the accuracy for individual classes varied from 50 per cent to 100 per cent. Low accuracy values were in case of two classes – trees whose pixels were classified into deciduous and coniferous forests. The class tree did not distinguish tree types. It was a class describing individual trees or tree groups out of forest areas. Rather wrong distinguishing can be found between very young forest and deciduous forest. It is difficult to distinguish these two classes even during visual interpretation especially for forest age younger than 40 years.

REFERENCES

Baatz, M. and Schäpe, A., 1999. *Object –oriented and multi-scale image analysis in semantic networks*. Proc. Of the 2nd International symposium on operationalization of remote sensing, August 16-20, , Enschede ITC.

Borisov, A.N., Kashin, V.B., Khlebopros, R.G., 1989. *Method for indication of horizontal structure of tree stands*. Doklady – Biological Sciences, 1989. 293: (1 – 6), 132 – 133.

Halounová, L., 2003. *Textural classification of B&W aerial photos for the forest classification*. Proc. of the 23rd symp. of EARSeL, Gent, Belgium, June 2-5, 2003: 173-179.

Halounová, L., 2004. Classification of B&W aerial photos and of radar data. Habilitation thesis. Czech Technical University Prague (*in Czech*)

Halounová, L. 2003. *Classification of B&W aerial photos*. The 4th Conference Actual Problems of Photogrammetry and RS, Prague, September 17-18.(*in Czech*)

Haralick, R.M., Shanmugan K. and Dinstein, I., 1973. *Textural Features for Image Classification*, IEEE Tr. on Systems, Man, and Cybernetics, Vol. SMC-3, No. 6, pp. 610-621, November 1973.

Haralick, R.M., 1979. *Statistical and Structural Approaches to Texture*, Proceedings of the IEEE, Vol. 67, No. 5, pp. 786-804, May 1979.

Jagtap, T.G., Untawale, A.G., Inamdar, S.N., 1994. *Study of mangrove environment of Maharashtra coast using remote sensing data*. Indian Journal of Marine Sciences. 23(2): 90-93.

Naesset, E., 1996. *Determination of number of stems in coniferous forest stands by means of aerial photo-interpretation*. Scandinavian Journal of Forest Research, 11(1): 76-84,

Žihlavník, S., Palaga, J., 1995. *Interpretation key-topographic elements and tree species on colour syntheses*, Lesnictví-Forestry, 41(10): 476-482.

ACKNOWLEDGEMENT

The processed data were used by the courtesy of the Ministry of Agriculture and the research part was processed within the GA ČR project 205/03/0218 research and Vyzkumny zamer 210000007.
The project presentation was financially supported by the Czech Technical University grant 84134-11153.