Annales Univ. Sci. Budapest., Sect. Comp. 23 (2004) 163-178

A METHOD FOR CLASSIFYING SATELLITE IMAGES USING SEGMENTS

I. László, T. Pröhle, I. Fekete and G. Csornai

(Budapest, Hungary)

Dedicated to Prof. Imre Kátai on his 65th birthday

Abstract. The thematic evaluation of remotely sensed images has continuously growing importance. Nowadays, very high resolution and hyperspectral images show the most spectacular development. FÖMI Remote Sensing Centre has an experience in the usage of satellite images of various kinds in different agricultural applications. There is a need for important summary statistics and survey maps where the usage of high spatial and spectral resolution images is not reasonable, because coarser resolution images contain enough information. In this case the large cost and work requirement of the finest images can be avoided.

In this study we deal with optical images with 20-30 m surface resolution. This kind of images still holds its leading position in the county- and country-scale land survey. As one of its fundamental applications, FÖMI RSC regularly reports crop area and yield predictions to the Ministry of Agriculture and Rural Development. To support it, a classification system has been working operationally for years, with a continuously developed method. Currently, it uses a basically traditional per-point approach. An improvement is described here, which uses a different paradigm: a segment-based, per-field classification.

After introducing basic concepts of remote sensing and the necessary preprocessing steps used in our practice, we outline the present classification method. Then, the principle and algorithm of segmentation is presented. The implemented procedure was elaborated using the methods described in the literature. The improvement was integrated into the operationally working system in a sophisticated way.

Advantages of the segment-based new method are the following: on

one hand, the applied cell formation filters the noise, on the other hand, the annexation algorithm assures the incorporation of spatial, geometric information, resulting in significant accuracy increase.

1. Introduction

As technology develops, *remote sensing (RS)* is continuously gaining importance: the quality of RS images is improving while unit costs are decreasing. Thus a need for handling large, sometimes enormous data sets is arising, which is mainly supported by high-speed computers. This paper concentrates on *agricultural applications* of remotely sensed images. In this section a brief summary of the related parts of remote sensing and image processing is given. For a detailed introduction to remote sensing and its applications, see e.g. Richards [1].

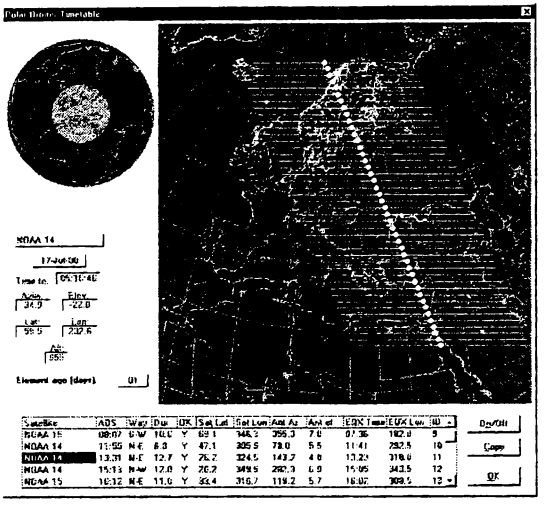


Figure 1. A pass of the satellite NOAA-14

Remotely sensed images convey information about a certain area of the Earth's surface. Satellites (e.g. Landsat, NOAA, SPOT, IRS series) surveying our environment detect the reflected radiation in various electromagnetic wavelength intervals. These data are stored in the form of *digital image*. Figure

1 shows the area covered by a pass of the satellite NOAA-14 as seen on the screen of an acquisition program.

A definite part of the whole electromagnetic spectrum, called optical band, is involved in RS. The radiation of the surface is measured by sensors, each capturing a given subinterval of the spectrum. A multispectral image can be regarded as a matrix, whose elements correspond to a given spot of the surface. These elements themselves, called *pixels*, are r-dimensional vectors, containing the intensity values recorded by the r sensors.

RS images are categorized by four principal parameters. *Pixel size* is the smallest distinguishable area on the surface. It varies between 1m by 1m and 1km by 1km. Most images in our practice (taken by Landsat TM and ETM+) have a pixel size of 30m by 30m. The second type of parameters are number of the *spectral bands* and the respective wavebands. The spectral bands are selected to requirements of applications. *Radiometric resolution* refers to the number of distinguishable intensity levels, yielding pixel values of 0..255. *Period of acquisition* (e.g. 16 days) is the time between the passes of a given satellite over the same area.

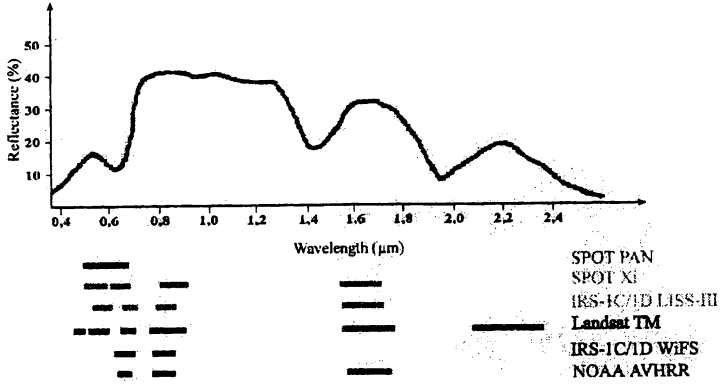


Figure 2. A typical reflectance function of vegetation and spectral bands of satellite sensors

Before starting the analysis of images, some *preprocessing steps* are necessary. Geometric and radiometric correction transform raw images into a uniform (spatial and spectral) system, which conforms to the way the Earth is usually pictured by man. The map projection used holds the area, but does not hold the distances. Nevertheless, the length distortion is not significant compared to the pixel size and the extent of the sample area. Effects disturbing detection and data transfer can be decreased by atmospheric correction and noise filtering.

The reflectance (i.e. the percentage of reflected radiation) of land covers heavily depends on the wavelength. The so-called *reflectance function* for a given land cover gives the reflectance along the optical band. Every land cover category has its own characteristic reflectance function, which depends on the phenological phase of the vegetation. This fact can be used to make distinction between land covers, which is a fundamental requirement of RS. Figure 2 shows the reflectance function of a typical cultivated crop. A representative sample of this function is taken by the sensors. Horizontal lines in the figure show a few examples of the spectral bands covered by some satellite sensors.

Unfortunately, there are time periods when spectral response of different crops nearly coincide. Therefore, a series of images from different dates have to be used in order to separate covers. Besides, they greatly help progress monitoring.

Straight application of determining land covers is *vegetation mapping*, which serves as a basis for area estimation. Similar technical and theoretical tools are used at flood and waterlog monitoring, or drought monitoring in another period of the year. The sophisticated task of early yield forecast can be solved with extensive use of RS image time series. For some representative applications in Hungary, see Csornai et al. [2].

The importance of RS is obvious: less field work is required, it gives immediate results and the accuracy and reliability is higher than those of human information resources.

2. The present classification method

The main purpose of classification is to produce a thematic map, the so-called *land cover map* of a certain area. Basically, the land cover map separates vegetation from surface elements of other kind, and more importantly, it makes a distinction between plant species. Unlike satellite images, which contain radiation intensities, pixel values of the thematic map refer to the categories of vegetation. The input for classification consists of satellite images and the spatial database of previously known surface areas (reference data). The sophisticated procedure is heavily based on clustering, and it intensively uses the human expert's knowledge.

As the reflectance of different plants often coincide in a given period, images from multiple dates are necessary to create an adequate classification. To start with, several images of the target area, having different acquisition dates, are "stacked together" to compose a multi-layer image.

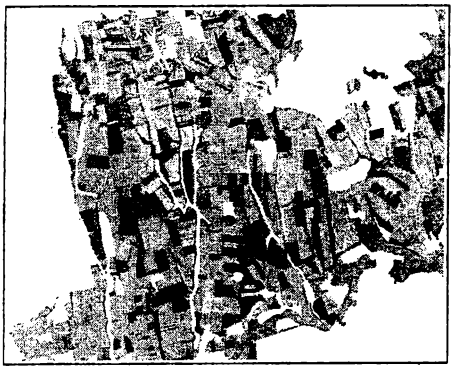
In the following, we are going to use the western part of Tolna County, Hungary, to demonstrate the methods. Figure 3 shows a composite image of this area taken in the late spring of 2000. The whole image used in classification is composed of 3 images (mid-spring, late spring, early summer), each having $r = 4$ bands.

Figure 3. The early summer satellite image of the demonstration site

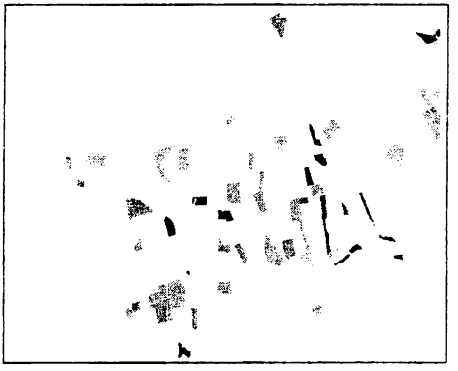
The principle of classification is the assumption that each vegetation category is the composition of appropriate *classes*, each class having multivariate normal distribution in the intensity space. Classes will be determined during the *teaching phase*, which consists of clustering followed by refinement of classes. In the *classification phase*, a decision rule is used to assign a class to each pixel. A pixel is mapped to the class having maximal density function value in the point defined by the pixel; this is called maximum likelihood (ML) decision rule. It must be noted here that this "per-point" classification method totally ignores spatial homogeneity, as it uses only intensities of pixels.

2.1. Steps of classification

The classification process will be described in four major steps.



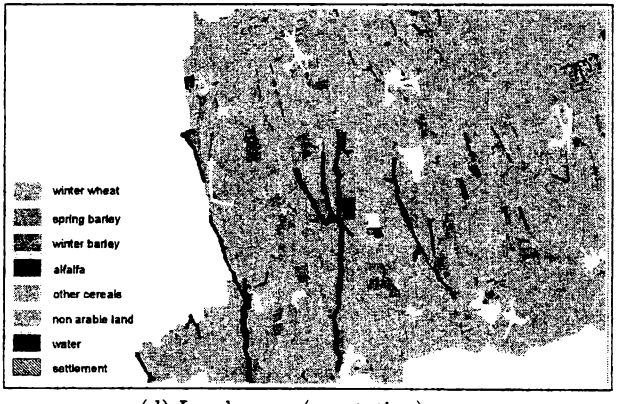
(a) Cluster map



(b) Training reference data

plant code	1	2	3	4	6	7	14	18	27	
plant	winter wheat	spring wheat	spring barley	winter barley	oats	triticale	alfalfa	rape	non arable	ref. data (ha)
növény	őszi búza	tavaszi búza	tavaszi árpa	őszi árpa	zab	tritikálé	lucerna	repce	áll. füves	ref. ter. (ha)
class1	0									0
class2	280		3	1		17				301
class3	252	0	0	0	0	0	1	1	1	255
class4	13			4				108	1	126
class5	0			80				1	2	83
class6	24		15				3		6	48
class7	0			93				1	1	95
class8	381		13				1	1	1	396
class9	2						109		12	123

(c) A partial summary matrix



(d) Land cover (vegetation) map

Figure 4. The steps of traditional classification process

1. The first step of determining classes is *clustering*. This is an unsupervised classification process, that is, no preliminary information from the target area is used; it is completely based on the characteristics of the image. It means, the evolution of classes is not restricted by reference data, but they are allowed to fit to the whole image. In our case, Isodata clustering is used with default initial cluster centres. As a result, we get *signatures* of classes, i.e. centres (averages) and convariance matrices. Fig. 4(a) shows the cluster map we obtained after classifying the original image using these clusters.
2. In the second major step reference data are used to derive the classes from the clusters. Reference data (or ground truth data) is a thematic map describing some previously known parts of the target area. This data set is divided into two parts. The training area is used to extract class parameters, while the test area forms the basis for accuracy assessment. The training area for our demonstration site can be seen in Fig. 4(b), where gray levels refer to land cover categories.

   In this step the aim is to search for correspondence between data classes and land cover categories on the trainig area. The intersection of each cluster with each category is calculated, which results in the summary matrix. The element in the ith row, jth column gives the number of pixels belonging to the ith cluster and jth category. The highlit row and column of matrix in Fig. 4(c) illustrate the intersection of the cluster drawn with black in 4(a) and winter wheat reference data in 4(b).

   In the ideal case conforming to the above mentioned assumption each class would dominantly belong to one category. One of the undesired cases is that a class does not belong significantly to any category. In that case the class will be discarded. The other case, where a class intersects with several categories, needs more sophisticated elaboration (decomposition), intensively exploiting human expertise.
3. Using the new, refined signature set obtained from the training process, a land cover map (Fig. 4(d)) is derived. It assigns a category to each pixel, determined by the decision rule and the class-to-thematic-category mapping.
4. Finally an accuracy assessment is done, using the test part of the reference data. In the event of high error rate we reconsider some of the former steps with revised parameter settings.

The main drawback of the per-point processing is that it ignores the similarity of neighbouring pixels. As it can be seen in Figure 6(a), the number of deviant pixels is higher than what is tolerable. The variation between neighbouring pixels is considerable. A possible improvement is to amalgamate

certain pixels to segments and adapt the rest of the classification to segments of pixels. This process will be described in the following.

3. Algorithm for segmentation

We have introduced a new initial step, called segmentation, which is regarded as the principle of segment based processing. A *segment* is a set of spectrally similar neighbouring pixels. Segments are formed from homogeneous rectangular *cells* of pixels as the smallest building blocks. In this new *per-field* approach, instead of pixels, segments become the units of image analysis. Schoenmakers [3] gives a good summary on segmentation methods in his dissertation. In this paper a region growing method is presented, which basically follows the principles described by Kettig and Landgrebe [4].

The input of the algorithm for segmentation is the digital image, the parameters are the size of a cell (usually $2 * 2$) and some statistical thresholds. The output is a thematic map, called the *segment map*, which is a matrix of non-negative integer labels. A matrix element assigns a segment number to a cell, where the value of zero designates inhomogeneous cell. The result highly depends on the mentioned statistical thresholds. Currently they have to be set up manually, by the examination of the frequency function of the image intensity and density function of the used similarity statistics.

3.1. Similarity tests

First, the image is divided into cells, as the elementary units of segmentation. Let $X = (x_1, x_2, \ldots, x_n)$ represent pixels of a homogeneous cell. The cell is considered homogeneous if the estimation of its normalized deviation (c_j in the jth band) falls below a certain predefined limit, i.e.

$$\forall j (1 \leq j \leq b) : c_j = \frac{s_j}{\overline{x}_j} \leq C_H,$$

where $\overline{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$, the mean in the jth band, and $s_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} x_{ij}^2 - \frac{n}{n-1} \overline{x}_j^2$, the empirical variance. This test selects the cells with pixels intensity disperses close to the mean. On the other hand, it discards the cells with high variability in proportion to the cell mean, among them, the cells with reflectance level differing from zero non-significantly.

In the following, statistically similar, neighbouring cells will be merged into segments. Inhomogeneous cells are ignored in the segmentation steps. The annexation criterion is a hypothesis test. It takes the cells in a prescribed order, detailed in 3.2.

Outline of the annexation test. Let $Y = (y_1, y_2, \ldots, y_m)$ be the pixels of a segment's current extent and X a neighbouring cell.

X can be connected to Y if the inequalities $L_1 \geq C_1$ and $L_2 \geq C_2$ hold, where C_1 and C_2 are predefined thresholds. L_1 is an ANOVA-like statistics, which measures the equality of means. L_2 measures the equality of covariances of X and Y. The comparison with the resulting L_1 tests for the hypothesis of equal mean vectors (first-order statistics), while L_2 is used for testing of equal covariances (second-order statistics). This way we assure the approximate equality of both the mean and the covariance.

L_1 and L_2 can be calculated independently. Let $\overline{x}$ and $\overline{y}$ mean the average of X and Y, respectively, and M denote the new segment centre (in the spectral space), i.e. $M = (n\overline{x} + m\overline{y})/(n+m)$. A_X, A_Y, B_X and B_Y are covariance-like quantities:

$$A_X = \sum_{i=1}^{n}(x_i - \overline{x})^2, \quad A_Y = \sum_{i=1}^{m}(y_i - \overline{y})^2,$$

$$B_X = \sum_{i=1}^{n}(x_i - M)^2, \quad B_Y = \sum_{i=1}^{m}(y_i - M)^2.$$

Using this notation, and $N = n + m - 2$, $A = A_X + A_Y$, $B = B_X + B_Y$:

$$L_1 = (A/B)^{N/2},$$

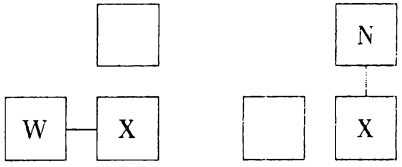
$$L_2 = \left(\frac{(A_X/n)^{n-1} * (A_Y/m)^{m-1}}{(A/N)^N} \right)^{1/2}$$

Once X is annexed to Y, their points will be treated together in the following.

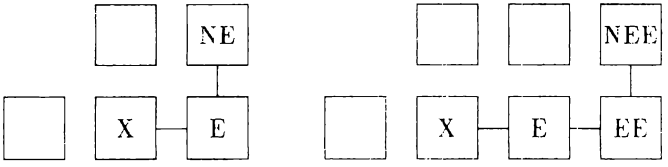
3.2. The strategy of the annexation

Merging starts from the upper left corner of the image, proceeds rowwise, and only homogeneous cells are examined. For each new cell, attempt is made to merge it to an existing segment, represented by the western (W), northern (N), first or second north-eastern (NE or NEE) neighbouring cell. Remote sensing images of the examined area have a special texture. Majority of the agricultural parcels appear as relatively simple polygons. The elaborated method eliminate the faults caused by the order of annexation.

First the procedure tries to merge the cell to its western or northern adjacent cell. These cells have already been assigned to a segment. If the cell can be connected to both segments (and they are not the same), the one with the closer centre is chosen.



If both cases fail, the algorithm investigates one or two cells to the east to find a way to another northern segment, as described by Fekete and Farkasfalvy [5]. The annexation criterion is examined for the east or second east neighbouring cells and their northern adjacent cells. If the criterion meets, an attempt is made to merge the cell X to this segment.



If all previous attempts fail, X will start a new segment.

4. Segment-based classification

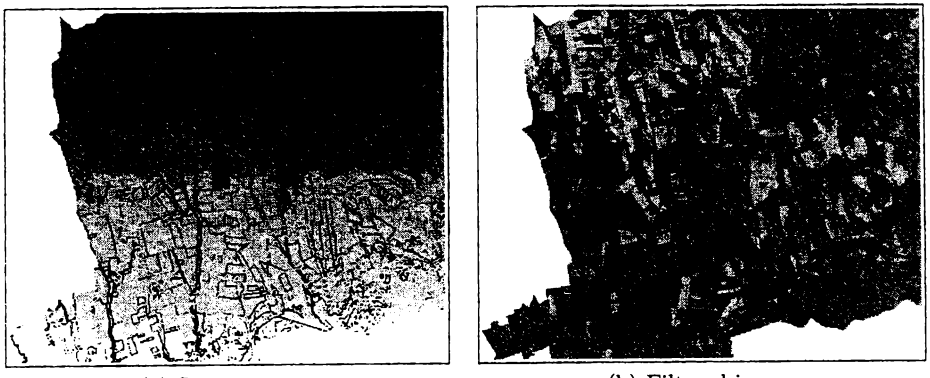
4.1. Integrating segmentation into classification

A new classification method has been developed, which exploits the benefits of segmentation. Instead of building a completely new, coherent segment-based classification system, steps of segmentation have been integrated into the existing program.

1. First, the image is divided into segments using the method described in Section 3. The principle of this approach is that pixels of the same segment represent the same land cover category. Accordingly, relatively strict criteria are used when forming segments from neighbouring pixels. This first step yields the segment map, as formerly mentioned. Gray levels

in Fig. 5(a) refer to segment numbers. Inhomogeneous cells, which we disregard in this phase, are marked with black.

2. In the second preparatory step the so-called *segment-based filtering* replaces pixel values with the average intensity of the corresponding segment, while pixels of inhomogeneous cells are usually temporarily ignored. As one can expect, this way of distribution of classes will not be disturbed by deviant pixels. Usually, a regular land parcel will be partitioned into few segments, or, in the ideal case, it will match a single one. The resulting image, which can be seen in Fig. 5(b), is similar to the original one, but it consists of spots having identical colour (intensity value).
3. Using this special image as input, we execute the same steps (1-4 in 2.1) as in the traditional per-point classification.



(a) Segment map (b) Filtered image

Figure 5. The result of segmentation

4.2. Fine-tuning of segmentation

As it has been mentioned, the success of segmentation highly depends on the setting of numerical parameters (C_H, C_1, C_2).

In the first part a decision is made about homogeneity of each cell. According to experience, the distribution of the normalized deviation of intensity in the bands are similar. This is why a uniform threshold can be used for all the r bands in the concurrent hypothesis test. The ratio of homogeneous cells does not significantly depend on the number of bands. Therefore it is sufficient to determine the threshold once for a given type of RS images and a chosen cell size.

C_1 and C_2 are the mean and covariance stability parameters of the segments. They are determined in iterative manner, using reference data. If,

for a given parametrization, some segments significantly intersect with several land cover categories, we will have to refine thresholds in order to obtain more segments and to eliminate this undesired symptom. According to this principle, we can start from a relatively coarse partitioning, and gradually, possibly in several steps, an appropriate segmentation can be achieved. Fortunately, the above mentioned criterion for the adequacy of C_1 and C_2 – about the error in intersection – can be easily checked.

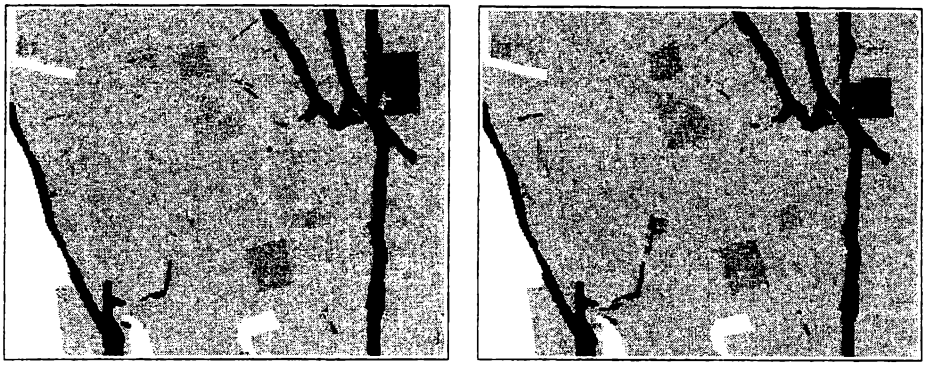
We have not emphasized yet that the segmentation method itself has a real multivariate (MV) and a multiple univariate (MUV) implementation. The formulas in subsection 3.1 apply in both cases with the only difference that variables are considered as vectors in the MV case. In the MV case, pixel values in different bands are treated together as a vector, while in the MUV case decision about annexation is made by bands, where every band has to fulfill the criterion.

The variation in C_1 and C_2 constants differently influences the number of segments in the two major approaches. Generally, the change in C_1 has greater impact on the results than that of C_2. For a given parametrization, the number of segments highly depends on the number of bands with the MUV model, while in the MV model this dependence cannot be observed. The MV approach, using the complete covariance matrix, treats all band values together, whereas MUV ignores the correspondence between the bands. The MUV model, approximating the covariance matrix, requires lower computational load.

There is also some flexibility in the treatment of pixels in inhomogeneous cells. Such a pixel can be merged with the adjacent segment (based on individual decision), or it can be ignored during the training process. Anyway, in the final classification step all pixels are involved.

4.3. Experience and development perspectives

The benefits of the per-field classification stems from incorporating spatial characteristics of images. Investigating the results of segmentation we found an interesting paradox. A segmented image seems to be somewhat artificial, consisting of too regular parts. Indeed, the method tends to simplify the real view: a large segment can "swallow up" a few strange points. However, in many cases such deviant points are present in the digital image because of noise. And even if there is not any noise, i.e. deviant pixels correspond to inhomogeneities in cultivated crop fields, accuracy of classification for large parcels significantly grows. The difference between land cover maps created with traditional (per-pixel) and segment-based (per-field) methods can be seen in images 6(a) and 6(b), respectively, both showing the same small area magnified.



(a) Traditional method (b) Segment-based method

Figure 6. Comparison of results I: magnified land cover maps

As segmentation highly reduces pixelwise noise, the overall accuracy increases. Contingency matrices in Fig. 7 demonstrate the increase in accuracy achieved by segmentation on the test area. The rows of the matrices refer to the resulting categories of classification while the columns are associated with reference categories. Entries show the number of pixels in the intersection of the given classification and reference category, i.e. row i, column j indicates the number of pixels classified as i but actually (according to reference data) being part of category j. Thus, the sum of the elements in the main diagonal gives the number of the correctly classified reference points. A description of the accuracy measures used in the classification can be found in the paper written by Nádor et al. [6].

Tests carried out so far show a significant improvement in accuracy for major crops (e.g. winter wheat, barley). However, in some cases, for less important crops we encountered inadequacy. For example, in the presented test case the traditional method proved to be more appropriate for alfalfa. The difference that can be seen in the upper right corner of Fig. 6(a) and (b) shows this phenomenon. (Here we omit the analyis of the causes, which stem from the small quantity of reference data.)

Our next aim is to involve the segment-based procedure into operational work. During the whole development, effort is made to raise the ratio of automatic steps, thus decreasing the need for human interaction. The final goal is to develop a uniform segment-based classification framework. A straightforward step is the introduction of the variability of cells' shape. The most important development direction is the improvement of the region growing strategy used

plant növény	winter wheat őszi búza	spring barley tavaszi árpa	winter barley őszi árpa	alfalfa lucerna	other egyéb	sum összesen
winter wheat	943	25	2	0	74	1044
spring barley	1	214	0	0	43	257
winter barley	0	0	323	0	5	328
alfalfa	5	0	1	72	31	108
other	70	31	9	3	472	585
Sum	1018	270	335	75	625	2323

overall 87%
kappa 86%

	Hellden	Short	H-S average
winter wheat	91%	84%	88%
spring barley	81%	68%	75%
winter barley	97%	95%	96%
alfalfa	78%	64%	71%

(a) Traditional method

plant növény	winter wheat őszi búza	spring barley tavaszi árpa	winter barley őszi árpa	alfalfa lucerna	other egyéb	sum összesen
winter wheat	994	19	2	0	63	1079
spring barley	6	244	0	0	57	307
winter barley	8	0	329	0	7	344
alfalfa	0	0	2	37	32	71
other	6	4	0	2	407	419
Sum	1015	267	333	39	566	2220

overall 91%
kappa 86%

	Hellden	Short	H-S average
winter wheat	95%	90%	93%
spring barley	85%	74%	80%
winter barley	97%	94%	96%
alfalfa	68%	51%	59%

(b) Segment-based method

Figure 7. Comparison of results II: contingency matrices

in the direction of the so-called Kriging Method (see e.g. Stein [7]). We also intend to systematically compare the results with those of the simulated annealing method.

References

[1] **Richards J.A.,** *Remote sensing, digital image analysis - An introduction,* Springer, 1993.

[2] **Csornai G., Dalia O., Farkasfalvy J. and Nádor G.,** Crop inventory studies using Landsat data on large area in Hungary, *Applications of remote sensing in agriculture*, Butterworths, 1990, 159-165.

[3] **Schoenmakers R.,** *Integrated methodology for segmentation of large optical satellite images in land applications of remote sensing,* Joint Research Centre, 1995.

[4] **Kettig R.L. and Landgrebe D.A.,** Classification of multispectral image data by extraction and classification of homogeneous objects, *IEEE Transactions on Geoscience Electronics*, **14** (1) (1976), 19-26.

[5] **Fekete I. and Farkasfalvy J.,** Automatic segmentation of multispectral digital images, *First Seminar on Artificial Intelligence, Visegrád, Hungary,* ELTE, Budapest, 1989, 121-126.

[6] **Nádor G., Csornai G. and Kocsis A.,** A tematikus pontosság mérésének módszerei a távérzékeléses növényterület felmérésben, *7th Seminar on Earth and Meteorological Observations by Satellites,* FÖMI, Budapest, 1997. (Methods of accuracy assessment of thematic vegetation mapping by remote sensing)

[7] **Stein M.L.,** *Interpolation of spatial data: Some theory for Kriging,* Springer, 1999.

(Received December 11, 2003)

I. László
Inst. of Geodesy, Cartography
and Remote Sensing
Remote Sensing Centre (FÖMI)
Bosnyák t. 5.
H-1149 Budapest, Hungary
istvan.laszlo@rsc.fomi.hu

T. Pröhle
Dept. of Probability Theory
and Statistics
Eötvös Loránd University
Pázmány P. s. 1/C
H-1117 Budapest, Hungary
prohlet@ludens.elte.hu

I. Fekete
Dept. of Algorithms and
Applications
Eötvös Loránd University
Pázmány P. s. 1/C
H-1117 Budapest, Hungary
fekete@ludens.elte.hu

G. Csornai
Inst. of Geodesy, Cartography
and Remote Sensing
Remote Sensing Centre (FÖMI)
Bosnyák t. 5.
H-1149 Budapest, Hungary
gabor.csornai@rsc.fomi.hu