5^{th} *International Conference on Applied Informatics*
Eger, Hungary, January 29 – February 2, 2001

A segment-based classification method for satellite images

István László[1,2], Gizella Nádor[2], István Fekete[1], Gábor Csornai[2], Attila Kocsis[2]

[1] Department of General Computer Science, Loránd Eötvös University
[2] Institute of Geodesy, Cartography and Remote Sensing (FÖMI), Remote Sensing Centre

Dedicated to Prof. Mátyás Arató and Prof. László Varga on their 70th birthday

Abstract

The importance of remotely sensed images increases parallel to the development of technology. The scope of this article is the processing of satellite images for agricultural purposes. FÖMI Remote Sensing Centre regularly reports plant area and yield predictions to the Ministry of Agriculture and Regional Development. To support this, a classification system has been working operationally for years, on a 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 literature. The improvement was integrated into the operationally working system in a sophisticated way.

Advantages of the new method are obvious for large land parcels, where the accuracy increased as a result of incorporating spatial characteristics.

Categories and Subject Descriptors: I.4.6 [Image processing]: Segmentation - *region growing, partitioning, pixel classification*

Key Words and Phrases: remote sensing, agricultural applications, crop monitoring, image classification, segmentation

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 the subject of RS, see e.g. Richards [1].

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.

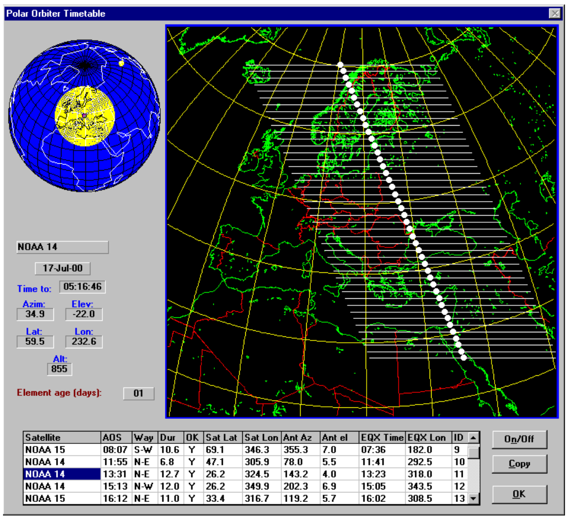


Figure 1: A pass of the satellite NOAA-14

A definite part of the whole elegtromagnetic spectrum, called optical band, is involved in RS. The radiation of the surface is measured by sensors, each capturing a given wavelength subinterval. 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 n-dimensional arrays, containing the intensity values recorded by the n sensors.

RS images have 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 number of the *spectral bands* and wavelength subintervals associated with them are fitted to the requirements of applications. *Radiometric resolution* refers to the number of distinguishable 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.

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. 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. This fact can be used to make distinction between land covers, which is a fundamental requirement of RS. In Figure 2 the reflectance function of a typical cultivated crop is shown. A representative sample of this function is taken

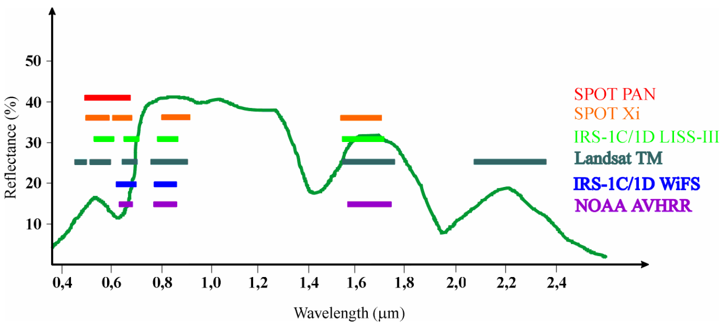


Figure 2: A typical reflectance function and spectral bands of satellite sensors

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.

Immediate application of determining land covers is *vegetation mapping*, which serves as a basis for area estimation. The same 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 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 Present classification method

2.1 Outline

The main purpose of classification is to produce the thematic 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. Satellite images and the database of previously known surface areas (reference data) are the input for classification. The sophisticated procedure is heavily based on clustering, and it intensively uses the human expert's knowledge.

To start with, a multi-band image of the target area is composed from several images having different acquisition dates. As it was mentioned, different dates are required to separate plants that are similar in a given period. In the following, we are going to use the eastern 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 4 layers.

The principle of classification is the assumption that each vegetation category is the

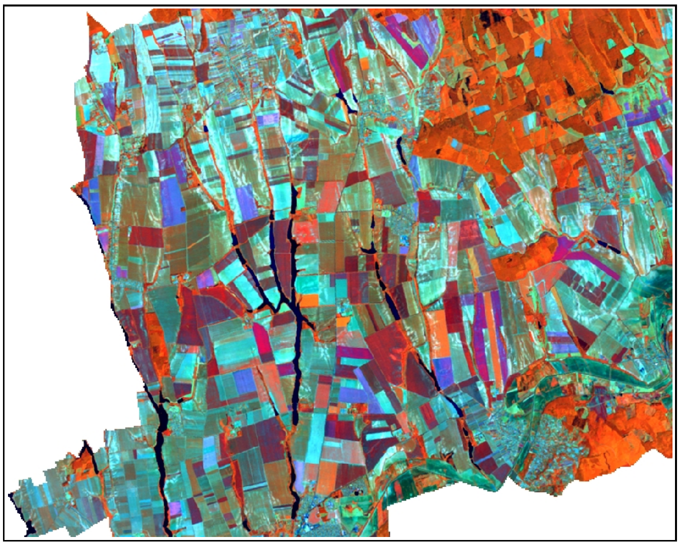


Figure 3: A snapshot of the demonstration site

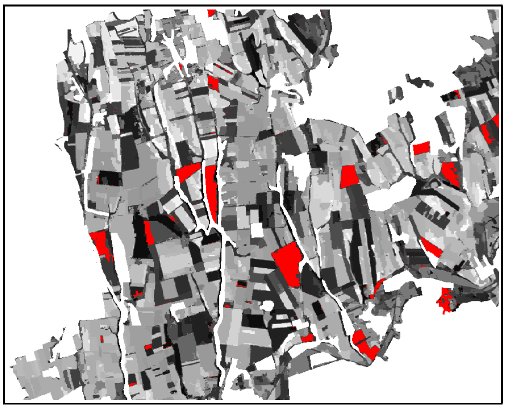
composition of appropriate classes, each class having multivariate normal distribution in the intensity space. Classes will be determined during the *teaching phase*, which consists of clustering followed by refinement of classes. In the *classification phase*, a decision rule is used to assign a class 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.

2.2 Steps of classification

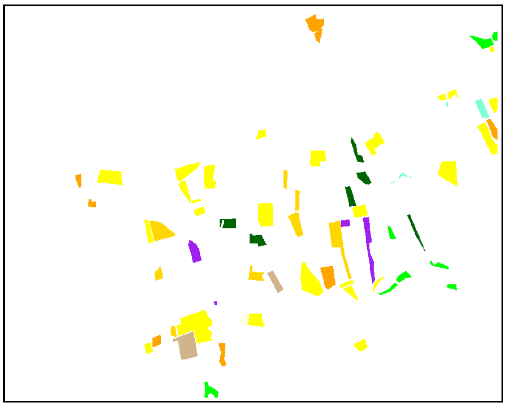
The classification process will be described in four major steps.

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. Classifying the original image using these classes, the cluster map, shown by Fig. 4(a), is obtained.

2. In the second major step reference data are used to refine the classes. 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).



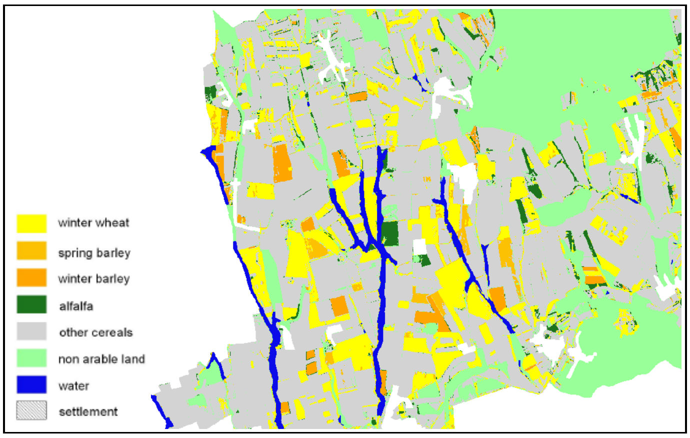
(a) Cluster map



(b) Training reference data

plant code plant növény	1 winter wheat őszi búza	2 spring wheat tav. búza	3 spring barley tav. árpa	4 winter barley őszi árpa	6 oats zab	7 triticale tritikálé	14 alfalfa lucerna	18 rape repce	27 non arable állandó füves	ground data (ha) refter (ha)
class1	0									0
class2	280		3	1		17				301
class3	252	0	0	0	0	0	1	1	1	254
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

Searching for correspondence between classes and categories on the trainig area, the intersection of each class 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 class and jth category. The matrix in Fig. 4(c) illustrates the intersection of the cluster highlighted with red in 4(a) and winter wheat reference data marked with yellow 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 has already been mentioned: it ignores the similarity of neighbouring pixels. A possible improvement is to adapt some parts of the classification to segments of pixels.

3 Algorithm for segmentation

3.1 Overview

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 [5] gives a good summary on segmentation methods in his dissertation. The method presented in this paper basically follows the principles desribed by Kettig and Landgrebe [3].

The input of the algorithm for segmentation are the digital image, the size of a cell (usually $2 * 2$) and some statistical thresholds. The output is a segment map, which is a matrix of integer labels. A matrix element assigns a segment number to a cell, where the value of zero designates inhomogeneous cell. The method highly depends on the mentioned statistical thresholds. Currently they have to be set up manually; in the following we will see possibilities how their determination is helped by the computer.

3.2 Similarity tests

First, the image is divided into cells, as the elementary units of segmentation. A 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.

$$\forall j (1 \leq j \leq r) : c_j = \frac{\sigma_j}{\overline{x}_j} = \frac{1}{\overline{x}_j} \left(\frac{1}{n-1} \sum_{i=1}^{n} x_{ij}^2 - \frac{n}{n-1} \overline{x}_j^2 \right)^{1/2} \leq C_H$$

Inhomogeneous cells will be ignored in the further segmentation steps.

Statistically similar, neighbouring cells will be merged into segments. The annexation criterion is a statistical hypothesis test.

Let $Y = (y_1, y_2, \ldots, y_m)$ be the pixels of a segment's current extent and $X = (x_1, x_2, \ldots, x_n)$ represent pixels of a homogeneous 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 and L_2 can be calculated independently:

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

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

$$N = n + m - 2$$

$$A = A_X + A_Y \qquad B = B_X + B_Y$$

Let x and y mean the average of X and Y, respectively, and M denote the new segment centre (in the spectral space), i. e. $M = (nx + my)/(n + m)$. A_X, A_Y, B_X and B_Y are covariance-like quantities:

$$A_X = \sum_{i=1}^{n} (x_i - x)^2 = \sum_{i=1}^{n} x_i - n * x^2$$

$$A_Y = \sum_{i=1}^{m} (y_i - y)^2 = \sum_{i=1}^{m} y_i - m * y^2$$

$$B_X = \sum_{i=1}^{n} (x_i - M)^2 = A_x - n * (x - M)^2$$

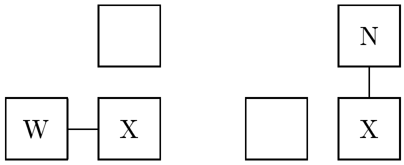
$$B_Y = \sum_{i=1}^{m} (y_i - M)^2 = A_y - m * (y - M)^2$$

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 covariance matrices (second-order statistics). If X is annexed to Y, the composition of their density function is calculated.

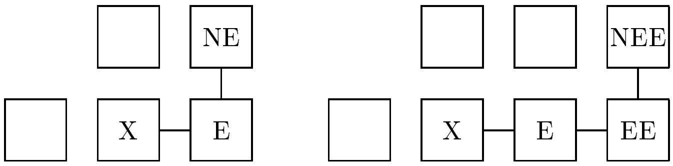
3.3 The strategy of the annexation

Merging starts from the upper left corner of the image, proceeds rowwise, and only homogeneous cells are examined.

First the procedure tries to merge the cell to its western or northern adjacent cell. These cells, if they are homogeneous, 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 segment, as described by Fekete and Farkasfalvy [4]. The annexation criterion is examined for the east or second east neighbouring cells and their northern adjacent cells. If the criterion meets, the cell X is also tried to be merged to this segment.



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

4 Segment-based classification

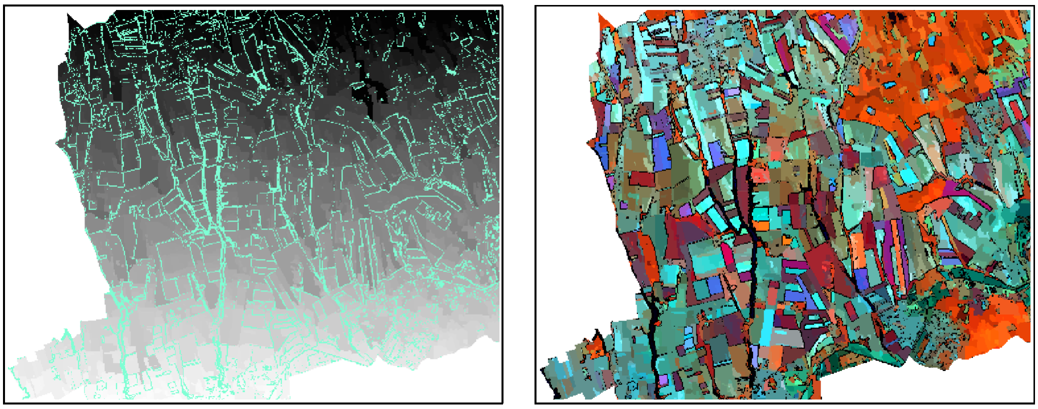
4.1 Integrating segmentation into classification

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 light green.

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 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, as 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 2.2) as in the traditional per-point classification.



(a) Segment map (b) Filtered image

Figure 5: Result of segmentation

4.2 Fine-tuning of segmentation

The segmentation method itself has a real multivariate (MV) and a multiple univariate (MUV) implementation. 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.

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.

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

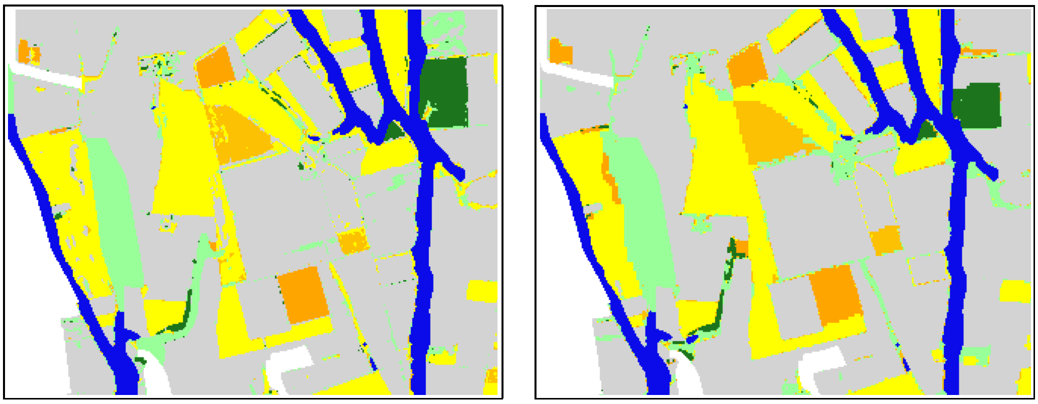
In the first part, where decision is made about homogeneity of cells, normalized deviation values in each band are calculated independently, and their distribution is similar. This is why a uniform threshold can be used for all bands. For a given type of RS images, it depends only on the cell size; ratio of homogeneous cells does not significantly depend on the number of bands. Therefore it is sufficient to determine the threshold once.

Reference data can help to set appropriate values for C_1 and C_2. 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 iteratively an appropriate segmentation can be achieved. As the criterion for refinement can be easily checked, setting of parameters can be done nearly automatically.

As it can be expected, the change in C_1 has greater influence on the results than that of C_2. These parameters differently affect the number of segments in the two major approaches. For a given parametrization, the number of segments highly depends on the number of bands in the MUV model, while in the MV model this dependence cannot be observed. Actually, MV approach literally conforms to the hypothesis test, whereas MUV is only an approximation requiring lower computational load. Experiments also confirm the adequacy of the MV model in general.

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 is increased. Contingency matrices in Fig. 7 demonstrate the increase in accuracy achieved by segmentation on the test area. Rows of a matrix refer to the result categories of classification, columns are associated with reference categories. Entries show 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.

Radical decrease in pixelwise noise also proves to be greatly helpful to the human expert. As it was mentioned, refinement of classes during the teaching phase heavily relies on human knowledge. Working with segmented images makes decisions easier.

	winter wheat	spring barley	winter barley	alfalfa	other	Sum
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

	winter wheat	spring barley	winter barley	alfalfa	other	Sum
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

After segment-based filtering the number of different intensity values enormously decreases, this makes Isodata clustering converge much faster. Beside, the processing time, while the amount of human effort will be also lowered. As a side effect of low number of intensities, dynamics of values within a class is weakened, which means classes with degenerate distribution can appear and it can cause difficulties in the teaching phase.

Exhaustive tests 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 above case the traditional method proved to be more appropriate for alfalfa. (Difference seen in the upper right corner of Fig.6(a) and (b) shows this phenomenon. Here we omit analyis of the causes, which stem from the small quantity of reference data.)

Our next aim is to introduce the segment-based procedure into operational work. To achieve this, a great deal of experience is required, included adequate parameter setting for robust operation. 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.

References

[1] Richards, J. A.: *Remote Sensing, Digital Image Analysis — An Introduction.* Springer-Verlag, 1993.

[2] Csornai G., Dalia O., Farkasfalvy J., Nádor G.: Crop inventory studies using Landsat data on large area in Hungary. In: *Applications of Remote Sensing in Agriculture,* pp. 159–165, Butterworths, 1990.

[3] Kettig, R. L., Landgrebe, D. A.: Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects. *IEEE Transactions on Geoscience Electronics*, 14(1):19–26, 1976.

[4] Fekete I., Farkasfalvy J.: Automatic segmentation of multispectral digital images. *5th Conference of Program Designers,* ELTE, Budapest, 1989.

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

Contact information

István Fekete
fekete@ludens.elte.hu

Department of General Computer Science
Eötvös Loránd University
1117 Budapest, Pázmány P. sétány 1/D
Hungary

István László
listv@rsc.fomi.hu

Gizella Nádor
g.nador@rsc.fomi.hu

Gábor Csornai
gabor.csornai@rsc.fomi.hu

Attila Kocsis
attila.kocsis@rsc.fomi.hu

FÖMI Remote Sensing Centre
1149 Budapest, Bosnyák tér 5.
Hungary