Watershed based image segmentation — an effective tool for detecting landscape structure

Miloš Šrámek[a] and Thomas Wrbka[b]

[a]Institute of Measurement Science, Slovak Academy of Sciences
Dúbravská cesta 9, 842 19 Bratislava, Slovakia

[b]University of Vienna, Dept. of Vegetation Ecology and Conservation Biology
Pbox 285, A-1091 Vienna, Austria

ABSTRACT

Satellite image segmentation with the aim to detect spatial units having an ecological meaning has become an important field in methodological research in modern landscape ecology. Within the theoretical framework, elaborated in,[1] a landscape can be defined as spatial arrangement of ecosystems. Regions, that are more or less homogeneous in that sense, become more and more important as land units for physical-planning purposes. Such spatial objects can of course differ in size, in Central Europe they usually can be deliminated at the scale of some square kilometers. Therefore such objects should be detectable on satellite images and can then be ecologically characterized by their most important features - structure, function and change. Elaborating operational procedures to do that for the Austrian territory is the aim of an multidisclipinary research project called *SINUS - Structural Features of Landscapes as Indicators for Sustainable Land Use*, which is financed by the Austrian Ministry for Science and Transportation an will be finished in 1999. This article can be regarded as one methodological output of this project.

Small working groups with limited computing resources are usually facing the problem of high computation costs due to time-consuming segmentation procedures. This was the motivation to develop a new effective tool for detecting landscape structure on satellite images based on the concept of watersheds. Therefore the segmentation and classification package `imagine-ws` was elaborated to create spatially homogeneous regions and classify them by means of externally provided class signatures. For each of the identified regions a set of attributes is computed and stored: size - defined as the number of pixels, elongatedness - based on the principal axis, orientation - computed as an angle of the main principal axis and irregularity - measured as a ratio of real contour length and the contour length of an ellipse with the same area, elongatedness and orientation. Both data sets - the regions and their attributes can be used for ecological interpretation. Analyzing landscape structure means the distinction of three major components: the matrix - as the element which covers most of a given region and is regarded to control major ecological processes; the corridors - which are linear elements conducting material and energy fluxes through a region and can serve as pathways for the movement of organisms; and at least the patches - small elements which differ from the surrounding matrix for historical or biogeochemical reasons. In Austrian cultural landscapes, which are heavily influenced by man since ages these three basic element types are usually found in typical combinations, so called “landscape types”. Examining a first test region in Eastern Austria the new segmentation method was proved to deliver reasonable results. Most of the classified regions detected by `imagine-ws` were found to be congruent with ecologically defined landscape-types deliminated by terrestrial mapping.

Keywords: Landscape ecology, Landscape structure, Scale-space concept, Watershed Transform, Region hierarchy, Region attributes

1. INTRODUCTION

Within the framework of the Austrian national research program “sustainable development of cultural landscapes” several projects are carried out to elaborate so-called “indicators for sustainability”. Among the different approaches for this kind of indication the use of information about landscape structure derived from satellite images is a very promising one. Therefore, satellite image segmentation with the aim to detect spatial units having an ecological meaning has become an important field

---

Email: sramek@savba.sk, wrbka@pflaphy.pph.univie.ac.at. Supported by the project *Sustainable Development of the Austrian Cultural Landscapes* of BMWV Austria, No. IN 2 in methodological research in modern landscape ecology. One of the problems in this field of research is the lack of region-based classification procedures, which are not disturbed by details and therefore reducing computation time. As a solution, the classification and segmentation package `imagine-ws` was developed, allowing for (i) segmentation of multiband Landsat TM images into spatially homogeneous regions and (ii) classification of such regions to different usage classes by means of externally provided class signatures.

2. SCOPE OF RESEARCH - DETECTING LANDSCAPE STRUCTURE

As stated above, the structure of landscapes can be regarded as most interesting feature when investigating them from an ecological point of view. Based on the concept of FORMAN & GODRON[1] a landscape can be defined as spatial arrangement of ecosystems. Regions, that are more or less homogeneous in that sense, become more and more important as land units for physical-planning purposes. Such spatial objects can of course differ in size, in Central Europe they usually can be deliminated at the scale of some square kilometers. Therefore such objects should be detectable on satellite images and can then be ecologically characterized by their most important features - structure, function and change. Elaborating operational procedures to do that for the Austrian territory is the aim of a multidisclipinary research project called *SINUS - Structural Features of Landscapes as Indicators for Sustainable Land Use*, which is funded by the Austrian Ministry for Science and Transportation an will be finished in 1999. This article can be regarded as one methodological output of this project.

How can we be sure that there is a relationship between the spatial arrangement of ecosystems and “sustainability”? As it was shown in[2] for the State of Georgia, there exists a close relationship between the complexity of landscapes, measured for instance by means of the fractal dimension of their land-units, and the intensity of land-use, measured for instance by the amount of agrochemicals used in the investigated regions. It could be shown that the fractal dimension of different land-use types was reduced significantly during the last decades, whereas the agricultural productivity and the consumption of fuel, fertilizers and pesticides were increased in the same period. This results lead us to the assumption that spatial pattern and intensity of land-use in Austrian cultural landscapes are linked in a similar way and therefore landscape structure could be a suitable “indicator for sustainability”. The empirical basis for that concept and its application to Austria was therefore elaborated in a case-study for a small region along the Hungarian border.[3] It turned out that for this area, in which a variety of different land-use systems occurs, the negative correlation between intensity of human influence and complexity of the shape and distribution pattern of landscape elements, is also true.

Encouraged by this preliminary studies, a methodology for investigating the whole Austrian territory - about 87.000 km - had to be developed. In our case a two way procedure combining a top-down with a bottom-up approach was chosen. This means that the analysis of the fine-grain landscape structure derived from field studies in 140 test regions has to be extrapolated to spatial units on a larger scale. This is done by classifying these test regions according to the spatial arrangement of their basic landscape elements and identifying these classes as landscape types. The top-down approach is based on the segmentation of multiband Landsat TM images into regions which are more or less homogeneous with respect to spectral attributes, followed by a classification of such regions by means of externally provided class signatures. Such regions can then also be treated as landscape-types with an ecological meaning. It is obvious, that a procedure like this requires a powerful tool for computing a large amount of satellite data in a reasonable time.

3. WATERSHED BASED SEGMENTATION OF LANDSAT-TM IMAGES

The basic idea of the proposed region-based approach is that classification of approximately homogeneous regions is less sensitive to superimposed noise than pixel-based classification, due to averaging of their properties over their area. The next advantage of the approach is that it enables to introduce *scale* into processing. Usually, large data sets are processed at larger scale levels, i.e. with suppression of small details, which enables us to concentrate to features relevant to the processed area, not being disturbed by the details. Introduction of scale into the problem thus significantly reduces computation cost by reducing number of processed primitives.

3.1. The Scale-Space Concept and the Watershed Transform

Objects in the world are perceived only over certain ranges of scale. A typical example of this fact is cartography: a map of the world depicts continents, big islands, rivers and maybe some of the major cities, while a city guide shows streets, buildings, parks and other details. From this point of view, atlas can be understood as a multiscale symbolic representation of the world around us.

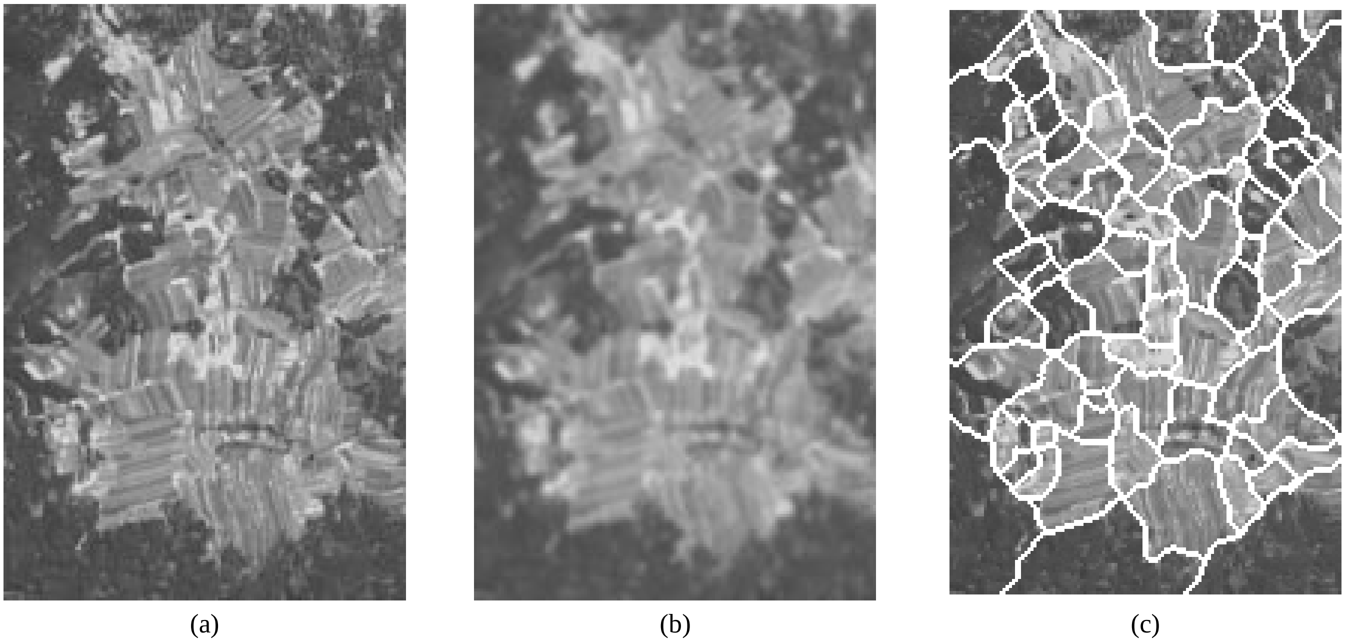


Figure 1. *(a) Original image, image smoothed by Gaussian with $\sigma = 2.0$(c) and the original image overlayed by region contours.*

Transferring this concept into the area of image processing and vision, we should represent an image as a sequence of smoothed images, showing gradually less details.[4] Theory has shown, that Gaussian

$$\mathcal{G}(x,y,\sigma) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{\sum_i x_i^2 + \sum_j y_j^2}{2\sigma^2}\right] \tag{1}$$

is the smoothing operator, that should be involved to generate this derived sequence, due to its linearity and spatial shift invariance as well as the fact that it introduces no new accidental structures (e.g. local extrema).[5] The level of smoothing can be characterized by the parameter σ, the standard deviation, approximately defining dimensions of suppressed details (Figure 1).

The goal of low-level vision operations is to identify important image features, usually edges or regions, that can be later used for image interpretation. Therefore, the linear smoothing by Gaussian kernel should be combined with some other operators in order to get some more explicit descriptors of the scene geometry. The Canny's edge detector is a well known example of this approach.[6]

Another possibility how to extract edges and regions from digital images, attracting interest of the vision community in the recent years, is the concept of *watersheds*, adopted from topography.[7] From the point of view of this concept, gray scale images are considered as topographic reliefs. Around each local minimum M of such image a *catchment basin* $C(M)$ is defined, such that each of its points can be connected with the minimum M by a descending path, called *downstream*. Lines, separating different catchment basins are called *watersheds*.

Regions, identified in the segmentation process should represent areas with some level of density homogeneity. Therefore, a gradient operator $\mathcal{E}$ is usually applied in order to enhance the inhomogeneities (edges) and this image is then subjected to the watershed transform. Thus, taking into account the aforementioned scale-space concept, the segmentation by watersheds can be expressed as

$$S = \mathcal{W} \star \mathcal{E} \star \mathcal{G}(\sigma), \tag{2}$$

where the standard deviation σ of the smoothing kernel $\mathcal{G}$ defines level of details and thus size of the detected regions.

3.2. Watershed hierarchies

As we have already mentioned in the previous section, smoothing by the Gaussian filter results in the desired larger regions. However, also region contours are affected: they are also smoother and do not follow edges exactly. To avoid this drawback, we proposed a technique, based on watershed hierarchies, i.e. on a sequence of regions obtained by smoothing the image by different Gaussians with increasing σ.

Let I_σ is a segmented image smoothed by a Gaussian with kernel size σ and $S(\sigma_1, \sigma_n)$ is a sequence of n segmented images with $\sigma \in \{\sigma_1 = \sigma_{min}, \sigma_2, \ldots, \sigma_n = \sigma_{max}\}$. We tested two possibilities, how to build such sequence:

1. $\sigma_{i+1} = \sigma_i + \sigma_s$ (additive sequence) and
2. $\sigma_{i+1} = \sqrt{2}\sigma_i$ (multiplicative sequence).

Experiments have shown that in the case of the multiplicative sequence size of the region approximately doubles in one step, which indicates merging of neighboring pairs. Since we liked this property, only multiplicative sequences were used.

Images with small σ have precise contours but small mean region size, while images with large σ value have larger regions with imprecise contours. The idea of the watershed hierarchy based segmentation is to transfer the (precise) contours from low levels to the large regions at the higher levels of the hierarchy by means of *region overlapping*.

Let us have 2 segmentations at levels σ_i and σ_{i+1}.

1. For each region j at level σ_i a region k at level σ_{i+1} with the largest number of common pixels is found. Since the region contours are between two neighboring levels shifted only moderately, there is usually one region k, which dominates in the number of common pixels.
2. A temporary image is created by assigning to each region j at the level σ_i the label of the corresponding region k of level σ_{i+1}. Since there are usually two or several regions at σ_i which overlap with the region k, a new region is defined, with size corresponding to k at level σ_{i+1} but with contours from level σ_i.
3. Replace segmentation at level σ_{i+1} by the temporary image. Thus, the precise contours are transferred to the level σ_{i+1}.

In order to build the region hierarchy $H(\sigma_1, \sigma_n)$ corresponding to the sequence $S(\sigma_1, \sigma_n)$, but with precise contours, we have to process all levels, starting from σ_1:

```
for($i = 1; i < n; i = i + 1$)
          $I_{i+1} = \mathrm{overlap}(I_i, I_{i+1});$
end for
```

It is necessary to point out that both i-th images in $S(\sigma_1, \sigma_n)$ and $H(\sigma_1, \sigma_n)$ have the same number of regions. Figure 2 shows an example of 2 segmentations of an image, with different smoothing and with and without the region overlapping. We can see the perfect coincidence of contours on the overlapped image and shifted contours on the image without the overlap test.

4. REGIONS AND THEIR ATTRIBUTES

In order to enable higher level processing of the segmented regions by the `Imagine` system, a set of attributes is computed and stored for each region: size, elongatedness, orientation and irregularity.

4.1. Size

Size S of the region $\mathcal{R}$ is defined as number of its pixels.

4.2. Elongatedness

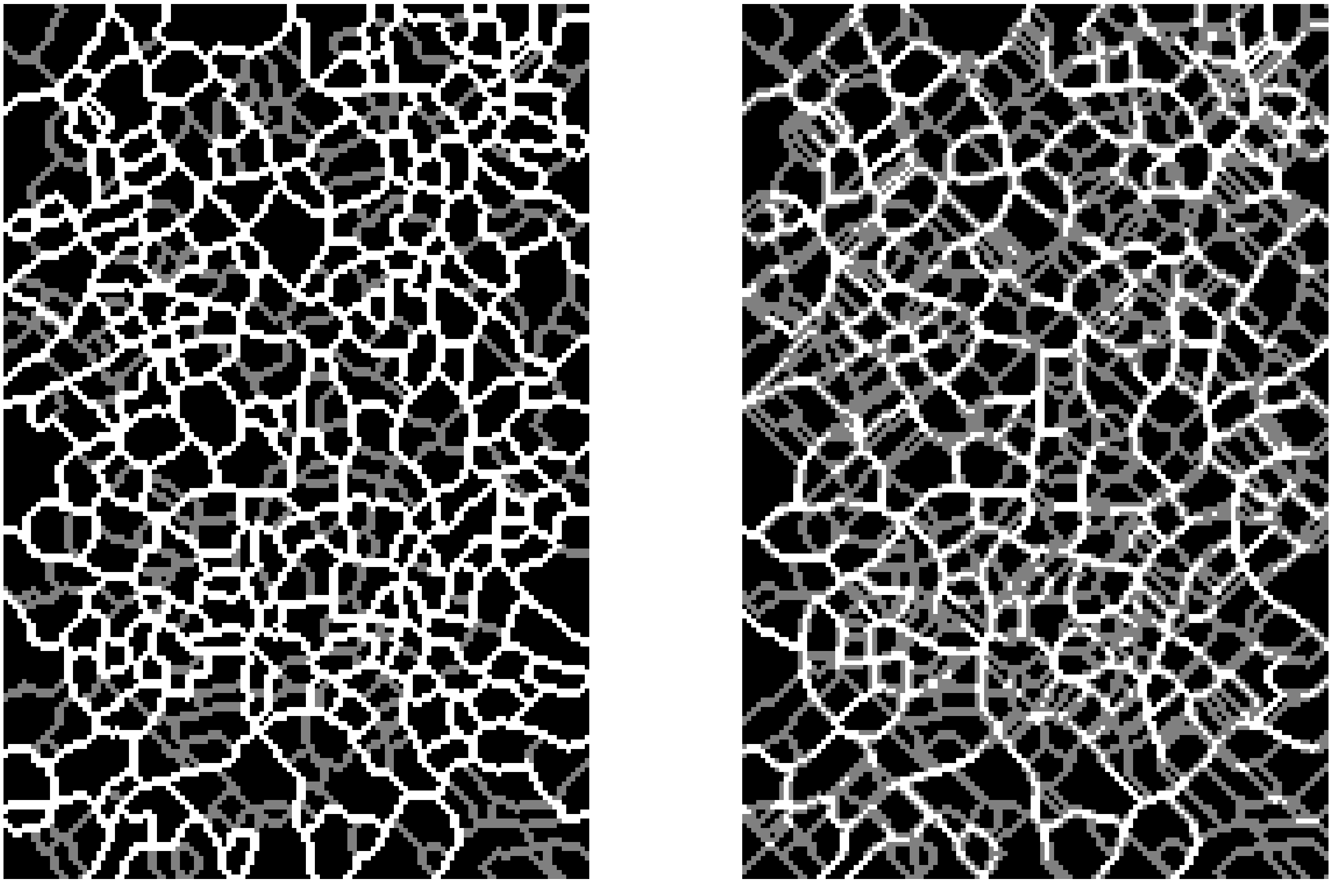
Computation of elongatedness of the region $\mathcal{R}$ is based on the principal axis transform:

$$L = \sqrt{\frac{\lambda_1}{\lambda_2}}, \tag{3}$$

where λ_1, λ_2 are sorted eigenvalues of the covariance matrix $\mathbf{C}$ of the region:

$$\mathbf{C} = \mathbf{E}(\mathbf{x} - \mathbf{m}_{\mathcal{R}})(\mathbf{x} - \mathbf{m}_{\mathcal{R}}). \tag{4}$$

where $\mathbf{x} = (\mathbf{x_i}, \mathbf{y_i})$ is coordinate of a pixel belonging to the region of interest and $\mathbf{m}_{\mathcal{R}}$ is mean of coordinates of all pixels in the region. For a perfectly symmetric region $L = 1$, for other shapes $L > 1$ (Figure 4 and Tab. 1, column *Elongatedness*).



overlapping no overlapping

Figure 2. *Coincidence of region contours smoothed with $\sigma = 2.8$ and $\sigma = 4$. Bright values represent pixels belonging to contours at both levels, darker only to one contour at one level.*

4.3. Orientation

Orientation of the region $\mathcal{R}$ is computed as an angle of the main principal axis with respect to the horizontal axis. It is computed by the principal axis transform again:

$$\varphi = \arcsin \frac{v_{y_1}}{\sqrt{v_{x_1}^2 + v_{y_1}^2}}, \tag{5}$$

where v_{x_1} and v_{y_1} are the x and y components of the first eigenvector of $\mathbf{C}$. φ is expressed in degrees, $\varphi \in < -90, 90 >$. If elongatedness L of $\mathcal{R}$ is near to or equal 1, φ is loosing its sense and can get arbitrary values (Figure 4 and Tab. 1, column *Orientation*).

4.4. Irregularity

Irregularity of the region $\mathcal{R}$ is measured as a ratio of real contour length l of the region $\mathcal{R}$ and contour length l_e of an ellipse $\mathcal{E}$ with the same area, elongatedness and orientation:

$$I = \frac{l}{l_e}. \tag{6}$$

The contour length l is obtained by counting of all pairs of neighboring pixels (p_1, p_2), where $p_1 \in \mathcal{R}$ and $p_2 \notin \mathcal{R}$.

Contour length l_e of the ellipse $\mathcal{E}$ is computed as follows:

1. Computation of the ellipse bounding box $(x_{max} - x_{min}) \times (y_{max} - y_{min})$ (Figure 3).

2. Computation of the contour length:

$$l_e = 2(x_{max} - x_{min}) + 2(y_{max} - y_{min}), \tag{7}$$

since there are $2(x_{max} - x_{min})$ horizontal (similarly for vertical) edge segments (edge defined 2 neighboring pixels) between the rightmost x_{max} and leftmost x_{min} point of the ellipse. The multiplier 2 represents upper and lower pair and left and right pair of spans of the ellipse.

Irregularity of a compact region is equal to 1, other region give values larger than 1 (Figure 4 and Tab. 1, column *Irregularity*).

		Elongatedness		Orientation		Irregularity	
Region	Size	theor	real	theor	real	theor	real
1	13085	1.00	1.00	-	-0.80	1.00	1.00
2	3313	1.00	1.00	-	88.94	1.00	1.00
3	861	1.00	1.00	-	-90.00	1.00	1.00
4	225	1.00	1.00	-	-0.00	1.00	1.00
4	225	1.00	1.00	-	-0.00	1.00	1.00
5	445	2.00	1.95	0.0	-0.00	1.00	0.99
6	877	4.00	3.79	0.0	-0.00	1.00	1.00
7	1749	8.00	7.56	0.0	-0.00	1.00	0.99
7	1749	8.00	7.56	0.0	-0.00	1.00	0.99
8	1760	8.00	7.63	22.5	22.44	1.00	1.00
9	1751	8.00	7.70	45.0	45.11	1.00	1.00
10	1774	8.00	7.51	67.5	67.61	1.00	1.00
11	1749	8.00	7.63	90.0	90.00	1.00	0.99
12	1759	8.00	7.71	-67.5	-67.57	1.00	0.99
13	1751	8.00	7.74	-45.0	-44.90	1.00	0.99
14	1774	8.00	7.48	-22.5	-22.41	1.00	1.00
15	423	1.00	1.01	-	-40.97	>1.00	2.35
16	1489	1.00	1.00	-	-1.42	>1.00	2.68
17	5622	1.00	1.00	-	84.47	>1.00	2.77
17	5622	1.00	1.00	90.0	84.47	>1.00	2.77
18	7881	1.41	1.41	90.0	89.98	>1.00	2.78
19	11089	2.00	2.01	90.0	89.89	>1.00	2.79

Table 1. *Theoretical and measured values of the proposed shape descriptors for regions in Figure 4.*

5. ECOLOGICAL INTERPRETATION OF SEGMENTS AND REGIONS

Analyzing landscape structure means the distinction of three major components: the matrix - as the element which covers most of a given region and is regarded to control major ecological processes; the corridors - which are linear elements conducting material and energy fluxes through a region and can serve as pathways for the movement of organisms; and at least the patches - small elements which differ from the surrounding matrix for historical or biogeochemical reasons. In Austrian cultural landscapes, which are heavily influenced by man since ages these three basic element types are usually found in typical combinations, so called “cultural landscape types”. These landscape types can be grouped according to their predominant land-use system and traced throughout the Austrian territory. As an example, cultural landscapes dominated by hay-meadows and pastures with a well-established orthogonal network of dense hedgerows occur at the northern fringe of the Alps and on steep slopes of alpine valleys, whereas cultural landscapes dominated by terraced vineyards intersected by stripes of semi-dry grassland are restricted to the warmest parts in the eastern Loess-region.

Examining a first test region in Eastern Austria the new segmentation method was proved to deliver reasonable results. When comparing the results of the segmentation procedure with the landscape types derived from terrestrial mapping, it turned out that most of the segments could be related to specific clusters of land-units. Many of the classified regions detected by `imagine-ws` were found to be congruent with ecologically defined landscape-types deliminated in the field, but the degree of

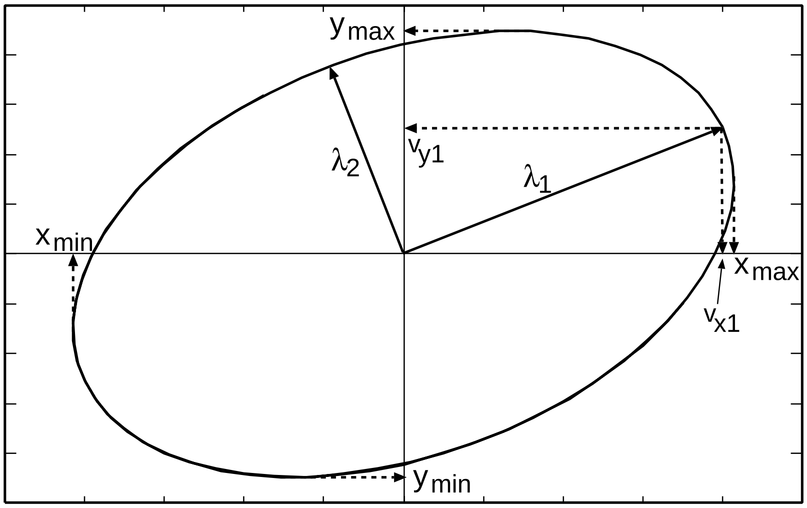


Figure 3. *An ellipse and its bounding box* $(x_{max} - x_{min}) \times (y_{max} - y_{min})$.

correspondence was found to be different according to the predominant land-use system of a given landscape. It was highest in landscapes dominated by large blocks of arable land and lowest in fine-grained vineyards interdigitated with settlements. On the level of singular patches or smaller landscape elements also some differences could be found. Some so-called "environmental resource patches", especially superficial water bodies or semi-dry grassland areas corresponded highly with segments of the satellite image. On the other hand fallow land, certain types of hay-meadows and broad leaved crops like sugar beet appeared to have an uncertain assignment to corresponding segments.

6. IMPLEMENTATION

The proposed segmentation procedure is implemented as a part of Landsat-TM image segmentation and classification package `imagine-ws`, which consists of the following programs:

Binaries `regWshed`, `regMerge` and `regClasSig` aimed at watershed hierarchy segmentation, region merging and region classification. All programs are written in the C language and are based on the image format and image processing routines provided by the image processing package `XITE`.[8] Of course, reading and writing of `Imagine` *img* files is also possible.

EML scripts, which represent a graphical user interface for the `Imagine` system, enabling to define parameters of the binaries.

A common feature of all binaries from the `imagine-ws` package are their enormous memory demands. E.g., the M28 meridian strip of Austria has, with 30 m pixel resolution, approximately 6700×6300 pixels and 300 MB. Since a floating point precision is necessary for many of the processing techniques, and that several copies of the image are simultaneously necessary, memory demands extend to hundreds of megabytes or even gigabytes. This amount of data often exceeds not only the real memory, but often also the virtual memory of a workstation and the program aborts. In the case when the memory demands are lower than the virtual memory of the workstation, but exceed its real one, the slow virtual memory page swapping can degradate throughput of the workstation, resulting into unacceptable processing times of tens of hours or even days.

Therefore, it was necessary to carefully design all implemented algorithms in order to minimize their memory demands. We proposed to process the image matrices in *strips*, with width equal to image width and with height depending on the actual resources of the workstation (defined during the compilation). The whole image matrix is stored in a temporary file buffer and each strip is either loaded from the corresponding buffer before processing or stored after processing. This procedure is, of course similar to virtual page swapping, but since it reflects structure of the algorithm, it is much more effective.

Since most of the operations involved in the package are region oriented, we must process the strips with some overlapping areas, in order to transfer the necessary information from one strip to another. Size of the overlaps depends on the given operation and tests have shown that no artifacts or errors are introduced.

Results obtained for three data sets of small, medium and large size are summarized in Table 2. We can see that the processing time is acceptable even for the largest image.

				Processing time [m:s]	
Image	Dimensions	Bands	Size [MB]	1($\sigma = 1.0$)	3($\sigma = 1.0, 1.41, 2.0$)
small	512×512	6	1.6	0:17	0:52
medium	2335×1688	6	33.6	3:55	13:07
large	6793×6340	7	307.8	48:11	52.0

Table 2. *Test data sets and processing time.*

7. SOME CONCLUSIONS CONCERNING THE USEFULNESS OF THE METHOD FOR LANDSCAPE ECOLOGICAL PURPOSES

As the analysis of landscape structure provides valuable information about the degree of human impact on cultural landscapes, it is necessary to combine the results of field mapping with appropriate remote sensing methods. If one is interested in the spatial patterns of land-use all over Austria, which is a highly diverse country, Landsat TM satellite images are an affordable source of data. To process them was a time-consuming procedure until now, required a very high level of computation and was therefore costly. Especially those working groups in smaller countries dealing with ecological research normally have smaller budgets and had therefore no access to such techniques.

The segmentation and classification package `imagine-ws` was developed to fill this gap and was tested successfully in a national research project. It turned out to perform well in landscapes with clear distinct features, which could be classified as

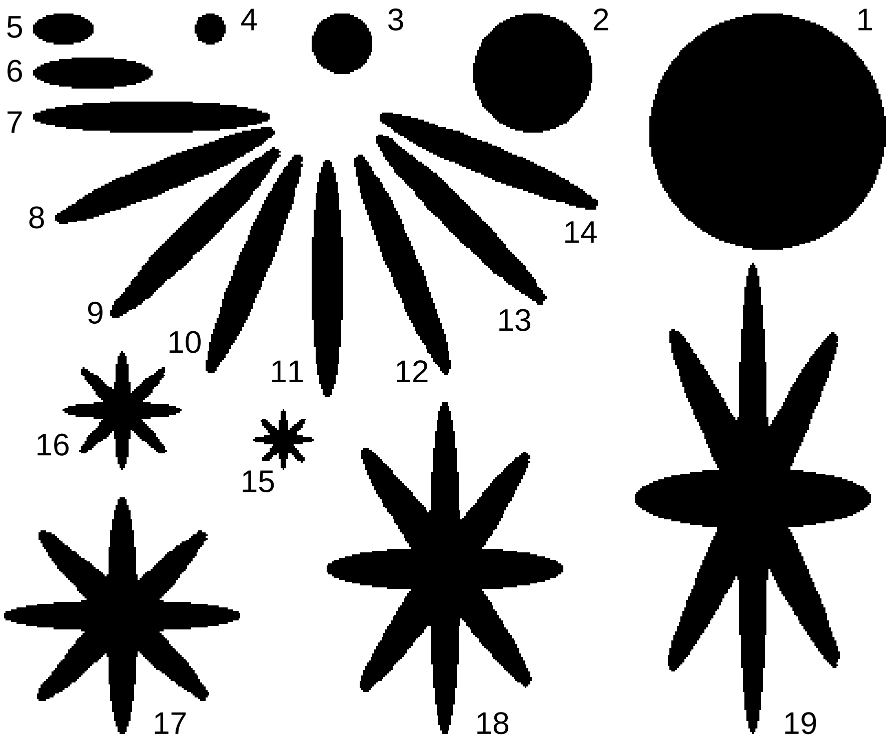


Figure 4. *Test regions.*

so-called checkerboard or scattered-patch landscapes.[9] In such areas, which are widespread in the Eastern Lowlands of Austria, a high degree of correspondence between segmentation and ground-truthing was achieved. The handling of the method in other landscapes with lesser distinct features, like areas with a combination of vineyards and settlements has to be optimized in the future.

REFERENCES

1. R. Forman and M. Godron, *Landscape Ecology*, Wiley & Sons, New York, 1986.
2. E. Odum and M. Turner, “The georgia landscape: a changing resource,” in *Changing Landscapes: An Ecological Perspective*, Z. IS and R. Forman, eds., pp. 137–164, Springer-Verlag, New York, 1990.
3. T. Wrbka, E. Szerencsits, and K. Reiter, “Classification of austrian cultural landscapes — implications for nature conservation and sustainable developement,” in *Sec.Conf.on Culture and Environment*, 1997.
4. J. J. Koenderink, “The structure of images,” *Biolg. Cybern.* **50**, pp. 364–370, 1988.
5. J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, “Uniqueness of the Gaussian kernel for scale-space filtering,” *T-PAMI* **8**(1), pp. 22–33, 1986.
6. J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-8**, pp. 679–698, Nov. 1986.
7. L. Vincent and P. Soille, “Watersheds in digital spaces: An efficient algorithm based on immersion simulation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**, pp. 583–598, June 1991.
8. T. L. nnestad and O. Milwang, “XITE – X-based Image Processing Tools and Environm ent,” in *World Scientific*, H. Christensen and J.L.Crowley, eds., pp. 63–88, ISBN 981-02-1510-X, Singapore, 1994.
9. R. Forman, *Land Mosaics - The Ecology of Landscapes and Regions*, Cambridge Univ.Press, 1995.