POSSIBILITIES OF AUTOMATED GEORELIEF SEGMENTATION FOR THE REQUIREMENTS OF GEOMORPHOLOGIC INFORMATION SYSTEM (GMIS)

*JAN PACINA**

Jan Pacina: Possibilities of automated georelief segmentation for the requirements of Geomorphologic information system (GmIS). Geomorphologia Slovaca et Bohemica, 9, 2009, 2, 15 figs., 16 refs.

Four different algorithms based on searching for discontinuities in fields of derived morphometric characteristics were implemented and tested for automatic delimitation of boundaries of elementary forms of georelief. The best results came from an algorithm based on the Canny edge detector and this was used for further testing. This algorithm produces segments of the boundaries of elementary forms. In this paper strict rules are specified for delimiting the areas of elementary forms from these segments. Application of the Canny-based algorithm should use data smoothed to reduce the density of edges detected. Ways for creating suitable surfaces are compared here.

Key words: GmIS, elementary forms, automatic delimitation of elementary form boundaries, georelief segmentation, smoothing

INTRODUCTION

A Geomorphologic Information System (GmIS) is a specific type of information system concerned with collecting, managing and analysing geomorphologic information. It provides an excellent tool for applied geomorphologic analysis (MENTLÍK et al. 2006) .

A GmIS is being developed in cooperation between the Department of Physical Geography and Geoecology, Faculty of Natural Sciences at Comenius University in Bratislava, Slovakia, the Department of Mathematics – Geomatics Section, Faculty of Applied Sciences and Department of Geography, Pedagogical Faculty at the University of West Bohemia in Pilsen, Czech Republic and the Department of Geoinformatics, Faculty of the Environment at the J. E. Purkyně University, Czech Republic.

The concept, database model and possibilities of geomorphologic analysis are described in JEDLIČKA and JEŽEK (2008), JEDLIČKA (2008), MENTLÍK et al. (2006) and MINÁR et al. (2005). The following text is concerned with the possibilities of automated georelief segmentation.

ELEMENTARY FORMS AND THEIR BOUNDARIES

In accordance with MINÁR and EVANS (2008) we can define three axioms which create the theoretical background for georelief segmentation.

1. The georelief may be considered as a continuous surface – the geometrical field of altitudes.
2. At a given scale it is possible to find discontinuities on the Earth's surface, which may be considered as natural boundaries of geomorphologic objects.
3. These discontinuities (and other characteristics of the Earth's surface) are results of geomorphologic processes, which depend or are influenced by gravitation.

Specific structural elements of the field create the natural base for its segmentation. Such elements may be called *singular lines* and *points*. Here we can encounter for example *extreme* points and lines (peaks, pits, saddle points and ridges), inflection points and discontinuities of the altitude field and other derived fields. (LASTOČKIN 1987).

Georelief may be viewed as consisting of segments (elementary forms), which are characterized by different types and levels of homogeneity. This is expressed ideally by a constant value of altitude or of derived morphometric characteristics. Discontinuity of these characteristics marks logical boundaries of the segments. Then we can define an ideal *elementary form* as *an element of the georelief with a constant value of altitude, or of two or more well interpretable morphometric characteristics, bound by lines of discontinuity.*

The representation of the whole Earth's surface with the help of elements with constant value of selected morphometric characteristic

* Fakulta životního prostředí, UJEP, Katedra informatiky a geoinformatiky, Králova Výšina 3132/7, 400 96 Ústí nad Labem, Czech Republic, e-mail: jan.pacina@ujep.cz

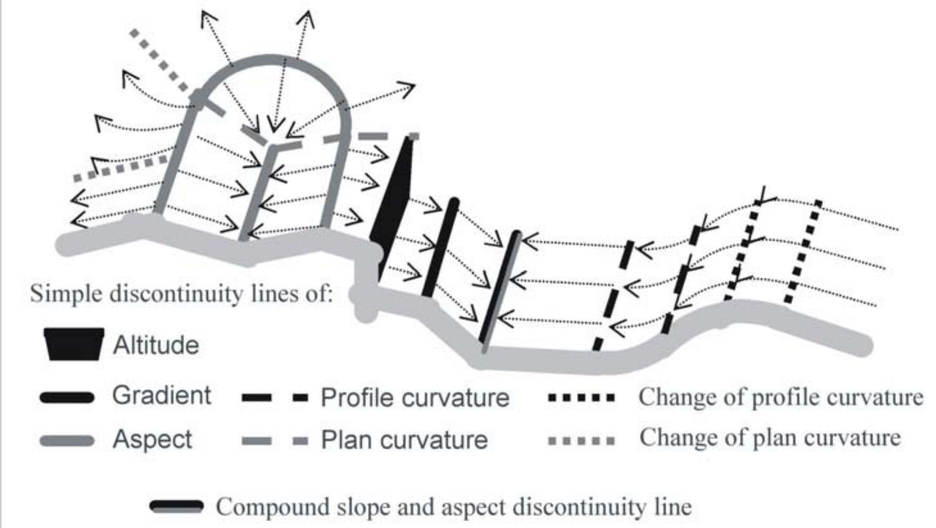

Fig. 1 Basic types of discontinuity lines (MINÁR and EVANS 2008)

has a strict theoretical consequence for delimitation of elementary forms. In a geometrically ideal case, the boundary of two elementary forms defined by different values of a morphometric characteristic must be a *line of discontinuity* - a sudden discrete change of value of some morphometric characteristic (**Fig. 1**).

For automatic delimitation of boundaries of elementary forms, we have to search for lines of discontinuity. We will use the following rule: a point of discontinuity of a derived surface of the first order (slope, aspect) matches a maximal value in the derived surface of second order (change of slope = curvature) and this matches the zero value in a derived surface of third order (change of curvature). That means, for example, that for discontinuities of the altitude field, we will search for local extremes in the field of slope.

Four different algorithms for automatic delimitation of boundaries of elementary forms were implemented and tested. Algorithms were

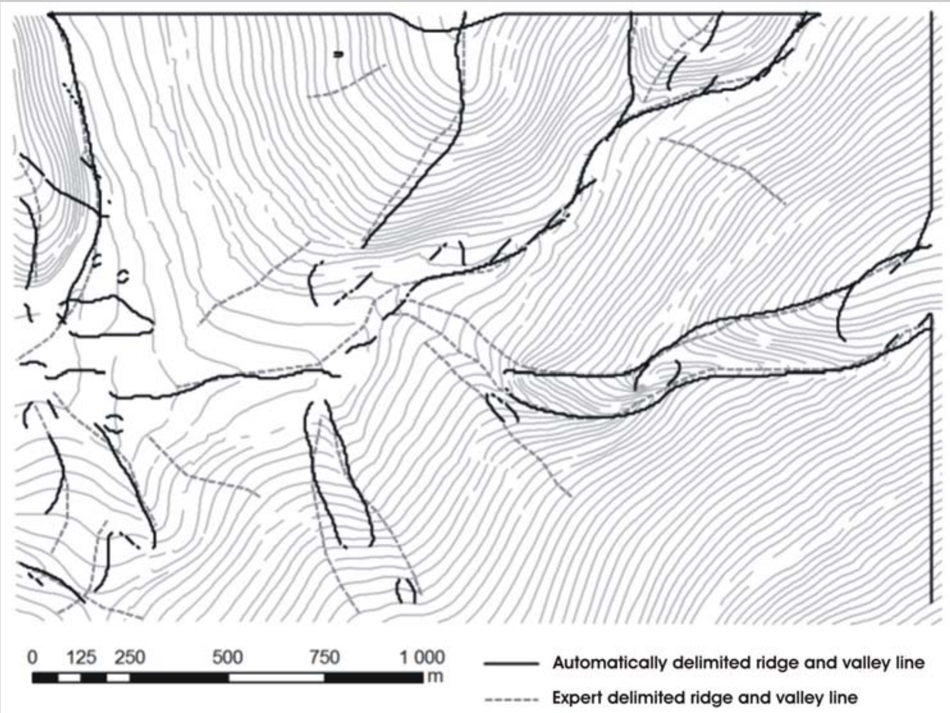

Fig. 2 Boundaries delimited by searching for the local maximum from the close neighbourhood

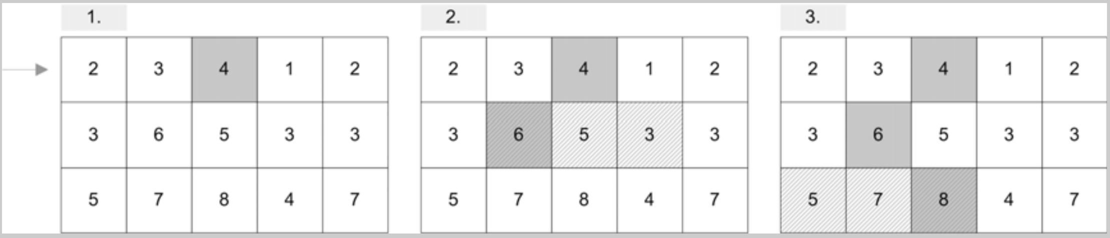

Fig. 3 Process of the floating mask algorithm

tested in the area of Šumava National Park (southern part of the Czech Republic). For easy evaluation of the results the four algorithms were applied initially to the altitude field. Such delimited lines will correspond to ridge and valley lines (maximum = ridge, minimum = valley line). According to the previous text, all algorithms are based on the principle of searching for local extremes (≈ elementary form boundaries) in the input data.

SEARCH FOR THE LOCAL MAXIMUM FROM THE CLOSE NEIGHBOURHOOD

The first and easiest method for finding local maxima is to compare the values in some *n*-neighbourhood. We consider the output grid as a matrix of dimension [*m,n*]. The algorithm goes through the matrix in rows and the actual cell $a_{i,j}$ is determined to be a local maximum if $a_{i,j} > a_{i,j-1}$ The same algorithm is than applied to columns with the condition $a_{i,j} > a_{i-1,j.}$ The algorithm as described above searches in the input data only for ridges. Valley lines would be delimited identically, but using opposite signs in the conditions described. **Figure 2** shows the ridge and valley lines delimited by this algorithm. Visual comparison shows that the algorithm found ridge and valley lines quite similar to those delimited by an expert geomorphologist.

SEARCH FOR THE LOCAL MAXIMUM WITH FLOATING MASK

The algorithm searches the input data in rows and searches for local maxima with the same principle as in the previous method. When the local maximum is found (part 1 of **Figure 3** - value 4) 3 adjacent cells in the next row are selected. In those 3 cells we again selected the maximum (part 2 of **Figure 3** – value 6). The process goes on until the conditions written in the source code of the algorithm are reached. On **Figure 4** are results of this algorithm. The delimited lines are exces-

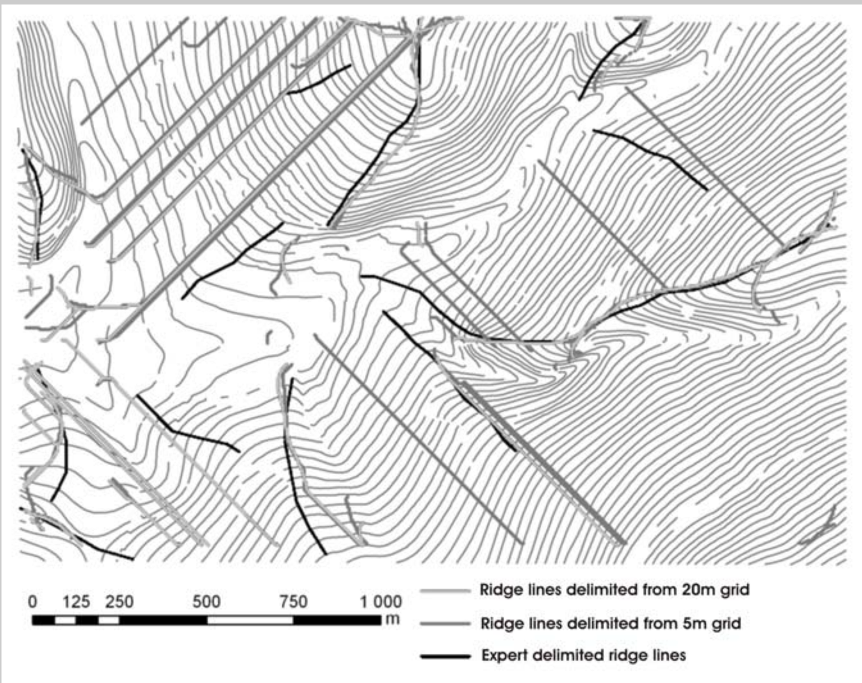

Fig. 4 Boundaries delimited by searching for the local maximum with floating mask

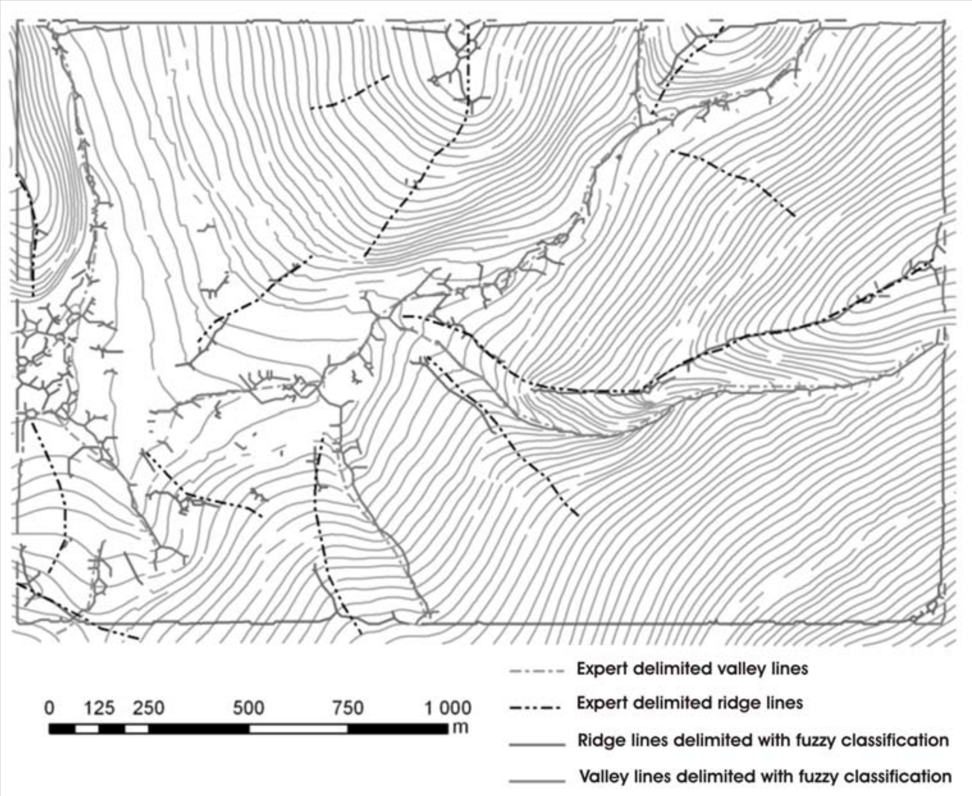

Fig. 5 Boundaries delimited by searching for the local maximum with the help of fuzzy classification

sively straight and parallel, have the character of slope lines, and differ a lot from expert delimited ridge lines.

SEARCH FOR THE LOCAL MAXIMUM WITH THE HELP OF FUZZY CLASSIFICATION

Fuzzy classification in GIS is attractive because it offers new possibilities to quantify spatial variability of categorical objects and analyze spatial confusion between the classes. In the case of fuzzy k-means classification (HENGL et al. 2003), each pre-ordained class is mapped separately as a single membership ranging from 0 (no membership to the class) to 1 (full membership).

For each cell of the input raster *slope*, K_r *(curvature of the contour lines)* and *ACV* (*Anisotropic Coefficient of Variation*) are calculated. From these results membership of each cell in each class is then computed. **Fi-**

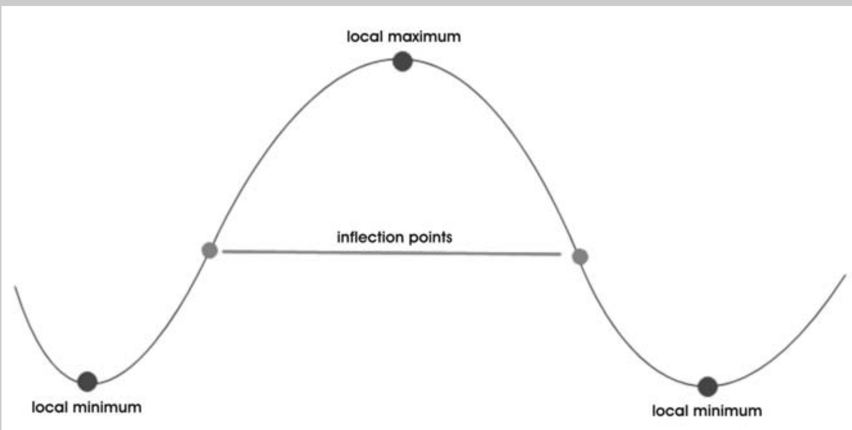

Fig. 6 Principle of the algorithm based on *Canny* edge detector

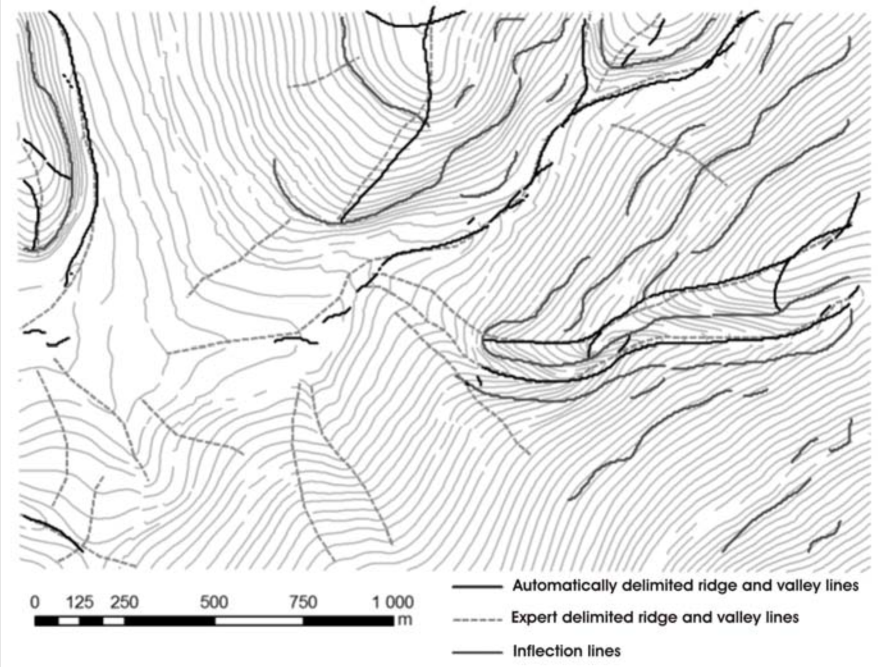

Fig. 7 Boundaries delimited by searching for the local maximum by algorithm based on Canny edge detector

gure 5 shows the result of fuzzy classification after thinning. This method is applicable only to the field of altitudes, because for derived fields of higher order stability problems affect the partial derivatives approximation.

SEARCH FOR THE LOCAL MAXIMUM WITH THE USE OF EDGE DETECTORS

In a digital image there are places with a sudden change of the brightness. Corresponding pixels are the *edges*. For automated delimitation of *edges* is used the image pre-processing method called edge detecting. In this paper we have tried to apply this methodology of digital image pre-processing to delimitation of elementary form boundaries. The *edges* delimited by the edge-detectors (in this case) applied on elevation data, correspond to inflection points in the field of altitudes. This is further used for selection of local extremes – see **Figure 6**. It appears that edge detectors are a very powerful method to solve this problem.

From the whole family of edge detectors was chosen the *Canny* edge detector. This edge detector is suitable[1] for processing the input data used for delimitation of elementary form boundaries[2]. The usage of the *Canny* edge detector is followed by application of noise elimi-

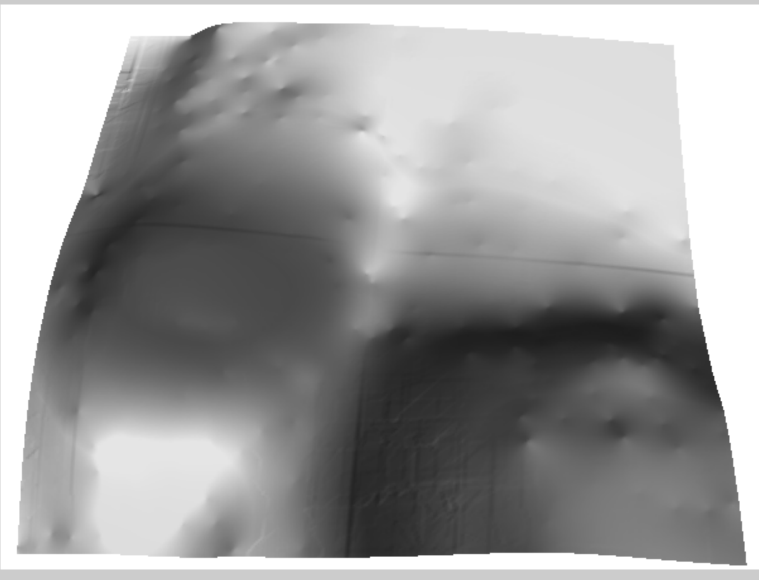

Fig. 8 Interpolation artifacts visible in the input data

nating filters in the input data and thresholding the gradient to eliminate weak edges. The whole process of the *Canny* edge detector is described for example in GREEN (2007).

In the first step of delimitation of the boundaries the *Canny* edge detector is applied to the input data. As written in the previous text, this detector searches for edges corresponding to inflection points in the input data.

In the next step local maxima and minima between these inflection points are sought (**Fig. 6**). During this computation it is easy to compute the importance and quality of the edge[3]. This quality is further used to eliminate the edges which are results of errors in input data, computational errors or which are not important for delimitation of elementary form boundaries. This means that only strong edges of high-importance are kept for further use.

On **Figure 7** are the results of the algorithm based on the *Canny* edge detector. With the dark gray colour are marked the inflection points delimited by the edge detector. Black marked lines are the ridge and valley lines delimited by the method shown on **Figure 6**.

INPUT DATA PREPROCESSING

The test area for automatic boundary delimitation was in the surroundings of Slovinec hill, in the Devínske Karpaty Mts. (part of the Malé Karpaty Mts.), Slovakia. This is the area where MINÁR and EVANS (2008) delimited elementary forms of georelief, so it will be possible to compare the results.

Input data were an elevation raster with 1m resolution, from MINÁR and EVANS (2008). This original raster was acquired by digitizing a 1:10 000 contour map supplemented by important terrain characteristics and then interpolated by *kriging*.

On **Figure 8** various interpolation artifacts (visible lines of segmentation, peaks …) are visible. This may be caused by errors in the input data or by improperly set interpolation parameters. Tests have shown that for delimitation of boundaries of elementary forms it is suitable to use a smoothed (generalized) digital elevation model, which keeps the main terrain characteristics. Terrain smoothing is required for the *Canny* edge detector and its thresholding. The *Canny* edge detector computes thresholds from the edges found (inflection points in our case)[4]. If we keep too many edges in the input data (artificial or natural) the *Canny* edge detector may evaluate the data in not desirable way. The unwanted artifacts can then be removed by lowering the density of input data and re-interpolating.

For interpolation, a non-commercial GIS GRASS v.6.3 was chosen. Among other interpolation algorithms, it offers RST interpolation (RST = Regularized Spline with Tension). The result of CHAPLOT et al. (2006) and PACINA (2008) is that RST interpolation is the most suitable for this type of input data.

The theory for RST computation is described for example in CEBECAUER et al. (2002) or NETELER and MITÁŠOVÁ (2004).

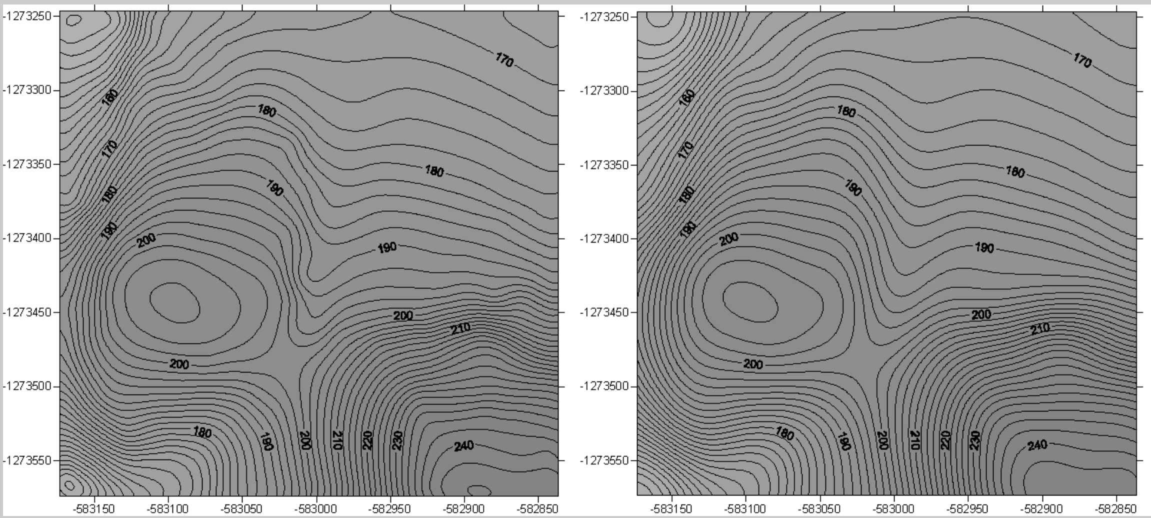

Fig. 9 Elevation surface interpolated from random points and with the use of dmin parameter

---

[1]More about Canny edge detector characteristics in HLAVÁČ and SEDLÁČEK (2005)
[2]More about input data pre-processing is later in the text
[3]See more in MINÁR and EVANS (2008), and PACINA (2008)
[4]See GREEN (2007) for the description of the *Canny* edge detector algorithm

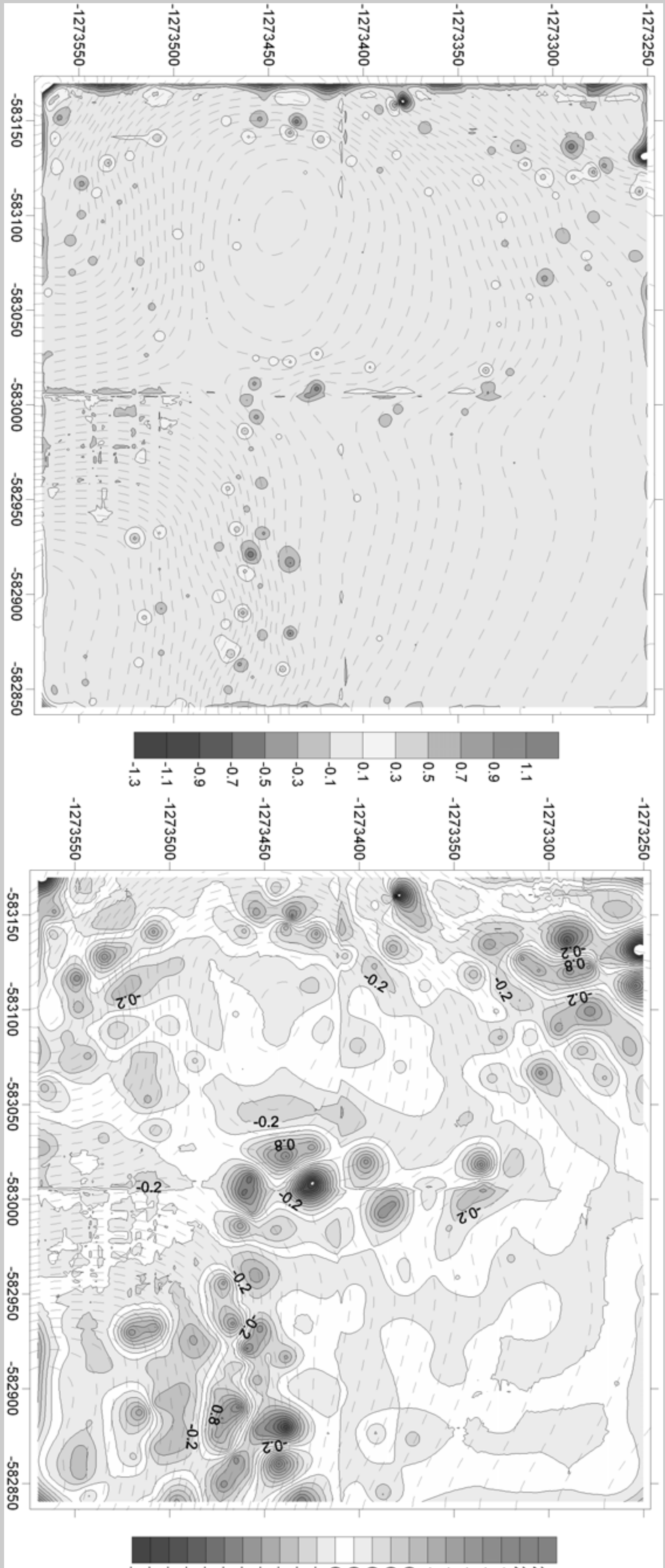

Fig. 10 Differences from the input data - elevation surface interpolated from random points and with dmin parameter

Theory for digital elevation model computation is given for example in BONK (2003) or PACINA (2005).

RST interpolation is driven mainly by the parameters *tension* and *smooth*. The *tension* parameter sets the toughness of the interpolated surface, between a thin steel plate and a rubber membrane. For noisy data the *smooth* parameter can be used. With *smooth* set to zero the interpolated surface passes exactly through the input data.

A *smoothed elevation model* can be obtained by one of the following methods:

1. From the input dataset select 3 – 10 % of the points at random and then re-interpolate from these random points. This method was presented in BONK (2003).
2. Dilution of input points during the re-interpolation by use of the *dmin* parameter. Within RST interpolation the *dmin* parameter can be set as the distance between points used for interpolation. Points which are closer than *dmin* are excluded from the interpolation.

In both cases the input data were interpolated with parameters *tension* = 20, *smooth* = 0.8, and for the second method *dmin* = 5 (distance between points is 5 m).

Figure 10 shows the distributions of differences between the input raster and both of the interpolated elevation rasters presented in **Figure 9**. When using the first method for data dilution we will obtain a different result every time the method is used. While using the *dmin* parameter for data dilution we will get an identical result every time – points generated in such a way create a regularly spaced grid. With use of the *dmin* parameter the differences are higher than when using the random point method (**Figure 10**). This gives evidence that the surface interpolated with *dmin* parameter is more suitable for application of the boundary delimitation algorithm based on *Canny* edge detector.

BOUNDARY DELIMITATION

Search for discontinuities in the field of derived morphometric characteristics is based on the following principle: a discontinuity in the field of order *n* equals an extreme value in the field of order *n+1*. For example - an ideal discontinuity in the field of altitudes corresponds to 90° in the field of slope. Similarly this can be applied to all other morphometric characteristics - for example discontinuity in slope = maximum curvature (where the ideal discontinuity corresponds to infinite curvature).

Boundaries were delimited on the following surfaces of derived morphometric characteristics:

- g_t – gradient change in the direction of contours,
- a_g – aspect change in the direction of slope (the normal change of gradient curvature),
- A_{Nt} – contour line curvature K_r,
- a_{gn} – change of the aspect change in the direction of slope ().

From the four implemented algorithms for automated boundary delimitation two were chosen for further testing. The algorithm based on *floating mask* was not producing sufficient results and because of approximation errors in partial derivatives the algorithm based on *fuzzy classification* can be applied only on the field of altitudes.

To evaluate the suitability of implemented algorithms for our purpose were set the following formal criteria:

1. delimited boundaries are continuous,
2. topological relationships between diffe-rent orders of delimited lines are expected,
3. existence of noise is minimal,
4. parallel or bifurcated lines are minimized,
5. spatial pattern of lines is suitable (tendency to enclose areas).

Algorithm based on *search for the local maximum from the close neighbourhood* was producing good results in the initial testing; even the algorithm itself is very simple. On **Figure 11** are results of this algorithm applied on the field of a_g. During further testing the cell-neighbourhood searched for local extremes was enlarged to 3x3 and data rotation was implemented to find edges (local extremes) in directions other than those of the *x* and *y* axes.

EVALUATION OF THE FORMAL CRITERIA:

1. Delimited boundaries (in relation to the noise amount) can not be considered to be continuous (**Fig. 11** and **Fig. 12**),
2. There are topological relations between lines delimited in the fields a_g and a_{gn},
3. Boundaries delimited in the field a_g contain much noise (**Fig. 11**). Boundaries in the field a_{gn} contain even more noise (**Fig. 12**),
4. In the resulting data are a lot of bifurcated lines,

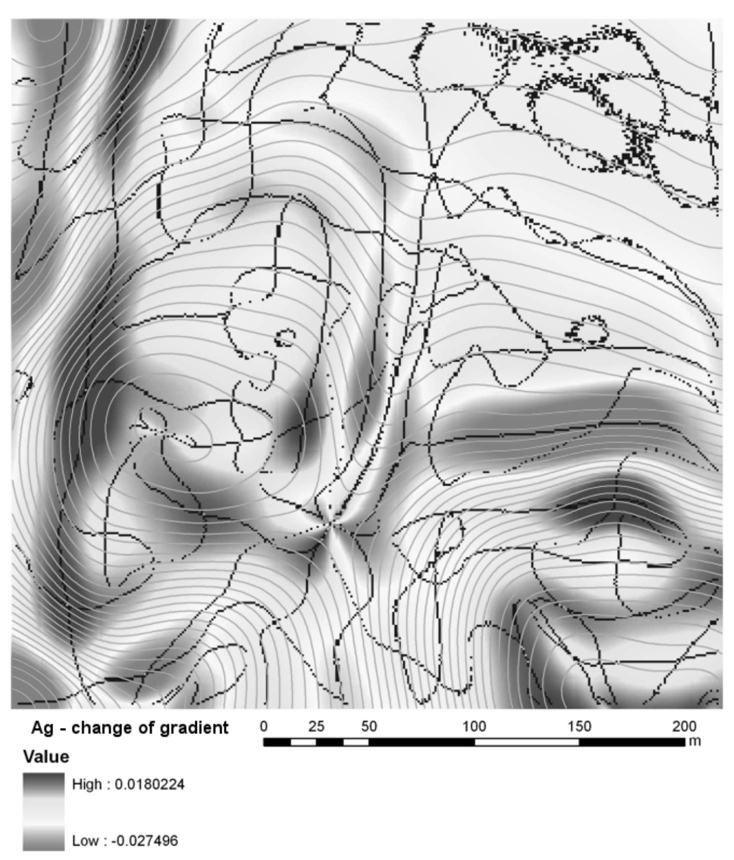

Fig. 11 Boundaries delimited in the field of a_g

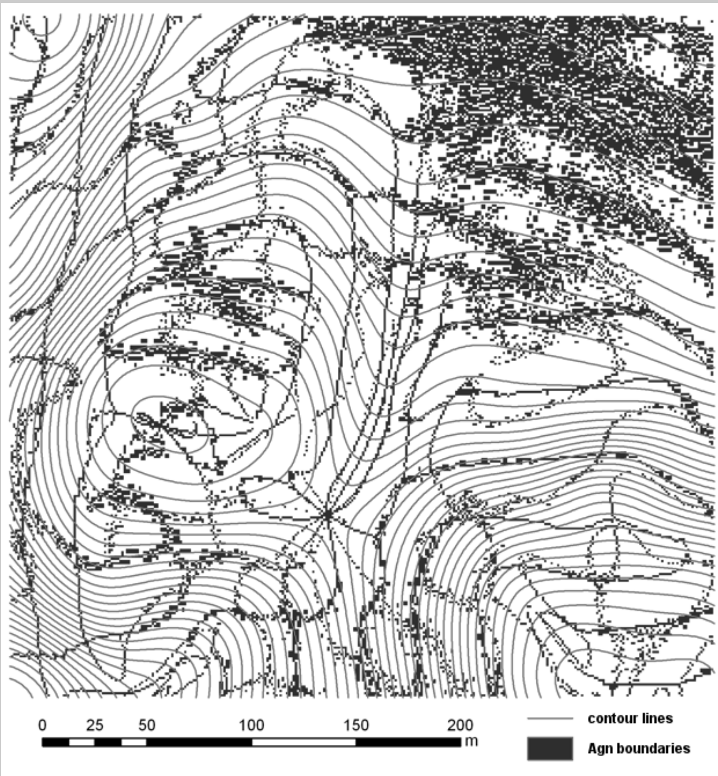

Fig. 12 Delimited boundaries and noise in the field of a_{gn}

5. Delimited boundaries tend to enclose areas.

From this evaluation of the formal criteria we may consider this algorithm unsuitable for further work.

For automatic delimitation of boundaries of elementary forms the algorithm based on the *Canny* edge detector was chosen. This algorithm fulfils the formal criteria introduced in the beginning of this section (**Fig. 13**).

On **Figure 13** is an example of automatically delimited discontinuities[5] in the field of a_g morphometric characteristic. The gray lines are the inflection points (lines) delimited by the *Canny* edge detector. The black lines are the local extremes (≈ lines of discontinuity) delimited in-between the inflection points. **Fig. 14** shows automatically evaluated quality and geomorphologic importance of delimited boundaries – dark colour = the highest importance, light colour the lowest. These lines of discontinuity correspond to segments of *protoform*[6] boundaries.

From the automatically identified segments of *protoform* boundaries the areas of *protoforms* should be delimited. Delimitation of these areas is not automated yet, so the areas have to be delimited with the help of an operator. The following rules are set for the delimitation of boundaries:

– we handle the fields of morphometric characteristics in hierarchical order – we consider that higher order ≈ higher importance and quality,

– for delimitation of boundaries in the field of a morphometric characteristic, we follow the selection criteria:

- quality of delimited boundary (in the meaning of **Fig. 14**),
- continuity of the line,
- length of the line,
- spatial consistency,

– even if part of the line is of lower quality, we use the entire line,

– we set the limit distance between lines – potential boundaries delimited in the field of morphometrical characteristics of different order tend to be parallel in some areas. This

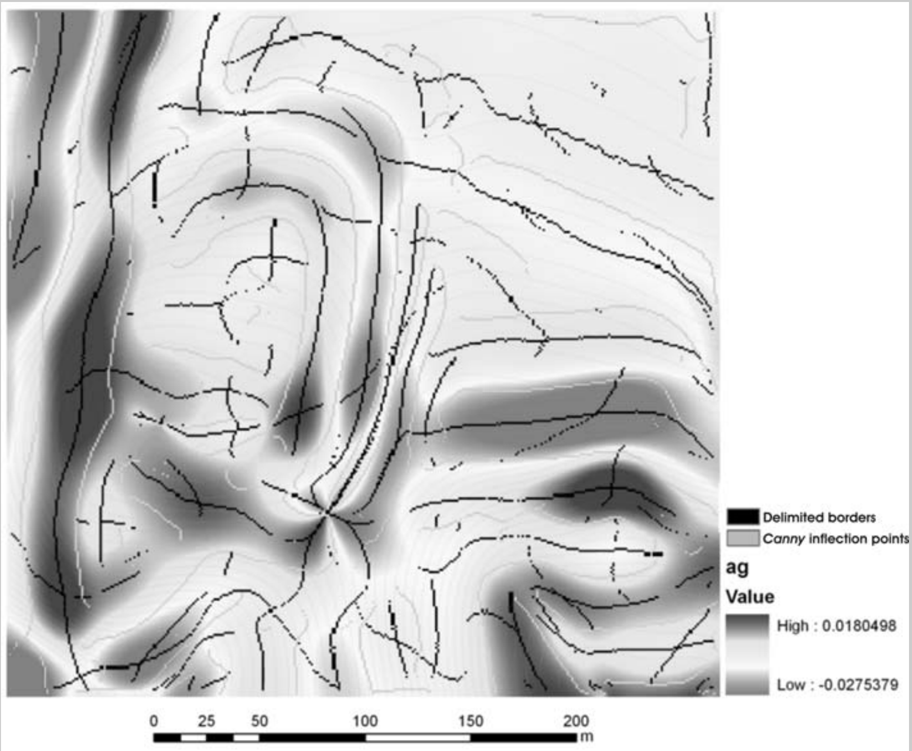

Fig. 13 Example of boundaries delimited on the surface of a_g and inflection points delimited by *Canny* edge detector

[5]Discontinuities in the meaning of this paper - discontinuity in the field of order *n* is the local maximum in the field of order *n+1*

[6]A *protoform* is an elementary form, which has not been specified yet

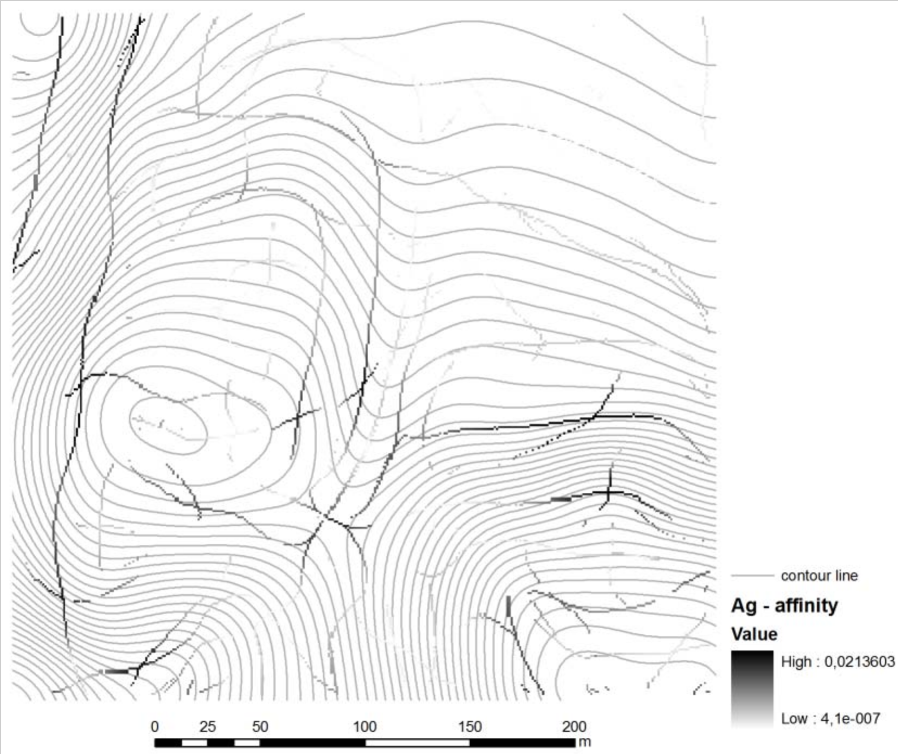

Fig. 14 Example of boundaries delimited on the surface of a_g with computed quality

rule is used to select the line to be a boundary or not,

- resulting boundaries in surroundings of singular points are unreliable,
- the area delimited by the boundaries has inner homogeneity (= the area is enclosed by the lines).

On **Figure 15** we can compare the results of elementary forms delimited by this semi-automated method with results from MINÁR and EVANS (2008) where elementary forms were delimited from the field of altitudes and fields of derived morphometrical characteristics of higher order based on the experience of the operator. For computation of partial derivatives common GIS methods were used, which led to approximation errors in the resulting derivatives.

The semi-automated method presented in this paper uses a robust method for approximation of partial derivatives of the 3rd order for the computation of the higher order morphometrical characteristic fields. This algorithm approximates the partial derivatives from 5x5 neighbourhood by a general polynomial of the third order, using the weighted least square method. The method is fully described in PACINA (2008) or PACINA (2009).

The automatically delimited boundary segments are partially similar to the boundaries delimited by the expert. The differences from the boundaries delimited in MINÁR and E-VANS (2008) are not errors, but the algorithm was able to detect the local extremes (boundaries) in the fields of derived morphometrical characteristics of higher order, which the expert was unable to recognize. The other reason may be the better quality of the input data which are not influenced by the derivative approximation errors.

The automatically delimited boundary segments led the operator to detailed elementarization. The amount of input information led to a model with higher detail. In contrast to expert delimited boundaries, we can see delimited valley bottoms and ridge areas as *proto-forms*.

CONCLUSIONS

This article has shown a possible way for automatic georelief segmentation for the requirements of GmIS. The automatic delimitation of elementary form boundaries should involve a data preprocessing to obtain the *smoothed surface* which is a generalized digital elevation model, which keeps the main terrain characteristics. An algorithm based on the

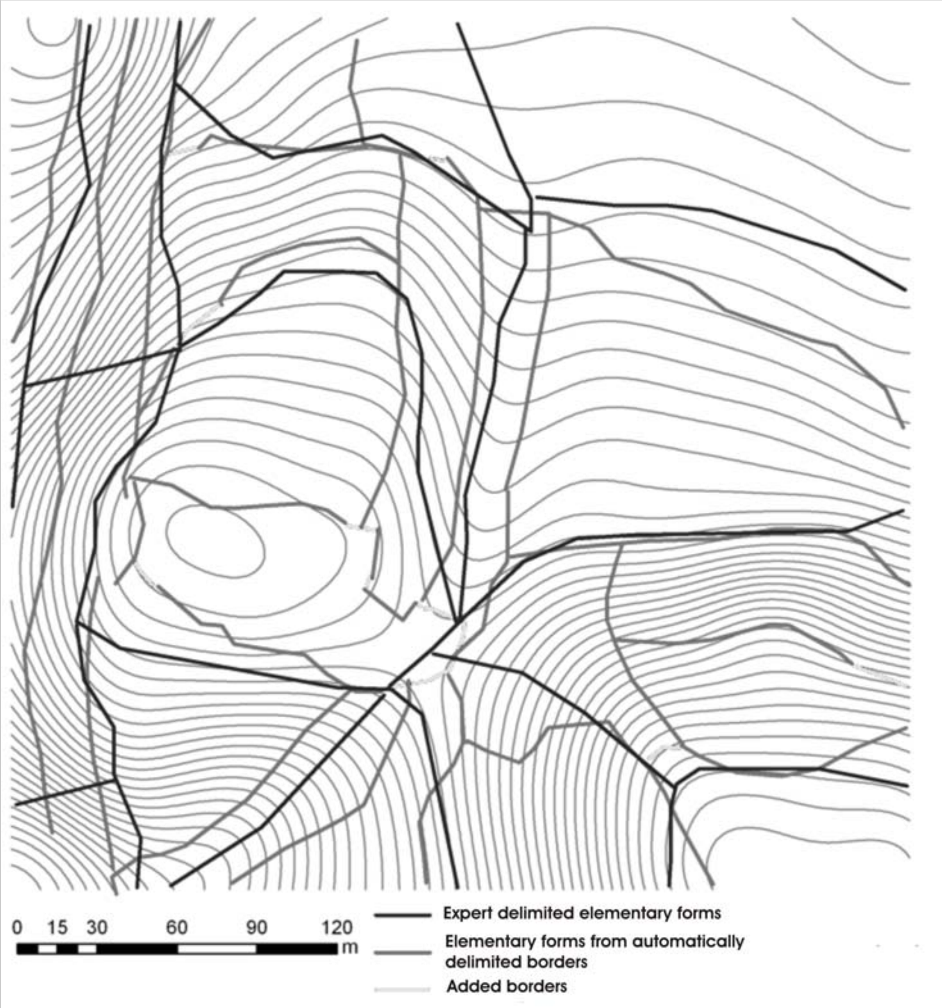

Fig. 15 Protoforms delimited by an expert and from automatically delimited boundaries

Canny edge detector delimits segments of form boundaries with high precision. It searches for inflection points in the input data and then finds the lines of discontinuity in-between them. The boundary segments delimited correspond with the boundaries of elementary forms of georelief, and they are even more precise than the geomorphologist-delimited boundaries in MINÁR and EVANS (2008). The next task would be to automate the delimitation of areas of elementary forms by the rules described above.

REFERENCES

BONK, R. (2003). *Scale-dependent impact of selected factors on morphometric parameters accuracy and automated geomorphological mapping.* Dissertation thesis, Faculty of Natural Sciences, Comenius University in Bratislava.

CEBECAUER, T., HOFIERKA, T., ŠÚRI, M. (2002). Processing digital terrain models by regularized spline with tension: tuning interpolation parameters for different input datasets. *Proceedings of the Open source GIS - GRASS users conference (Trento, Italy, 11 – 13 September 2002).*

CHAPLOT, V., DARBOUX, F., BOURENNANE, H. (2006). Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landorm types and data density. *Geomorphology*, 77, 1 – 2, 126 – 141.

GREEN, B. (2007). *Canny Edge Detection Tutorial*.<http://www.pages.drexel.edu/~weg22/can_tut.html> On-line [November 2nd 2009].

HENGL, T., GRUBER, S., SHRESTHA, D. P. (2003). *Digital terrain analysis in ILWIS*. <http://www.itc.nl/library/Papers_2003/misca/

hengl_digital.pdf> On-line [November 2nd 2009].

HLAVÁČ, V., SEDLÁČEK, M. (2005). *Zpracování signálů a obrazů*. ČVUT, Praha.

JEDLIČKA, K. (2008). Geomorphologic Information system - Use Cases. *Sborník - GIS Ostrava 2008 (Ostrava, Česká republika, 27-30 January 2008)*. Tanger, Ostrava. <http://gis.vsb.cz/GIS_Ostrava/GIS_Ova_2008/sbornik/index.htm> On-line [November 2nd 2009].

JEDLIČKA, K., JEŽEK, J. (2008). Geo-related open source software development at University of West Bohemia. In Bandrova, T. L. *Proceedings 1. - 2nd International Conference on Cartography and GIS (Borovets, Bulgaria, 21-24 January 2008)*, 125 – 131.

LASTOČKIN, A. N. (1987). *Morfodinamičeskij Analiz*. Nedra, Leningrad.

MENTLÍK, P., JEDLIČKA, K., MINÁR, J., BARKA, I. (2006). Geomorphological information system: physical model and options of geomorphological analysis. *Geografie – Sborník České geografické společnosti*, 111, 1, 15 – 32.

MINÁR, J., MENTLÍK, P., JEDLIČKA, K., BARKA, I. (2005). Geomorphological information system: idea and options for practical implementation. *Geografický časopis*, 57, 3, 247 – 266.

MINÁR, J., EVANS, I. S. (2008). Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping. *Geomorphology*, 95, 3 – 4, 236 – 259.

NETELER, M., MITÁŠOVÁ, H. (2004). *Open Source GIS: a GRASS GIS approach (2nd Edition)*. Kluwer Academic Publishers/Springer, Boston, USA, 420 p.

PACINA, J. (2005). *Popis principů fungování vybraných prostorových analýz a jejich implementace v nekomerčním GIS GRASS pro názornou výuku*. Master thesis, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, Plzeň, Česká republika.

PACINA, J. (2008). *Methods for automatical delimitation of elementary forms of georelief as a part of Geomorphologic information system*. PhD thesis, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, Plzeň, Česká republika.

PACINA, J. (2009). Advanced methods for georelief segmentation. *Sborník - GIS Ostrava 2009 (Ostrava, Česká republika, 25-28 January 2009)*. Tanger, Ostrava. <http://gis.vsb.cz/GIS_Ostrava/GIS_Ova_2009/sbornik/index.htm> On-line [November 2nd 2009].