Article

Canscan — An Algorithm for Automatic Extraction of Canyons

Nebojsa Balic * and Barbara Koch

Institute for Remote Sensing and Landscape Information Systems, Tennenbacherstr. 4, 79106, Freiburg, Germany; E-Mail: barbara.koch@felis.uni-freiburg.de

* Author to whom correspondence should be addressed; E-Mail: nebojsa.balic@felis.uni-freiburg.de; Tel.: +49-761-203-8644; Fax: +49-761-203-3701

Abstract: This article introduces a novel algorithm for automatic extraction of canyons which was developed in the Department of the Remote Sensing and Land Information Systems at the University of Freiburg (FELIS). The algorithm detects canyons by means of user-defined dimension parameters and elevation information provided in a Digital Terrain Model (DTM). The extraction procedure is based on the geometric interpretation of canyons through which the input dimension parameters are identified. The dimension parameters are used for identifying cross-sections across DTM on the basis of which canyons are extracted. In addition to the detailed description of the extraction algorithm, this paper includes the results obtained in test regions as well as a thorough discussion.

Keywords: automatic extraction; canyons; DTM; LIDAR

1. Introduction

The automatic mapping of canyons and valleys is vitally important in a wide range of environmental applications. In case presented here, mapping has been conducted for the purpose of nature conservation, focusing on the forest complexes growing along steep cliffs of canyons. The terrestrial mapping of these forests would have been time consuming and limited both by the cost constraints and inaccessibility of canyon areas. Moreover, the analysis of aerial photos would yield no better results, being error-prone and time demanding.

In order to detect canyon areas, the algorithm for their delineation has been developed based on the user-defined valued of identified dimension parameters. Digital Terrain Models (DTM) interpolated

from LIDAR data have been chosen as the most appropriate data basis for obtaining detailed terrain information even under vegetation cover.

Automatic landform extraction through image segmentation gains increasing importance [1]. The common approach is to aggregate pixels in Digital Elevation Models (DEM) according to per-pixel parameters (slope, aspect, curvature etc.) calculated from the DEM [2]. On the basis of these parameters terrain can be classified into homogenous units representing landforms. However, in contrast to regions in DEM, landforms in reality do not exhibit sharp boundaries but are separated with areas in which they mutually overlap. In order to mitigate the problem of vague boundaries between landforms, the fuzzy classification of landforms is introduced [3]. As an alternative to these methods, the proposed method aims at detecting the landform utilizing a user-defined threshold for each of the identified dimension parameters. The target landform is the canyon, a landform subjected to the quantitative analysis [4] and classification efforts [5], but never, at least to the authors' knowledge, subjected to automatic extraction techniques.

2. Materials

2.1. Study Area and Datasets

The LIDAR data used for testing the algorithm has been provided by the Bavarian State Institute of Forestry. The data have been acquired in May 2006 by using OPTECH ALTM mounted on a Pilatus Porter airplane flying at an average altitude of 1,000 m above the ground. The swath width was 810 m while the sidelap amounted to 360 m. The sensor system acquires 71,000 laser pulses per second within the scan angle of $\pm$ 22° relative to the nadir axis.

The study area, "Wertachbruch", is situated in southern Germany. It is characterized by steep canyons inhabited by rare forest trees and bird species. Two test sites were selected in the northern and southern part of the study area, spanning about 7.7 and 2.5 km^2, respectively.

The DTMs have been generated by applying the "Active Surface Algorithm" for the interpolation of first and last LIDAR implemented in the TreesVis software [6]. The first and last pulse returns with the average point spacing of 2.6 points per m^2 were used for the generation of two DTMs with 1 m grid and 0.01 m vertical resolution. In order to reduce computation time, this original resolution has been coarsened into 16 m.

2.2. Implementation of the Algorithm

The algorithm has been implemented as a module in the TreesVis software developed in the FELIS department [6] by using entirely commercial solution such as the machine vision software HALCON (http://www.mvtec.com/halcon) and Microsoft's Visual Studio as the development environment. This implementation has been coded in C++ programming language.

The same extraction method is deployed as a java-based web application compliant with the WebProcessingService (WPS) Specification [7] adopted by Open Geospatial Consortium (OGC). This application utilized entirely open source solutions ImageJ (http://rsbweb.nih.gov/ij/) for image processing tasks and Eclipse (http://www.eclipse.org/) as a development platform.

3. Methods

Four phases can be distinguished in the process of developing the algorithm for the automatic extraction of canyons. These are:
- Identification of input parameters
- Recognition of cross-sections of the canyon
- Merging pixels in DEM representing these cross-sections into regions
- Length assessment of the regions composed of recognized cross-sections.

3.1. Identification of Input Parameters

The algorithm is based on the assumption that a canyon is a landform surrounded with steep slopes. As the result of geometric interpretation the following dimension parameters have been identified (Figure 1):
- minimal inclination of slopes α_{min}
- maximal width (the horizontal length of the longest cross-section along the canyon) W_{max}
- minimal width (the horizontal length of the shortest cross-section along the canyon) W_{min}
- minimal depth (the height of the shallowest cross-section along the canyon) H_{min}
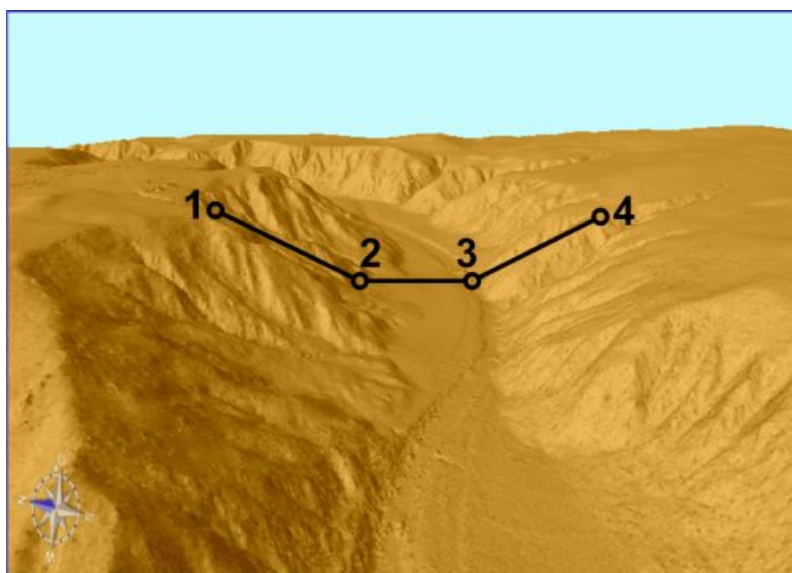- minimal length L_{min}

Figure 1. A characteristic cross-section of a canyon defined by the slope inclination α, depth H and width W. The white line is centreline of the canyon along which its length (L) is measured.

3.2. Recognition of the Cross-Sections

The characteristic cross-sections of the canyon have been identified on the basis of their four characteristic points (Figure 2), two of which are endpoints (points 1 and 4) lying on upper parts and two more points (points 2 and 3) lying in the lower section of each opposite slope.

Figure 2. A cross-section of a canyon with its four characteristic points.

Geometric relationships between four characteristic points can be derived by means of the input parameters (Figure 3). The maximal horizontal difference d_h between endpoints (1 and 4) is defined by the parameter of the maximal width W_{max}:

$$d_h(1,4) \leq W_{max} \qquad (1)$$

The minimal elevational difference between two characteristic points placed on the same slope (points 1 & 2 and points 3 & 4) is defined by the parameter of minimal depth H_{min}:

$$h(1) - h(2) \geq H_{min} \qquad (2)$$

And:

$$h(4) - h(3) \geq H_{min} \qquad (3)$$

The maximal horizontal difference between two characteristic points on the same slope must not exceed the search radius r expressed by the ratio of the minimal depth parameter to the tangent value of the minimal inclination α_{min}:
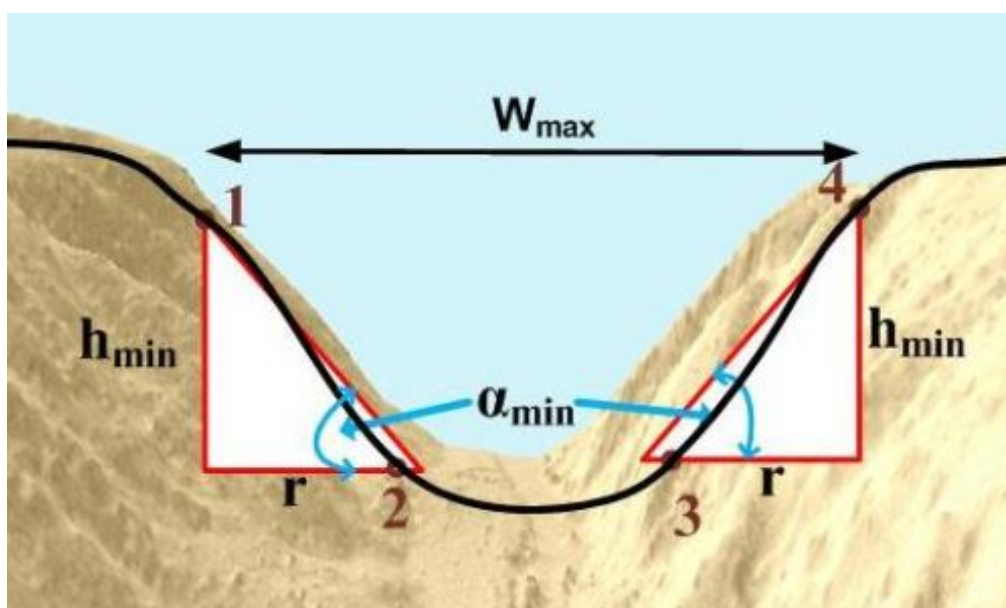
$$\mathbf{r = \frac{h_{min}}{tg(\alpha_{min})}} \qquad (4)$$

That is:

$$d_h(1,2) \leq r \qquad (5)$$

$$d_h(3,4) \leq r \qquad (6)$$

Figure 3. Geometric relationships between four characteristic points (1, 2, 3 and 4) defined by the parameters of the maximal width W_{max}, the minimal height h_{min} and the minimal inclination angle α_{min}.
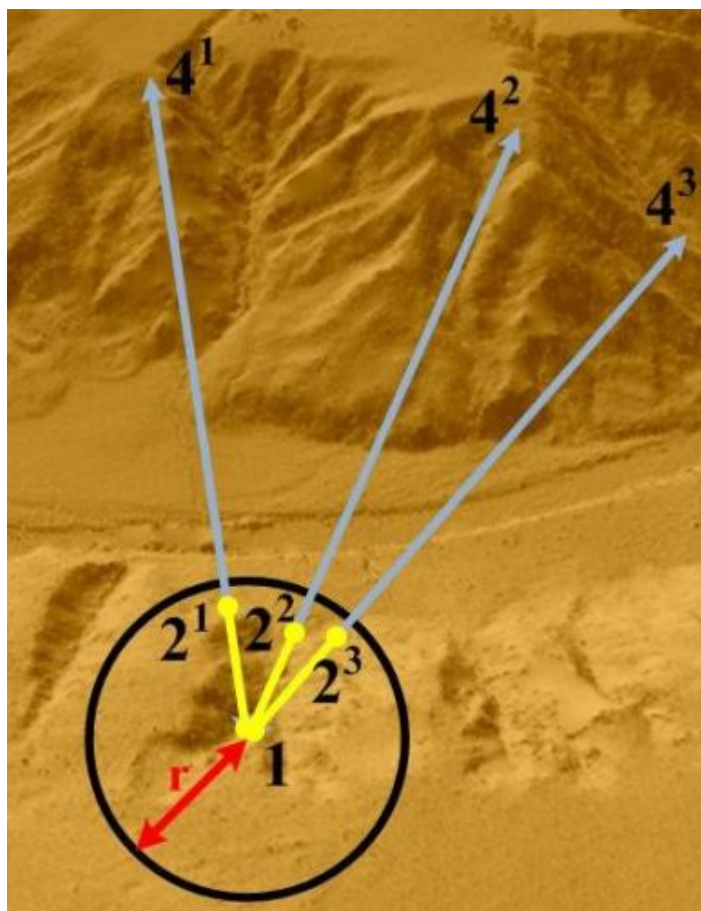

Bearing these relationships in mind, a procedure for detection of the characteristic points is devised based on the ray tracing method. Therefore, a circle with search radius r given in Equation 4 is constructed around each DTM point (point 1 in Figure 4). Then, all DTM pixels lying within this circle are identified. Among these pixels, only those lying below the tested DTM pixel (point 1) and vertically separated from it for at least the amount of the parameter of the minimal depth (points 2^1, 2^2 and 2^3) are considered as the potential candidates for the second characteristic points of cross-sections. Through each of these pixels, rays (1-4^1, 1-4^2 and 1-4^3) are constructed originating in the DTM pixel under consideration with the length equal to the parameter of the maximal width.

$$d_h(1,4^1) = W_{max} \qquad (7)$$

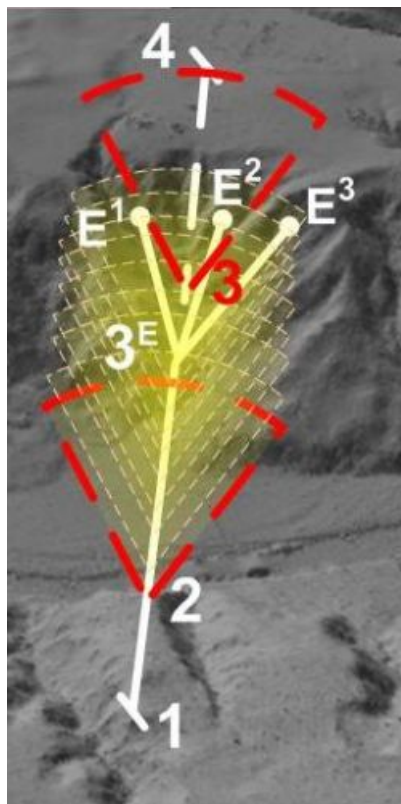$$d_h(1,4^2) = W_{max} \qquad (8)$$

$$d_h(1,4^3) = W_{max} \qquad (9)$$

Figure 4. Constructed rays (1-4^1, 1-4^2, 1-4^3) with the length corresponding to the parameter of the maximal width defined by the DTM point under consideration (point 1) and points classified as potential candidates for the second characteristic points (2^1, 2^2, 2^3).

In order to be classified as cross-sections, rays need to contain a pair of characteristic points representing the slope opposite to that defined by the DTM point under consideration and the candidate pixels for the second characteristic points. Instead along the path of the constructed ray, these characteristic points of the opposite slopes are traced within a sequence of circular sectors (Figure 5) constructed in each pixel belonging to the segment within the ray between the candidate pixel for the second characteristic point (2) and the point (3) search radius away from the endpoint (4) of the ray. The radius of the circular sectors is equal to the search radius r while the size of the central angle of sectors is a predefined value. The search terminates with the circular sector with the centre (3^E) closest to the second characteristic point within which at least one DTM pixels has been found satisfying spatial requirements for the endpoint of the cross-section. If more pixels have been identified as endpoints (E^1, E^2 and E^3) of cross-sections in the circular sector, opposite slope is presented with segments joining each of these endpoints with the centre (3^E) of the circular sector within which they have been detected.

Figure 5. The identification of the characteristic points on the opposite slope within circular sectors constructed along the section of the ray between the second characteristic point 2 and the point 3 that is horizontally separated from the ray endpoint 4 with the search radius r. Segments joining the endpoints (E^1, E^2 and E^3) with the centre (3^E) of the circular sector within which they are identified represent the opposite slope.

3.3. Merging of the DTM Pixels Representing Cross-Sections in to Regions
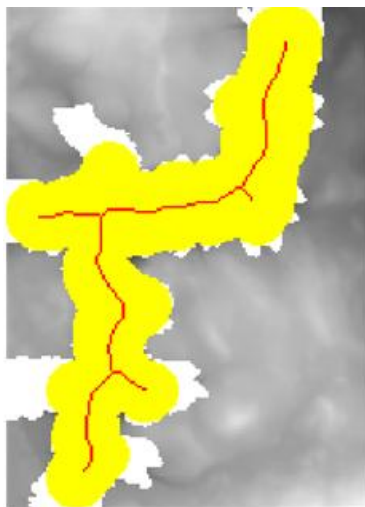
The aggregation of pixels representing identified cross-sections is based on the 8-connectivity criterion. Regions created in this manner were pre-processed before their length is to be estimated – holes within regions have been filled out and branches signifying isolated identified cross-sections have been eliminated using the pruning technique.

3.4. The Estimation of the Length of the Regions Composed of Identified Cross-Sections

The use of two software applications for digital image processing, differing considerably in the choice of available operators, entailed devising two different approaches for calculating the length of the regions composed of identified cross-sections.
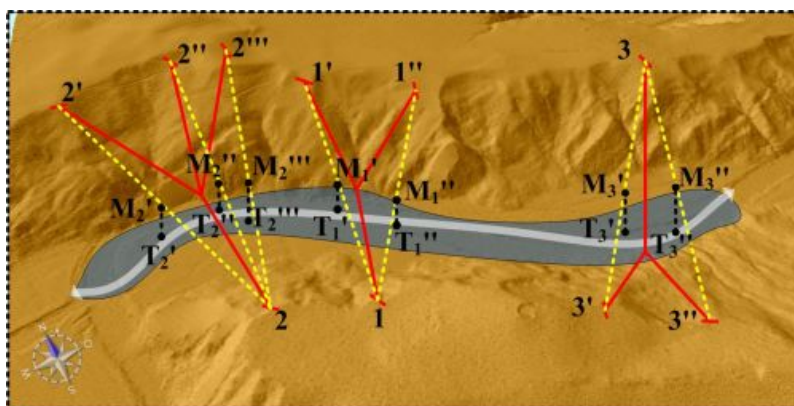
In the first approach, developed with the digital image processing software HALCON (http://www.mvtec.com), the regions' boundaries were smoothed by using the circle with the radius equal to the parameter of the minimal width (Figure 6). The smoothed regions are thereafter skeletonised. The length of acquired skeleton lines is taken as the approximate length of the canyon.

Figure 6. Smoothing of regions composed of identified cross-sections – white areas represent parts of the region excluded by smoothing, the yellow area is the result of smoothing and the red line is the skeleton line used for the length assessment.

In the second approach, implemented using the open source software for image processing ImageJ, the length assessment is based on the midpoints of identified cross-sections not shorter than the parameter of the minimal width (Figure 7). The coordinates of these points have been calculated by averaging endpoints' coordinates of the selected cross-sections. In the next step, the midpoints are merged into regions on the basis of the 8-connectivity principle. Among these regions, the one with the largest area is selected for skeletonization. The length of the skeleton line is considered as the calculated length of the region composed of pixels representing the identified cross-sections.

Figure 7. Several identified cross-sections (red) with their midpoints (M) and their projections on the terrain surface (T). Midpoints of the cross-sections longer than the parameter of the minimal width are merged into the region (grey) which skeleton line (white) is used for the length assessment.

In the final stage, the estimated lengths of regions containing cross-sections have been compared to the parameter of the minimal length. The regions with calculated lengths that are not shorter than the minimal length parameter are classified as the planimetric representations of canyons.

4. Results

The results of the extractions are shown in Figures 8 and 9. Visual inspection conducted using TreesVis software revealed that the majority of pixels representing valleys and steep slopes of canyons have been successfully identified. Draping extracted canyons onto contour maps derived from DTMs by using ArcView software, clearly show how their border lines envelope the densely spaced contour lines representing steep slopes of canyons.

Figure 8. The first DTM with 16 m resolution and size of 148 × 204 pixels (left) and the extracted canyon (middle – white area) The border line of the extracted canyon (red) overlaid onto the contour map (right) with the equidistance of 25 m derived from the first DTM.
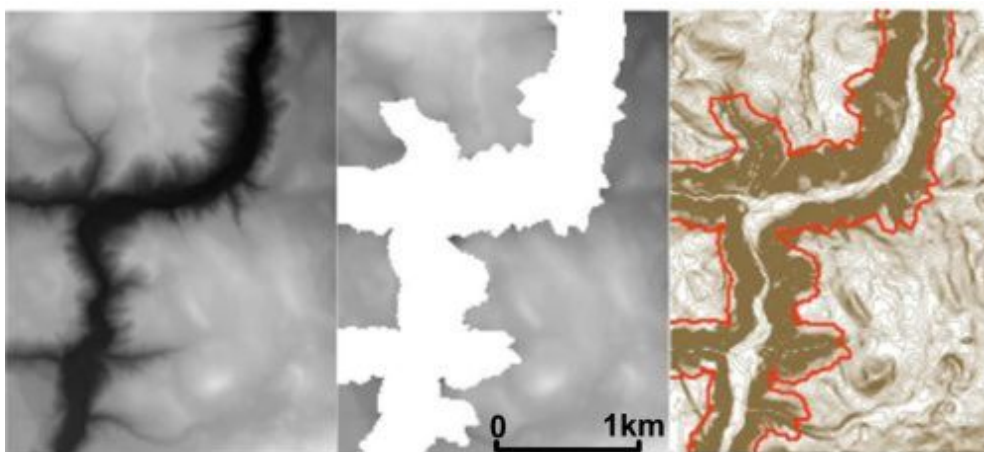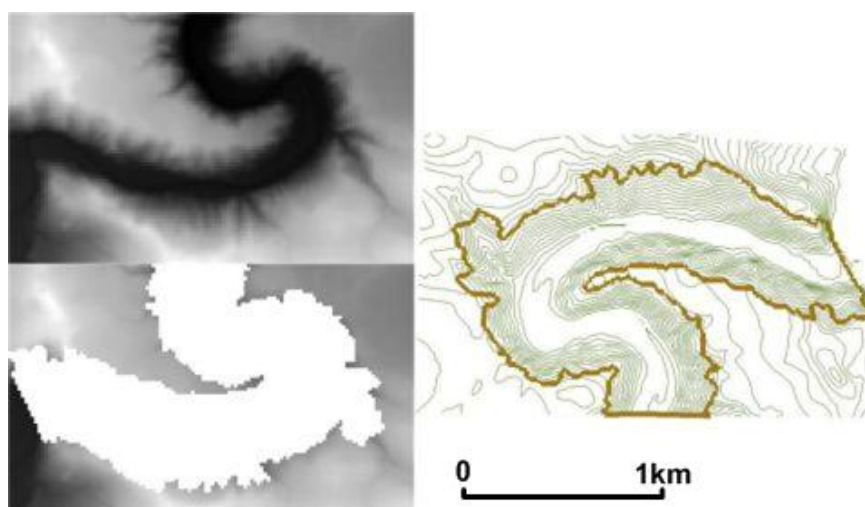

Figure 9. The second DTM (16 m resolution) used for testing (up left) and with draped extraction results (down left – white area). The boundary of the extracted canyon superimposed on the contour map with the equidistance of 10 m (right).

The results presented in the Figure 8 are obtained by selecting following input parameters: maximal width = 350 m, minimal width = 150 m, minimal inclination = 10°, minimal height = 10 m, minimal length = 1,500 m.

The following input parameters have been utilized for the extraction of the canyon from the DTM shown in the Figure 9: maximal width = 500 m, minimal width = 200 m, minimal inclination = 15°, minimal height = 10 m, minimal length = 1,500 m.

The selection of input parameters has been based on measurements of characteristic sections of the canyon's area in DTM such as the canyon width in the narrow and wide sections as well as the inclinations of milder slopes in the canyon.

5. Discussion

An accuracy assessment has not been undertaken due to the lack of reference data. The use of quantitative indicators proposed for evaluating results of binary classification such as correctness, completeness and quality [8] is difficult to estimate and apply for landforms since they seldom exhibit crisp boundaries [9]. An alternative assessment method could be ascertaining the degree of coincidence between canyons delineated by the proposed algorithm and human experts [10].

The selection of the input parameters has a major impact on the extraction results. It can be guided by user's conception of canyon's morphometry. The given dimensions of a canyon can be acquired upon testing the algorithm repeatedly by choosing various threshold values of the input parameters until the expected extraction result is obtained.

Needless to say, the extraction results are affected by the scale effect. However, since the extraction is based on dimension parameters, its results do not change considerably when tested on the DTMs with various resolutions (Figure 10), as long as the resolution does not exceed the smallest selected value for any of the given dimension of the canyon. However, improvement of the DTM resolution has a negative effect on computation intensity resulting in the increased number of DTM pixels needed to be checked during the identifications of cross-sections.

Figure 10. Results of canyon extractions performed on DTMs with the resolution of 8 (left) and 16 (right) meters.
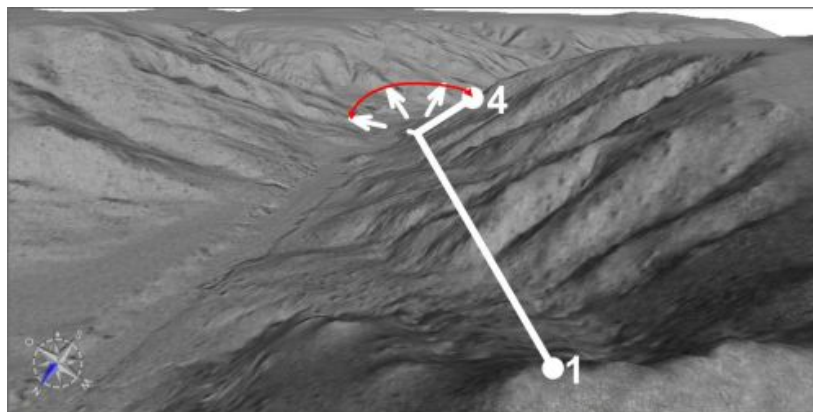

The proposed algorithm enables automatic delineation of a canyon in an entirely automated manner. It provides a canyon's borderline without directly searching after it. Thus, this extraction procedure represents an alternative to various approaches for the reconstruction of breaklines of slopes commonly based on the edge detection methods.

Results of the extraction depend on a degree of coverage of the canyon's area with identified cross-sections. This number is decreased in the case that slopes of canyons are intensively fractured, eroded or exhibit high variability in their morphology. A lack of identified of cross-sections can lead to partial identification of canyon's area and underestimation of the length.
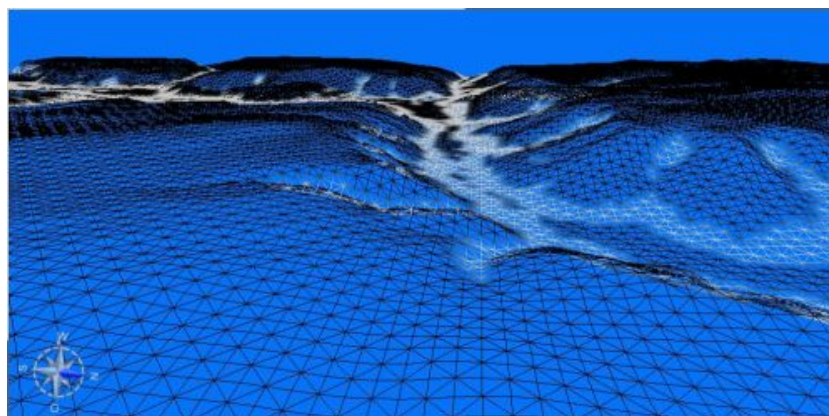
Another important factor that influences the results is the size of the central angle of circular sectors within which endpoints of cross-sections are searched. Circular sectors have been introduced in order to improve the identification of cross-sections in sections of canyons where discontinuities on slopes are pronounced and where the valley of the canyon intensively meanders. Nevertheless, the selection of the larger values for this angle carries a risk of identifying rays stretching across a slope as cross-sections (Figure 11). Therefore, it is worth considering inclusion of this angle in the list of input parameters, allowing the user to select a recommended value for the angle depending on a type of terrain.

Figure 11. A constructed ray (1-4, full white line) traversing a slope is classified as a cross-section due to the large value of the central angle (red) of circular sectors used for detecting endpoints of cross-sections.

The algorithm permits the identification of deepest DTM pixels along each of the identified cross-sections (Figure 12). Since the algorithm can be implemented on DTMs representing seafloors or river beds, these points can be used for the construction of the thalweg which is valuable topographic information.

Figure 12. Pixels with the lowest elevation (white) along identified cross-sections draped onto the wired mesh of DTM.

The ray tracing method for identification of cross-section can be applied for hills or mountains since geometric relationships given in 3.2 can be also applied if altitudes of four characteristic points along a cross-section are inverted (Figure 13). The points with maximal elevations along of individual cross-sections can be used for detections of ridgelines.

Figure 13. Four characteristic points of a cross-section along a hill.

6. Conclusions

The proposed algorithm for automatic extraction of canyons based on dimension parameters yielded promising results. By being able to select values of input parameters, a user is allowed to shape morphology of a landform to be extracted. In this manner, various types of depressions can be delineated. Moreover, the algorithm can be applied for extraction of hills and mountains. Thus, using this procedure provides great flexibility. Additional work is necessary to validate the suitability of applying the employed approach in various types of environmental or engineering studies.

Acknowledgements

References and Notes

1. Strobl, J. Segmentation-Based Terrain Classification. In *Advances in Digital Terrain Analysis* ; Zhou Q., Lees, B., Tang, G, Eds.; Springer Verlag: Berlin, Germany, 2008; pp. 125-139.
2. Schmidt, J; Dickau, R. GIS and DEMs – Semantics, Methods, Future Needs. In *GIS for Earth Surface Systems: Analysis and Mode lling the Natural Environment* ; Dikau, R., Saurer, H., Eds.; Gebruder Borntraeger Verlagsbuchhandlung: Stuttgart, Germany, 1999; pp. 153-173.
3. Dragut, L.; Blaschke, T. Automated classification of landform elements using object-based image analysis. *Geomorphology* **2006**, *81*, 330-344.

4. Babonneau, N.; Savoye, B.; Cremer, M.; Klein, B. Morphology and architecture of the present canyon and channel system of the Zaire deep-sea fan. *Mar. Petrol. Geol.* **2002**, *19*, 445-467.

5. Goff, J.A. Quantitative classification of canyon system on continental slopes and a possible relationship to slope curvature. *Geophys. Res. Lett.* **2001**, *28*, 4359-4362.

6. Weinacker, H.; Koch, B.; Weinacker. R. TREESVIS - A Software System for Simultaneous 3D-Real-Time Visualization of DTM, DSM, Laser Row Data, Multi-Spectral Data, Simple Tree and Building Models. In *Proceedings of the ISPRS Working Group on Laser-Scanners for Forest and Landscape Assessment*, Freiburg, Germany, October 2004; pp. 90-95.

7. OGC: Web Processing Service – Open Geospatial Consortium – the official web site. Available Online: http://www.opengeospatial.org/standards/wps (accessed on 21 February 2009).

8. Heipke, C.; Mayer, H.; Wiedemann, C.; Jamet, O. Evaluation of Automatic Road Extraction. *Int. Arch. Photogramm. Remote Sens.* **1997**, *32*, 47-56.

9. Smith, B.; Mark, D.M. Do mountains exist? Towards and ontology of landforms. *Environ. Plan. B - Plan. Design* **2003**, *30*, 411-427.

10. Fried, J.S.; Brown, D.G.; Zweifler, M.O.; Gold, M.A. Mapping Contributing Area for Stormwater Discharge to Streams Using Terrain Analysis. In *Terrain Analysis: Principles and Applications* Wilson, J.P., Gallant J.C., Eds.; John Wiley & Sons: New York, NY, USA, 2000; pp. 183-204.