Fractal perspectives of GIScience based on the leaf shape analysis

P. Tuček^{1,2}, L. Marek¹, V. Paszto¹, Z. Janoška¹ and M. Dančák³

¹Department of Geoinformatics, Faculty of Science, Palacký University in Olomouc, Tř. Svobody 26, 771 46 Olomouc, Czech Republic Telephone: +420 585 634 513

Fax: +420 585 034 513 Fax: +420 585 225 737 Email: <u>vit.vozenilek@upol.cz</u>

²Regional Centre of Advanced Technologies and Materials, Palacky University in Olomouc, Department of Mathematical Analysis and Applications of Mathematics, 17. Listopadu 12, 771 46, Olomouc, Czech Republic Telephone: +420 585 634 521 Fax: +420 585 225 737 Email: <u>pavel.tucek@upol.cz</u>

³ Department of Botany, Faculty of Science, Palacký University in Olomouc, Šlechtitelů 11, 783 71 Olomouc, Czech Republic Telephone: +420 585 634 805 Fax: +420 585 634 824 Email: <u>martin.dancak@upol.cz</u>

1. Introduction

Since Mandelbrot (1967) published its basics, fractal geometry and fractal dimension (non-integer dimension) is well known as a valuable tool for describing the shape of objects. It gained large popularity in many fields of natural sciences (Batty and Longley 1994, Goodchild 1980, Hastings and Sugihara 1994, Kitchin and Thrift 2009, Peitgen et al. 1992), including e.g. ecology, geography, GIScience, where the measures of object's shape are essential.

One of the major principles in fractal geometry is self-similarity and self-affinity. The most theoretical fractal objects, such as Barnsley's fern, are self-similar (any part of the object is exactly similar to the whole) and self-affine (transformed self-similar objects). And typical fractal objects like leaves are very suitable to test our methods for possible use on geodata. And for geospatial fractal-based analysis, the various drainage systems were acquired and examined.

For leaves, we show that discrimination based on only two fractal features has more than 90% accuracy. This notion is important, because it proves that automated classification can be based also on complexity of shapes and not only on their qualitative measurements. Fischer discriminant analysis is used to distinguish between families and species with satisfactory results. Leaf skeleton is very similar especially to river and road network and thus we examine different types of river drainage network and show that their complexity differs significantly.

2. Methods

There exist a number of methods for estimating fractal dimension and as e.g. Reynoso (2005) shows, results obtained by different methods often differ significantly. Also not only the method itself, but the software, which calculates the fractal dimension, may contribute to the differences (Reynoso 2005). All the calculations were accomplished in free software Fractalyse, easily downloadable from <u>www.fractalyse.org</u>.

2.1 Box-counting method

The box-counting method was used for modified data – binary pictures. Box-counting dimension of a subset X of the plain is defined by counting number of unit boxes which intersects X: for any $\Delta s > 0$, let $N(\Delta s)$ denote the minimum number of *n*-dimensional cubes of linear scale Δs (side length) needed to cover X. Than X has box dimension D if $N(\Delta s)$ satisfies (according to Hastings and Sugihara 1994, Theiler 1990):

$$N(\Delta s) \approx c(1/\Delta s)^{D}, \qquad (1)$$

where $\Delta s \rightarrow 0$, *c* is a constant and box-counting dimension of *X* is *D*. Formula (1) is called power law. Dimension *D* is then be computed by:

$$D = \lim_{\Delta s \to 0} \left[-\log N(\Delta s) / \log \Delta s \right],$$
(2)

According to formula (2), calculation of box-counting dimension is simple. For a sequence of cell size $\Delta s > 0$, the number of cells $N(\Delta s)$ needed to cover the set S is calculated.

2.2 Linear Discriminant Analysis (LDA)

Discriminant analysis is used in situations where the clusters are known a priori. The aim of discriminant analysis is to classify an observation, or several observations, into these known groups (Hardle and Simar, 2007). The classification rule is often a linear function of measurements that maximizes the separation between groups relative to their withingroup variability (Johnson and Wichern 2007). Discriminant scores are results of the LDA.

3. Data processing

Unique dataset of leaves was available thanks to Department of Botany of Faculty of Science in Olomouc. At first, possibilities of automated data classification were tested on dataset of leaves. A unique dataset containing 133 samples of leaves from 7 different species belonging to 3 families was available for scanning into raster digital format (fig. 1 left). Examined plants can be divided to two groups (tab. 1).

First group	Second group			
(Angiosperms, Eudicots)	(Pteridophytes, Polypodiopsida)			
Roseaceae family:	Dryopteridaceae family:	Blechnaceae family:		
Alchemilla vulgaris (simple lobed leaves), Rubus wimmerianus (compound 5-foliolate leaves), Fragaria moschata (compound trifoliolate leaves)	Polystichum aculeatum (compound 2-pinnate leaves), Dryopteris filix- mas (compound 1-2 pinnate leaves), Dryopteris carthusiana (compound 2- 3-pinnate leaves)	<i>Blechnum spicant</i> (simple pinnatifid sterile leaves and compound 1-pinnate fertile leaves)		

Table 1. Two main groups of examined plants (with particular species).

Scanned leaves were transformed into two datasets. First examined data-set was acquired by transforming raster pictures into binary raster picture in order to perform fractal analysis of leaf area (fig. 1 middle). Then the leaf skeleton was digitalized and extracted from the raster pictures for further fractal dimension analysis (fig. 1 right).



Figure 1. Fern leaves and particular steps in their processing.

Fractal dimensions of both, the leaf area and skeleton, were calculated. These computed fractal properties together with an affiliation to family (or species) served as basis attributes for linear discriminant analysis. R project was used as a computational environment where LDA was applied, visualized and compared.

The LDA was applied using the relation *group~area.FD+skeleton.FD* as the basic formula. Two-dimensional space in which single points belonged to one of three families (or 7 species) was the result of the analysis.

Predicted affiliations to single groups (based on both above mentioned methods) was visualized and compared with the real belonging to the group (fig. 2).

River drainage network was obtained from free Internet source. Ten examples from each selected drainage system were selected and examined. Data are available from DIVA-GIS website (<u>http://www.diva-gis.org/Data</u>) and were used to select appropriate areas with typical drainage systems. Drainage system is governed by many factors, most importantly by topography, geology and preceding (and contemporary) geomorphologic processes (Zernitz 1932). There are several types of drainage system, each of them typical for a certain type of relief (Knighton 1998).

Type of drainage system	Description	Example	
Dendritic	Most common, river follows the slope, in V-shaped valleys	Mississippi	
Parallel	Steep, uniformly sloping relief	Angola (Moxico, Cuando Cubango)	
Trellis	Folded mountains with strike valleys	Appalachian Mountains	
Rectangular	Rocks with uniform resistance to erosion, but with two directions of jointing at approximately right	Western Iran	

	angles	
Radial	Typical for volcanoes, craters, radial depressions	Kauai Island
Deranged	No coherent pattern of rivers and lakes,	Canadian Shield

Table 2. Examined river drainage systems with descriptions and examples (according
to Knighton 1998, Lambert 1998, Ritter 2006).

4. Results

In the case of LDA, the result of analysis is two-dimensional space, where points belong to the predicted groups (fig. 2). The success of the classification to the family was 93.2 % and 64.7 % in the classification to the species. Overview of classification into the families is in the table 3.



Figure 2. Classification to families using LDA.

	А	В	С	Sum
	LDA	LDA	LDA	
А	62	0	2	64
В	3	11	1	15
С	3	0	51	54

Table 3. Comparison of classification to particular families.

Cluster analysis showed that species from the same family are more similar than species from different families (fig. 3) and shows unsupervised classification of species, which corresponds with actual taxonomical classification.



Figure 3. Dissimilarity of different species based on the cluster analysis.

Then, for each selected area of different river drainage system, fractal dimension of river network was computed. Results are summarized in figure 4.



Figure 4. Graphical overview of fractal dimension values of different drainage system types.

Analysis of variance (ANOVA) was used to verify that mean fractal dimension for each drainage system differs. Bartlett's test, however, did not refuse the equality of variances. Since the only one characteristic of the drainage system was measured, data were not suitable for classification purposes. Despite of this fact, cluster analysis was conducted based on mean and standard deviation of values for each drainage network in order to examine the similarity of drainage systems.

The most alike are parallel and rectangular drainage systems, which both embody quite similar patterns. Both of them (and also trellis drainage system) show regular patterns and their fractal dimension is the lowest. Radial and dendritic drainage systems embody more irregular shapes and therefore their fractal dimension is higher. The most irregular patterns are observed in deranged drainage system. Meanders, lakes and flood plane lobes are typical for deranged drainage systems and it is difficult to observe any distinguished structure in the shape of river network. The complexity of this type of drainage system is the greatest and analysis shows that the fractal dimension is also the highest. Cluster analysis shows that this type of drainage network is also the most dissimilar from any other drainage system types.



Figure 5. Dissimilarity of different drainage systems based on their fractal dimension.

5. Conclusion

We showed that just only two characteristics based on fractal dimension measurement (without any additional geometric features) are sufficient for accurate object classification. Although the analyses were applied to leaves, it is possible to use the methods for various geographical analyses based on geodatasets.

Although leaves characteristics were used in this study, the perspectives of using analogical methods in GIScience are bidding themselves.

We examined different types of drainage systems by means of their fractal dimension. Results show that mean fractal dimension for the six most common drainage systems differ significantly, and that the more regular drainage system, the lower fractal dimension (and vice versa). Complexity of fluvial network can point to different origin of the rivers and therefore this could be useful for geologist to semi-automatically evaluate the drainage systems. Also river complexity can be of vital importance in precipitationrunoff evaluation of the water basin and could be correlated to different processes, e.g. during flood event.

Further research concerning fractal measurements upon geospatial data is nowadays intensively conducted on Department of Geoinformatics, Palacky University in Olomouc, Czech Republic.

6. Acknowledgements

This work has been supported by the Operational Program Education for competitiveness – European Social Fund (CZ.1.07/2.2.00/15.0276) and the by the Ministry of Education, Youth and Sports of the Czech Republic.

7. References

- Batty M. and Longley P., 1994, Fractal Cities: A Geometry of Form and Function, Academic Press Ltd., London, San Diego, 1994, 394 s.
- Goodchild M.F., 1980, Fractals and the accuracy of geographical measures. Math. Geol., Vol 12, pp 85–98.
- Härdle W., Simar L., 2007, Applied Multivariate Statistical Analysis. Springer Berlin, 2nd edition.
- Hastings H. M., Sugihara, G., 1994, Fractals: A User's Guide for the Natural Sciences. Oxford : Oxford University Press. 235 s.
- Hothorn T., Hornik K., Zeileis A., 2006, Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics, 15(3), 651–674.
- Johnson R., Wichern D., 1982, Applied Multivariate Statistical Analysis, Prentice-Hall, Englewood Cliffs, NJ.
- Kitchin R., Thrift N., 2009, International Encyclopedia of Human Geography. United Kingdom : Elsevier Science. 8250 s. (hardcover).
- Knighton, D., 1998, Fluvial forms and processes: a new perspective. Oxford University Press, Inc., New York.
- Lambert D., 1998, The Field Guide to Geology. Checkmark Books. pp. 130–131.
- Mandelbrot B. B., 1967, How long is the coast of Britain? Statistical self-similarity and fractional dimension, Science 155, str. 636-638.
- Peitgen H.-O., Jürgens H., Saupe D., 1992, Chaos and Fractals : New Frontiers of Science. New York : Springer. 984 s.
- R Development Core Team, 2010, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <u>http://www.R-project.org</u>.
- Reynoso C., 2005, The impact of chaos and complexity theories on spatial analysis problems and perspectives. 24th Research Symposium: Reading Historical Spatial Information from around the World: Studies of Culture and Civilization Based on GIS Data, Kyoto Japan, 7-11 February. , 63(2):1037-1068.
- Ritter M. E., 2006 The Physical Environment: an Introduction to Physical Geography. Date visited.

http://www.uwsp.edu/geo/faculty/ritter/geog101/textbook/title_page.html

Zernitz, E. R., 1932, Drainage patterns and their significance, J. Geol., 40, 498-521.