How does the enumeration unit size influence spatial pattern recognition in choropleth maps? User study

Izabela Karsznia [a, *]

[a] Department of Geoinformatics, Cartography and Remote Sensing, Faculty of Geography and Regional Studies, University of Warsaw, i.karsznia@uw.edu.pl

* Corresponding author

Keywords: enumeration unit size, spatial pattern recognition, choropleth map, user study

Abstract:

The recognition and preservation of the characteristic elements of visualization are one of the fundamental principles of map design, especially in map generalization. As underlined by numerous authors, the importance of preserving or even highlighting the context of the visualization constitutes one of the more significant challenges in cartographic generalization (Mackaness W. A., Ruas A., and Sarjakoski L. T., 2007; Burghardt D., Duchêne C., and Mackaness W. A., 2014). Cartographic generalization, which tailors a map and other graphics to the map's purpose, detail level, as well as the intended audience's requirements; is a difficult but necessary task for the map designer. Without careful and meaningful generalization, maps will not convey the essential patterns contained in the data (Maceachren and Ganter, 1990; Mackaness and Edwards, 2002; Zhang and Guilbert, 2016; Yu 2018).

The research on cartographic generalization and pattern recognition concerning topographic maps is very rich. Since the 1960s, when the period of automated generalization began, the issues connected with pattern recognition and preservation were also being tackled from many aspects. These concerned point pattern preservation, including building typification (Regnauld 2001; Sester 2001; Ai and Liu 2002; Burghardt and Cecconi 2003; Yan and Weibel 2008; Yan and Li 2013; Bereuter and Weibel 2017; Yu 2018), line pattern recognition (Heinzle and Anders 2007; Yang B., Luan X., Li Q.,2010; Touya 2010; Wang 2012; Brewer et al. 2013; Shoman W., and Gülgen F., 2017), and polygon pattern analysis (Steiniger S., et al 2006; Atwood Williams and Wentz 2008; Sayidov and Weibel 2016).

On the other hand, in thematic mapping, especially in choropleth map design, cartographic generalization and pattern recognition also play an important role as they help to preserve and convey spatial information (Maceachren 1982). The effectiveness with which thematic maps communicate spatial information is reflected in previous research (Maceachren 1982). It is often associated with map complexity, which is influenced by many factors, including the number of classes, the degree of autocorrelation present in the data, the choice of class intervals, and the shape and size of the enumeration units (Mersy 1990). Map complexity is related to spatial pattern recognition and is often 'regarded as the visual intricacy of the spatial patterns formed by the map symbols' (Mersy 1990: 15) which form specific patterns that are important to preserve in the thematic map generalization process.

In thematic maps, generalization takes place when either the number of classes is decreased or the enumeration unit size is increased. While aspects related to optimal data classification methods and the number of classes have been extensively examined (Chang 1976; Chang 1977; Cox 1976; Dent 1999; Slocum et al. 2008; Cabello et al. 2010), the size of the enumeration unit as well as its influence on pattern recognition has not, so far, been the subject of in-depth empirical studies. Robinson (1984) notes that an enumeration unit which is too small in size in thematic map, results in a dispersed image and difficulty in perceiving characteristic elements in the data. A generalization level which is too small causes the image to be blurred and diffuse. As a result the user loses the picture of the whole, the context, and the characteristic elements of the structure – he or she is no longer able to see the relevant spatial patterns. At the same time, a generalization level which is too large, as expressed by the size of the enumeration unit in choropleth maps, also causes the characteristic elements in the data to be lost, as the visualization is too general. In both cases, a too strong simplification of the map pattern might result in the loss of valuable information for the map reader (Bregt and Wopereis 1990). Thus the enumeration unit size for thematic maps should be designed in such a way as to preserve specific global and local spatial patterns.

The challenge, therefore, is to find the optimal generalization level whereby the spatial patterns present in the data are adequately transformed and can be distinguished by the user. Thus the research goal is an attempt to answer the question as to whether specific presentation methods, namely choropleth maps with particular enumeration unit sizes, result in a more effective recognition of the characteristic patterns or trends in data. We also want to answer the question: do users prefer methods which have specific enumeration unit sizes? The presented research aims to answer where the approximate limit may lie, expressed by the enumeration unit size after which the user loses the possibility of correctly and effectively exploring the spatial patterns presented on thematic maps.

To answer this research question an empirical user study is being designed. In this study, we would like to investigate the effectiveness of conveying spatial patterns on choropleth maps using different enumeration unit sizes. Within the study, choropleth map variants will be compared to a symbol map presenting the phenomena in a discreet way. The symbol map being used as a reference for the choropleth maps can be treated as a presentation of the raw data. We will consider choropleth maps that have several hexagonal enumeration unit shape. Four enumeration unit sizes will be assessed. The users will be asked to evaluate the similarity of the choropleth map variants to the reference symbol map, and nominate their preferred map. An example of the stimuli to be presented to users is shown in figure 1.
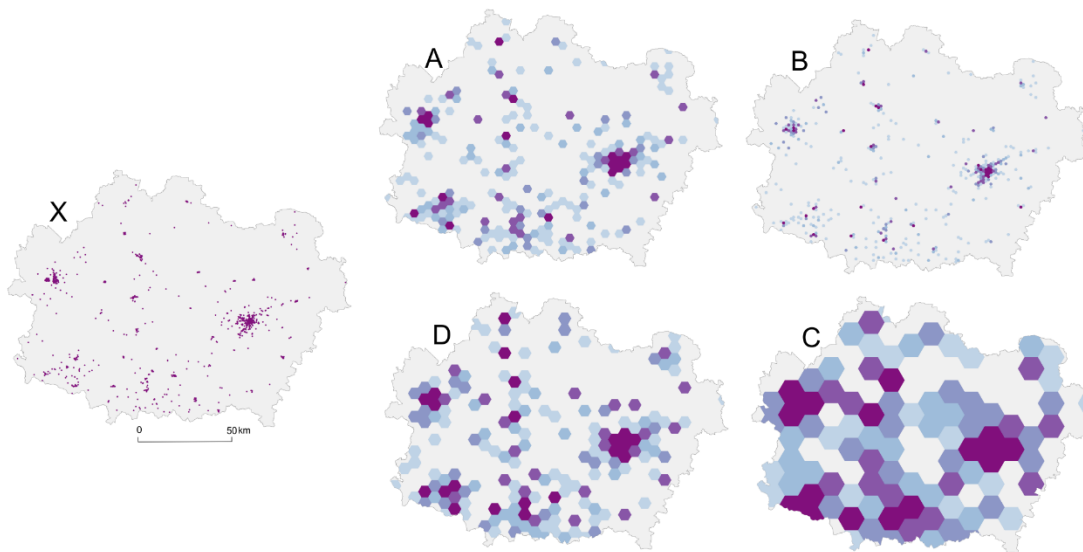

Figure 1. An example of stimuli to be used in the empirical user study on the general map reading level.

The empirical study will be conducted with over 100 high school students. The main assumptions, as well as the preliminary results of the user study, will be discussed. Within the user study, among other tasks, two task types concerning map reading at both a general and a detailed reading level are being planned. One task type concerns map reading, especially the possibilities for conveying spatial patterns at a general reading level. The second task type will concern map comprehension at a detailed reading level. We believe it will allow us to verify which size of enumeration unit works better at a particular map reading level. We also want to explore which enumeration unit size works better for a particular task type. The assumptions will be verified based on the usability performance metrics: the correctness, the speed of the users' responses as well as the users' preferences.