

DO THE VISUAL COMPLEXITY ALGORITHMS MATCH THE GENERALIZATION PROCESS IN GEOGRAPHICAL DISPLAYS?

A. Brychtová^{a*}, A. Çöltekin^a, Vít Pászto^b

^a GIVA, Department of Geography, University of Zurich, Switzerland - (alzbeta.brychtova, arzu.coltekin)@geo.uzh.ch

^b Department of Geoinformatics, Palacký University Olomouc, Czech Republic – vit.paszto@upol.cz

Commission II, TeS: WG II/6 Geovisualization and Virtual Reality

KEY WORDS: Visual complexity, feature congestion, subband entropy, generalization, maps

ABSTRACT:

In this study, we first develop a hypothesis that existing quantitative visual complexity measures will overall reflect the level of cartographic generalization, and test this hypothesis. Specifically, to test our hypothesis, we first selected common geovisualization types (i.e., cartographic maps, hybrid maps, satellite images and shaded relief maps) and retrieved examples as provided by Google Maps, OpenStreetMap and SchweizMobil by swisstopo. Selected geovisualizations vary in cartographic design choices, scene contents and different levels of generalization. Following this, we applied one of Rosenholtz et al.'s (2007) visual clutter algorithms to obtain quantitative visual complexity scores for screenshots of the selected maps. We hypothesized that visual complexity should be constant across generalization levels, however, the algorithm suggested that the complexity of small-scale displays (less detailed) is higher than those of large-scale (high detail). We also observed vast differences in visual complexity among maps providers, which we attribute to their varying approaches towards the cartographic design and generalization process. Our efforts will contribute towards creating recommendations as to how the visual complexity algorithms could be optimized for cartographic products, and eventually be utilized as a part of the cartographic design process to assess the visual complexity.

1. INTRODUCTION

Several algorithmic approaches have been proposed in interdisciplinary scientific literature to measure the complexity levels of visual displays (Batty et al., 2014; MacEachren, 1982; Moacdieh & Sarter, 2015; Rosenholtz et al., 2007). When applied to maps, the question arises whether these algorithmic measures match the semantically enhanced cartographic generalization approaches, given that 'simplification' operations to reduce complexity dominate the generalization processes. In this project, to assess and validate quantitative approaches to measuring visual complexity in a cartographic context, we compare results from a selected set of visual complexity algorithms. In this paper, we present results from one of these visual complexity algorithms, which we applied to various map types (i.e., 2D cartographic maps, hybrid maps, aerial/satellite images, shaded relief maps). For the 2D cartographic map types, we extended the selection and applied the algorithm for a set of maps with various levels of generalization. We believe such a comparison would be helpful in validating and improving the aforementioned algorithmic measures, and optimizing them for geographic visualizations, which, in turn, could be used as interim measures for cartographic design as the display size and the zoom levels change.

Despite the great progress in automation efforts since the beginning of the computer cartography (Brassel & Weibel, 1988), generalization process as a whole remains highly 'human' (i.e., human decision-making is involved in many of the steps), offering geometric as well as semantic approaches in managing the visual complexity of (geo)visualizations. With generalization, cartographers and other visualization designers aim to summarize information without removing essential details, making the display more legible, and highlighting the

most relevant information (e.g., Shea & McMaster, 1989). Therefore, generalization essentially reduces visual complexity through simplification operations in order to improve legibility, or visually designing the objects and their attributes for better discriminability (such as shape, color, mutual distance, see e.g., Brychtová & Çöltekin, A., 2014, 2016), so that the human eye can reasonably distinguish them. A further goal when working with complex visualizations related to level of detail management is to reduce computational processing time and bandwidth requirements, especially when dealing with large datasets, e.g., large map collections, imagery, 3D graphics (e.g., Çöltekin et al., 2011). Visual complexity algorithms can serve as supporting tools for generalization and level of detail management, especially when we are concerned with perceptual issues by determining if a display is too complex (thus would require a (re)consideration of generalization and level of detail management for the given scale). However, most modern visual complexity measures are not tested, thus also not optimized, for geographic visualizations. To fill this gap, as a first step, we contrast the results from the generalization process to the results from the selected visual complexity measures; examine *if* and *when* the complexity measures reflect the generalization levels and study the reasons in cases where they do not.

2. METHODS

Because many visual complexity measures are based on psychophysiological literature on how attention works (i.e., they account for certain optical, perceptual and cognitive characteristics based on, broadly, vision science), we hypothesize that these complexity measures will overall reflect the levels of map generalization. At the same time, we believe, that measuring the visual complexity of geographic displays could help to evaluate the quality of generalization process. In a

* Corresponding author

related study, Pászto et al. (2015) examined geometry generalization using shape metrics, which provide quantitative evaluation of the simplification degree of the shape (e.g., building footprints), demonstrating that shape metrics could help reveal discrepancies in the process of semi-automated generalization.

In this study, similarly to Schnur et al. (2010), we selected examples of common geovisualization types and computed their *Feature Congestion* (FC), which is a metric proposed by Rosenholtz et al. 2007 to describe visual complexity. Feature Congestion was originally designed to describe the clutter of visual displays. It involves color, orientation and luminance contrast as three features of the visual clutter while considering their local and global covariance in multiple scales (for more details, see Rosenholtz et al. 2007). We believe that *Feature Congestion* can be applied in cartography to assess either the quality of symbolization (e.g., if a lot of contrasting black is used, such as in swisstopo maps, FC scores will be high), and amount of visualized information, which is closely related to generalization.

Our study was conducted in two phases. At first, we aimed at analyzing the effect of the scale on the visual complexity. According to Ruas (2008) for a given size of map sheet, nearly the same quantity of information should be given for different scales. Thus we hypothesized that the visual complexity of consecutive generalization levels should remain equal. In the second phase, we analyzed how the visual complexity is influenced by the content (information density) of geovisualizations. We expected that visual complexity will be positively correlated with the amount of mapped features.

We selected example displays to analyze from the widely accessible geovisualizations provided by Google Maps (Map data ©2016 Google & various imagery providers), OpenStreetMap (Map data ©2016 OpenStreetMap contributors) and SchweizMobil (Geodaten ©2016 swisstopo).

Phase 1: First, we analyzed 5 sets of geovisualizations of various *types* as shown in Figure 1: OpenStreetMap (OSM)'s cartographic map, and four types from Google Maps: 'roadmap' (the standard 2D cartographic map), terrain, hybrid and satellite visualizations. For each type we distinguished 6 *zoom levels* (corresponding to Google Maps zoom levels 8, 9, 10, 12, 14 and 16) each with 12 *samples* (Figure 1). Therefore, in total we worked with 360 maps (5 types × 6 zoom levels × 12 samples). All samples were of the same size (1600×1600 px) and their centers were placed to 12 predefined (but arbitrary) coordinates within Europe. The scene content was not particularly controlled in this phase, however all scenes displayed a rural landscape without any dominant features such as big cities.

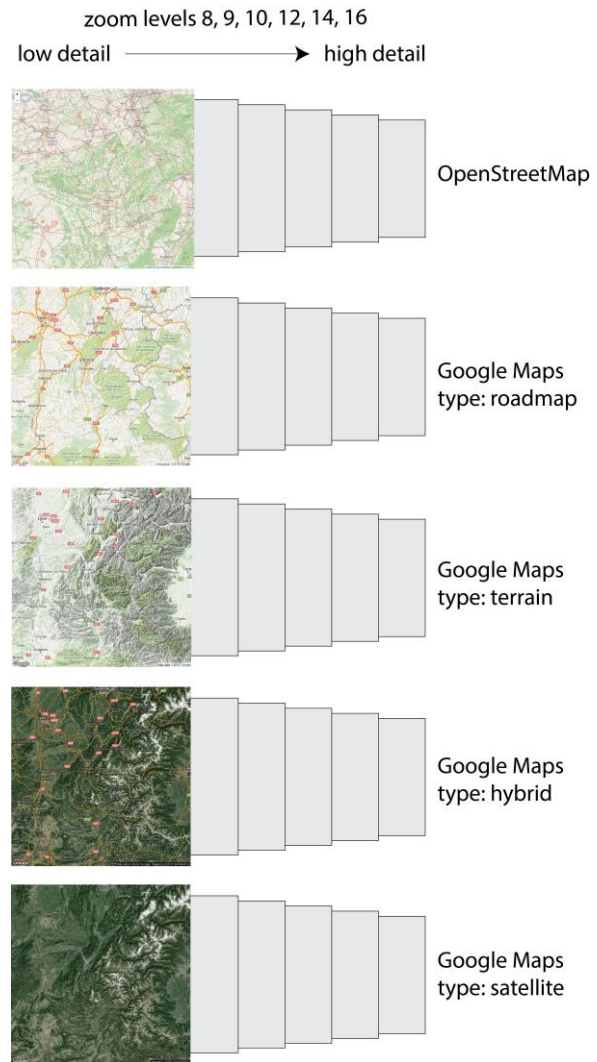


Figure 1. Analyzed geovisualization types in the Phase 1.

Phase 2: We selected two sets of 'triplet' city maps from each map provider. The screenshots are 'triplets' because they are identical in area (e.g., part of the Zurich city, as shown in Figure 2), and scale (corresponding to Google Maps zoom level 14). First set showed the city center and the second set an arbitrarily selected suburban area. As stated earlier, the maps were taken as screenshots from Google Maps, OpenStreetMap and SchweizMobil (Figure 2). In this phase, the scene (map) content was balanced to show areas of higher and lower urban development (city center and suburban area).

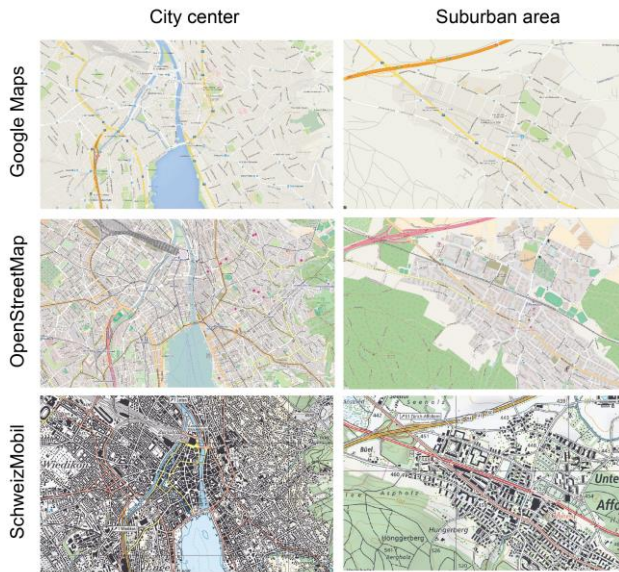


Figure 2. Analyzed city maps in the Phase 2.

3. RESULTS

Figure 3 shows the Feature Congestion (FC) scores for the studied geovisualization types (as shown in Figure 1) in the Phase 1. The highest mean visual complexity $Mdn=4.85$ (measured with FC) was found for a set of OpenStreetMap displays while the lowest was (counterintuitively) for Google Maps' satellite displays ($Mdn=2.11$). On the other hand, as expected, the visual complexity of satellite displays were also the least sensitive on the zoom level and mapped content ($sd=0.37$) in comparison to the other visualization types whose standard deviation (sd) is much higher ($sd_{OSM}=1.93$, $sd_{roadmaps}=1.83$, $sd_{terrain}=1.83$, $sd_{hybrid}=1.26$).

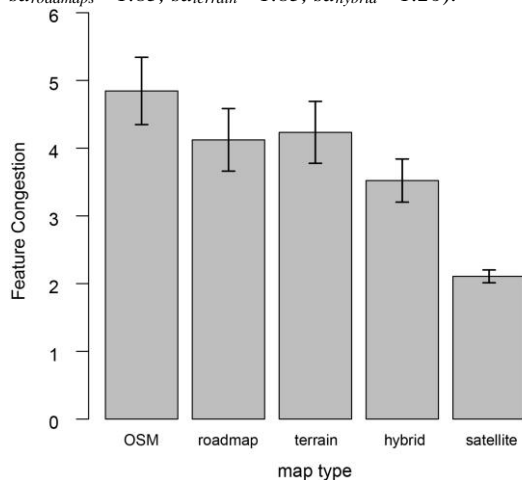


Figure 3. Mean Feature Congestion of examined types of geovisualizations. Error bars correspond to $M \pm 2 \times SE$.

The smaller the scale of the map, the less detail should be displayed per square kilometer. Conversely, the larger the scale, the more detailed is the area mapped for the same map size (Ruas, 2008). However, as can be seen in Figure 4, by studying the Feature Congestion of individual zoom levels, we found that all geovisualizations types, except satellite images, manifest increase of visual complexity as the level of map detail decreases. The generalization process is typically executed from high detailed data by removing/suppressing some information in

less detailed displays of larger areas. Thus, the relative visual complexity should be ideally constant across all zoom levels. However, our results, as presented in the Figure 3, show that less detailed displays of smaller scale (e.g., zoom level 8) contains more feature congestion ('more information'), than highly detailed displays of large scale (e.g., zoom level 16), contradicting our expectations. On the other hand, these results, especially given that the results for the satellite images remain more or less constant over scale changes, suggests that the algorithm is working as it should in this case.

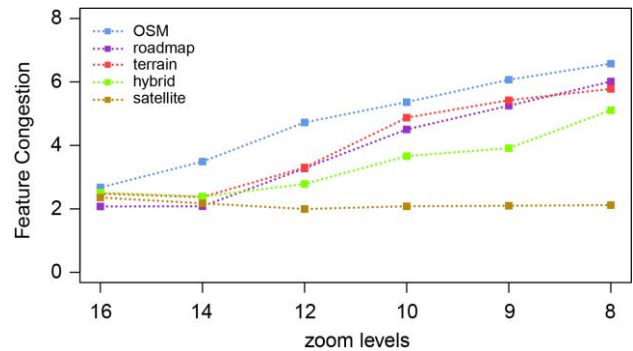


Figure 4. Mean Feature Congestion of geovisualizations types across examined zoom levels.

In the Phase 2, to track the effect of map content (area) on visual complexity we studied six maps displaying two areas of the same scale (see Figure 3). The visual complexity of different map providers considerably varied (see Table 1). The highest FC was for SchweizMobil maps in both areas, while the lowest was for Google roadmaps. Differences between individual providers can be assigned both to the applied cartographic symbology (visual variables), and to the amount of displayed symbols. In this point, we are not able to say what level of FC related to the amount of visualized information is optimal, and it remains for future investigation with a larger sample set. In all three cases studied in Phase 2, the visual complexity was lower for suburban areas maps, which is caused by the naturally lower amount of spatial information in these areas – validating the algorithm's ability to pick up on the level of detail. Since all maps display the same area (represents the same reality, thus they contain the same amount of input information to be generalized) with the same cartographic symbology, the differences should be the same regardless the map provider. However, the biggest difference between the two areas was observed for OpenStreetMap ($\Delta_{FC}=3.77$) and the lowest for Google Maps roadmap ($\Delta_{FC}=1.79$). This can only be explained by different approaches of the cartographers (or the generalization algorithms) when selecting information to be displayed. It could either mean that in OSM the amount of information displayed in suburban areas are less dense compared to city center (which is plausible), or that Google Maps exaggerates suburban area information. Again, without having more knowledge about optimal FC, we could only scarcely explain this result.

| map | Feature Congestion | | |
|--------------|--------------------|---------------|------------|
| | city centre | suburban area | difference |
| GM roadmap | 5.25 | 3.46 | 1.79 |
| OSM | 8.76 | 4.99 | 3.77 |
| SchweizMobil | 13.36 | 10.87 | 2.48 |

Table 1. Visual complexity measured with Feature Congestion for maps of the city center and suburban area from various providers.

4. CONCLUSION AND FUTURE PLANS

The larger aim of our project is to test a hypothesis that visual complexity measures will overall reflect the level of cartographic generalization. In this paper, we specifically hypothesized that visual complexity should be constant across generalization levels; however we found that the complexity of large scale (low detail) displays was higher than those of small scale (high detail) ones as expressed with the FC metric. We also observed vast differences of visual complexity among various maps providers, which we attribute to their different approach towards the cartographic design and generalization process. Previously it was found that perceived visual complexity might be correlated to algorithmically measured complexity (Schnur et al. 2010). Even though people can assess and report the level of visual complexity quite good by mere intuition, the Feature Congestion metric can be helpful to automatically reveal poor results of generalization or cartographic design over large sample of data. We will continue to test our hypotheses with other existing visual complexity computations (e.g., Ciolkosz-Styk & Styk, 2013; Harrie & Stigmar, 2009, Krejtz et al., 2014) and assess their suitability for various geographical displays, and thus we will contribute to the understanding of their usefulness in a geographic context. We also wish to gain more knowledge about relation between visual complexity and usability of geographical displays, thus we plan to perform series of user studies.

REFERENCES

- Batty, M., Morphet, R., Masucci, P., & Stanilov, K. (2014). Entropy, complexity, and spatial information. *Journal of Geographical Systems*, 16(4), 363–385. doi:10.1007/s10109-014-0202-2
- Brassel, K. E., & Weibel, R. (1988). A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Science*, 2(3), 229–244.
- Brychtová, A., & Çöltekin, A. (2014). An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. *The Cartographic Journal*, 1743277414Y.000. doi:10.1179/1743277414Y.0000000103
- Brychtová, A., & Çöltekin, A. (2016). The effect of spatial distance on the discriminability of colors in maps. *Cartography and Geographic Information Science*, 1–17. doi:10.1080/15230406.2016.1140074
- Ciolkosz-Styk, A., & Styk, A. (2013). Advanced image processing for maps graphical complexity estimation. In *ICC2013*.
- Çöltekin, A., Reichenbacher, T. (2011). High Quality Geographic Services and Bandwidth Limitations. *Future Internet*, 3(4), 379–396. doi:10.3390/fi3040379
- Google. (n.d.) [various roadmap, terrain, hybrid and satellite displays]. Retrieved March 29, 2016, from <http://maps.google.com>
- Harrie, L., & Stigmar, H. (2009). An evaluation of measures for quantifying map information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(3), 266–274. doi:10.1016/j.isprsjprs.2009.05.004
- Krejtz, K., Duchowski, A. T., & Çöltekin, A. (2014). High-Level Gaze Metrics From Map Viewing: Charting Ambient/Focal Visual Attention. In *Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research, GIScience2014*. Vienna, Austria.
- MacEachren, A. M. (1982). Map Complexity: Comparison and Measurement. *Cartography and Geographic Information Science*, 9(1), 31–46. doi:10.1559/152304082783948286
- Moacdieh, N., & Sarter, N. (2015). Display Clutter: A Review of Definitions and Measurement Techniques. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(1), 61–100. doi:10.1177/0018720814541145
- OpenStreetMap. (n.d.) [various map displays]. Retrieved March 29, 2016, from <http://www.openstreetmap.org/>
- Pászto, V., Brychtová, A. & Marek, L. (2015). Modern Trends in Cartography, (Brus, J., Vondrakova, A. and Vozenilek, V.,Eds.) Modern Trends in Cartography, Lecture Notes in Geoinformation and Cartography, Springer International Publishing, Cham, Lecture No., doi:10.1007/978-3-319-07926-4.
- Rosenholtz, R., Li, Y., Nakano, L., Yuanzhen, L., Nakano, L., Li, Y., ... Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, 7(2), 17 (pp.1–22). doi:10.1167/7.2.17
- Ruas, A. (2008). Map Generalization. In *Encyclopedia of GIS*, Springer, ISBN 978-0-387-35973-1, 631–632.
- Schnur, S., Bektas, K., Salahi, M., Çöltekin, A. (2010). A Comparison of Measured and Perceived Visual Complexity for Dynamic Web Maps. In *GIScience 2010*.
- Shea, K. S., & McMaster, R. B. (1989). Cartographic generalization in a digital environment: When and how to generalize. In *AutoCarto*, 56–67.
- SwitzerlandMobility. (n.d.). Map hiking in Switzerland. Retrieved March 29, 2016, from <http://map.wanderland.ch/>
- Weibel, R., & Editor, G. (1998). Computational Perspectives on Map Generalization. *Information Systems*, 314, 307–315.