An Efficient Method for Exploratory Data Visualization of Big Spatial Data on Commodity Hardware

Presentation of work originally published in Geoinformatica (2019)

Christian Beilschmidt[1], Michael Mattig[1], Thomas Fober[1], Bernhard Seeger[1]

Abstract: The exploratory and interactive visualization of big spatial data is becoming increasingly important in business, science, and many other application areas. In this paper, we discuss the Circle Merging Quadtree, an efficient method for aggregating and visualizing big spatial point data on commodity hardware.

Keywords: Data Visualization, Biodiversity Data Analytics, Big Spatial Data Analysis

1 Summary

The exploratory and interactive visualization of big spatial data is becoming increasingly important in business, science, and many other application areas. For example, the demanding challenges in biodiversity requires new data-driven approaches to extract the information from an increasing amount of available heterogeneous data sources. Because scientific tasks generally result in complex analytical workflows with the necessity of having a researcher in the loop, efficient methods for the visualization of intermediate results are of utmost importance. While many of these methods are designed for dedicated hardware, there is currently a tendency that scientists make use of mobile devices instead. Thus, commodity hardware like tablets and smartphones are becoming the target end-user device for both, user interaction and scientific visualization. Therefore, methods for data visualization should not only offer low latency results, but also adapt to the limitations of the underlying devices like network bandwidth, battery power and screen resolution.

In this paper, we discuss an efficient method for aggregating and visualizing big spatial point data on commodity hardware. Our Circle Merging Quadtree (CMQ) method [Be19] offers a low latency, and thus supports users to explore big spatial data in an interactive manner. For two-dimensional point collections, our method uses a small set of non-overlapping circles such that (i) they follow the distribution of the points, (ii) they represent the cardinality of the underlying point subset by the circle area, (iii) they reveal hot spots while simultaneously keeping outliers. Fig. 1 exemplifies the transformation from raw points to aggregated circles.

[1] University of Marburg, Faculty of Mathematics and Computer Science, Hans-Meerwein-Straße, Marburg, Germany, {beilschmidt,mattig,thomas,seeger}@mathematik.uni-marburg.de

doi:10.18420/inf2019_34

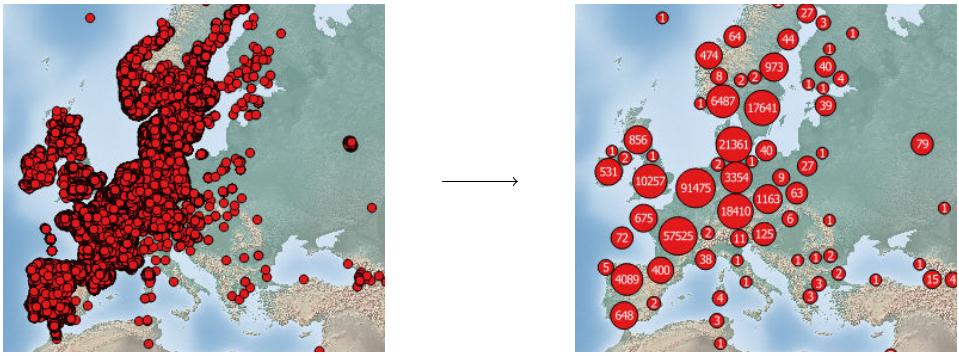

Fig. 1: This figure shows the discrepancy between mapping raw points (left) of the black alder and showing an aggregated view (right) of the points [Be17].

Based on a grid-based pre-aggregation step and the management of the resulting circles in a modified quadtree, our algorithm computes the final non-overlapping circles in linear time with respect to the number of points. Experimental results confirm its excellent runtime and quality in comparison to competitors. For instance, Jänicke et al. [JHS13], who provided one of the first algorithm to this problem that avoids overlaps, performs significantly slower in runtime complexity and practical performance.

2 Outlook

In geographic information systems, it is common to display multiple datasets as multiple layers on a map. This again leads to a cluttered representation with overlapping points and circles. As a future task, we will extend CMQ to process multiples point datasets at once into a joint representation. Furthermore, we aim at extending our method to compute aggregates of non-spatial attributes in the data.

References

[Be17] Beilschmidt, C.; Fober, T.; Mattig, M.; Seeger, B.: A Linear-Time Algorithm for the Aggregation and Visualization of Big Spatial Point Data. In: SIGSPATIAL '17. ACM, New York, NY, USA, 73:1–73:4, 2017.

[Be19] Beilschmidt, C.; Fober, T.; Mattig, M.; Seeger, B.: An Efficient Aggregation and Overlap Removal Algorithm for Circle Maps, GeoInformatica, 2019.

[JHS13] Jänicke, S.; Heine, C.; Scheuermann, G.: GeoTemCo: Comparative Visualization of Geospatial-Temporal Data with Clutter Removal Based on Dynamic Delaunay Triangulations. In: VISIGRAPP 2012. Vol. 359, Springer, Berlin, Heidelberg, Germany, pp. 160–175, 2013.