Eurographics Conference on Visualization (EuroVis), Posters Track (2016)
T. Isenberg and F. Sadlo (Guest Editors)

Coloring interactive compositional dot maps

M. Tennekes[1] and E. de Jonge[1]

[1] Statistics Netherlands

Abstract

We propose an algorithmic color scheme for zoom-able compositional dot maps. Contrary to existing methods, it uses density, composition, and zoom level to color the pixels of the resulting dot map. We describe the method and its application.

Categories and Subject Descriptors (according to ACM CCS): [Human-centered computing]: Geographic visualization—

1. Introduction

John Snows map of the 1854 cholera outbreak in London [CJ83] is a famous example of a dot distribution map. It shows that density of cholera deaths is correlated with the location of infected water pumps, by plotting the location of each casualty with a dot. The Racial Dot Map [Cab13] extends the dotmap by not only showing the US population density of census blocks, but also their ethnic composition: a color-coded dot for each person in the USA is plotted, revealing racial segregation patterns in US cities. While this map is aesthetically pleasing, the used color codes and alpha transparency seem to have been chosen arbitrarily.

We propose an algorithmic color scheme for plotting compositional dot maps, which takes density, composition, as well as zoom level into account. In a dot map, the location of each dot may be accurate, but more often is either jittered, to protect privacy, or randomly located within an administrative region, such as a census block or neighborhood. The latter option allows for drawing dot maps from administrative regional data, which is widely available. An important design choice for a dot map is whether each dot represents one unit or multiple units.

Compositional data [Ait82] are part-whole data: they denote for each category of a variable its fraction of the population e.g. female (sex) , child (age-group) or Asian (ethnicity). While small multiples with a dot map for each different category is a reasonable choice to show their distribution, it is very hard to see the geographical composition of the different sub populations. A compositional dotmap is designed to accommodate that.

2. Related work

The Racial Dot Map [Cab13] is the inspiration for the work in this paper. We are not aware of a scientific paper describing their method, but a web page with an overall description is available, as well as the source code of an implementation of their algorithm. A dot is plotted for each person in a census block of the US with fixed alpha-transparency α_z per zoom level. In the resulting map, the color of each pixel depends on the drawing order: later drawn dots weigh more. Therefore, color blending does not correctly reflects density nor composition. The choice of colors for the different races is undocumented and seems to be aesthetically motivated. The earlier segregation dot map of the NY Times [BCM10] and its update [BCG15] make different choices: each dot represents multiple persons N_z of same race. N_z depends on zoom level and uses rounded values 40, 60, 100, and 200. The algorithm to generate these values is not documented, but seems to prevent occlusion for the different dots. The resulting maps are coarse at every zoom level. [HSKIH07] provide a detailed paper including an user experiment on using dot maps for multivariate geographic data: plotting multiple variables in the same dot map. They compare different methods for combining the spatial distributions of two or more different variables: plotting each region with one blended color (blending) vs dots of different color within that region (weaving). Their finding is that weaving performed better then blending. Our method combines the two visualization approaches. Their experiment does not take composition into account, nor the (interactive) zoom level. Various research on color blending includes [CWM09, KGZ*12], describing hue preserving blending for nomimal data and [GC04, WGM*08] using perception based color blending for illustrative visualisation.

3. Method

We apply a colorization scheme to cartographic bitmap tiles. By convention, such interactive tiles are 256 by 256 bitmap images, which are generated for about 20 different zoom levels. The lowest zoom level, 0 , displays the whole world in one tile. Both dimensions are doubled at each zoom level. For a compositional dot map, we start with detailed base-level z^b in which a location is retrieved or generated for each person/unit in the population. Each pixel of a tile at zoom level z^b is assigned a tuple $(n_1,\ldots,n_k)$ with n_i the number of persons of category i in that pixel and k the number of

DOI: 10.2312/eurp.20161132

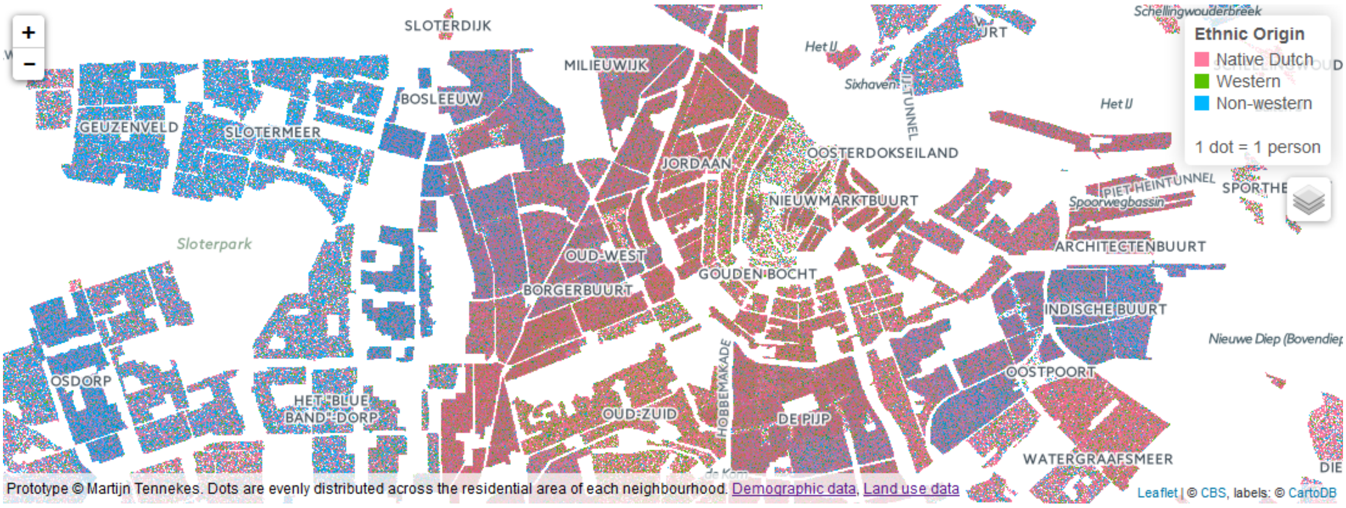

Figure 1: Interactive compositional dot map (zoom level 13).

different categories. A pixel of a tile in level z^a with $a < b$ has a tuple which is the sum of its corresponding pixels at base level z^b. Note that for practical purposes, the user zoom levels are typically restricted to levels that are relevant for the topic displayed. This setup generates a map with dots at the detailed level z^b, which are gradually transformed into color blended maps at lower zoom levels. Since dots at base level z^b may be hard to observe, especially on small devices with high resolution, tiles are generated for level z^c, with $c > b$, by zooming in from z^b without anti-aliasing.

The colors used for pixels are generated from the HCL color space, which is known for its well-balanced perceptual properties [Iha03, ZHM09]. The HCL color space model is depicted in Figure 2, where hue $\mathsf{H} \in [0,360]$ is the angle, chroma $\mathsf{C} \in [0,100]$ the radius, and luminance $\mathsf{L} \in [0,100]$ the height.

For pixel p, its tuple $(n_1,\ldots,n_k)$ and density $N_p = \sum_i^k n_i$, the following color scheme is used:

- *Luminance* encodes the density N_p of a pixel

$$\mathsf{L}_p = 80 - 60(N_p/w \cdot \delta^{z^b - z})$$

  with w an upperbound to the density and δ the overall luminance for a zoom level z. The resulting L values are in the range $[80,20]$. Choosing $\delta = 1$ creates identical overall luminance at every zoom level z.
- *Hue* and *chroma* encode the composition of the categories. For each category, a point is allocated on the circlular cross section of the HCL space at luminance L_p, which is depicted in Figure 3. These points are equaly spread on the circle border. The values of H_p and C_p are determined by the weighted mean of the coordinates of these points with weights $(n_1/N_p,\ldots,n_k/N_p)$.
  For $k = 3$, the category points are located at the corners of the triangle depicted in Figure 3. The resulting mixed points will be allocated in the triangle shown in Figure 3. Pixels with a dominant category will have high chroma values, while pixels where all categories are equally present will be gray ($\mathsf{C}_p = 0$).
  For $k \leq 3$, this results in unique and reversible colors. For $k > 3$, this scheme seems usable when some regions have dominating category fractions. For cardinalities larger than 10, we expect usability to deteriorate.

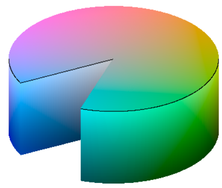

Figure 2: HCL color space.

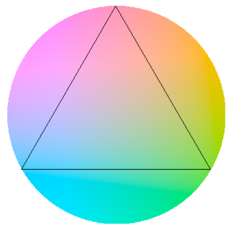

Figure 3: Cross section with constant L. For $k = 3$, colors are mixed within the triangle.

4. Application

Figure 1 shows a screenshot of the interactive compositional dot map, which displays ethnic origin of the population of the city of Amsterdam [Ten16]. The segregation between native Dutch people and people with a non-western background becomes apparent, although these sub populations are evenly mixed in a couple of neighborhoods. People with a western background are a minority, and predominantly live in the city center among native Dutch people.

5. Conclusions

The proposed method of coloring composition dot maps reflects both spatial density and composition of the population of interest. Further research is needed to tune the algorithm and its parameters. A user study is recommended to compare this method with the alternative methods described in paragraph 2.

References

[Ait82] AITCHISON J.: The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)* (1982), 139–177. 1

[BCG15] BLOCH M., COX A., GIRATIKANON T.: Mapping seggration. *The New York Times* (2015). 1

[BCM10] BLOCH M., CARTER S., MCLEAN A.: Mapping america: Every city, every block. *The New York Times* (2010). 1

[Cab13] CABLE D.: The Racial Dot Map. http://www.coopercenter.org/demographics/Racial-Dot-Map, 2013. 1

[CJ83] CAMERON D., JONES I. G.: John snow, the broad street pump and modern epidemiology. *International journal of epidemiology 12*, 4 (1983), 393–396. 1

[CWM09] CHUANG J., WEISKOPF D., MÖLLER T.: Hue-preserving color blending. *Visualization and Computer Graphics, IEEE Transactions on 15*, 6 (2009), 1275–1282. 1

[GC04] GOSSETT N., CHEN B.: Paint inspired color mixing and compositing for visualization. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on* (2004), IEEE, pp. 113–118. 1

[HSKIH07] HAGH-SHENAS H., KIM S., INTERRANTE V., HEALEY C.: Weaving versus blending: a quantitative assessment of the information carrying capacities of two alternative methods for conveying multivariate data with color. *Visualization and Computer Graphics, IEEE Transactions on 13*, 6 (2007), 1270–1277. 1

[Iha03] IHAKA R.: Colour for presentation graphics. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna Austria* (2003). 2

[KGZ*12] KÜHNE L., GIESEN J., ZHANG Z., HA S., MUELLER K.: A data-driven approach to hue-preserving color-blending. *Visualization and Computer Graphics, IEEE Transactions on 18*, 12 (2012), 2122–2129. 1

[Ten16] TENNEKES M.: Prototype colored dot map. http://research.cbs.nl/colordotmap/, 2016. 2

[WGM*08] WANG L., GIESEN J., MCDONNELL K. T., ZOLLIKER P., MUELLER K.: Color design for illustrative visualization. *Visualization and Computer Graphics, IEEE Transactions on 14*, 6 (2008), 1739–1754. 1

[ZHM09] ZEILEIS A., HORNIK K., MURRELL P.: Escaping RGBland: Selecting colors for statistical graphics. *Comput. Stat. Data Anal. 53*, 9 (2009), 3259–3270. 2