Hierarchical Color Palettes

Martijn Tennekes and Edwin de Jonge

Fig. 1. Treemap with hierarchical colors.

Abstract— Color is an important means to display categorical data in statistical graphics. Categories are often hierarchically structured in a classification tree, but most color palettes do not take this hierarchy into account. We present a method to map tree structures to colors from the Hue-Chroma-Luminance (HCL) color model. The HCL color space is known for its well balanced perceptual properties. Our study suggest that hierarchical qualitative color palettes are very useful: not only for improving standard hierarchical visualizations such as trees and treemaps, but also for showing tree structure in non-hierarchical visualizations.

Index Terms—Color palettes, statistical graphics, hierarchical data.

◆

1 INTRODUCTION

Hierarchical data are of crucial importance in official statistics. Most official data are published using hierarchically structured categories, for instance geographic regions or economic activities. Several data visualization methods are useful to explore and analyze hierarchical statistical data, for instance treemaps [4, 6]. Color palettes reflecting the hierarchical structure would be very useful in supporting visual analysis.

Assigning colors to categories is far from trivial. On the one hand, qualitative colors should be distinct, but on the other hand they should not suggest non-existent order or proximity and introduce perceptual bias. The selection of color palettes for categorical data first depends on the type of data. For nominal data, such as gender or nationality, qualitative color palettes are used, while for ordinal data, such as

- *Martijn Tennekes, Statistics Netherlands. E-mail: m.tennekes@cbs.nl.*
- *Edwin de Jonge, Statistics Netherlands. E-mail: e.dejonge@cbs.nl.*

level of urbanization, sequential or diverging palettes are used [1, 8]. However, for hierarchical categories there are no specific guidelines for selecting color palettes, to the best of our knowledge.

Although many tree visualizations are proposed in literature [3], most of them use color to a small extent. A visualization technique that uses color as a major attribute is the InterRing [7], a navigation tool with a radial layout. The leaf nodes are assigned to a different hue values. The color of a parent node is derived from averaging the colors of its children, where larger branches have more weight. An implicit effect is that colors of higher hierarchical levels are less saturated, except for one-child-per-parent branches.

We present a method to map tree structures to color palettes, which will be refered to as Hierarchical Color Palettes (HCP's). The used color space is the Hue-Chroma-Luminance (HCL) space, which is a transformation of the CIELUV color space, is designed with the aim to control human color perception [2]. Colors with different hue values are perceptually uniform in colorfulness and brightness, which does not hold for the popular Hue-Saturation-Value (HSV) color space [8].

This paper is outlined as follows. In Section 2 we describe the proposed method. We provide several applications of statistical graphics that use HCP's in Section 3. The conducted user survey to evaluate the method is described in Section 4. We conclude with a discusssion in

2 METHOD

Our method maps a tree structure on colors in HCL space, such that it reflects the hierarchical properties of the tree. We use H, with range $[0, 360]$, for the tree structure, where the hue of each child node resembles the hue of its parent. C and L, both with range $[0, 100]$, are used to discriminate the different hierarchical levels.

In this paper we illustrate our method with the European classification system of economic activity NACE. Figure 2 show a small part of the NACE: section F (Construction). Figure 2 is a typical tree visualization using a radial layout, but is improved by using nodes with colors from our hierarchical color palette.
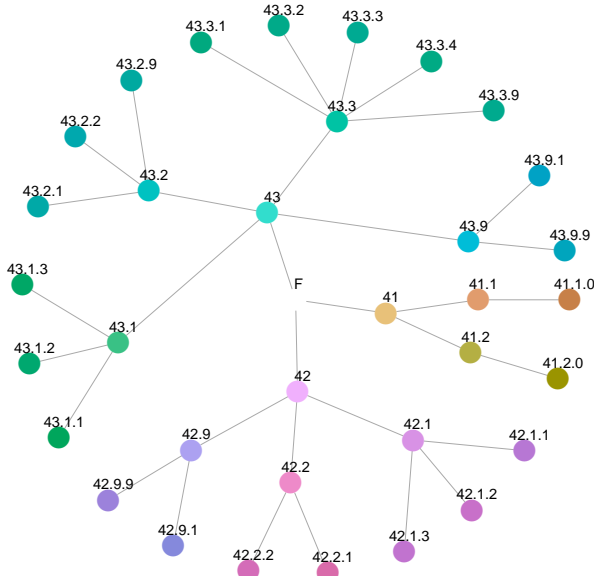
Fig. 2. Tree structure of economic sector F of NACE.

For selecting hue values we use the following recursive algorithm. It will assign to each node v of a tree structure a hue value h and a hue value range r. We start with the root node, which has by default hue range $[0, 360]$:

AssignHue(v, r)

1. assign the middle hue value in r to h [1]
2. N is number of child nodes of v, if $N > 0$:

   i   divide r in N equal parts r_i;

   ii   permute the r_i's and assign them to the child nodes

   iii reduce each r_i by keeping the middle fraction f;

   iv for each child node v_i DO AssignHue(v_i, r_i)

This division of the hue range is illustrated in Figure 3: in (a) the full hue range (for a constant $C = 60$ and $L = 80$) is divided and permuted among the three children of the root, in (b) the middle fractions are kept, in (c) and (d) these steps are recursively taken for the deepest two hierarchical layers.

Ad 2ii) In most hierarchical structures, there is no order between siblings. When the nodes in such structure are plotted in a linear or radial layout, the colors of the siblings should not introduce a perceptual order. Therefore, the assigned hue ranges are permuted among the siblings. The used permutation order is based on the five-elements-permutation $[1, 3, 5, 2, 4]$. Furthermore, the permutation within even numbered branches is reversed to differentiate between branches. Note the labeling of the color wheel that shows that the assignment of colors is permuted.

Ad 2iii) The fraction is needed to introduce a 'hue gap' between nodes with a different parent. The fraction f is by default set to 0.75. This choice is a trade-off between discriminating different main branches and discriminating different leaf nodes.

---

[1]In most cases the root node itself will not be drawn.

(a) Hue range equally split in three

(b) Middle fractions assigned to first layer nodes

(c) Recursively applied to second layer nodes

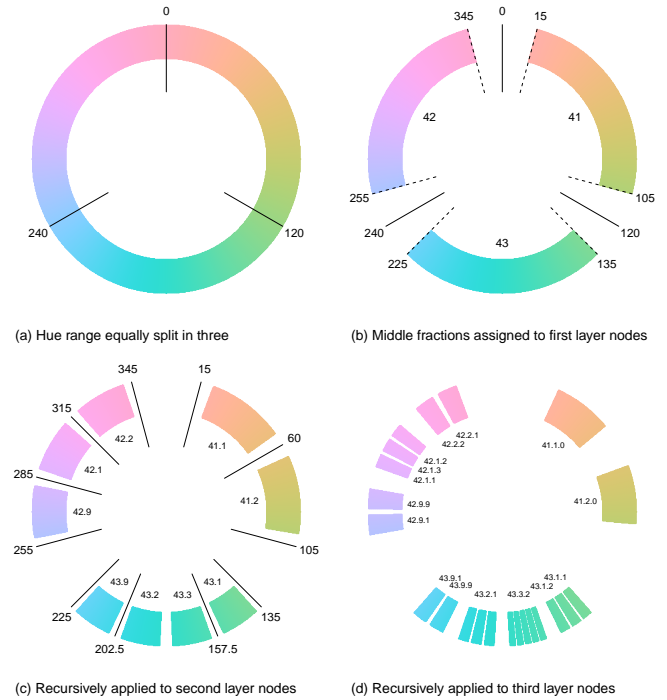(d) Recursively applied to third layer nodes

Fig. 3. Assignment of hue values.

In order to show depth, we let C and L values only depend on the depth of the corresponding nodes in the tree structure. We let the L decrease linearly with depth and C increase: having more intense colors helps in discriminating leaf nodes.
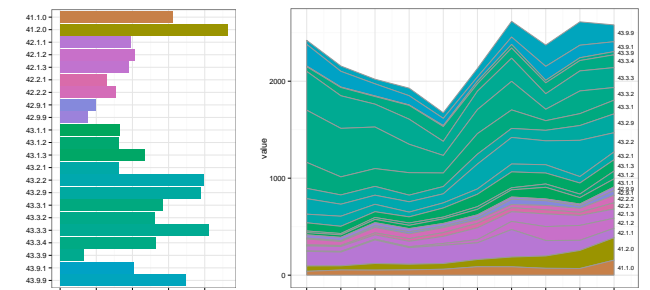

Fig. 4. Treemap with hierarchical colors

Fig. 5. Bar chart and stacked area chart with hierarchical colors

3 APPLICATION

The hierarchical colors can be applied to enhance standard tree visualizations, as we saw in Figure 2. Strictly speaking this is redundant color usage, but in our opinion it can improve many tree visualizations, because branches can be distinguished more easily.

A second example of improvement is depicted in Figure 4. It shows a treemap depicting (fictious) turnover values in Construction (NACE

F). In official statistics, turnover is available for each business enterprise in a business register, and aggregated according to the NACE tree. The hierarchical color palette is used to differentiate between different aggregated groups, that makes it possible to compare turnover values at different hierarchical layers. Although the colors of higher NACE layers are only used for the text label backgrounds, they are also resembled by the colors of the lowest NACE layers. This treemap is created with the free and open source R package treemap [5].

Hierarchical colors can also improve visualizations without explicit tree structure. The colors hint at the underlying tree structure. To illustrate this, a bar chart and a stacked area chart of fictive turnover data are depicted in Figure 5. Such graphics could be useful when the hierarchical structure will not be the main focus in the conducted analyses. The bar chart can be used to compare turnover values of all leaf node sectors in Construction (NACE F), and the stacked area plot to analyse turnover values in time.

4 USER STUDY

5 DISCUSSION

In our opinion, the proposed method to create hierarchical color palettes improves statistical visualizations, both hierarchically and non-hierarchically structured. The pre-condition that colors in the same hierarchical layer should be similar in terms of colorfulness and brightness it satisfied. This property is especially important in statistical visualizations, since they aim to visualize data as objectively as possible. The downside of the proposed method is that some leaf node colors will still be hard to distinguish.

We recommend a user study to evaluate the obtained hierarchical color palettes when applied in various statistical visualizations. The main aim of this user study would be to find out whether hierarchical palettes are useful in statistical analysis.

REFERENCES

[1] C. A. Brewer, G. W. Hatchard, and M. A. Harrower. Colorbrewer in print: A catalog of color schemes for maps. *Cartography and Geographic Information Science*, 30(1):5–32, 2003.

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

[3] H.-J. Schulz. Treevis.net: A tree visualization reference. *IEEE Comput. Graph. Appl.*, 31(6):11–15, 2011.

[4] B. Shneiderman. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Trans. Graph.*, 11:92–99, 1992.

[5] M. Tennekes. *treemap: Treemap visualization*, 2013. R package version 2.0.1.

[6] M. Tennekes and E. de Jonge. Top-down data analysis with treemaps. In *Proceedings of the International Conference on Information Visualization Theory and Applications, IVAPP 2011*, 2011.

[7] J. Yang, M. O. Ward, and E. A. Rundensteiner. Interring: An interactive tool for visually navigating and manipulating hierarchical structures. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, INFOVIS '02, pages 77–84, 2002.

[8] A. Zeileis, K. Hornik, and P. Murrell. Escaping rgbland: Selecting colors for statistical graphics. *Comput. Stat. Data Anal.*, 53(9):3259–3270, 2009.