Hierarchical Color Palettes

Martijn Tennekes and Edwin de Jonge

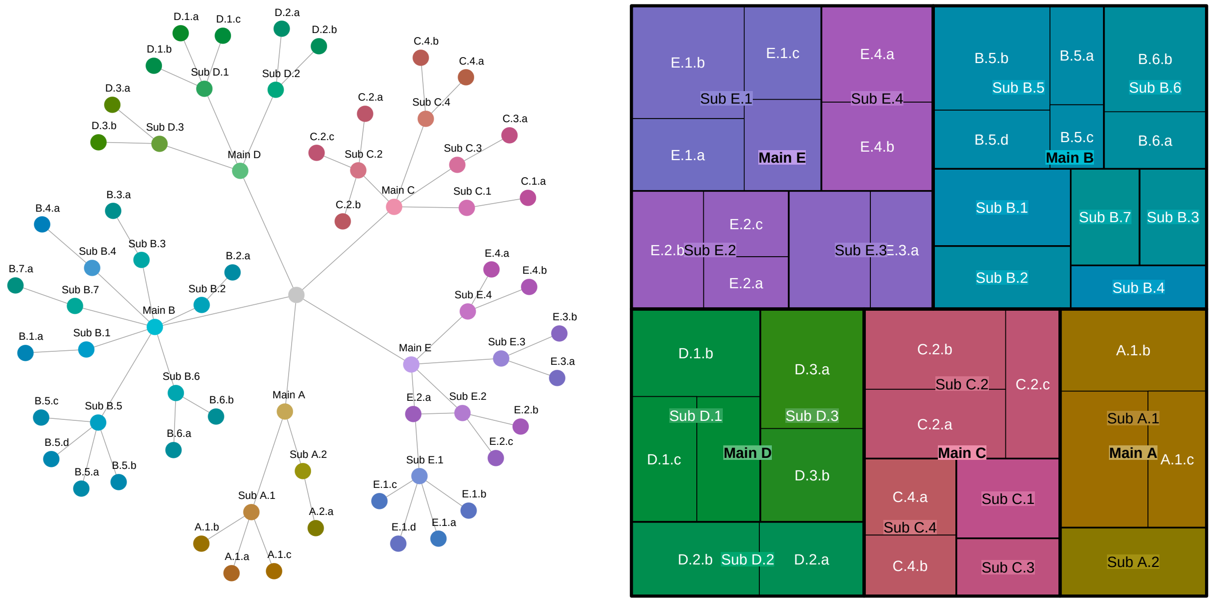


Fig. 1. Treemap with hierarchical colors.

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 [5, 7]. 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 level of urbanization, sequential or diverging palettes are used [1, 9]. 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 [4], most of them use color to a small extent. A visualization technique that uses color as a major attribute is the InterRing [8], 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. Hierarchical color schemes are also applied to the Hyperbolic Wheel [3], an exploration tool for hierchical data. These color schemes are abstracted from the Hue-Saturation-Lightness (HSL) space, where brightness decreases proportional from root to leaf nodes, and where child nodes inherit the hue values from their parent nodes and add small hue values to distinct them from their siblings. However, hue values of nodes in the same hierarchical layer may be overlapping.

Our purpose is to map tree structures to color palettes with the following properties. First, the hierarchical depth of a node should be perceived with color. The second property is that branches should be clearly distinguishable in all parts of the tree. In family trees, there is a specific order among siblings, namely by age. However, in the many hierarchical structures there is no such order apart from labeling purposes. Therefore, the third property is that the colors of sibling nodes should not introduce a perceptial order. Last, each node should have a theoretical unique color, that is also perceived that as unique as much as possible.

The color palettes that are obtained by our proposed method are called Tree Colors. As a starting point, we the Hue-Chroma-

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

Manuscript received 31 March 2013; accepted 1 August 2013; posted online 13 October 2013; mailed on 4 October 2013.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

Luminance (HCL) space, a transformation of the CIELUV color space, that 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) and HSL color spaces [9].

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 discussssion in Section 5.

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.

We illustrate our method with a tree structure that is depicted in Figure 2. The layout of the graph already highlights the tree structure, but the HCP colors of the nodes extra emphasize the tree structure in our opinion.

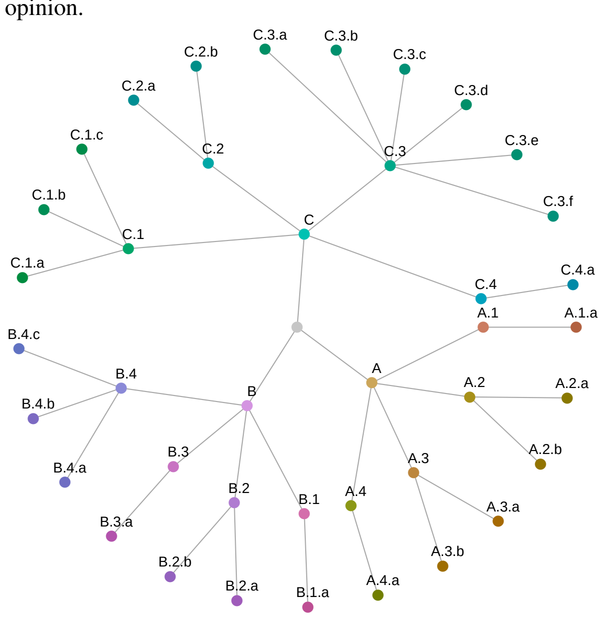


Fig. 2. Tree with HCP colors.

2.1 Hue values

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 $[0, 360]$:

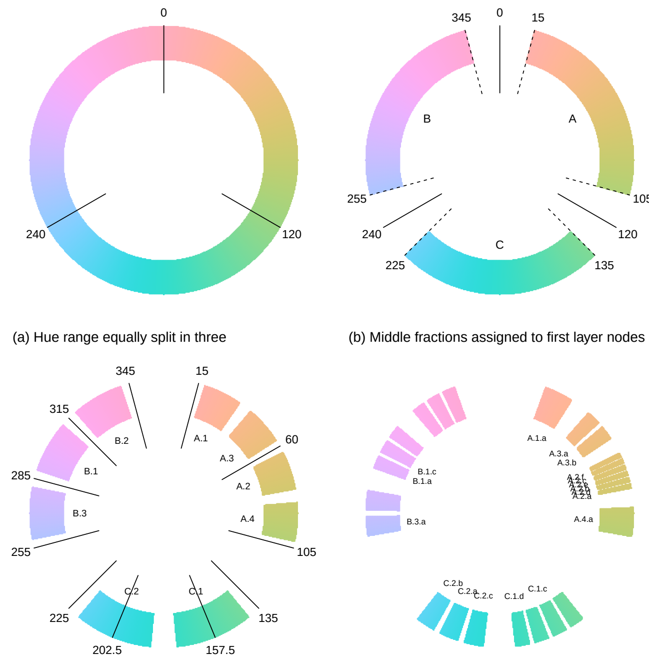
AssignHue(v, r)

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 $L = 70$) 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.

Permutations In most hierarchical structures, there is no order between siblings. When the nodes in such structure are plotted in a linear

[1]The root node itself is colored grey, so its hue is irrelevant.



(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.

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 $[1,3,5,2,4]$, where the distance between any two adjacent items in the original order is 2/5 of the full range in the permuted order. For the 3 and 4 siblings case we use the permutations $[1,3,2]$ and $[1,3,2,4]$ respectively. 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.

Hue fraction The fraction is needed to introduce a 'hue gap' between nodes with a different parent. This choice is a trade-off between discriminating different main branches and discriminating different leaf nodes. If $f = 0$, the hue ranges are diminished to single hue points, which implies that each main branch is assinged a constant hue. On the other end of the extreme, if $f = 1$, the full hue range is available at each hierarchical layer, which makes leaf nodes easier to distinguish, but harder to take apart from leaf nodes of different branches. The choice of f therefore depends on the application, on the used visualization method, and on the dimensions of the hierarchical data. We propose the following guidelines.

Representation	Children per node	3	4	5 or more
Implicit		0.75	0.60	0.50
Explicit		0.75	0.95	1

In the color wheel of Figure 3, the fraction f is set to 0.75.

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. The parameters for the chroma and luminance for the first hierarchical level below the root are by default set as $C_1 = 60$ and $L_1 - 70$. For each other hierarchical level $i \in \{2,\dots,k\}$, the chroma and luminance values are given by

$$C_i = (i-1)\beta^C + C_1 \tag{1}$$

and

$$L_i = (i-1)\beta^L + L_1. \tag{2}$$

By default, the slope parameters are set as follows: $\beta^C = 5$ and $\beta^L = -10$. Since the root node itself does not have a suitable hue value, its color is specified by $< H_0 = 0, C_0 = 0, L_0 = L_1 - \beta^L >$.

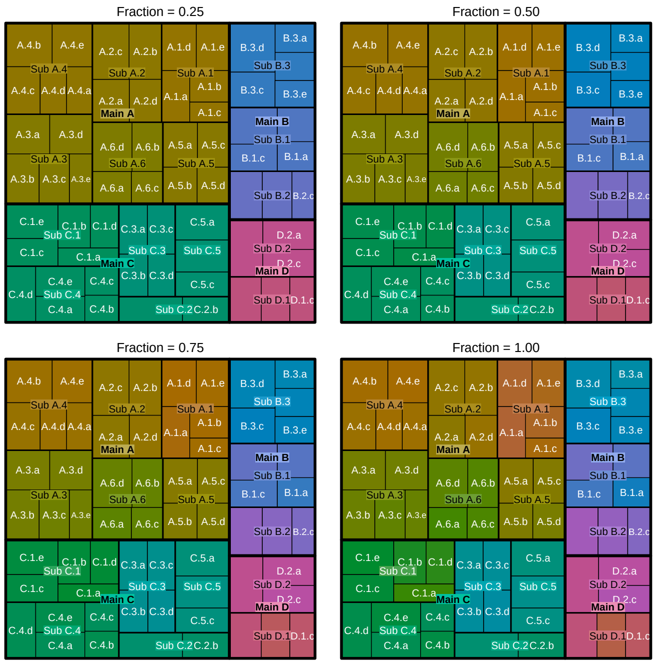


Fig. 4. Treemaps with different fraction values.

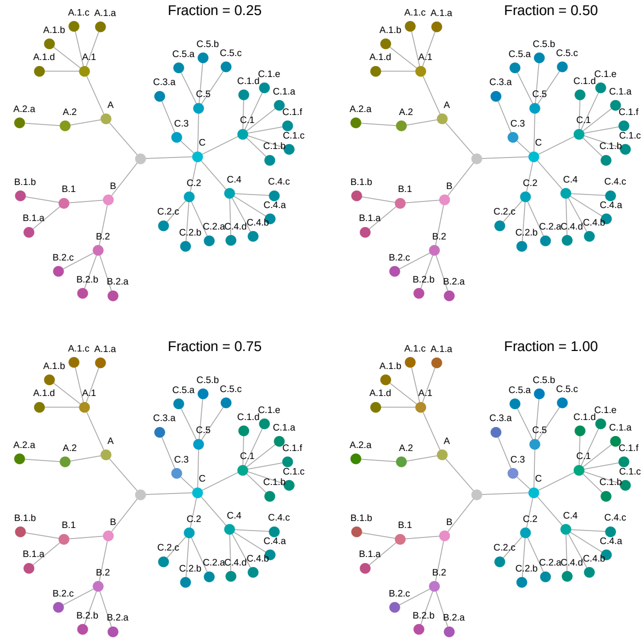


Fig. 5. Graphs with different fraction values.

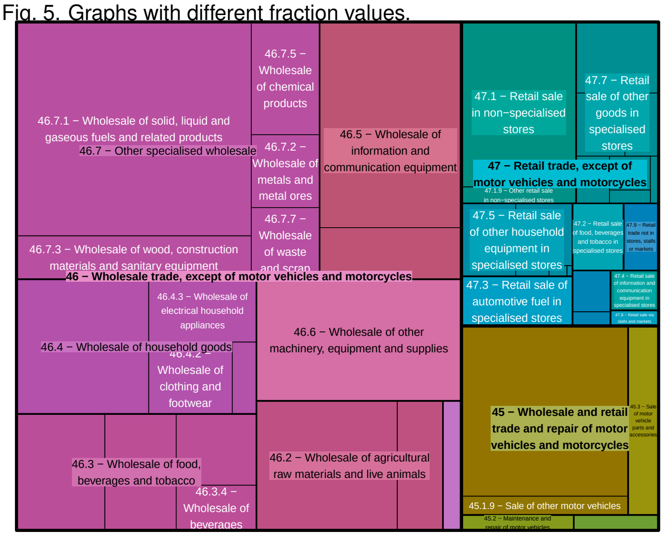


Fig. 6. Turnover among Dutch manufacturing enterprises in 2011

3 Application

The hierarchical colors can be applied to enhance standard tree visualizations, as we saw in Figure **??**. 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 **??**. 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 [6].

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 7. 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.

Acknowledgments

The authors wish to thank their collegues at Statistics Netherlands who participated in the user survey.

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.-C. Lam and A. I. . . p. . Ivo D. Dinov, vol. 2012. Hyperbolic wheel: A novel hyperbolic space graph viewer for hierarchical information content. *ISRN Computer Graphics*, 2012:Article ID 609234, 2012.

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

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

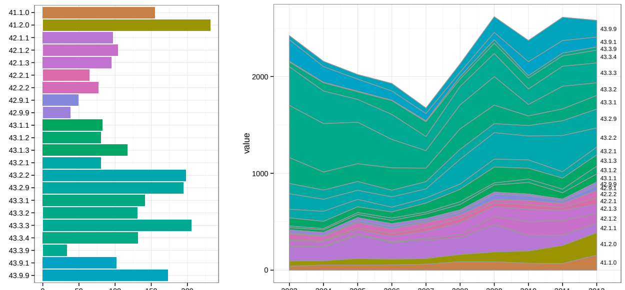


Fig. 7. Bar chart and stacked area chart with hierarchical colors

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

[7] 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.

[8] 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.

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