MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Equal-Height Treemaps for Multivariate Data

Wittenburg, K.B.; Lee, T.-Y.

TR2018-063 July 11, 2018

Abstract

A well-known limitation of classic continuous treemaps is that they generally provide two (or at most a few) visual mappings for data variables apart from the hierarchical relationships. Typically, one variable maps to cell area; another maps to color. However, many data-centric tasks require human users to consider multiple variables simultaneously. The current work introduces the concept of equal-height, variable-width cells in treemaps, which affords the packing of multiple variables into the cell areas of the terminals of the hierarchy. We demonstrate how color and some largely widthinvariant graphs can be utilized in the cell areas to add additional visual information in a multi-variate treemap. Examples come from machine learning and from finance applications.

International Working Conference on Advanced Visual Interfaces (AVI)

Equal-Height Treemaps for Multivariate Data

Kent Wittenburg
Mitsubishi Electric Research Laboratories, Inc.
201 Broadway
Cambridge, MA 02139
USA
kwittenburg@acm.org

Teng-Yok Lee
Mitsubishi Electric Research Laboratories, Inc.
201 Broadway
Cambridge, MA 02139
USA
tlee@merl.com

ABSTRACT

A well-known limitation of classic continuous treemaps is that they generally provide two (or at most a few) visual mappings for data variables apart from the hierarchical relationships. Typically, one variable maps to cell area; another maps to color. However, many data-centric tasks require human users to consider multiple variables simultaneously. The current work introduces the concept of equal-height, variable-width cells in treemaps, which affords the packing of multiple variables into the cell areas of the terminals of the hierarchy. We demonstrate how color and some largely width-invariant graphs can be utilized in the cell areas to add additional visual information in a multi-variate treemap. Examples come from machine learning and from finance applications.

CCS CONCEPTS

• **Human-centered computing~Visualization techniques**

KEYWORDS

Treemaps; information visualization; hierarchical data visualization; multivariate data visualization

1 INTRODUCTION

Treemaps have demonstrated their value over the years by intuitively visualizing hierarchical data. They indicate hierarchical relationships through graphical containment, show proportional distribution of one variable through graphical area, and show value of a second variable through color [5][7]. However, many data-centric problems involve the consideration of multiple variables simultaneously, and a general solution to multi-variate treemaps has not emerged despite a number of interesting proposals, e.g., [6][8]. The primary reason that classical treemaps do not lend themselves to multivariate visualization is that the rectangles at the leaves of the hierarchical trees are of varying widths and heights. Color is invariant to the aspect ratio of the containing rectangles, so color provides a natural mapping for one variable. Patterned fills have sometimes provided a second mapping. However, further regularization is needed for comparison across elements [3]. Researchers have proposed to regularize the shapes and positions of cells at the leaves of treemaps [1][8] and also insert 2D charts in constant sized leaf cells [3]. However, we have not encountered in research or in practice the proposal to fix one dimension of the leaf rectangles to accommodate multiple fixed-height colorable canvases while varying the other dimension to accommodate the area mapping.

Figure 1: A set of classic treemaps from finance. The variable "Value" maps to area in all 4 cases; a different variable maps to color in each case. (Courtesy of Google Sheets.)

To illustrate the basic problem with classic treemaps, we show in Fig. 1 a series of treemaps from a financial dataset where each treemap shares the same categorical hierarchy and the same area mapping from the variable "Value." They differ only in which variable defines the color mapping. Variables include gain/loss dollars and percentages, total and annualized, and the difference from the benchmark results for that asset class. Color mapping indicates distribution across the range of values with more saturated

AVI '18, May 29-June 1, 2018, Castiglione della Pescaia, Italy

ACM ISBN 978-1-4503-5616-9/18/05.
https://doi.org/10.1145/3206505.3206591

colors indicating the largest difference from the mean. An investor evaluating the performance of a portfolio would want to consider all the variables in relation to each asset. However, visual scanning and memory required to integrate individual asset performance information across rectangles from each of these four different treemaps is a cumbersome task. For example, in order to assess the performance of the single corporate bond holding, the largest rectangle, a user would have to find it in all four treemaps to realize that, although it has a total gain above the mean of other assets, other measures of performance are relatively negative.

2 EQUAL-HEIGHT TREEMAPS

Our proposal groups visual mappings of multiple variables belonging to terminal assets in the same tiered area. We realize the four treemaps from Fig. 1 as the single equal-height treemap in Fig. 2. Color, as before, indicates value relative to a reference of the variable distribution. We have chosen black as neutral and zero as the reference value, typical in finance applications.

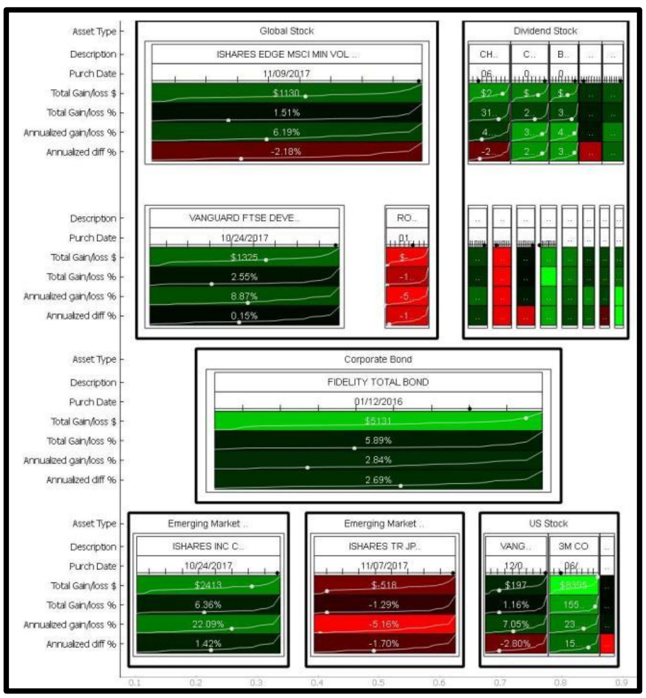

Figure 2: Equal-height treemap incorporating all information (and more) from investment data treemaps in Figure 1.

There are three advantages of our proposal: (1) it is easy to evaluate the terminal entities in a multi-variate sense since multiple attributes are rendered in the same graphical area (cf. corporate bond holding in Figs. 1 and 2); (2) comparison across entities per variable is made easy since the variable renderings are horizontally aligned in regularized positions across rows; and (3) width-invariant 2D graphics can be incorporated in cells that have the space. For instance, note the shapes of line graphs and relative point positions in Fig. 2. They are interpretable regardless of the width of the graph. Other graphics such as bar charts, timelines, or histograms could also be appropriate in the terminal cell areas [9].

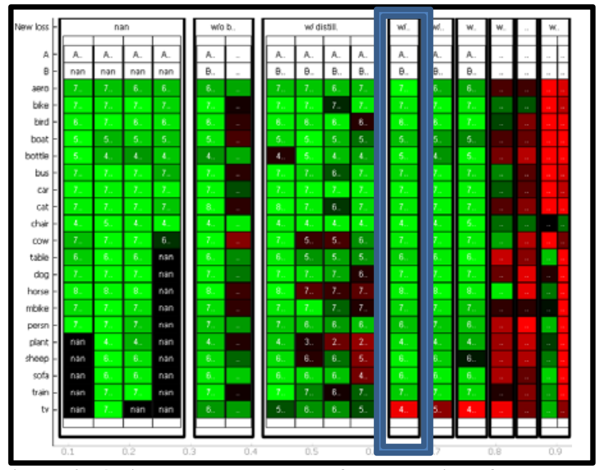

Figure 3: Object detector results from a series of test runs with different parameter regimes.

The primary disadvantage of our proposal is that it does not admit a classic continuous space-filling layout. Whitespace is inevitable if non-terminal rectangles remain whole. This problem is already addressed for fixed sizes for terminal rectangles in [3].

An additional example appears in Fig. 3 from the domain of machine learning. This example visualizes the published recognition results of an object detector utilizing different training regimes and parameters [4]. Each test run is shown as a terminal represented by a single column with stacked cells. The individual cells represent a series of variables, primarily class-specific recognition accuracy. Color represents accuracy of each object class relative to the mean. Overall accuracy of the trial run maps to terminal width, which does not vary greatly in most of the dataset. Different options in the training regimes determine the hierarchy, which is simple in this case; the entire 2-level hierarchy fits within a single row of terminals, something like a fish-eye table. Larger datasets would of course require additional rows.

Insights enabled from this visualization include that the trial runs of the left-most group are superior in general and also that the trial run shown in outline has the best but also the worst recognition results for the object classes at the top and at the bottom, respectively--an anomaly perhaps worth investigating. Although this example is not illustrative of the layout issues involved in larger datasets with more complex hierarchies, it still indicates that equal-height treemaps are capable of handling 10s of variables and that they have applicability in disparate domains.

3 CONCLUSION

The primary idea introduced here is to use fixed-height, variable-width rectangular areas for the terminal cells of a treemap. Multi-variate visualizations with fixed height can thereby be stacked and colored. The variable-width cells may also accommodate other graphics. The trade-off is that equal-height variable-width treemaps cannot in general fill the space completely. We are exploring algorithms that minimize whitespace.

REFERENCES

[1] Bederson, B., Shneiderman, B., and Wattenberg, M. 2002. Ordered and quantum treemaps: Making effective use of 2D space to display hierarchies. *ACM Trans. Graph.* 21, 4 (October 2002), 833-854. DOI=http://dx.doi.org/10.1145/571647.571649

[2] Bruls M., Huizing K., van Wijk J.J. (2000) Squarified Treemaps. In: de Leeuw W.C., van Liere R. (eds) *Data Visualization 2000. Eurographics*. Springer, Vienna.

[3] Schreck, T., Keim, D., and Mansmann, F.. 2006. Regular TreeMap layouts for visual analysis of hierarchical data. In *Proceedings of the 22nd Spring Conference on Computer Graphics* (SCCG '06). ACM, New York, NY, USA, 183-190. DOI=http://dx.doi.org/10.1145/2602161.2602183

[4] Shmelkov, K., Schmid, C., and Alahari, K. 2017. Incremental Learning of Object Detectors without Catastrophic Forgetting. *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, 2017, pp. 3420-3429.

[5] Shneiderman, B., 1992. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on Graphics*, 11,1, 92-99.

[6] Slingsby, A., Dyles, J, and Wood, J. 2008. Using Treemaps for Variable Selection in Spatio-Temporal Visualization. *Information Visualization 7, 3-4*, 210-224.

[7] Treemaps for space-constrained visualization of hierarchies: Including the History of Treemap Research at the University of Maryland. http://www.cs.umd.edu/hcil/treemap-history/ [Visited March 13, 2018]

[8] Vliegen, R., van Wijk, J. J., and van der Linden, E. 2006. Visualizing Business Data with Generalized Treemaps. *Visualization and Computer Graphics, IEEE Transactions on*, 12, 5 (September 2006), 789-796.

[9] Wittenburg, K., and Turchi, T. 2016. Treemaps and the Visual Comparison of Hierarchical Multidimensional Data. In *Proceedings of AVI'16*, June 7-10, 2016, Bari, Italy, 64-67.