Treemaps for space-constrained visualization of hierarchies

by Ben Shneiderman
Started Dec. 26th, 1998, last updated March 24th, 2005

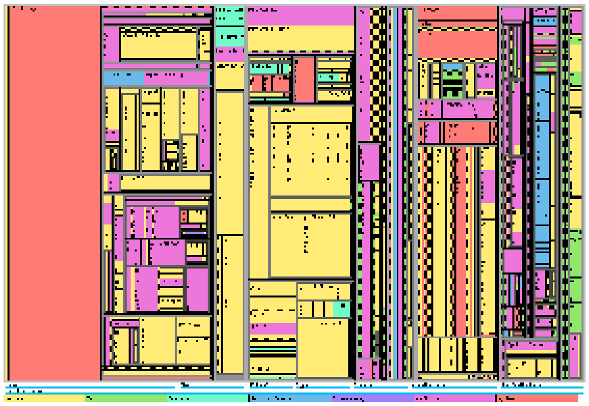
Our treemap products:

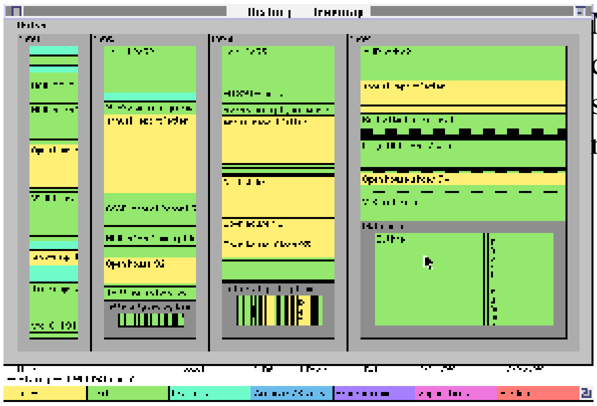
- Treemap 4.0: General treemap tool (Free demo version, plus licensing information for full package)
- PhotoMesa: Zoomable image library browser (Free demo version, plus licensing information for full package)
- Treemap Algorithms and Algorithm Animations (Open source Java code)

A History of Treemap Research at the University of Maryland

During 1990, in response to the common problem of a filled hard disk, I became obsessed with the idea of producing a compact visualization of directory tree structures. Since the 80 Megabyte hard disk in the HCIL was shared by 14 users it was difficult to determine how and where space was used. Finding large files that could be deleted, or even determining which users consumed the largest shares of disk space were difficult tasks.

Tree structured node-link diagrams grew too large to be useful, so I explored ways to show a tree in a space-constrained layout. I rejected strategies that left blank spaces or those that dealt with only fixed levels or fixed branching factors. Showing file size by area coding seemed appealing, but various rectangular, triangular, and circular strategies all had problems. Then while puzzling about this in the faculty lounge, I had the Aha! experience of splitting the screen into rectangles in alternating horizontal and vertical directions as you traverse down the levels. This recursive algorithm seemed attractive, but it took me a few days to convince myself that it would always work and to write a six line algorithm. This algorithm and the initial designs led to the first Technical Report (HCIL TR 91-03) in March 1991 which was published in the ACM Transactions on Graphics in January 1992 (http://www.acm.org/pubs/citations/journals/tog/1992-11-1/p92-shneiderman/). Choosing the right name took probably as long, but the term 'treemap' described the notion of turning a tree into a planar space-filling map.



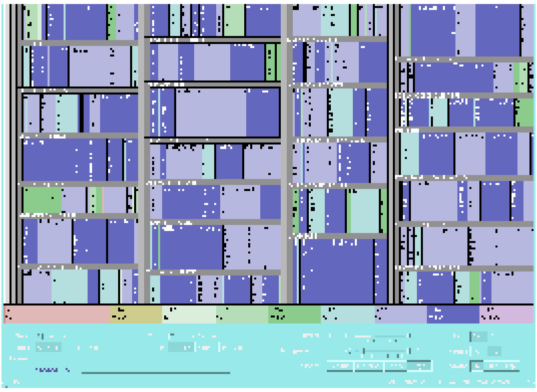


My initial design simply nested the rectangles, but a more comprehensible design used a border to show the nesting. Finding an effective visualization strategy took only a few months but producing a working piece of software took over a year. Brian Johnson implemented the algorithms and refined the presentation strategies while preserving rapid performance even with 5,000 node hierarchies. The TreeViz application ran on color Macintosh models and led to the widely cited paper (HCIL TR 91-06) jointly authored paper in the October 1991 IEEE Conference on Visualization. This paper was reprinted in Readings in Information Visualization. I think treemaps are a convenient representation that has unmatched utility for certain tasks. The capacity to see tens of thousands of nodes in a fixed space and find large areas or duplicate directories is very powerful. I still use TreeViz for cleaning up my Macintosh. It does take some learning for novices to grasp the tree structure layout in treemaps, but the benefits are great.

PhD student Brian Johnson's implementation added many other interesting features such as zooming, sound (as a redundant or independent code, for example, larger files had a lower pitched sound), hue/saturation control, many border variations, and labeling control. We struggled to deal with the problem of many small files in some directories, but wound up showing only a blackened area that invited closer examination by zooming. We knew that encoding a linear variable such as file size as an area was breaking a graphic design guideline, but the benefits of seeing a large range of file sizes seemed like a compensation. We also knew that visually comparing long narrow rectangles to squarish ones was problematic, but cursoring over the boxes produced the exact file size on the bottom of the display.

My excitement about treemaps was great and like many innovators I thought millions of users would be using this tool within a few years. Our minds were not focused on getting a patent, since I thought this was more of a concept that a product. Brian's implementation of TreeViz was registered with the University's Office of Technology Liaison which sought to distribute TreeViz.

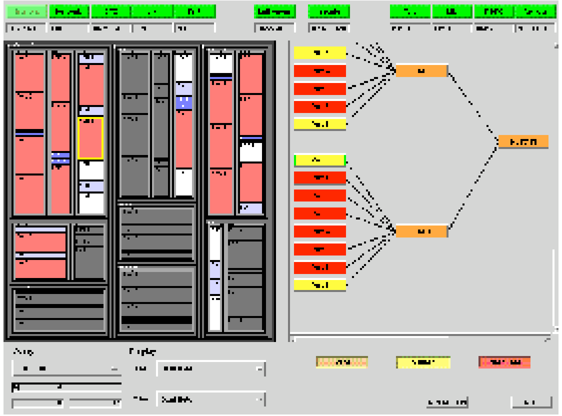
We found that new users took 10-15 minutes to get acquainted with the treemap display, so we began to explore improvements and training methods. We were impressed to examine thousands of nodes at 5-7 levels at once on the screen, but novices did better seeing 20-50 nodes at 1-3 levels. We had to bring our training times up to about 15 minutes in order to demonstrate the strong benefits of treemaps. Brian Johnson's dissertation (HCIL TR 94-04) reports on two studies, that were never published, that sort out the benefits of his treemaps. The 1992 HCIL Video Reports and the 1993 HCIL Video Reports showed TreeViz in action. TreeViz for Apple Macintosh is available for free downloading.



Masters student David Turo also built a treemap system on the Sun workstation and to make it more comprehensible we chose a fixed-level hierarchy. We used an appealing and familiar sports application: 453 basketball players, organized into the 27 teams in four leagues of the National Basketball Association. We had 48 statistics about each player for the 1991-92 season, so users could chose color and area coding from points scored, fouls, free throws, etc. . The 1993 HCIL Video Reports showed Turo's system with the basketball data. His Masters thesis (Unpublished!) describes his implementations and an empirical study. Johnson and Turo cooperated on a paper describing improvements they made to the visual presentation (HCIL TR 92-06).

By now we were pushing ahead on several application domains. A German visitor, Alexander Jungmeiseter worked with Dave Turo's implementation and built a stock portfolio visualization that showed clients, portfolios, industry groups, stocks and trades (HCIL TR 92-14) . Size might indicate worth of the holdings and color might indicate the degree of increase/decrease in value. I still believe that a worthwhile application would be a stock market monitor that would show the current daily trade activity. It could present the 30 Dow Jones Industrials, the Standard and Poor's 500, or all 2700 companies on the New York Stock Exchange. They would be grouped by industry (airlines, chemicals, drugs,...), area coded by volume of trading, and color coded by increase/decrease.

A Japanese visitor, Asahi Toshiyuki, built his own innovative treemap interface to implement the Analytical Hierarchy Process in decision making (HCIL TR 94-08). Users could express their opinions of the relative merits of a decision choice (such as which site to chose for a factory) by pumping up areas for their preferred choices, and pumping up the areas for importance of costs, availability of labor, tax breaks, etc. (Figure 5). The video demonstrates these processes (HCIL Video Reports 1994, and HCIL TR 95-04) and an empirical study showed users could succeed with this tool.



Another success story for treemaps was their inclusion in a satellite management system for Hughes Network Systems (HCIL Video Reports 1994 and HCIL Report TR 94-07). The three-level hierarchy showed each node of a network as a fixed size and color was used to indicate available capacity. The engineering-oriented community of ground station operators grasped this simplified version quickly.

In my travels to lecture about our work, treemaps became a major topic. However, I ran into resistance when showing still images of our hard disk directories with thousands of nodes. Once at the University of Washington after my talk produced a mixed reaction about treemaps, I asked my audience to follow me down to one of their labs. I installed TreeViz and examined their hard disk directories. I immediately spotted a problem, and with a few clues they could see for themselves that there were three copies of the same C compiler installed on this machine. The x-ray vision metaphor had proven to be effective on this occasion. Similarly, at Apple Computers, my audience much preferred the dynamic queries demos of the HomeFinder, but I gave copies of TreeViz to several interested attendees. The next day one of them reported finding many megabytes of useless information on their network servers.

While TreeViz was appreciated for the Apple Macintosh, we were getting requests for a Windows version. Graduate student Marko Teittinen took up the challenge and used Galaxy from Visix to produce a Windows 3.1 implementation called WinSurfer. Marko's carefully scripted video (HCIL Video Reports 1995) showed the features which were meant to match the Windows Explorer. WinSurfer allowed users to view, delete, copy, move, rename and run files. It worked nicely, but novices were often struggling to understand the layout that might show 5000 or more files at 7 or more levels. Simplifying the initial screen presentation to only 2 levels and allowing simpler user control never got implemented. In fact, we never produced a Windows 95 version and Galaxy is no longer a viable commercial product.

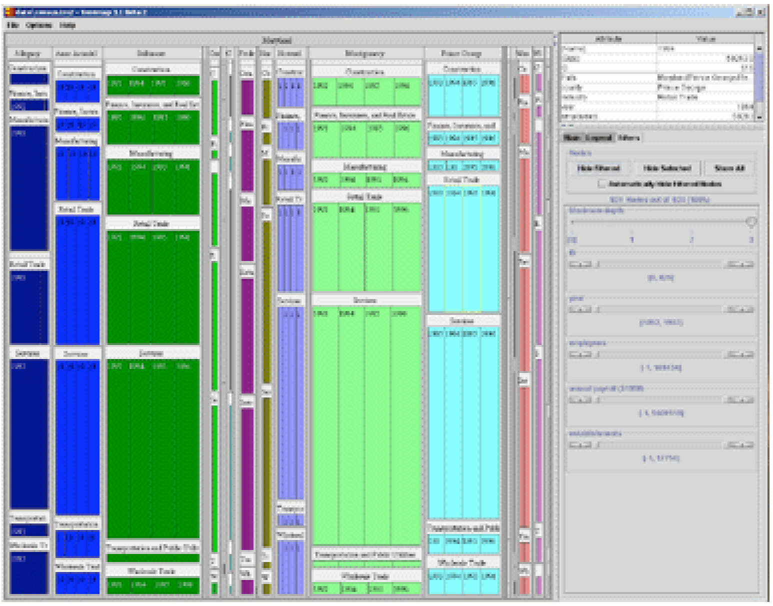
However, we did make a temporary patch so as to enter WinSurfer in the Browse-Off competition at ACM SIGCHI's 1997 Conference. This event drew six software tools for exploring hierarchies. There was no clear cut winner and WinSurfer was a leader in some tasks.

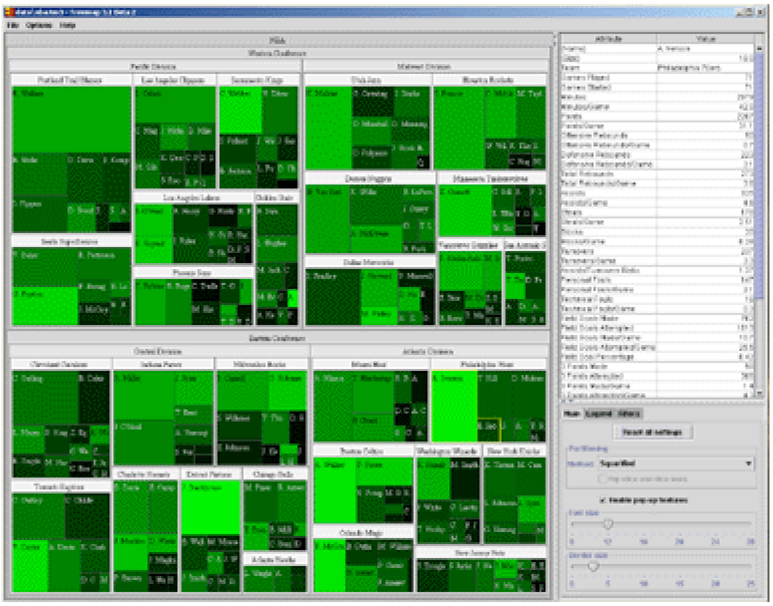
Micro Logic Corp, a New Jersey company, sells a commercial product, called DiskMapper http://www.miclog.com/dmdesc.htm for Windows machines based on the treemap idea. They have received great press attention and awards for their product. The University of Maryland receives a modest royalty on DiskMapper by way of a license agreement with the Office of Technology Liaison.

Statisticians point out that the mosaic display, shown by Bertin, and others is similar to the treemap concept. For fixed level hierarchies there is a great similarity, but the gist of the treemap idea was intimately tied to the computerized implementation and user control panel for setting attributes. A discussion of mosaic displays and an implementation that allows rapid switching of the hierarchy is in a Univ of Maryland graduate student project: CatTrees: Dynamic Visualization of Categorical Data Using Treemaps.

During 1997 an implementation was done in Delphi by Univ. of Maryland grad students Jerome Brown and Shaun Gittens. It is called TreeMap97 (http://www.otal.umd.edu/Olive/Class/Trees/). This general purpose version of slice-and-dice treemaps was revised by Chris North. You can download the software (updated version) and run the sample data set or import your own data sets.

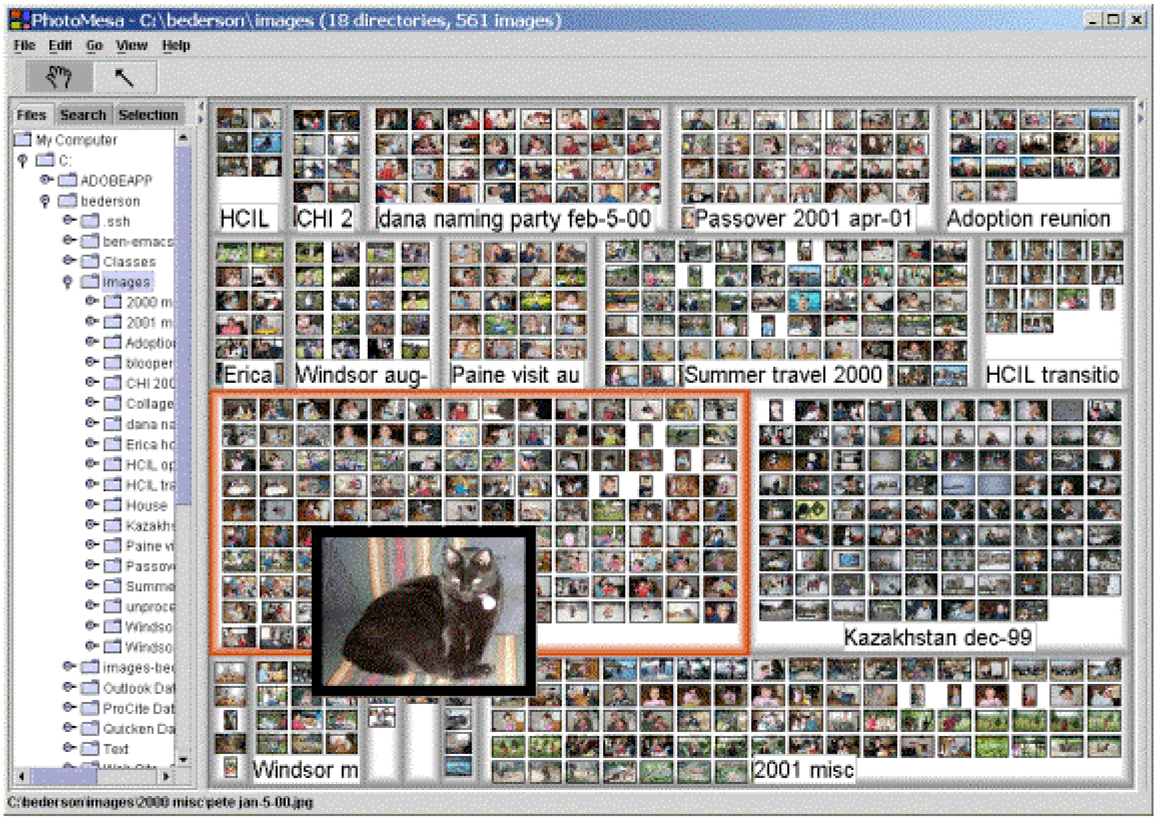
During Summer 2000, the HCIL resumed work on treemaps, when Raghuveera Chalasani developed a Java version (http://www.cs.umd.edu/hcil/treemaps/treemap2000) that included dynamic queries sliders. By early 2001 we polished his implementation into Treemap 3.0 http://www.cs.umd.edu/hcil/treemap3 which is licensed by the Univ. of Maryland's Office of Technology Liaison http://www.otc.umd.edu (screen shot). Treemap 3.0 includes dynamic queries to filter out unwanted items, squarified layout (as well as slice-and-dice), improved input, infotips, and color/font controls. A demo version with five data sets is available for free. A useful feature of Treemap 3.0 is that it also allows you to visualize the contents of your Windows PC directory structure.





The cluster and squarified treemap algorithms are visually appealing in part because the rectangles are more square like. They avoid the thin rectangles in the slice-and-dice algorithm, but sacrifice the alphabetic ordering of nodes. This creates an additional problem when leaf node sizes change because the position of rectangles can alter dramatically. The goal of ordered treemaps was achieved by a novel algorithm that nicely balanced square-like nodes (aspect ratios close to one) while preserving order. This paper, "Ordered Treemap Layouts" was presented at the October 2001 IEEE Symposium on Information Visualization.

However, we found still more ways to improve the layout by organizing the screen space into horizontal (or vertical) strips. This idea helped keep the square-like aspect ratio and made for an easier to follow ordering. We demonstrated this with a study of 20 users. Another innovation, quantum treemaps, satisfy the need to accommodate fixed shape items like page thumbnails or photos. The paper, "Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies" appeared in ACM Transactions on Graphics (October 2002, vol 21, no 4, pp 833-854) (http://portal.acm.org/citation.cfm?doid=571647.571649). The figure below shows a quantum strip treemap of a photo image browser, PhotoMesa.

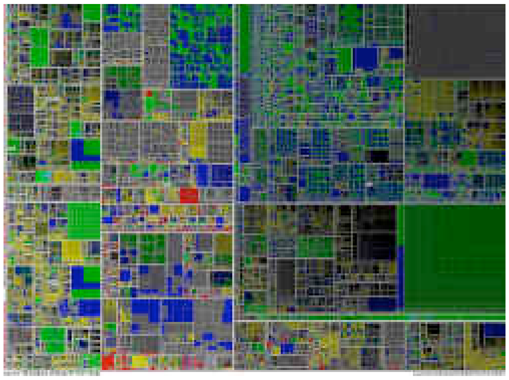


To support researchers and students, Ben Bederson and Martin Wattenberg wrote an open source Java 1.1 library of five Treemap Algorithms. They are generic, and should be readily usable by arbitrary applications. Each takes a list of numbers and an input rectangle, and generates a set of sub-rectangles that are proportional in area to the input numbers, and partition the space of the input rectangle.

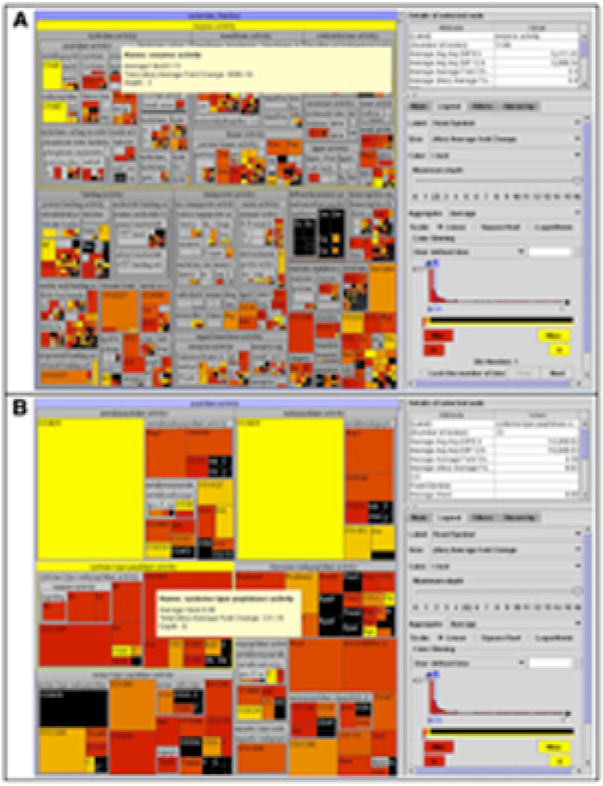
The 5 treemap algorithms implemented are:

- BinaryTree - Partially ordered, not very good aspect ratios, stable
- Ordered - Partially ordered, medium aspect ratios, medium stability
- SliceAndDice - Ordered, very bad aspect ratios, stable
- Squarified - Unordered, best aspect ratios, medium stability
- Strip - Ordered, medium aspect ratios, medium stability

This animated demonstration shows the strengths and weaknesses of these algorithms.



Jean-Daniel Fekete developed his treemap implementation to support large data sets of a million items and explore the uses of animation (http://www.cs.umd.edu/hcil/VisuMillion/).



Treemap 4, released in February 2003, includes a flexible hierarchy, color binning, improved color setting, aggregation, directory mapping, and export. Our applications include treemaps to show oil production data (SPE paper), project management (PMI paper), and a gene ontology with 14,000 nodes at 16 levels. Our work with biologist Eric Baehrecke led to a special version of Treemap to handle the Gene Ontology and gene expression data from microarray studies (see paper in BMC Bioinformatics 2004, 5:84) In August 2003, substantial code modifications were made to speed loading, permit use as an applet, and enhance usage of tm3 files. In addition, online documentation has been added in the usual textual format, and also as a set of 13 user selectable QuickDemos (animated with sound) that run from 11 to 75 seconds. A 9-minute video overview is also available.

Related Treemap Research and Development

John Stasko at Georgia Tech produced a nice X-Windows version, and his colleagues Sougata Mukherjea, James D. Foley, and Scott Hudson produced a ACM CHI95 paper that used treemaps: Visualizing Complex Hypermedia Networks through Multiple Hierarchical Views (http://info.sigchi.acm.org/sigchi/chi95/Electronic/documnts/papers/sm_bdy.htm)

Pedro Szekeley of University of Southern California's Information Sciences Institute cleverly built a quick and dirty version using his user interface building tools. Treemaps began popping up in surprising applications such as the visualization of a tennis match (Liqun Jin and David C. Banks, IEEE Information Visualization 1996), software module design (Marla J. Baker and Stephen G. Eick, 1995), and information search results in the Forager for the Information Super Highway (FISH) system (Mitre

Corp videotape) (http://www.mitre.org/resources/centers/it/g063/qt.html).

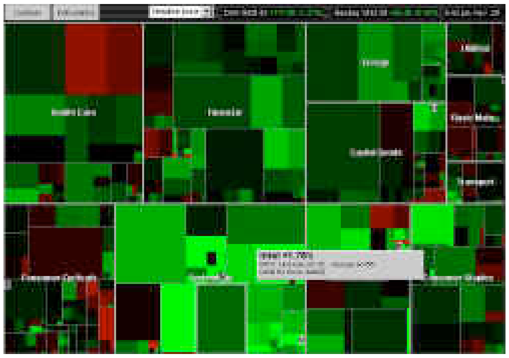
The SeeDiff software code viewer from Thomas Ball of Lucent Bell Labs organizes hierarchical programs by color coded treemaps: http://www.bell-labs.com/~eick/figs/seediff/vz-diff-2.gif

The Storyspace hypertext authoring system from Eastgate offers a treemap viewer http://www.eastgate.com

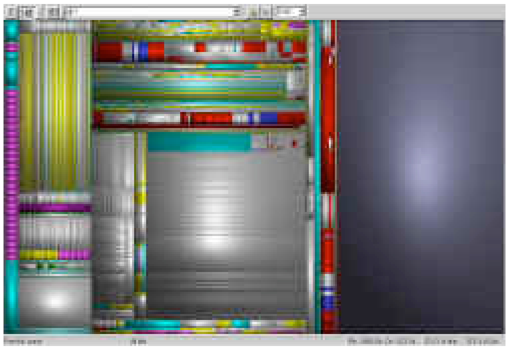
A well-written tutorial on treemaps by Chris Jones appears at http://orcs.bus.okstate.edu/jones98/treemaps.htm

A Tcl/Tk widget for treemaps is available at http://www-dse.doc.ic.ac.uk/~np2/itcl_widgets/treemap.html

A well done implementation of treemaps shows 535 popularly held stocks, organized by industry groups, size-coded by market capitalization, and color-coded to show rise or fall: http://www.smartmoney.com/marketmap/



This clever variation (created by Martin Wattenberg), which I call cluster treemaps, ensures low aspect ratio rectangles (most rectangles are square-ish and there are very few thin rectangles), but gives up the lexicographic ordering in slice-and-dice treemaps. Wattenberg describes his method in an ACM CHI99 short paper, in the Conference Companion. Marketmap was written in Java and the software is available for licensing. The SmartMoney website also offers a for-fee premium service, SmartMoney Select, with advanced treemap features such as dynamic query filters and other data search services. SmartMoney's software tools are marketed with demos at http://www.smartdatavision.com/.

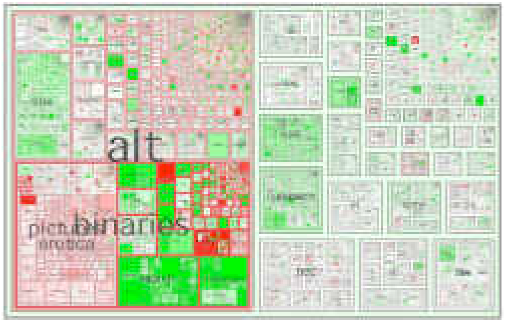


A visually intriguing variation was developed by the Visualization Group of the Technische Universiteit Eindhoven, headed by Jarke J. van Wijk. Their "cushion treemaps" show depth of nesting by shadows on cushion-like 3-D mounds. It was presented at the IEEE Symposium on Information Visualization, October, 1999: http://www.win.tue.nl/~vanwijk/ctm.pdf. They also created a new layout strategy called "squarified treemaps" that avoids high aspect ratio rectangles by using an alternative to Wattenberg's algorithm. This is presented at the Eurographics/IEEE TVCG 2000 Symposium: http://www.win.tue.nl/~vanwijk/stm.pdf. They developed an excellent free program, SequoiaView, to show your Windows PC file directories using cushion treemaps. This colorful representation reveals large files that waste space and duplicate directories that could be deleted.

Interesting treemap applications began sprouting during fall 2000. At TruePeers http://www.truepeers.com/ you can compare personalities with a cluster-style PeopleMap, and they include a special year 2000 election comparison among the 42 US Presidents. At Peets Coffee you'll find a neat treemap that shows their 32 coffees with area showing price and color (from light to dark) showing flavor http://www.peets.com/tast/11/coffee_selector.asp.

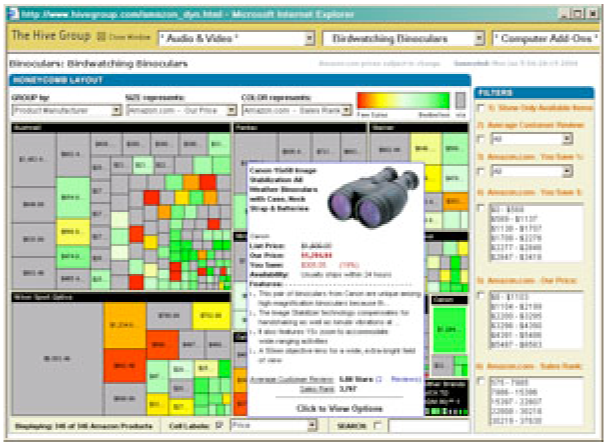
Another set of refinements to nested treemaps was made by Frederic Vernier in "Modifiable treemaps containing variable-shaped units" http://iihm.imag.fr/publs/2000/Visu2K_Vernier.pdf. His parametrized rectangles can be square-like or set to ensure that they are wider than high, facilitating the placement of readable text labels. Borders and labels for non-leaf nodes start out large at the root and shrink on descending in the tree. Also, when users click on a leaf node, the path to the root is highlighted.

The University of Maryland and Smartmoney organized a day-long treemap workshop on May 31, 2001 http://www.cs.umd.edu/hcil/about/events/open-house-2001/w2.shtml with 35 attendees to discuss the growing number of applications and implementations. Computerworld’s story, “Treemaps Bloom” described the excitement and innovations presented at workshop.



Marc Smith’s work at micro$oft Research applies treemaps in the Netscan project, http://netscan.research.micro$oft.com to data mine and visualize Usenet, one of the largest collections of social cyberspaces. Their initial goal is to provide an overview of the range of variation in Usenet and to highlight indicators and frequencies of different types of groups. Marc Smith and Andrew Fiore presented their work "Visualization components for persistent conversations" at ACM SIG CHI 2001. micro$oft Research makes .NET treemap components available.

Christophe Bouthier, a graduate student in Nancy, France, maintains a free Java library at http://treemap.sourceforge.net/.



The Hive Group (*www.hivegroup.com*), San Mateo, CA, develops and licenses *Honeycomb*, a treemapping software suite suitable for large-scale enterprise deployments. Honeycomb automates the development, deployment and management of treemaps including the definition and maintenance of users, roles, and security. The Hive Group's customers include Intel, Sun Microsystems, Merck, and The US Marine Corps. The Marine Corps Equipment Readiness Information Tool (MERIT) is used daily by 3000 people to manage logistics for material readiness. It has received numerous awards and commendations. See descriptions and news releases. The story was written up in Government Computer News, see: "Marine Corps logistics systems are better on MERIT". Another Hivegroup licensee, Matrikon, has built treemaps into their process control tools and reports on it in Finding a Needle in the Haystack- An Innovative Means of Visualizing Control Performance

Problems. A second paper has further examples and explanations: A Picture Worth a Thousand Control Loops. The HiveGroup software was also used in a controlled experimental evaluation by Joe Goldberg of Oracle (proc. HFES 2005) . He compared treemaps to tabular data for 8 tasks and found that treemaps were significantly faster for all tasks. His conclusion is that: "These results suggest that treemaps should be included as a standard graphical component in enterprise-level data analysis and monitoring applications."

The Smithsonian Institution developed a treemap exhibit during summer 2001, History Wired, with help from Smartmoney's Martin Wattenberg. The 450 items are clustered into groups such as home, clothing, business, computers,... and linked to attributes such as politics, medicine, and science. Users can click to get more details, search by attributes or filter by time period. This novel web site invites users to record their level of interest for items, which grow in size as they get higher scores.

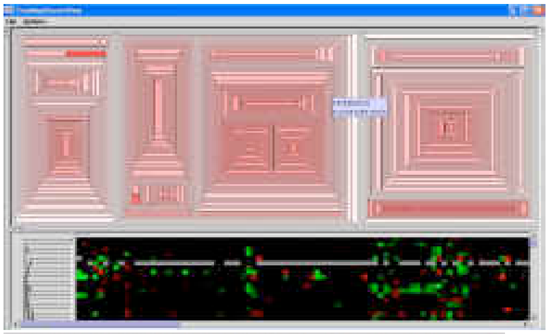
Lucernex, is promoting "Visualization software for program management" with its first success story in real estate development project management: http://www.lucernex.com/Walkthrough/programmap.htm.

For Java programmers, a tool called JavaTreeProfiler http://jcoverage.sourceforge.net/ shows execution time profiles in a zoomable treemap with area for each method and color coding for the method types.

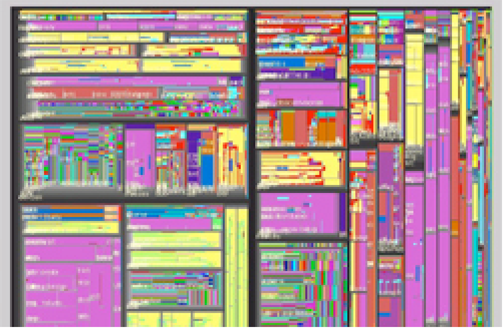
Panopticon (*www.panopticon.com*), Stockholm, Sweden, sells financial treemap software, and offers toolkits for Java, ActiveX and .Net developers.

The German Research Center for Artificial Intelligence uses treemaps in a decision support system, based on the Analytical Hierarchy Process, called Maut Machine. Their demo and research paper are available at *http://www2.dfki.de:8080/mautmachine/html/index.html.*

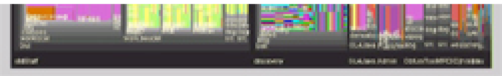
IBM Research developed a treemap version for network management, described in March/April 2003 issue of *IEEE Computer Graphics & Applications* "Treemaps for Workload Visualization".

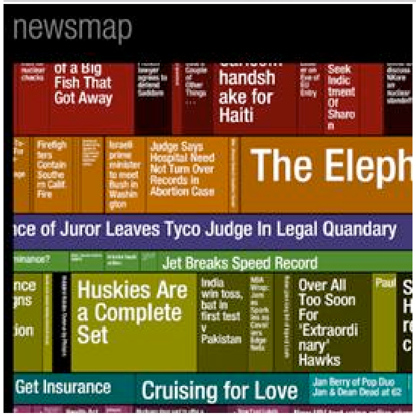


Patrick McConnell of Duke University Medical Center has applied treemaps for showing gene expression results from hierarchical clustering. His paper, "Applications of Tree-Maps to hierarchical biological data," appears in the journal *Bioinformatics*. And the software is available for TreeMapClusterView.



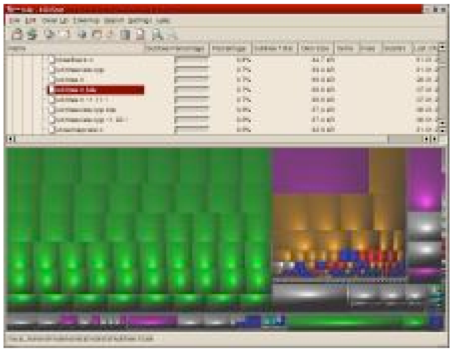
In October 2003, a major supplier of optimization and visualization components, ILOG , began offering a free test version of a major visualization tool kit called Discovery. This toolkit includes treemaps, and allows users to show any hierarchy, such as a directory of 50000 files, colored according to their file type.



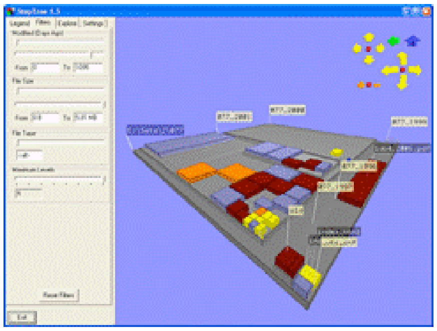

newsmap
of a Big Fish That Got Away
handshake for Haiti
Seek Indictment Of Sharon
Firefighters Contain Southern Calif Fire
Israeli prime minister to meet Bush in Washington
Judge Says Hospital Need Not Turn Over Records in Abortion Case
The Eleph
ice of Juror Leaves Tyco Judge In Legal Quandary
Jet Breaks Speed Record
Huskies Are a Complete Set
India win toss, bat in first test v Pakistan
Over All Too Soon For 'Extraordinary' Hawks
Get Insurance
Cruising for Love
Jan Berry of Pop Duo Jan & Dean Dead at 62

An innovative application by Marcos Weskamp in April 2004, was to take the Google News output and show it with a newsmap. Color codes the type of news story (sports, business, national, etc.) and size indicates how many news stories there are on this topic. The newsmap provoked a lively discussion at Slashdot

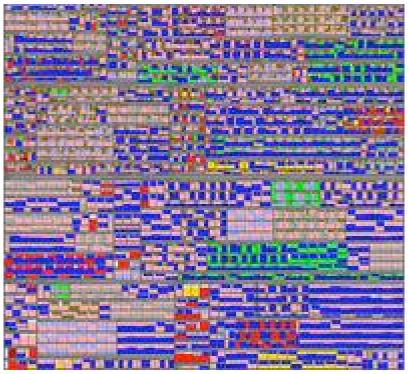
In 2003, Solvern Innovation licensed Treemap 4 from the University of Maryland and in 2004 is offering its own product called "Snowflake" based on Treemap.



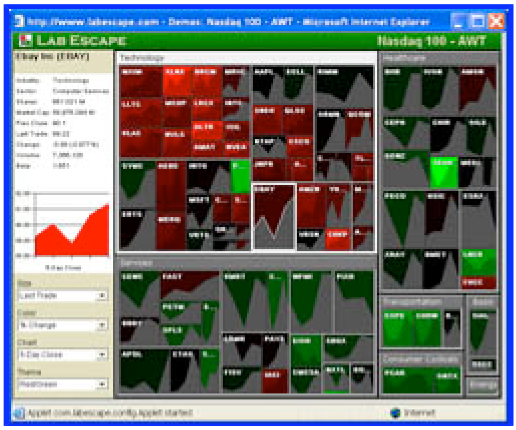
Stefan Hundhammer used the cushion treemap strategy to build KDirStat, a graphical disk usage utility, available (August 2003) at http://kdirstat.sourceforge.net/kdirstat.html



A group at Lulea University of Technology in Sweden developed a 3D treemap for file browsing that shows depth in tree as the height of steps (http://www.sm.luth.se/csee/csn/visualization/filesysvis.php) Their study with 20 participants showed benefits for the 3D layout for the task that asked to locate the deepest directory:

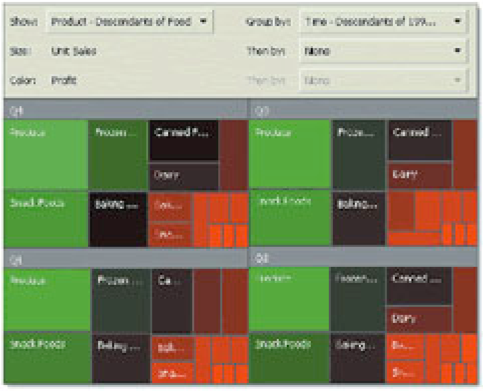


Bladh, T., Carr, D., Scholl, J. Extending tree-maps to three dimensions: a comparative study, *Proceedings of the 6th Asia-Pacific Conference on Computer-Human Interaction* (APCHI 2004), M. Masoodian, S. Jones, and B. Rogers, editors. Rotorua, New Zealand, June 29-July 2, 2004.



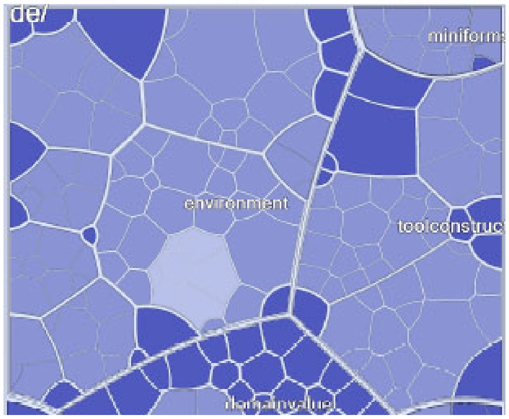
Jean-Daniel Fekete used his InfoVis Toolkit to generate several treemaps of the results of the April 2004 French elections in high resolution.
http://www.lri.fr/~fekete/elections2004

LabEscape's Tree Map 2.0 is sold to the financial community as a way to "explore data in new ways... discover impressive insights... make better decisions." One novel feature is to include stock price graph in the rectangle for each stock.



Show: Product - Descendants of Food
Group by: Time - Descendants of 199...
Size: Unit Sales
Then by: None
Color: Profit
Then by: None

ProClarity Performance Map is also based on Treemaps. See article in intelligence enterprise.



Voronoi Treemaps: To display software metrics Michael Balzer developed this variation of the Treemap algorithm which uses arbitrary polygons instead of rectangles. The layouts are computed by the iterative relaxation of Voronoi tessellations. See the SoftVis 05 paper.

taglet
util

- -.