Part 2: Color, Space, Statistical Graphics

Information Visualization Mini-Course
TECS Week 2008

Tamara Munzner

UBC Computer Science

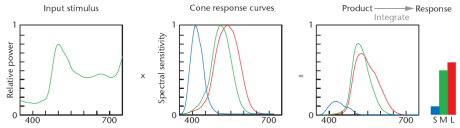
7 January 2008

Mini-Course Outline

- Part 1: Monday morning
  - Intro
  - Design Studies
  - Models
  - Perception and Memory
- Part 2: Monday afternoon
  - Color
  - Space, Layers, and Ordering
  - Statistical Graphics
- Part 3: Thursday afternoon
  - Multiples and Interaction
  - Navigation and Zooming
  - Focus+Context
- Part 4: Friday morning
  - High Dimensional Data
  - Graphs and Trees
  - User Studies

Trichromacy

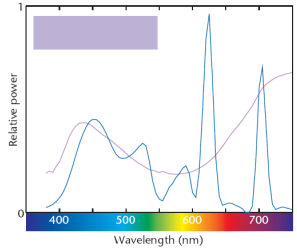
- different cone responses area function of wavelength
- for a given spectrum
  - multiply by response curve
  - integrate to get response



[Stone, Representing Color As Three Numbers, CG&A 25(4):78-85, www.stonesc.com/pubs/Stone%20CGA%2007-2005.pdf]

Metamerism

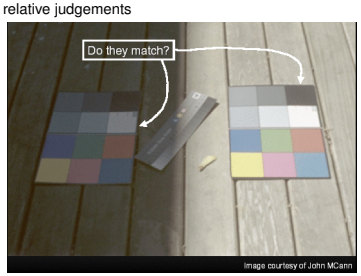
- brain sees only cone response
- different spectra appear the same



[Stone, Representing Color As Three Numbers, CG&A 25(4):78-85, www.stonesc.com/pubs/Stone%20CGA%2007-2005.pdf]

Color Constancy

- relative judgements



[courtesy of John McCann, from Stone 2001 SIGGRAPH course graphics.stanford.edu/courses/cs448b-02-spring/04cdrom.pdf]

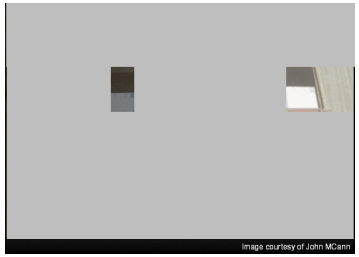
Color Constancy

- relative judgements

[courtesy of John McCann, from Stone 2001 SIGGRAPH course graphics.stanford.edu/courses/cs448b-02-spring/04cdrom.pdf]

Color Constancy

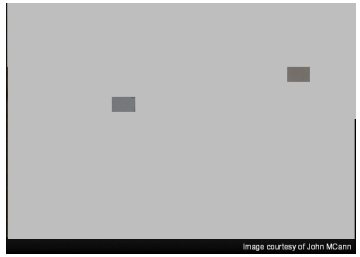
- relative judgements



[courtesy of John McCann, from Stone 2001 SIGGRAPH course graphics.stanford.edu/courses/cs448b-02-spring/04cdrom.pdf]

Color Constancy

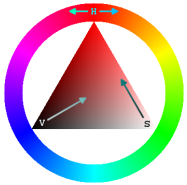
- relative judgements



[courtesy of John McCann, from Stone 2001 SIGGRAPH course graphics.stanford.edu/courses/cs448b-02-spring/04cdrom.pdf]

Color Spaces

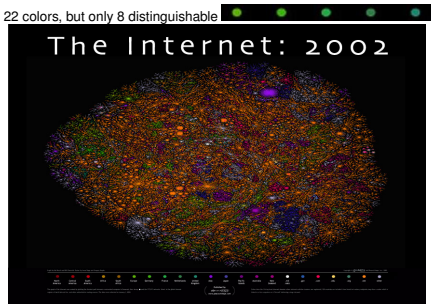
- HSV/HSB is more intuitive than RGB
- hue: color (dominant wavelength)
- saturation: amount of color vs. white
  - pink is less saturated than red
- value/brightness: amount of color vs. black
  - maroon is lower brightness than red



[http://upload.wikimedia.org/wikipedia/commons/1/1b/Triangulo_HSV.png]

Coloring Categorical Data

22 colors, but only 8 distinguishable



[www.peacockmaps.com, research.lumeta.com/ches/map]

Coloring Categorical Data

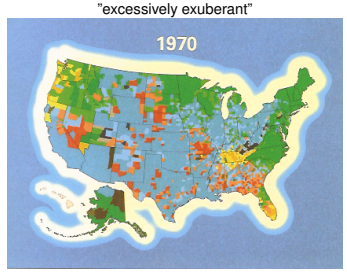
- discrete small patches separated in space
- limited distinguishability: around 8-14
  - channel dynamic range: low
  - choose bins explicitly for maximum mileage
- maximally discriminable colors from Ware
  - maximal saturation for small areas

[Colin Ware, Information Visualization: Perception for Design. Morgan Kaufmann 1999. Figure 4.21]

Minimal Saturation For Large Areas

- avoid saturated color in large areas

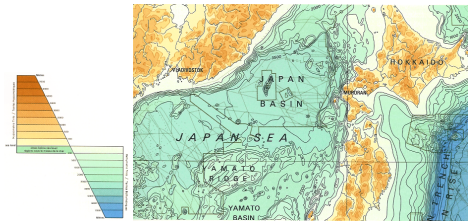
"excessively exuberant"



[Edward Tufte, Envisioning Information, p.82] [Colin Ware, Information Visualization: Perception for Design. Morgan Kaufmann 1999. Figure 4.20]

Minimal Saturation For Large Areas

- large continouous areas in pastel
  - diverging colormap (bathymetric/hypsometric)

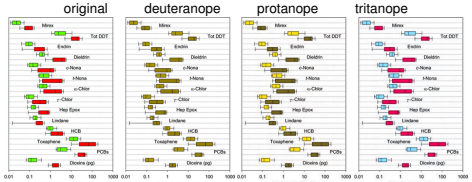


[Tufte, Envisioning Information, p. 91]

Color Deficiency

- deutanope
- protanope
  - has red/green deficit
  - 10% of males, so 5% of population!
- tritanope
  - has yellow/blue deficit
  - much less common
- http://www.vischeck.com/vischeck
  - test your images

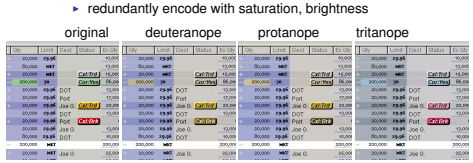
Color Deficiency Examples: vischeck



[www.cs.ubc.ca/~tmm/courses/cpsc533c-04-spr/a1/dmitry/533a1.html, citing Global Assessment of Organic Contaminants in Farmed Salmon, Hites et al, Science 2004 303:226-229.]

Designing Around Deficiencies

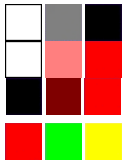
- red/green could have domain meaning
- then distinguish by more then hue alone
  - redundantly encode with saturation, brightness



[Courtesy of Brad Paley]

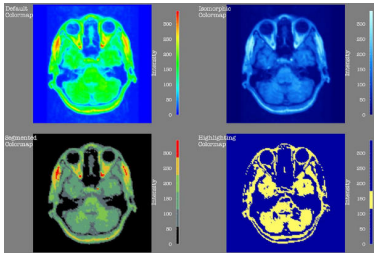
Coloring Ordered Data

- innate visual order
  - greyscale/luminance
  - saturation
  - brightness
- unclear visual order
  - hue



Choosing Colormaps

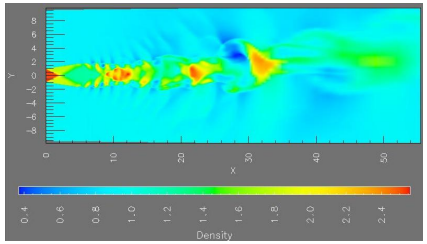
- rainbow popular but tricky



[Rogowitz and Treinish, How NOT to Lie with Visualization, www.research.ibm.com/dx/proceedings/pravda/truevis.htm]

Rainbow Colormap Advantages

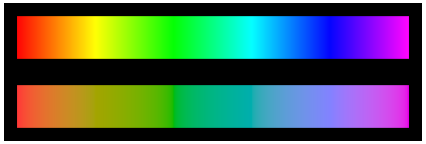
- low-frequency segmentation
  - the red part, the orange part, the green part, ...



[Rogowitz and Treinish, Why Should Engineers and Scientists Be Worried About Color? http://www.research.ibm.com/people/l/lloydt/color/color.HTM]

Rainbow Colormap Disadvantages

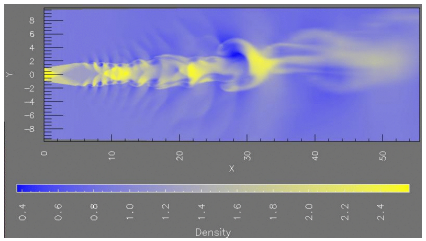
- segmentation artifacts
  - popular interpolation perceptually nonlinear!
- one solution: create perceptually linear colormap
  - but lose vibrancy



[Kindlmann, Reinhard, and Creem. Face-based Luminance Matching for Perceptual Colormap Generation. Proc. Vis 02 www.cs.utah.edu/~gk/lumFace]

Non-Rainbow Colormap Advantages

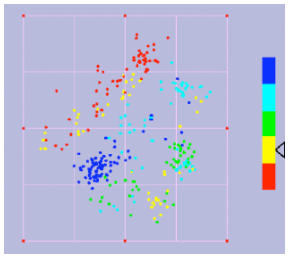
- high-frequency continuity
  - interpolating between just two hues



[Rogowitz and Treinish, How NOT to Lie with Visualization, www.research.ibm.com/dx/proceedings/pravda/truevis.htm]

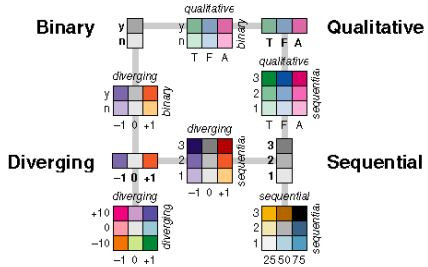
Segmented Rainbow

- explicit segmentation safer if need multiple bins



[Tory, Sprague, Wu, So, and Munzner. Spatialization Design: Comparing Points and Landscapes. IEEE TVCG 13(6):1262–1269, (Proc. InfoVis 07), 2007. webhome.cs.uvic.ca/~mtory/publications/infovis2007.pdf]

Cartographic Color Advice, Brewer



[Brewer, www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html]

Color: Readings

Representing Colors as Three Numbers, Maureen Stone, IEEE CG&A 25(4):78-85, Jul 2005. http://www.stonesc.com/pubs/Stone%20CGA%2007-2005.pdf

Information Visualization: Perception for Design. Chapter 4: Color. Colin Ware. Morgan Kaufmann, 2004 (2nd edition).

Edward Tufte, Envisioning Information. Chapter 5: Color and Information. Graphics Press, 1990.

How Not to Lie with Visualization, Bernice E. Rogowitz and Lloyd A. Treinish, Computers In Physics 10(3) May/June 1996, pp 268-273. http://www.research.ibm.com/dx/proceedings/pravda/truevis.htm

Color Use Guidelines for Mapping and Visualization. Cindy Brewer. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html

Color: Further Reading

A Field Guide To Digital Color, Maureen Stone, AK Peters 2003.

Information Visualization: Perception for Design. Chapter 3: Lightness, Brightness, Contrast, and Constancy. Colin Ware. Morgan Kaufmann, 2004 (2nd edition).

Face-based Luminance Matching for Perceptual Colormap Generation. Gordon Kindlmann, Erik Reinhard, Sarah Creem. IEEE Visualization 2002. http://www.cs.utah.edu/~gk/papers/vis02

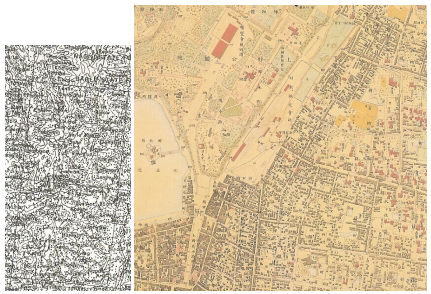
Color use guidelines for data representation. C. Brewer, 1999. http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/ASApaper.html

Why Should Engineers and Scientists Be Worried About Color? Bernice E. Rogowitz and Lloyd A. Treinish, http://www.research.ibm.com/people/l/lloydt/color/color.HTM

Mini-Course Outline

- Part 1: Monday morning
  - Intro
  - Design Studies
  - Models
  - Perception and Memory
- Part 2: Monday afternoon
  - Color
  - Space, Layers, and Ordering
  - Statistical Graphics
- Part 3: Thursday afternoon
  - Multiples and Interaction
  - Navigation and Zooming
  - Focus+Context
- Part 4: Friday morning
  - High Dimensional Data
  - Graphs and Trees
  - User Studies

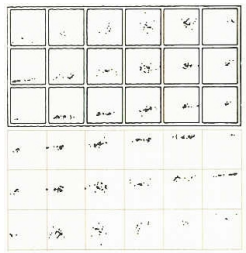
Layering And Separation



[Tufte, Envisioning Information, Chap 3]

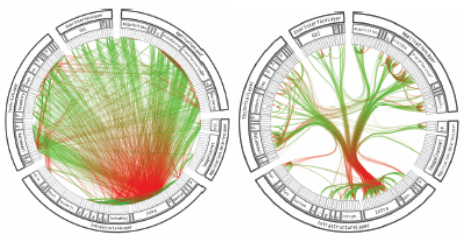
Visual Clutter

- subtler background than foreground



[Tufte, Envisioning Information, Chap 3]

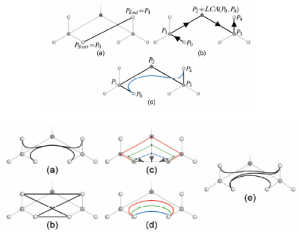
Hierarchical Edge Bundles



[Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten, Proc. InfoVis06.]

Hierarchical Edge Bundles

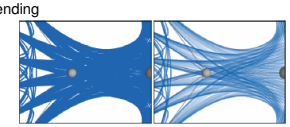
- bundle by hierarchy using splines



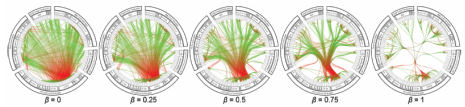
[Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten, Proc. InfoVis06.]

Hierarchical Edge Bundles

- alpha blending



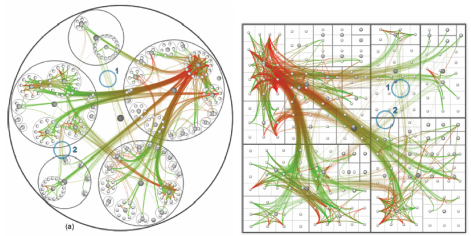
- bundling strength



[Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten, Proc. InfoVis06.]

Hierarchical Edge Bundling

- (mostly) agnostic to layout



[Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten, Proc. InfoVis06.]

Critique

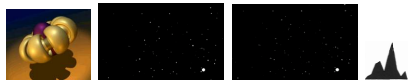
- flexible and general idea
- simple - after you see it
- successful example of creating foreground layer

Space vs. Time: Showing Change

literal ⟵ ⟶ abstract

time for time — space for time

- animation: show time using temporal change
  - good: show process
  - good: compare by flipping between two things
  - bad: compare between many things
    - interference from intermediate frames



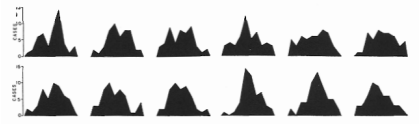
[www.geom.uiuc.edu/docs/outreach/oi/evert.mpg]
[www.astroshow.com/ccdpho/pluto.gif]

Space vs. Time: Showing Change

literal ⟵ ⟶ abstract

time for time — space for time

- small multiples: show time using space
  - overview: show each time step in array
  - compare: side-by-side easier than temporal
    - external cognition instead of internal memory
  - general technique, not just for temporal changes



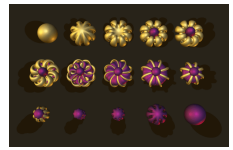
[Edward Tufte. The Visual Display of Quantitative Information, p 172]

Space vs. Time: Showing Change

literal ⟵ ⟶ abstract

time for time — space for time

- small multiples: show time using space
  - also can be good for showing process



[www.geom.uiuc.edu/graphics/pix/Video_Productions/Outside_In/postcard.comp.html]

Animation vs. Small Multiples

- Tversky argument: intuition that animation helps is wrong
  - meta-review of previous studies
  - often more info shown in animation view so not a fair comparison
  - carefully chosen segmentation into small multiples better than animation if equivalent information shown

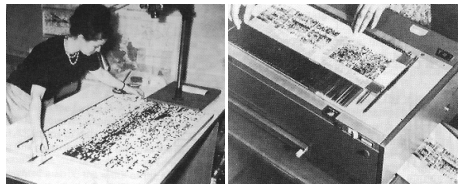
[Animation: Can It Facilitate? Barbara Tversky, Julie Morrison, Mireille Betrancourt. International Journal of Human Computer Studies 57:4, pp 247-262, 2002.]

Sorting and Ordering

- derived spaces for ordering
- spatial position as strongest perceptual cue
- finding the right order
  - automatically
  - through exploration

Manual Ordering: Bertin

- reorderable matrices - manually!



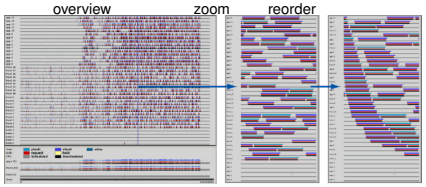
[Bertin, Graphics and Graphic Information Processing, p 34]

Interactive Ordering: Table Lens

- click to sort by columns
- also, is focus+context approach
- demo: www.inxight.com/products/sdks/tl

Interactive Ordering: Rivet

- performance analysis of parallel system
  - order: machine name vs. lock acquisition time



[Bosch, Performance Analysis and Visualization of Parallel Systems Using SimOS and Rivet: A Case Study, HPCA6, 2000. graphics.stanford.edu/papers/rivet_argus]

Space: Readings

Edward Tufte, Envisioning Information. Chapter 3: Layering and Separation. Graphics Press, 1990.

Edward Tufte, Envisioning Information. Chapter 6: Narratives of Space and Time Graphics Press, 1990.

Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten, IEEE TVCG 12(5):741–748 (Proc. InfoVis 06), 2006. http://www.win.tue.nl/~dholten/papers/bundles_infovis.pdf

Barbara Tversky, Julie Morrison, Mireille Betrancourt. Animation: Can It Facilitate? International Journal of Human Computer Studies 57:4, pp 247-262, 2002.

Ramana Rao and Stuart K. Card. The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus + Context Visualization for Tabular Information. Proc SIGCHI '94, pp. 318-322.

Mini-Course Outline

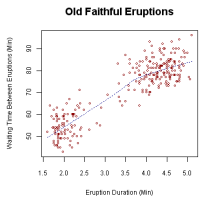
- Part 1: Monday morning
  - Intro
  - Design Studies
  - Models
  - Perception and Memory
- Part 2: Monday afternoon
  - Color
  - Space, Layers, and Ordering
  - Statistical Graphics
- Part 3: Thursday afternoon
  - Multiples and Interaction
  - Navigation and Zooming
  - Focus+Context
- Part 4: Friday morning
  - High Dimensional Data
  - Graphs and Trees
  - User Studies

Statistical Graphics

- long history for paper-based views of data
  - springboard for infovis
- interacting with scatterplots
  - interactive dynamic queries
  - matrix of scatterplots, level of indirection
  - linked views
- improving line charts

Scatterplots

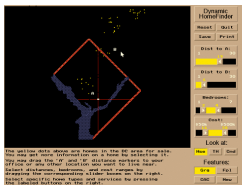
- encode two input variables with spatial position
- show positive/negative/no correllation between variables



[http://upload.wikimedia.org/wikipedia/commons/0/0f/Oldfaithful3.png]

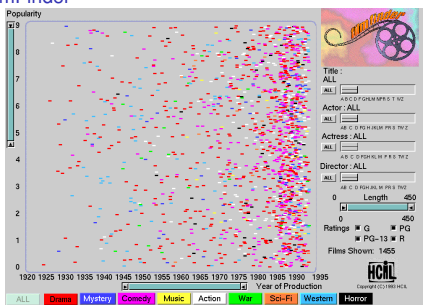
Dynamic Queries on Scatterplots

- tight coupling: immediate feedback after action
- starfield = interactive scatterplot
- dynamic queries as lightweight visual exploration
  - vs. composing SQL query



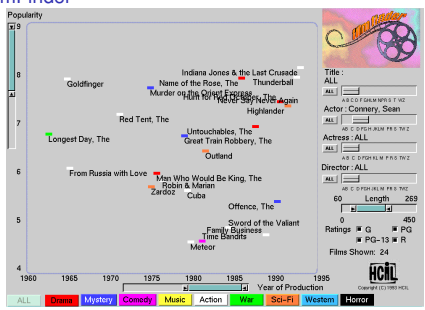
[Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Chris Ahlberg and Ben Shneiderman, Proc SIGCHI '94, p 313-317] [http://www.cs.umd.edu/hcil/pubs/screenshots/FilmFinder/]

FilmFinder



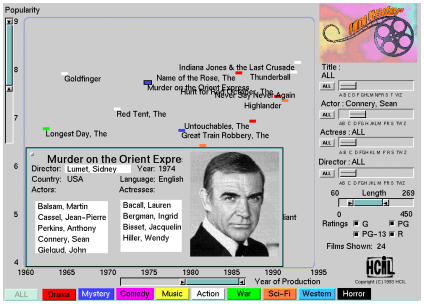
[Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Chris Ahlberg and Ben Shneiderman, Proc SIGCHI '94, p 313-317] [http://www.cs.umd.edu/hcil/pubs/screenshots/FilmFinder/]

FilmFinder



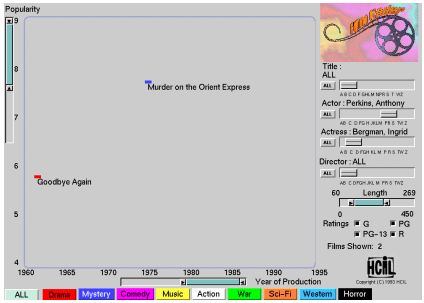
[Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Chris Ahlberg and Ben Shneiderman, Proc SIGCHI '94, p 313-317] [http://www.cs.umd.edu/hcil/pubs/screenshots/FilmFinder/]

FilmFinder



[Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Chris Ahlberg and Ben Shneiderman, Proc SIGCHI '94, p 313-317] [http://www.cs.umd.edu/hcil/pubs/screenshots/FilmFinder/]

FilmFinder



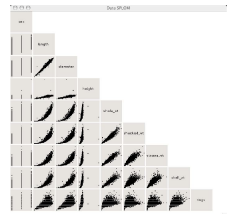
[Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Chris Ahlberg and Ben Shneiderman, Proc SIGCHI '94, p 313-317] [http://www.cs.umd.edu/hcil/pubs/screenshots/FilmFinder/]

Critique

- clear successes
  - fast, lightweight visual queries
  - details on demand
  - easy to use for novices
- more arguable: alphasliders
  - other techniques: data vis sliders, fisheye menus, speed-dependent automatic zooming

SPLOM: Scatterplot Matrix

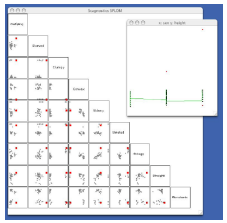
- show all pairwise variable combos side by side
  - matrix size grows quadratically with variable count



[Graph-Theoretic Scagnostics. Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

Graph-Theoretic Scagnostics

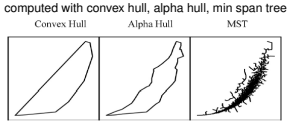
- reduce problem to constant size
  - overview matrix of 9 geometric metrics
- meta-SPLOM: each point represents scatterplot
  - detail on demand to see individual scatterplots



Graph-Theoretic Scagnostics. Leland Wilkinson, Anushka Anand, and Robert Grossman. Proc InfoVis 05.

Measuring Scatterplots

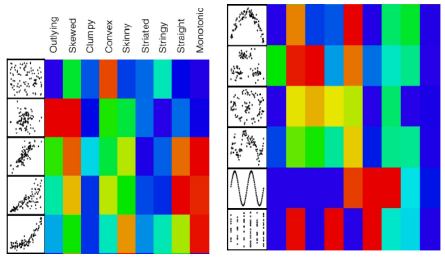
- aspects and measures
  - outliers: outlying
  - shape: convex, skinny, stringy, straight
    - computed with convex hull, alpha hull, min span tree



  - trend: monotonic
  - density: skewed, clumpy
  - coherence: striated

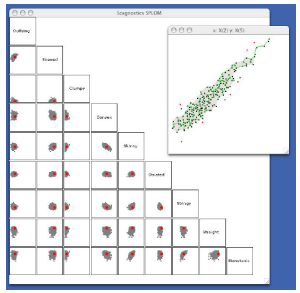
[Graph-Theoretic Scagnostics. Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

Measuring Scatterplots



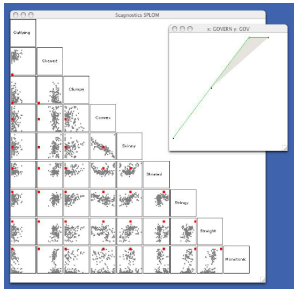
[Graph-Theoretic Scagnostics. Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

Results



[Graph-Theoretic Scagnostics. Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

Results



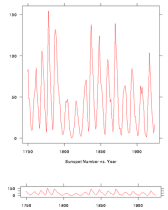
[Graph-Theoretic Scagnostics. Wilkinson, Anand, and Grossman. Proc InfoVis 05.]

Critique

- very powerful and elegant method
  - curse of dimensionality is hard problem
- abstraction level clearly appropriate for experts

Banking to 45 Degrees

- evangelized by Cleveland
- perceptual principle: most accurate angle judgement at 45 degrees
- pick aspect ratio (height/width) accordingly



[www.research.att.com/~rab/trellis/sunspot.html]

Multiscale Banking to 45

- frequency domain analysis
- find interesting regions at multiple scales

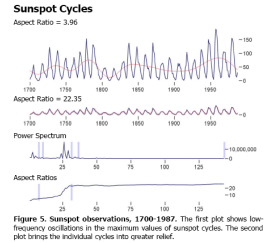
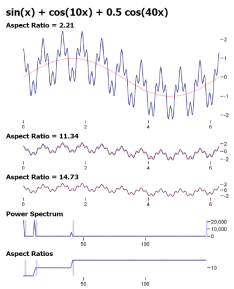


Figure 5. Sunspot observations, 1700-1987. The first plot shows low-frequency oscillations in the maximum values of sunspot cycles. The second plot brings the individual cycles into greater relief.

[Multi-Scale Banking to 45 Degrees. Heer and Agrawala, Proc InfoVis 2006 vis.berkeley.edu/papers/banking]

Choosing Aspect Ratios

- FFT the data, smooth by convolve with Gaussian
- find interesting spikes/ranges in power spectrum
- cull nearby regions if too similar, ensure overview shown
- create trend curves for each aspect ratio



Multiscale Banking to 45

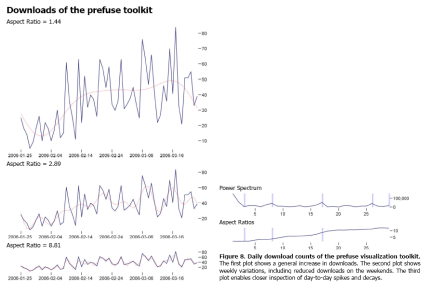


Figure 8. Daily download counts of the prefuse visualization toolkit. The first plot shows a general increase in downloads. The second plot shows weekly variations, including reduced downloads on the weekends. The third plot enables closer inspection of day-to-day spikes and decays.

[Multi-Scale Banking to 45 Degrees. Heer and Agrawala, Proc InfoVis 2006 vis.berkeley.edu/papers/banking]

Critique

- very nice generalization of old idea
- does not require interactivity to reap benefits

Statistical Graphics: Readings

Visual information seeking: Tight coupling of dynamic query filters with starfield displays. Chris Ahlberg and Ben Shneiderman, Proc SIGCHI '94, pages 313-317

Graph-Theoretic Scagnostics. Leland Wilkinson, Anushka Anand, and Robert Grossman. Proc InfoVis 05

Multi-Scale Banking to 45 Degrees. Jeffrey Heer, Maneesh Agrawala. IEEE TVCG 12(5) (Proc. InfoVis 2006), Sep/Oct 2006, pages 701-708.

SG: Further Readings

Metric-Based Network Exploration and Multiscale Scatterplot. Yves Chiricota, Fabien Jourdan, Guy Melancon. Proc. InfoVis 04, pages 135-142.

The Visual Design and Control of Trellis Display. R. A. Becker, W. S. Cleveland, and M. J. Shyu Journal of Computational and Statistical Graphics, 5:123-155. (1996).

The Elements of Graphing Data. William S. Cleveland. Hobart Press 1994.