University of British Columbia
CPSC 314 Computer Graphics
Jan-Apr 2013

Tamara Munzner

Visualization

http://www.ugrad.cs.ubc.ca/~cs314/Vjan2013

Nonspatial/Information Visualization

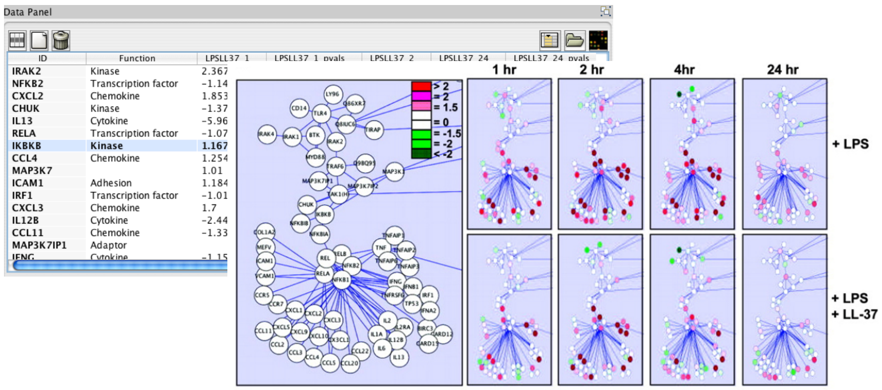
Reading

- FCG Chap 27
  - N/A 2nd edition, available online at

    http://www.cs.ubc.ca/labs/imager/tr/2009/VisChapter

Why Do Visualization?

- pictures help us think
  - substitute perception for cognition
  - external memory: free up limited cognitive/memory resources for higher-level problems



Information Visualization

- interactive visual representation of abstract data
  - help human perform some task more effectively
- bridging many fields
  - computer graphics: interact in realtime
  - cognitive psychology: find appropriate representation
  - HCI: use task to guide design and evaluation
- external representation
  - reduces load on working memory
  - offload cognition
  - familiar example: multiplication/division
  - infovis example: topic graphs

External Representation: Topic Graphs

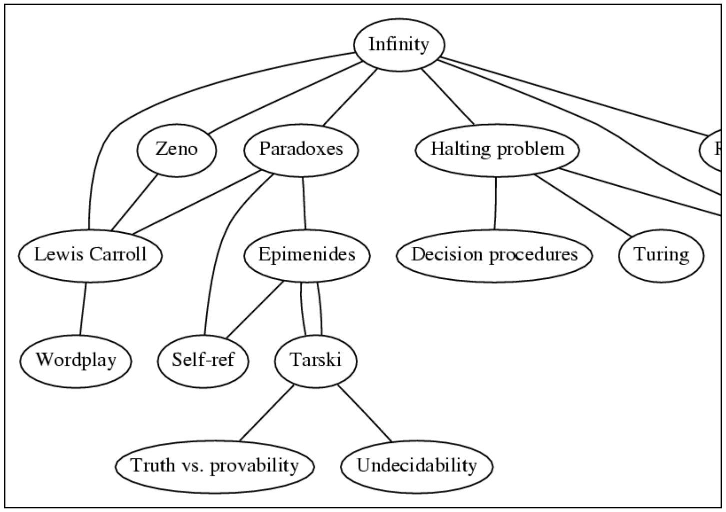
- hard to find topics two hops away from target

[Godel, Escher, Bach: The Eternal Golden Braid. Hofstadter 1979]

- **Paradoxes** - Lewis Carroll
- Turing - Halting problem
- Halting problem - Infinity
- Paradoxes - Infinity
- Infinity - Lewis Carroll
- Infinity - Unpredictably long searches
- Infinity - Recursion
- Infinity - Zeno
- Infinity - Paradoxes
- Lewis Carroll - Zeno
- Lewis Carroll - Wordplay
- Halting problem - Decision procedures
- BlooP and FlooP - AI
- Halting problem - Unpredictably long searches
- BlooP and FlooP - Unpredictably long searches
- BlooP and FlooP - Recursion
- Tarski - Truth vs. provability
- Tarski - Epimenides
- Tarski - Undecidability
- Paradoxes - Self-ref
- [...]

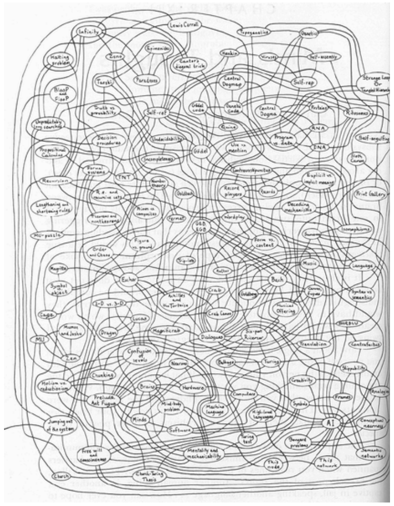
External Representation: Topic Graphs

- offload cognition to visual system



Automatic Node-Link Graph Layout

- manual: hours, days



[Godel, Escher, Bach. Hofstadter 1979]

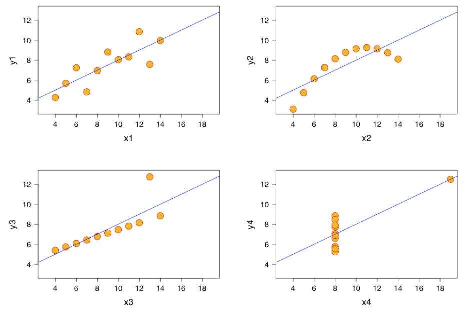
- automatic: seconds

[dot, Gansner et al, 1973.]

When To Do Vis?

- need a human in the loop
  - augment, not replace, human cognition
  - for problems that cannot be (completely) automated
- simple summary not adequate
  - statistics may not adequately characterize complexity of dataset distribution

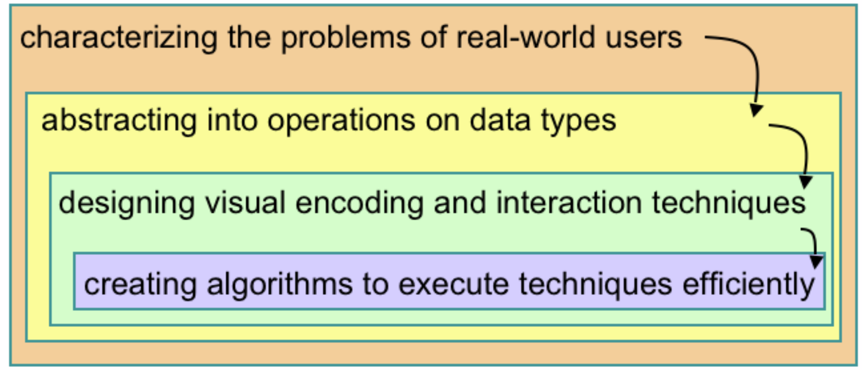
Anscombe’s quartet: same
- mean
- variance
- correlation coefficient
- linear regression line



http://upload.wikimedia.org/wikipedia/commons/b/b6/Anscombe.svg

Visualization Design Layers

- depends on both data and task



Visual Encoding

marks: geometric primitives

points lines areas

attributes

position

size

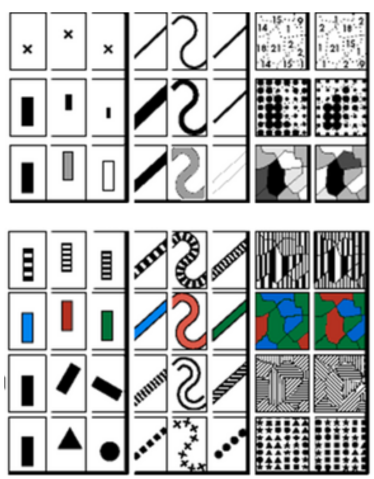
grey level

texture

color

orientation

shape

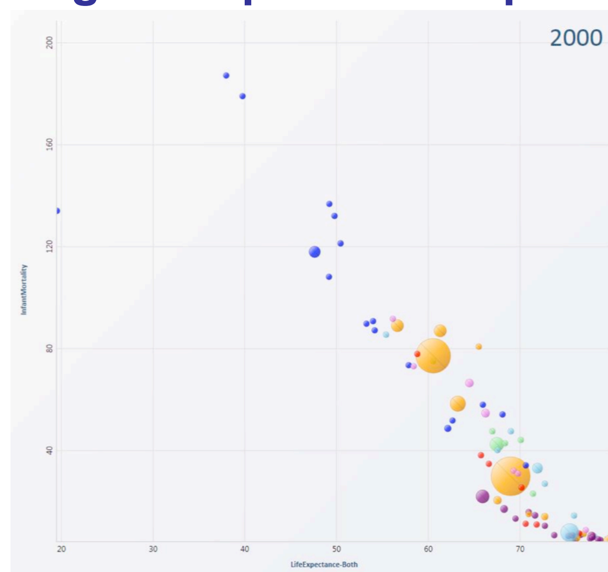


- attributes
  - parameters control mark appearance
  - separable channels flowing from retina to brain

Semiology of Graphics. Jacques Bertin, Gauthier-Villars 1967, EHESS 1998

Visual Encoding Example: Scatterplot

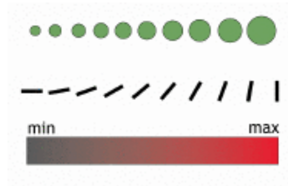
- x position
- y position
- hue
- size



Robertson et al. Effectiveness of Animation in Trend Visualization. IEEE TVCG (Proc. InfoVis08) 14:6 (2008), 1325–1332.

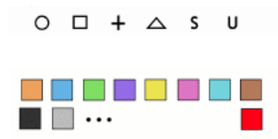
Data Types

- quantitative
  - lengths: 10 inches, 17 inches, 23 inches



- ordered
  - sizes: small, medium, large
  - days: Mon, Tue, Wed, ...

- categorical
  - fruit: apples, oranges, bananas



[Stolte and Hanrahan. Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Proc InfoVis 2000. graphics.stanford.edu/projects/polaris/]

Channel Ranking Varies By Data Type

Quantitative	Ordered	Categorical
Position	Position	Position
Length	Lightness	Hue
Angle	Saturation	Texture
Slope	Hue	Connection
Area	Texture	Containment
Volume	Connection	Lightness
Lightness	Containment	Saturation
Saturation	Length	Shape
Hue	Angle	Length
Texture	Slope	Angle
Connection	Area	Slope
Containment	Volume	Area
Shape	Shape	Volume

[Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM TOG 5:2, 1986]

Integral vs. Separable Dimensions

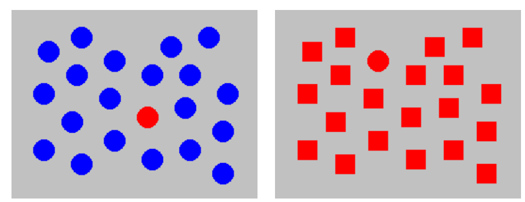
- not all dimensions separable

color location | color motion | color shape | size orientation | x-size y-size | red-green yellow-blue

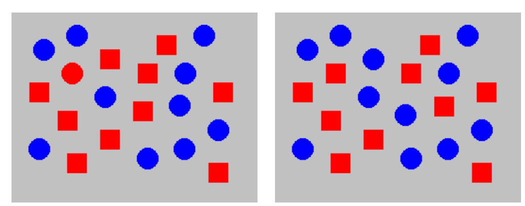
[Colin Ware, Information Visualization: Perception for Design. Morgan Kaufmann 1999.]

Preattentive Visual Channels

- color alone, shape alone: preattentive



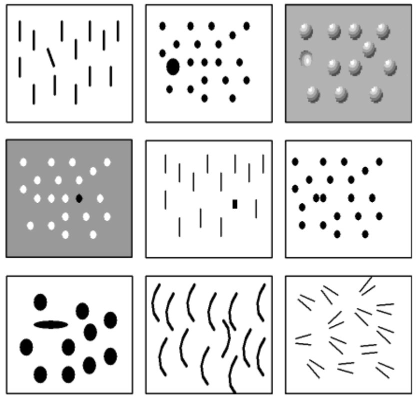
- combined color and shape: requires attention
  - search speed linear with distractor count



[Christopher Healey, [www.csc.ncsu.edu/faculty/healey/PP/PP.html]

Preattentive Visual Channels

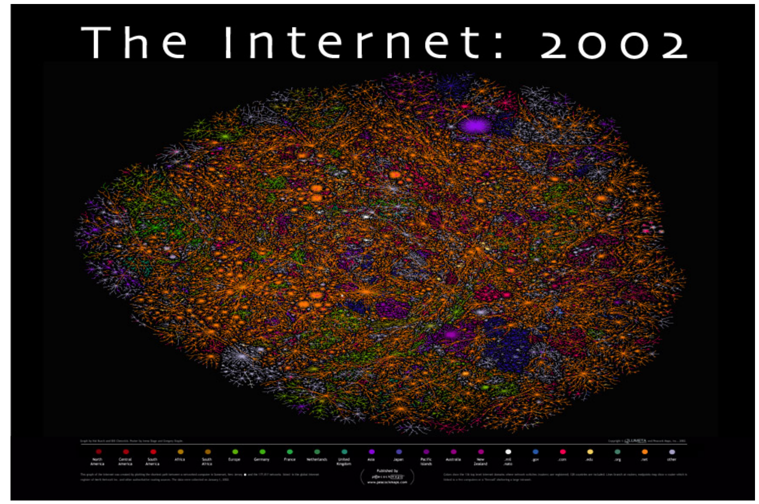
- preattentive channels include
  - hue
  - shape
  - texture
  - length
  - width
  - size
  - orientation
  - curvature
  - intersection
  - intensity
  - flicker
  - direction of motion
  - stereoscopic depth
  - lighting direction
  - many more...



[Healey, [www.csc.ncsu.edu/faculty/healey/PP/PP.html]

Coloring Categorical Data

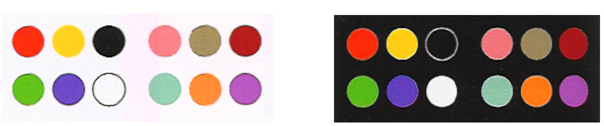
- 22 colors, but only ~8 distinguishable



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

Coloring Categorical Data

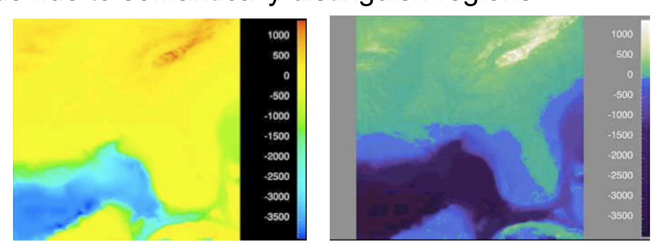
- discrete small patches separated in space
- limited distinguishability: around 8-14
  - channel dynamic range low
  - best to choose bins explicitly
- maximal saturation for small areas



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

Quantitative Colormaps

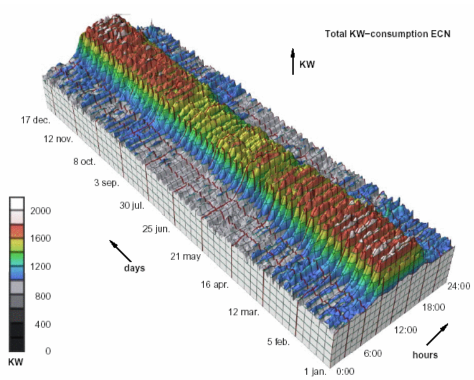
- dangers of rainbows
  - perceptually nonlinear
  - arbitrary not innate ordering
- other approaches
  - explicitly segmented colormaps
  - monotonically increasing/(decreasing) luminance, plus hue to semantically distinguish regions

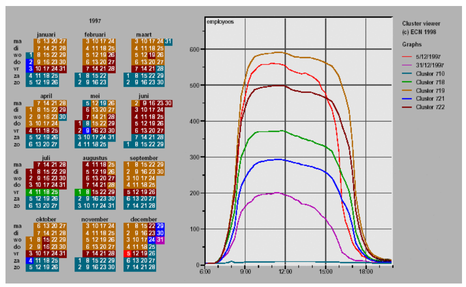


Rogowitz and Treinish. Data Visualization: The End of the Rainbow. IEEE Spectrum 35(12):52-59, Dec 1998.

3D vs 2D Representations

- curve comparison difficult: perspective distortion, occlusion
  - dataset is abstract, not inherently spatial
  - after data transformation to clusters, linked 2D views of representative curves show more





[van Wijk and van Selow, Cluster and Calendar based Visualization of Time Series Data, InfoVis99

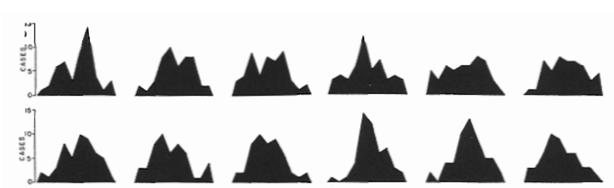
Space vs Time: Showing Change

- animation: show time using temporal change
  - good: show process
  - good: flip between two things
  - bad: flip between between many things
    - interference between intermediate frames

[Outside In excerpt. www.geom.uiuc.edu/docs/outreach/oi/evert.mpg]
[www.astroshow.com/ccdpho/pluto.gif]
[Edward Tufte. The Visual Display of Quantitative Information, p 172]

Space vs Time: Showing Change

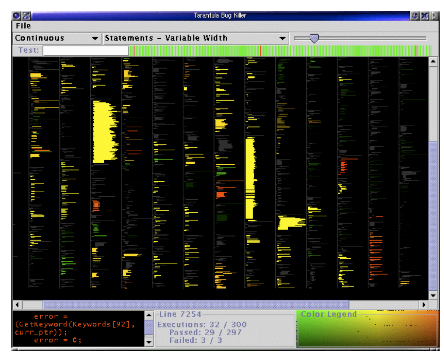
- small multiples: show time using space
  - overview: show each time step in array
  - compare: side by side easier than temporal
    - external cognition vs internal memory
  - general technique, not just for temporal changes



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

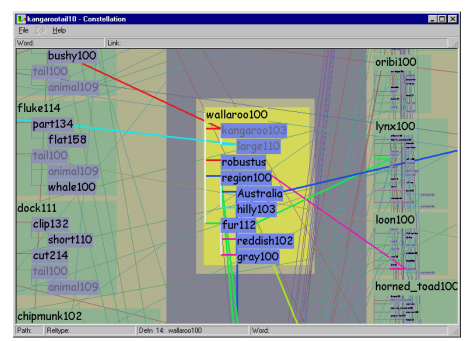
Composite Views

- pixel-oriented views
  - overviews with high information density



[Jones, Harrold, and Stasko. Visualization of Test Information to Assist Fault Localization. Proc. ICSE 2002, p 467-477.]

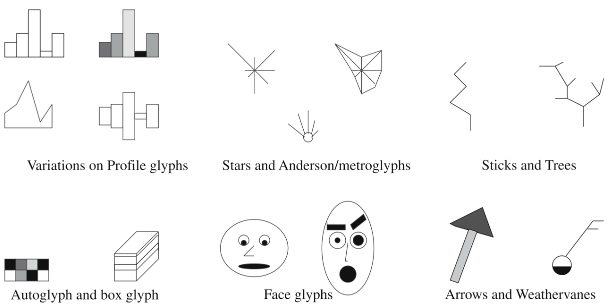
- superimposing/layering
  - shared coordinate frame
  - redundant visual encoding

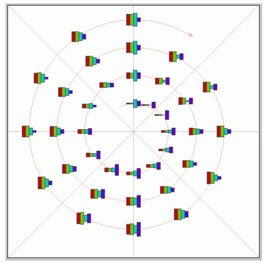


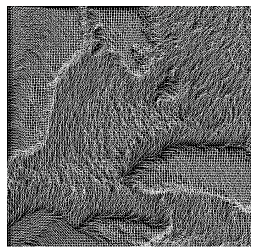
[Munzner. Interactive Visualization of Large Graphs and Networks. Stanford CS, 2000]

Composite Views: Glyphs

- internal structure where subregions have different visual channel encodings





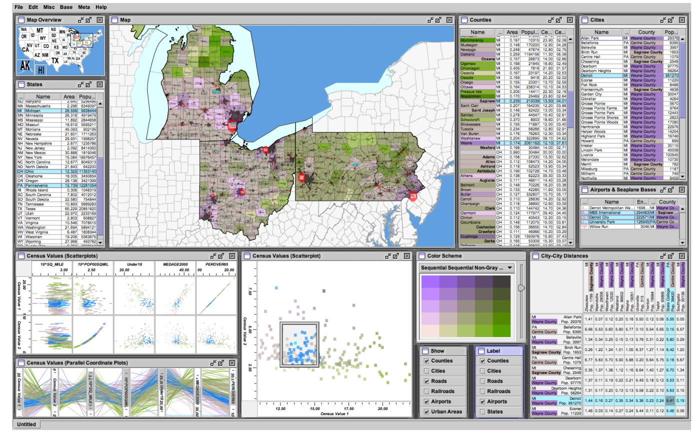


[Ward. A Taxonomy of Glyph Placement Strategies for Multidimensional Data Visualization. Information Visualization Journal 1:3-4 (2002), 194--210.]

[Smith, Grinstein, and Bergeron. Interactive data exploration with a supercomputer. Proc. IEEE Visualization, p 248-254, 1991.]

Adjacent: Multiple Views

- different visual encodings show different aspects of the data
- linked highlighting to show where contiguous in one view distributed within another



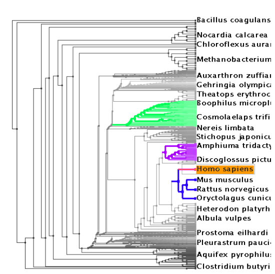
[Weaver. http://www.personal.psu.edu/cew15/improvise/examples/census]

Adjacent Views

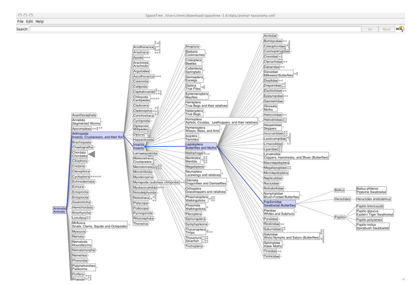
- overview and detail
  - same visual encoding, different resolutions
- small multiples
  - same visual encoding, different data

Data Reduction

- overviews as aggregation
- focus+context
  - show details embedded within context
    - distortion: TreeJuxtaposer video
    - filtering: SpaceTree demo



[Munzner et al. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. Proc SIGGRAPH 2003, p 453-462]



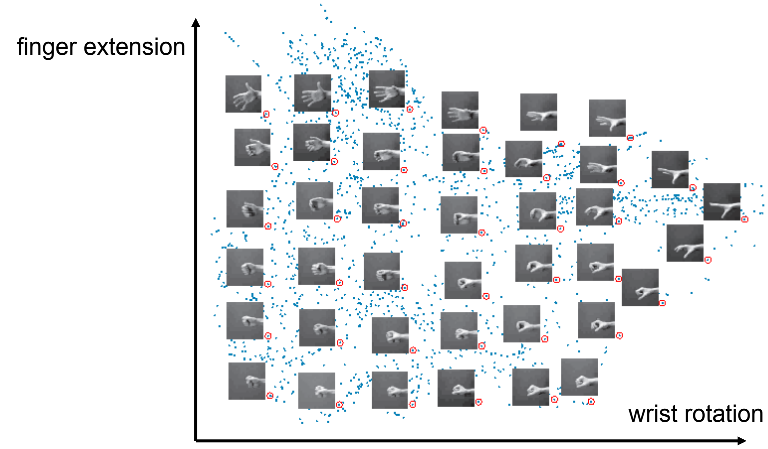
[Plaisant, Grosjean, and Bederson. SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation. Proc. InfoVis 2002

Dimensionality Reduction

- mapping from high-dimensional space into space of fewer dimensions
  - generate new synthetic dimensions
- why is lower-dimensional approximation useful?
  - assume **true/intrinsic** dimensionality of dataset is (much) lower than measured dimensionality!
    - only indirect measurement possible?
      - fisheries: want spawn rates.
        have water color, air temp, catch rates...
    - sparse data in verbose space?
      - documents: word occurrence vectors.
        10K+ dimensions, want dozens of topic clusters

DR Example: Image Database

- 4096 D (pixels) to 2D (hand gesture)
  - no semantics of new synthetic dimensions from alg.
    - assigned by humans after inspecting results



[A Global Geometric Framework for Nonlinear Dimensionality Reduction. Tenenbaum, de Silva and Langford. *Science* 290 (5500): 2319-2323, 2000, isomap.stanford.edu]

DR Technique: MDS

- multidimensional scaling
  - minimize differences between interpoint distances in high and low dimensions
- minimize objective function: stress

$$stress(D,\Delta) = \sqrt{\frac{\sum_{ij}(d_{ij}-\delta_{ij})^2}{\sum_{ij}\delta_{ij}^2}}$$

D: matrix of lowD distances d_{ij}

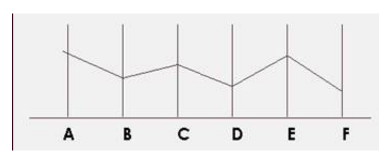
Δ: matrix of hiD distances δ_{ij}

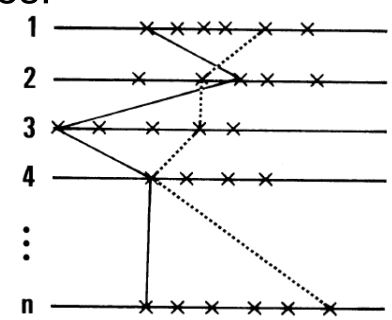
- Glimmer: MDS on the GPU

[Ingram, Munzner, Olano. Glimmer: Multiscale MDS on the GPU. IEEE TVCG 15(2):249-261, 2009.

Parallel Coordinates

- only two orthogonal axes in the plane
- instead, use parallel axes!

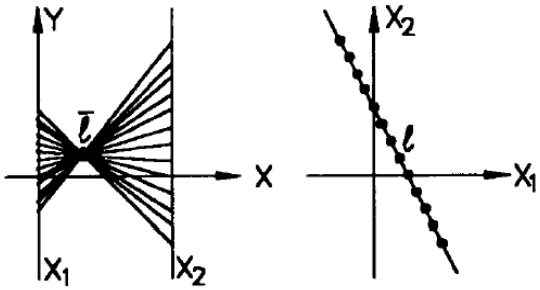




[Hyperdimensional Data Analysis Using Parallel Coordinates. Edward J. Wegman. Journal of the American Statistical Association, Vol. 85, No. 411. (Sep., 1990), pp. 664-675.]

Parallel Coordinates

- point in Cartesian coords is line in par coords
- point in par coords is line in Cartesian n-space



[Inselberg and Dimdale. Parallel Coordinates: A Tool for Visualizing Multi-Dimensional Geometry. IEEE Visualization '90.]

Par Coords: Correllation

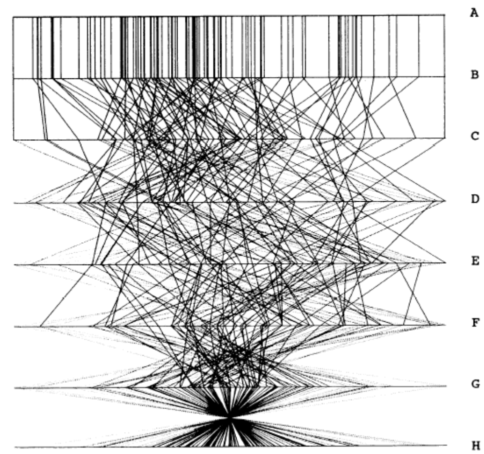
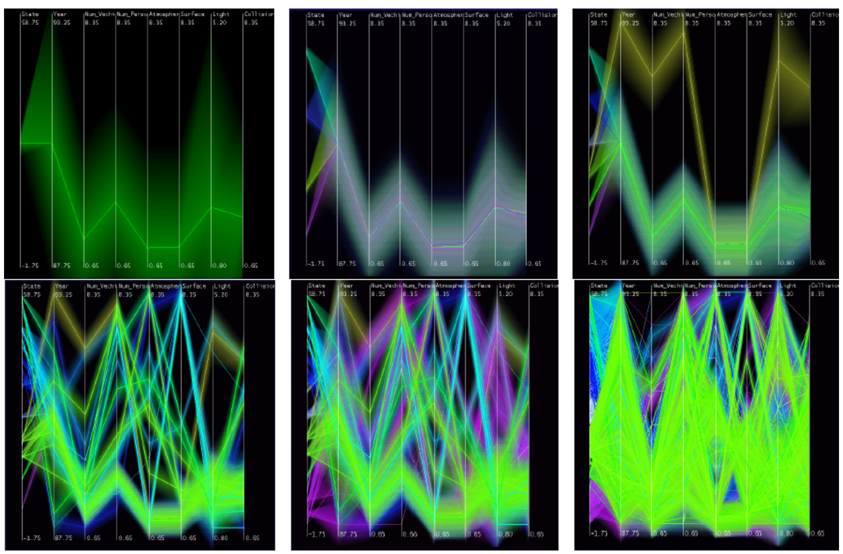


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of ρ = 1, .8, .2, 0, −.2, −.8, and −1.

[Hyperdimensional Data Analysis Using Parallel Coordinates. Wegman. Journal of the American Statistical Association, Vol. 85, No. 411. (Sep., 1990), pp. 664-675.]

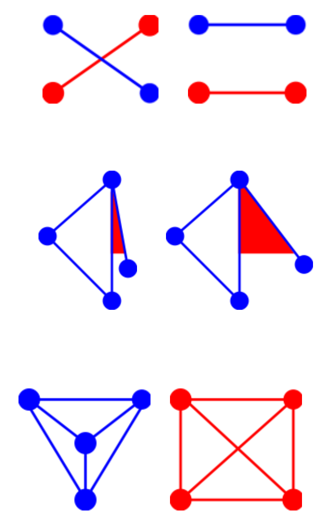
Hierarchical Parallel Coords: LOD



[Hierarchical Parallel Coordinates for Visualizing Large Multivariate Data Sets. Fua, Ward, and Rundensteiner. IEEE Visualization '99.]

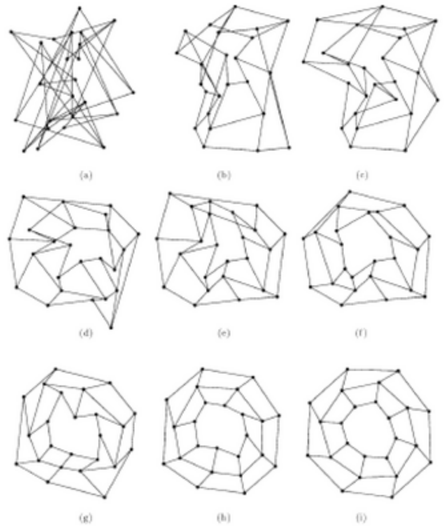
Node-Link Graph Layout

- minimize
  - crossings, area, bends/curves
- maximize
  - angular resolution, symmetry
- most criteria individually NP-hard
  - cannot just compute optimal answer
  - heuristics: try to find something reasonable
- criteria mutually incompatible



Force-Directed Placement

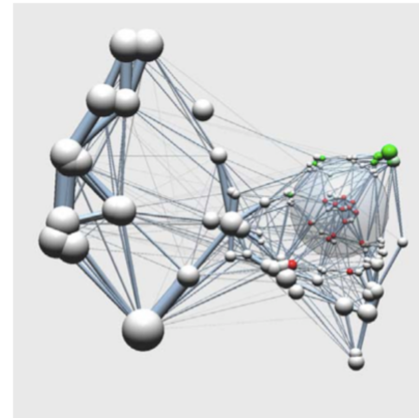
- nodes: repel like magnets
- edges: attract like springs
  - start from random positions, run to convergence
- very well studied area!
  - many people reinvent the wheel



[www.csse.monash.edu.au/~berndm/CSE460/Lectures/cse460-7.pdf]

Interactive Graph Exploration

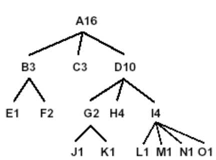
- geometric and semantic fisheye



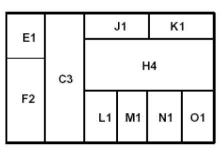
van Ham and van Wijk. Interactive Visualization of Small World Graphs. Proc. InfoVis 2005

Treemaps

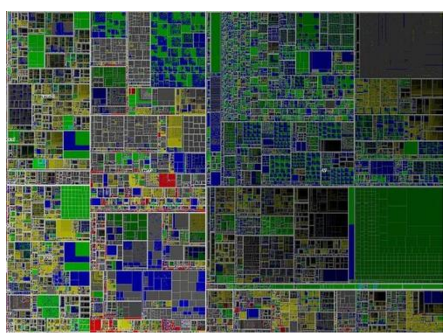
- containment rather than connection
  - emphasize node attributes, not topological structure



Node and link diagram



Treemap

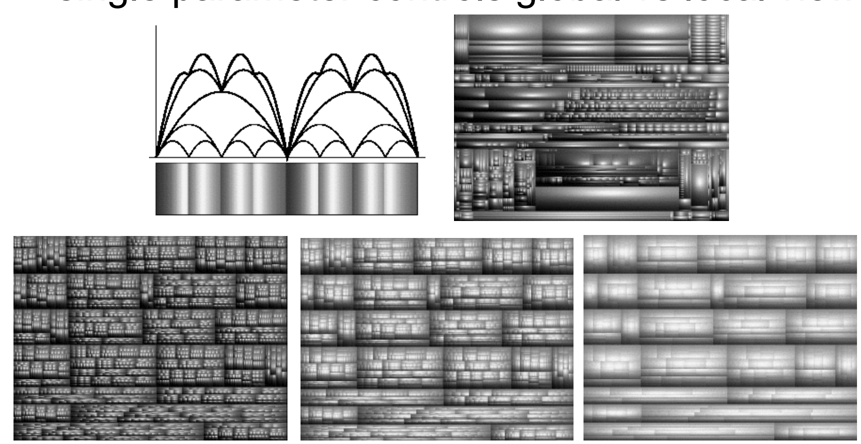


[van Wijk and van de Wetering. Cushion Treemaps. Proc InfoVis 1999]

[Fekete and Plaisant. Interactive Information Visualization of a Million Items. Proc InfoVis 2002.

Cushion Treemaps

- show structure with shading
  - single parameter controls global vs local view



[van Wijk and van de Wetering. Cushion Treemaps. Proc InfoVis 1999]

Now What?

Beyond 314: Other Graphics Courses

- 424: Geometric Modelling
  - was offered this year
- 426: Computer Animation
  - will be offered next year

- 514: Image-Based Rendering - Heidrich
- 526: Algorithmic Animation - van de Panne
- 533A: Digital Geometry - Sheffer
- 533B: Animation Physics - Bridson
- 547: Information Visualization - Munzner

Beyond UBC CS

- SIGGRAPH conference back in Vancouver August 2014!
  - 15K-20K people: incredible combination of research, entertainment, art
  - Electronic Theater, Exhibit, ETech, ...
  - pricey: but student rate, student volunteer program
- local SIGGRAPH chapter
  - talk series, SPARK FX festival, ...
  - http://siggraph.ca