Visualization Principles

Tamara Munzner
Department of Computer Science
University of British Columbia

Cytoscape Symposium on Network Biology 2012
13 Dec 2012

http://www.cs.ubc.ca/~tmm/talks.html#networkbio12

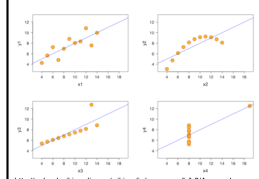
Defining visualization

computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively

Defining visualization

computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively

- human in the loop needs the details



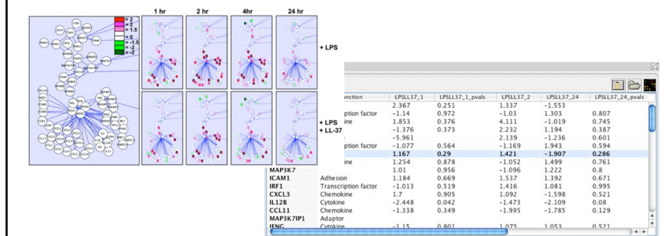
Identical statistics	
x mean	9.0
x variance	10.0
y mean	7.50
y variance	3.75
x/y correlation	0.816

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

Defining visualization

computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively

- human in the loop needs the details
- external representation: perception vs cognition



Defining visualization

computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively

- human in the loop needs the details
- external representation: perception vs cognition
- intended task

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

Defining visualization

computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively

- human in the loop needs the details
- external representation: perception vs cognition
- intended task
- measureable definitions of effectiveness

Defining visualization

Computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively.

These visualization systems are often but not always interactive. Resource limitations include the capacity of computers, of humans, and of displays.

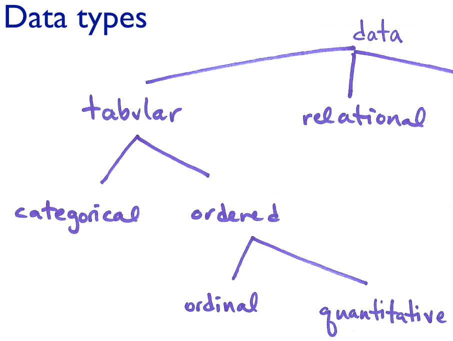
Visualization design space

- huge space of design alternatives
  - tradeoffs abound
- many possibilities now known to be ineffective
  - avoid random walk through parameter space
  - avoid some of our past mistakes
  - extensive experimentation has already been done
- guidelines continue to evolve
  - we reflect on lessons learned in design studies
  - iterative refinement usually wise

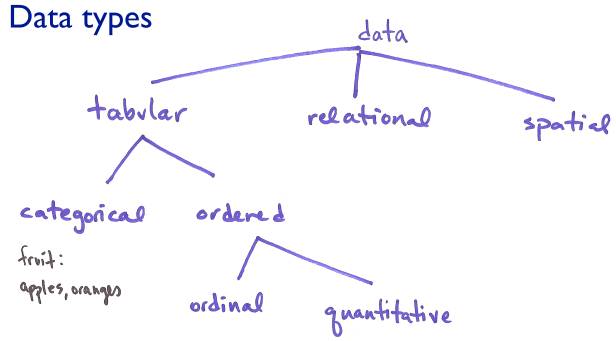
Principles

- know your visual channel types and ranks
- categorical color constraints
- power of the plane
- danger of depth
- resolution beats immersion
- eyes beat memory
- validate against the right threat

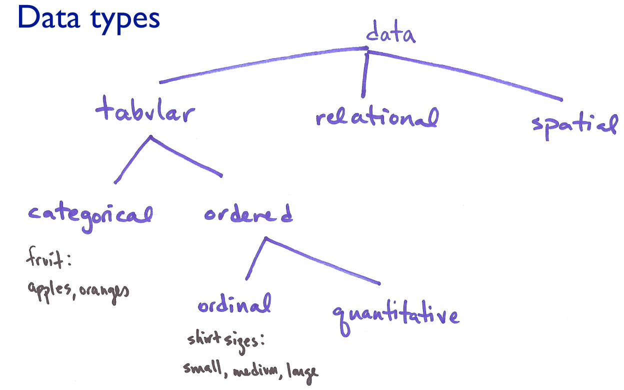
Data types



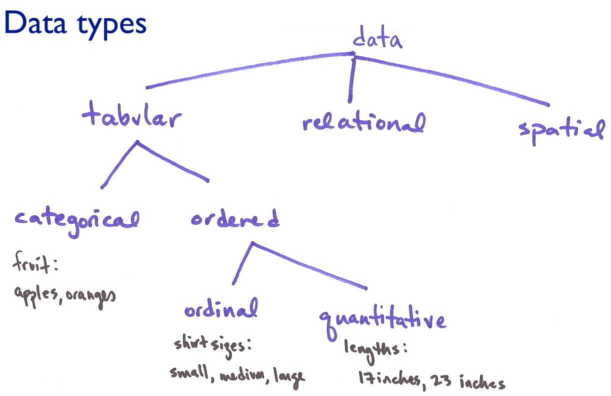
Data types



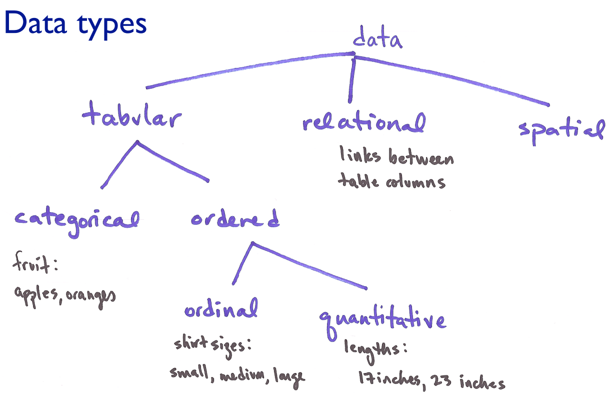
Data types



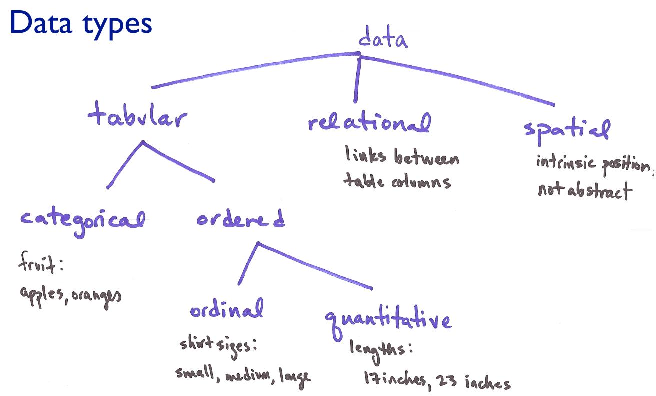
Data types



Data types

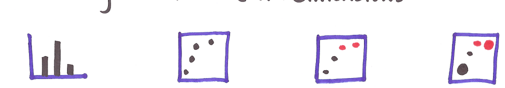


Data types



Visual encoding

- analyze
  showing abstract data dimensions



Visual encoding

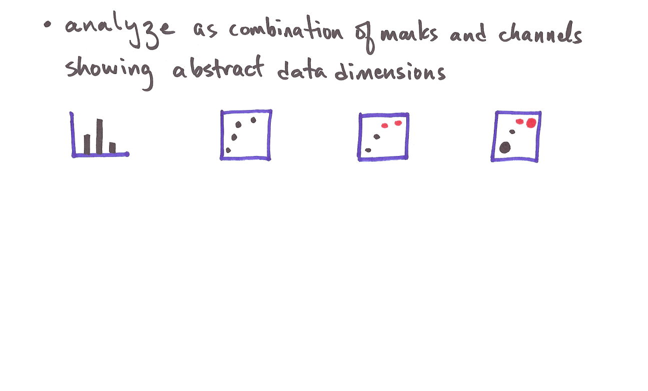
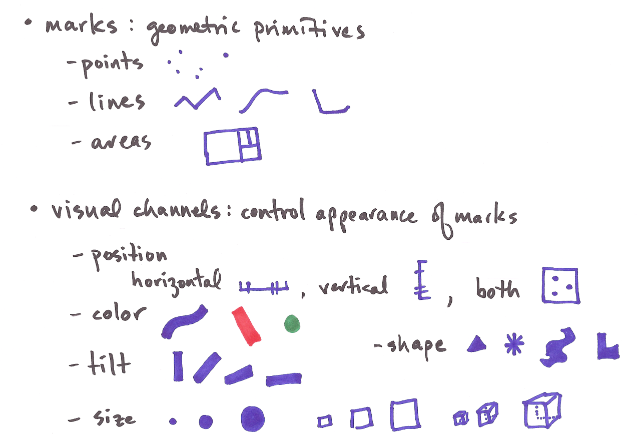
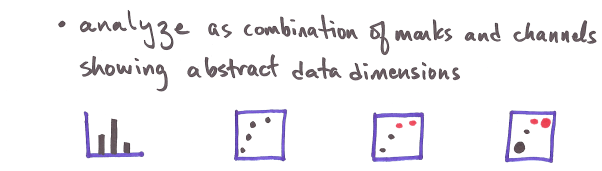


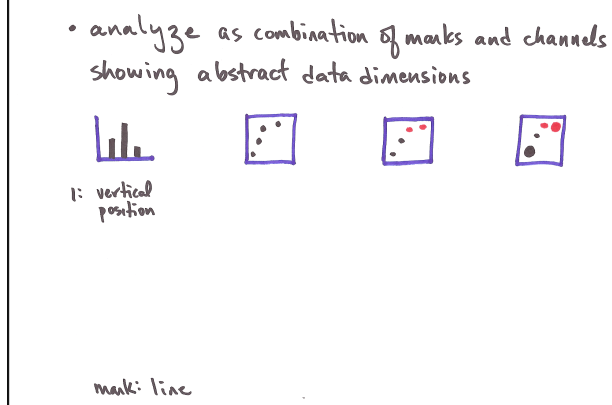
Image theory



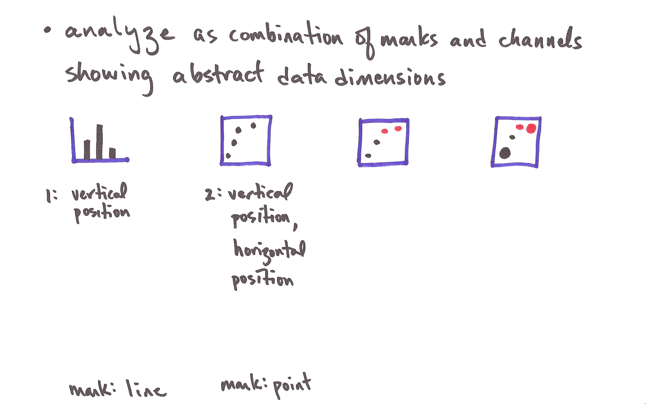
Visual encoding



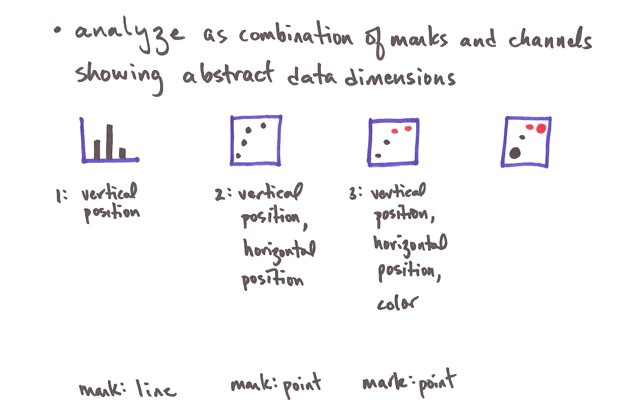
Visual encoding



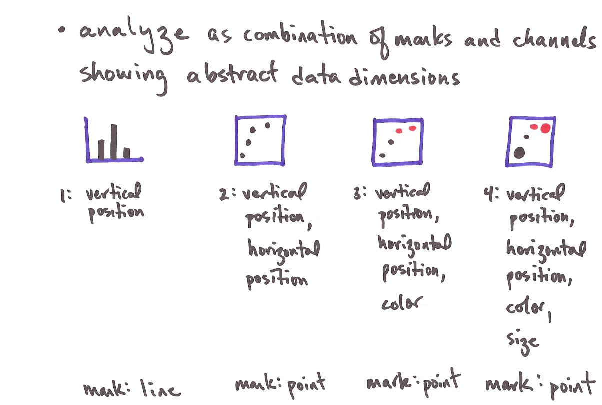
Visual encoding



Visual encoding

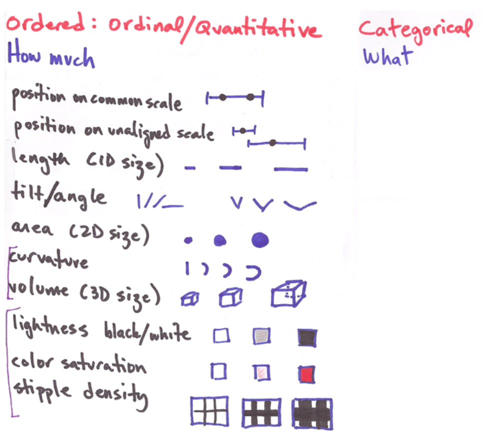


Visual encoding

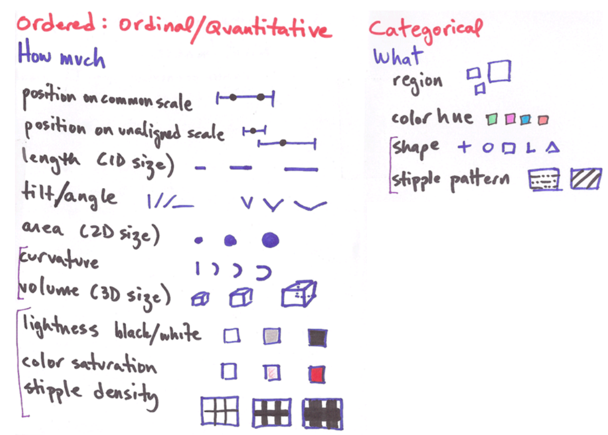


Visual channel types and rankings

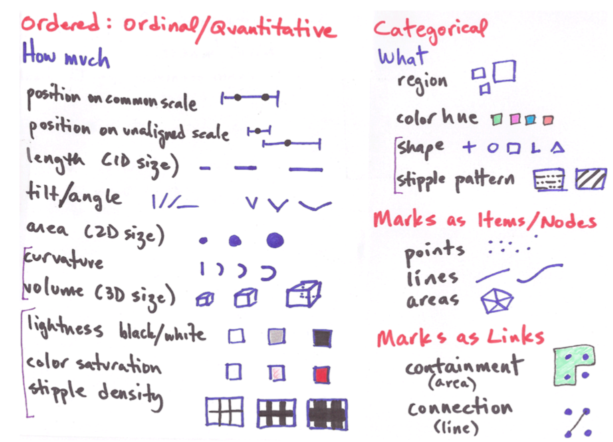
Visual channel types and rankings



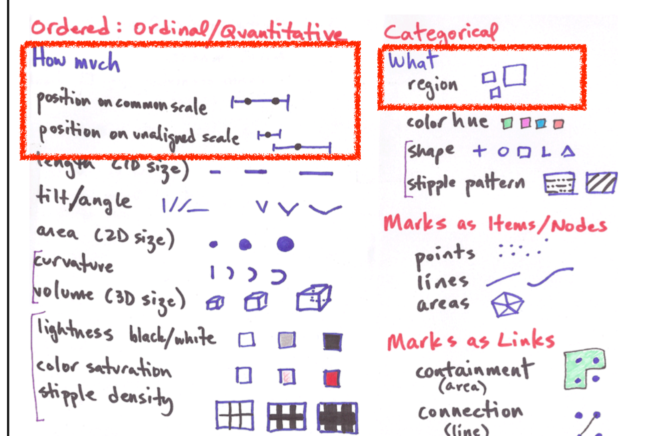
Visual channel types and rankings



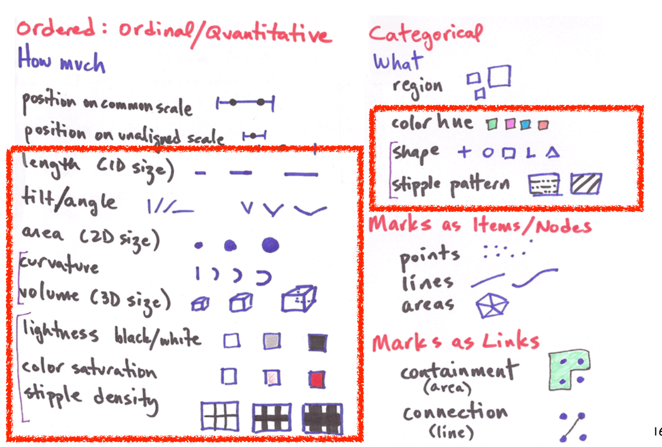
Visual channel types and rankings



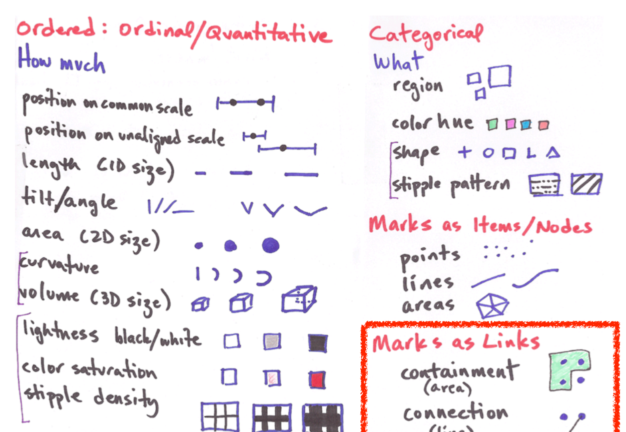
Power of the plane: only position works for all!



Ranking differs for all other channels



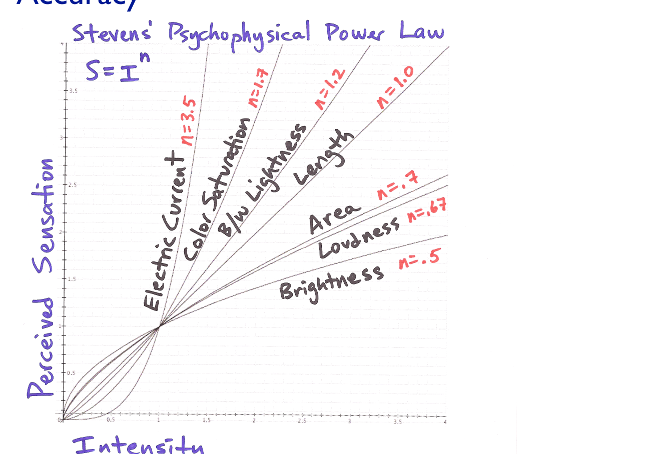
Networks: special case of general principles



Channel rankings

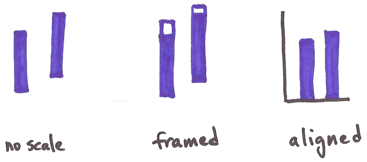
- effectiveness principle: encode most important attributes with highest ranked channels [Mackinlay 86]
- where do rankings come from?
  - accuracy, discriminability, separability, popout

Accuracy



Accuracy

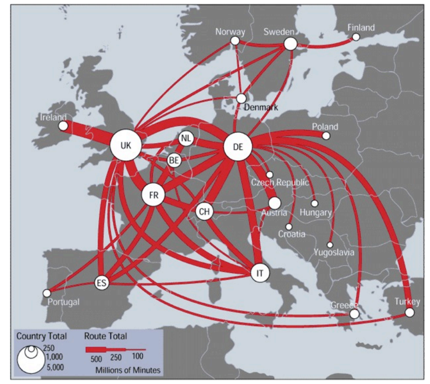
- position along common scale



- frame increases accuracy [Cleveland 84]
- Weber's Law: relative judgements
  - filled rectangles differ by 1:9
  - white rectangles differ by 1:2

Discriminability: How many usable steps?

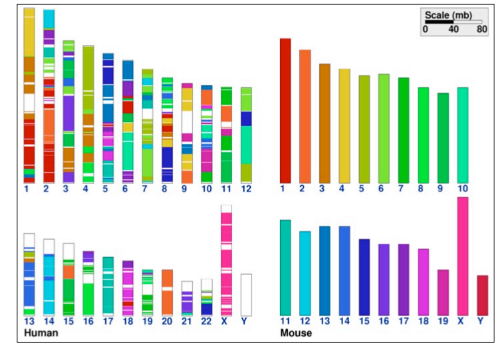
- linewidth: only a few



[mappa.mundi.net/maps/maps 014/telegeography.html]

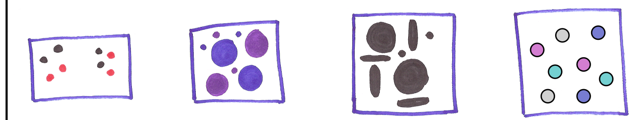
Discriminability: Categorical color constraints

- noncontiguous small regions of color: only 6-12 bins

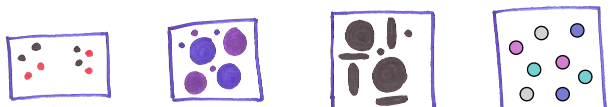


Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. Bioinformatics 2007

Separability vs. integrality



Separability vs. integrality

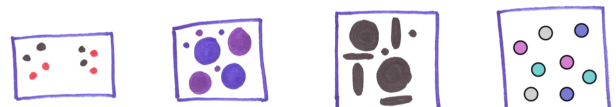


position
hue (color)

fully separable

2 groups each

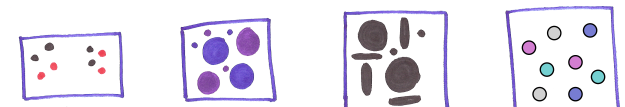
Separability vs. integrality



position, hue (color) — fully separable — 2 groups each

size, hue (color) — some interference — difficult to discriminate small items — (2 groups each)

Separability vs. integrality

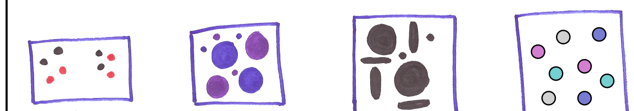


position, hue (color) — fully separable — 2 groups each

size, hue (color) — some interference — difficult to discriminate small items — (2 groups each)

size: width, size: height — some/significant interference — integral percept: area (planar size) — 3 groups

Separability vs. integrality



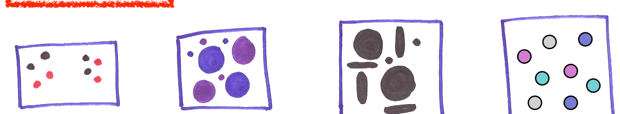
position, hue (color) — fully separable — 2 groups each

size, hue (color) — some interference — difficult to discriminate small items — (2 groups each)

size: width, size: height — some/significant interference — integral percept: area (planar size) — 3 groups

red, green — major interference — integral percept: color/hue — 4 groups

Separability vs. integrality



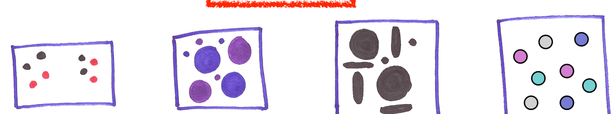
position, hue (color) — fully separable — 2 groups each

size, hue (color) — some interference — difficult to discriminate small items — (2 groups each)

size: width, size: height — some/significant interference — integral percept: area (planar size) — 3 groups

red, green — major interference — integral percept: color/hue — 4 groups

Separability vs. integrality



position, hue (color) — fully separable — 2 groups each

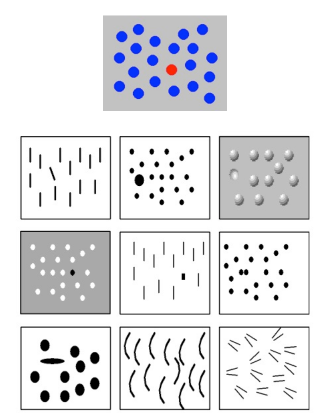
size, hue (color) — some interference — difficult to discriminate small items — (2 groups each)

size: width, size: height — some/significant interference — integral percept: area (planar size) — 3 groups

red, green — major interference — integral percept: color/hue — 4 groups

Popout: Most channels

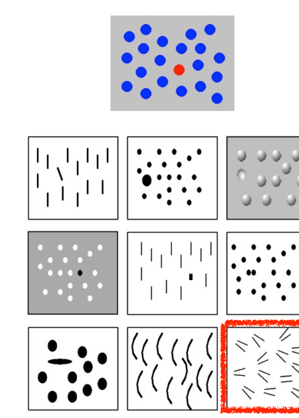
- parallel processing on most channels
  - sufficiently different item noticed immediately, independent of distractor count
- some channels have no popout: serial search required



Healey. Perception in Visualization
http://www.csc.ncsu.edu/faculty/healey/PP/

Popout: Most channels

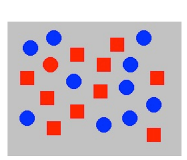
- parallel processing on most channels
  - sufficiently different item noticed immediately, independent of distractor count
- some channels have no popout: serial search required



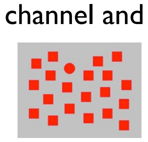
Healey. Perception in Visualization
http://www.csc.ncsu.edu/faculty/healey/PP/

Popout limits

- only one channel at a time
  - combination searches are serial
    - most channel pairs
    - all channel triplets, etc



- within channel, speed depends on which channel and how different item is from surroundings
  - 'sufficiently different': context dependent

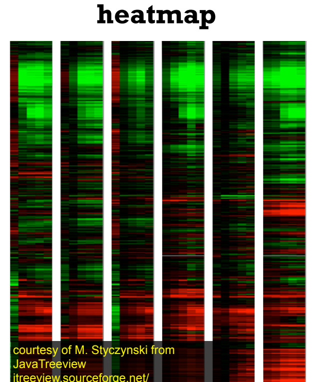


Healey. Perception in Visualization
http://www.csc.ncsu.edu/faculty/healey/PP/

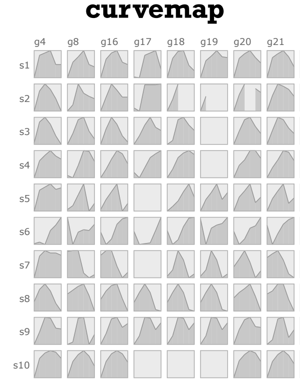
Encoding example: Heatmaps vs. curvemaps

- color traditional, but spatial position outranks it

heatmap

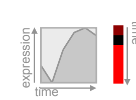


curvemap



Curvemap

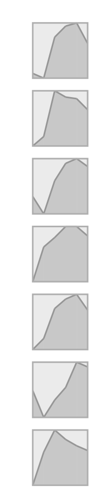
- shape perception easier for filled framed line charts than colored boxes



Pathline: A Tool for Comparative Functional Genomics.
Meyer, Wong, Styczynski, Munzner, Pfister. EuroVis 2010.

Curvemap

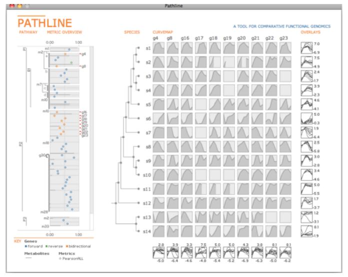
- shape perception easier for filled framed line charts than colored boxes

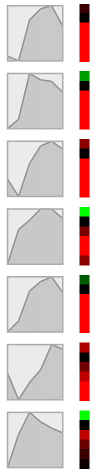


Pathline: A Tool for Comparative Functional Genomics.
Meyer, Wong, Styczynski, Munzner, Pfister. EuroVis 2010.

Curvemap

- shape perception easier for filled framed line charts than colored boxes

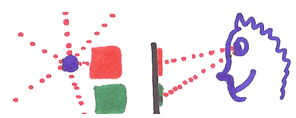




Pathline: A Tool for Comparative Functional Genomics. Meyer, Wong, Styczynski, Munzner, Pfister. EuroVis 2010.

Dangers of depth

- rankings for **planar** spatial position, not depth!
- we don't really live in 3D: we **see** in 2.05D
  - up/down and sideways: image plane
    - acquire more info quickly from eye movements
  - away: depth into scene
    - only acquire more info from head/body motion

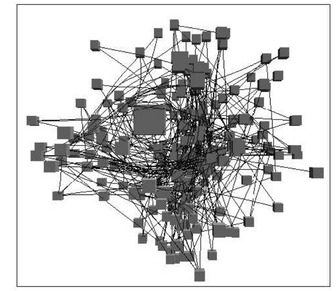


- further reading

Visual Thinking for Design (Chap 5). Colin Ware. 2008

Dangers of depth: difficulties of 3D

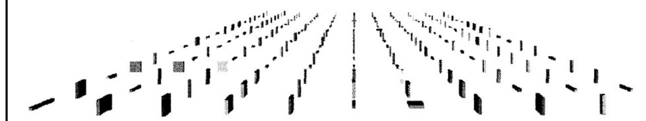
- occlusion
- interaction complexity



Distortion Viewing Techniques for 3D Data. Carpendale et al. InfoVis1996.

Dangers of depth: difficulties of 3D

- perspective distortion
  - interferes with all size channel encodings
  - power of the plane is lost!

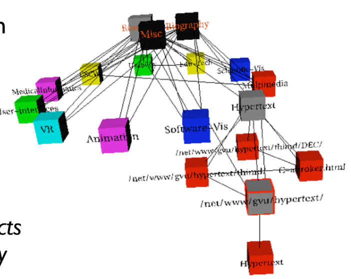


Visualizing the Results of Multimedia Web Search Engines. Mukherjea, Hirata, and Hara. InfoVis 96

Dangers of depth: difficulties of 3D

- tilted text isn't legible
  - far worse when tilted from image plane
- further reading

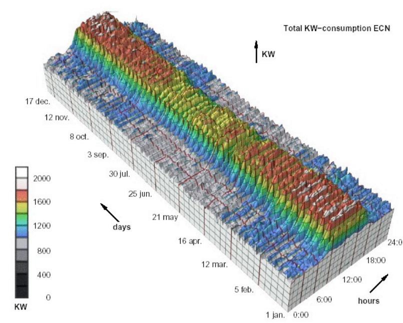
Exploring and Reducing the Effects of Orientation on Text Readability in Volumetric Displays. Grossman et al. CHI 2007



Visualizing the World-Wide Web with the Navigational View Builder. Mukherjea and Foley. Computer Networks and ISDN Systems, 1995.

Dangers of depth example

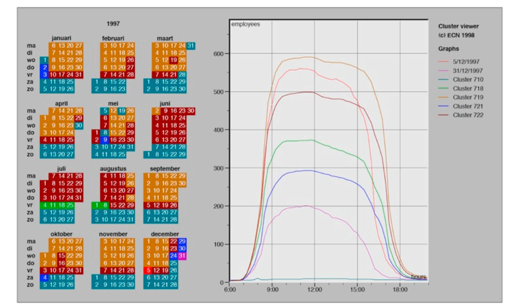
- extruded curves: detailed comparisons impossible



Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc InfoVis 99.

Transformation to suitable abstraction

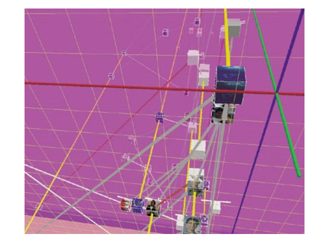
- derived data: clusters
- multiple views: calendar, superimposed 2D curves



Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc InfoVis 99.

Dangers of depth: must justify

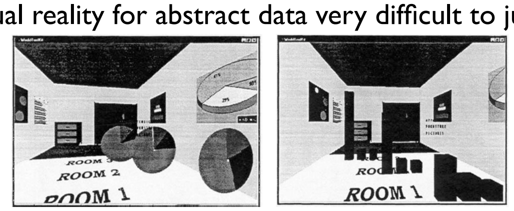
- 3D legitimate for true 3D spatial data
- 3D needs very careful justification for abstract data
  - enthusiasm in 1990s, but now skepticism
  - be especially careful with 3D for point clouds or networks



WEBPATH-a three dimensional Web history. Frecon and Smith. InfoVis 1999

Pixels are precious: Resolution beats immersion

- immersion typically not helpful for abstract data
  - do not need sense of presence or stereoscopic 3D
- resolution much more important
  - pixels are the scarcest resource
  - desktop also better for workflow integration
- virtual reality for abstract data very difficult to justify



Development of an information visualization tool using virtual reality. Kirner and Martins. Symp Applied Computing 2000

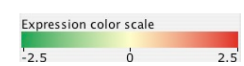
Eyes beat memory

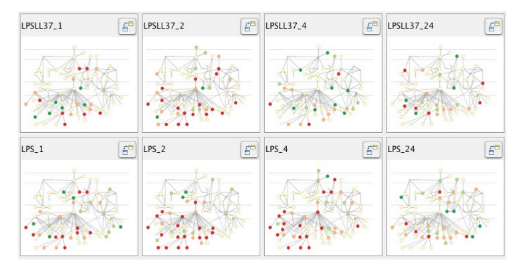
- principle: external cognition vs. internal memory
  - easy to compare by moving eyes between side-by-side views
  - harder to compare visible item to memory of what you saw
- implications for animation
  - great for choreographed storytelling
  - great for transitions between two states
  - poor for many states with changes everywhere
    - consider small multiples instead

literal	abstract
animation	small multiples
show time with time	show time with space

Small multiples example: Cerebral

- small multiples: one graph instance per experimental condition
  - same spatial layout
  - color differently, by condition

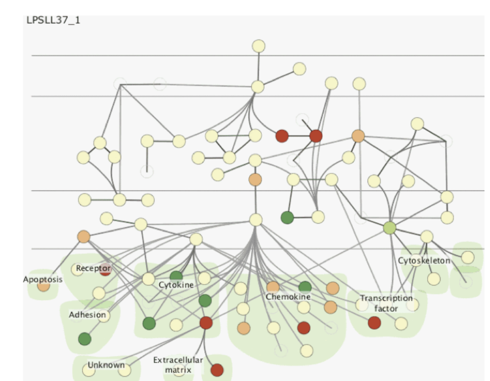




Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, Kincaid. IEEE InfoVis 2008.

Why not animation?

- global comparison difficult



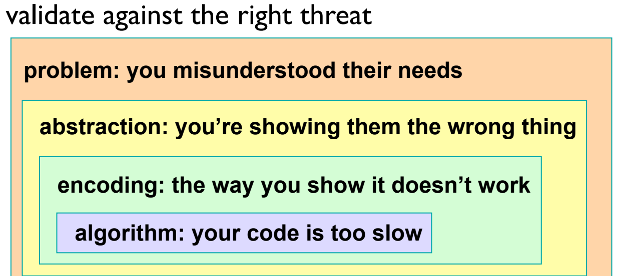
Why not animation?

- further reading

Animation: can it facilitate? Tversky et al. Intl Journ Human-Computer Studies, 57(4):247-262, 2002.

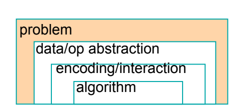
Beyond encoding and interaction

- three more levels of design questions
  - different threats to validity at each level
- validate against the right threat



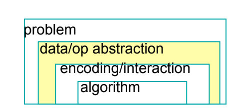
A Nested Model for Visualization Design and Validation. Munzner. IEEE InfoVis 2009.

Characterizing problems of real-world users



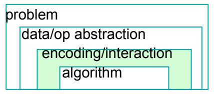
- identify a problem amenable to vis
  - provide novel capabilities
  - speed up existing workflow
- validation
  - immediate: interview and observe target users
  - downstream: notice adoption rates

Abstracting into operations on data types



- abstract from domain-specific to generic
- operations
  - sorting, filtering, browsing, comparing, finding trend/outlier, characterizing distributions, finding correlation...
- data types
  - tables, networks, spatial
  - transform into useful configuration: derived data
- validation
  - deploy in the field and observe usage

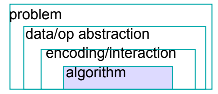
Designing visual encoding, interaction techniques



- visual encoding: drawings they are shown
- interaction: how they manipulate drawings
- validation
  - immediate: careful justification wrt known principles
  - downstream: qualitative or quantitative analysis of results
  - downstream: lab study measuring time/error on given task

- focus of this talk

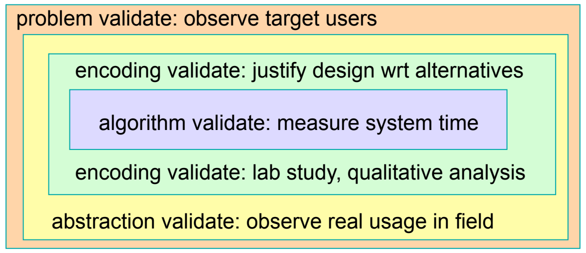
Creating algorithms to execute techniques



- automatically carry out specification
- validation
  - immediate: complexity analysis
  - downstream: benchmarks for system time, memory

Danger of validation mismatch

- cannot show encoding good with system timings
- cannot show abstraction good with lab study



Principles recap

- know your visual channel types and ranks
- categorical color constraints
- power of the plane
- danger of depth
- resolution beats immersion
- eyes beat memory

- validate against the right threat

More information

- this talk
  http://www.cs.ubc.ca/~tmm/talks.html#networkbio12

- papers, videos, software, talks, courses
  http://www.cs.ubc.ca/~tmm

- vis intro book chapter
  - principles in more depth
  - also, techniques!
    http://www.cs.ubc.ca/~tmm/papers.html#akpchapter

- textbook to appear early 2014
  - Visualization Analysis and Design: Abstractions, Principles, and Methods