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
- Measurable definitions of effectiveness

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

- Tabular
- Relational
- Spatial
- Categorical
- Ordered
- Quantitative

Visual encoding

- Analyze as combination of marks and channels showing abstract data dimensions
Image theory

Visual encoding

- analyze as combination of marks and channels showing abstract data dimensions

Visual channel types and rankings

What/where | How much
---|---
plan position | 1
are 2
size 3
line 4
angle 5
shape 6

Visual channel types and rankings

Categorical

- position
- size
- color
- shape
- line
- angle

Ordinal

- plan position
- area
- volume
- lightness
- color saturation
- line pattern

Power of the plane: only position works for all!

Ranking differs for all other channels

Channel rankings

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

**Popout limits**
- only one channel at a time
  - combination searches are serial
    - most channel pairs
    - all channel triples, etc
- within channel, speed depends on which channel and how different item is from surroundings
  - 'sufficiently different': context dependent

**Accuracy**
- position along common scale
  - no scale, framed, aligned
  - frame increases accuracy [Cleveland 99]
  - Weber's Law: relative judgements
    - filled rectangles differ by 1:3
  - white rectangles differ by 1:2

**Discriminability: How many usable steps?**
- linewidth: only a few

**Discriminability: Categorical color constraints**
- noncontiguous small regions of color: only 6-12 bins

**Separability vs. integrality**
- position hue (color)
  - fully separable

**Encoding example: Heatmaps vs. curvemaps**
- color traditional, but spatial position outranks it
  - heatmap
  - curvemap

**Popout limits**
- only one channel at a time
  - combination searches are serial
    - most channel pairs
    - all channel triples, etc
- within channel, speed depends on which channel and how different item is from surroundings
  - 'sufficiently different': context dependent
Dangers of depth: difficulties of 3D
- perspective distortion
  - interferes with all size channel encodings
  - power of the plane is lost!

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 scarce resource
  - desktop also better for workflow integration
- virtual reality for abstract data very difficult to justify

Why not animation?
- global comparison difficult

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

Why not animation?
- further reading
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 of numbers, relational 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

- vis intro book chapter
  - principles in more depth
  - also, techniques!
  - http://www.cs.ubc.ca/~tmm/papers.html#akpchapter
- papers, videos, software, talks, courses
  - http://www.cs.ubc.ca/~tmm
- this talk
  - http://www.cs.ubc.ca/~tmm/talks.html#vizbi11