Visualization & Journalism: Four Vignettes

Tamar Munzner
Department of Computer Science
University of British Columbia
November 2016, Stanford, CA

Overview: current version

Overview: evolution: rational?

Overview: early version

Starting point: Dimensionality reduction for document datasets

Design Study Methodology: Reflections from the Trenches and from the Stacks

Design study methodology: 9-stage framework

Design study methodology: definitions

PF-1 premature advance: jumping forward over stages...
Collaboration incentives
- why do CS/vis people need to understand journalism’s problems?
  - we work with you to understand the problems deeply
  - we observe how you use them
  - CS vis research teaches stories
  - journalism vis access to better tools
- we develop guidelines on how to build better tools in general
- CS wins: research progress in visualization

The general case for curation
- Manual creation process
  - time required for each task
- Structured creation process
  - time required for each task

Threats to validity differ at each level
- Domain situation
  - who are the target users?
  - vital to allow domain to continuously evolve
- Data/task abstraction
  - build mapping from the engineering thing
- Visual encoding/interaction idiom
  - the more you show the less work
- Algorithm
  - efficient computation
  - key feature handles sparseness appropriately

Vignette 2: Vis Tool for Journalistic Presentation

Manual creation process
- Browse
- Extract
- Format
- Show
- Update

Structured creation process
- Browse
- Extract
- Format
- Show
- Update

Origin story: Tedium in the newsroom
- Johanna Fulda: interactive infographics developer, Sueddeutsche Zeitung
  - then Munich CS master’s student, visiting UBC

The general case for curation
- build for human in the loop as continuing need
- automatic processing to accelerate not replace
- assume computational results good but not perfect
- for the interactive feature:
- visual feedback to accelerate
**The importance of being brisk**

- sexy use case: eureka moment
  - enable what was impossible before
  - vis tools for new insights & discoveries
- workhorse use case: workflow speedup
  - vis tools to accelerate what you’re already doing
  - sometimes enables the previously infeasible

**TLC use cases**
- started with speedup use case, for presentation
- make this doc into a timeline now!
- two other use cases nudge towards exploration
- comparison between multiple timelines
- speculative browsing

**Vignette 3: Challenges of Color (A Cautionary Tale)**

**Decomposing color**

- first rule of color: do not talk about color!
  - color is confusing if treated as monolithic
- decompose into three channels
  - ordered can show magnitude
  - luminance
  - saturation
  - categorical can show identity
  - hue

- channels have different properties
  - what they convey directly to perceptual system
  - how much can they convey? how many discriminable bins can we use?

**BallotMaps: deriving data**

- bias exists in regions where systematic structure in bar lengths visible
  - yes in some
  - no in others

**Challenges of Color**

- what is wrong with this picture?
- visualizations that make no sense

**Vignette 4: Difficulties of Depth (Another Cautionary Tale)**

**Four strategies to handle complexity**
- derive new data to show within view
- change view over time
- focus across multiple views
- reduce items/attributes within single view

**Channels: Matching expressiveness**

- Magnitude Channels: Ordered Attributes
- Identity Channels: Categorical Attributes

- expressiveness principle
- match channel and data characteristics

**Channels: Ranking effectiveness**

- expressiveness principle
- match channel and data characteristics
- effectiveness principle
- encode most important attributes with highest ranked channels

**Channels**

- Magnitude Channels: Ordered Attributes
- Identity Channels: Categorical Attributes

- expressiveness principle
- match channel and data characteristics

**Categorical color: limited number of discriminable bins**

- human perception built on relative comparisons
  - great if color contiguous
  - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
  - fewer bins than you were 
  - rule of thumbs: 6-12 bins

**Munzner. AK Peters Visualization Series, CRC Press, 2014.**

**Challenges of Color**

- what is wrong with this picture?
- visualizations that make no sense
Visual encoding: 2D vs 3D

- 2D good, 3D better?
  - not so fast...

● 3D legitimate for true 3D spatial data
● 3D needs very careful justification for abstract data
  - ... 3D for point clouds or networks

Occlusion hides information

• occlusion
• interaction complexity

Perspective distortion loses information

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

3D vs 2D bar charts

• 3D bars never a good idea!

Justified 3D shape perception

• benefits outweigh costs when task is shape perception for 3D spatial data
  - interactive navigation supports synthesis across many viewpoints

No unjustified 3D example: Time-series data

• extruded curves: detailed comparisons impossible

Justified 3D: Economic growth curve

• 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

Wrap-up

• a tale of two tools
  - collaboration between CS and journalism: methods & rewards
  - reasoning about four levels of vis design
  - presentation: TimelineCurator
  - visual curations of imperfect computational results
  - the importance of being brisk: speedup vs eureka moment

• two cautionary tales
  - guidance on color & 3D from vis literature
  - ... two cautionary tales

More Information

- this talk
  - http://amberleyromo.com/images/Bookcover/Animal-Farm.png
  - book
    - http://www.cs.ubc.ca/group/infovis
    - 20% off promo code,
      - presentation: TimeLineCurator

- papers, videos, software, talks, courses
  - http://www.cs.ubc.ca/~tmm/vadbook
  - book+ebook combo: HVN17