Ch 4/5/6: Validation, Marks & Channels, Rules of Thumb
Paper: Artery Vis

Tamara Munzner
Department of Computer Science
University of British Columbia

CPSC 547, Information Visualization
Week 3: 25 September 2017

http://www.cs.ubc.ca/~tmm/courses/547-17F
News

• marks out for last week (Sep 19)
  – most got 5/5 (1 for each of 4 readings, 1 for responses)
  – a few lower who gave only one general comment rather than per-reading specific comments

• today
  – some discussion
  – exercise: Decoding
  – more discussion
Ch 4: Validation
VAD Ch 4: Analysis: Four Levels for Validation

Domain situation

Data/task abstraction

Visual encoding/interaction idiom

Algorithm
Four levels of design and validation

- four levels of design problems
  - different threats to validity at each level

- **Domain situation**
  - You misunderstood their needs

- **Data/task abstraction**
  - You're showing them the wrong thing

- **Visual encoding/interaction idiom**
  - The way you show it doesn’t work

- **Algorithm**
  - Your code is too slow
Validation by level

- **Domain situation**
  Observe target users using existing tools

- **Data/task abstraction**

  - **Visual encoding/interaction idiom**
    Justify design with respect to alternatives

  - **Algorithm**
    Measure system time/memory
    Analyze computational complexity

  - Analyze results qualitatively
  - Measure human time with lab experiment *(lab study)*
  - Observe target users after deployment *(field study)*

  - **Measure adoption**

- mismatch: cannot show idiom good with system timings
- mismatch: cannot show abstraction good with lab study
Directionality & scope

Domain situation

Data/task abstraction

Visual encoding/interaction idiom

Algorithm

problem-driven work

technique-driven work
Ch 5: Marks & Channels
Definitions: Marks and channels

• marks
  – geometric primitives

• channels
  – control appearance of marks
Encoding visually with marks and channels

• analyze idiom structure
  – as combination of marks and channels

1: vertical position
mark: line

2: vertical position
horizontal position
mark: point

3: vertical position
horizontal position
color hue
mark: point

4: vertical position
horizontal position
color hue
size (area)
mark: point
Channels

Position on common scale

Position on unaligned scale

Length (1D size)

Tilt/angle

Area (2D size)

Depth (3D position)

Color luminance

Color saturation

Curvature

Volume (3D size)

Spatial region

Color hue

Motion

Shape
### Channels: Rankings

#### Magnitude Channels: Ordered Attributes
- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

#### Identity Channels: Categorical Attributes
- Spatial region
- Color hue
- Motion
- Shape

• **effectiveness principle**
  - encode most important attributes with highest ranked channels

• **expressiveness principle**
  - match channel and data characteristics
Accuracy: Fundamental Theory

Steven’s Psychophysical Power Law: \( S = I^N \)
Accuracy: Vis experiments

Discriminability: How many usable steps?

- must be sufficient for number of attribute levels to show
  - linewidth: few bins

[mappa.mundi.net/maps/maps 014/telegeography.html]
Separability vs. Integrality

<table>
<thead>
<tr>
<th>Position</th>
<th>Size</th>
<th>Width</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Hue (Color)</td>
<td>+ Hue (Color)</td>
<td>+ Height</td>
<td>+ Green</td>
</tr>
<tr>
<td>Fully separable</td>
<td>Some interference</td>
<td>Some/significant interference</td>
<td>Major interference</td>
</tr>
</tbody>
</table>

- 2 groups each
- 2 groups each
- 3 groups total: integral area
- 4 groups total: integral hue

- Fully separable
- Some interference
- Some/significant interference
- Major interference
Popout

• find the red dot
  – how long does it take?

• parallel processing on many individual channels
  – speed independent of distractor count
  – speed depends on channel and amount of difference from distractors

• serial search for (almost all) combinations
  – speed depends on number of distractors
• many channels: tilt, size, shape, proximity, shadow direction, ...
• but not all! parallel line pairs do not pop out from tilted pairs
Grouping

• containment
• connection

Marks as Links

- Containment
- Connection

Identity Channels: Categorical Attributes

• proximity
  – same spatial region
• similarity
  – same values as other categorical channels
Relative vs. absolute judgements

- perceptual system mostly operates with relative judgements, not absolute
  - that’s why accuracy increases with common frame/scale and alignment
  - Weber’s Law: ratio of increment to background is constant
    - filled rectangles differ in length by 1:9, difficult judgement
    - white rectangles differ in length by 1:2, easy judgement

A \hspace{1cm} B

length

A \hspace{1cm} B

position along unaligned common scale

A \hspace{1cm} B

position along aligned scale

Relative luminance judgements

• perception of luminance is contextual based on contrast with surroundings

http://persci.mit.edu/gallery/checkershadow
Relative color judgements

- color constancy across broad range of illumination conditions

http://www.purveslab.net/seeforyourself/
Ch 6: Rules of Thumb
VAD Ch 6: Rules of Thumb

• No unjustified 3D
  – Power of the plane, dangers of depth
  – Occlusion hides information
  – Perspective distortion loses information
  – Tilted text isn’t legible

• No unjustified 2D

• Eyes beat memory

• Resolution over immersion

• Overview first, zoom and filter, details on demand

• Function first, form next

• (Get it right in black and white)
No unjustified 3D: Power of the plane

• high-ranked spatial position channels: **planar** spatial position – not depth!

**Magnitude Channels: Ordered Attributes**

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)

Steven’s Psychophysical Power Law: $S = I^N$
No unjustified 3D: Danger of depth

• we don’t really live in 3D: we see in 2.05D
  – acquire more info on image plane quickly from eye movements
  – acquire more info for depth slower, from head/body motion

We can only see the outside shell of the world
Occlusion hides information

- occlusion
- interaction complexity

Perspective distortion loses information

• 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]
Tilted text isn’t legible

• text legibility
  – far worse when tilted from image plane

• further reading


No unjustified 3D example: Time-series data

- extruded curves: detailed comparisons impossible

[Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc. InfoVis 99.]
No unjustified 3D example: Transform for new data abstraction

• derived data: cluster hierarchy
• juxtapose multiple views: calendar, superimposed 2D curves

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

- 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

No unjustified 2D

- consider whether network data requires 2D spatial layout
  - especially if reading text is central to task!
  - arranging as network means lower information density and harder label lookup compared to text lists
- benefits outweigh costs when topological structure/context important for task
  - be especially careful for search results, document collections, ontologies
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

35
Eyes beat memory example: Cerebral

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

Why not animation?

• disparate frames and regions: comparison difficult
  – vs contiguous frames
  – vs small region
  – vs coherent motion of group

• change blindness
  – even major changes difficult to notice if mental buffer wiped

• safe special case
  – animated transitions
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

Overview first, zoom and filter, details on demand

• influential mantra from Shneiderman


• overview = summary
  – microcosm of full vis design problem

 Query

 Identify

 Compare

 Summarise

• nuances
  – beyond just two levels: multi-scale structure
  – difficult when scale huge: give up on overview and browse local neighborhoods?

Function first, form next

• start with focus on functionality
  – straightforward to improve aesthetics later on, as refinement
  – if no expertise in-house, find good graphic designer to work with

• dangerous to start with aesthetics
  – usually impossible to add function retroactively
Artery Visualizations for Heart Disease Diagnosis
HemoViz: Design study + evaluation

- formative study with experts
  - task taxonomy
- HemoViz design
- deploy attempt fails
  - experts balk: demand 3D and rainbows
- quantitative user study
  - med students, real data
  - 91% with 2D/diverging vs 39% with 3D/rainbows
  - experts willing to use

[Fig 1. Borkin et al. Artery Visualizations for Heart Disease Diagnosis. Proc InfoVis 2011.]

Study results: Error
Study results: Time

[Graph showing time comparison between 2D and 3D for different color schemes: All Colors, Diverging, Rainbow. The graphs indicate statistical significance marked with asterisks (*) indicating differences between conditions.]
Critique

• many strengths
  – careful and well justified design, convincing human-subjects experiment
    • bringing visualization best practices to medical domain

• limitation
  – paper does not clearly communicate why colormap is diverging not sequential
    • answer by email
    • doctors care about extremely high and extremely low ESS (scalar) values
      – high values (top of scale, dark grey): extreme blood flow patterns may relate to heart malfunctions - but not imminently life threatening and don't indicate plaque locations
      – low values (bottom of scale, dark red): very diseased regions with lots of plaque, docs care a lot!
      – much debate from doctors on where is boundary between “normal” and “low” ESS values
        » most think below 3 Pa are indicative of disease but many argue other values in the 2-4 range.
        » all docs agree that values below 2 Pa are increasingly dangerous disease levels.
        » thus map has transition at 3 Pa for the diverging point and truly red below 2 Pa

• why continuous not segmented?
  – doctors gain tremendous insight by seeing the subtle patterning of the ESS values
  – particularly varying values in red region - patterns help them understand disease progression and severity
    » especially useful for deciding what types of interventions to prescribe for the patient