Data Visualization Pitfalls to Avoid

Tamara Munzner
Department of Computer Science
University of British Columbia

Why are we faced with these resource limitations?
• computational limits
  – processing time
  – system memory
• human limits
  – attention and memory
  – time
• display limits
  – pixels are precious resource

What visualization techniques can we use?
• domain abstraction
  – translate from specifics to general
• domain visualization
  – translate from general to specifics

Beyond these basics, we need to:

• understand the task
• understand the data
• create the optimal visual representation

Visualizations can help people carry out tasks more effectively

What can we evaluate?

• success of visualization technique
• success of visualization

Conclusion

Future research

References

Why use an external representation?

Why represent all the data?

Summary

• representations reduce information, details matter

• user needs, user interface

• confirm expectations and unexpected patterns

Evaluation

• success of visualization

Future work

• more human factors

• automatic creation of visualizations

• user modeling and adaptation

References

Why use an external representation?

Why represent all the data?

Summary

• representations reduce information, details matter

• user needs, user interface

• confirm expectations and unexpected patterns

Evaluation

• success of visualization

Future work

• more human factors

• automatic creation of visualizations

• user modeling and adaptation

References

Why use an external representation?

Why represent all the data?

Summary

• representations reduce information, details matter

• user needs, user interface

• confirm expectations and unexpected patterns

Evaluation

• success of visualization

Future work

• more human factors

• automatic creation of visualizations

• user modeling and adaptation

References
Color/Lightness constancy: Illumination conditions

- human perception built on relative comparisons
  - great if color contiguous
  - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
  - fewers bins than you want
  - rule of thumb: 6-12 bins, including background and highlights
  - alternatives? this afternoon!

Ordered color: Rainbow is poor default

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
  - alternatives
  - large-scale structure fewer hues
  - fine structure: multiple hues with monotonically increasing luminance (e.g. viridis, Kyvos)

Viridis
- colorful, perceptually uniform, colorblind-safe, monotonically increasing luminance

Ordered color: Rainbow is poor default

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
  - alternatives
  - large-scale structure fewer hues
  - fine structure: multiple hues with monotonically increasing luminance (e.g. viridis, Kyvos)
  - segmented rainbows for binned or compositional

Protanope simulation

- color constancy: simultaneous contrast effect
- color constancy: simultaneous contrast effect
- color constancy: simultaneous contrast effect
- color constancy: simultaneous contrast effect

Categorical color: limited number of discriminable bins

- problems
  - perceptually unordered
  - perceptually nonlinear
- benefits
  - fine-grained structure visible and nameable
- alternatives
  - large-scale structure fewer hues
  - fine structure: multiple hues with monotonically increasing luminance (e.g. viridis, Kyvos)

Ordered color: Blue-Orange is safe

- color channel interactions
  - size heavily affects salience
  - small regions need high saturation
  - large need low saturation
  - saturation & luminance 3-4 bins max
  - also not separable from transparency
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!

No unjustified 3D example: Time-series data
• extruded curves: detailed comparisons impossible

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

No unjustified 3D example: Transform for new data abstraction
• derived data: cluster hierarchy
• juxtapose multiple views: calendar, superimposed 2D curves

Justified 3D: Economic growth curve
(WIRED.com-three-dimensional-web-history-frenes-smith-proc.icitvis.99)

More Information
• this talk
http://www.cs.ubc.ca/~tmm/talks.html#vad17can-morn
• afternoon session in more depth
www.cs.ubc.ca/~tmm/talks.html
• book
http://www.cs.ubc.ca/~tmm/vadbook
– 20% off promo code, book+ebook combo: HVN17
• papers, videos, software, tabs, courses
www.cs.ubc.ca/~tmm/vadcourses