AVL Visualization Workshop Series – Fall 2018
Wednesdays @ 4:00pm – Luddy Hall Vis Lab (IQ-Wall)

Date   Topic                                                      Presenter
Aug. 29 Introduction to Scientific Visualization                William Sherman
Sep.  5 Scientific Visualization with Perseus                    William Sherman
Sep. 12 Scientific Visualization                               William Sherman
Sep. 19 Augmented Reality for Visualization                    William Sherman
Oct.  2 Visualizing Data with Large Data Visualization          Alex. Potts
Oct.  9 Augmented Reality Art                                    Andrew Frueholtz
Oct. 16 Interactive Visualizations                              William Sherman
Oct. 30 Image-Based, Volumetric & Scatter Displays for Research  Cheng Wei
Nov.  7 3D Visualization, History, and Arch-It 10             Carly Bonier & Bruce Aal
Dec.  5 Brain Diffusion Imaging with DTI                          David Reagan

Scientific Visualization Workshop Series

Introduction to Scientific Visualization
(Data, Perception, Design)

William Sherman
Advanced Visualization Lab
Indiana University
August 29, 2018

Visualisation is not new...

Al-Biruni – time series visualization phases of the moon in art (c. 1020)

Herford map – largest surviving map of the Middle Ages (1280)

John Snow – cases of cholera in London (1849)
John Snow - cases of cholera in London (1849)

Visualization for Scientific Computing

The Visualization Process
(for computational science)
The Visualization Process (for computational science)

Stages:
- Phenomenal
- Mathematical
- Procedural
- Numerical
- Graphical

The Visualization Process (for computational science)

Stages:
- Phenomenal
- Mathematical
- Procedural
- Numerical
- Graphical

The Visualization Process (for computational science)

Stages:
- Phenomenal
- Mathematical
- Procedural
- Numerical
- Graphical

The Visualization Process (for computational science)

Stages:
- Phenomenal
- Mathematical
- Procedural
- Numerical
- Graphical
Qualitative Insights

- Anscombe's Quartet

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>8.0</td>
<td>9.0</td>
<td>7.5</td>
<td>8.0</td>
</tr>
<tr>
<td>y</td>
<td>9.0</td>
<td>10.0</td>
<td>7.5</td>
<td>8.0</td>
</tr>
</tbody>
</table>

n = 11
mean x = 9.0
mean y = 7.5
variance x = 11.0
variance y = 4.12
correlation = r = 0.816
regression line: y = 3 + 0.5x

Visualization to the rescue

- Visualization = "trading zone" (Peter Galison, Harvard): similar to how different cultures are able to exchange goods, despite differences in language and culture.
- Might need the help of an agent = someone with enough knowledge of both languages/cultures to be able to translate and facilitate the trade.
- Trading zone between:
  - data and scientist
  - scientist (biologist) and scientist (engineer)
  - scientist and public

Visual Analytics for exploration

- Filters what is presented to the participant
- Beyond the sensorial
- Cognitive processes that enhance comprehension:
  - Generalization
  - Semiotics
  - Ideas taking Shape (Representing Ideas)
  - Form: The Gestalt approach to perception
Generalization

- **generalize**: grouping together objects and concepts with similar characteristics
- **analogy**: extending the relationship between grouped objects/concepts to conclude that other characteristics true of one are true of the other

Semiotics

- **semiotics**: the study of signs and symbols
- **sign**: a mark having a conventional meaning and used in place of words to represent a complex notion
- **symbol**: an arbitrary or conventional sign used in writing or printing related to a particular field to represent operations, quantities, elements, relations, or qualities


Semiotics

- The written word is a complex semiotic system
- New symbols are always being developed
- It is best for the meaning of a new symbol to be easily discernable upon first viewing
- Symbols are culturally dependent:
  - $$/\£/¥/€$

Ideas taking shape

- Representations of the realistic to the symbolic
  - **P-reps**: representations tied to a physical (or imagined physical) entity
  - **C-reps**: representations of things that can only be conceptualized such as emotions, risk, potential, danger, etc.
- Images have meanings & implications:
  - **denotations**: the literal meanings of objects in a scene
  - **connotations**: the implications suggested by the scene that are not explicitly presented

Gestalt approach to perception

- Early psychological school of thought
  - Analysis of perception of form → **gestalt** (in German)
  - No longer considered scientifically rigorous
  - Still works well as principles of design
  - “The Whole is other than the sum of the parts”
  - The relationship between parts of some object (within a holistic context) leads to emergent pattern
  - **Gestalt**
  - NOTE: often mis-quoted
Gestalt approach to perception

- Four properties of form perception:
  - Emergence: “The whole is identified before the parts”
  - Reification: “The mind fills in the gaps”
  - Multi-stability: “The mind seeks to avoid uncertainty”
  - Invariance: “We’re good at recognizing similarities and differences”

Gestalt approach

- Four properties of form perception:
  - Emergence
  - Reification
  - Multi-stability
  - Invariance

Gestalt approach to design

- Thirteen Principles of Gestalt Perception/Design:
  - Law Of Prägnanz
  - Closure
  - Symmetry and Order
  - Uniform Connectedness
  - Common Regions
  - Proximity
  - Continuation
  - Common Fate (Synchrony)
  - Parallelism
  - Similarity
  - Focal Points
  - Past Experiences

Gestalt approach

- Thirteen Principles of Gestalt Perception/Design (1-4):
  - Law Of Prägnanz: “People will perceive and interpret ambiguous or complex images as the simplest form(s) possible.”
  - Closure: “When seeing a complex arrangement of elements we tend to look for a single, recognizable pattern.”
  - Symmetry and Order: “People tend to perceive objects as symmetrical shapes that form around their center.”
  - Figure vs. Ground: “Elements are perceived as either figure (the element in focus) or ground (the background on which the figure rests)”

Gestalt approach

- Thirteen Principles of Gestalt Perception/Design (5-8):
  - Uniform Connectedness: “Elements that are visually connected are perceived as more related than elements with no connection.”
  - Common Regions: “Elements that are perceived as part of a group if they are located within the same closed region.”
  - Proximity: “Objects that are closer together are perceived as more related than objects that are further apart.”
  - Continuation: “Elements arranged on a line or curve are perceived as more related than elements not on the line or curve.”

Gestalt approach

- Thirteen Principles of Gestalt Perception/Design (9-11):
  - Common Fate (Synchrony): “Elements that move in the same direction are perceived as more related than elements that are stationary or that move in different directions.”
  - Parallelism: “Elements that are parallel to each other are seen as more related than elements not parallel to each other.”
  - Similarity: “Elements that share similar characteristics are perceived as more related than elements that share less of the characteristics.”
Gestalt approach

  - Focal Points – “Elements with a point of interest, emphasis or difference will capture and hold the viewer’s attention”
  - Past Experiences – “Elements tend to be perceived according to an observer’s past experience”

[Images of a digital interface and a diagram showing a cognitive task => perceptive task]
Limitations of preattentive vision

1. Speed depends on which channel (use one that is good for categorical)

2. Combining pre-attentive features does not always work: we would need to resort to "serial search" (most channel: pain; color channel: shape)

   e.g., is there a red square in this picture?
Apophenia
- Human tendency to perceive meaningful patterns within random data
- Leads you to believe, wrongly, that you have evidence to support a position when you don’t

Mapping
- Originates from cartography: “mappa mundi”
- "mapping": relating specific data values to a perceptible form
  - Red indicates rapid flow
  - Sound (“beep beep beep”) indicates a machine moving in reverse
  - Red indicates an oxygen atom
- Similar perceptible forms may have different meaning in different cultures
  - Astrophysics: red → increased temperature
  - Geology: red → decreased density

Quantitative vs. Qualitative Representations
- Quantitative:
  - Requires a means to access the numbers
  - Provides the details
- Qualitative:
  - Useful for showing general trends and overviews of massive collections of data
  - Used to simplify data in the extreme

Quantitative Measurements
- Different “levels” of how to quantify particular collections of data:
  - Nominal: name or identity of elements
  - Ordinal: elements with a rankable ordering
  - Interval: measurable units
  - Ratio: measurable units w/ absolute zero

Quantitative vs. Qualitative Representations
- Quantitative:
- Qualitative:
Qualitative Insights

- Anscombe's Quartet

<table>
<thead>
<tr>
<th></th>
<th>x1</th>
<th>y1</th>
<th></th>
<th>x2</th>
<th>y2</th>
<th></th>
<th>x3</th>
<th>y3</th>
<th></th>
<th>x4</th>
<th>y4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>8.04</td>
<td></td>
<td>10</td>
<td>9.14</td>
<td></td>
<td>10</td>
<td>7.46</td>
<td></td>
<td>8</td>
<td>6.58</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6.95</td>
<td></td>
<td>8</td>
<td>8.14</td>
<td></td>
<td>8</td>
<td>6.77</td>
<td></td>
<td>8</td>
<td>5.76</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>7.58</td>
<td></td>
<td>13</td>
<td>8.74</td>
<td></td>
<td>13</td>
<td>12.74</td>
<td></td>
<td>8</td>
<td>7.71</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>8.81</td>
<td></td>
<td>9</td>
<td>8.77</td>
<td></td>
<td>9</td>
<td>7.11</td>
<td></td>
<td>8</td>
<td>8.84</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>8.33</td>
<td></td>
<td>11</td>
<td>9.26</td>
<td></td>
<td>11</td>
<td>7.81</td>
<td></td>
<td>8</td>
<td>8.47</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>9.96</td>
<td></td>
<td>14</td>
<td>8.10</td>
<td></td>
<td>14</td>
<td>8.84</td>
<td></td>
<td>8</td>
<td>7.04</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>7.24</td>
<td></td>
<td>6</td>
<td>6.13</td>
<td></td>
<td>6</td>
<td>6.08</td>
<td></td>
<td>8</td>
<td>5.25</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>4.26</td>
<td></td>
<td>4</td>
<td>3.10</td>
<td></td>
<td>4</td>
<td>5.39</td>
<td></td>
<td>19</td>
<td>12.50</td>
</tr>
<tr>
<td>9</td>
<td>12</td>
<td>10.84</td>
<td></td>
<td>12</td>
<td>9.13</td>
<td></td>
<td>12</td>
<td>8.15</td>
<td></td>
<td>8</td>
<td>5.56</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>4.82</td>
<td></td>
<td>7</td>
<td>7.26</td>
<td></td>
<td>7</td>
<td>6.42</td>
<td></td>
<td>8</td>
<td>7.91</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>5.68</td>
<td></td>
<td>5</td>
<td>4.74</td>
<td></td>
<td>5</td>
<td>5.73</td>
<td></td>
<td>8</td>
<td>6.8</td>
</tr>
</tbody>
</table>

- Anscombe's Quartet

$\begin{align*}
\text{correlation } x \text{ & } y &= 0.816 \\
\text{regression line: } y &= 3 + 0.5x \\
\text{mean } x &= 9.0 \\
\text{mean } y &= 7.5 \\
\text{variance } x &= 11.0 \\
\text{variance } y &= 4.12 \\
\end{align*}$
Scientific Visualization & Information Visualization

Scientific Data is of course Information, but “sci-vis” tends to use a specific subset of data types and visualization methods.

- **Scientific visualization**: concerned with situations where spatial position is *given* with the dataset. A central concern is handling continuous data appropriately within the mathematical framework of signal processing.
- **Information visualization** (info-vis): concerned with situations where the use of space in a visual encoding is *chosen* by the designer. A central concern is determining whether the chosen idiom is suitable for the combination of data and task.

Spatial Fields

Spatial Fields are typically at least one form for which scientific visualizations are based.

Munzner:
- **Spatial field**: cell structure is based on sampling at spatial positions.
- E.g. a dataset generated with a medical imaging instrument.

Good Visualization
Rules of thumb

(partly based on Munzner)
1. maximize data-to-ink ratio (Edward Tufte)
2. beware of the lie-factor
3. no unjustified 3D
4. eyes beat memory
5. focus + context
6. overview first, zoom & filter, details on demand
7. don’t overengineer

1. maximize data-to-ink ratio

*data-to-ink ratio* = proportion of graphical that can be read

2. beware of the lie-factor

3. no unjustified 3D

4. eyes beat memory

animation vs side-by-side views:

- switch between different views that are visible at same time = lower cognitive load than consulting memory to compare current view with what was seen before
- try to represent dynamic processes in a static way
5. focus & context

Show selected regions in greater detail (focus)
Preserve global view at reduced detail (context)
No occlusion (all information is visible simultaneously)
(Kaehler, 2003)

6. overview first, zoom & filter, details on demand

Task taxonomy Ben Schollemann:
- Overview: see overall patterns in data
- Zoom: see a subset of data
- Filter: see a subset based on values
- Detail on demand: see values of items
- Relate: compare values
- History: keep track of actions
- Extract: mark and capture

7. “underengineer”, if possible

Important advantage of human in the loop vs. algorithm: you can take shortcuts
- Keep interaction simple (see mouse position & data filter flight patterns: blog - hands-on visualisation using p5)
- Might be OK to not handle edge cases in fast prototyping
- Simple raw data visualization can have emergent properties

#badgraphoftheday
**Visualization dichotomy**

- **Batch vs. Interactive**
  - Interactive:
    - Expansive
    - Immediate feedback
    - Performance constraints
  - Batch:
    - Pre-determined
    - Big data / big processing
    - HPC ready

**Visualization process**

- **Interactive vs. Batch**
  - Interactive:
    - Explore
    - Reduced data
  - Batch:
    - Gated by exploration
    - Full data / high res

**Example Batch Renderings**

- Some examples of end results after all the interacting
  - Pulsar with a planet?
  - Smog in LA
  - Vi-o-Matic

**Visualization Tools**

[available at IU]

[i.e. getting ready for next week]
IU HPC Systems

- Research Desktop (aka RED) on Carbonate
  - Requires Carbonate Account
  - https://access.iu.edu/Account
  - Requires "pivince" software
  - https://access.iu.edu/pivince
  - Available for:
    - MS Windows
    - Apple OS X
    - Linux
- Provides
  - Ready-to-use software
  - Desktop interface to HPC system

Next Week: ParaView Visualization Tool

- ParaView
  - Available on Carbonate Research Desktop
  - Binaries (or Source) available for:
    - MS Windows
    - Apple OS X
    - Linux
  - Participants can follow along, but we won’t have time for waiting for software to be installed, etc.

Fin