Uncertainty Part II:
Node-Edge Graphs & Terrain

Today

• “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”
  – Visual Variables (originally from Bertin)
• “Visualizing Uncertain Information”, MacEachren, Cartographic Perspectives, 1992
  – "Quantitative Texton Sequences for Legible Bivariate Maps"
  – Prof Franklin’s work on Observer Siting
• Homework 7
• Readings for Next Time
• Institutional Review Board (IRB)
• Visual.ly's Code of Ethics for Data Visualization Professionals
  • Society of Professional Journalists' Code of Ethics
Reading for Today

• “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”
  Guo, Huang, and Laidlaw, IEEE TCVG 2015

<table>
<thead>
<tr>
<th>Bertin’s Original Visual Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Position</strong></td>
</tr>
<tr>
<td><strong>Size</strong></td>
</tr>
<tr>
<td><strong>Shape</strong></td>
</tr>
<tr>
<td><strong>Value</strong></td>
</tr>
<tr>
<td><strong>Colour</strong></td>
</tr>
<tr>
<td><strong>Orientation</strong></td>
</tr>
<tr>
<td><strong>Texture</strong></td>
</tr>
</tbody>
</table>

Table 1: These are Bertin's visual variables

• Visual Variables: color based, focus based, geometry
    • location/position, size/width, color hue, color value/brightness, grain, orientation, and shape
  – J. L. Morrison [1974]
    • Color saturation, arrangement
  – Cleveland & McGill [1984]
    • Angle, volume curvature
  – A. M. MacEachren [1992]
    • Focus/fuzziness, resolution, transparency
  – M. Carpendale [2003]
    • Motion, depth, occlusion

• Encode “strength” (placeholder data)
  – Width, hue, or saturation

• Encode uncertainty using a visual variable
  – Lightness, fuzziness, grain, or transparency

• Their conclusions only apply to line based marks
• Which visual variables are most salient? Are most discriminable?
• Disassociativity of each pair of visual variables: Can you differentiate changes in one variable while ignoring changes in another variable?
• Are some visual variables more appropriate (more natural/intuitive?) for certain data?
• Evaluate by studying response time, accuracy
• How much can we (should we?) pack into a single visualization?
• Tangent: Are we good at multitasking?

• Random graphs
• Each edge one of 5 values for “strength” and “uncertainty”
• Locate one edge of a specific value (max or min) of strength or uncertainty that must be identified (or determined to be missing)
  – “find extremum”
  – “retrieve value”
  – “visual search”
• Which graph has overall higher strength or uncertainty?
  – “characterize distribution”
  – “identification-comparison”
  – “visual aggregation”
• Varied the relative discriminability of the two variables
  – Perception of the weaker one is better when they are more similar
• Large number of hypotheses
• 20 participants, 1 hour each, 5760 trials
  – Short teaching/training session with feedback on correctness
  – (personally don’t want to have to administer a user study!)
• Provide explicit design recommendations, useful reference
• Not surprised that lightness interferes with hue and width confused with fuzziness; Surprised that grain performed well
• Well written
  – clearly state hypotheses, justified their conclusions well
  – I could recreate the results from this paper
  – “open questions” instead of future work
• How would the results be different with people with visual training (not novices)?
• How would the results be different for colorblind users?
• Would have liked to see a real-world example of this graph style. And specifically high density graphs (requires thin edges).

Today

• “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”
  – Visual Variables (originally from Bertin)
• “Visualizing Uncertain Information”, MacEachren, Cartographic Perspectives, 1992
  – "Quantitative Texton Sequences for Legible Bivariate Maps"
  – Prof Franklin’s work on Observer Siting
• Homework 7
• Readings for Next Time
• Institutional Review Board (IRB)
• Visual.ly's Code of Ethics for Data Visualization Professionals
• Society of Professional Journalists' Code of Ethics
GIS=Geographic Information System
Are maps & computers infallible? Less fallible than humans?
Data “quality” (when referring to error)? Better term is “uncertainty”?
- Incomplete (census response rate),
- attribute inaccuracy (misunderstood survey question or deliberate misinformation),
- spatial inaccuracy (typos introduced by census taker),
- temporal uncertainty
What is the importance of uncertainty relative to map data?
“...few GIS users are trained in cartographic symbolization & design…”
Proposals for map uncertainty:
- Contour crispness, fill clarity, fog, resolution or Dynamic/Animation (e.g., blinking)
User Interface: Map pairs, sequential presentation, bivariate maps
Evaluation/Experiments needed to confirm prior work and test hypotheses of proposed changes
- Type I Visualization Error: tendency to see patterns that do not exist
- Type II Visualization Error: failure to notice patterns or relationships
Related Reading

Today

- “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”
  - Visual Variables (originally from Bertin)
- “Visualizing Uncertain Information”, MacEachren, Cartographic Perspectives, 1992
  - "Quantitative Texton Sequences for Legible Bivariate Maps"
  - Prof Franklin’s work on Observer Siting
- Homework 7
- Readings for Next Time
- Institutional Review Board (IRB)
- Visual.ly's Code of Ethics for Data Visualization Professionals
- Society of Professional Journalists' Code of Ethics

Reading for Today


Figure 3. The pattern of grid cells included in viewsheds derived from alternative methods of inferring elevations.
• Software quality & different/inconsistent outcomes from different GIS products
• Simple computation/algorithm (calculate slope) vs. more complex algorithms (calculate drainage network, representation conversion raster/grid->vector/mesh)
• Viewshed challenges:
  – Vegetation
  – Earth curvature
  – Small errors can result in large changes
  – “Ground truth” comparison not possible
• How to compute line of sight?
  – Bilinear interpolation
  – Convert to triangle mesh (but which diagonal?)
  – Subgrid w/ interpolation (is this the same as bilinear?)
  – Stepped (constant height within each cell), a.k.a. nearest neighbor
  – Point-to-point, point-to-cell, cell-to-point, cell-to-cell?
• Rounding errors in the geometry computation can yield inconsistent outcomes with what should be geometrically identical comparisons

“Siting Observers on Terrain”
Wm Randolph Franklin, RPI ECSE
Observers have a specified maximum straight line sight distance
Some observer placements see more (black)
Some are occluded

Classic terrain height visualization
Red = higher elevations
blue = lower elevations
• Let’s place “observers” (e.g., cell phone towers) on a complex terrain
  – Where should they be placed to maximize coverage?
    • What if the observers need to see each other? (form a connected network for communication)
  – How much error is introduced because of the original sensor measurements (discrete sample points might miss significant ridges or valleys)?
  – How much error is introduced if the dataset is compressed for storage or transmission and then lossily reconstructed?
    • Erroneous visibility, Erroneous occlusions
  – Knowing the terrain and placement of “red team” observers what path should the “blue team” take to avoid being seen?
  – Knowing that it will be used to do siting tasks, can you design a better compression algorithm that reduces lossy artifacts that cause significant errors?
Regular grid of height samples

Query for occlusions along sight line

Data interpolation might be incorrect!

Hue = terrain height
Value (binary color/black) = visibility
Grain/texture = uncertainty

If height is changed by epsilon, the visibility flips!

The visibility of one half of the points in uncertain!
Today

- “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”
  - Visual Variables (originally from Bertin)
- “Visualizing Uncertain Information”, MacEachren, Cartographic Perspectives, 1992
  - "Quantitative Texton Sequences for Legible Bivariate Maps"
  - Prof Franklin’s work on Observer Siting

- Homework 7
- Readings for Next Time
- Institutional Review Board (IRB)
- Visual.ly’s Code of Ethics for Data Visualization Professionals
- Society of Professional Journalists’ Code of Ethics

Assignment 7: Final Project Ideas

- Invent 2 different Final Project Ideas
  - “Who” (audience), "why" (research question), "what" (the finished visualization)
  - One technical challenge for the project. What makes it difficult? What is a potential “risk” for completion? For example:
    - acquiring the data,
    - working with very large data,
    - implementing a new visualization design,
    - implementing a novel interaction scheme, or
    - revising the visualization design to validate your hypothesis.
  - Do you already have a partner?
- Make Submitty forum post by Thursday 3/8 @ midnight
- Read & reply to 3 other students by Monday 3/19 (after break)

I will post projects from past semesters on course webpage!
Readings for Friday


Today

• “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”  
  – Visual Variables (originally from Bertin)
• “Visualizing Uncertain Information”, MacEachren, Cartographic Perspectives, 1992  
  – "Quantitative Texton Sequences for Legible Bivariate Maps"
  – Prof Franklin’s work on Observer Siting
• Homework 7
• Readings for Next Time
  • Institutional Review Board (IRB)
  • Visual.ly's Code of Ethics for Data Visualization Professionals
  • Society of Professional Journalists' Code of Ethics
• World War II: German physicians conducted medical experiments on concentration camp prisoners without their consent. Tested blood clotting (shooting them), vaccines (infecting them), effectiveness of poison bullets, and effects of high altitude and low oxygen.

• In the 1950's, thalidomide given to pregnant women to help with sleep and nausea, but they did not know it was experimental nor did they give consent.

• Tuskegee, Alabama (1940s-1970s): Low-income African-American males with high incidence of syphilis infection were given free medical examinations, but not told about their disease, and researchers intervened to prevent treatment.

• 1961, Milgram obedience study (the shock machine): lack of proper attention to debriefing, didn't reveal the purpose of the study, didn't comfort subjects ethical qualms about having inflicted pain on a fellow human, didn't offer his participants an opportunity to opt out of the study.

• Zimbardo's prison experiment (Stanford): The study did meet the criteria of his IRB in 1973!

• In the 70's various federal regulations established IRB at all research institutions.

Institutional Review Board (IRB)

• Privacy, Confidentiality, Anonymity, and Informed Consent

• Reduce risk (physical/mental/privacy) to the participants engaged in research
Institutional Review Board (IRB)

Rensselaer Polytechnic Institute is committed to protecting the rights and welfare of human subjects of research conducted on the campus or sponsored by the institute.

Rensselaer subscribes to the basic ethical principles that underlie the conduct of biomedical and behavioral research involving human subjects as set forth in the Belmont Report, and in accordance with Title 45, Code of Federal Regulations, Part 46.

The Institutional Review Board (IRB) has the responsibility and authority to review, approve, disapprove, or require changes in research activities involving human subjects. This policy applies to all faculty, staff, and student projects, regardless of whether the project is funded externally, internally, or receives no funding support.

Researchers should refer to Rensselaer's Guidelines for Human Subjects Research to determine whether or not their research is indeed human subjects research, and/or if their research satisfies the requirements for expedited review by the IRB.

IRB Training Requirements

As federally mandated and required by the Rensselaer IRB, all investigators must complete a self-study course in human subject protection via the CITI Training Program. Each investigator on a research project involving human subjects is required to certify that they have completed the required course(s) before seeking IRB approval.

https://oasis.cs.rpi.edu/
Spatially Augmented Reality (SAR) Projection

camera detects design geometry

6 projectors augment design

design sketched with foam-core walls
Today

• “Representing Uncertainty in Graph Edges: An Evaluation of Paired Visual Variables”
  – Visual Variables (originally from Bertin)
• “Visualizing Uncertain Information”, MacEachren, Cartographic Perspectives, 1992
  – "Quantitative Texton Sequences for Legible Bivariate Maps"
  – Prof Franklin’s work on Observer Siting
• Homework 7
• Readings for Next Time
• Institutional Review Board (IRB)
• Visual.ly’s Code of Ethics for Data Visualization Professionals
  • Society of Professional Journalists’ Code of Ethics
Society of Professional Journalists' Code of Ethics

• Is this actually used? Impractical (won’t have a long career). Seems like many media outlets do the opposite
  – Examples?
• First amendment, expose negative qualities/actions of people in power
• Create emotions that can influence viewers – Tufte says to avoid garnishes (for this reason?)

http://visual.ly/about/ethics

As an organization that both practices and recognizes quality data-journalism, Visual.ly subscribes to the code of ethics of the Society of Professional Journalists and agrees to abide by all of its principles.

We also agree to the following principles to support data analysis and visualization:
Data will be accurate and verifiable - Visual.ly will not "lie with statistics."
Proper Sourcing & Attribution - Visual.ly will always give credit where due and will do its own reporting.
Best Practices in Visual Representation - Visual.ly will not exploit idiosyncrasies of the human visual system to exaggerate or misrepresent data.
Most succinctly stated, Visual.ly’s policy is one that encompasses accuracy, honesty, and transparency.
While Visual.ly will do our best to promote these standards, the policy applies only to the visualizations we create ourselves and those we feature as staff picks, not to those uploaded by members of the community.
Visual.ly's Code of Ethics for Data Visualization Professionals

• Data analysis is important
• Too narrow focus or omission can lead to bias
• Incorrect analysis must be avoided
• Be open to criticism, learn from past work