Information Visualization

Intro, Time Series Exercise
Tamara Munzner
Department of Computer Science
University of British Columbia
12 September 2017

Visualization (vis) defined & motivated
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively. Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- human in the loop needs the details
  - doesn’t know exactly what questions to ask in advance
  - long-term exploratory analysis
- speed up through human-in-the-loop visual data analysis
- presentation of known results
- stepping stone towards automation: refining, troubleshooting
- intended task, measurable definitions of effectiveness

Audience
- no prerequisites
- many areas helpful but not required
  - human–computer interaction (CSCE 544 this term)
  - computer graphics, cognitive psychology, machine learning, statistics, algorithms, graphic design, <application domain>
- open to non-CS people
  - if no programming background, can do analysis or survey project
- open to advanced undergrads
- talk to me
- open to informal auditors
  - some or all days of readings/discussion/exercises, as you like
  - you get out of it what you put into it...

Readings
- textbook
- readings posted by 6 days before class
- ~4 each session: mix of chapters & papers

Marking
- 50% Project
  - 15% Intermediate Milestones (pass/fail)
    - extensive feedback along the way
    - great: 100%
    - ok: 75%
    - poor: 67%
  - 15% Final Presentation
  - 25% Final Report
  - 5% Course Survey
- 25% Presentations (pass/fail)
  - 75% Content: Analysis 35%, Critique 25%
  - 25% Delivery
- 75% Participation
  - 50% In-Class Exercises
  - 15% Discussion

Schedule, big picture
- once/week, 2-3pm Tuesdays, 12 sessions
- Sep 5, no class: CS 542 class grades, orientation events only
- Sep 12, first class today!
- Oct 3, no class: annual VIS conference
- Dec 5, last class: one week past usual time

Class sessions
- first part: read & participate (30%)
  - before class:
    - you do readings (~4, mix of chapters & papers)
    - you submit comments before class
  - during class:
    - sometimes I lecture (briefly) and we discuss
    - frequent in-class work/exercises/lecture/tactique
  - maybe: presentations (20%)
    - before one of the classes: you read paper I assign on topic of your choice
    - during class that you present it to everybody else (~10-15 min)
    - TBD depending on final enrollment

Projects (50%)
- groups of 2, 3, or 4
- common case (I will only consider supervising students who do these)
  - four types
    - problem-driven design studies (very specific tasks)
    - technique-driven (explore design choice space for encoding or interaction idiom)
    - algorithm implementation (as described in previous paper)
    - algorithm exploration (like today’s exercise)
- analysis
  - use existing tools on dataset
  - detailed domain survey
  - particularly suitable for non-CS students
- survey
  - very detailed domain survey
  - particularly suitable for non-CS students

Class participation
- in-class group/individual exercises
  - workshopping/lecture for projects
- principal part of course, attendance expected

Projects
- groups of 2, 3, or 4
  - common case (I will only consider supervising students who do these)
    - four types
      - problem-driven design studies (very specific tasks)
      - technique-driven (explore design choice space for encoding or interaction idiom)
      - algorithm implementation (as described in previous paper)
      - algorithm exploration (like today’s exercise)
- analysis
  - use existing tools on dataset
  - detailed domain survey
  - particularly suitable for non-CS students
- survey
  - very detailed domain survey
  - particularly suitable for non-CS students

Marking: Previous
- 50% Project
  - 2% Pitches
  - -10% Proposal
  - -4% Source/Writings
  - -4% Project Peer Reviews
  - -12% Final Presentation
  - -18% Final Report
  - -50% Course Survey
- 20% Presentations
  - 75% Course Survey 50%, Analysis 35%, Critique 25%
  - 25% Delivery
- 30% Participation
  - 60% Written Questions
  - 40% In-Class Discussion/Exercises

Projects
- programming
  - common case (I will only consider supervising students who do these)
- analysis
  - use existing tools on dataset
  - detailed domain survey
  - particularly suitable for non-CS students
- survey
  - very detailed domain survey
  - particularly suitable for non-CS students

Class participation
- comments or questions
  - fine to be less formal than written report
  - correct grammar and spelling still expected
  - be concise: one paragraph is good
  - should be thoughtful, show you’ve read and reflected
  - please ask something vital to look up
  - ask to fork for clarification of genuinely confusing section
- comments on what you’re thinking carefully about what you read
  - great to point out something that I haven’t seen before
- examples on http://www.cs.ubc.ca/~tmm/courses/infovis/structure.html

Logistics
- email is the best way to reach me: tmm@cs.ubc.ca
- office hours Tue right after class (5-6pm)
  - will try to catch me by dropping by usually either in meeting or elsewhere
- X661 (W-Xing of ICICS/CS bldg)
- course page is font of all information
  - don’t forget to refresh, frequent updates
  - http://www.cs.ubc.ca/~tmm/courses/547-17F

Comments submission & marking
- written comments on reading in advance, in two rounds
  - round 1 due 9am (5 hrs before class), 90% of comment mark
  - I for each reading
  - bring printout or laptop with you, springboard for discussion
- round 2 due 1:30pm (30 min before class), 10% of comment mark
  - written responses to at least 2 comments per session/week
  - you can only read comments from others after you put your own
- stay as pass/fail marking, see how it goes
  - switch to explicit marking if quality concerns
Projects: Design studies

• BYOD (Bring Your Own Data)
  – you (or your teammates) have your own data to analyze
    • the research topic
    • personal interest
    • dovetail with another course (sometimes works, but timing may be tricky)
    • FDOI (Find Data Of Interest)
      – many existing datasets, see resource page to get started
      - http://www.cs.ubc.ca/~tmm/courses/547-17F/projectdesc.html
        – can be tricky to determine reasonable task

Case A: 3D Approach (Not Recommended)
• extruded curves: detailed comparisons impossible

Case A: Cluster-Calendar Solution
• derived data: cluster hierarchy
• juxtapose multiple views: calendar, superimposed 2D curves

Case B: Stack Zooming
https://youtu.be/dK0De4XPm5Y

Case C: ChronoLenses
http://youtu.be/k7pI8ikczqk

Case D: RankExplorer
http://youtu.be/rdgn1qcZ2A4

Case E: LiveRAC
http://youtu.be/ld0c3H0VSkw

Project examples

• http://www.cs.ubc.ca/~tmm/courses/547-17F/projectdesc.html

Presentations [20%]

• maybe - depends on final enrollment! TBD
• present, analyze, and critique one paper
  – send me topic choices, I will assign papers accordingly
  • expectations
  • slides required
  • summary/description important, but also your own thoughts
  • analysis according to book framework
  • critique of strengths and weaknesses
  • timing
    – exacts times TBD depending on enrollment
      – likely around 10 minutes each
  • topics at http://www.cs.ubc.ca/~tmm/courses/infosys/presentations.html

Next Time
• to read
  – VAD book, Ch 1: What’s Vis, and Why Do It?
  – VAD book, Ch 2: What: Data Abstraction
  – VAD book, Ch 3: Why: Task Abstraction
  – paper: Design Study Methodology

Break

Now: In-class design exercise, in small groups

• Five time-series scenarios
  – A: every 5 min, duration 1 year, 1 thing: building occupancy rates
  – B: every 5 min, 1 year, 2 things: currency values (exchange rate)
  – C: several years and several things: 5 years, 10 currencies
  – D: 1 year, many things: CPU load across 1000 machines
  – E: 1 year, several parameters, many things: 10 params on each of 1000 machines

• Small-group exercise: 15-20 min
  – one group per table (4-5 people/group)
  – discuss/sketch possible visual encodings appropriate for your assigned scenario
  – Reportback: 20-30 min
  – 3 min from each group
  – Design space examples/discussion: 20-30 min

Case A: 3D Approach (Not Recommended)
• extruded curves: detailed comparisons impossible

Case E: LiveRAC video
http://youtu.be/dK0De4XPm5Y

Break

Case B: Stack Zooming
https://youtu.be/dK0De4XPm5Y

Case C: ChronoLenses
http://youtu.be/k7pI8ikczqk

Next Time
• to read
  – VAD book, Ch 1: What’s Vis, and Why Do It?
  – VAD book, Ch 2: What: Data Abstraction
  – VAD book, Ch 3: Why: Task Abstraction
  – paper: Design Study Methodology

Case D: RankExplorer
http://youtu.be/rdgn1qcZ2A4

Presentations [20%]

• maybe - depends on final enrollment! TBD
• present, analyze, and critique one paper
  – send me topic choices, I will assign papers accordingly
  • expectations
  • slides required
  • summary/description important, but also your own thoughts
  • analysis according to book framework
  • critique of strengths and weaknesses
  • timing
    – exacts times TBD depending on enrollment
      – likely around 10 minutes each
  • topics at http://www.cs.ubc.ca/~tmm/courses/infosys/presentations.html

Next Time
• to read
  – VAD book, Ch 1: What’s Vis, and Why Do It?
  – VAD book, Ch 2: What: Data Abstraction
  – VAD book, Ch 3: Why: Task Abstraction
  – paper: Design Study Methodology