Lecture 2: Design Studies

Information Visualization
CPSC 533C, Fall 2009

Tamara Munzner

UBC Computer Science

Mon, 14 September 2009

News

- questions were due today at 10am by email
  - one question per reading
  - plain (ASCII) text not Word/PDF/etc
- reserve books
  - in CS reading room (ICICS/CS 2nd floor)
  - www.cs.ubc.ca/local/reading/about
  - 24-hr fob access possible
- proxy server
  - note instructions on course page for IEEE/ACM DL access

Required Reading

Cluster and Calendar based Visualization of Time Series Data. Jarke J. van Wijk and Edward R. van Selow. Proc. InfoVis 99, pp 4-9. http://www.win.tue.nl/~vanwijk/clv.pdf

Visual Exploration and Analysis of Historic Hotel Visits. Chris Weaver, David Fyfe, Anthony Robinson, Deryck W. Holdsworth, Donna J. Peuquet and Alan M. MacEachren. Information Visualization (Special Issue on Visual Analytics), Feb 2007. http://www.cs.ou.edu/~weaver/academic/publications/weaver-2007b.pdf

MizBee: A Multiscale Synteny Browser. Miriah Meyer, Tamara Munzner, and Hanspeter Pfister. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09), to appear 2009. http://www.mizbee.org/More_Info_files/mizbee.pdf

Design Study

- describe/characterize task
- abstract up from domain-specific issues
- justify solution

- often: refine until satisfied

- not necessarily new algorithms/techniques

Design Study Definition

Design study papers explore the choices made when applying infovis techniques in an application area, for example relating the visual encodings and interaction techniques to the requirements of the target task. Although a limited amount of application domain background information can be useful to provide a framing context in which to discuss the specifics of the target task, the primary focus of the case study must be the infovis content. Describing new techniques and algorithms developed to solve the target problem will strengthen a design study paper, but the requirements for novelty are less stringent than in a Technique paper.

[InfoVis03 CFP, infovis.org/infovis2003/CFP]

Cluster-Calendar, van Wijk and van Selow

- data: N pairs of (value, time)
  - N large: 50K
- tasks
  - find standard day patterns
  - find how patterns distributed over year, week, season
  - find outliers from standard daily patterns
  - want overview first, then detail on demand
- limitations of previous work
  - predictive mathematical models
    - details lost, multiscale not addressed
  - scale-space approaches (wavelet, fourier, fractal)
    - hard to interpret, known scales lost
  - 3D mountain: x hours, y value, z days
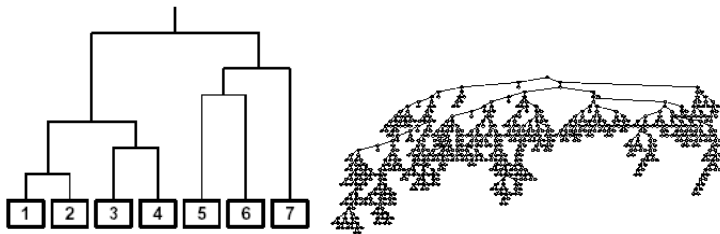
3D Time-series Data

- 3D extrusion pretty but not useful
  - daily, weekly patterns hard to see
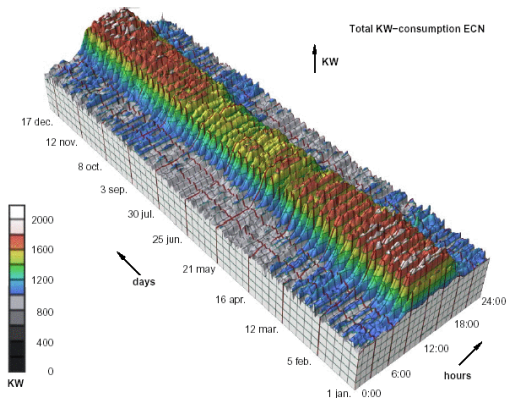
[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, http://www.win.tue.nl/~vanwijk/clv.pdf]
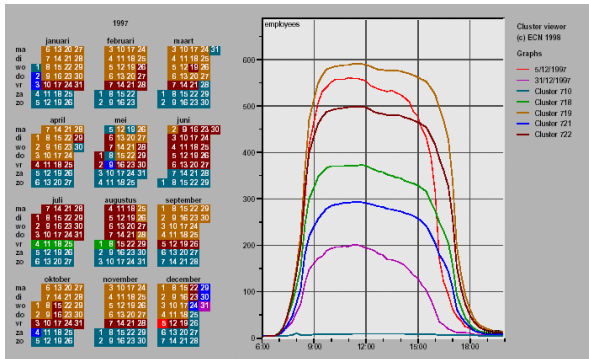
Data Transformation: Hierarchical Clustering

- start with all M day patterns
  - compute mutual differences, merge most similar: M-1
  - continue up to 1 root cluster
- result: binary hierarchy of clusters
- choice of distance metrics
- dendrogram display common
  - but shows structure of hierarchy, not time distribution

[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, http://www.win.tue.nl/~vanwijk/clv.pdf]
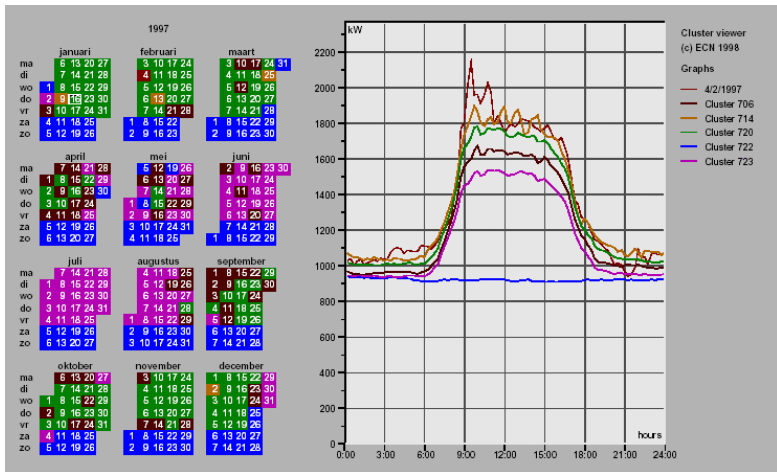
Linked Views: Clusters and Calendar

- single curve for entire cluster as aggregate representation
- calendar for temporal patterns (count of people in building)
  - office hours, fridays in/and summer, school break
  - weekend/holidays, post-holiday, santa claus

[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, http://www.win.tue.nl/~vanwijk/clv.pdf]

Power Consumption

[van Wijk and van Selow, Cluster and Calender based Visualization of Time Series Data, InfoVis99, http://www.win.tue.nl/~vanwijk/clv.pdf]

Key Ideas

- clusters: data transformation to create
- calendar: good existing visual representation for time
- power of linking two different views
  - interactive exploration
- clear task analysis guided choices
  - reject standard 3D extrusion
  - reject standard dendrogram
- critique

Key Ideas

- clusters: data transformation to create
- calendar: good existing visual representation for time
- power of linking two different views
  - interactive exploration
- clear task analysis guided choices
  - reject standard 3D extrusion
  - reject standard dendrogram
- critique
  - color choice not so discriminable
  - especially legend

Historic Hotel Visits, Weaver et al.

- domain
  - historical geography
- data
  - guest name(s)
  - guest occupations (sometimes)
  - geographical location of hotels
  - geographical location where guests live
  - time of visit (day/week/season/year)
- tasks: find visitation patterns
  - periodic temporal patterns
  - commercial, cultural connectivity patterns

Hotel Vis Video
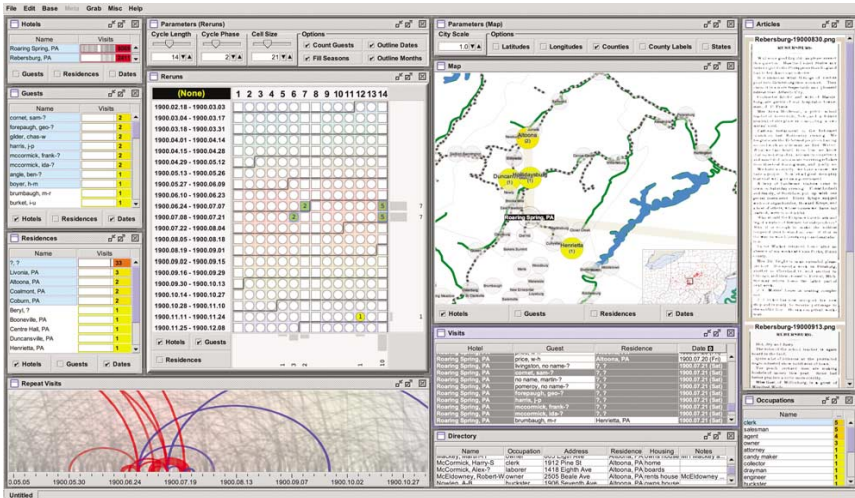
[Fig 4. Weaver et al. Visual Exploration and Analysis of Historic Hotel Visits. Information Visualization 6(1):89–103, 2007.]

Hotel Vis Views

- multilayer map, detail+overview
  - hometowns, railroads, rivers
- many linked sortable tables
  - hotels, guest names, cities, jobs, ...
- arc diagram
  - sequences of guest/group visits
- reruns - cyclic patterns
  - easily change cycle lengths
  - summary histograms
    - horizontal: cycle period
    - vertical: day

Techniques

- coordinated multiple views
  - each view has different strengths
  - linked highlighting across views (brushing)
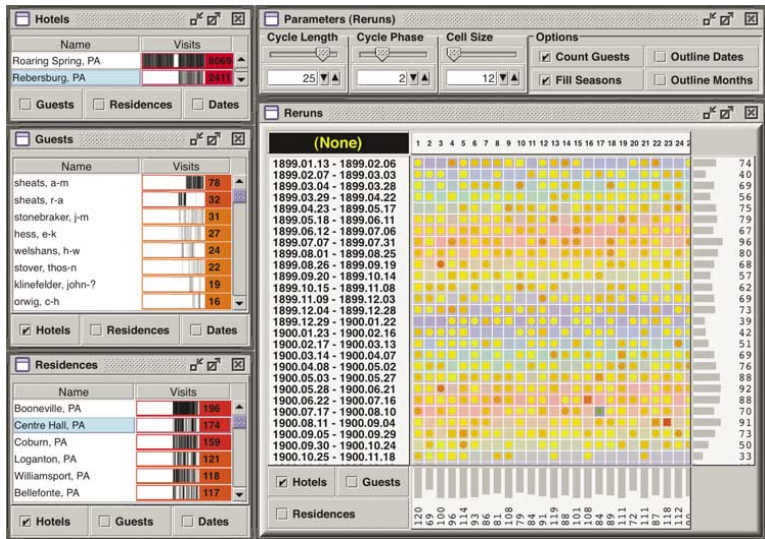 - overview+detail
- grouping
- sorting
- filtering

- iterative refinement
  - many versions over 9 months
  - Improvise: tool for quickly building CMVs

Visit Patterns
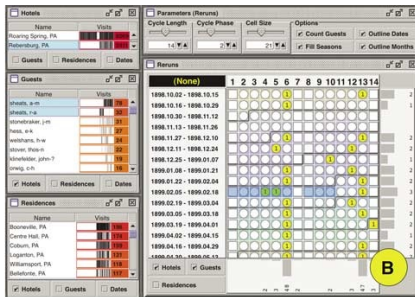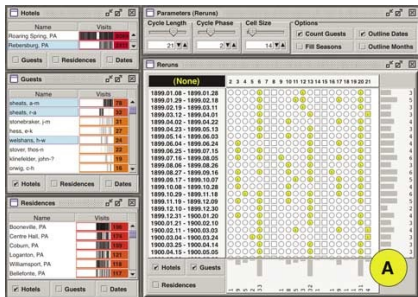
[Fig 5ab. Weaver et al. Visual Exploration and Analysis of Historic Hotel Visits.
Information Visualization 6(1):89–103, 2007.]

Seasonal Variation

[Fig 6. Weaver et al. Visual Exploration and Analysis of Historic Hotel Visits.
Information Visualization 6(1):89–103, 2007.]

Evaluation - Qualitative

- round 1: suggest improvements
- round 2: assess by precepts
  - worldview (was strongly supported)
    - create knowledge
    - find correlations
    - support hypothesis generation
  - rationale (was weakly supported)
    - expose uncertainty
    - present concrete outcomes
    - show possible causation
- round 3: suggest improvements for rationale goals

Key Ideas

- power of linking many views
- power of resortable lists/tables
- arc view technique (from previous work)
- reruns: interactively explore to find interesting cycles
- iterative tool refinement with domain specialists

- critique

Key Ideas

- power of linking many views
- power of resortable lists/tables
- arc view technique (from previous work)
- reruns: interactively explore to find interesting cycles
- iterative tool refinement with domain specialists

- critique
  - Improvise very powerful, but how much learning curve for people besides tool author to get these results?

MizBee, Meyer/Munzner/Pfister

- domain
    - comparative genomics
- data
    - levels: genome, chromosome, block, feature
- task
    - synteny relationships: features on same chromosome
        - proximity/location
        - size
        - orientation
        - similarity

MizBee Video
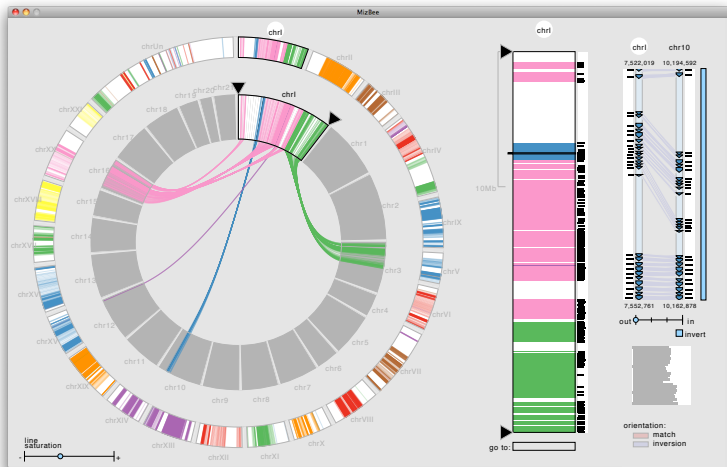
[Fig 1. Meyer, Munzner, and Pfister. MizBee: A Multiscale Synteny Browser. IEEE TVCG 15(6) (Proc. InfoVis 2009).]
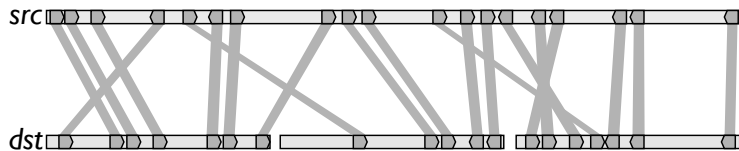
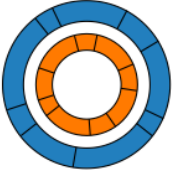# Visual Encoding

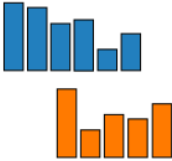- color limits: no info about destination
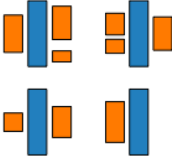 - < 12 distinguishable colors

- connection limits: visual clutter

[Fig 3. Meyer, Munzner, and Pfister. MizBee: A Multiscale Synteny Browser. IEEE TVCG 15(6) (Proc. InfoVis 2009)]
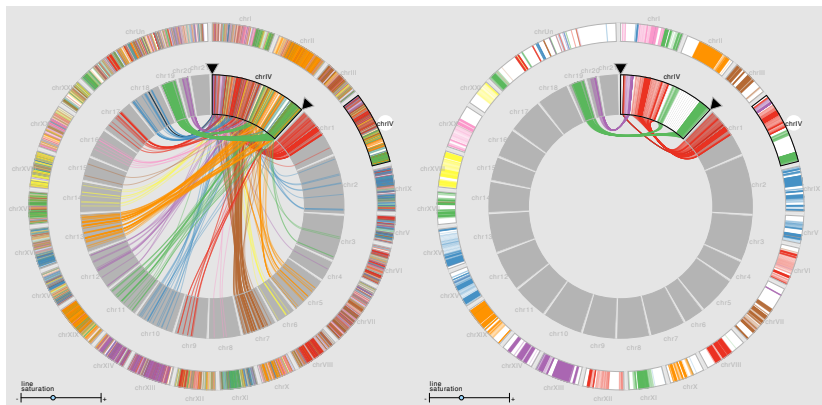
Taxonomy

[Table 2. Meyer, Munzner, and Pfister. MizBee: A Multiscale Synteny Browser. IEEE TVCG 15(6) (Proc. InfoVis 2009)]

Techniques

- linked views: 3 levels to drill down
    - genome: separate-circular, color and connection
        - edge bundling (Lecture 8)
    - chromosome: rectangular, color
        - more screenspace for details
        - histograms for block stats
        - annotations marking feature positions
    - block: connection
        - separate+contiguous histograms for feature stats

Stickleback/Pufferfish Case Study

[Fig 5. Meyer, Munzner, and Pfister. MizBee: A Multiscale Synteny Browser. IEEE TVCG 15(6) (Proc. InfoVis 2009)]

Key Ideas

- power of linked views for multiscale
- abstracting from domain to generic problems
- visual encoding choices according to known limitations
- clutter reduction with edge bundles
- two levels of task: block reliability vs. higher-level science

- critique?

Readings For Next Time

Chapter 1, Readings in Information Visualization: Using Vision to Think. Stuart Card, Jock Mackinlay, and Ben Shneiderman, Morgan Kaufmann 1999. **Handed out in class**

A Nested Model for Visualization Design and Validation. Tamara Munzner. IEEE TVCG 15(6) (Proc. InfoVis 2009), to appear.

MatrixExplorer: a Dual-Representation System to Explore Social Networks. Nathalie Henry and Jean-Daniel Fekete. IEEE Trans. Visualization and Computer Graphics (Proc InfoVis 2006) 12(5), pages 677-684, 2006.

Low-Level Components of Analytic Activity in Information Visualization. Robert Amar, James Eagan, and John Stasko. Proc. InfoVis 05, pp. 111-117.

Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases (extended paper) Chris Stolte, Diane Tang and Pat Hanrahan, IEEE Transactions on Visualization and Computer Graphics, Vol. 8, No. 1, January 2002.