Vorlesung
Advanced Topics in HCI
(Mensch-Maschine-Interaktion 2)

Ludwig-Maximilians-Universität München
LFE Medieninformatik
Heinrich Hußmann & Albrecht Schmidt
WS2003/2004
http://www.medien.informatik.uni-muenchen.de/

Chapter 2: Information Visualization

Table of Content

"Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space."

-- Edward R. Tufte

Tufte – Principles of Graphical Excellence

- Graphical excellence
  - the well-designed presentation of interesting data – a matter of substance, of statistics, and of design
  - consists of complex ideas communicated with clarity, precision and efficiency
  - is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space
  - **requires telling the truth about the data**.

Hearst, 2003

Tufte Principle

Maximize the data-ink ratio
(Avoid "chart junk")

$$\text{Data-ink ratio} = \frac{\text{data ink}}{\text{total ink used in graphic}}$$

Hearst, 2003

Tufte's Graphical Integrity

- Some lapses intentional, some not

$$\text{Lie Factor} = \frac{\text{size of effect in graph}}{\text{size of effect in data}}$$

- Misleading uses of area
- Misleading uses of perspective
- Leaving out important context
- Lack of taste and aesthetics

Hearst, 2003

Lie factor

$$\text{lie factor} = \frac{\text{size of effect shown in graph}}{\text{size of effect in data}}$$

where

$$\text{size of effect} = \frac{|\text{second value - first value}|}{\text{first value}}$$

A lie factor that is either much higher or much lower than one is bad. A **high** lie factor **exaggerates** differences between values. A **low** lie factor **obscures** differences between values.

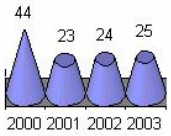
A common example of a **high** lie factor occurs when both dimensions of a two-dimensional figure are made proportional to the same data, so that the size of the figure is proportional to the square of the data; for instance,

Year	Books circulated
2001	100
2002	141
2003	200

where the lie factor is about 2.4.

http://instruct.uwo.ca/fim-lis/504/504gra.htm

An example of a **low** lie factor can be seen in the "Cones" custom chart format in Microsoft Excel.

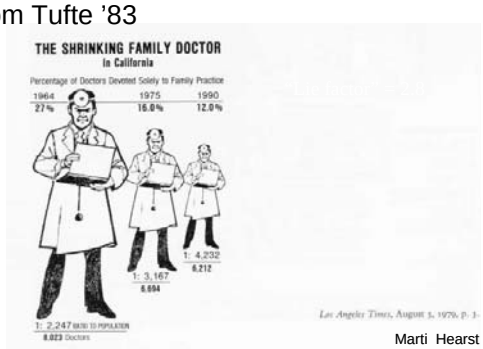


The heights of the (truncated) cones are proportional to the data, but their areas on the screen and their apparent volumes make the larger data values seem relatively small.

Charting on a **logarithmic** scale can also produce a low lie factor.

http://instruct.uwo.ca/fim-lis/504/504gra.htm

How to Exaggerate with Graphs
from Tufte '83



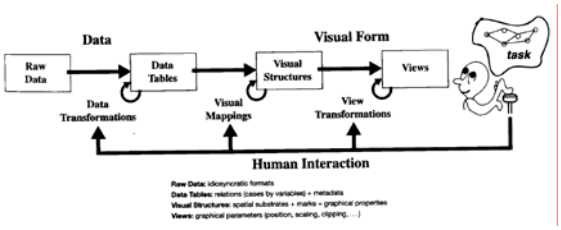
Marti Hearst

How to Exaggerate with Graphs
from Tufte '83

Error: Shrinking along both dimensions

Marti Hearst

Visualization Reference Model
Human Interaction



- Raw Data → Data Table
  filtering
- Data Table → Visual Structure
  pick mappings
- Visual Structure → Views
  probes, viewpoints, distortions

(Storey, 2004)

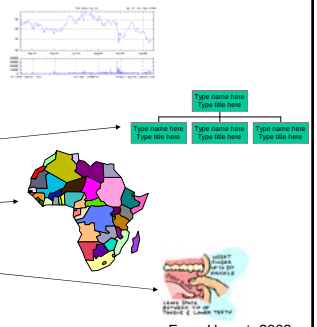
Visualization Reference Model
Overview

DATA TABLES	VISUAL STRUCTURES	VIEWS	HUMAN INTERACTION	TASKS	LEVEL
Cases Variables Values Metadata	Spatial Substrate Marks Graphical properties	Location Probes Viewpoint Controls Distortion	Data Tables Visual Structures Views	Forage for Data Problem Solving Search for Schema Instantiate Schema Author, Decide, or Act	Infosphere Workspace Visual Knowledge Tools Visual Objects
Specific Techniques					
Spatial (Scientific) Geographic Documents Time Database Hierarchies Networks World Wide Web	Position: NOQ Marks: PLAV Properties: Connection, Enclosure, Retinal, Time Axes: Composition Alignment Folding Recursion Overloading	Brushing Zooming Overview + Detail Focus + Context	Dynamic Queries Direct Manipulation Magic Lens	Overview Zoom Filter Details-on-Demand Browse Search Read Fact Read Comparison Read Pattern Manipulate Create	Delete Reorder Cluster Class Promote Average Abstract Instantiate Extract Compose Organize

(Storey, 2004)

Basic Types of Symbolic Displays
(Kosslyn 89)

- Graphs
- Charts
- Maps
- Diagrams



From Hearst, 2003

Basic Types of Data

- Nominal (qualitative)
  - (no inherent order)
  - city names, types of diseases, ...
- Ordinal (qualitative)
  - (ordered, but not at measurable intervals)
  - first, second, third, …
  - cold, warm, hot
- Nominal/Interval (quantitative)
  - list of integers or reals

Hearst, 2003

Data Types - Overview

- **Generic**
  - **entity, relationship,**
  - **Attribute to entity or relationship**
  - **operation**
- **Specific**

1-D Linear	Document Lens, SeeSoft, Info Mural, Value Bars
2-D Map	GIS, ArcView, PageMaker, Medical imagery
3-D World	CAD, Medical, Molecules, Architecture
Multi-Dim	Parallel Coordinates, Spotfire, XGobi, Visage, Influence Explorer, TableLens, DEVise
Temporal	Perspective Wall, LifeLines, Lifestreams, Project Managers, DataSpiral
Tree	Cone/Cam/Hyperbolic, TreeBrowser, Treemap
Network	Netmap, netViz, SeeNet, Butterfly, Multi-trees

Shneiderman, 2003

Information Visualization Mantra

…

Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand
Overview, zoom & filter, details-on-demand

…

Shneiderman, 2003

Information Visualization Tasks

- **Overview** Gain an overview of the entire collection
- **Zoom** Zoom in on items of interest
- **Filter** Filter out uninteresting items
- **Details-on-demand** Select an item or group and get details when needed
- **Relate** View relationships among items
- **History** Keep a history of actions to support undo, replay, and progressive refinement
- **Extract** Allow extraction of sub-collections and of the query parameters

Shneiderman, 2003

Example: PhotoMesa

Information Visualization: Design Guidelines

Direct manipulation strategies

- Visual presentation of query components
- Visual presentation of results
- Rapid, incremental and reversible actions
- Selection by pointing (not typing)
- Immediate and continuous feedback
- Reduces errors
- Encourages exploration

Shneiderman, 2003

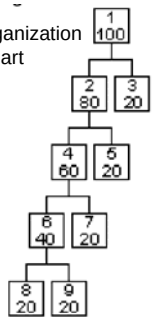
Basic Visualization Techniques

- Finding appropriate visualization for data structures
- Example: trees / graphs

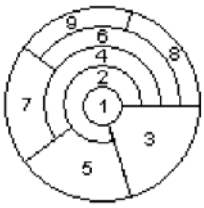
  (n1;(n2;n3,n4),(n5;(n9;n10,(n11;(n12;n14,n15),n13),(n6;n7,n8))
- Is there a more readable way to show it?

Alternative Tree Visualization

Organization Chart

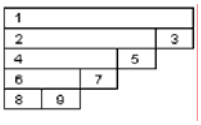


Tree Ring

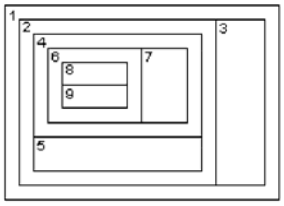


Alternative Tree Visualization

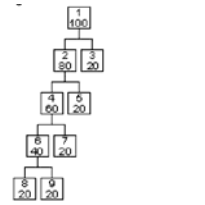
Icicle Plot

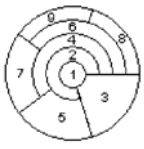


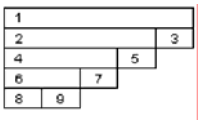
Tree Map

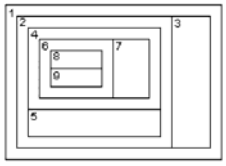


Comparing Visualizations









Typical Tasks for viewing Trees

- Determine the type of tree, e.g.
  - Binary
  - N-ary
  - Balanced
  - Unbalanced
- Find relations, e.g.
  - Deepest common ancestor
- Size of the tree, e.g.
  - How many levels
  - How many leaves
- Details about leaves, e.g.
  - Larges leaf
- Different representation may be better for a given task, e.g.
  - To find out if a tree is balanced or how many levels exist the Icicle Plot is good

More details see:

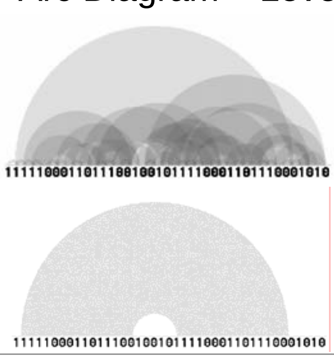
Barlow et al. “A Comparison of 2-D Visualizations of Hierarchies” INFOVIS’01
http://www.sims.berkeley.edu/courses/is247/s02/readings/barlow.pdf

Arc Diagrams

- Visualization method
  - For representing complex patterns of repetition in string data.
  - Arc diagrams scale efficiently for strings that contain many instances of the same subsequence.
  - idea of visualizing only a subset of all possible pairs of matching substrings.
  - highlight just the subsequences essential to understanding the string's structure

Arc Diagrams - Basics

Arc Diagram – Level of Detail

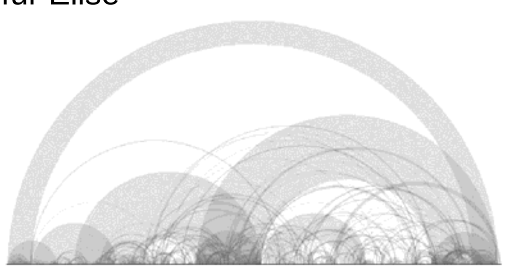


Applied to

- Music
- DNA
- Web pages
- Byte code

Arc Diagram applied to Music

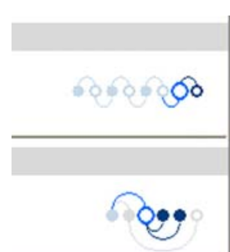
Arc Diagram applied to Music "für Elise"



- More details
  Martin Wattenberg. Arc Diagrams: Visualizing Structure in Strings
  IBM Watson Research Center, Technical report 2002-11
  http://domino.research.ibm.com/cambridge/research.nsf/0/e2a83c4986332d4785256ca7006cb621?OpenDocument

Thread Arcs

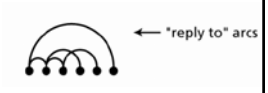
- Thread Arcs combine the chronology of messages with the branching tree structure of a conversational thread
- Benefits
  - Chronology.
  - Relationships
  - Stability:
  - Compactness:
  - Attribute Highlighting:
  - Scale:
  - Interpretation/Sense

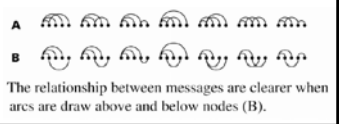


- http://www.research.ibm.com/remail/threadarcs.html

Thread Arcs for Emails

- **Visualization**
  - linear layout of message nodes connected by relationship arcs.
  - each circular node represents a message in the thread.
  - *chronology* of the thread is encoded by the position
  - The width of a Thread Arc is a linear function of the size of the thread
  - *compact visualization* if height is constrain





The relationship between messages are clearer when arcs are draw above and below nodes (B).

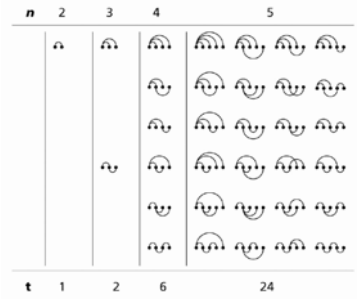
Pseudo code for drawing a thread arc

To make a Thread Arc

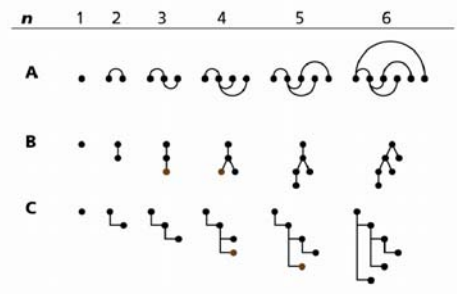
```
sort all messages chronologically
find the generation depth of each message

for each message
    if the message is the root message then
        place the node at the starting position
        don't draw an arc
    else
        place the message to the right of the last message
        if the message generation depth is odd then
            draw an arc above the line to the message's parent
        else
            draw an arc below the line to the message's parent
next message
```


Possible Thread Arcs that can be built with 2 to 5 messages.



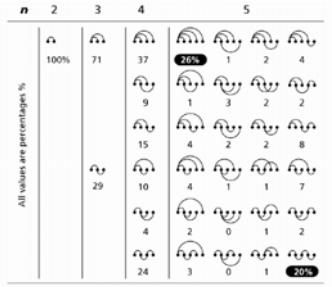
Stability of Thread Arcs



Chronological Information in the Thread Arcs

Example Email Client using Thread Arcs

Distribution of distinctive Thread Arcs of 2 to 5 messages



More details: http://www.research.ibm.com/remail/publications.html

References

- A Review and Taxonomy of Distortion-Oriented Presentation Techniques, Leung & Apperley, 1994
- Barlow et al. "A Comparison of 2-D Visualizations of Hierarchies" INFOVIS'01 http://www.sims.berkeley.edu/courses/is247/s02/readings/barlow.pdf
- Martin Wattenberg. Arc Diagrams: Visualizing Structure in Strings IBM Watson Research Center, Technical report 2002-11 http://domino.research.ibm.com/cambridge/research.nsf/0/e2a83c4986332d4785256ca7006cb621?OpenDocument
- Thread Arcs http://www.research.ibm.com/remail/threadarcs.html
- Focus+Context Taken Literally, Robert Kosara, Silvia Miksch, Helwig Hauser, 2000
- Marti Hearst, http://bailando.sims.berkeley.edu/talks/chi03-tutorial.ppt
- Storey, http://www.cs.uvic.ca/~mstorey/teaching/infovis/course_notes/introduction.pdf
- Shneiderman, http://www.cs.ubc.ca/~tmm/courses/cpsc533c-03-spr/readings/shneiderman96eyes.pdf

Techniques

- Focus & Context
- Zoom & Pan

Background

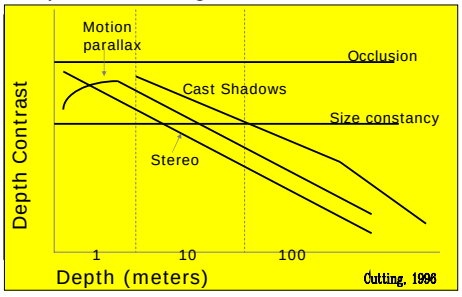
- Useful Field of View (UFOV)
  - expands searchlight metaphor
  - size of region from which we can rapidly take information
  - maintains constant number of targets
- Tunnel Vision and Stress
  - UFOV narrows as cognitive load/stress goes up
- Role of Motion in Attracting Attention
  - UFOV larger for movement detection

Depth Perception Theory

- Perceived Depth = Weighted sum of all Depth Cues
- Rank the cues in importance
  - e.g.
    - Occlusion
    - Motion Parallax
    - Stereo
    - Size constancy
    - Etc.

Depth Perception Theory

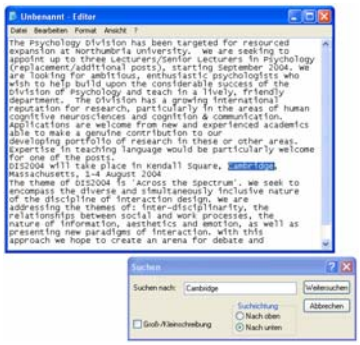
- Importance changes with distance



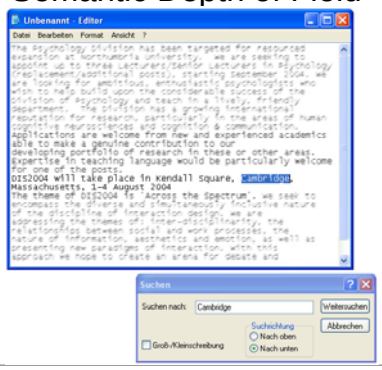
Depth of Field

- Guiding user attention by blurring less relevant parts of an image
- Keeping the context
- Semantic Depth of field = blurring objects based on their relevance

Semantic Depth of Field - Example



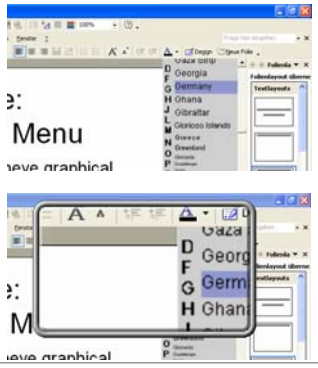
Semantic Depth of Field - Example



Techniques

- Focus & Context
- Zoom & Pan

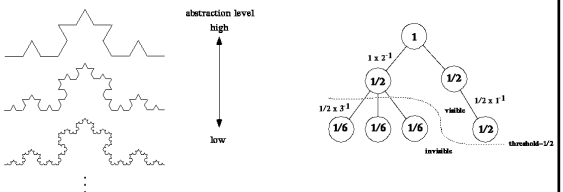
Magnifying Glass



- Magnifying glass hides context!
- This is not focus+context

Fractal Views

- Information structures regarded as complex objects
- Abstraction of objects
- Controlling amount of information that is displayed



- http://www.vogue.is.uec.ac.jp/~koike/papers/tois95/tois95.html
- **Hideki Koike. Fractal Views: A Fractal-Based Method for Controlling Information Display** *ACM Transaction on Information Systems*, Vol. 13, No. 3, July, pp.305-323, ACM, 1995.

Alternate Geometry

- Euclidean geometry – we use it since primary school…
  - 3 angles of a triangle add up to?
  - Shortest distance between two points?
- Spherical geometry
  - Geographical view of the world
    - What is the shortest way from Moscow to San Francisco?
    - Sum of angles of a triangle between Paris, NY, and Cape Town?
  - http://math.rice.edu/~pcmi/sphere/
- Hyperbolic Geometry / Space
  - Theory of Relativity
  - The "fifth" dimension
  - Can be projected into 2-D as a *pseudosphere*
  - Key: As a point moves away from the center towards the boundary circle, its distance approaches *infinity*
  - http://cs.unm.edu/~joel/NonEuclid/ (Applet)

Focus + Context

- Basic Idea:
  - Show selected regions of interest in greater detail (*focus*)
  - Preserve global view at reduced detail (*context*)
  - NO occlusion - All information is visible simultaneously
- Techniques
  - Fisheye views
  - Fisheye lens
  - Continuously variable zoom
  - Nonlinear magnification
  - Hyperbolic views
  - Distortion viewing
  - Rubber sheet views

Focus + Context

- Often combined with distortion
  - E.g. fisheye
  - Data not in focus is suppressed and distorted
  - Data of interest is larger and clearer
- "Allows dynamic interactive positioning of the local detail without severely compromising spatial relationships."
  - *Leung & Apperley*
- "One challenge in navigating through any large dataspace is maintaining a sense of relationship between what you are looking at and where it is with respect to the rest of the data."
  - *Bederson & Hollan*

Distorted vs. Non-distorted

- Non-distorted
  - Display only a selection at a time
  - Scrolling
  - Paging access
  - hierarchical structure
  - Structure-specific presentation
- Distorted
  - See the following slides

Leung & Apperley: Distortion

Unified theory of distortion techniques

- "…stretchable rubber sheet mounted on a rigid frame"
- Stretching = Magnification
- Stretching one part must equal shrinkage in other areas

Taxonomy for presentations and distortions

Leung & Apperley

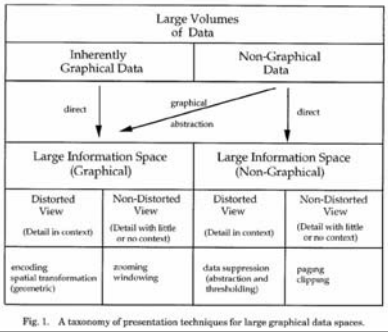


Fig. 1. A taxonomy of presentation techniques for large graphical data spaces.

Distortions

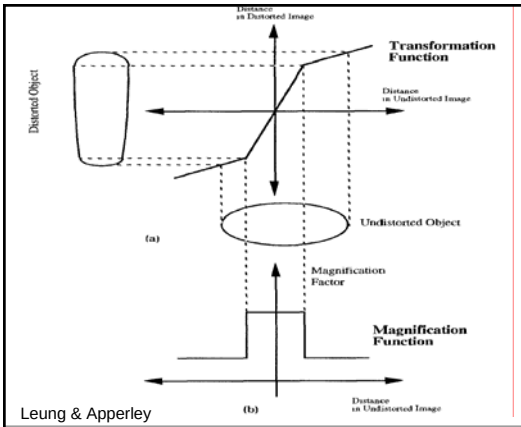
- Method
  - Post-Process: Modify results after primary graphical mapping
  - In-Process: Distortion during the primary graphical mapping
- Types
  - Focus+context – change display size relative to focus
  - Highlighting – change display type relative to focus

Distortion-based Techniques

- Bifocal Display
- Polyfocal Display
- Perspective Wall
- Fisheye View
- Graphical Fisheye View

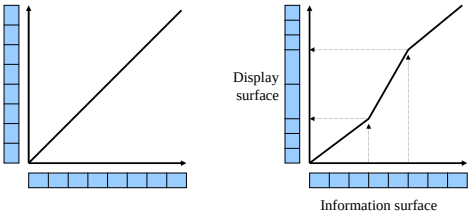
Distortion

- A distorted view is created by applying a transformation function to an undistorted image.
- A magnification function, provides a profile of the magnification factors for the entire area of image.



Leung & Apperley

Visual Transfer Functions

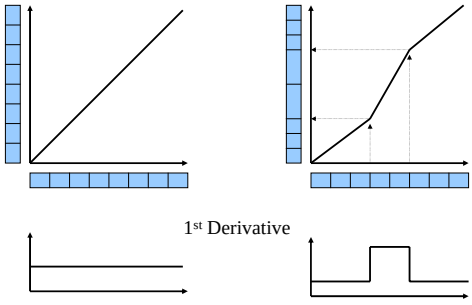


Identity function = normal flat overview

Bifocal

From http://people.cs.vt.edu/~north/infoviz/lecture11.ppt

Magnification Functions



From http://people.cs.vt.edu/~north/infoviz/lecture11.ppt

References

- Leung, Y., and Apperley, M. *A Review and Taxonomy of Distortion-Oriented Presentation Techniques ACM Transactions on Computer-Human Interaction*, 1994 1, 2, 126-160.
- B.-B. Bederson and J. D. Hollan. Pad++: *A zooming graphical interface for exploring alternate interface physics*. In Proc. ACM UIST'94, pages 17--26. ACM, ACM Press, 1994.