Lecture 3

Information Visualization

- Origins
- Data Types, Display Variables and Ranking of Visual Properties
- Mappings + Timings
- Key Design Principles
- InfoVis Toolbox
- Design + Interaction

Illustration of Key Design Principles

- Using Classic InfoVis tools (see Video on Lectures page)

Hierarchical Data Visualization

Focus + Context Visualization

Graded Discussion 1 – Analyze Visualization

Information Visualization - Problem Statement

- **Scientific Visualization**
  - Show abstractions, but based on physical space
- **Information Visualization**
  - Information does not have any obvious spatial mapping
- **Fundamental Problem**

  How to map non–spatial abstractions into effective visual form?
- **Goal**

  Use of computer-supported, interactive, visual representations of abstract data to Amplify Cognition

Goal of Information Visualization

- **Use human perceptual capabilities**
  to **gain insights** into **large data sets**
  that are **difficult to extract**
  using standard query languages

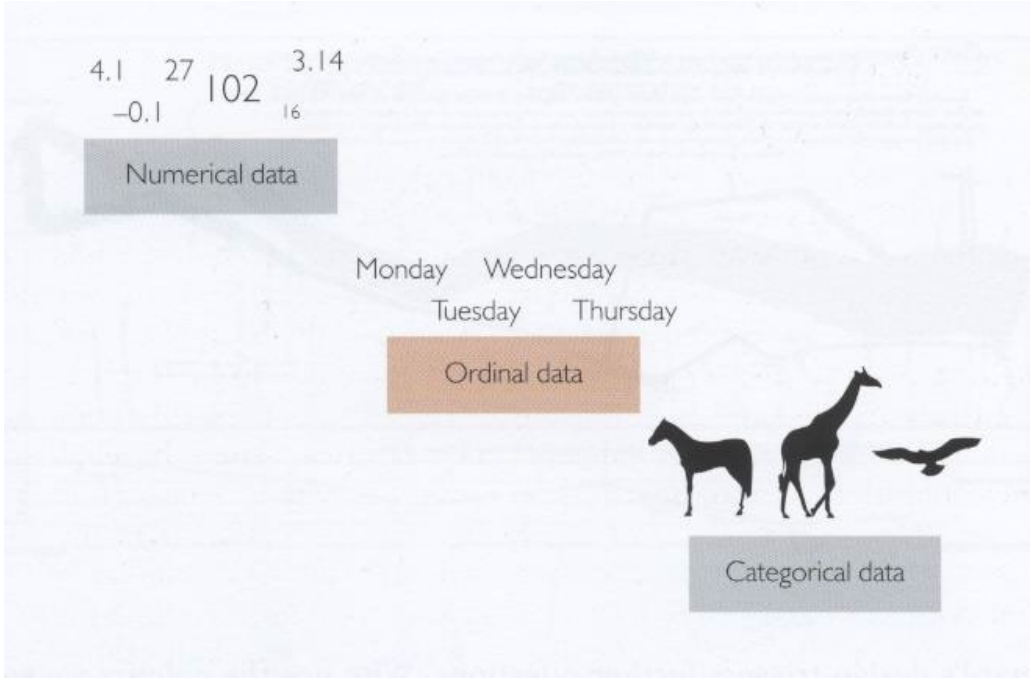
- **Exploratory Visualization**
  - Look for structure, **patterns**, trends, anomalies, relationships
  - Provide a **qualitative** overview of large, complex data sets
  - Assist in **identifying region(s) of interest** and appropriate parameters for more focussed quantitative analysis

- **Shneiderman's Mantra:**
  - **Overview first, zoom and filter, then details-on-demand**
  - Overview first, zoom and filter, then details-on-demand
  - Overview first, zoom and filter, then details-on-demand

Data Types, Data Sets and Marks

Date Types

- **Quantitative** (can perform arithmetics)
- **Ordinal** (obeys ordering relations)
- **Nominal** (equal or not equal to other values)



Abstract Data Sets

- Symbolic
- Tabular
- Networked
- Hierarchical
- Textual information

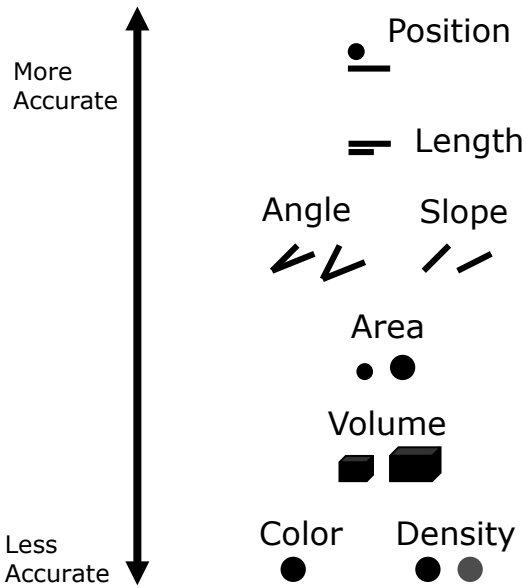
Marks

- **Points** (position, color, size)
- **Lines** (location, length, width, color)
- **Areas** (uniform / smoothed shading)
- **Volumes** (resolution, translucency)

Mapping Data to Display Variables

- Position (2)
- Orientation (1)
- Size (spatial frequency)
- Motion (2)++
- Blinking?
- Color (3)

Accuracy Ranking for Quantitative Perceptual Tasks



Ranking of Visual Properties for Different Data Types

QUANTITATIVE	ORDINAL	NOMINAL
• Position	Position	Position
• Length	Density	Color Hue
• Angle	Color Saturation	Texture
• Slope	Color Hue	Connection
• Area	Texture	Containment
• Volume	Connection	Density
• Density	Containment	Color Saturation
• Color Saturation	Length	Shape
• Color Hue	Angle	Length

Information Visualization – Key Design Principles

- **Interactivity** delegate calculator democrats
- Immediate Feedback
- **Linked Displays** stacked scatterplot
- Overview ➔ Zoom+Filter ➔ Details-on-Demand
- **Dynamic Queries** parallel coordinates
- Focus + Context network data
- **Animate Transitions** bubbles **Animation** motion chart
- Increase Information Density treemap

Information Visualization – "Toolbox"

Perceptual Coding

Position	
Size	
Orientation	
Texture	
Shape	
Color	
Shading	
Depth Cues	
Surface	
Motion	
Stereo	
Proximity	
Similarity	
Continuity	
Connectedness	
Closure	
Containment	

Interaction

Direct Manipulation	
Immediate Feedback	
Linked Displays	
Animate Shift of Focus	
Dynamic Sliders	
Semantic Zoom	
Focus+Context	
Details-on-Demand	
Output ➔ Input	

Information Density

Maximize Data-Ink Ratio	
Maximize Data Density	
Minimize Lie factor	

Mapping Data to Visual Form

1. **Variables Mapped** to **"Visual Display"**
2. **Variables Mapped** to **"Controls"**

➔ "Visual Display" and "Controls" **Linked**

Interaction Responsiveness

"0.1" second

- ➔ Perception of Motion
- ➔ Perception of Cause & Effect

"1.0" second

- ➔ "Unprepared response"

"10" seconds

- ➔ Pace of routine cognitive task

Information Visualization – Origins

1 Thought Leaders

- **Bertin**, French cartographer, "The Semiology of Graphics (1967/1983)
- **Tufte** (1983) emphasizes maximizing the density of useful information

2 Statistical Visualization

- Tukey (1977) "Exploratory Data Analysis": rapid statistical insight into data
- Cleveland and McGilll (1988) "Dynamic Graphics for Statistics"
- Analysis of multi–dimensional, multi–variable data

3 Scientific Visualization

- Satellites sending large quantities of data ➔ how to better understand it?

4 Computer Graphics and Artificial Intelligence

- Mackinlay (1986) formalized Bertin's design theory; added psychophysical data, and used to generate automatic design of data

5 User Interface and Human Computer Interaction

- Card, Robertson & Mackinlay (1989) coined "Information Visualization" and used animation and distortion to interact with large data sets in a system called the "Information Visualizer"

Toward a InfoVis Toolbox – Problem Statement & Goal

Information Visualization

- **Information does not have any obvious spatial mapping**

Fundamental Problem

How to map non–spatial abstractions into effective visual form?

Goal

Use of computer-supported, interactive, visual representations of abstract data to **amplify cognition**

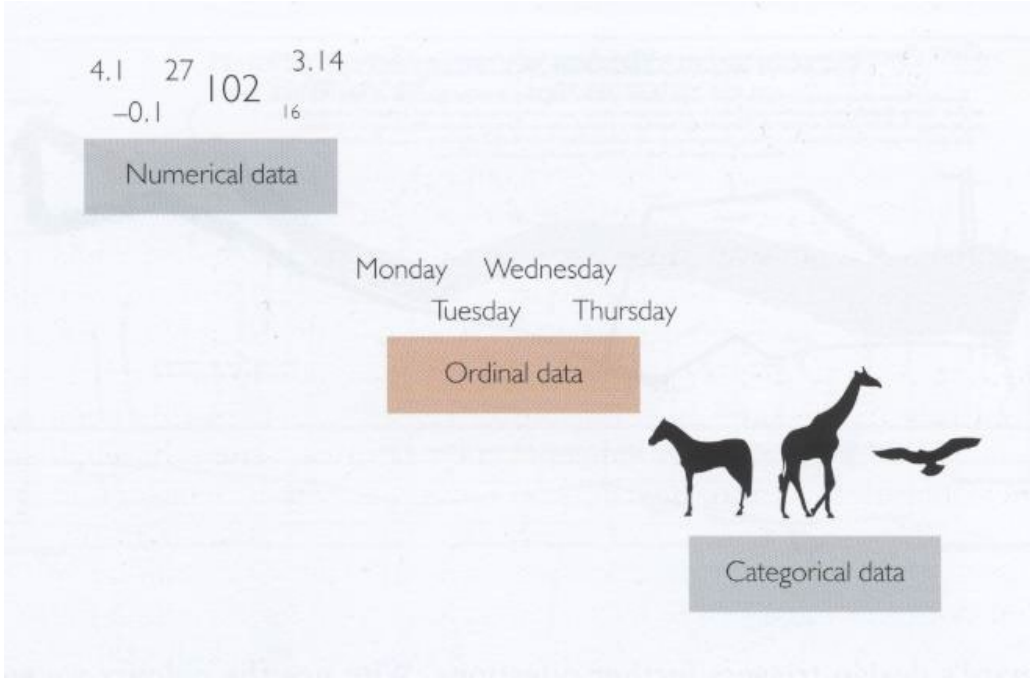
Data Types, Data Sets and Marks

Data Types

Numerical (can perform arithmetics)

Ordinal (obeys ordering relations)

Categorical (equal or not equal to other values)



Abstract Data Sets

- Symbolic
- Tabular
- Networked
- Hierarchical
- Textual information
- …

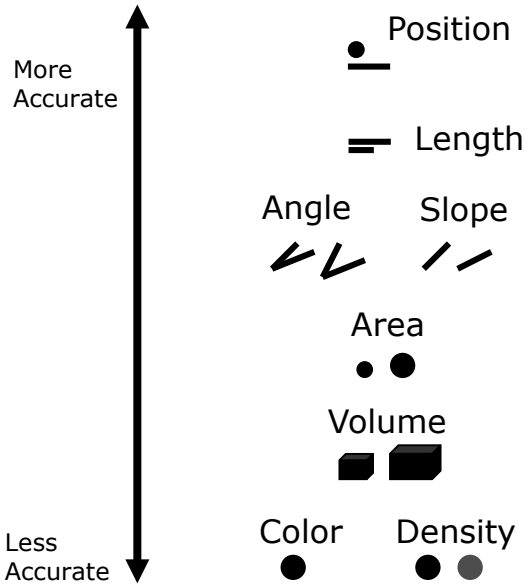
Marks

- **Points** (position, color, size)
- **Lines** (location, length, width, color)
- **Areas** (uniform / smoothed shading)
- **Volumes** (resolution, translucency)

Mapping Data to Display Variables

- Position (2)
- Orientation (1)
- Size (spatial frequency)
- Motion (2)++
- Blinking?
- Color (3)

Accuracy Ranking for Quantitative Perceptual Tasks



Ranking of Visual Properties for Different Data Types

NUMERICAL	ORDINAL	CATEGORICAL
Position	Position	Position
Length	Density	Color Hue
Angle	Color Saturation	Texture
Slope	Color Hue	Connection
Area	Texture	Containment
Volume	Connection	Density
Density	Containment	Color Saturation
Color Saturation	Length	Shape
Color Hue	Angle	Length

Mapping Data to Visual Form

1. **Variables Mapped** to **"Visual Display"**
2. **Variables Mapped** to **"Controls"**

➔ "Visual Display" and "Controls" **Linked**

Interaction Responsiveness

"0.1" second

- ➔ Perception of Motion
- ➔ Perception of Cause & Effect

"1.0" second

- ➔ "Unprepared response"

"10" seconds

- ➔ Pace of routine cognitive task

Information Visualization – Key Design Principles

- Direct Manipulation
- Immediate Feedback
- Linked Displays
- Dynamic Queries
- Tight Coupling Output ➔ Input
- Overview ➔ Zoom+Filter ➔ Details-on-Demand
- Provide Context + Focus
- Animate Transitions
- Increase Information Density

Information Visualization – "Toolbox"

Perceptual Coding

Position	
Size	
Orientation	
Texture	
Shape	
Color	
Shading	
Depth Cues	
Surface	
Motion	
Stereo	
Proximity	
Similarity	
Continuity	
Connectedness	
Closure	
Containment	

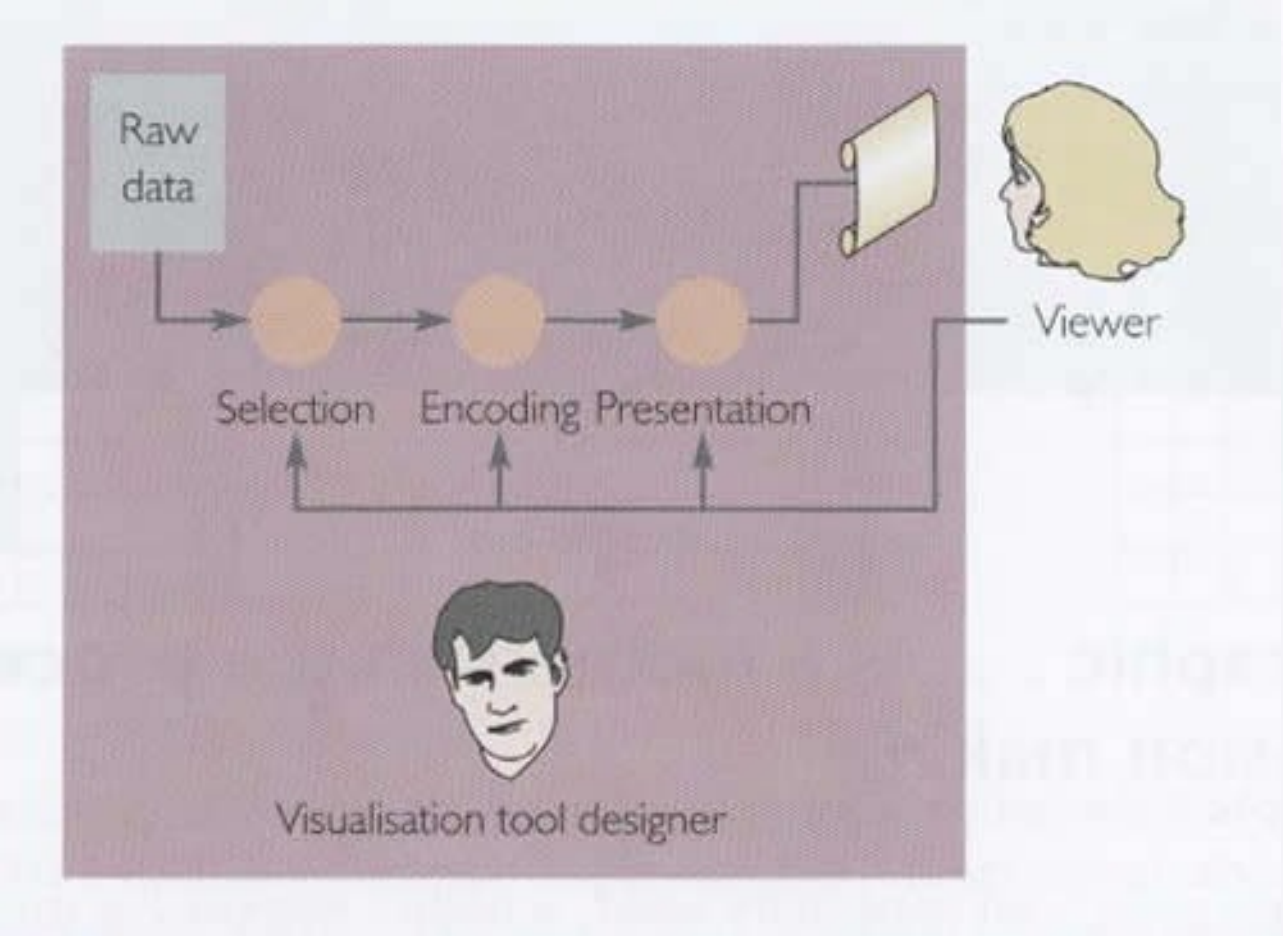
Interaction

Direct Manipulation	
Immediate Feedback	
Linked Displays	
Animate Shift of Focus	
Dynamic Sliders	
Semantic Zoom	
Focus+Context	
Details-on-Demand	
Output ➔ Input	

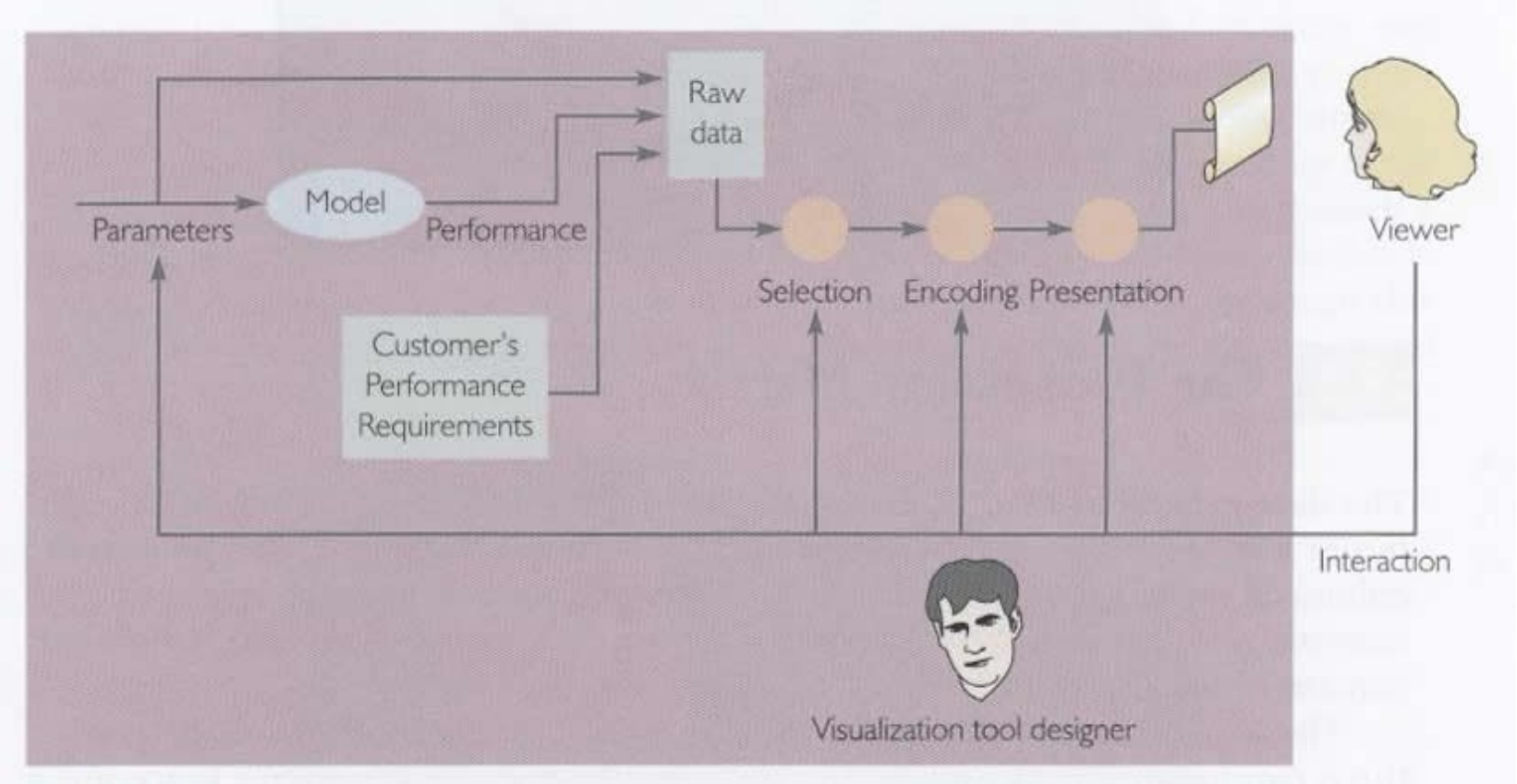
Information Density

Maximize Data-Ink Ratio	
Maximize Data Density	
Minimize Lie factor	

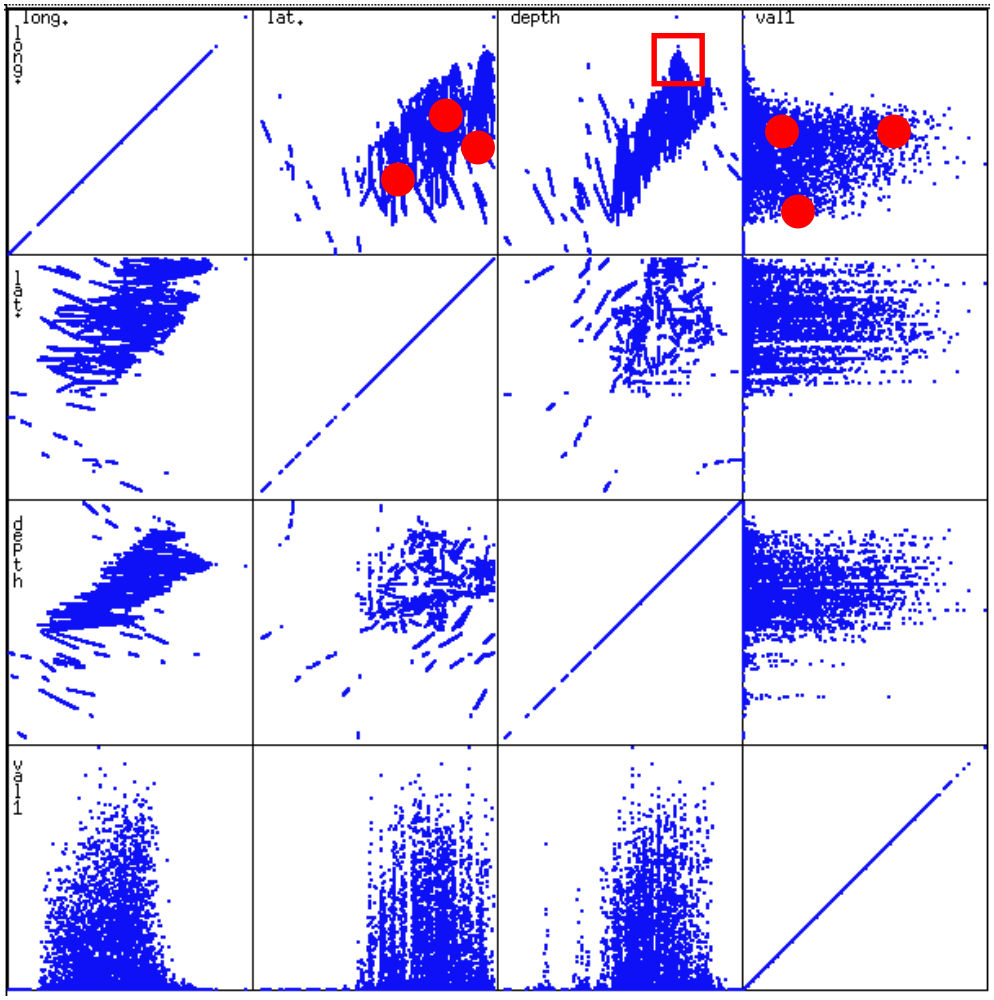
Information Visualization – Design & Interaction



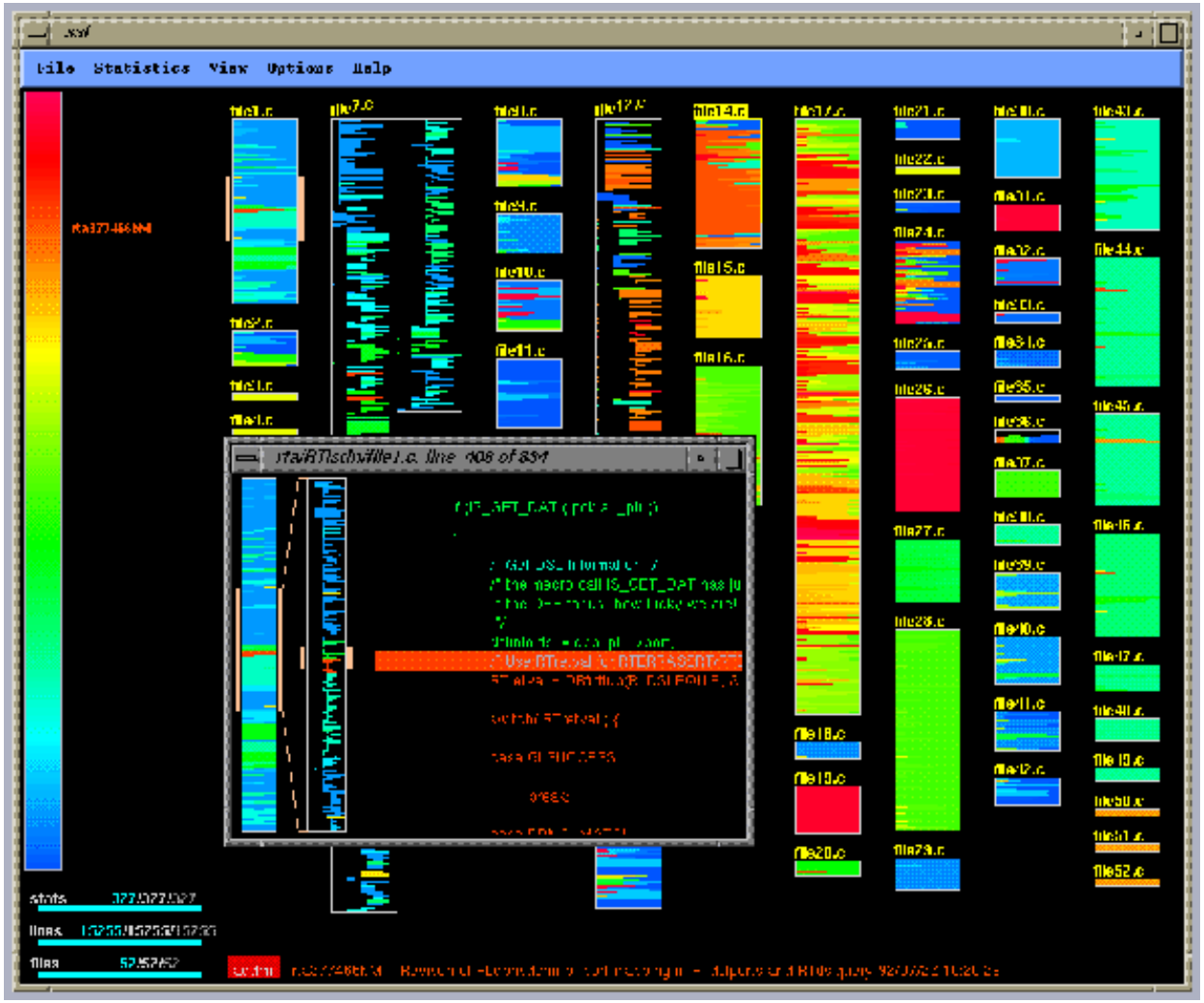
Information Visualization – Design & Interaction



Stacked Scatterplots – Brushing ➔ **Linked Displays**

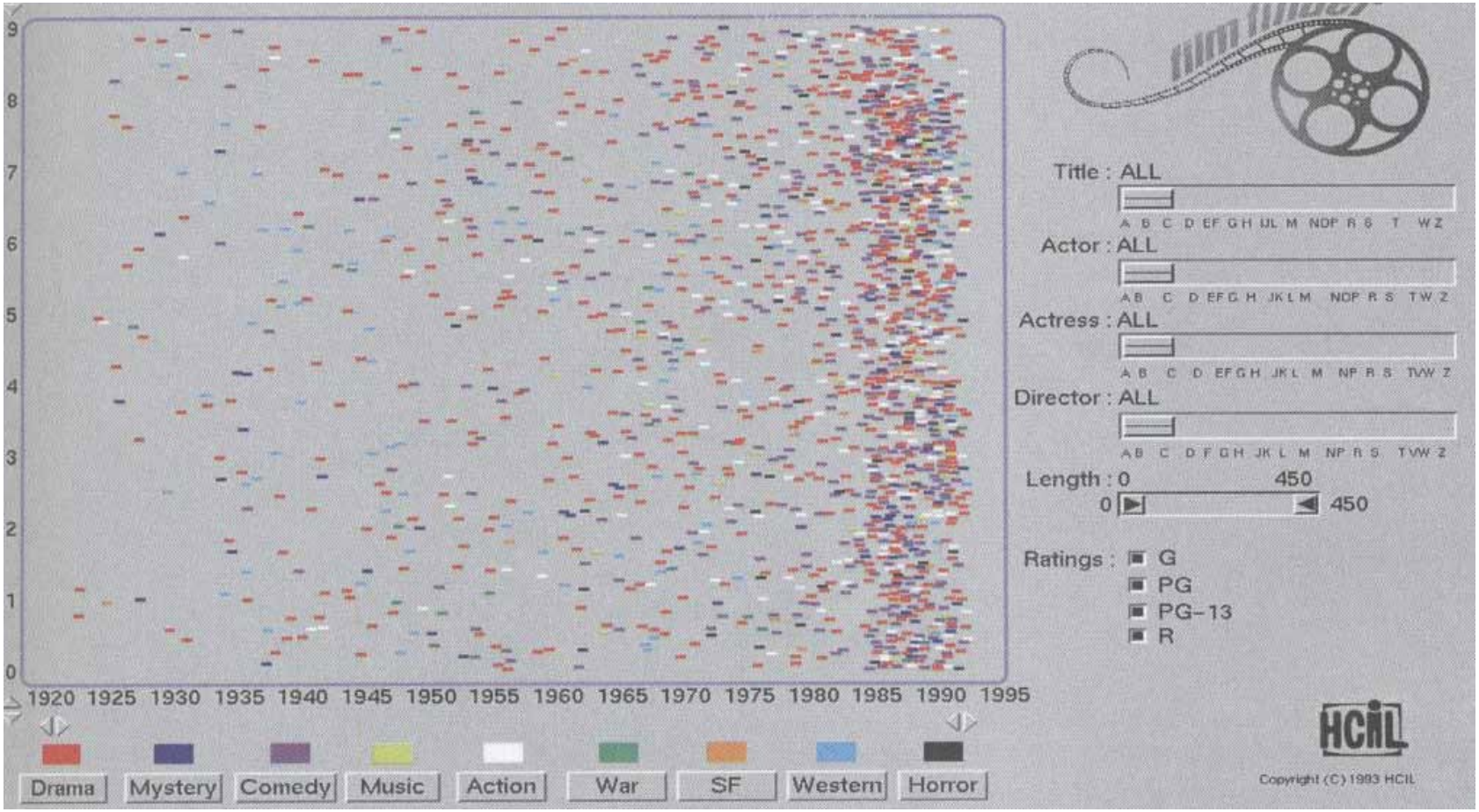


SeeSoft – Software Visualization ➔ **Linked Displays**



Line = single line of source code and its length **Color** = different properties

FilmFinder & Starfields Display ➔ **Dynamic Queries**



Two Most Important Variables Mapped to **"Scatterplot"**

Other Variables Mapped to **"Controls"**

"Visual Display" and "Controls" **Linked**

FilmFinder & Starfields Display

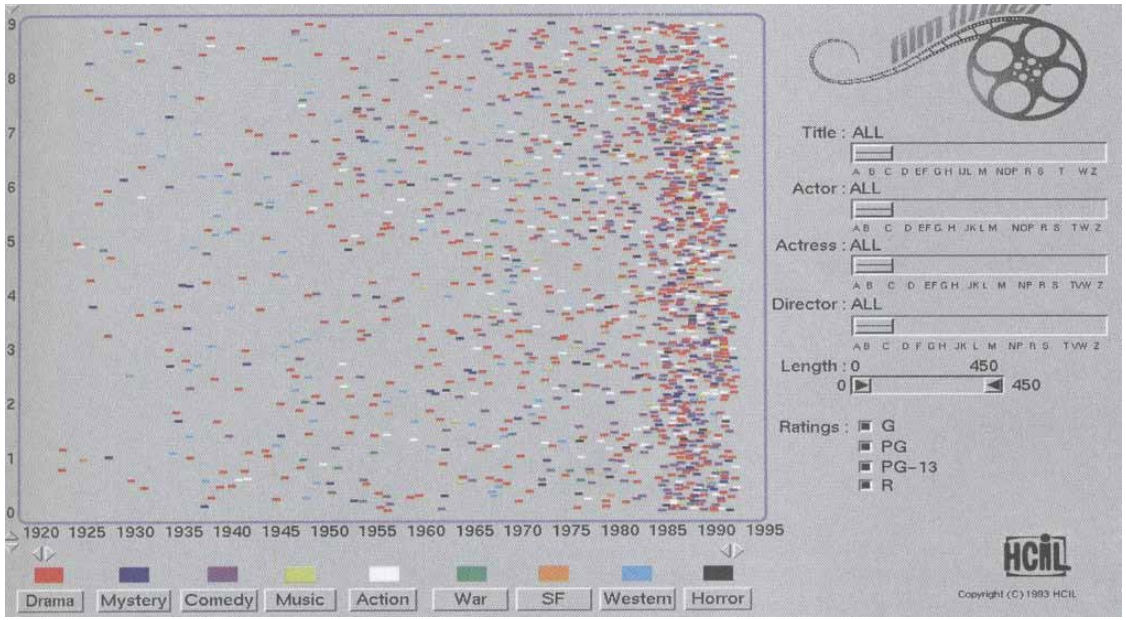
Advantages of Dynamic Queries over traditional query language such as SQL

- ➔ Make Query Formulation Easy = Interact with Sliders and Visual Objects
  (SQL = Structured Query Language is difficult to master)
- ➔ **Support Rapid, Incremental and Reversible Exploration**
- ➔ Shift Cognitive Load to Perceptual System
- ➔ Selection by Pointing

Tight Coupling of Interface Components

- ➔ Immediate Visual Feedback
- ➔ Linked Display and Controls
- ➔ **Avoid "Null set"** by having current selection limit further query refinement
- ➔ Progressive Query Refinement
- ➔ Details on Demand

Starfields



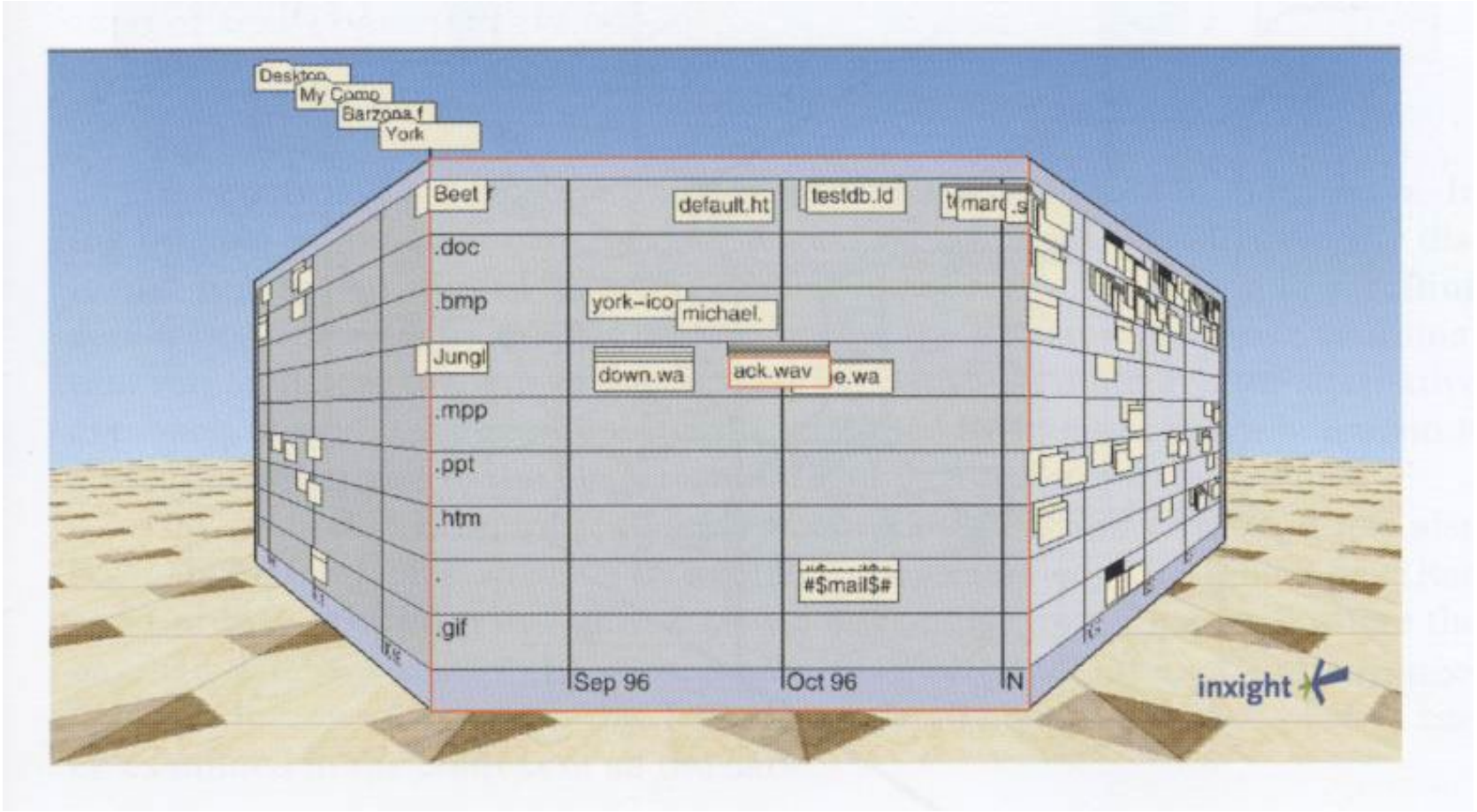
Perceptual Coding

Position	**Yes**
Size	
Orientation	
Texture	
Shape	
Color	**Yes**
Shading	
Depth Cues	
Surface	
Motion	**Yes**
Stereo	
Proximity	**Yes**
Similarity	**Yes**
Continuity	
Connectedness	
Closure	
Containment	

Interaction

Direct Manipulation	**Yes**
Immediate Feedback	**Yes**
Linked Displays	**Yes**
Logarithmic Shift of Focus	
Dynamic Sliders	**Yes**
Semantic Zoom	**Yes**
Focus+Context	
Details-on-Demand	**Yes**
Output ➔ Input	**Yes**

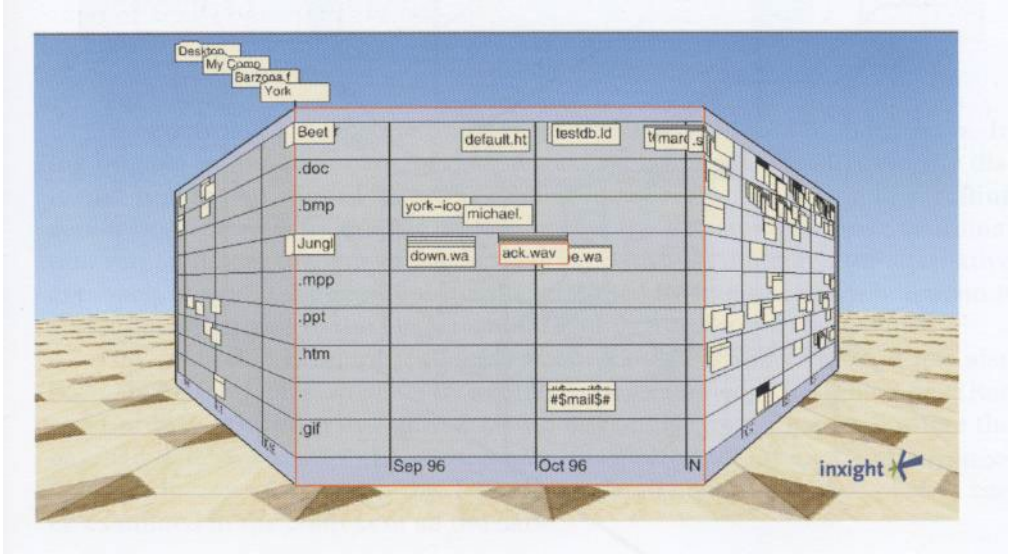
Perspective Wall ➔ Focus + Context



Fisheye Distortion to Increase Information Density

PerspectiveWall

Data = Temporal / Linear



Perceptual Coding

Position	**Yes**
Size	**Yes**
Orientation	
Texture	
Shape	**Yes**
Color	**Yes**
Shading	
Depth Cues	**Yes**
Surface	Yes
Motion	**Yes**
Stereo	
Proximity	**Yes**
Similarity	**Yes**
Continuity	
Connectedness	
Closure	
Containment	Yes

Interaction

Direct Manipulation	**Yes**
Immediate Feedback	**Yes**
Linked Displays	
Logarithmic Shift of Focus	**Yes**
Dynamic Sliders	**Yes**
Semantic Zoom	
Focus+Context	**Yes**
Details-on-Demand	
Output ➔ Input	

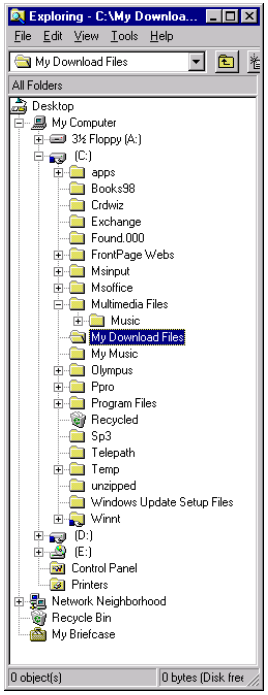
Hierarchical Information

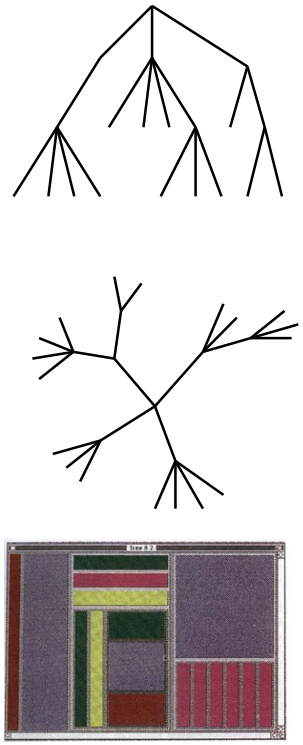
Pervasive

- File / Directory systems on computers
- Classifications / Taxonomies / Controlled Vocabularies
- Software Menu structure
- Organization charts
- …

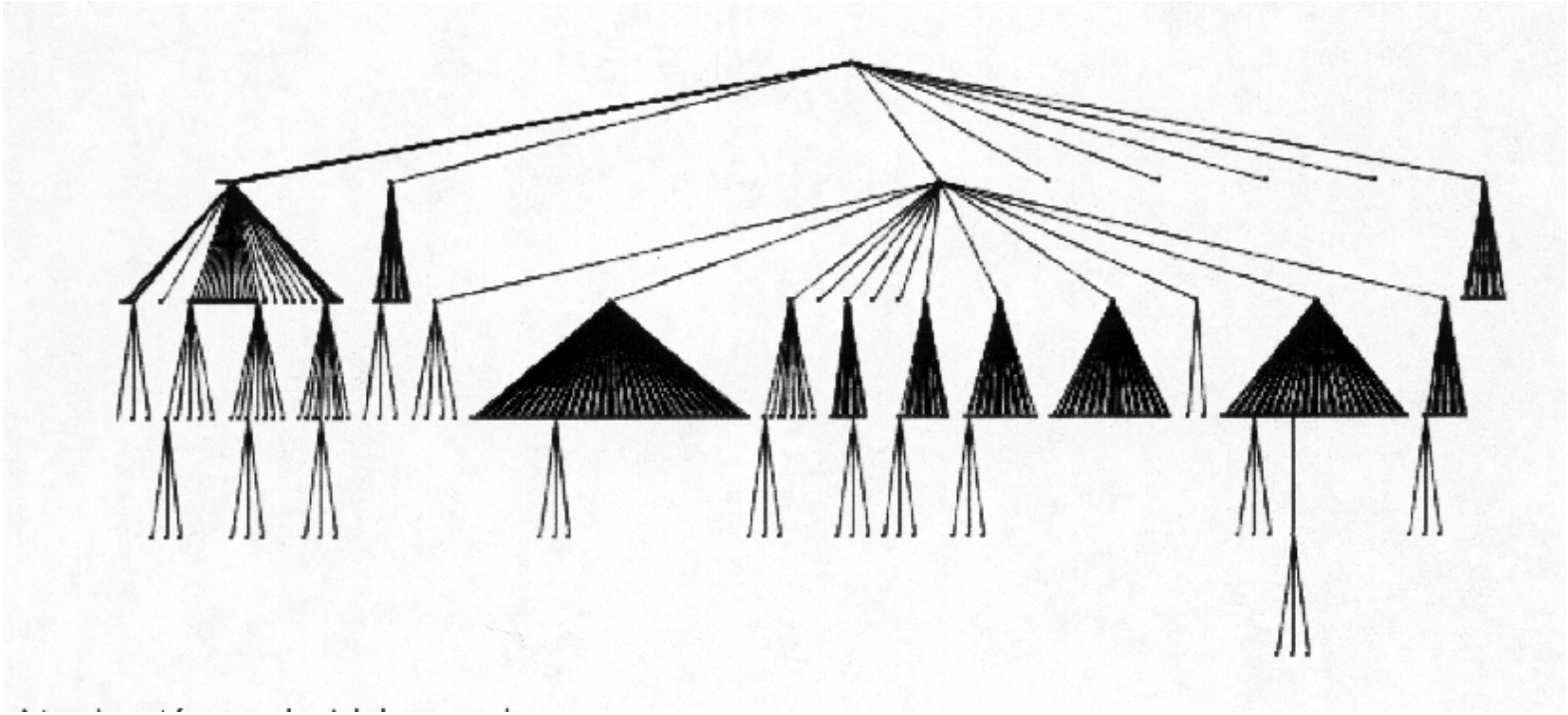
Main Visualization Schemes

- **Indented Outlines**
  - Good for Searching Bad for Structure
- **Node-Link Trees**
  - Top-to-Bottom Layout
    - 2D
    - 3D : ConeTree
  - Radial Layout
    - 2D : SunBurst, Hyperbolic Trees
    - 3D : H3 & Walrus
- **Space-Filling Treemaps**





Hierarchical Data – Traditional Node-Link Layout



Allocate Space proportional to # of Children at Different Levels

Hierarchical Data – **3D ConeTree**

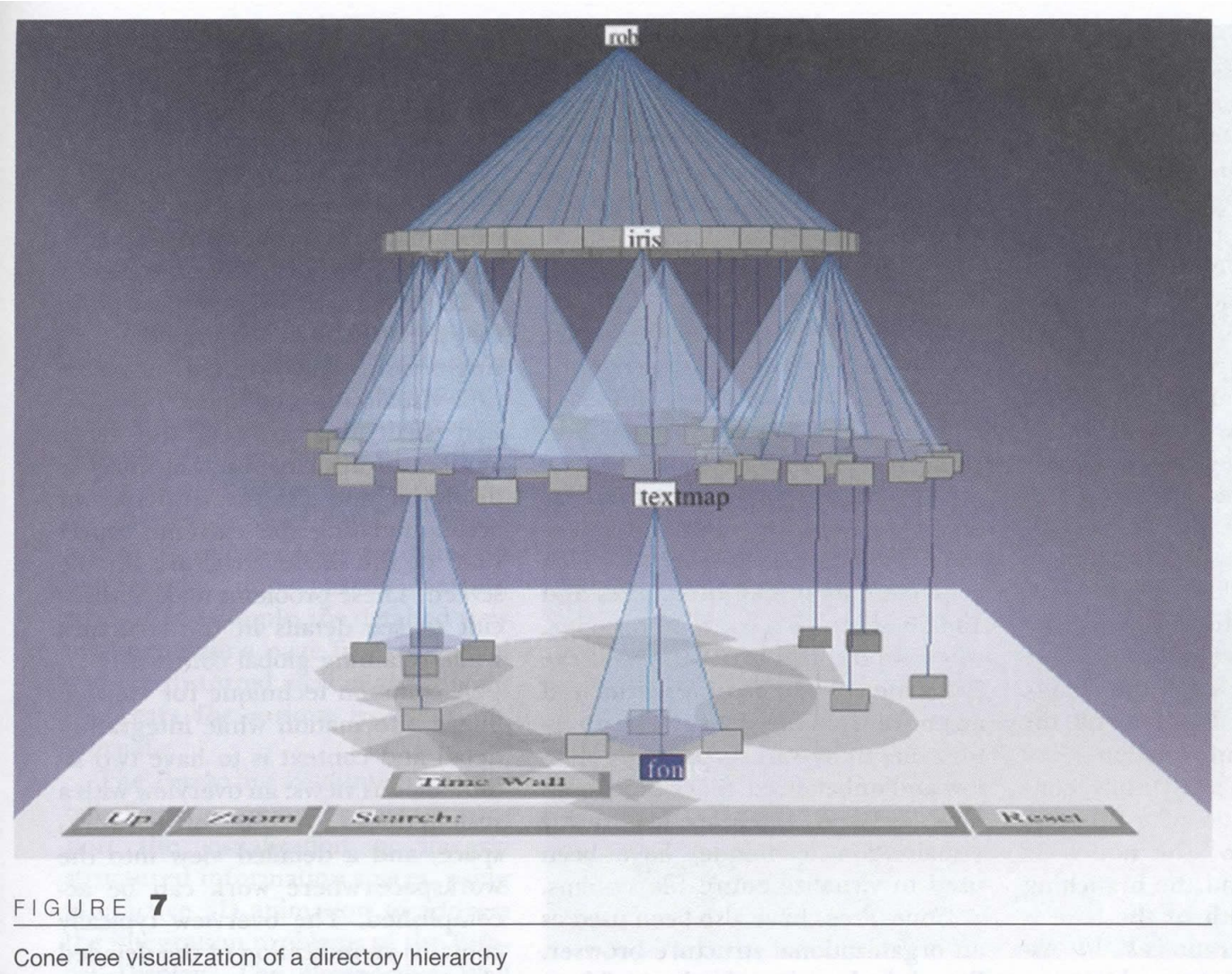


FIGURE **7**

Cone Tree visualization of a directory hierarchy

Hierarchical Data – **3D ConeTree** (cont.)

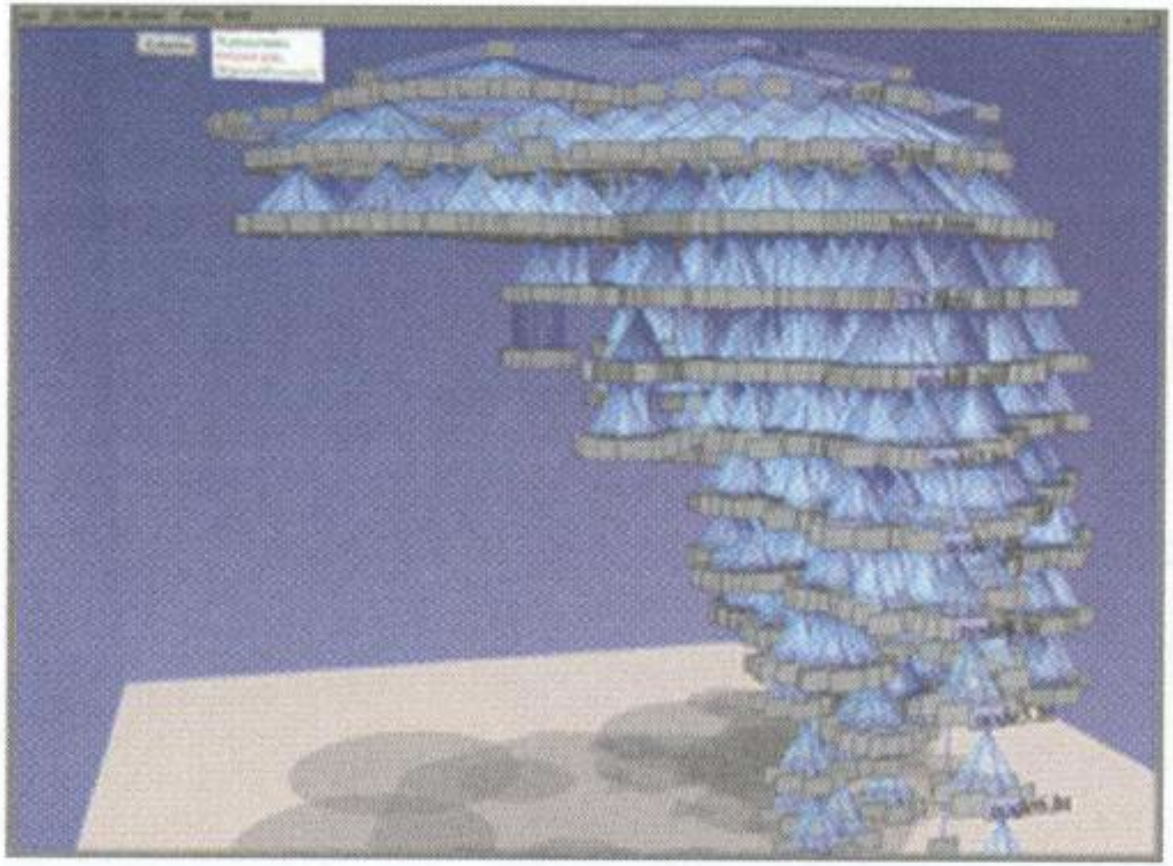
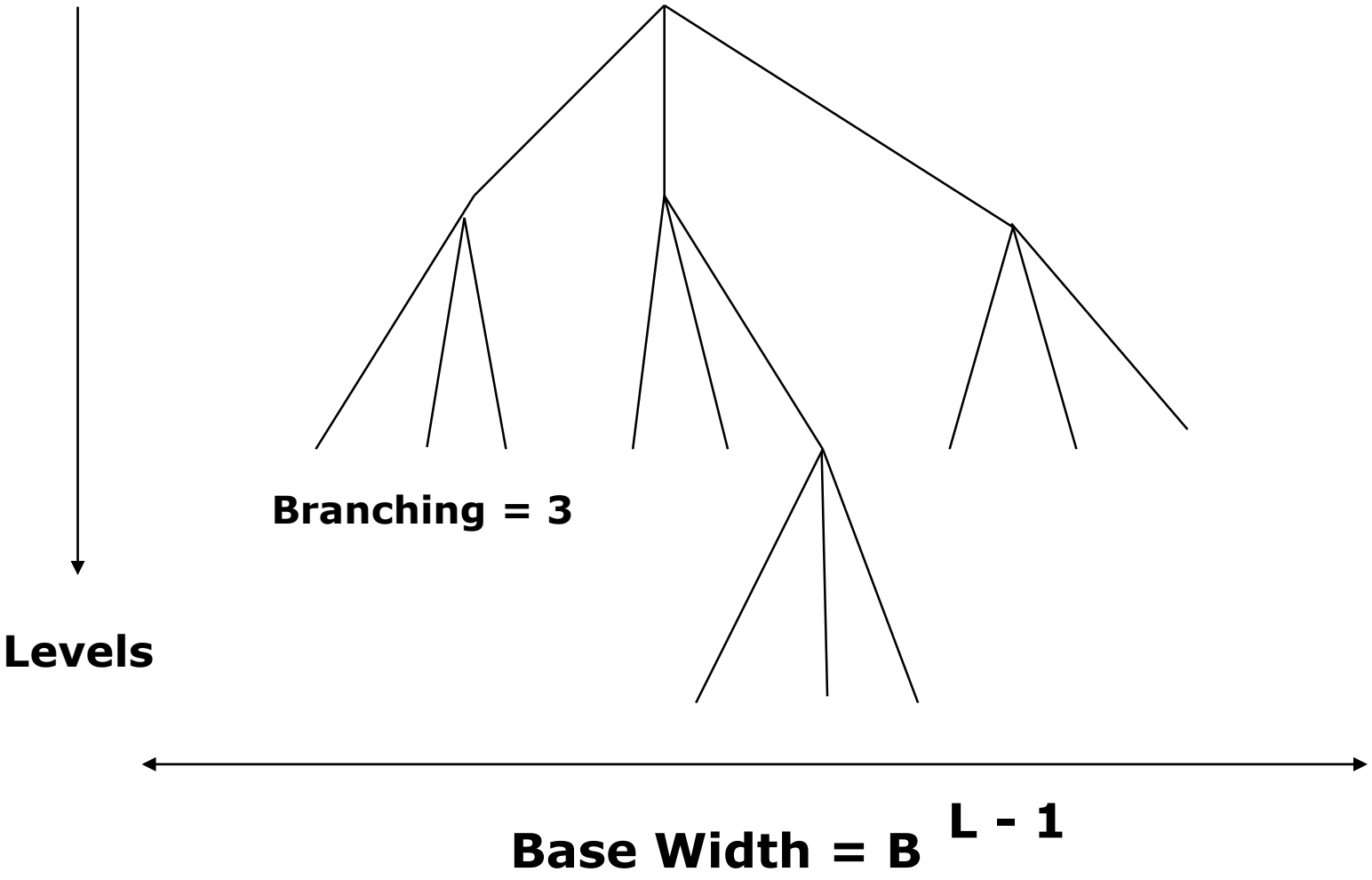


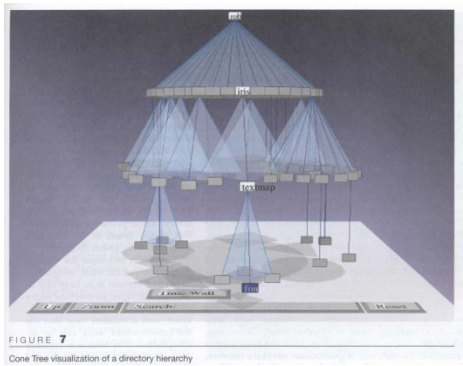
FIGURE **2.13**

Cone tree of 10,000 nodes in the Xerox PARC Web.

Hierarchy – Exponential Growth of Nodes



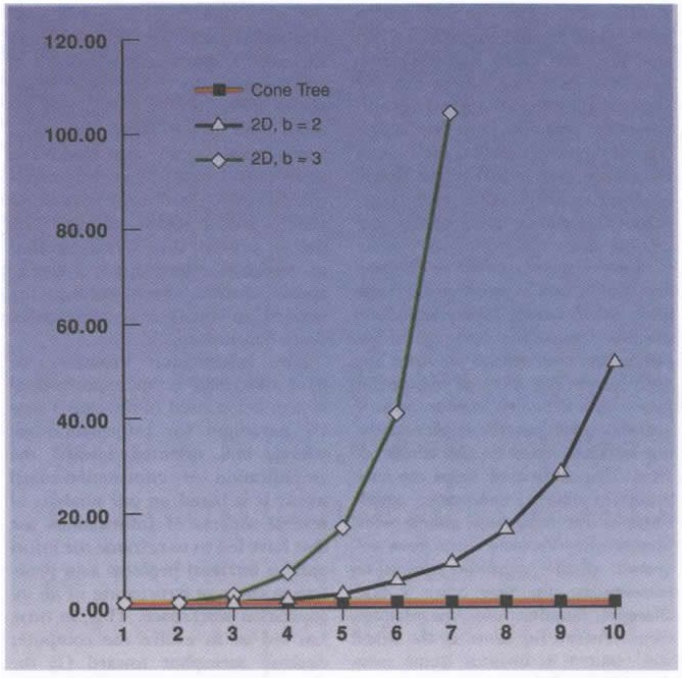
Hierarchical Data – **3D ConeTree** (cont.)

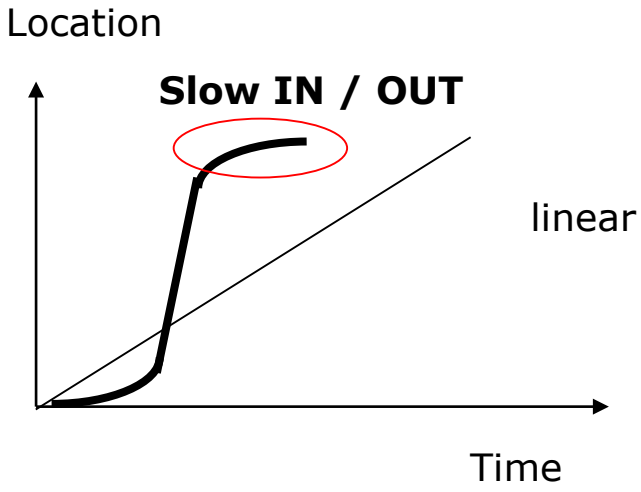


How to manage exponential growth of nodes?

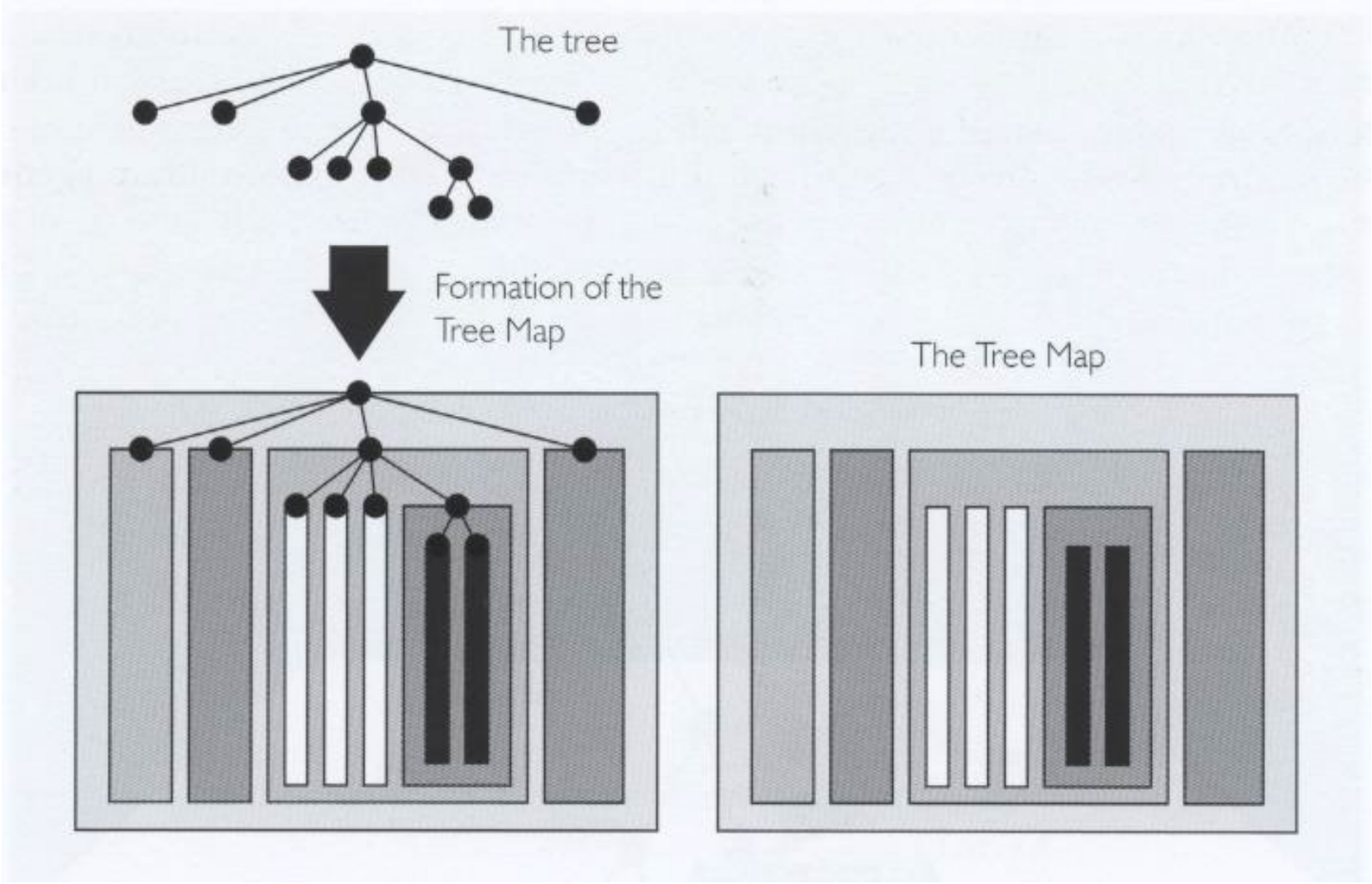
➔ Use 3D to "linearize" problem – width fixed

➔ Use "Slow IN / OUT" animation of object or point of interest to create "**Object Constancy**"

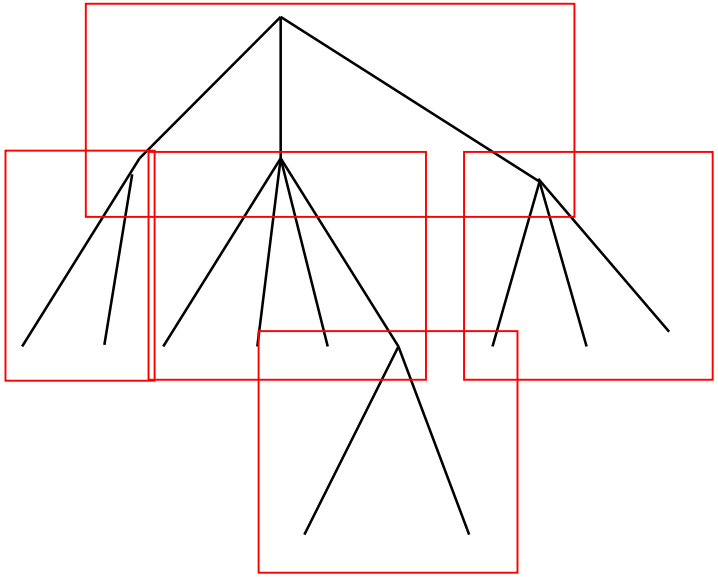


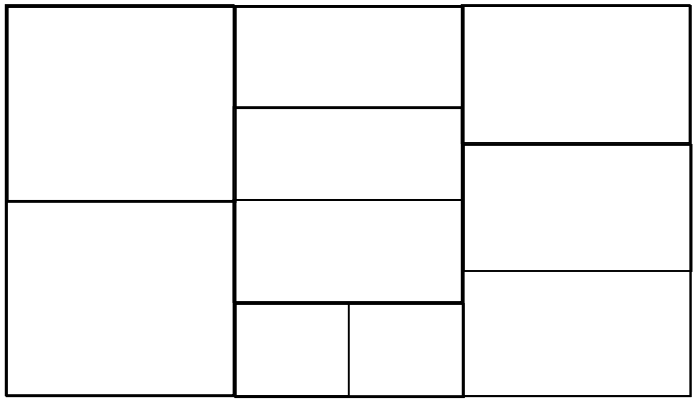


Treemaps ➔ Space-Filling Design

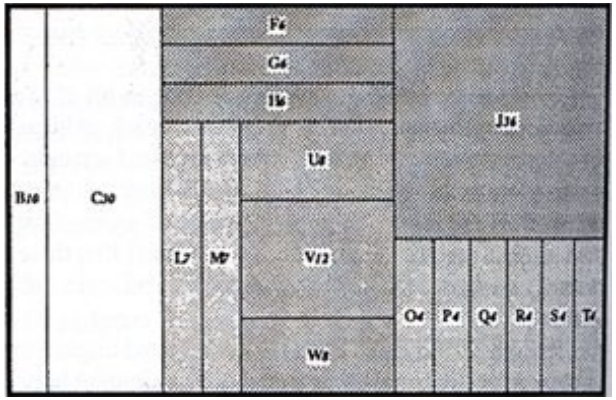


Treemaps – “Slice & Dice”

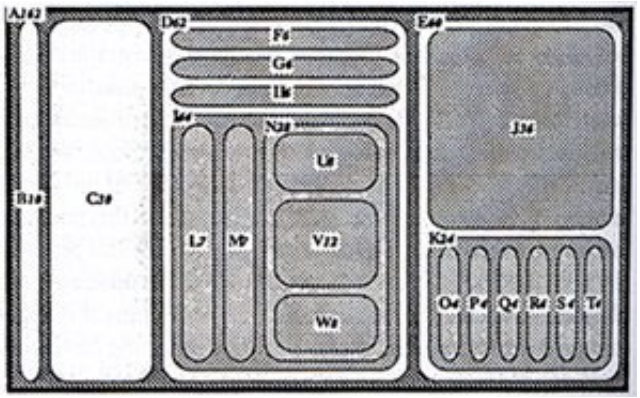




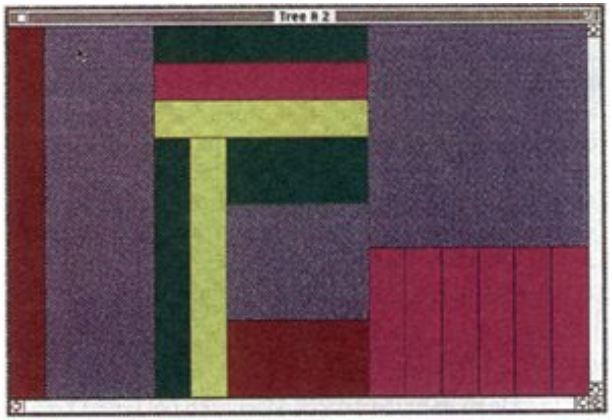
Treemaps – Nested vs. Non-nested



Non-nested Tree-Map



Nested Tree-Map



Treemaps

Which Problem do Treemaps aim to address?

➔ Visualize hierarchical structure as well as content of (atom) nodes

What are Treemaps' main design goals?

➔ **Space–filling** (High Data / Ink Ratio)

➔ "Structure" is represented using **Enclosure / Containment**

➔ "Content" is represented using **Area**

Pre–attentive, Early Visual Processes Used?

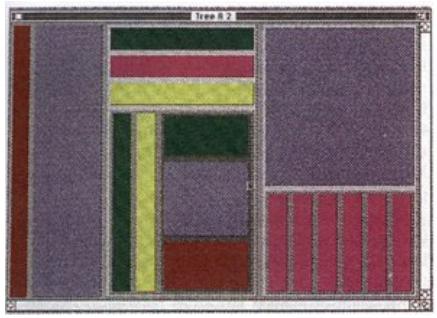
➔ Position, Size = Area, Color and Containment

Treemap

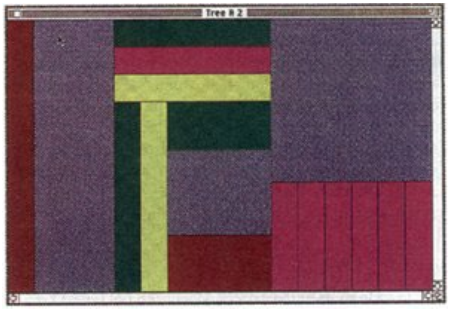
Data = Hierarchy

Perceptual Coding

Position	**Yes**
Size	**Yes**
Orientation	
Texture	Yes
Shape	
Color	**Yes**
Shading	
Depth Cues	
Surface	
Motion	**Yes**
Stereo	
Proximity	Yes
Similarity	
Continuity	
Connectedness	
Closure	
Containment	**Yes**



Nested

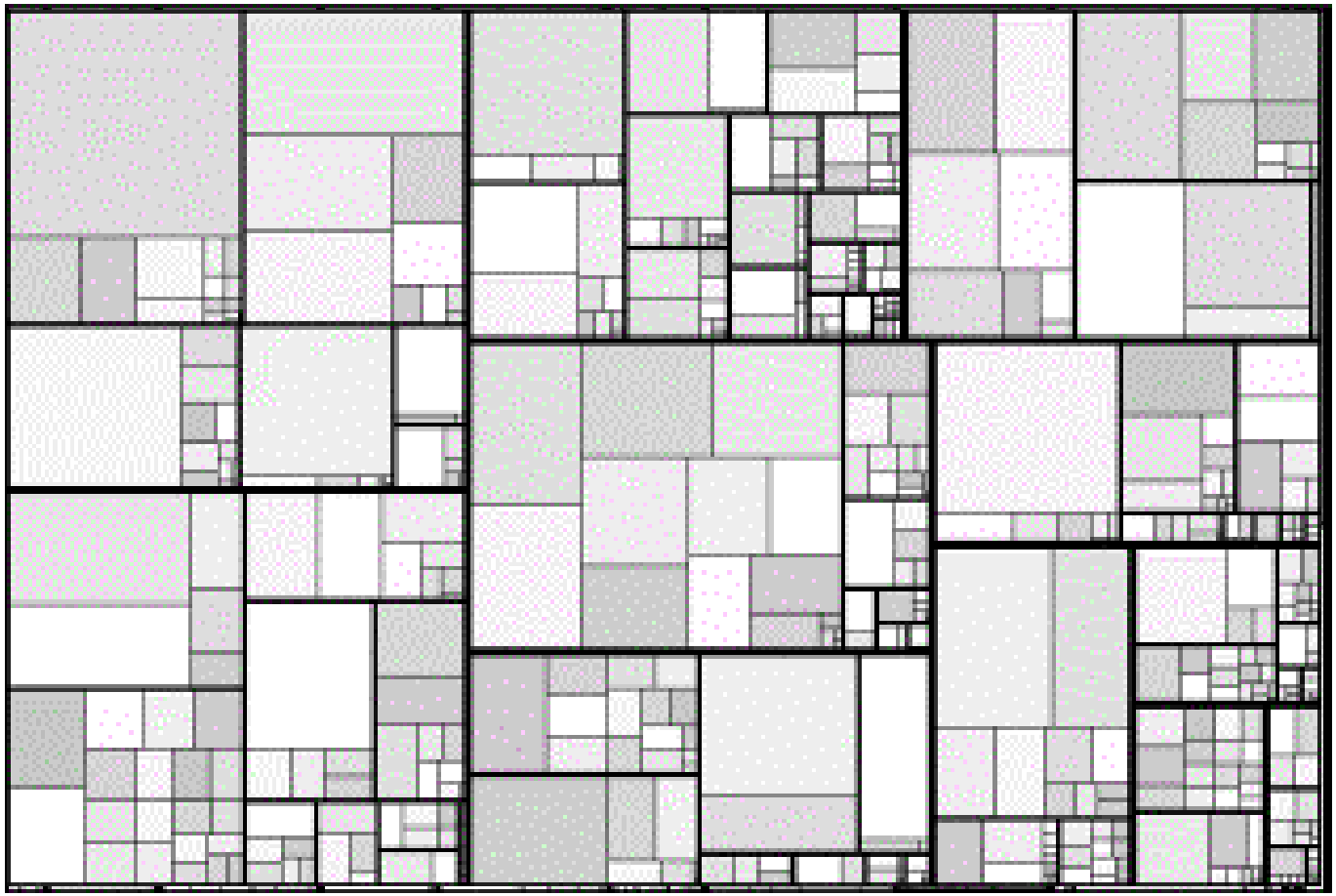


Non-nested

Interaction

Direct Manipulation	**Yes**
Immediate Feedback	**Yes**
Linked Displays	**Yes**
Logarithmic Shift of Focus	
Dynamic Sliders	**Yes**
Semantic Zoom	Yes
Focus+Context	
Details-on-Demand	**Yes**
Output ➔ Input	

Treemaps – Other Layout Algorithms ➔ Better Aspect Ratio

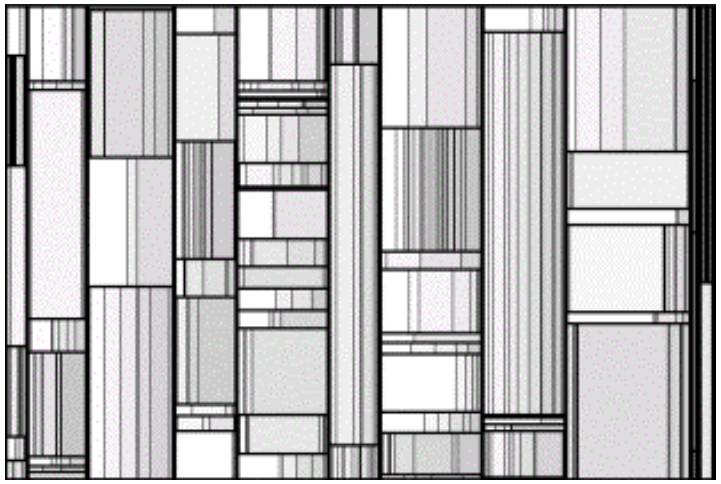


Slice-and-dice

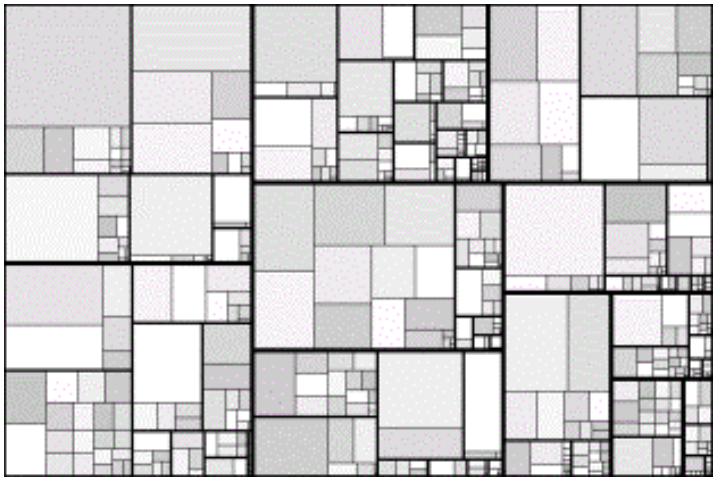
Squarified

Treemaps – Other Layout Algorithms

Hard to **Improve Aspect Ratio** and **Preserve Ordering**

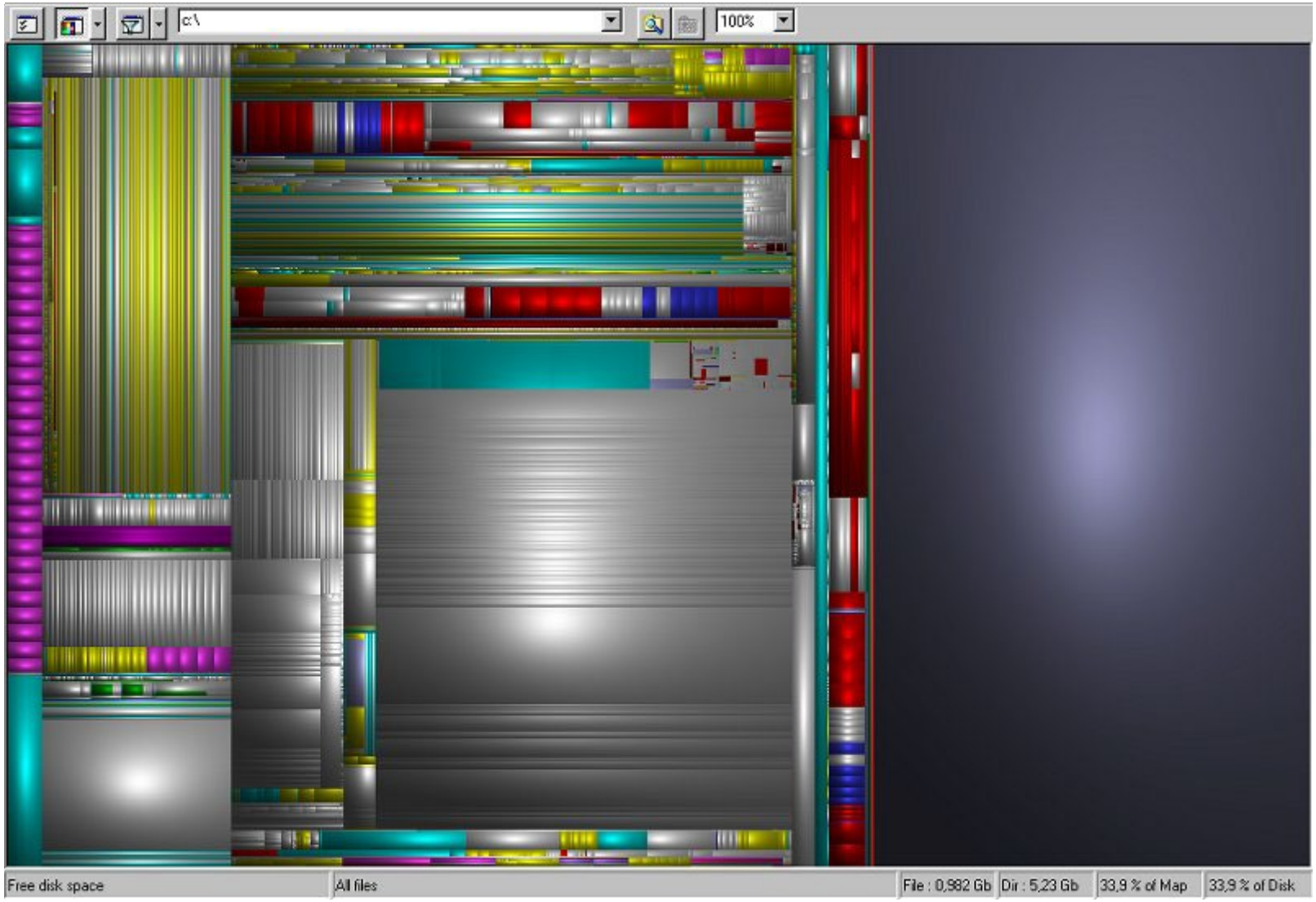


Slice-and-dice
Ordered,
very bad aspect ratios
stable

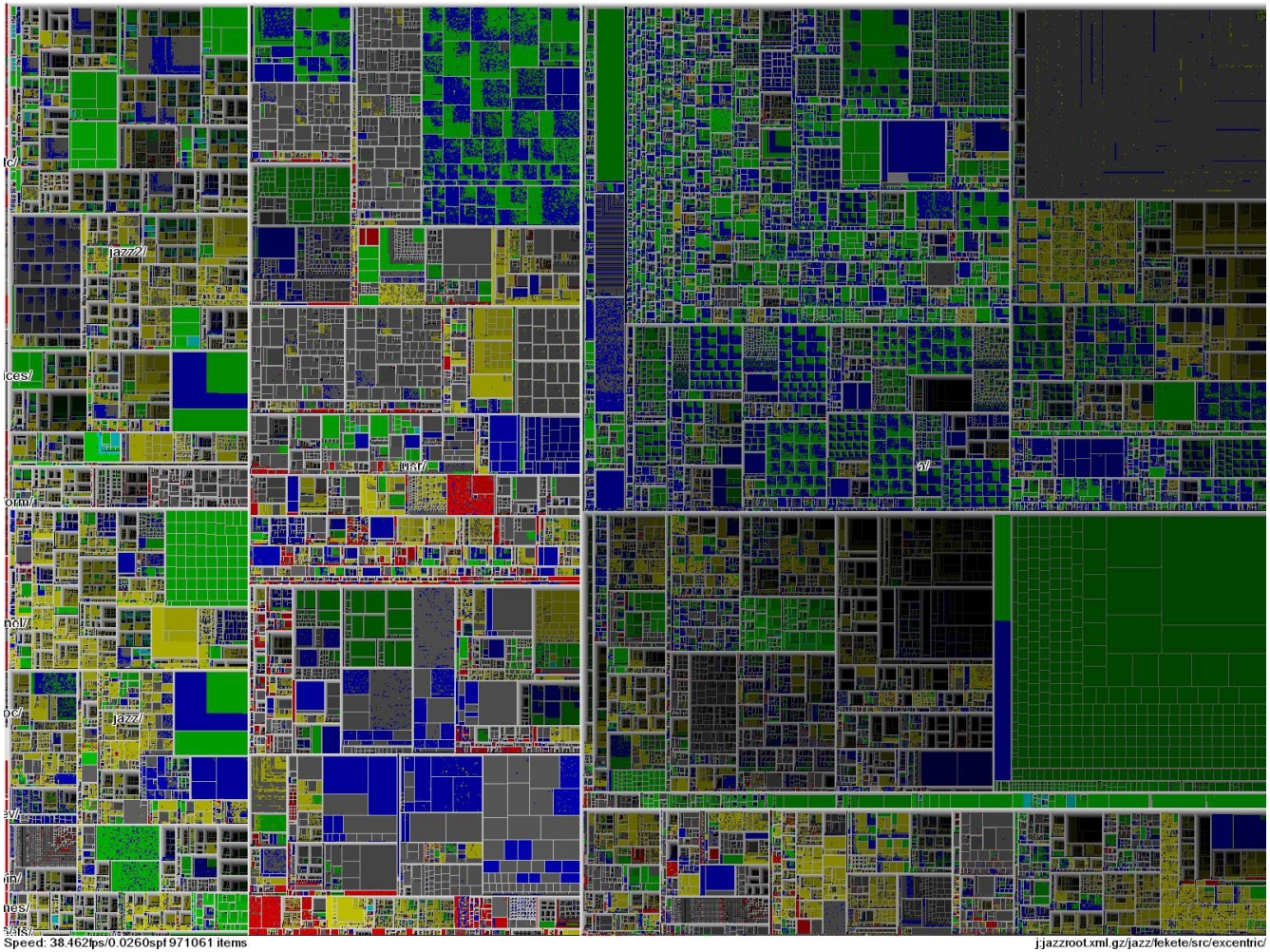


Squarified
Unordered
best aspect ratios
medium stability

Treemaps – Shading



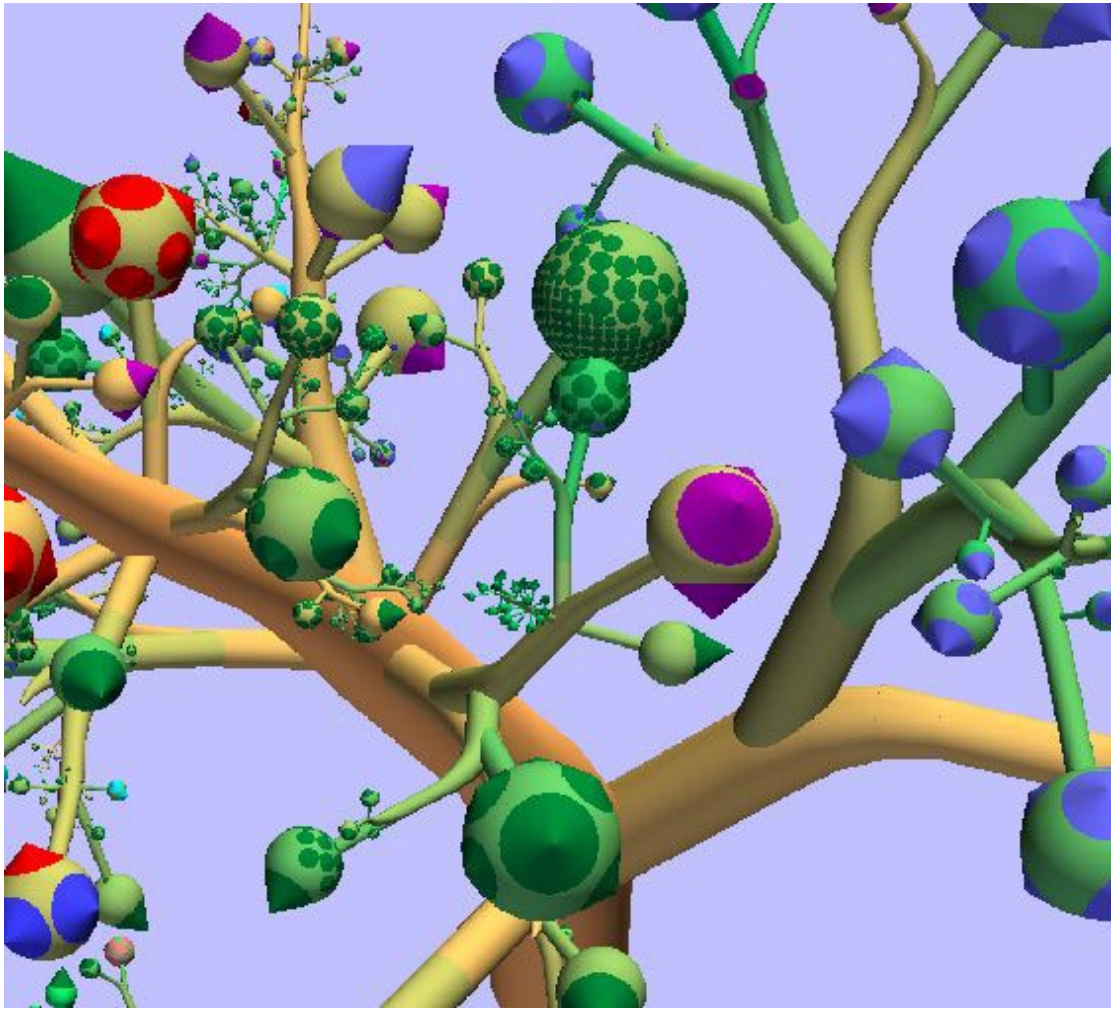
Treemaps – 1,000,000 items



http://www.cs.umd.edu/hcil/VisuMillion/

Botanical Visualization of Huge Hierarchies

Visualization Group - Technical University of Eindhoven http://www.win.tue.nl/vis/



Botanical Visualization of Huge Hierarchies

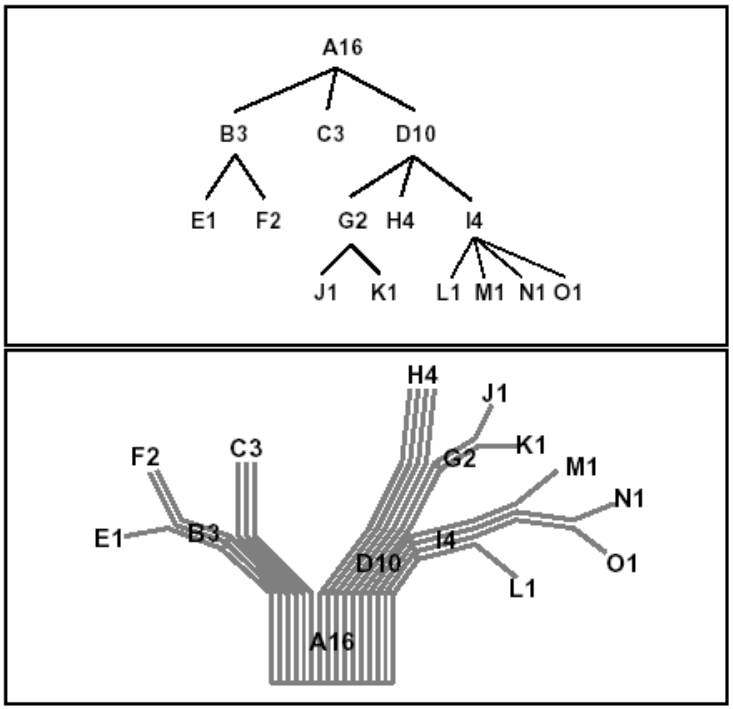


Figure 2. Node and link diagram (t) and corresponding strands model (d).

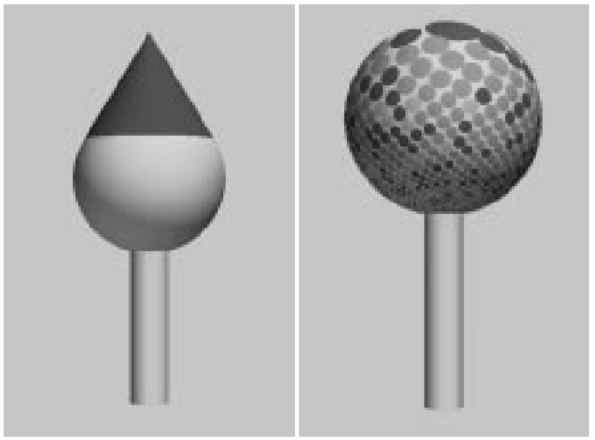


Figure 8. Phi-ball with one (l) and many (r) files.

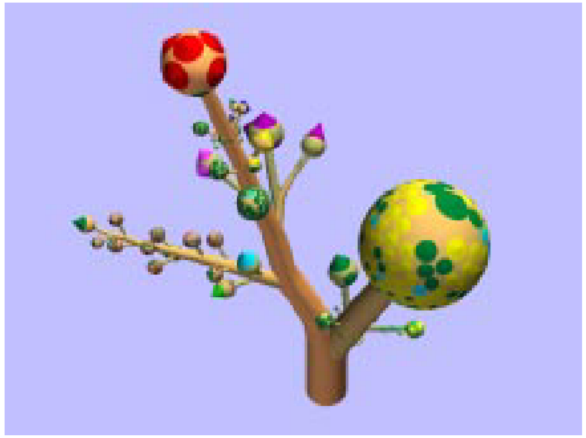


Figure 9. Final model with contraction, extrusion, and phi-balls.

Botanical Visualization of Huge Hierarchies

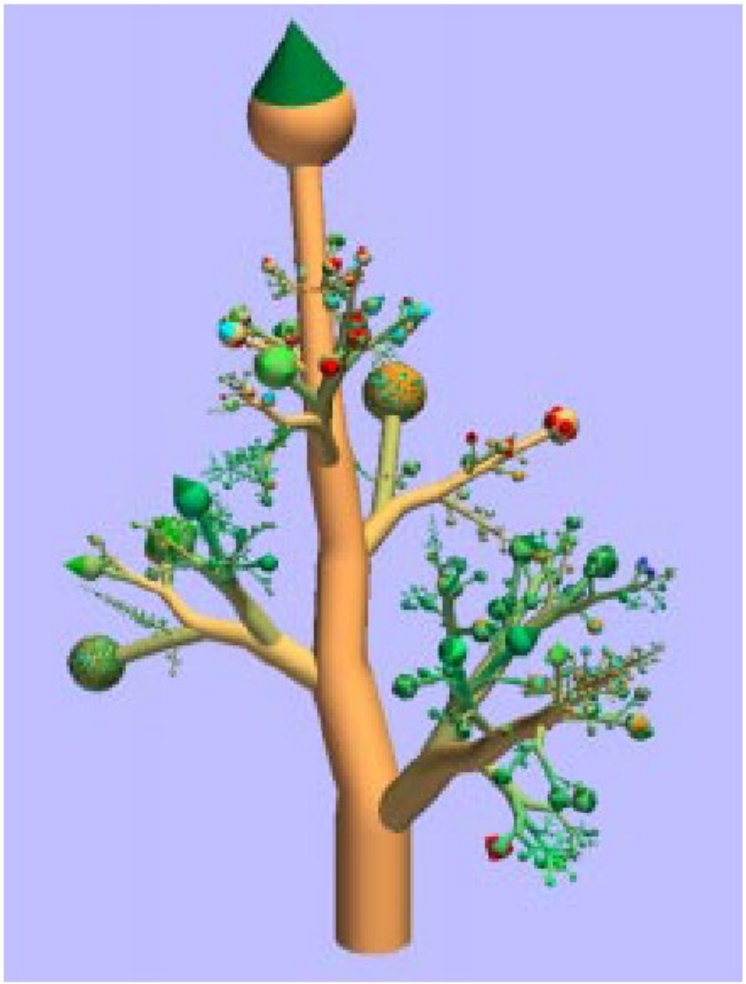


Figure 10. Complete hard disk with $\alpha = 45$ and $\beta = 360/\varphi$.

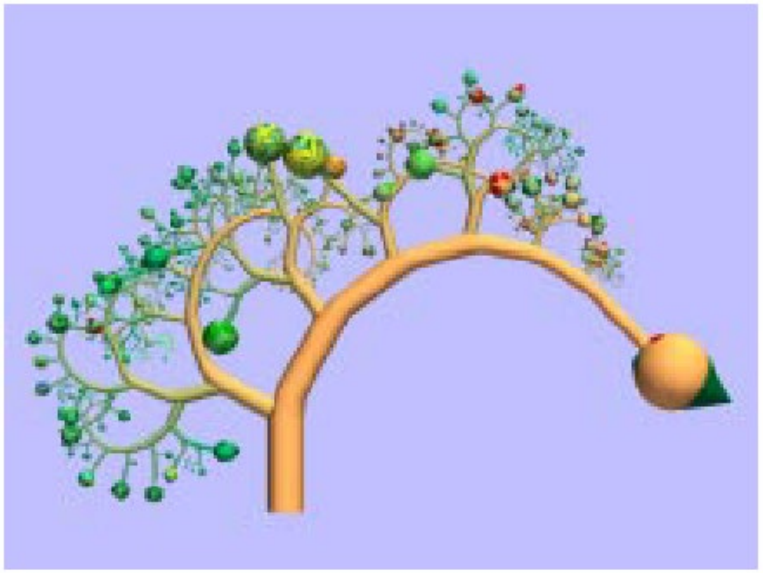
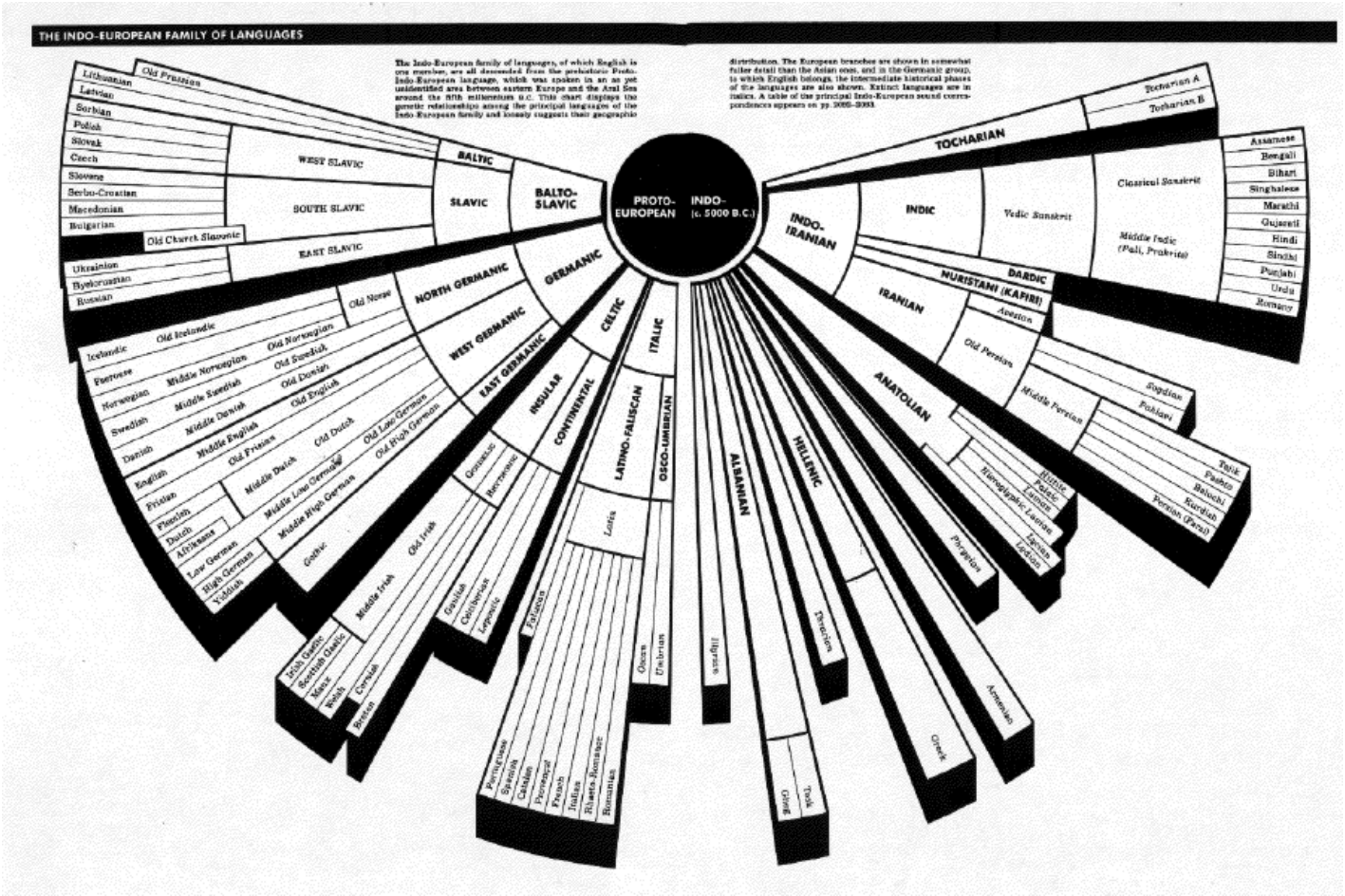


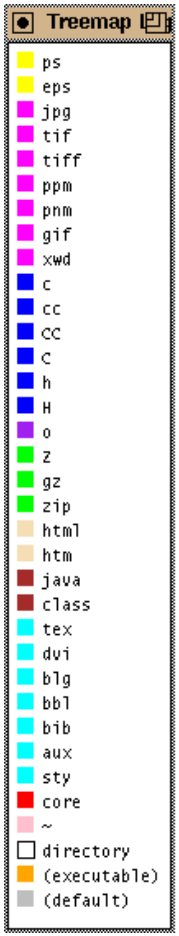
Figure 11. Complete harddisk with $\alpha = 90$ and $\beta = 0$.

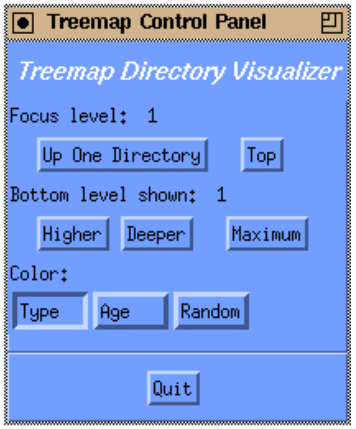
Hierarchical Data – Radial Space-Filling

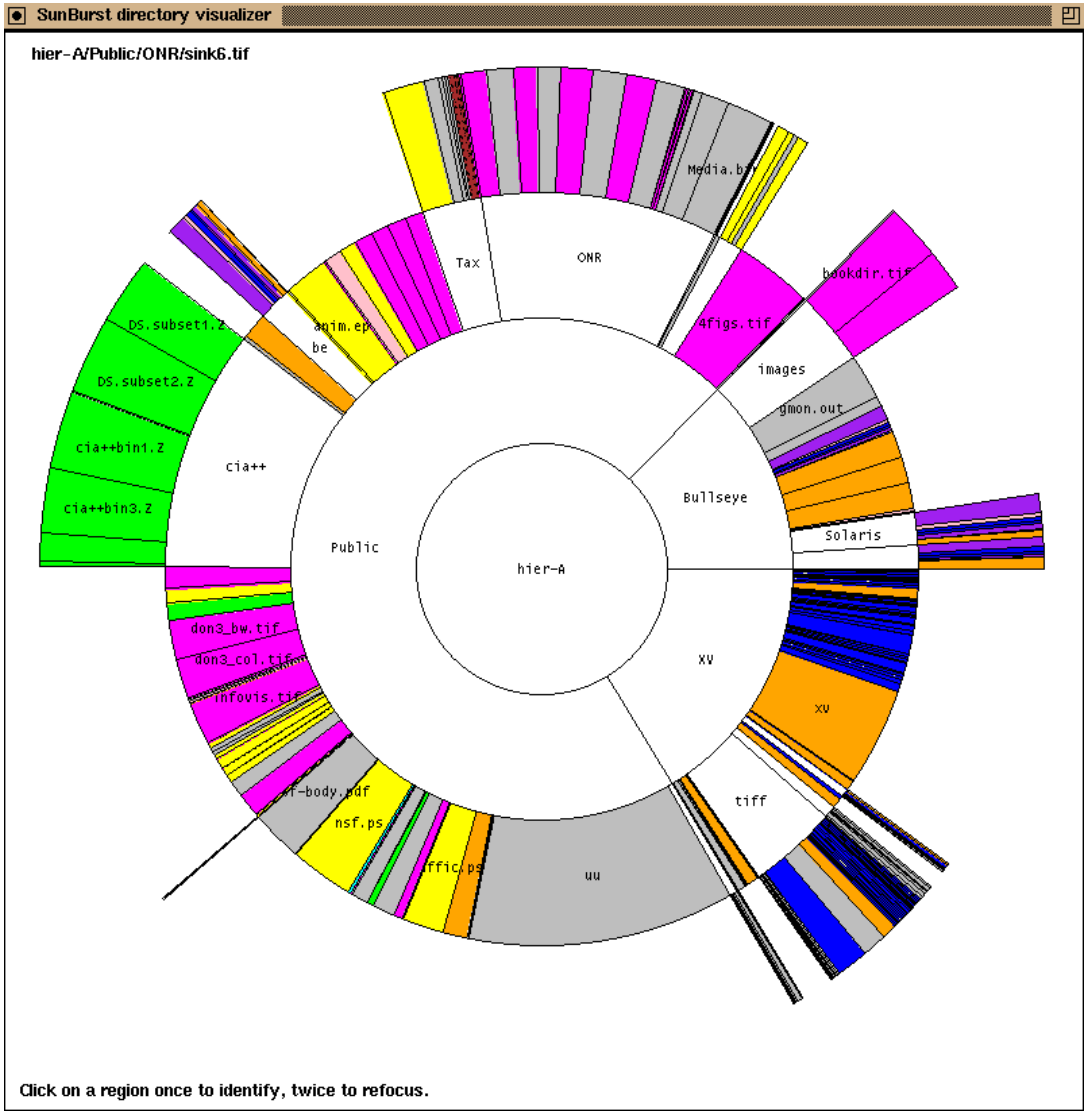


American Heritage Dictionary, 3rd Ed. Houghton Mifflin, 1992

Hierarchical Data – Radial Space-Filling ➔ SunBurst

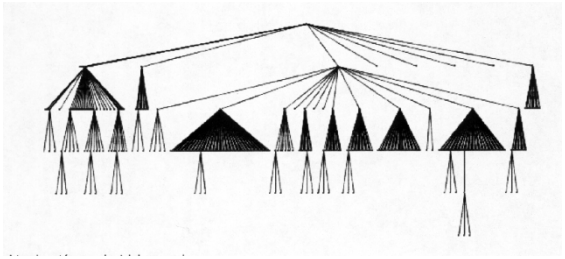




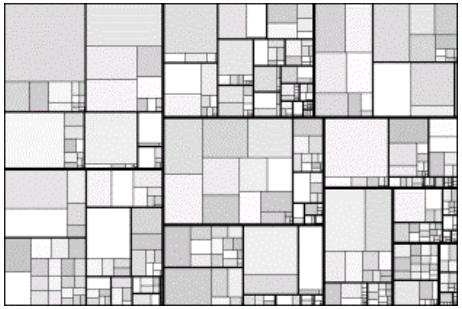


http://www.cc.gatech.edu/gvu/ii/sunburst/

Hierarchical Information – Recap



Traditional



Treemap

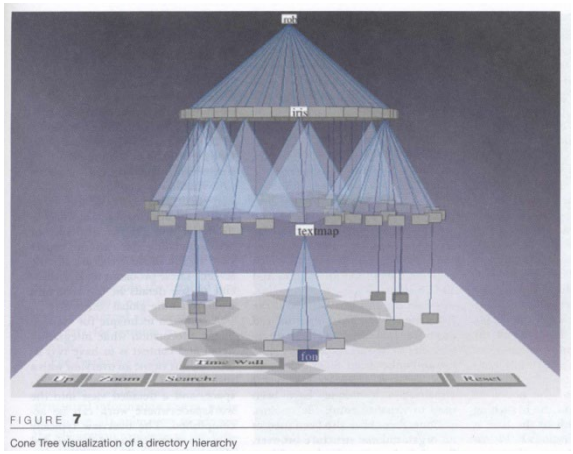
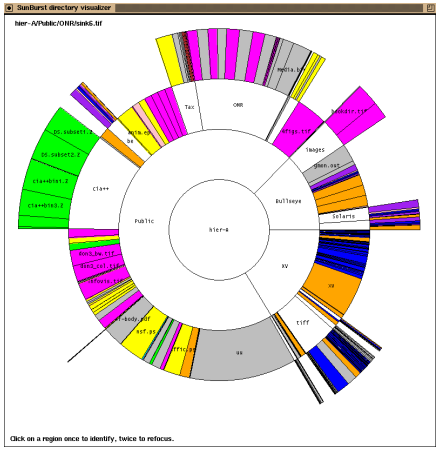
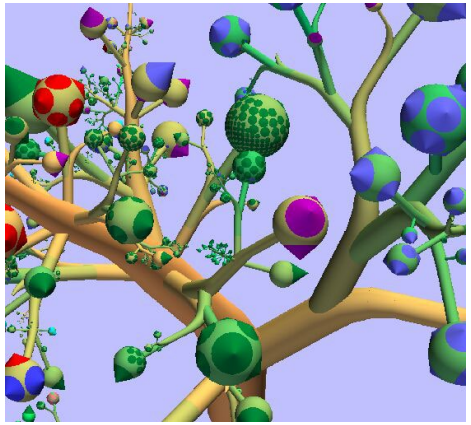


FIGURE 7
Cone Tree visualization of a directory hierarchy

ConeTree



SunTree



Botanical

Focus+Context Interaction

Nonlinear Magnification InfoCenter

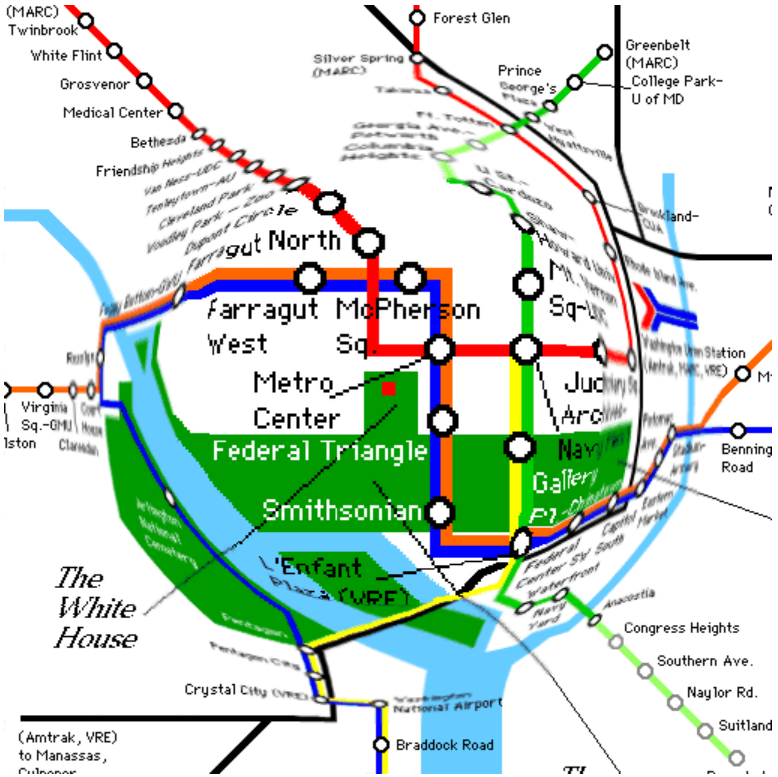
- http://www.cs.indiana.edu/~tkeahey/research/nlm/nlm.html

Nonlinear Magnification

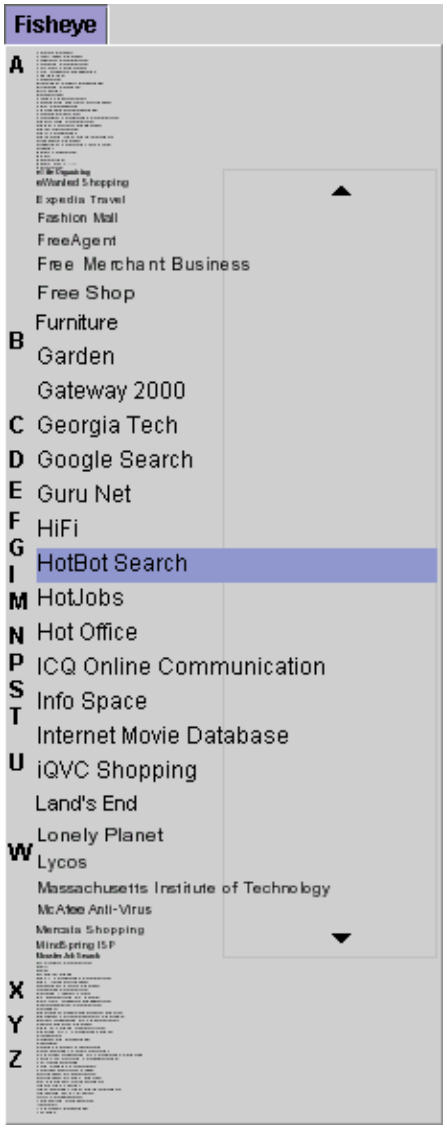
= "Fisheye Views"

= "Focus+Context"

Preserve **Overview**
enable **Detail Analysis**
in **same view**



Fisheye Menus



B. Bederson

- HCI Lab, Uni. of Maryland

Demo

http://www.cs.umd.edu/hcil/fisheyemenu/fisheyemenu-demo.shtml

Table Lens

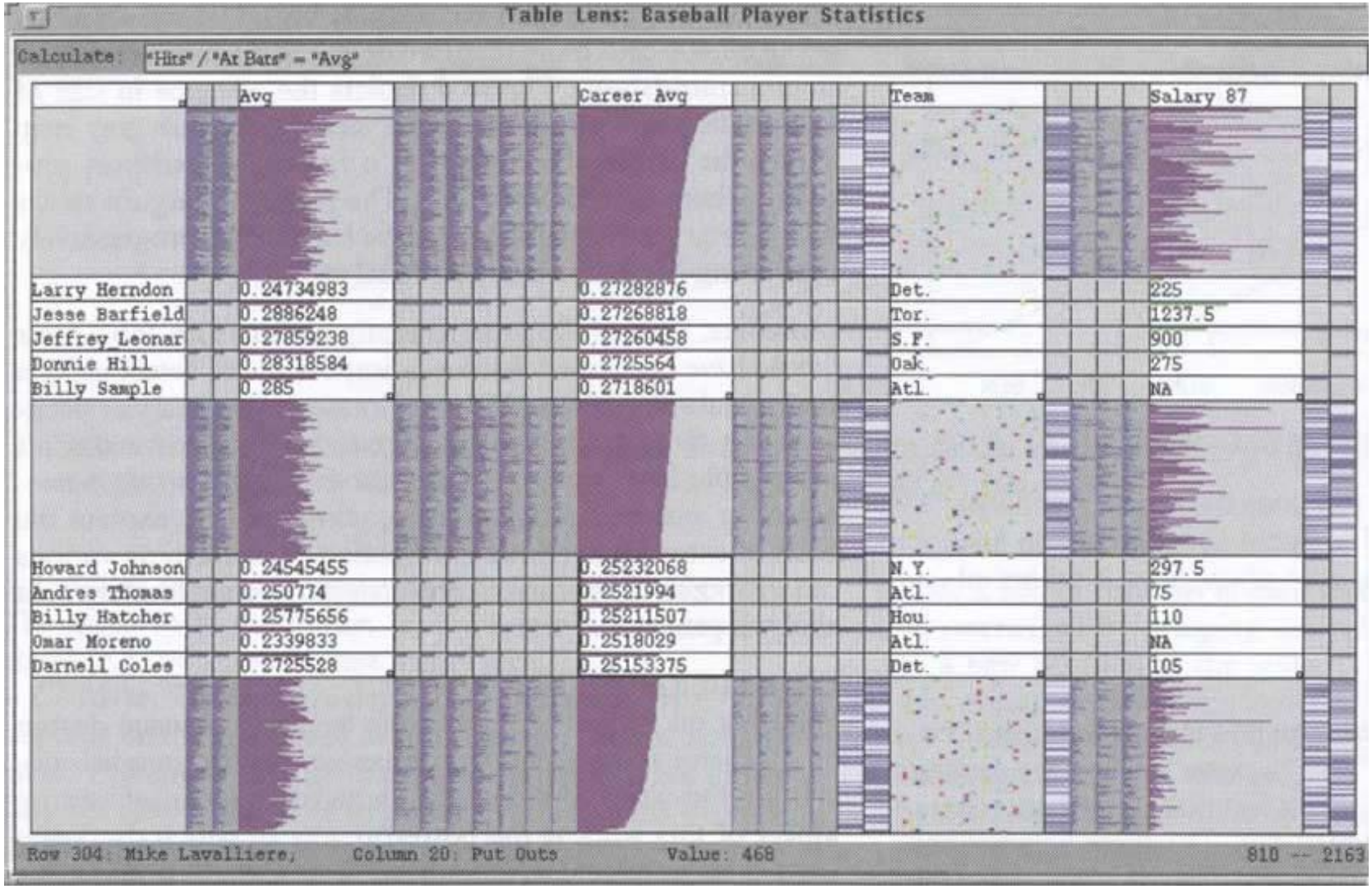


Table Lens – Focus+Context

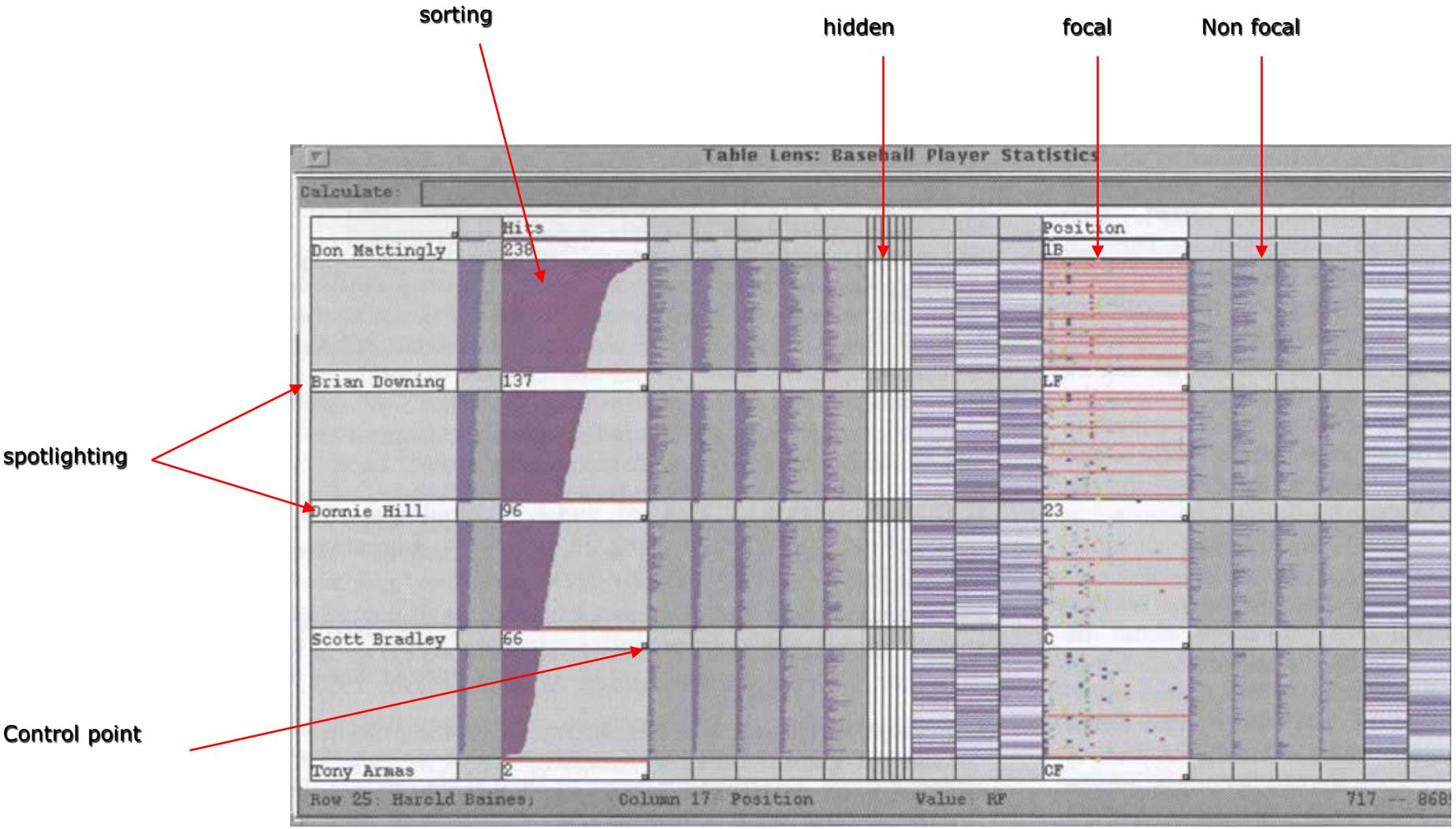


Table Lens (cont.)

SHAPE

- Pattern detection and comparison

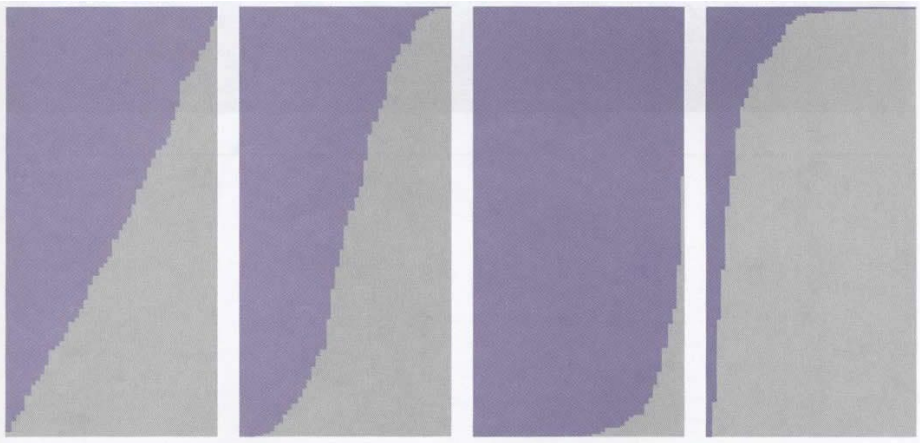


FIGURE 2

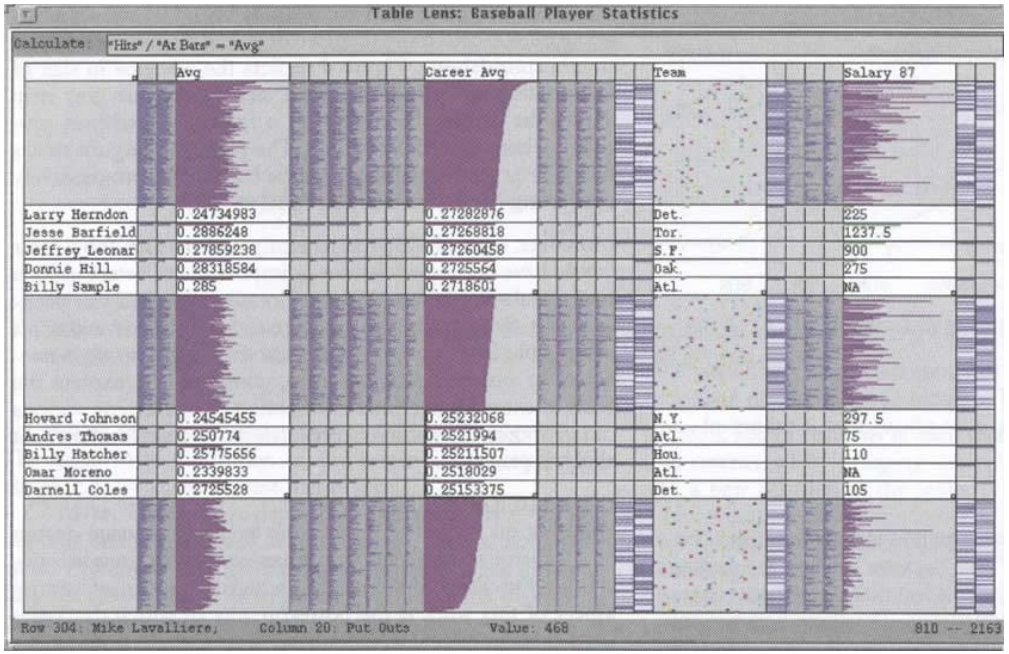
Prototypical distributions as they appear in the Table Lens.

OUTLIERS

- Detect extreme values
- Sort to see MAX and MIN

Table Lens

Data = Multi– Variate



Perceptual Coding

Position	**Yes**
Size	**Yes**
Orientation	
Texture	
Shape	
Color	**Yes**
Shading	
Depth Cues	
Surface	
Motion	**Yes**
Stereo	
Proximity	**Yes**
Similarity	**Yes**
Continuity	**Yes**
Connectedness	
Closure	
Containment	Yes

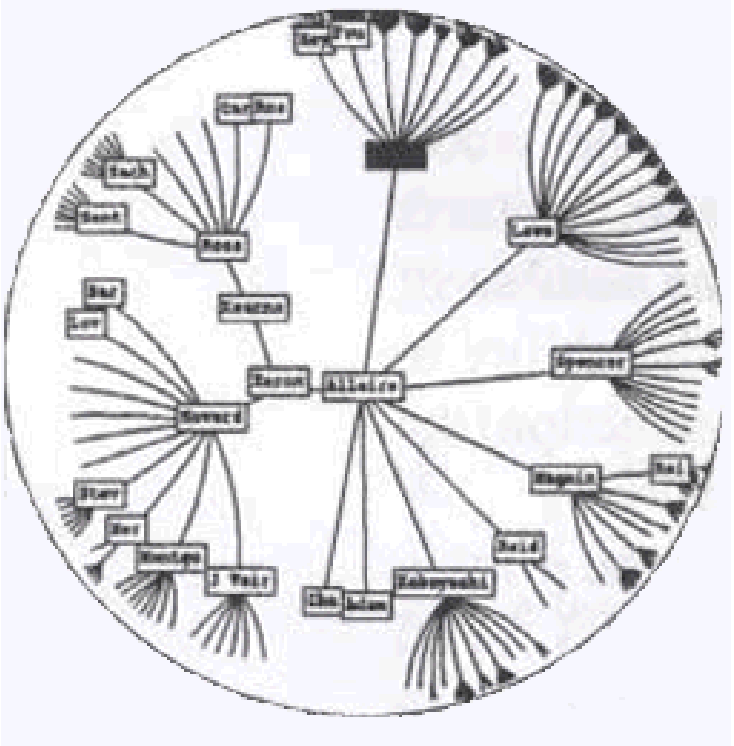
Interaction

Direct Manipulation	**Yes**
Immediate Feedback	**Yes**
Linked Displays	**Yes**
Logarithmic Shift of Focus	
Dynamic Sliders	
Semantic Zoom	**Yes**
Focus+Context	**Yes**
Details-on-Demand	
Output ➔ Input	

Hyperbolic Trees

Visualize Hierarchical Data

Focus + Context Technique



Inxigth StarTree Browser

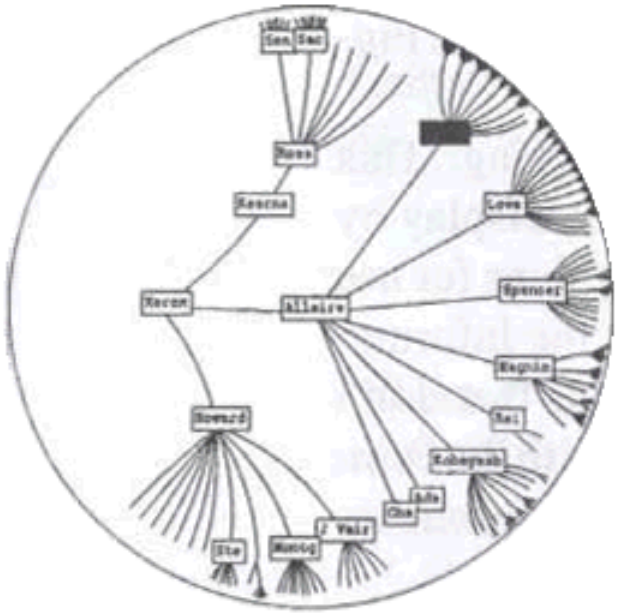
http://www.flashkit.com/search/sitemap/index.shtml

Comparison

- **Standard 2D Browser: 100 nodes** (3 character text strings)
- **Hyperbolic Browser: 1000 nodes** (50 nearest to focus can show from 3 to dozens of characters)

Hyperbolic Trees

Data = Hierarchy



Perceptual Coding

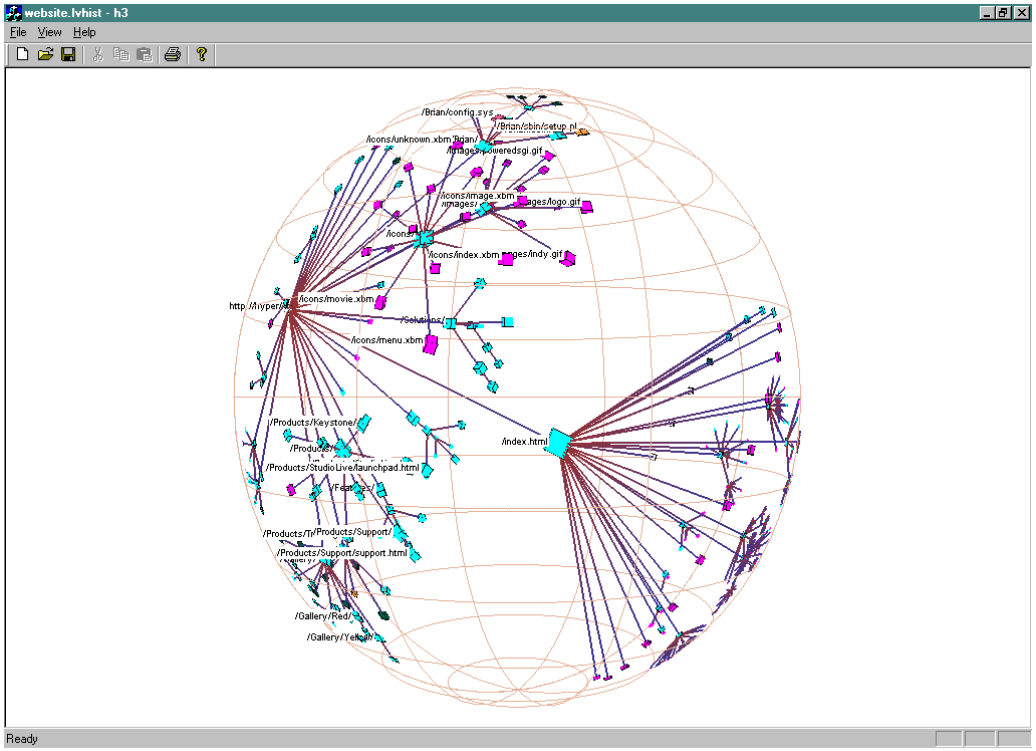
Position	**Yes**
Size	**Yes**
Orientation	
Texture	
Shape	Yes
Color	**Yes**
Shading	
Depth Cues	
Surface	
Motion	**Yes**
Stereo	
Proximity	**Yes**
Similarity	
Continuity	
Connectedness	**Yes**
Closure	
Containment	

Interaction

Direct Manipulation	**Yes**
Immediate Feedback	**Yes**
Linked Displays	
Logarithmic Shift of Focus	**Yes**
Dynamic Sliders	
Semantic Zoom	**Yes**
Focus+Context	**Yes**
Details-on-Demand	**Yes**
Output ➔ Input	

Hyperbolic Tree ➔ 3D

Munzner's H3 / H3 Viewer http://graphics.stanford.edu/videos/h3/



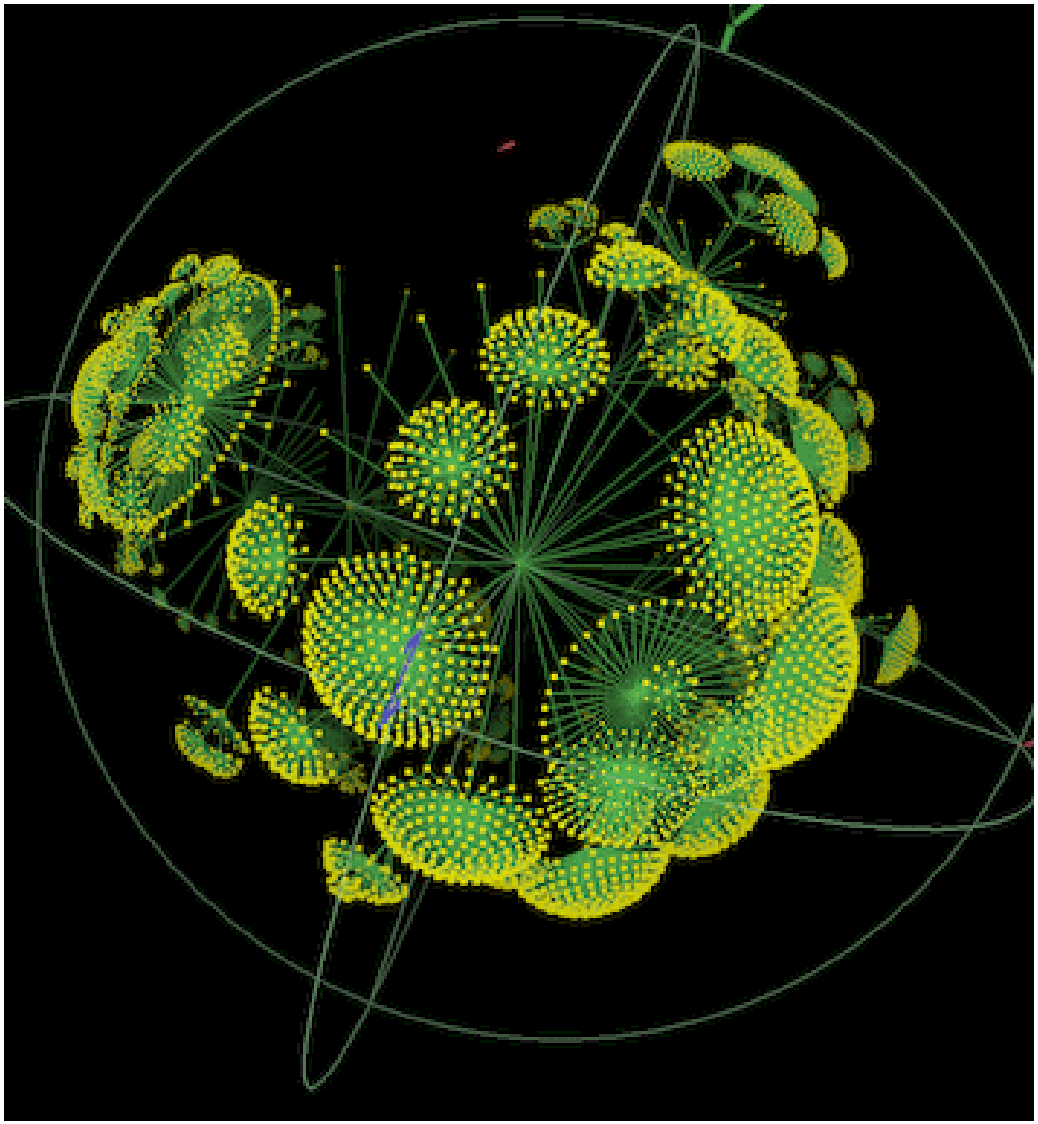
Hyperbolic Browser

- Projection onto sphere rather than circle
- Handles graphs as well as trees

ConeTree

- Distributes child nodes on surface of hemisphere rather than circle circumference

3D Hyperbolic Browser ➔ Walrus



Interaction Benefits

Direct Manipulation	Reduce Short-term Memory Load
Immediate Feedback	Permit Easy Reversal of Actions
Linked Displays	Increase Info Density
Animated Shift of Focus	Offload work from cognitive to perceptual system Object Constancy and Increase Info Density
Dynamic Sliders	Reduce Errors
Semantic Zoom	➔ O(LOG(N)) Navigation Diameter
Focus+Context	➔ O(LOG(N)) Navigation Diameter
Details-on-Demand	Reduce Clutter & Overload
Output ➔ Input	Reduce Errors

Graded Discussion 1 – Analyze Visualization

Data **Quantitative** | Categorical: **Ordinal** | **Nominal** # Data Variables?

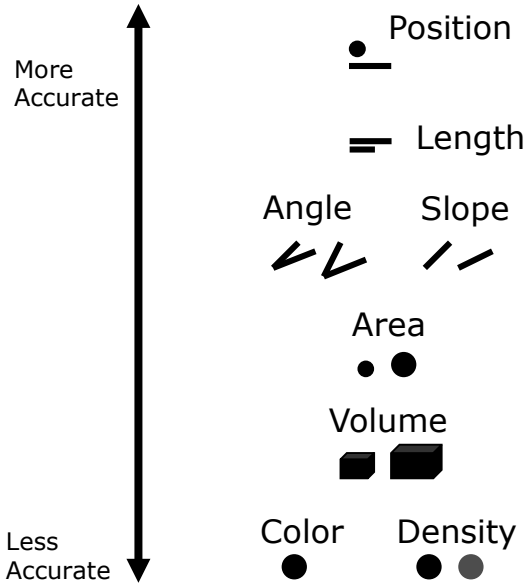
Marks **Points** (position, color, size) | **Lines** (position, angle, length, width, color)

Areas (uniform / smoothed shading) | **Volumes** (resolution, translucency)

Visual Coding

Position
Size
Orientation
Texture
Shape
Color
Shading
Depth Cues
Surface
Motion
Visual Pop Out ?

Accuracy Ranking



Gestalt Principles

Proximity
Similarity
Continuity
Connectedness
Closure
Containment

Tufte's Principles

Avoid Distortion / Chart Junk
Maximize Data-Ink Ratio
Use Direct Labeling
Encourage Comparison
Show Causality