# An Introduction to Information Visualization Techniques for Exploring Large Database

Jing Yang

Fall 2005

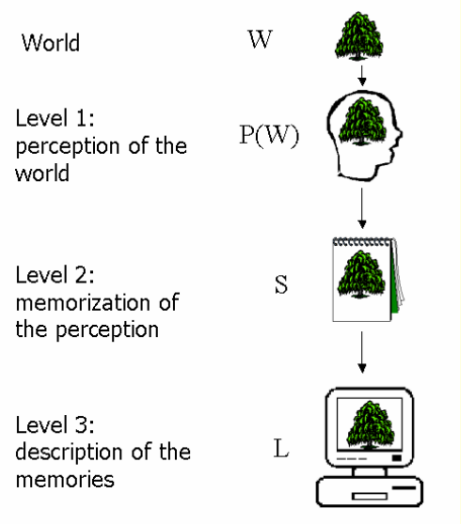
1

# A Multi-resolution Framework

Class 5, Part B

2

# The KRA Model




S. Musti`ere, J.-D. Zucker, and L. Saitta. Cartographic generalization as a combination of representing and abstracting knowledge. *ACM-GIS*, pages 162–164, 1999.

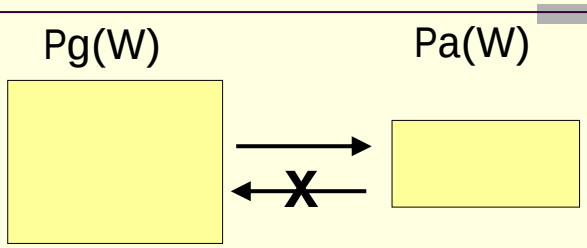
3

# The KRA Model

- *World* W: W is the world where the objects reside.
- *Perception, or View* P(W): P(W) is the perception that an observer has of the world. The perception exists only for the observer and only during its being perceived.
- *Structure* S: S is an extensional representation of the perceived world, in which stimuli relating one to another are stored together.
- *Language* L: L is a language that allows reasoning about the perceived world and communication with others. L allows the perceived world to be described intentionally.

4

# Abstraction and Simplification

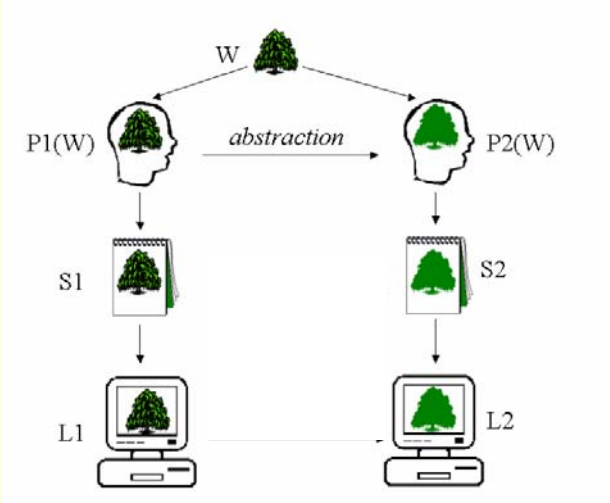



- An abstraction is a functional mapping: Pg(W) → Pa(W)
- Abstraction is a mapping between the views (perceptions) of the world
- The modifications of the structures and the languages, which are called simplification, are side effects that are necessary for describing what happens at the level of the perceived world

L. Saitta and J.-D. Zucker. Semantic abstraction for concept representation and learning. *Proc. Symposium on Abstraction, Reformulation and Approximation*, 1998.

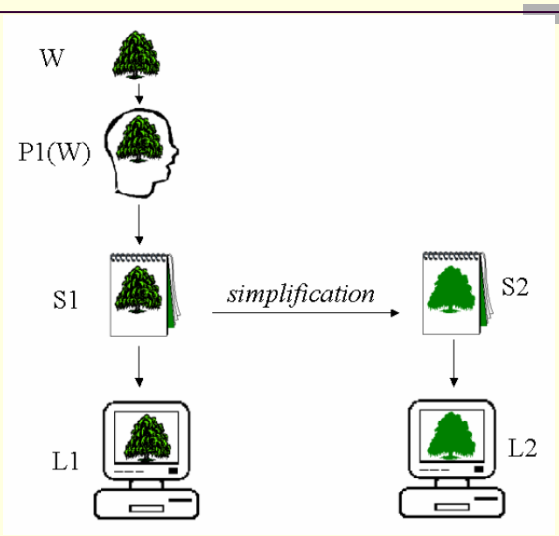
5

# Abstraction



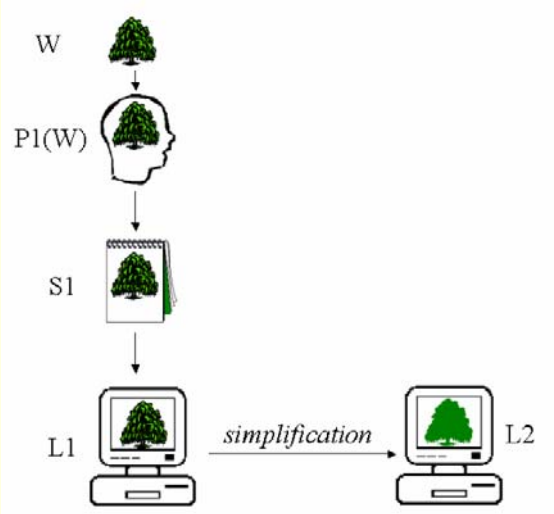

6

# Structure Level Simplification



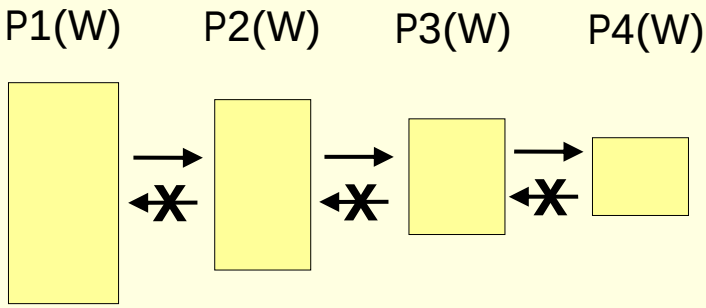

7

# Visualization Level Simplification




8

## Abstraction Hierarchy




Level of detail (LOD):

LOD of P1(W) > LOD of P2(W) > LOD of P3(W) > LOD of P4(W)

9

## Definition of MRV System

- MRV system
  - A visualization system that visually represents abstraction hierarchies of views and allows users to interactively navigate among the views
- Essential features of a MRV system:
  - Hierarchies of views
    - Problem: they are often not provided to the MVT systems
  - Interactive visualization

10

## General Framework for MRV

- View simulation
  - A MRV system simulates abstraction hierarchies of views through simplification if they are not provided
    - Simplification operators
    - Simplification operands (spaces)
- Interactive visualization
  - A MRV system visually presents the views and allows users to interactively navigate among them
    - MRV interfaces
    - MRV interactions

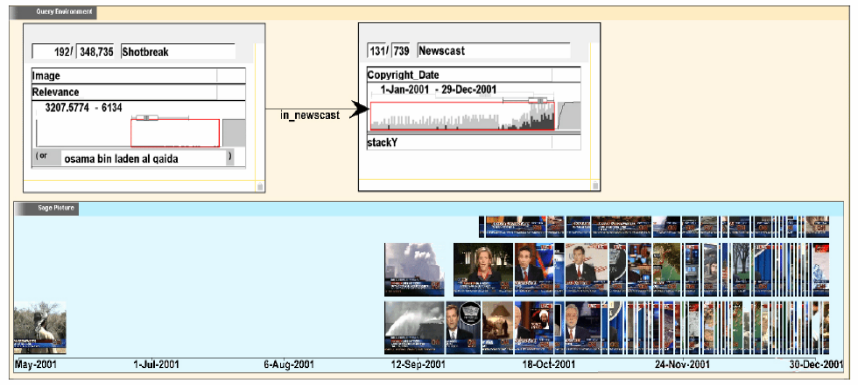
11

## Sampling Operator

- Sampling - the process of selecting some part of a population to observe so that one may estimate something about the whole population [Tho92].
- Example sampling techniques:
  - Simple random sampling
  - Biased sampling [KGKB03]
  - Dynamic sample selection [BCD03]
  - Multi-level sampling [HK04]

12

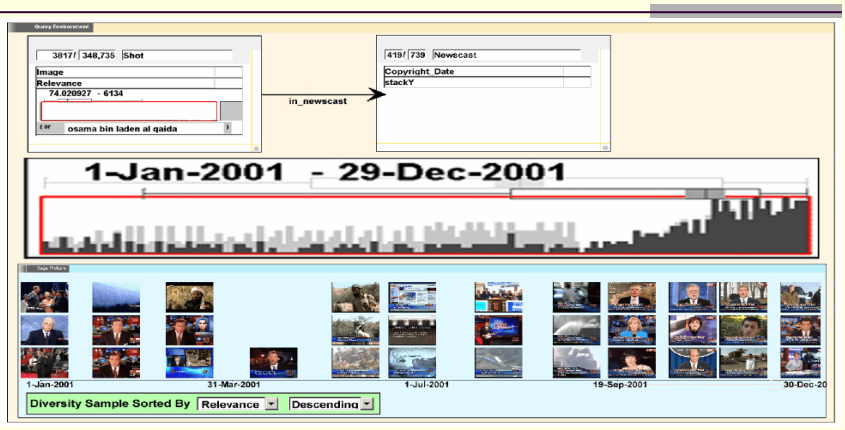
# Sampling in MRV Application




Example: query results to a video library are represented using a collage of representative keyframes [DCHW03]. Occlusion happens in this display.

13

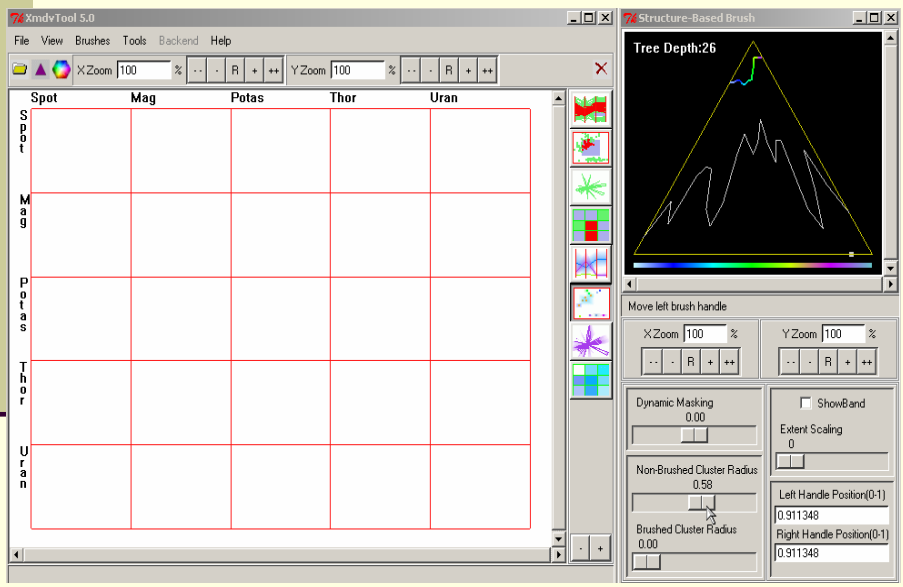
# Sampling in MRV Application




Sampling is used to reduce the number of keyframes shown in the display to make sure that there is no occlusion [DCHW03].

14

# Sampling in MRV Application




Jing' work. Not published yet.

# Aggregation Operators

- Aggregation - a simplification in which a relationship between objects is regarded as a higher level object [SS77].
- Aggregation operators include:
  - Clustering
  - Histograms
  - Other operators

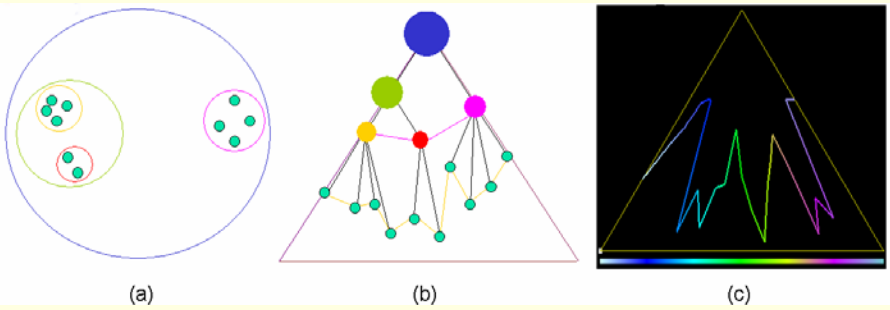
16

## Clustering Operators

- Clustering - a division of data into groups of similar objects. Each group, called a cluster, consists of objects that are similar among themselves and dissimilar to objects of other groups [Ber02].
- Example clustering approaches [Ber02]:
  - Hierarchical clustering
  - Partitioning clustering
  - Grid-based clustering
  - Human-computer clustering

17

## Clustering




(a) Clustering. (b) Hierarchy generated. (c) Approximation display [FWR99a, FWR99b, YWR03b]

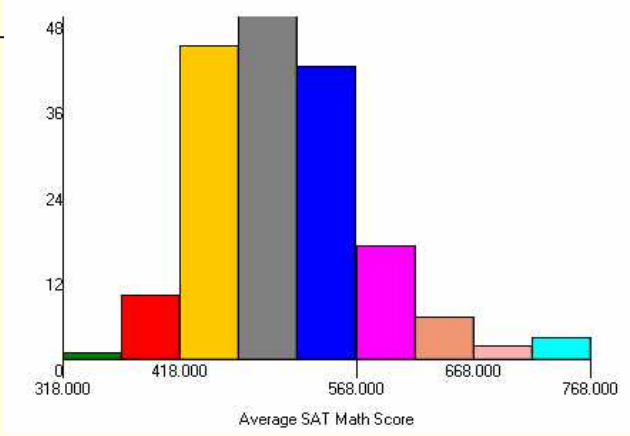
18

# Histogram Operators

- A histogram partitions the data space into buckets. In each bucket, the data distribution is often assumed uniform and recorded using simple statistic data.
- Example histograms:
  - 1-D histograms
  - Multi-dimensional histograms
  - Dynamic histograms

19

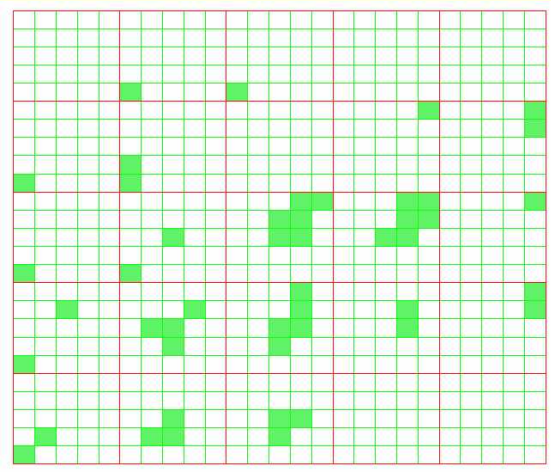
# 1-D Histogram in MRV Application




An equi-width histogram of the SAT math scores of students in some colleges. These figures were captured from [NIS05].

20

# M-D Histogram in MRV Application



A multi-dimensional histogram with equal sized bins visualized in Dimensional Stacking[WLT94].

21

# M-D Histogram in MRV Application

- Possible usage of M-D histograms in MRV systems:
  - Generating approximate displays of multi-dimensional data sets
  - Indicating potential interesting areas in the data sets to users
  - Estimating extent of clutter of a display to be generated

22

## Approximation Operators

- Approximation - a simplification in which objects and relationships are represented by fewer objects and simpler relationships without explicit many to 1 mappings.
- Approximation operators include:
  - Proximity positioning operators
  - Wavelet operators
  - Other operators

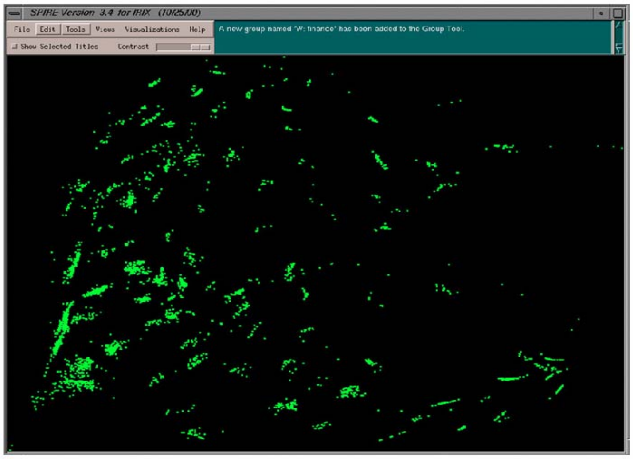
23

## Proximity Positioning Operators

- Proximity positioning
  - Proximity positioning generates a topology preserving map of a collection that gives an overview of similarity among the objects within the collection. In the map, similar objects are positioned close to one another, and far from dissimilar ones [Bas01].
- Example proximity positioning operators:
  - Multi-Dimensional Scaling (MDS)
  - Principal Component Analysis (PCA)
  - Self Organizing Map (SOM)
  - Pathfinder Network Scaling (PNS)

24

## Proximity Positioning Operators in MRV Application



A document collection visualized in the SPIRE Galaxies visualization [NHT01].

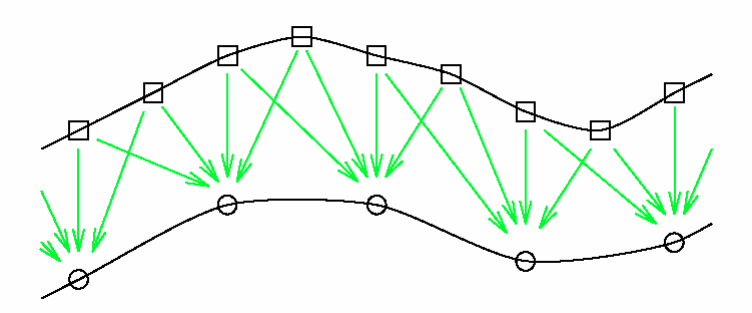
25

## Wavelets Operators

- Wavelets - filter matrices that accept a data stream with items, and generate items of approximations and items of details. The approximation is a coarse summary of the original data, and details contain the data loss during the decomposition [WB96].

26

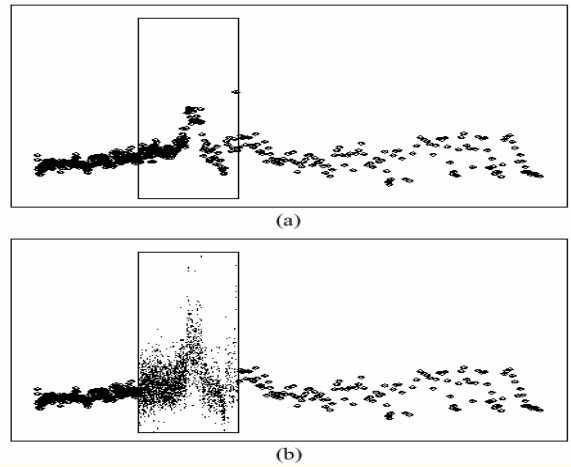
## Wavelets Illustration



A fine data curve is downsized to a coarse one using wavelet decomposition [WB96].

27

## Wavelets in MRV Application

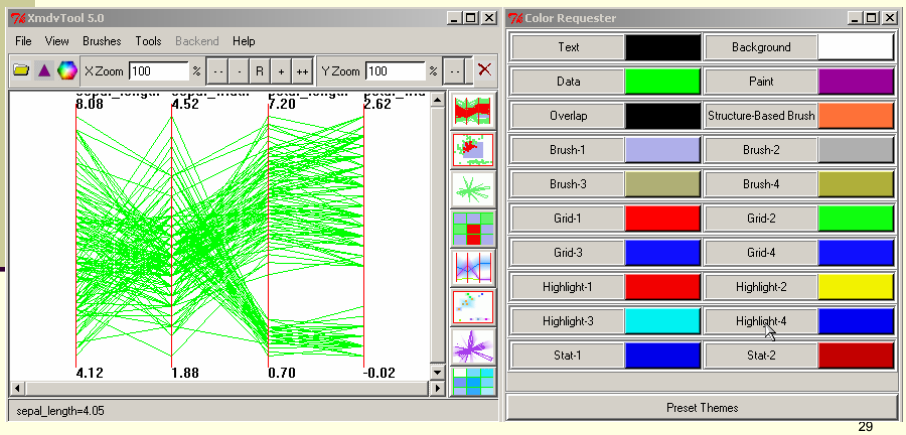



The wavelet brushing example [WB96]. (a) The brush data is defined. (b) Fine brushed data is painted over a coarse data background.

28

# Generalization Operators

- Generalization - a simplification in which a set of similar objects is regarded as a generic object [SS77].




29

# Generalization in MRV Application


30

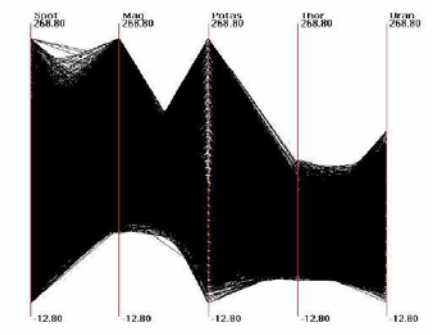
# Simplification Operands

- Data space (structure level)
  - Data item space
  - Dimension space
  - Topology space
- Visualization space (language level)
  - Visualization structure space
  - Visual encoding space
  - Screen space

31

# Data Item Space

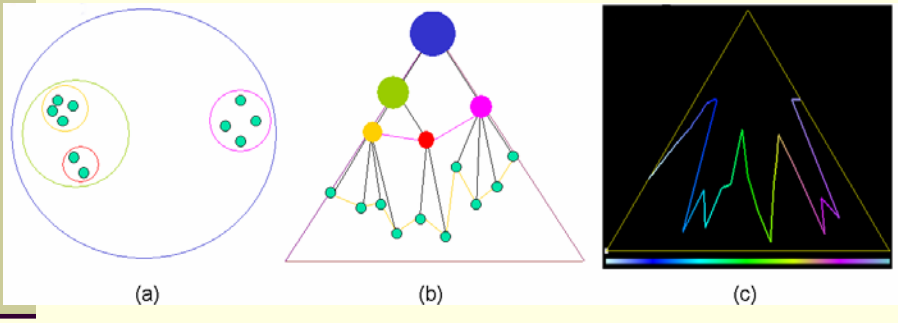
- Data items - individual objects contained in a data set




The clutter problem of data item space [YWR03b]. The Out5d data set (size: 16384 data items, 5 dimensions) visualized with parallel coordinates.

32

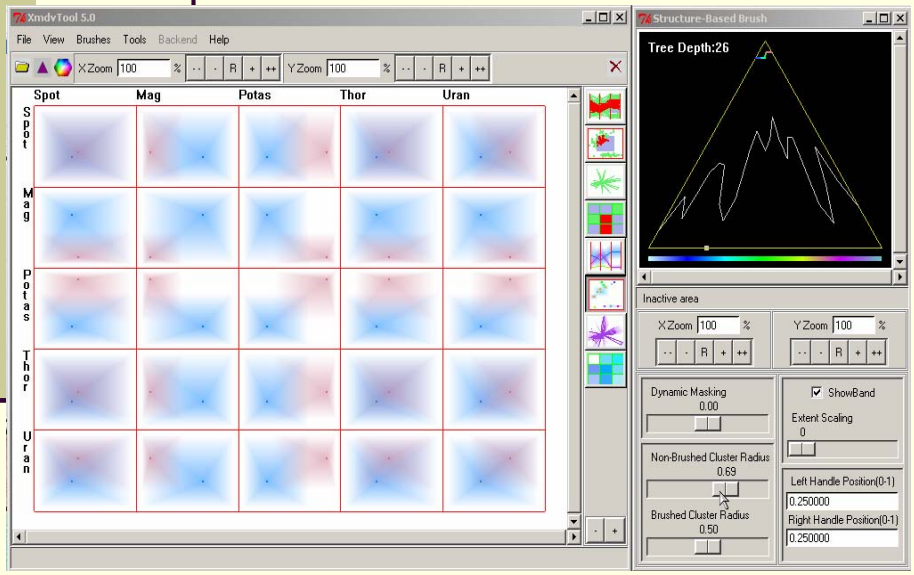
# Data Item Space Simplification Example – IHD framework




(a) Clustering. (b) Hierarchy generated. (c) Approximation display [FWR99a, FWR99b, YWR03b]

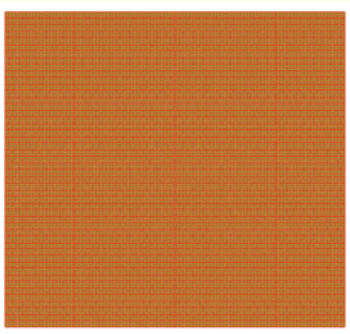
33

# Data Item Space Simplification Example – IHD framework

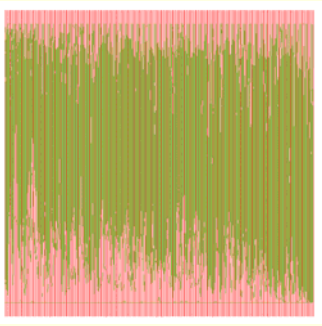


# Dimension Space

- Dimensions - individual attributes of objects contained in a data set



46,225 plots

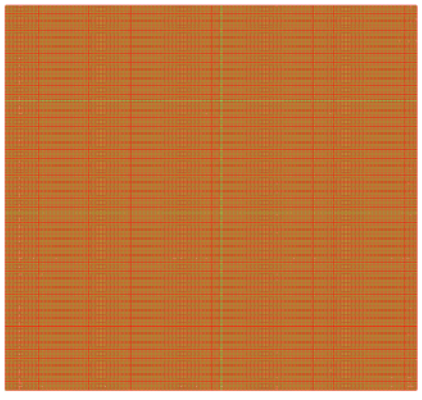


215 axes

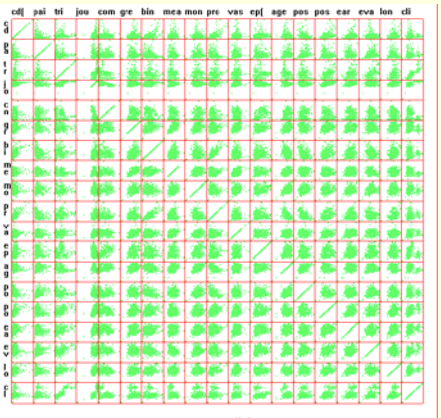
The clutter problem of dimension space. The OHSUMED dataset (215 dimensions) visualized with a scatterplot matrix and parallel coordinates.

35

# Dimension Space Simplification Example – VHDF Framework



(a)



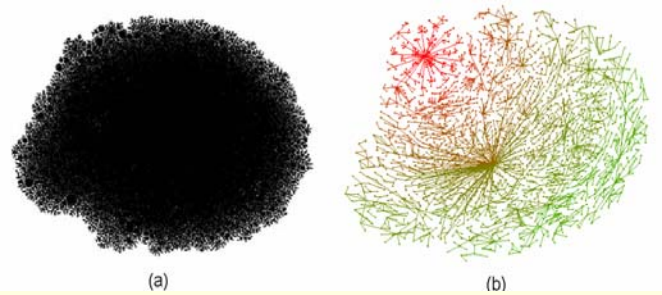
(b)

Dimension filtering in scattterplot matrices [YPWR03]. The OHSUMED data set is shown. (a) Before filtering (b) After filtering.

36

# Topology Space

- Topology - geometric, spatial, temporal, or logical relationships among objects in a data set.

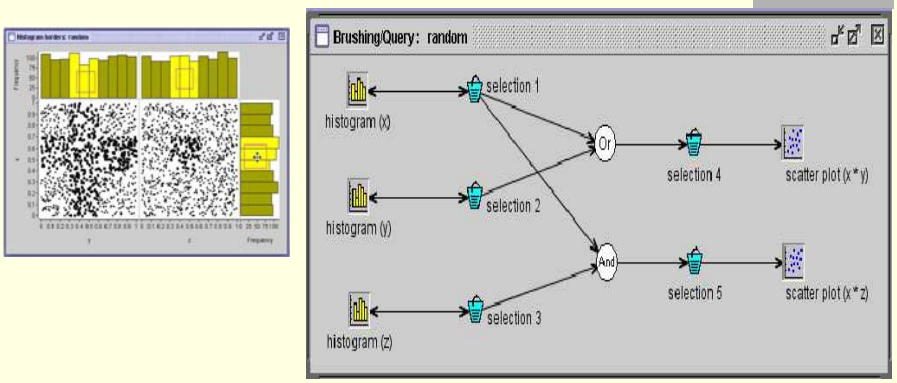



A graph that contains 87,931 vertices and 87,930 edges [GKN04]. (a) The display is seriously cluttered. (b) The same graph displayed in lower levels of detail.
Simplification operator: hierarchical clustering

37

# Visualization Structure Space

- Visualization structure - the organization of a visualization

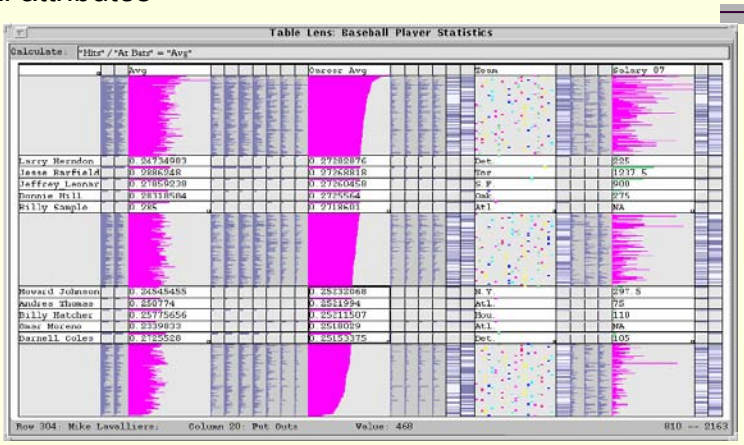



The Compound Brush [Che03]. Users can easily modify the brushing process by manipulating visual entities in the right window.
Simplification operator: generalization

38

# Visual Encoding Space

- Visual encoding - the mappings between data and visual attributes



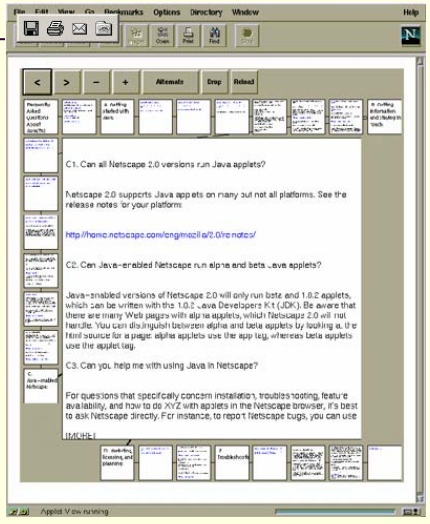

A basketball player statistics data set shown in the Table Lens [RC94].

39

# Screen Space

- Screen space - the pixels composing a display

The Zoom Browser displays context web pages as tiles around the focus page [Hol97]. Each tile contains a thumbnail or a summary of a context web page.




40

# Interactive Visualization

- Common interfaces of MRV systems:
  - Zoomable interface
  - Overview + detail interface
  - Focus + context interface
- Common interactions of MRV systems:
  - Zooming/panning
  - Selection
  - Distortion
  - Overlap reduction
  - Preview
  - Animation

41

# Zoomable Interface and Zooming/Panning

- Zoomable interface - an interface in which objects are organized in space and scale and users directly interact with the information space mainly through panning and zooming [HBP02].
- Zooming in - the interaction that changes the current display from a view of a lower level of detail to a view of a higher level of detail.
- Zooming out - the interaction that changes the current display from a view of a higher level of detail to a view of a lower level of detail.
- Panning - the interaction that changes the current display from a subregion of a view to an adjacent sub-region of the same view. There can be overlaps between the two regions.

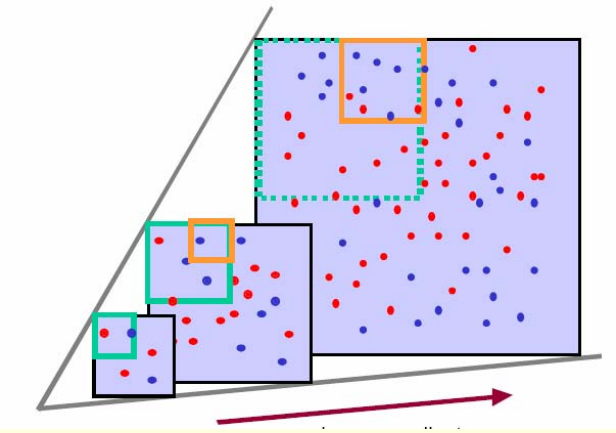
42

# Zooming in Screen-Space (Pixels)

- Example: Zooming in XmdvTool
- Demo

43

# Zooming in Data Space

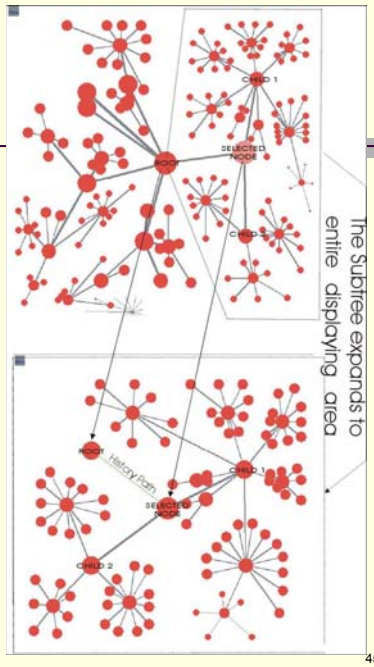


Astral Telescope Visualiser
By Chance - Enhancing Interaction with Large Data Sets through Statistical Sampling (A. Dix and G. Ellis. 2002)

44

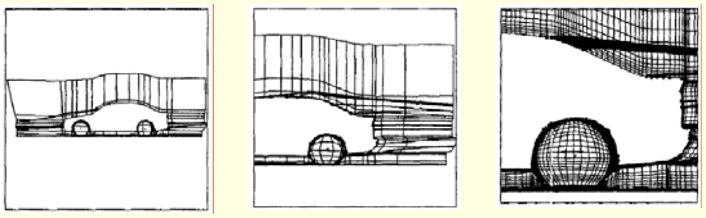
## Zooming in Data Space

- Semantic zooming in Space-Optimized Tree




45

## Zooming in Data Space



Wavelets over Curvilinear Grids
(G.M. Nielson, I.-H. Jung, and J. Sung. 1998)

46

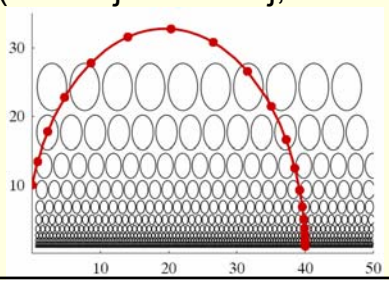
# Zooming in Data Space


Interactive map from www.mapquest.com

47

# Panning and Zooming

- Panning in high levels of detail can be time consuming
  - Solution: zoom out, pan, and zoom in
  - Drawbacks: jitter in the process
- Improvement: Smooth and Efficient Zooming and Panning (van Wijk and Nuij, Infovis 03)




48

## Panning and Zooming

- "Speed-Dependent Automatic Zooming for Browsing Large Documents" Igarashi & Hinckley, Proc. UIST'00, pp. 139-148.
  - Keep constant perceptual scrolling speed
  - Scale X Speed = Constant

  Video!

49

## Zooming and Panning

- SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation Grosjean, Plaisant and Bederson, InfoVis 2002
  - A zooming environment that dynamically lays out branches of a tree to best fit and available screen space
  - Video

50

## Overview + Detail Interface and Selection

- Overview + detail interface - an interface composed of multiple windows. Some windows provide context and easy navigation (overview windows) for other windows (detail windows).
- Selection - the interaction that isolates a subset of entities on a display for further operations [Wil96].
- Concepts: LOD, Sub-region

51

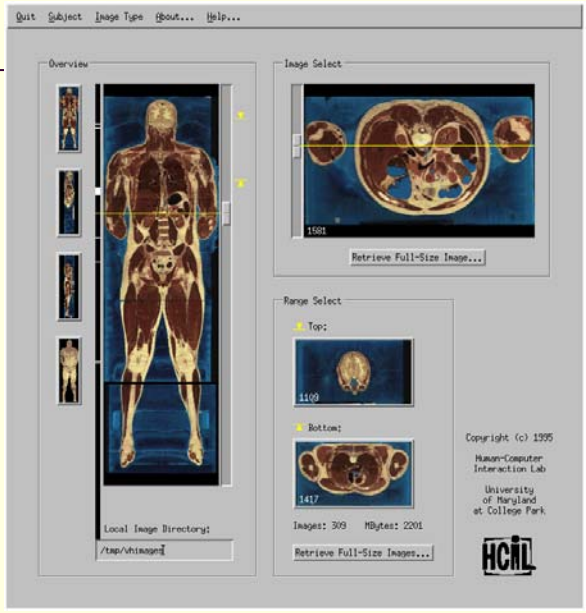
## Different Types of O+T Interfaces

Type 1

- Overview window - a view of a lower LOD
  - Provide context
  - Allow users to select sub-regions in detail window
  - Jump in detail window is allowed (compared to panning)
- Detail window – a view of a higher LOD

52

# Example – Visible Human Explorer




53

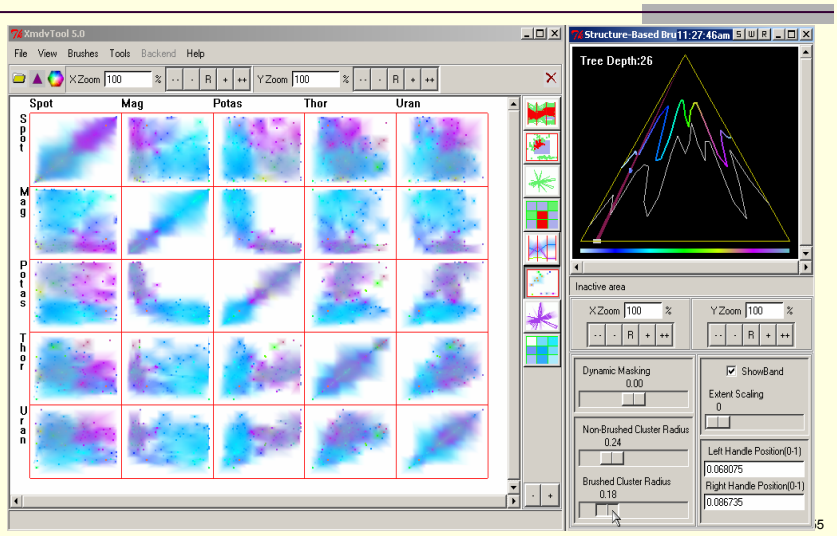
# Different Types of O+T Interfaces

Type 2

- Overview window
  - Users are allowed to select both LOD and sub-region of detail window
- Detail window – a view of a higher LOD

54

# Example – Hierarchical Parallel Coordinates



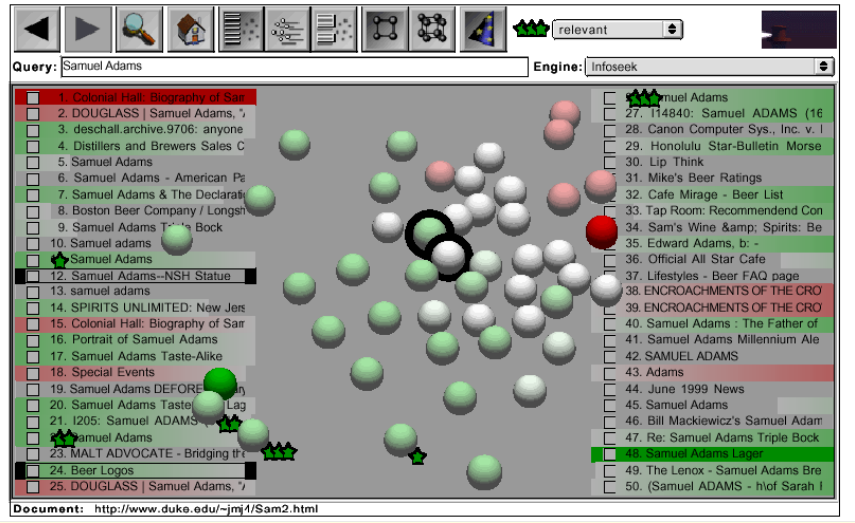

# Different Types of O+T Interfaces

Type 3

- Focus can be changed in both windows
- Hard to tell which window is overview

56

# Example – Lighthouse System




57

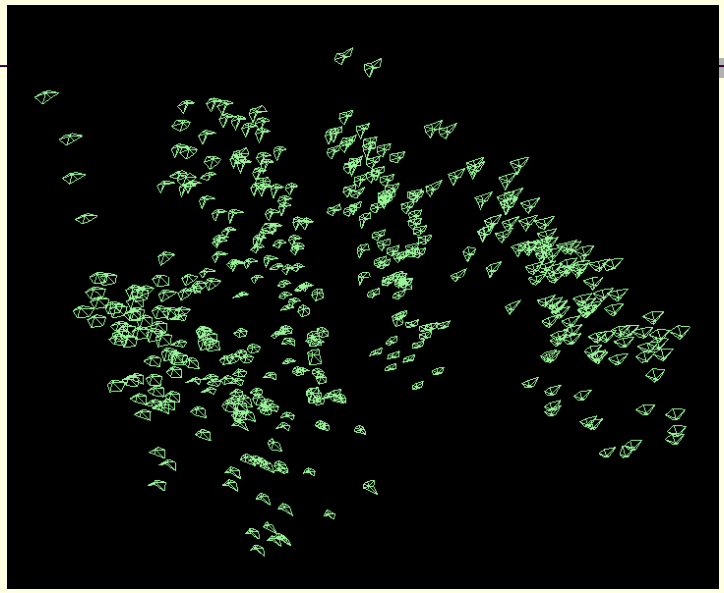
# Different Types of O+T Interfaces

Type 4

- Overview and detail view are mixed in one display
  - Compact view
  - Easy navigation

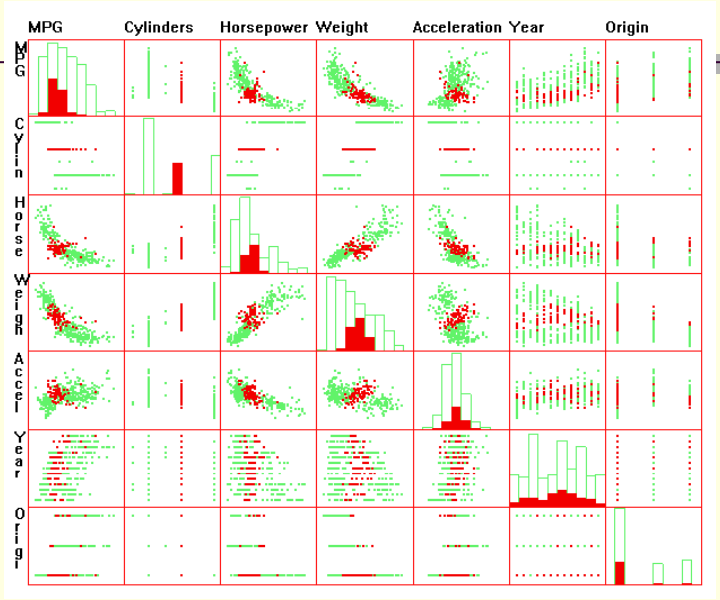
58

# Example – XmdvTool



59

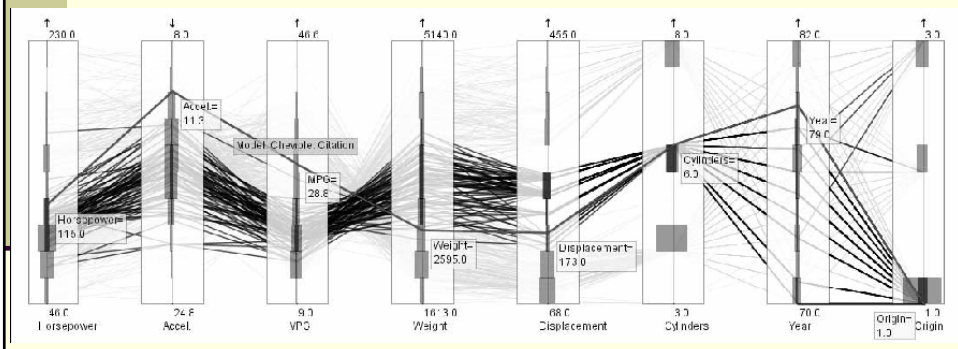
# Example – XmdvTool




60

# Example – Extended Parallel Coordinates

- Hauser et. al. Infovis 2002




61

# Refresh of Detail Window

- Alternative approaches:
  - Immediately refresh detail window after selection in overview window
  - Wait until users trigger the refresh
- Discussion: Compare these two approaches

62

# Features of O+T Interfaces

- Benefits
  - Efficient navigation
  - Context
  - Feeling of control
- Drawbacks
  - Spatially indirect relation between O and T
  - Require more screen space

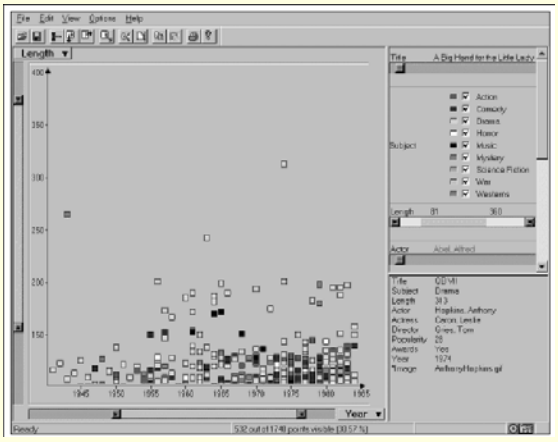
Reference: Hornbæk, Bederson and Plaisant, ACM Trans. Comput.-Hun. Interact. , 9(4): 362-389, 2002

63

# Dynamic Query [Tanin at. el. Infovis 97]

- A mechanism for specifying queries and visualizing their results

Example: spotfire



64

## Selection

- Selection - the interaction that isolates a subset of entities on a display for further operations

65

## N-D Brushes in XmdvTool

- Features
  - Move-over
  - Ramp
  - Multiple brushes
  - Boolean operations of multiple brushes

66

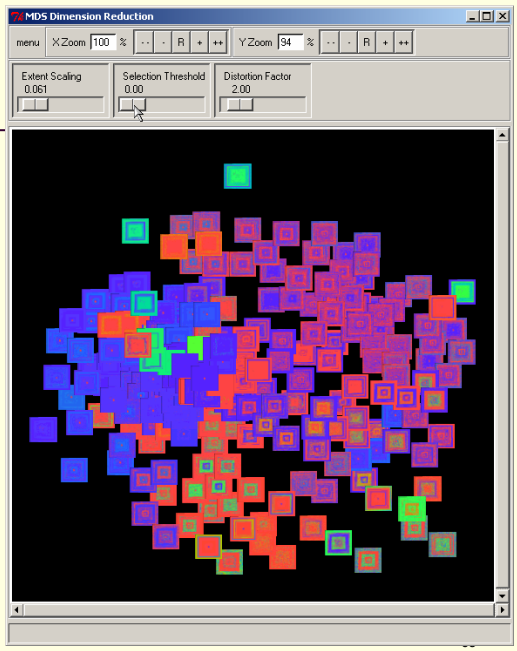
# Semi-automatic Selection

- Why?
- How?
- Example1: InterRing
- Example2: hierarchical parallel coordinates

67

# Semi-automatic Selection

- Example 3: VaR display

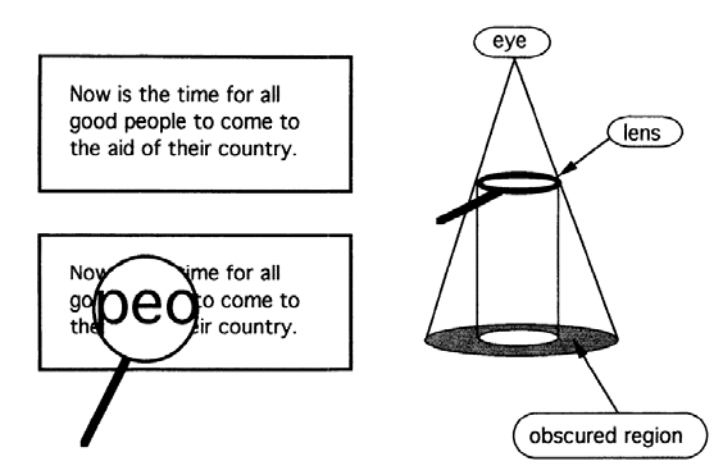


## Focus + Context Interface and Distortion

- Focus + context interface - an interface where views of different levels of detail are mixed together in the same display.
- Distortion - an operation that increase the screen space allocated to some objects in the display while decreasing the screen space allocated to other objects.

69

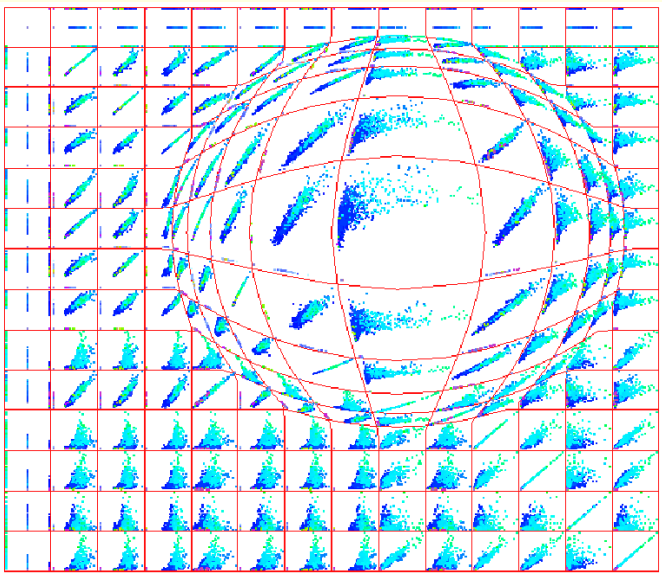
## Magnifier Lens




70

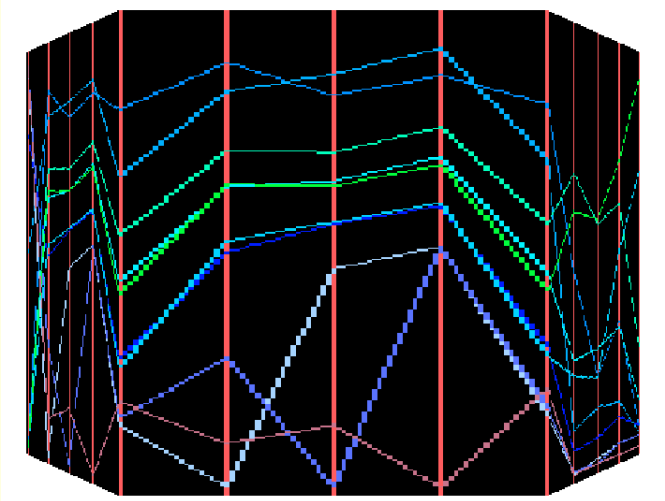
Figure from [Robertson & Mackinlay UIST 93]

## FishEye Lens [Fur86]



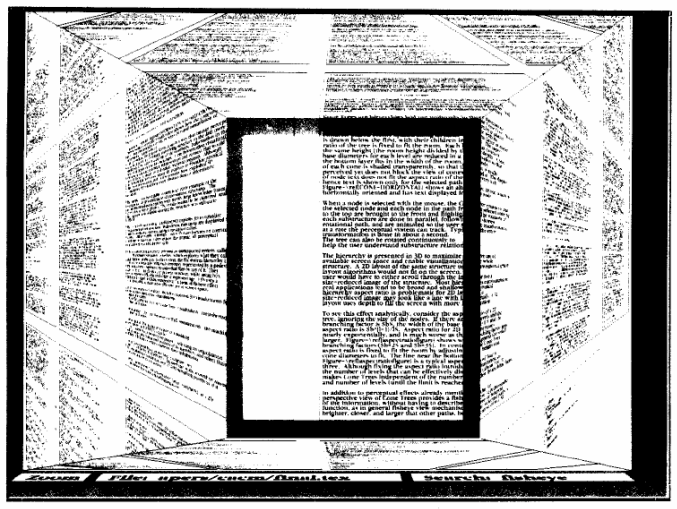
71

## Perspective Wall [MRC91]



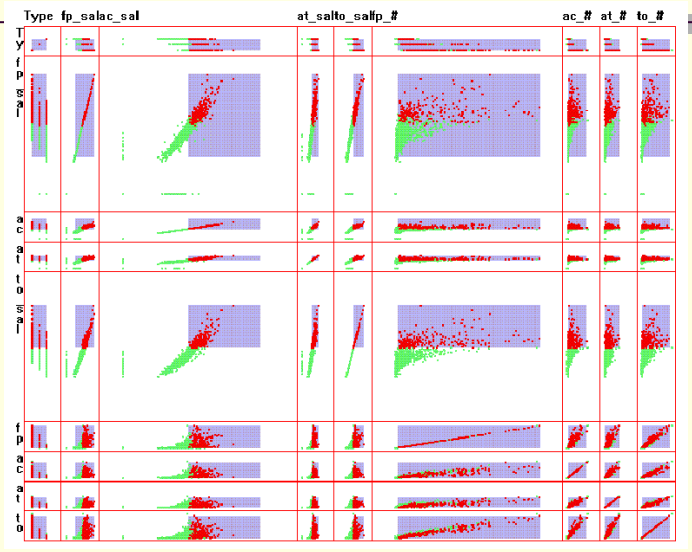
72

# Document Lens
# [Robertson & Mackinlay UIST 93]



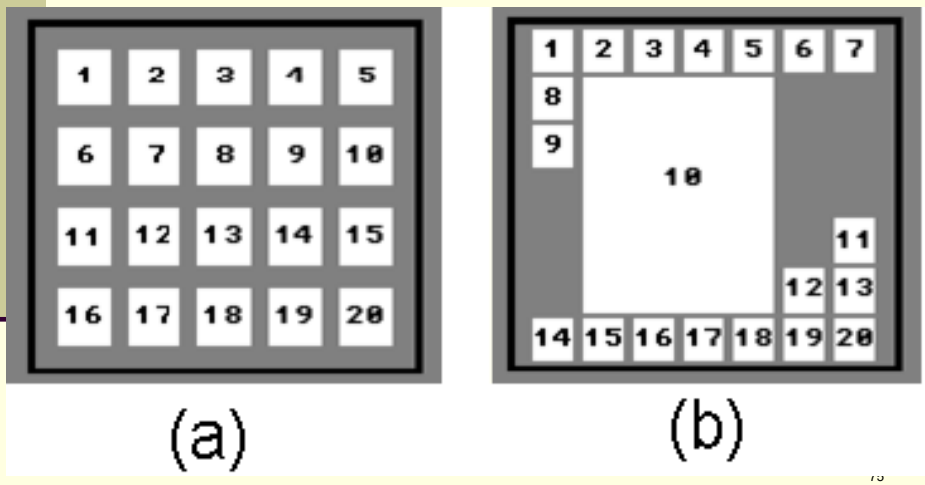
73

# Table Lens [RC95]

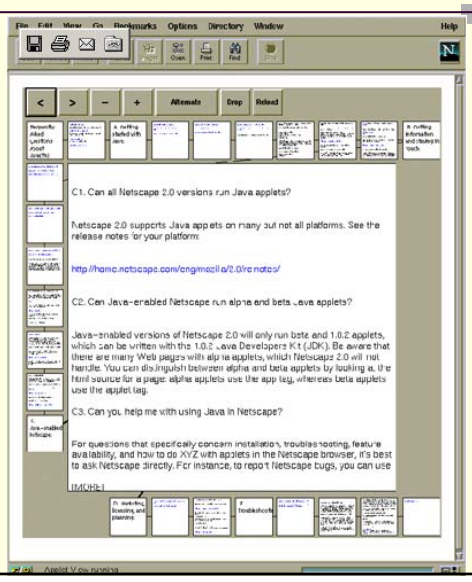



74

# Flip Zooming [Holmquist SIGCHI 97]

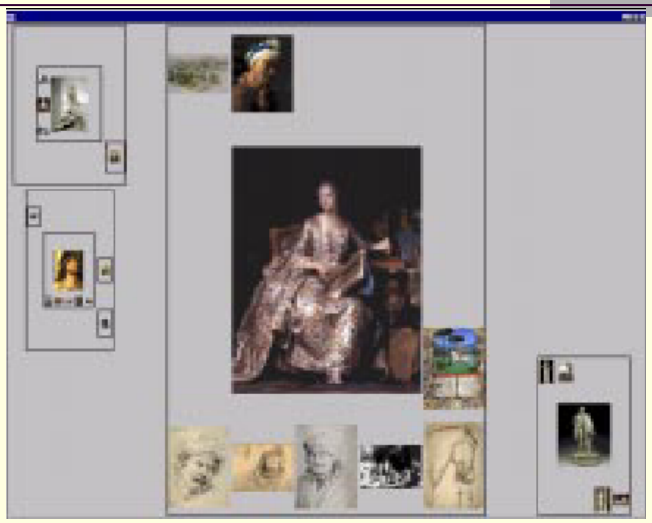



# Flip Zooming




76

## Hierarchical Image Browser [Holmquist and Bjőrk SIGGRAPH 98]



77

## EdgeLens [Wong at. el. Infovis 03]

- Video (InfoVis 03)

78

## MoireGraph [Jankun-Kelly and Ma Infovis 03]

- Video (Infovis 03)

79

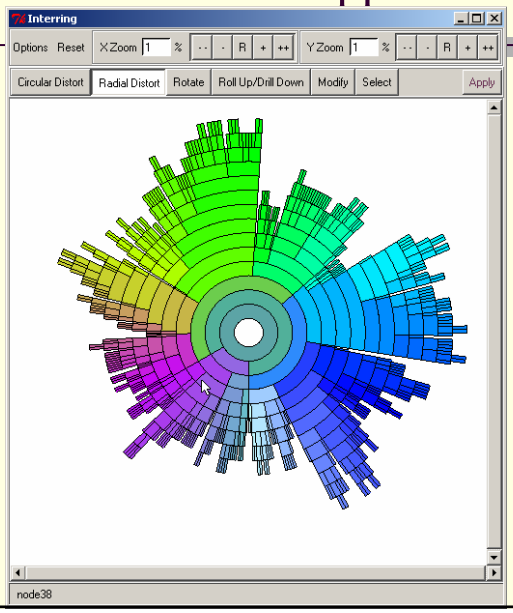
## Datelens

- Video
- http://www.cs.umd.edu/hcil/datelens/

80

# Focus + Context Interface and Distortion in MRV Application

InterRing [YWR02, YWRP03]




81

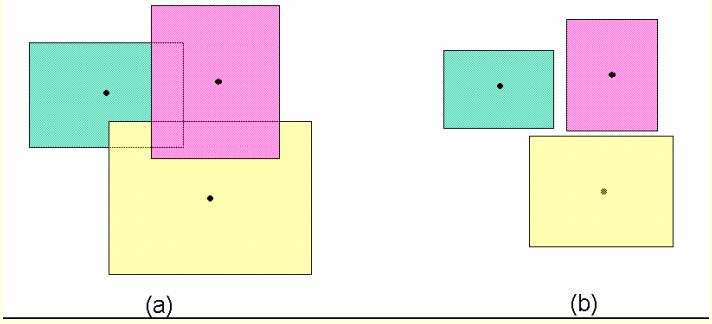
# Overlap Reduction

- Reloaction – adjust the location of objects
- Manual
  - Automatic
  - Position shifting
  - Demo (XmdvTool)

82

# Overlap Reduction

- Scaling – proportionally change the sizes of all objects
- Demo (xmdvtool)

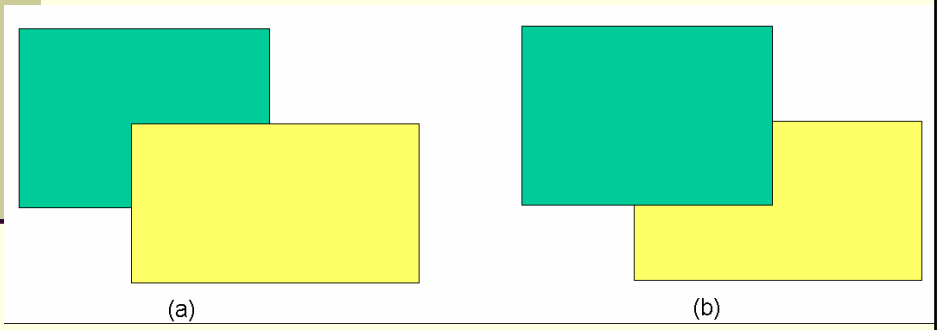



83

Illustration

# Overlap Reduction

- Layer reordering – changes the order in which objects are drawn




84

Illustration

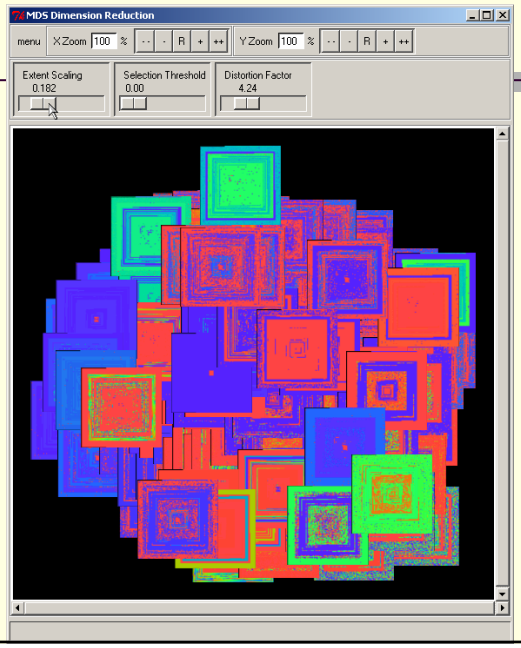
# Overlap Reduction

- Masking – hide some objects in the display
- Demo (xmdvtool)

85

# Overlap Reduction in MRV Application

- VaR display [YPH+04]
  - Extent Scaling
  - Dynamic Masking
  - Zooming and Panning
  - Showing Names
  - Layer Reordering
  - Manual Relocation
  - Automatic Shifting



# Previewing

- Motivation
- Preview – a displayed generated of relative small cost that provides summarization info or simple judgments of a display to be generated

87

# Animation

- A smooth transition that relates the old display to the new one when display changes in an interface
- A commonly held belief
  - Animation helps users maintain object constancy and thus helps users to relate the two states of the system
- A reported user study [Bederson and Boltman Infovis99]:
  - Increased users' ability to reconstruct info space
  - No penalty on task performance
    - Cost extra in response time vs. Relate two states fasters

88

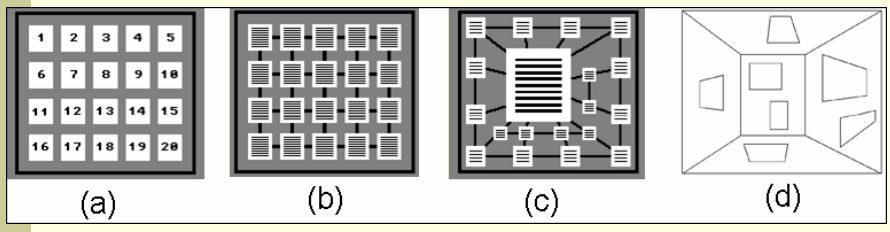
# Reference

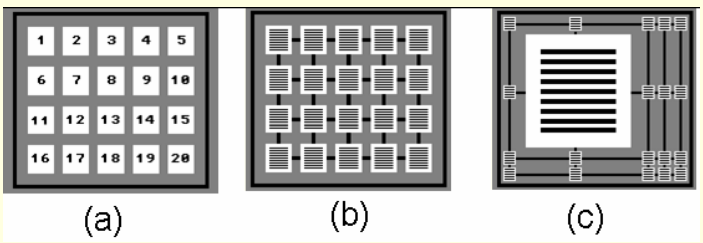
- Jing Yang's dissertation

(It is in course webct/References)

89

# A Case Study



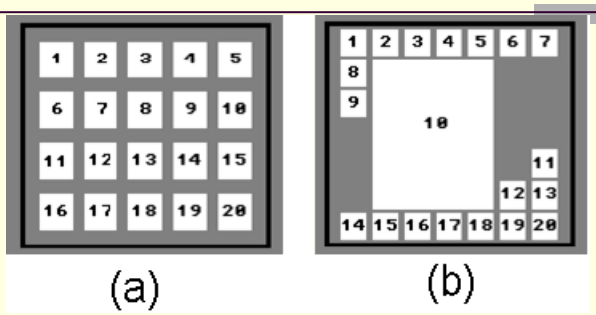

90

# First Law of Geography

- Every thing is related to everything else, but closer things are more closely related.

91

# A Case Study




92