Introduction

Databases and Visualization

Database Support for Visualization Systems

Visualization Support for Database Systems

Data Visualization

Query Visualization

Schema Visualization

Exploratory Analysis

Confirmatory Analysis

Presentation

Introduction

Database Support of Visualization Systems: Specific Requirements

- Specific Data Types (extensibility important)
  (e.g., 2D- & 3D-grids, arbitrary graphs, ...)
- Specific Operations (extensibility important)
  - ⇨ data selection (e.g., sampling, subsetting, ...)
  - ⇨ data manipulation (e.g., filter, interpolation, ...)
  - ⇨ data representation (e.g., attribute mapping, color mapping, ...)
  - ⇨ viewing (e.g., zoom, rotation, ...)
- Specific Transaction Types (read-only)
- Specific Performance Requirements (real-time)

Introduction

Objectives of Visualization Support of Database Systems

- ❑ Visualization of the Database Schema
  (for database design)
- ❑ Visualization of Queries
  (for query specification)
- ❑ **Visualization of the Data**
  (for database exploration)

Introduction

Visual Query Interfaces

- ❑ Tabular Query Interfaces
  (e.g., QBE [Zlo 77], Example-based Graphical Query Interfaces [OW 93])
- ❑ Graphical Support of Textual Query Languages
  (e.g., GRADI [KL 93])
- ❑ Schema-oriented Query Interfaces (e.g., SNAP [BH 86])
- ❑ Slider-based Query Interfaces (e.g., IVEE [AW 95], VisDB [KK 95])

Problem: specification of complex queries is difficult

⇨ example of a query interface for complex queries: InfoCrystal [Spo 93]

Introduction

Goals of Data Visualization

- **Explorative Analysis**
  - starting point: data without hypotheses about the data
  - process: interactive, usually undirected search for structures, trends, etc.
  - result: visualization of the data, which provides hypotheses about the data
- Confirmative Analysis
  - starting point: hypotheses about the data
  - process: goal-oriented examination of the hypotheses
  - result: visualization of the data, which allows the confirmation or rejection of the hypotheses
- Presentation
  - starting point: facts to be presented are fixed a priori
  - process: choice of an appropriate presentation technique
  - result: high-quality visualization of the data presenting the facts

Introduction

Historical Overview of Exploratory Data Visualization Techniques (cf. [WB 95])

- pioneering work of Tufte [Tuf 83, Tuf 90] and Bertin [Ber 81] focuses on
  - ⇨ visualization of data with inherent 2D-/3D-semantics
  - ⇨ general rules for layout, color composition, attribute mapping, etc.
- development of visualization techniques for different types of data with an underlying physical model
  - ⇨ geographic data, CAD data, flow data, image data, voxel data, etc.
- development of visualization techniques for arbitrary multidimensional data (without an underlying physical model)
  - ⇨ applicable to databases and other information resources

Introduction

Classification of Data Visualization Techniques

- Geometric Techniques: Scatterplots, Projection Pursuit, Prosection Views, Hyperslice, ***Parallel Coordinates***, ...
- Icon-based Techniques: Chernoff Faces, ***Stick Figures***, Shape-Coding, Grouping Technique...
- Pixel-oriented Techniques: ***Recursive Pattern Technique*****,** Spiral- & Axes-Techniques, ...
- Hierarchical Techniques: Dimensional Stacking, Hierarchical Plotting, Worlds-within-Worlds, ***Treemap***, Cone Trees, ...
- Graph-Based Techniques: SeeNet, ***Hygraphs***, Fisheye Views, 3D-Graphs (Narcissus), ...
- 3D-Techniques: ***Perspective Wall***, Hyperbox, Landscapes, Cone Trees, 3D-Graphs (Narcissus), ...
- Dynamic Techniques:Dyn. Projections, ***Dyn. Linking & Brushing***, Dyn. Environment, ***Dyn. Zooming***, Dyn. Detail on Demand, Dyn. Data-to-Visualization Mapping, ...
- Hybrid Techniques: arbitrary combinations from above

Data Preprocessing Techniques

- **Techniques for Dimension Reduction**
  (Set of d-dim Data Items -> Set of k-dim. Data Items; k << d)
  - Principal Component Analysis [DE 82]
    Determines a minimal set of principal components (linear combinations of the original dimensions) which explain the main variations of the data.
  - Factor Analysis [Har 67]
    Determines a set of unobservable common factors which explain the main variations of the data. The original dimensions are linear combinations of the common factors.
  - Multidimensional Scaling [SRN 72]
    Uses the similarity (or dissimilarity) matrix of the data as defining coordinate axes in multidimensional space. The Euclidean distance in that space is a measure of the similarity of the data items.
  - Fastmap [FL 95]
    Fastmap also operates on a given similarity matrix and iteratively reduces the number of dimensions while preserving the distances as much as possible.

Data Preprocessing Techniques

- **Subsetting Techniques**
  (Set of Data Items -> Subset of Data Items)
  - Sampling (determines a representative subset of the database)
  - Querying (determines a certain, usually a-priori fixed subset of the database)
- **Segmentation Techniques**
  (Set of Data Items -> Set of (Set of Data Items))
  - Segmentation based upon attribute values or attribute ranges
- **Aggregation Techniques**
  (Set of Data Items -> Set of Aggregate Values)
  - Aggregation (sum, count, min, max, ...) based upon
    - attribute values
    - topological properties, etc.
  - Visualizations of Aggregations:
    - Histograms
    - Pie Charts, Bar Charts, Line Graphs, etc.

Geometric Techniques

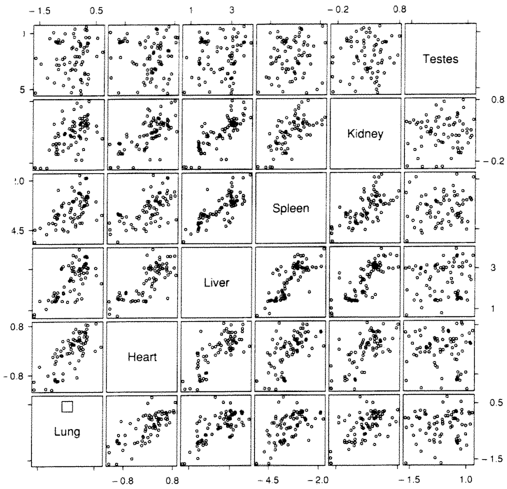
Basic Idea: Visualization of geometric transformations and projections of the data.

Overview

- **Scatterplot-Matrices [And 72, Cle 93]**
- **Projection Pursuit Techniques [Hub 85]**
  (⇨ techniques for finding meaningful projections of multidimensional data)
- **Prosection Views [FB 94, STDS 95]**
- **Hyperslice [WL 93]**
- **Parallel Coordinates [Ins 85, ID 90]**

Geometric Techniques

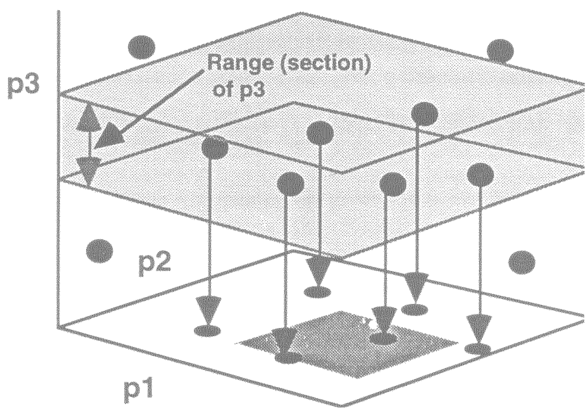
Scatterplot-Matrices [Cle 93]



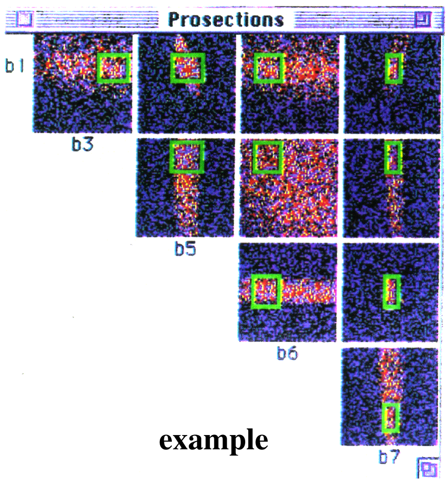
⇨ matrix of scatterplots (x-y-diagrams) of the k-dim. data [total of ($k^2/2$ - k) scatterplots]

Geometric Techniques

Prosection Views [FB 94, STDS 95]



schematic representation

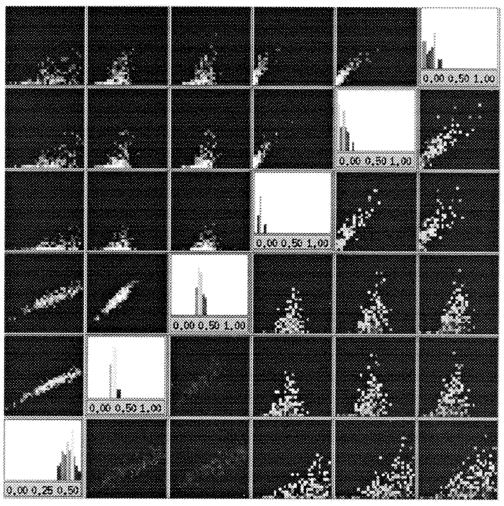


example

⇨ matrix of all orthogonal projections of a certain multidimensional range (combination of selections and projections)

Geometric Techniques

Hyperslice [WL 93]



⇨ matrix of k^2 slices through the k-dim. data (the slices are determined interactively)

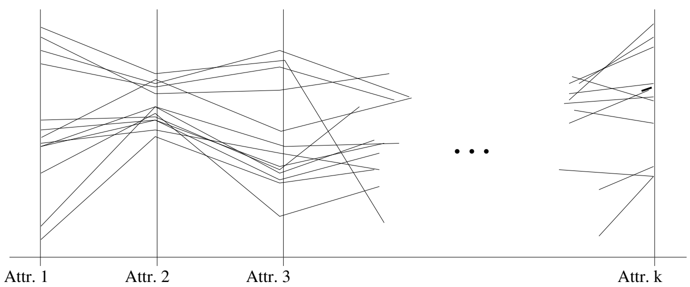
Geometric Techniques

Parallel Coordinates [Ins 85, ID 90]

⇨ n equidistant axes which are parallel to one of the screen axes and correspond to the attributes

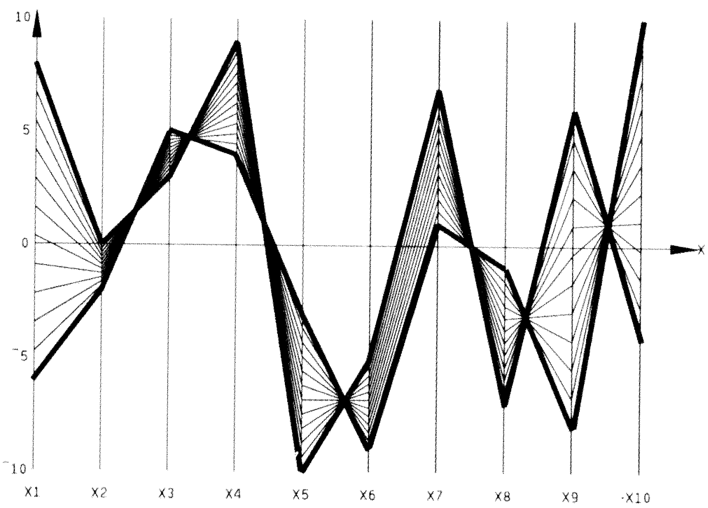
⇨ the axes are scaled to the [minimum, maximum] - range of the corresponding attribute

⇨ every data item corresponds to a polygonal line which intersects each of the axes at the point which corresponds to the value for the attribute



Geometric Techniques

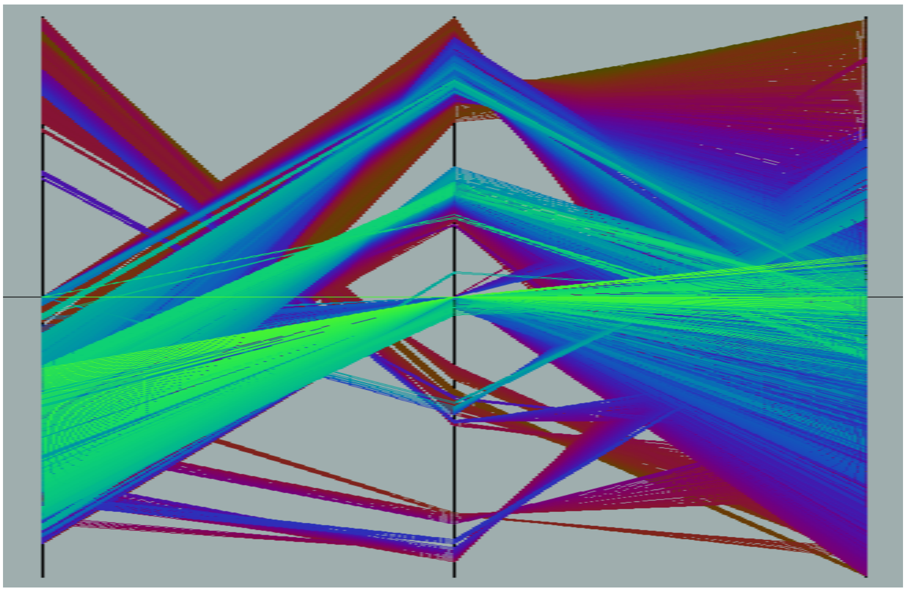
Parallel Coordinates (cont'd)



points on a line in 10-dim. space

Geometric Techniques

Parallel Coordinates (cont'd)



example of 3-dim. data

Icon-based Techniques

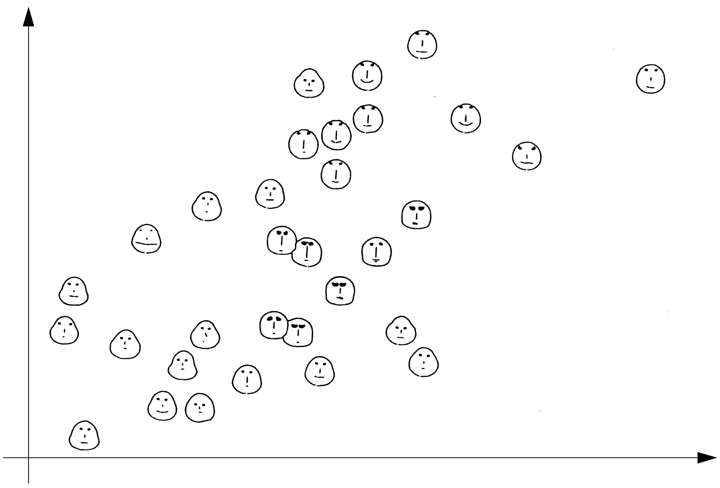
Basic Idea: Visualization of the data values as features of icons.

Overview

- Chernoff-Faces [Che 73, Tuf 83]
- Stick Figures [Pic 70, PG 88]
- Shape Coding [Bed 90]
- Grouping Technique [KK 94]

Icon-based Techniques

Chernoff-Faces [Che 73, Tuf 83]

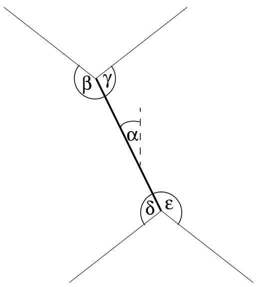


⇨ visualization of the multidim. data using the properties of a face icon (shape of nose, mouth, eyes, and the shape of the face itself)

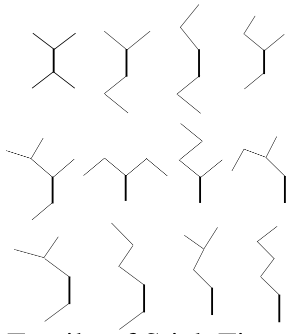
Icon-based Techniques

Stick Figures [Pic 70, PG 88]

- visualization of the multidimensional data using stick figure icons
- two attributes of the data are mapped to the display axes and the remaining attributes are mapped to the angle and/or length of the limbs
- texture patterns in the visualization show data characteristics



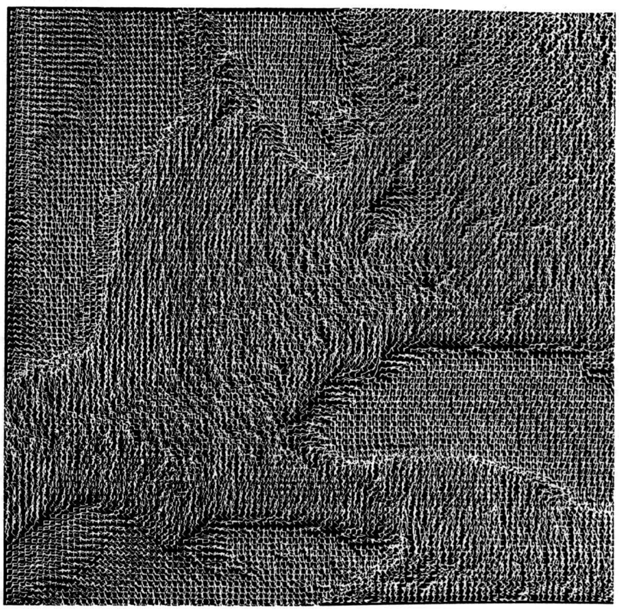
Stick Figure Icon



A Family of Stick Figures

Icon-based Techniques

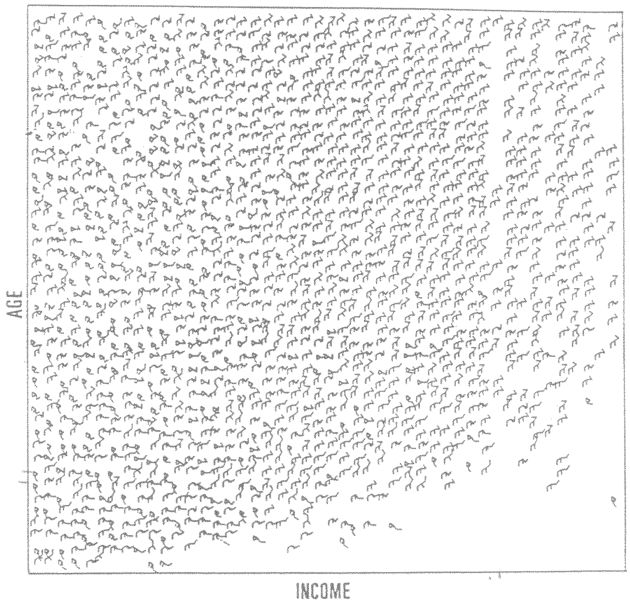
Stick Figures (cont'd)



5-dim. image data from the great lake region

Icon-based Techniques

Stick Figures (cont'd)

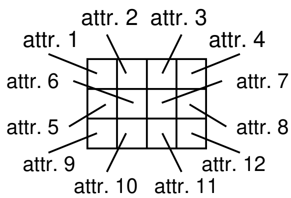


census data showing age, income, sex, education, etc.

Icon-based Techniques

Shape Coding [Bed 90]

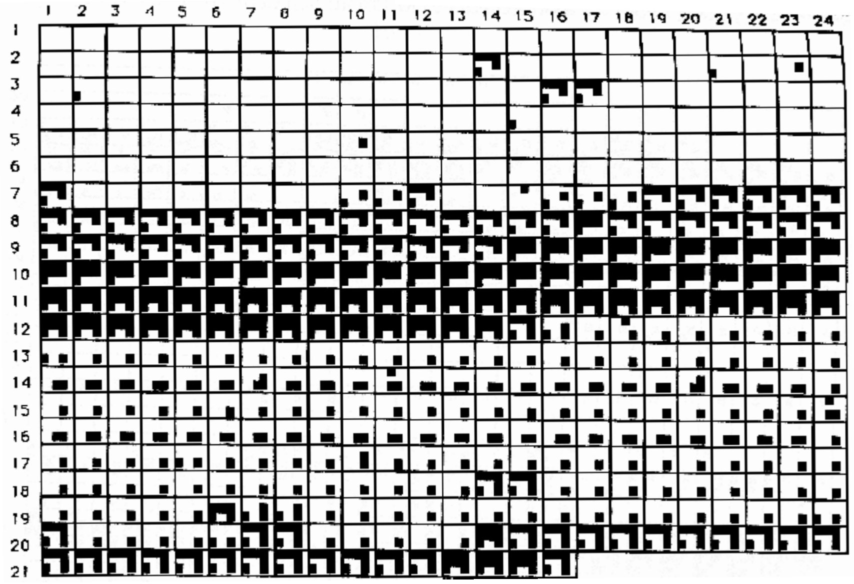
- the data are visualized using small arrays of fields
- each field represents one attribute value
- arrangement of attribute fields (e.g., 12-dimensional data):



- arrays are arranged line-by-line according to a given sorting (e.g., the time attribute for time-series data)

Icon-based Techniques

Shape Coding (cont'd)



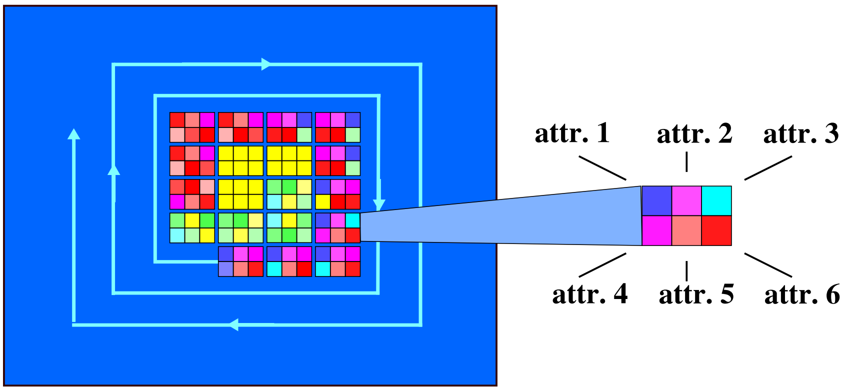
time series of NASA earth observation data

Icon-based Techniques

Grouping Technique [KK 94]

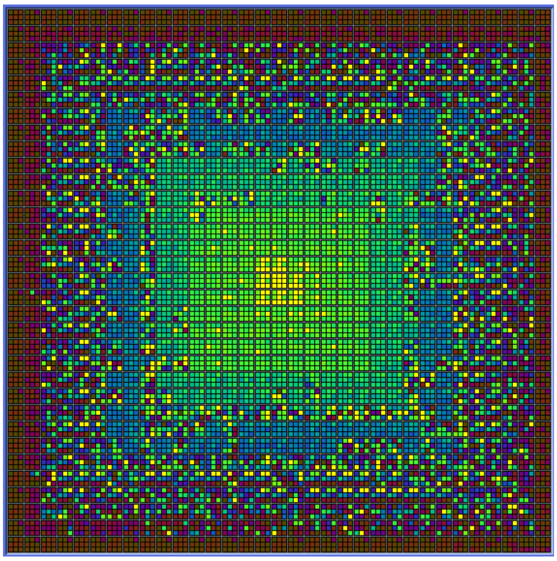
- visualization of the data using color icons
- color icons are array of color fields representing the attribute values
- arrangement is query-dependent (e.g., spiral)

schematic representation of 6-dim. data



Icon-based Techniques

Grouping Technique (cont'd)



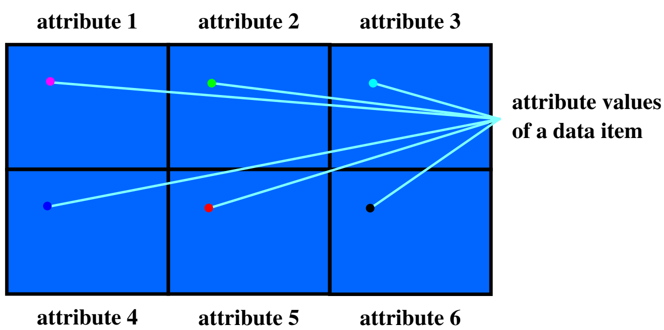
random data containing several clusters

Pixel-oriented Techniques

Basic Idea

- each attribute value is represented by one colored pixel
  (⇨ the value ranges of the attributes are mapped to a fixed colormap)
- the attribute values for each attribute are presented in separate subwindows

visualization of six-dim. data



Pixel-oriented Techniques

Overview

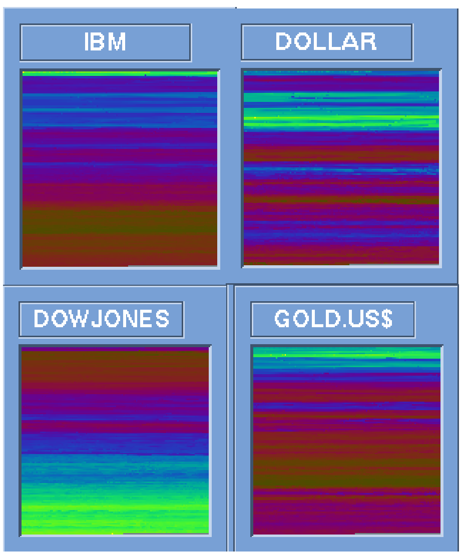
Query-Independent Techniques
- **Simple Techniques**
  - line-by-line
  - column-by-column ...
- **Space-Filling Curves**
  - Peano-Hilbert
  - Morton ...
- **Recursive Pattern Technique**

Query-Dependent Techniques
- **Spiral Technique**
- **Axes Technique**

Pixel-oriented Techniques

Query-Independent Techniques:

Simple Arrangement (Line-by-Line)

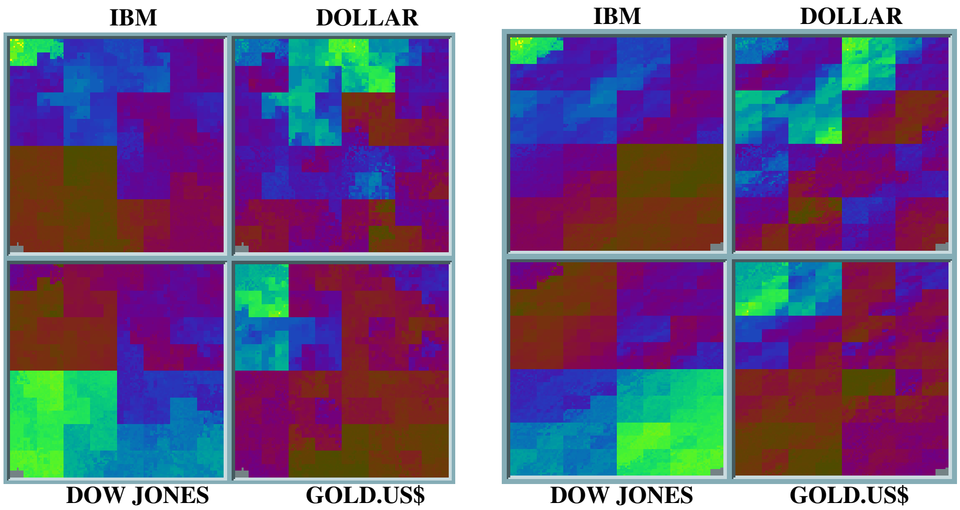


time series of financial data

Pixel-oriented Techniques

Query-Independent Techniques:

Space-Filling Curve Arrangements



Peano-Hilbert   Morton (Z-Curve)

Pixel-oriented Techniques

Query-Independent Techniques:

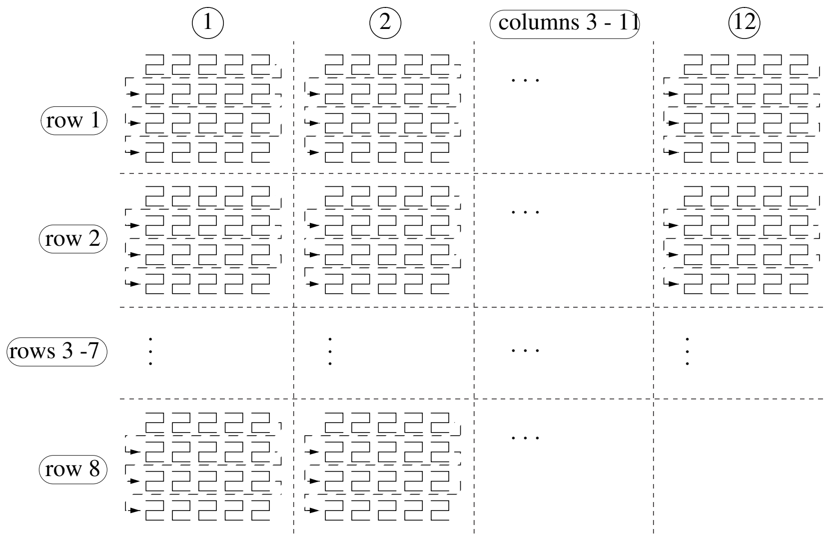
Recursive Pattern Technique

- recursive generalization of an iterated line- and column-based arrangement
- the user may specify the height h_i and width w_i for each recursion level
- on recursion level *i*, w_i patterns of recursion level *(i-1)* are drawn in left-right direction and this is repeated h_i times in top-down direction

=> the pattern on recursion level *i* consists of $w_i \times h_i$ patterns of recursion level *(i-1)*

Pixel-oriented Techniques

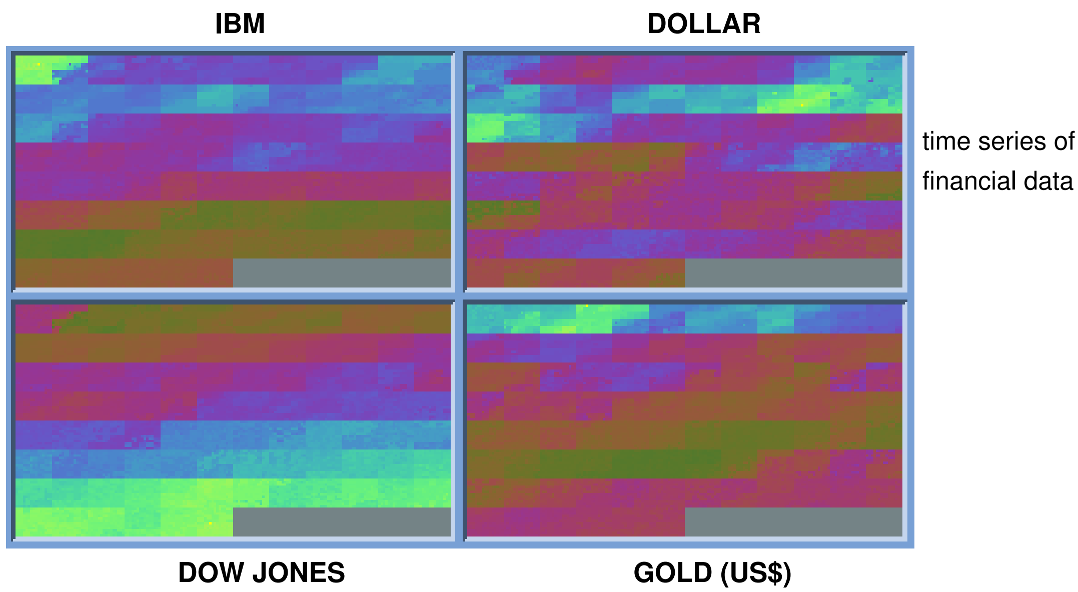
Recursive Pattern: Example of a Structured Arrangement



$(w_1, h_1) = (3, 3), (w_2, h_2) = (5,1), (w_3, h_3) = (1, 4), (w_4, h_4) = (12, 1)$ and $(w_5, h_5) = (1, 8)$

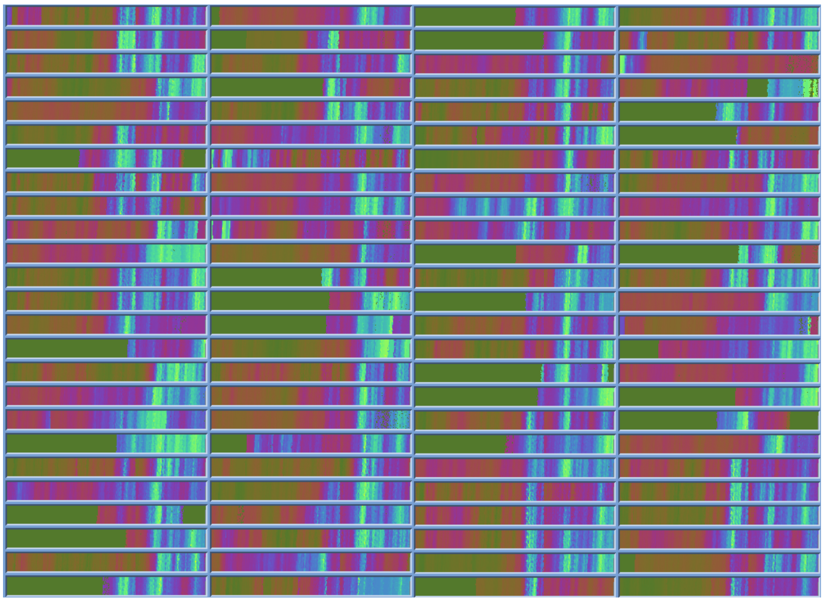
Pixel-oriented Techniques

Recursive Pattern: Example of Financial Data



Pixel-oriented Techniques

Recursive Pattern: FAZ-Index (Jan. '74 - Apr. '95)



time series of stock data

Pixel-oriented Techniques

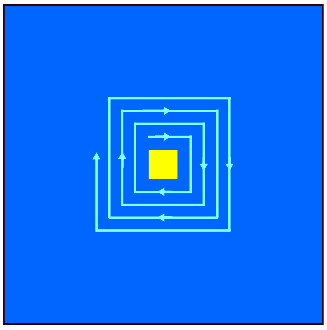
Query-Dependent Techniques: Basic Idea

- data items $(a_1, a_2, ..., a_m)$ & query $(q_1, q_2, ..., q_m)$

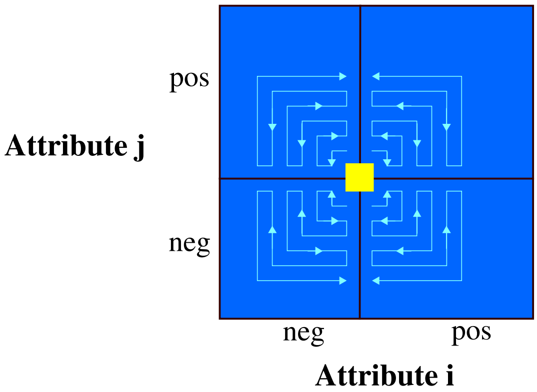
  => distances $(d_1, d_2, ... d_m)$
- extend distances by overall distance (d_{m+1})
- determine data items with lowest overall distances
- map distances to color (for each attribute)
- visualize each distance value d_i by one colored pixel

Pixel-oriented Techniques

Query-Dependent Techniques



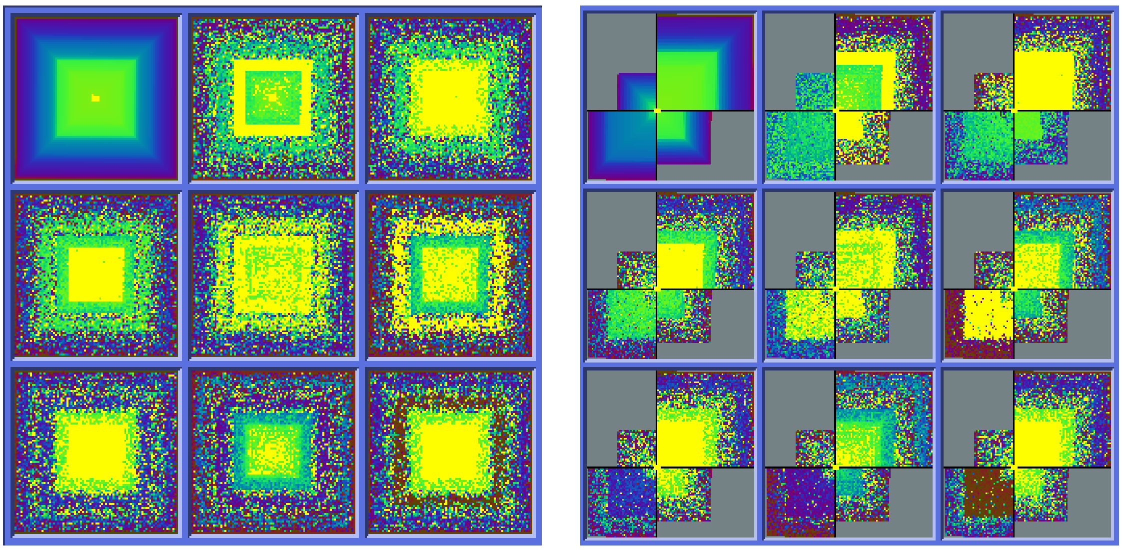
Spiral Technique



Axes Technique

Pixel-oriented Techniques

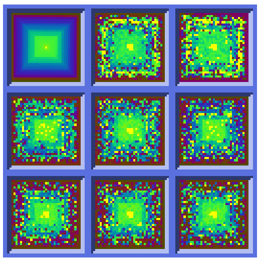
Spiral and Axes Techniques



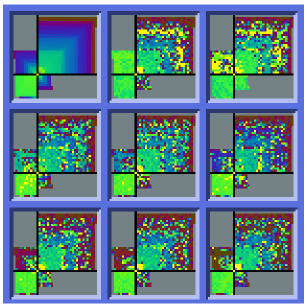
random data containing several clusters

Pixel-oriented Techniques

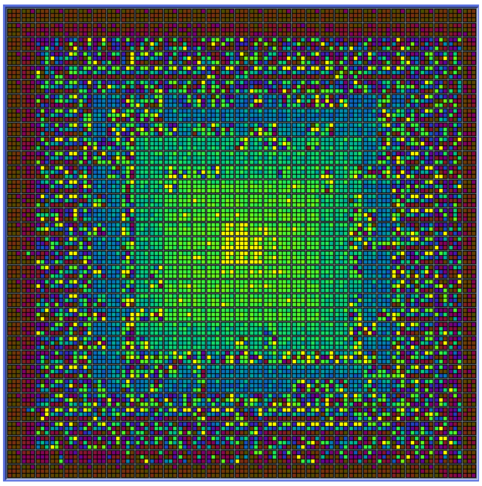
Spiral, Axes, and Grouping Techniques



Spiral Technique



Axes Technique



Grouping Technique

Hierarchical Techniques

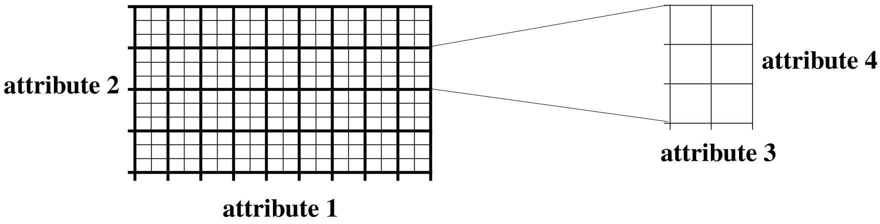
Basic Idea: Visualization of the data using a hierarchical partitioning into two- or three-dimensional subspaces.

Overview

- **Dimensional Stacking [LWW 90]**
- **Hierarchical Plotting [MGTS 90]**
  (⇨ hierarchical drawing of an n-dimensional grid)
- **Worlds-within-Worlds [FB 90]**
- **Treemap [Shn 92, Joh 93]**
- **Cone Trees [RMC 91]**
  (⇨ 3D-Techniques)

Hierarchical Techniques

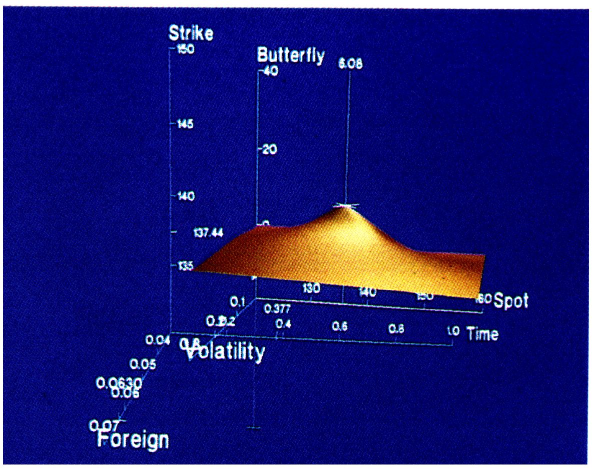
Dimensional Stacking [LWW 90]



- ⇨ partitioning of the n-dimensional attribute space in 2-dimensional subspaces which are 'stacked' into each other
- ⇨ partitioning of the attribute value ranges into classes
- ⇨ the important attributes should be used on the outer levels
- ⇨ adequate especially for data with ordinal attributes of low cardinality

Hierarchical Techniques

Worlds-within-Worlds [FB 90]



visualization of a six-dim. function

- ⇨ partitioning of the n-dim. space into 3-dim. subspaces (e.g., a six-dim. object is displayed by having a new coordinate system for the last three dimensions sit inside the coordinate system for the first three)

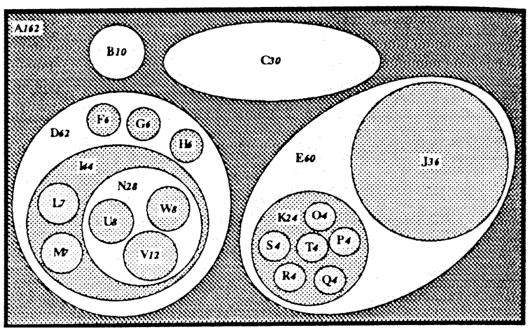
Hierarchical Techniques

Treemap [Shn 92, Joh 93]

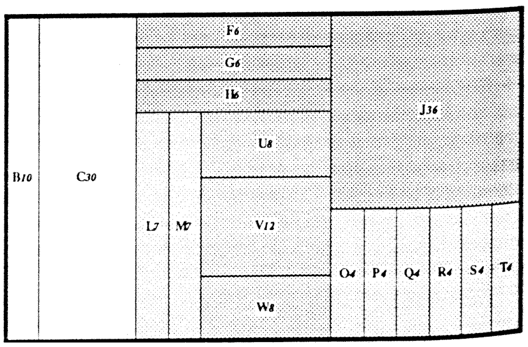
- ⇨ screen-filling method which uses a hierarchical partitioning of the screen into regions depending on the attribute values
- ⇨ the x- and y-dimension of the screen are partitioned alternately according to the attribute values (the attribute value ranges have to be partitioned into classes)
- ⇨ the attributes used for the partitioning and their ordering are user-defined (the most important attributes should be used first)
- ⇨ the color of the final regions may correspond to an additional attribute
- ⇨ suitable to get an overview over large amounts of hierarchical data (e.g., file system) and for data with multiple ordinal attributes (e.g., census data)

Hierarchical Techniques

Treemap (cont'd)



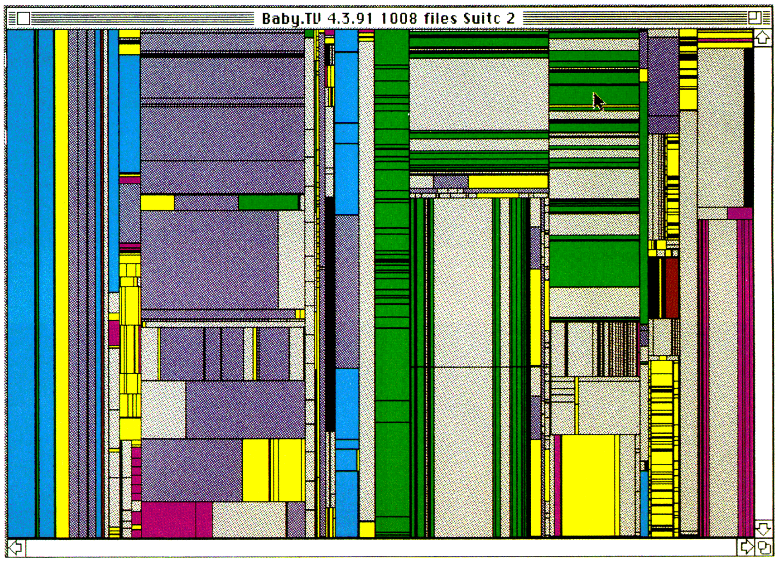
Venn Diagram



Tree-Map

Hierarchical Techniques

Treemap (cont'd)



treemap of a file system containing about 1000 files

Graph-based Techniques

Basic Idea: Visualization of large graphs using techniques to avoid clutter.

Overview

- **SeeNet [EW 93, BEW 95]**
- **Hygraphs [CM 93, Con 94]**
- **Fisheye Views [Fur 86, SB 94]**
- **3D-Graphs (Narcissus [HDWB 95])**
  (⇨ 3D-Techniques)

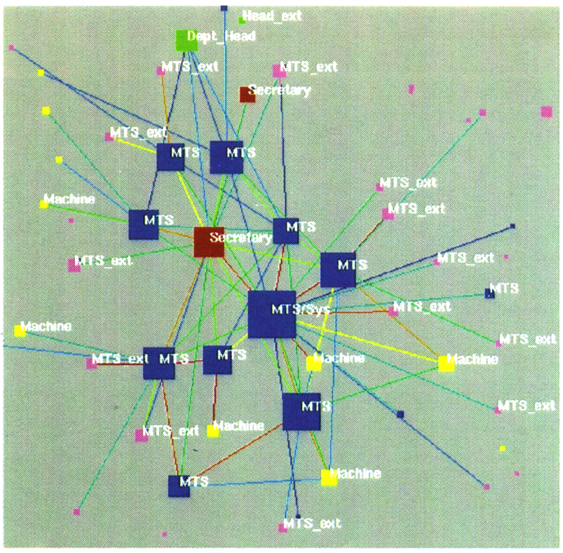
Graph-based Techniques

SeeNet [EW 93, BEW 95]

- ⇨ visualization of hierarchical networks with weighted links
- ⇨ special features of SeeNet:
  - semantic node placement
    (minimizing the distance of nodes with high-weighted links)
  - attributes are mapped to size and color of nodes and links
  - interactivity for
    - changing the mappings
    - expanding or collapsing nodes within the hierarchy
    - getting additional information, etc.
- ⇨ mappings in the example:
  - size of nodes: number of e-mail messages of a person
  - color of nodes: function of staff members
  - size of links: number of e-mail messages of the link
  - color of links: blue for few through green and yellow to red for many messages

Graph-based Techniques

SeeNet (cont'd)



visualization of all
e-mail connections
in a department
over a period of time

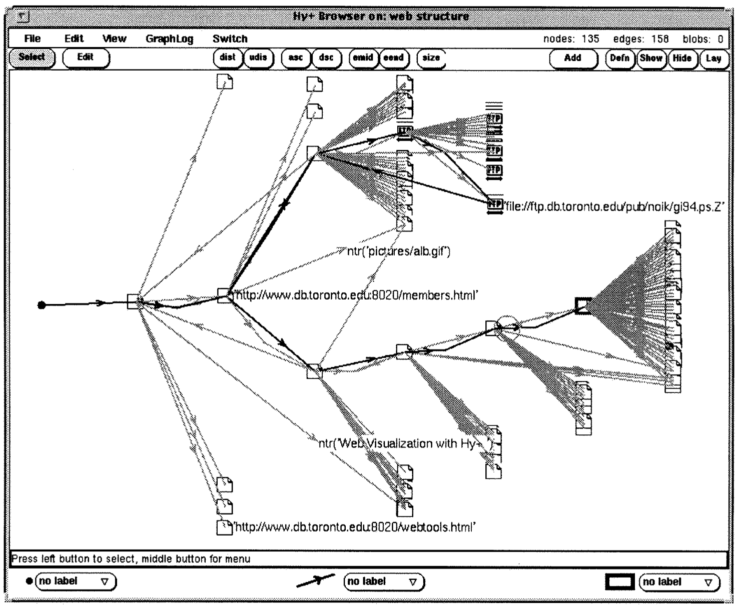
Graph-based Techniques

Hygraphs [CM 93, Con 94]

- Hy+ is an integrated query and visualization system
- Hy+ visualizes structural data as hygraphs
- hygraphs are a generalization of several diagrammatic notations (e.g., labelled graphs are a special case of hygraphs)
- Hy+ queries are described as expressions of the GraphLog query language
- Hy+ queries may be used to define new relationships but may also be used to reduce the data to be visualized
- data may be presented at different levels of detail (e.g., the user may interactively collapse nodes to avoid clutter)

Graph-based Techniques

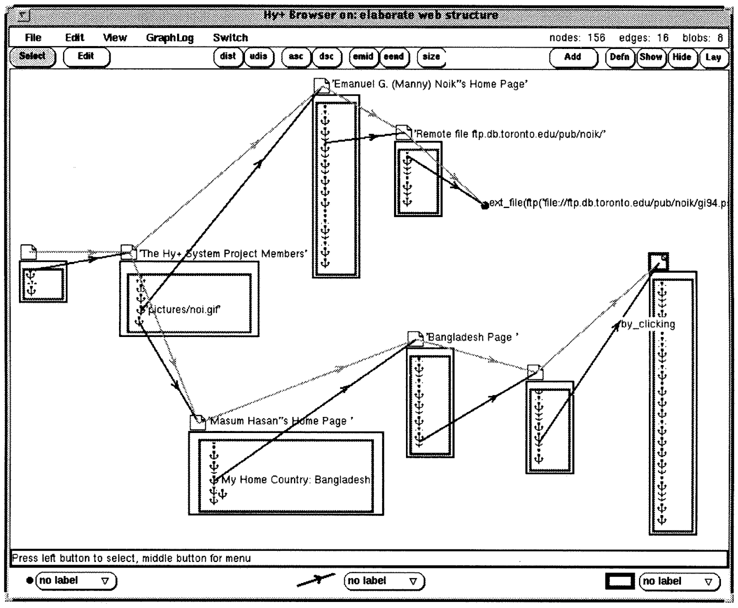
Hygraphs (cont'd)



visualization of a web browsing session

Graph-based Techniques

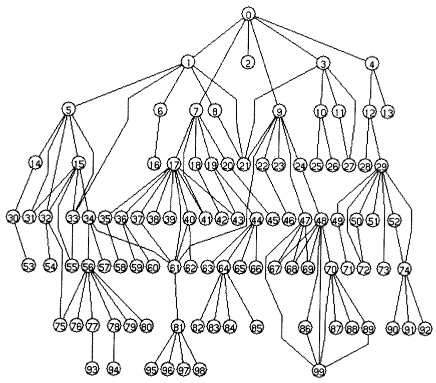
Hygraphs (cont'd)



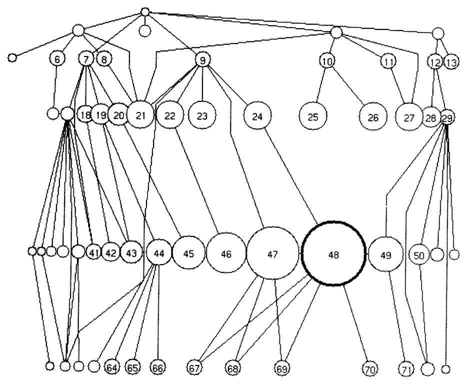
web browsing session with collapsed nodes (called blobs in Hy$^+$)

Graph-based Techniques

Fisheye View [Fur 86, SB 94]



original graph



fisheye view of the graph

- ⇨ graph visualization using a fisheye perspective
- ⇨ shows an area of interest quite large and with detail and the other areas successively smaller and in less detail

3D-Techniques

Basic Idea: Visualization of large data sets using 3D graphics.

Overview

- **Perspective Wall [MRC 91]**
- **Hyperbox [AC 91]**
- **Landscapes [Wri 95]**
- **Cone Trees [RMC 91]**
- **3D-Graphs (Narcissus [HDWB 95])**

3D-Techniques

Perspective Wall [MRC 91]

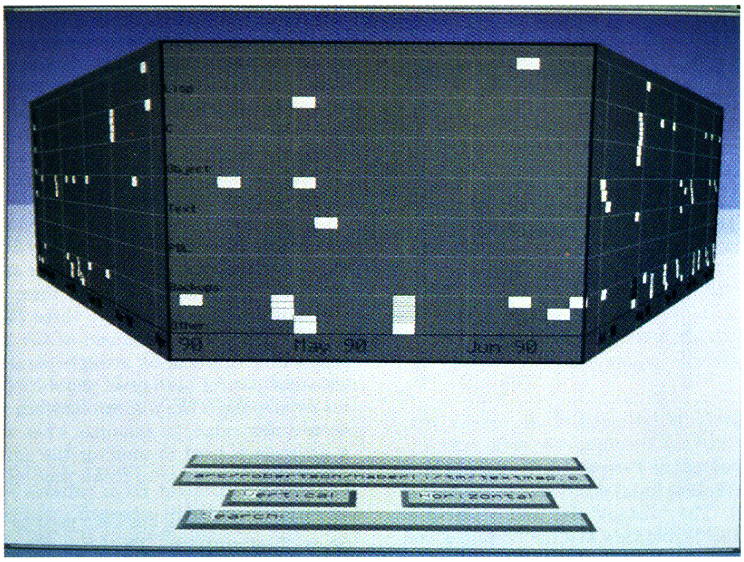
⇨ presentation of the data on a perspective wall

⇨ the data outside the focal area are perspectively reduced in size

⇨ the perspective wall is similar to the fisheye view (⇨ Graph-based Techniques); it is a variant of the bifocal lens display [SA 82] which horizontally compresses the sides of the workspace by direct scaling

3D-Techniques

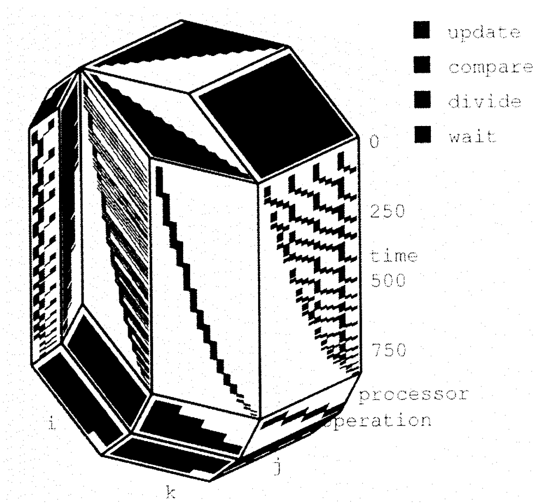
Perspective Wall (cont'd)



a number of documents arranged on a perspective wall

3D-Techniques

Hyperbox [AC 91]

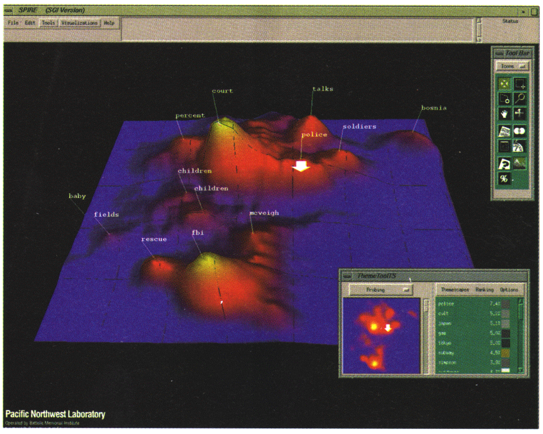


parallel processor performance data visualized as a hyperbox

⇨ mapping of scatterplots onto a hyperbox

3D-Techniques

Landscapes [Wri 95]

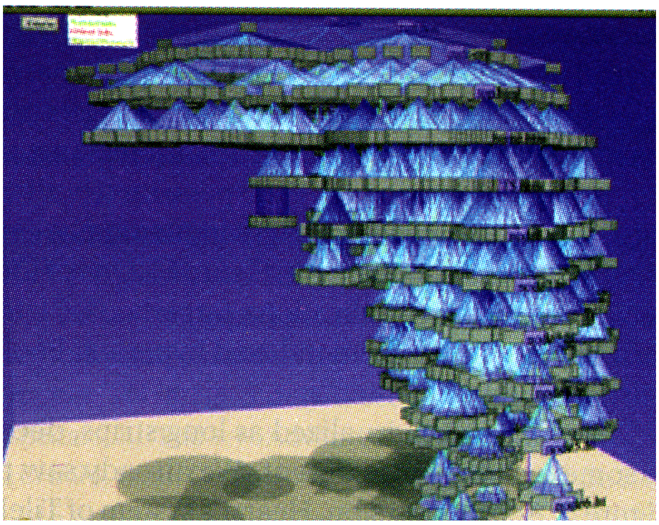


news articles visualized as a landscape

- ⇨ visualization of the data as landscapes
- ⇨ the data needs to be transformed into a spatial representation which preserves the characteristics of the data

3D-Techniques

Cone Trees [RMC 91]

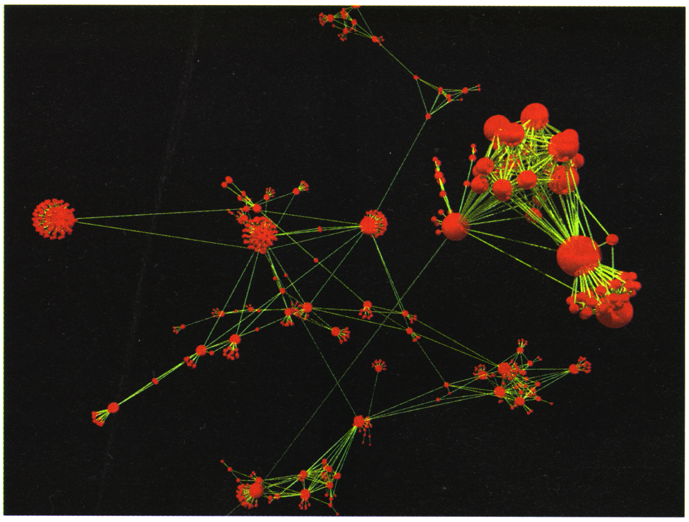


20.000 web pages visualized as a cone tree

- ⇨ animated 3D visualizations of hierarchical data

3D-Techniques

3D-Graphs (Narcissus [HDWB 95])



visualization of
a large number
of web pages

⇨ visualization of complex highly interconnected data (e.g., graphs such as the web)

Dynamic Techniques

Basic Idea: Dynamic interaction with the visualization of the data for a more effective exploration of the data.

Overview

- ❑ **Dynamic Projections**
- ❑ **Dynamic Linking & Brushing**
- ❑ **Dynamic Environment**
- ❑ **Dynamic Zooming**
- ❑ **Dynamic Detail on Demand**
- ❑ **Dynamic Data-to-Visualization Mapping**

Dynamic Techniques

- **Dynamic Projections**
  - ⇨ interactive or automatic variation of the projections
  - ⇨ visualization of the remaining parameters in 2D or 3D
  - ⇨ automatic variation results in an animation of the data
  - ⇨ examples:
    - GrandTour [Asi 85]
    - S Plus [BCW 88]
    - XGobi [SCB 92, BCS 96]

Dynamic Techniques

- **Dynamic Linking & Brushing**
  - ⇨ prerequisite: multiple visualizations of the same data (e.g., visualizations of different projections)
  - ⇨ interactive changes made in one visualization are automatically reflected in the other visualizations
  - ⇨ examples:
    - S Plus [BCW 88]
    - XGobi [SCB 92, BCS 96]
    - DataDesk [Vel 92, WUT 95]

Dynamic Techniques

❑ Dynamic Zooming

- ⇨ visualization of large amounts of data in reduced form to provide an overview of the data
- ⇨ variable zooming of the data with automatic changes of the visualization modes to present more details
- ⇨ examples:
  - PAD++ [Bed 94]
  - IVEE [AW 95]

Dynamic Techniques

❑ Dynamic Detail on Demand

- ⇨ the possibility to interactively obtain more details of the visualized data
- ⇨ details are, for example, the attribute values corresponding to an icon or additional attribute values of a data item
- ⇨ examples:
  - Table Lens [RC 94, PR 96]
  - VisDB [KK 94, KK 95]
  - IVEE [AW 95]

Dynamic Techniques

❑ **Dynamic Environment**

⇨ interactive movements in the visualization space which causes the visualization to change accordingly

⇨ combination of a partial dynamic zooming with a detail on demand strategy for the focal point

⇨ examples:
- Fisheye Views [Fur 86, SB 94]
- Perspective Wall [MRC 91]
- Dyn. Landscapes [LRP 95]

Dynamic Techniques

❑ **Dynamic Data-to-Visualization Mapping**

⇨ dynamic mapping of the data attributes to the parameters of the visualization

⇨ parameters of the visualization are
- x-, y-, and z-axes
- color and size of icons, links, etc.

⇨ examples:
- AutoVisual [BF 93]
- IVEE [AW 95]
- SeeNet [EW 93]
- S Plus [BCW 88]
- XGobi [SCB 92, BCS 96]

Hybrid Techniques

Basic Idea: Integrated use of multiple techniques in one or multiple windows to enhance the expressiveness of the visualizations.

- ⇨ linking diverse visualization techniques may provide additional information
- ⇨ virtually all visualization techniques are combined with dynamics & interactivity

Examples: IVEE [AW 95] uses *Starfield Displays* [AS 94] which are scatterplots of icons with dynamic zooming and mapping
(combination of geometric, icon-based, and dynamic techniques)

XmDv [War 94] allows to dynamically link and brush scatterplot matrices, star icons, parallel coordinates, and dimensional stacking
(combination of geometric, icon-based, hierarchical and dynamic techniques)

Comparison of the Techniques

Criteria for Comparison

comparison of the described data visualization techniques based on their suitability for certain

- ⇨ data characteristics
(e.g., no. of variates, no. of data items, categorical data, ...)
- ⇨ task characteristics
(e.g., clustering, multi variate hot spots, ...)
- ⇨ visualization characteristics
(e.g., visual overlap, learning curve, ...)

Note: The following comparison table expresses my personal opinion obtained from reading the literature and experimenting with several of the described techniques. Many of the ratings are arguable and largely depend on the considered data, the exploration task, experience of the user, etc. In addition, implementations of the techniques in real systems usually avoid the drawbacks of a single technique by combining it with other techniques, which is also not reflected in the ratings.

Comparison of the Techniques

Comparison: A First Attempt

		cluster-ing	multi-variate hot spot	no. of variates	no. of data items	cate-gorical data	visual overlap	learning curve
Geometric Techniques	Scatterplot Matrices	++	++	+	+	-	o	++
	Prosection Views	++	++	+	+	-	o	+
	Hyperslice	+	+	+	+	-	o	o
	Parallel Coordinates	o	++	++	-	o	--	o
Icon-based Techniques	Stick Figure	o	o	+	-	-	-	o
	Shape Coding	o	-	++	+	-	+	-
	Grouping	o	-	++	+	-	+	-
Pixel-oriented Techniques	Query-Independent	+	+	++	++	-	++	+
	Query-Dependent	+	+	++	++	-	++	-
Hierarchical Techniques	Dimensional Stacking	+	+	o	o	++	o	o
	Worlds-within-Worlds	o	o	o	+	o	o	o
	Treemap	+	o	+	o	++	+	o
Graph-based Techniques	SeeNet	+	+	o	+	+	+	+
	Hygraphs	+	+	-	+	+	+	+
	Fisheye Views	+	+	-	+	+	++	+
3D-Techniques	Perspective Wall	+	+	o	+	+	++	+
	Hyperbox	+	+	+	+	o	o	o
	Landscapes	+	+	-	o	o	+	+
	Cone Trees	+	+	o	+	o	+	+
	3D-Graphs (Narcissus)	+	+	-	+	o	+	+

Database Visualization Systems

Overview

- **Statistics-oriented Systems**
- **Visualization-oriented Systems**
- **Database-oriented Systems**
- **Special Purpose Systems**

Database Visualization Systems

- **Statistics-oriented Systems**
  - S Plus [BCW 88]
    (-> generic system for statistical analysis and visualization)
  - XGobi [SCB 92, BCS 96]
    (-> extensible lisp-based system for statistical analysis and visualization)
  - Data Desk [Vel 92, WUT 95]
    (-> commercial system for statistical analysis and visualization;
    features: dyn. linking & brushing of scatterplots and histograms)
  - Diamond (SPSS)
    (-> commercial system for statistical analysis and visualization;
    features: dyn. linking & brushing of scatterplots, parallel coordinates, etc.)
  - DataSpace [ADLP 95]
    (-> 3D-arrangement of a large number of arbitrary visualizations)

Database Visualization Systems

- **Visualization-oriented Systems**
  - ExVis [GPW 89]
    (-> features: stick figure and other icon-based techniques)
  - Parallel Visual Explorer (IBM)
    (-> features: parallel coordinate technique with query-based coloring, etc.)
  - XmDv [War 94, MW 95]
    (-> features: scatterplot matrices, star icons, parallel coordinates,
    dimensional stacking, dynamic linking and brushing)
  - Influence & Attribute Explorer [STDS 95, SDTS 95]
    (-> features: scatterplot and prosection matrices, histograms,
    dynamic linking and brushing)
  - Information Visualizer (Xerox)
    (-> features: diverse information visualization techniques including
    perspective wall, table lens, cone trees)

Database Visualization Systems

❑ Database-oriented Systems

- ⇨ Hy+ [CM 93]
  (-> features: query and visualizations of hygraphs)
- ⇨ TreeViz [Joh 93]
  (-> features: treemap technique)
- ⇨ Table Lens [RC 94, PR 96]
  (-> features: graphical display of tabular data with detail-on-demand capabilities)
- ⇨ VisDB [KK 94, KK 95]
  (-> system for interactive slider-based exploration of very large databases
  features: stick figure, parallel coordinate, and pixel-oriented techniques)
- ⇨ IVEE [AW 95a,b]
  (-> commercial system for database exploration;
  features: generic interactive slider-based visualization environment)

Database Visualization Systems

❑ Special Purpose Systems

- ⇨ Margritte [Vas 94]
  (-> graph visualization for CAD and Software Engineering databases)
- ⇨ Narcissus [HDWB 95]
  (-> system for visualizing complex highly interconnected data such as the web)
- ⇨ WebBook and WebForager [CRY 96]
  (-> systems for accessing and exploring the web)
- ⇨ SeeNet [EW 93] (-> system for visualizing networks with weighted links)
- ⇨ SeeSoft [ESS 92] (-> system for visualizing structured text such as software)
- ⇨ Information Retrieval interfaces such as VIBE [Ols 93]
  (a listing of Information Retrieval Interfaces can be found under
  “http://www-cui.darmstadt.gmd.de/visit/Activities/Viri/visual.html”)
- ⇨ ...

Summary and Conclusions

Summary

- there are a number of recently developed visualization techniques which are applicable to database exploration
- there are different techniques for different types of data (relational tables, hierarchies, graphs, etc.)
- many of the techniques also applicable to traditional relational databases
- there are a number of research prototypes and commercial systems which are available

Summary and Conclusions

Research Issues

- new applications of visualization technology in the database area such as
  Visual Database Exploration for Knowledge Discovery
- development of integrated database visualization and exploration systems
- evaluation and comparison of visualization techniques for database exploration (-> possibilities for improvement)

References

[AC 91] Alpern B., Carter L.: *'Hyperbox',* Visualization '91, San Diego, CA, 1991, pp. 133-139.

[ADLP 95] Anupam V., Dar S., Leibfried T., Petajan E.: *'DataSpace: 3-D Visualization of Large Databases',* Proc. Int. Symp. on Information Visualization, Atlanta, GA, 1995, pp. 82-88.

[And 72] Andrews D. F.: *'Plots of High-Dimensional Data',* Biometrics, Vol. 29, 1972, pp. 125-136.

[AS 94] Ahlberg C., Shneiderman B.: *'Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays'*, Proc. ACM CHI Int. Conf. on Human Factors in Computing (CHI94), Boston, MA, 1994, pp. 313-317.

[Asi 85] Asimov D.: *'The Grand Tour: A Tool For Viewing Multidimensional Data',* SIAM Journal of Science & Stat. Comp., Vol. 6, 1985, pp. 128-143.

[AW 95a] Ahlberg C., Wistrand E.: *'IVEE: An Environment for Automatic Creation of Dynamic Queries Applications',* Proc. ACM CHI Conf. on Human Factors in Computing (CHI95), Demo Program, 1995.

[AW 95b] Ahlberg C., Wistrand E.: *'IVEE: An Information Visualization and Exploration Environment',* Proc. Int. Symp. on Information Visualization, Atlanta, GA, 1995, pp. 66-73.

[Bed 90] Beddow J.: *'Shape Coding of Multidimensional Data on a Mircocomputer Display',* Visualization '90, San Francisco, CA, 1990, pp. 238-246.

[Bed 94] Bederson B.: *'Pad++: Advances in Multiscale Interfaces'*, Proc. ACM CHI Int. Conf. on Human Factors in Computing (CHI94), Boston, MA, 1994, p. 315.

[Ber 81] Bertin J.: *'Graphics and Graphic Information Processing'*, Berlin, 1981.

[BCS 96] Buja A., Swayne D.F., Cook D.: *'Interactive High-Dimensional Data Visualization',* Journal of Computational and Graphical Statistics, Vol. 5, No. 1, 1996, pp. 78-99.

[BCW 88] Becker R., Chambers J. M., Wilks A. R.: *'The New S Language'*, Wadsworth & Brooks/Cole Advanced Books and Software, Pacific Grove, CA, 1988.

References

[BEW 95] Becker R. A., Eick S., Wilks A. R.: *'Visualizing Network Data'*, Trans. on Visualization and Computer Graphics, Vol. 1, No. 1, 1995.

[BF 93] Beshers C., Feiner S.: *'AutoVisual: Rule-Based Design of Interactive Multivariate Visualizations',* IEEE Computer Graphics and Applications, Vol. 13, No. 4, 1993, pp. 41-49.

[BH 86] Bryce D., Hull R.: *'SNAP: A Graphics Based Schema Manager',* Proc. 2nd Int. Conf. on Data Engineering, Los Angeles, CA, 1986.

[BMMS 91] Buja A., McDonald J. A., Michalak J., Stuetzle W.: *'Interactive Data Visualization Using Focusing and Linking'*, Visualization '91, San Diego, CA, 1991, pp. 156-163.

[Che 73] Chernoff H.: *'The Use of Faces to Represent Points in k-Dimensional Space Graphically',* Journal Amer. Statistical Association, Vol. 68, pp. 361-368.

[Cle 93] Cleveland W. S.: *'Visualizing Data',* AT&T Bell Laboratories, Murray Hill, NJ, Hobart Press, Summit NJ, 1993.

[CM 93] Consens M.P., Mendelzon A.O.: *'Hy+: A Hygraph-based Query and Visualization System'*, Proc. ACM SIGMOD Int. Conf. on Management of Data, Washington, DC, 1993.

[Con 94] Consens M.P., Eigler F.Ch., Hasan M.Z., Mendelzon A.O., Noik E.G., Ryman A.G., Vista D.: 'Architecture and Applications of the *Hy+ Visualization System'*, IBM Systems Journal, Vol. 33, No. 3, 1994, pp. 458-476.

[CRY 96] Card S. K., Robertson G. G., York W.: *'The WebBook and the WebForager: An Information Workspace for the World Wide Web'*, Proc. ACM CHI Int. Conf. on Human Factors in Computing (CHI96), 1996.

[Eic 94] Eick S.: *'Data Visualization Sliders'*, Proc. ACM UIST'94, 1994.

[ESS 92] Eick S. G., Steffen J. L., Sumner E. E.: *'SeeSoft - A Tool for Visualizing Software'*, IEEE Trans. on Software Engineering, Vol. 18, No. 11, 1992, pp. 957-968.

[EW 93] Eick S., Wills G.J.: *'Navigating Large Networks with Hierarchies'*, Visualization '93, San Jose, CA, 1993, pp. 204-210.

References

[FB 90] Feiner S., Beshers C.: *'Visualizing n-Dimensional Virtual Worlds with n-Vision'*, Computer Graphics, Vol. 24, No. 2, 1990, pp. 37-38.

[FB 94] Furnas G. W., Buja A.: *'Prosections Views: Dimensional Inference through Sections and Projections',* Journal of Computational and Graphical Statistics, Vol. 3, No. 4, 1994, pp. 323-353.

[FL 95] Faloutsos C., Lin K.: *'Fastmap: A fast Algorithm for Indexing, Data-Mining and Visualization of Traditional and Multimedia Datasets',* Proc. ACM SIGMOD Int. Conf. on Management of Data, San Jose, CA, 1995, pp. 163-174.

[Fur 86] Furnas G.: *'Generalized Fisheye Views',* Proc. ACM CHI Int. Conf. on Human Factors in Software (CHI'86), pp. 16-23, 1986.

[GPW 89] Grinstein G, Pickett R., Williams M. G.: *'EXVIS: An Exploratory Visualization Environment',* Proc. Graphics Interface '89, London, Ontario, Canada, 1989.

[Har 67] Harman H. H.: *'Modern Factor Analysis',* University of Chicago Press, 1967.

[HDWB 95] Hendley R. J., Drew N. S., Wood A. M., Beale R.: *'Narcissus: Visualizing Information',* Proc. Int. Symp. on Information Visualization, Atlanta, GA, 1995, pp. 90-94.

[Hub 85] Huber P. J.: *'Projection Pursuit',* The Annals of Statistics, Vol. 13, No. 2, 1985, pp. 435-474.

[Ins 85] Inselberg A.: *'The Plane with Parallel Coordinates, Special Issue on Computational Geometry',* The Visual Computer, Vol. 1, 1985, pp. 69-97.

[ID 90] Inselberg A., Dimsdale B.: *'Parallel Coordinates: A Tool for Visualizing Multi-Dimensional Geometry'*, Visualization '90, San Francisco, CA, 1990, pp. 361-370.

[Joh 93] Johnson B.: *'Visualizing Hierarchical and Categorical Data'*, Ph.D. Thesis, Department of Computer Science, University of Maryland, 1993.

[KK 94] Keim D. A., Kriegel H.-P.: *'VisDB: Database Exploration using Multidimensional Visualization',* Computer Graphics & Applications, Sept. 1994, pp. 40-49.

References

[KK 95] Keim D. A., Kriegel H.-P.: *'VisDB: A System for Visualizing Large Databases',* System Demonstration, Proc. ACM SIGMOD Int. Conf. on Management of Data, San Jose, CA, 1995, p. 482.

[KKA 95] Keim D. A., Kriegel H.-P., Ankerst M.: *'Recursive Pattern: A Technique for Visualizing Very Large Amounts of Data',* Proc. Visualization '95, Atlanta, GA, 1995, pp. 279-286.

[KKS 94] Keim D. A., Kriegel H.-P., Seidl T.: *'Supporting Data Mining of Large Databases by Visual Feedback Queries',* Proc. 10th Int. Conf. on Data Engineering, Houston, TX, 1994, pp. 302-313.

[KL 93] Keim D. A., Lum V.: *'GRADI: A Graphical Database Interface for a Multimedia DBMS',* Proc. Int. Workshop on Interfaces to Databases, Glasgow, England, 1992, in: Workshops in Computing, Springer, 1993, pp. 95-112.

[LRP 95] Lamping J., Rao R., Pirolli P.: *'A Focus + Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies',* Proc. ACM CHI Conf. on Human Factors in Computing (CHI95), 1995, pp. 401-408.

[LWW 90] LeBlanc J., Ward M. O., Wittels N.: *'Exploring N-Dimensional Databases',* Visualization '90, San Francisco, CA, 1990, pp. 230-239.

[MGTS 90] Mihalisin T., Gawlinski E., Timlin J., Schwendler J.: *'Visualizing A Scalar Field on an N-dimensional Lattice',* Visualization '90, San Francisco, CA, 1990, pp. 255-262.

[MRC 91] Mackinlay J.D., Robertson G.G., Card S.K.: *'The Perspective Wall: Detail and Context Smoothly Integrated'*, Proc. ACM CHI Int. Conf. on Human Factors in Computing (CHI'91), 1991.

[MW 95] Martin A. R., Ward M. O.: *'High Dimensional Brushing for Interactive Exploration of Multivariate Data',* Visualization '95, Atlanta, GA, 1995, pp. 271-278.

[Ols 93] Olsen K.A. et al.: *'Visualization of a Document Collection: The VIBE System',* Information Processing and Management, Vol. 29, No. 1, 1993, pp. 69-81.

References

[OW 93] Ozoyoglu G., Wang H.: *'Example-Based Graphical Database Query Languages'*, Computer, Vol. 26, No. 5, 1993, pp. 25-38.

[PG 88] Pickett R. M., Grinstein G. G.: *'Iconographic Displays for Visualizing Multidimensional Data',* Proc. IEEE Conf. on Systems, Man and Cybernetics, IEEE Press, Piscataway, NJ, 1988, pp. 514-519.

[Pic 70] Pickett R. M.: *'Visual Analyses of Texture in the Detection and Recognition of Objects',* in: Picture Processing and Psycho-Pictorics, Lipkin B. S., Rosenfeld A. (eds.), Academic Press, New York, 1970.

[PR 96] Pirolli P., Rao R.: *'Table Lens as a Tool for Making Sense of Data'*, Proc. Conf. on Advanced Visual Interfaces (AVI-96), Gubbio, Italy, 1996.

[RC 94] Rao R., Card S.K.: *'The Table Lens: Merging Graphical and Symbolic Representation in an Interactive Focus+Context Visualization for Tabular Information'*, Proc. ACM CHI Int. Conf. on Human Factors in Computing (CHI'94), 1994, pp. 318-322.

[RMC 91] Robertson G., Mackinlay J., Card S.: *'Cone Trees: Animated 3D Visualizations of Hierarchical Information',* Proc. ACM CHI Int. Conf. on Human Factors in Computing (CHI'91), pp. 189-194.

[SA 82] Spence R., Apperley M.: *'Data Base Navigation: An Office Environment for the Professional',* Behaviour and Information Technology, Vol. 1, No. 1, pp. 43-54.

[SB 94] Sarkar M., Brown M.: *'Graphical Fisheye Views',* Communications of the ACM, Vol. 37, No. 12, 1994, pp. 73-84.

[SCB 92] Swayne D.F., Cook D., Buja A.: *'User's Manual for XGobi: A Dynamic Graphics Program for Data Analysis',* Bellcore Technical Memorandum, 1992.

[SDTS 95] Su H., Dawkes H., Tweedie L., Spence R.: *'An Interactive Visualization Tool for Tolerance Design',* Technical Report, Imperial College, London, 1995.

References

[Shn 92] Shneiderman B.: *'Tree Visualization with Treemaps: A 2D Space-Filling Approach',* ACM Trans. on Graphics, Vol. 11, No. 1, pp. 92-99, 1992.

[Spo 93] Spoerri A.: *'InfoCrystal: A Visual Tool for Information Retrieval',* Visualization '93, San Jose, CA, 1993, pp. 150-157.

[STDS 95] Spence R., Tweedie L., Dawkes H., Su H.: *'Visualization for Functional Design',* Proc. Int. Symp. on Information Visualization (InfoVis '95), Atlanta, GA, 1995, pp.4-10.

[SRN 72] Shepard R. N., Romney A. K., Nerlove S. B.: *'Multidimensional Scaling',* Seminar Press, New York, 1972.

[Tuf 83] Tufte E. R.: *'The Visual Display of Quantitative Information',* Graphics Press, Cheshire, CT, 1983.

[Tuf 90] Tufte E. R.: *'Envisioning Information',* Graphics Press, Cheshire, CT, 1990.

[Vas 94] Vasudevan V.: *'Supporting High Bandwidth Navigation in Object-Bases',* Proc. 10th Int. Conf. on Data Engineering, Houston, TX, 1994, pp. 294-301.

[Vel 92] Velleman P. F: *'Data Desk 4.2: Data Description'*, Ithaca, NY, 1992.

[War 94] Ward M. O.: *'XmdvTool: Integrating Multiple Methods for Visualizing Multivariate Data',* Visualization'94, Washington, DC, 1994, pp. 326-336.

[WB 95] Wong P.C., Bergeron R. D.: *'30 Years of Multidimensional Multivariate Visualization',* Proc. Workshop on Scientific Visualization, IEEE Computer Society Press, 1995

[WL 93] van Wijk J. J., van Liere R.. D.: *'Hyperslice',* Visualization '93, San Jose, CA, 1993, pp. 119-125.

[Wri 95] Wright W.: *'Information Animation Applications in the Capital Markets'*, Proc. Int. Symp. on Information Visualization, Atlanta, GA, 1995, pp. 19-25.

[WUT 95] Wilhelm A., Unwin A.R., Theus M.: *'Software for Interactive Statistical Graphics - A Review'*, Proc. Int. Softstat '95 Conf., Heidelberg, Germany, 1995.

[Zlo 77] Zloof M. M.: *'Query-By-Example: A Data Base Language',* IBM Systems Journal, Vol. 4, 1977, pp. 324-343.