Sensor Signal Processing

Prof. Dr.-Ing. Andreas König
Lehrstuhl Integrierte Sensorsysteme

FB Elektrotechnik und Informationstechnik
Technische Universität Kaiserslautern

Fall Semester 2005

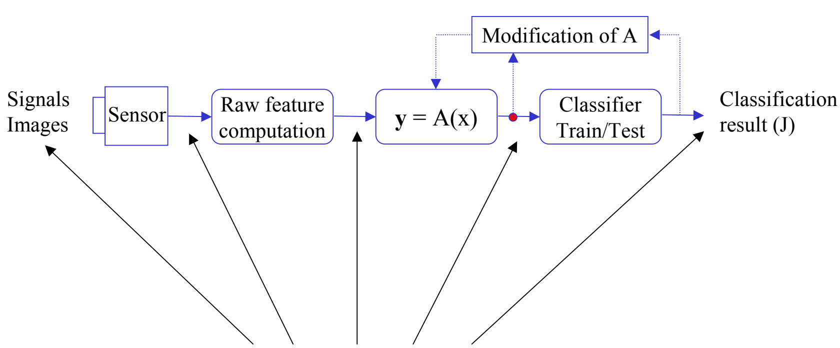
Course Contents

Chapter Contents

Sensor Signal Processing
Data Visualization

Motivation

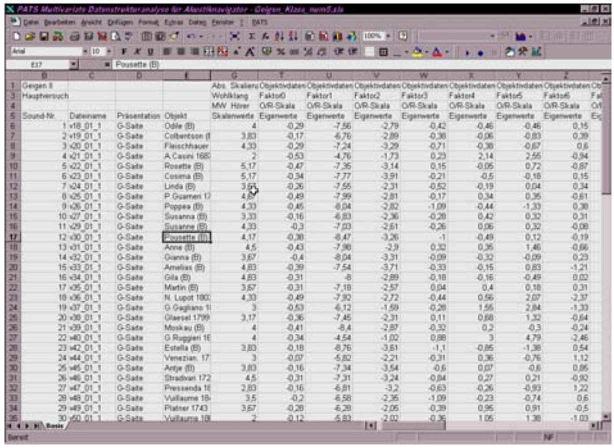
- Design decisions at multiple steps of system design require insight into the potentially complex nature of underlying data:



Human observation & understanding
of high-dimensional, large data

Motivation

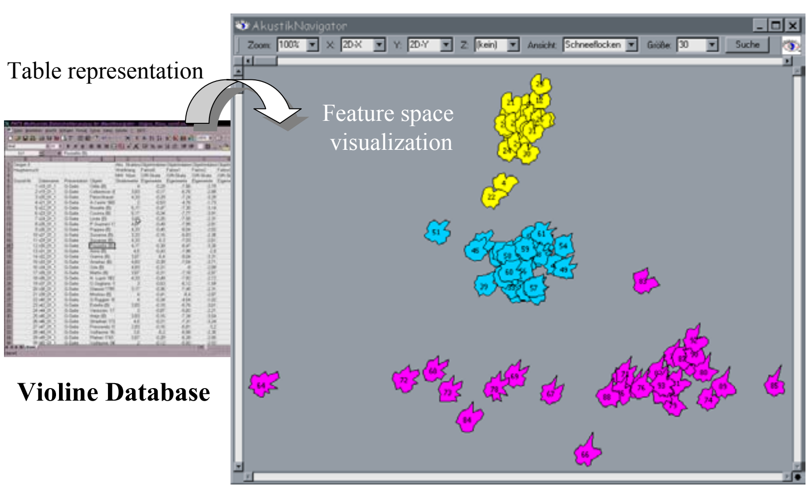
- Representation must be transparent, intuitive, and ergonomic:



- **This representation** is not well adapted to human structure perception capabilities

Motivation

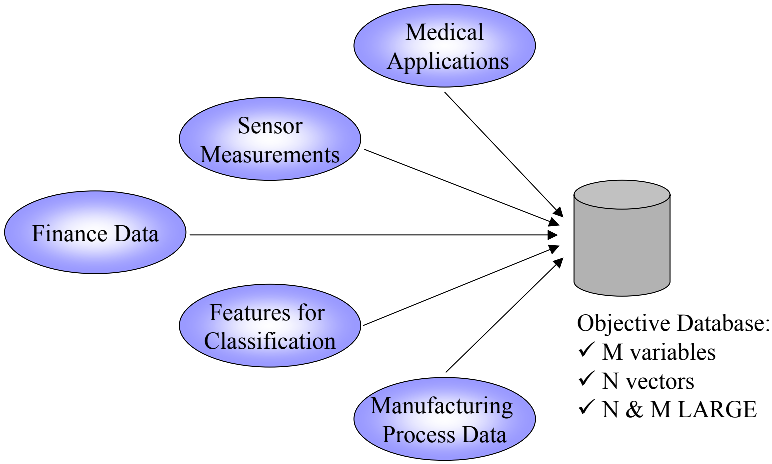
- Better approach:



Violine Database

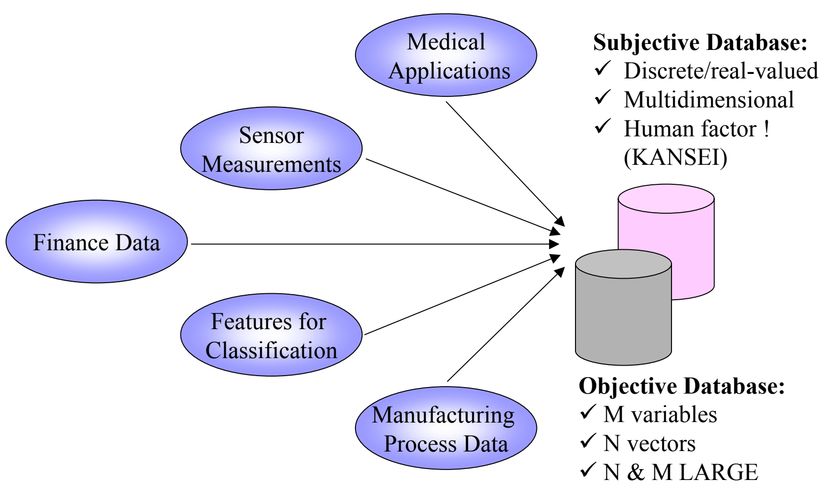
Motivation

- Relation to general data analysis:
- Ubiquitousness of the dimensionality reduction problem



Motivation

- Inclusion of additional attribute information

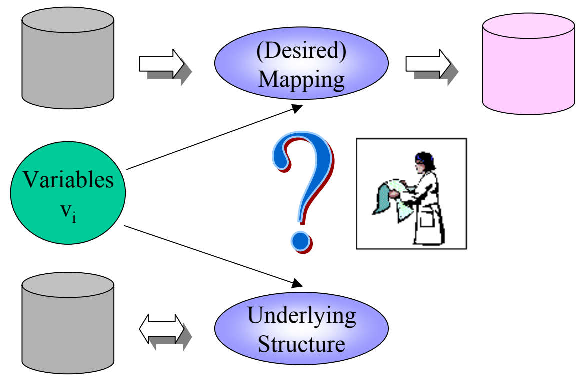


Motivation

➢ Finding the Relevant Information – Choice of Approach

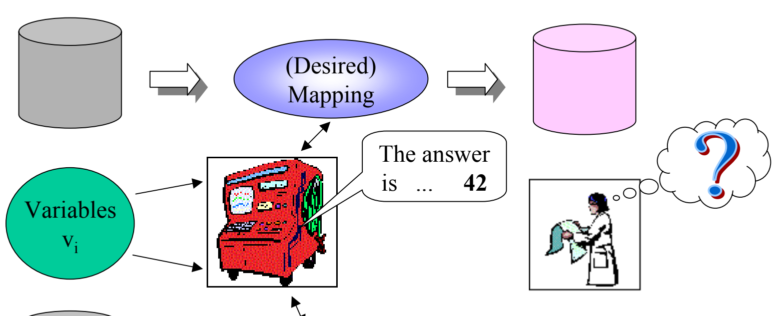
Means to determine salient variables v_i that are

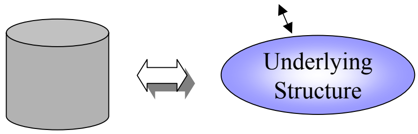
- significant
- redundant
- irrelevant
- correlated



Motivation

➢ Finding the Relevant Information – Automatic Machine-Centered Approach



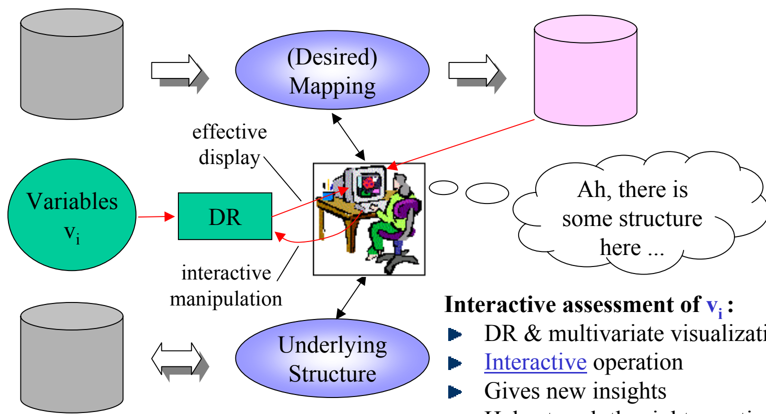


Automatic assessment of v_i :

- Statistical approaches
- Clustering
- Rule extraction

Motivation

- Finding the Relevant Information – Interactive Human-Centered Approach



Interactive assessment of v_i :

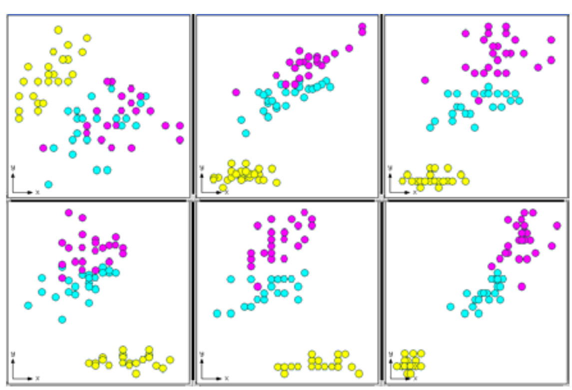
- DR & multivariate visualization
- Interactive operation
- Gives new insights
- Helps to ask the right questions ...
- Effective > 3 decades (OLPARS)

Interactive Visualization

- Dimensionality reduction and visualization by manual feature selection
- Visualization by multiple (linked) pairwise plots, six for Iris data:

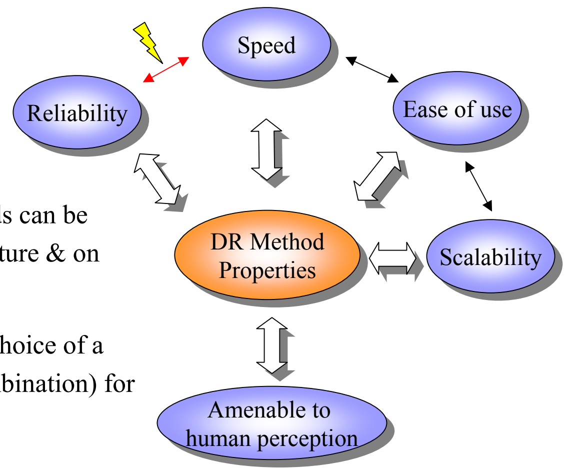
- Sideviews of hyper-cube
- Views grow with
  $$\frac{M(M-1)}{2}$$
- Limited applicability
- Gets intractable for data with $M >> 1$
- A priori knowledge required

- **Better:** Automatic feature selection or projection, e.g., NLM



Interactive Visualization

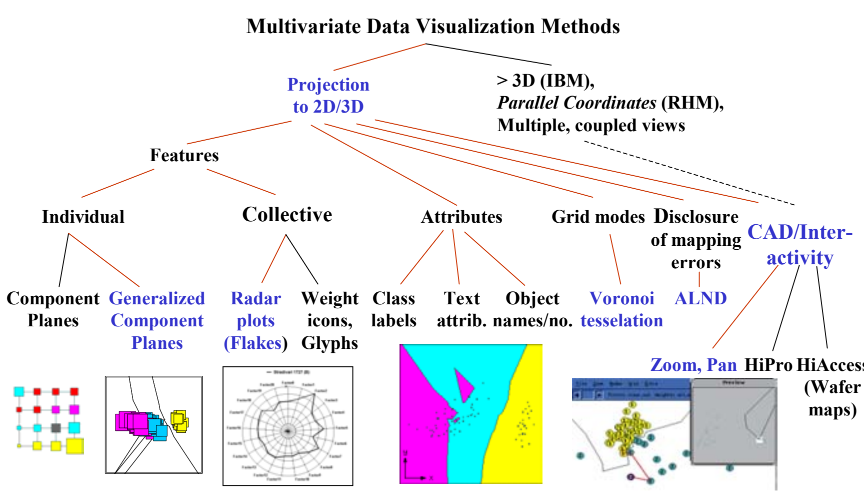
- Applicability of dimensionality reduction methods for visualization:



- A plethora of methods can be observed in the literature & on the tool market
- What motivates the choice of a method (method combination) for application ?

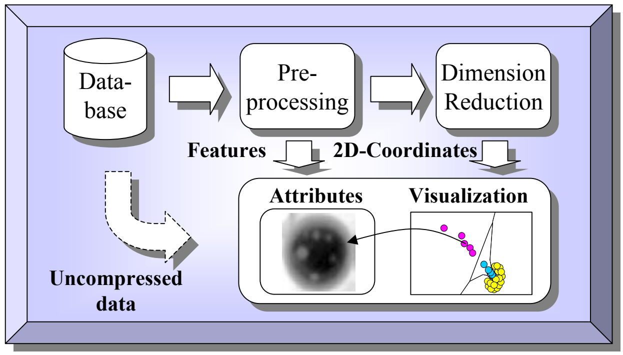
Interactive Visualization

- Taxonomy of Visualization Techniques:



Interactive Visualization

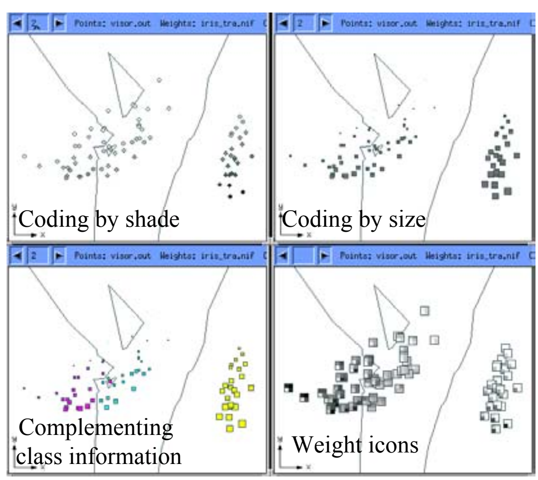
- Interactive Data Visualization and Explorative Analysis Architecture



- Visualization standard: Static (2D-) scatter plot
- Advanced Visualization by enhanced user interaction & CAD-functionality

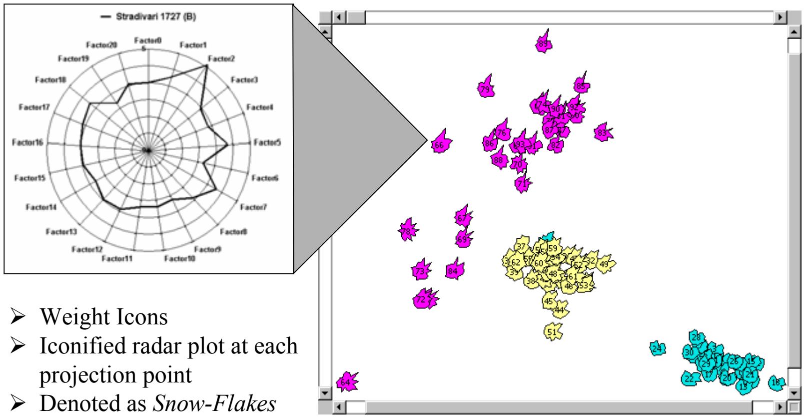
Component Planes

- Generalization from SOM component planes
- The values of one feature are displayed at the projection points
- Like Hinton diagrams
- Several planes can be displayed and analysed
- Correlating and salient features can be detected
- Multicomponent plots exploit human texture perception for analysis



Component Planes

- Extension to Multi-component plot based on radar-plot concept:



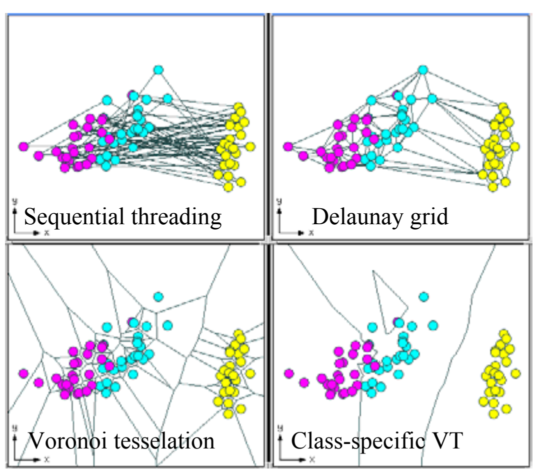
- Weight Icons
- Iconified radar plot at each projection point
- Denoted as *Snow-Flakes*

Sensor Signal Processing
Data Visualization

Grid Modes

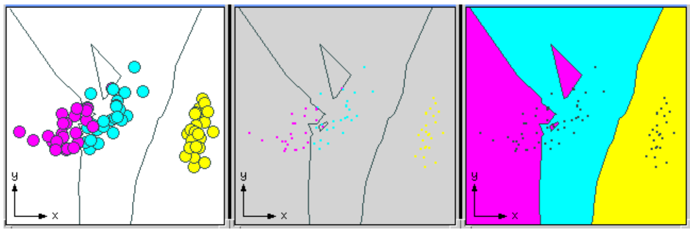
- Different grid modes for auxiliary (approximated) boundary information:

- Grid modes comple-menting the SOM mesh
- Sequence order of data vectors can be displayed
- Voronoi tesselation give exact quantization and 1-NN classification boundaries for 2D-data
- An adequate idea is given for projection of multidimensional data



Grid Modes

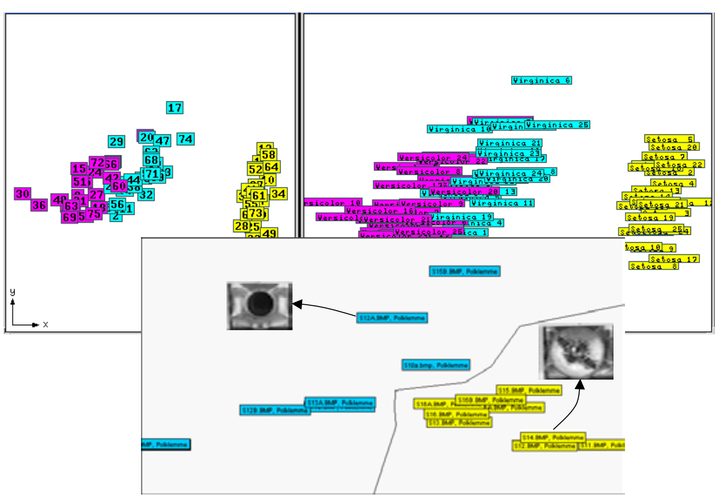
- Detection of Outliers and Identical Vectors by Visualization
- Outliers and identical vectors are pathological cases for system design
- At high point densities mutual occlusion of points becomes a problem



- Variable-sized fields and colored class-specific VT can alleviate this problem
- Outliers and isolated vectors in different class regions can easily be identified
- Identical vectors are marked by an hexagonal shape
- Stacking by mouse click rotates identical or strongly occluding vectors
- **Fast troubleshooting:** Pathological vectors can be tracked-back in the database

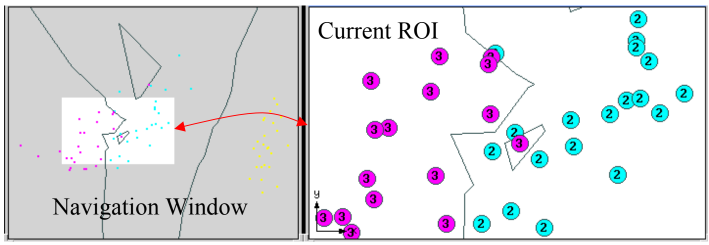
Attributes & Hierarchical Access

- Attribute Display and Hierachical Database Access



Interactive Navigation

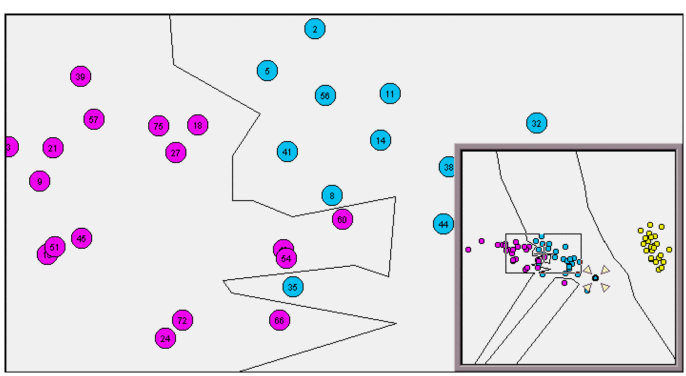
- Interactive Navigation and Hierarchical Analysis
- The methods described so far alleviate first insight in the data structure
- For very large data sets with considerable local density variations, analysis on a single scale will not be feasible
- Global to local analysis across scales (zoom/pan) offers superior performance:



- Focus of interest can conveniently be shifted across scales by zooming in or out and in the same scale by shifting the navigation window
- Descending down to local detail, mapping quality and errors become an issue

Interactive Navigation

- Different implementation in QuickCog with pop-up navigation window and drag&drop navigation

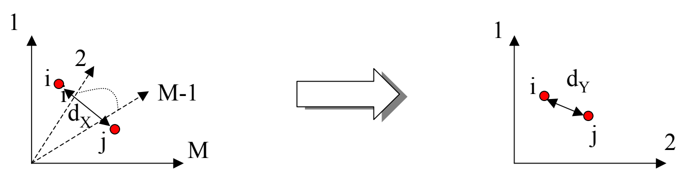


- Scale is preset from menu entry or interactively by mouse operation

Mapping Reliability

- Assessment of Mapping Quality and Mapping Errors
- With increasing intrinsic dimension, all mapping methods will introduce mapping errors
- Visualization results may be subject to twists and distortions that can lead to invalid conclusion drawing & corrupt knowledge aquisition
- How can mapping quality be measured and mapping faults be disclosed ?
- **First proposed measure:** Distance preservation E with

$$E = \frac{1}{c}\sum_{j=1}^{N}\sum_{i=1}^{j}\frac{(d_{X_{ij}} - d_{Y_{ij}})^2}{d_{X_{ij}}} \qquad (6.1)$$



Mapping Reliability

- **Second proposed measure:** Topology preservation q_m with

$$q_m = \frac{1}{3n * N}\sum_{i=1}^{N} q_{m_i} \qquad (6.2)$$

- Local quality q_{mi} is computed by determining the n-nearest-neighbors NN_{ji} of the ith pattern $(i \in [1,n], j \in [1,N])$
- Rank order of the nearest-neighbors in **X** and **Y** are compared and q_{mi} values are assigned according to the following simple assignment scheme:

$q_{mi} = 3$ if NN_{ji} in $\mathbf{X} = NN_{ji}$ in $\mathbf{Y}$
$q_{mi} = 2$ if NN_{ji} in $\mathbf{X} = NN_{jl}$ in $\mathbf{Y}$ with $l \in [1,n]; j \neq i$
$q_{mi} = 1$ if NN_{ji} in $\mathbf{X} = NN_{jt}$ in $\mathbf{Y}$ with $t \in [n,m]; n < m$
$q_{mi} = 0$ else

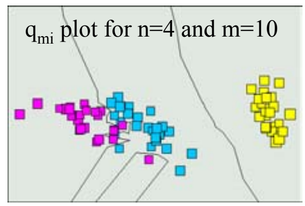
- The topology preservation measure q_m returns 1.0 for perfect preservation

Mapping Reliability

- For typical parameter settings n=4 and m=10 an application example for ***Iristrain*** and ***Mech$_1$*** data is given:

Data set	NLM		Visor	
	q_m	E	q_m	E
Iristrain	0.6667	0.0098	0.6711	0.0093
Mech$_1$	0.4527	0.0933	0.4207	0.1628

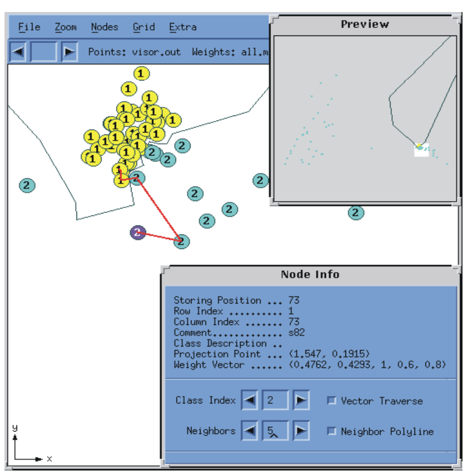
- In addition to the global measures E and q_m, local information for each point in the projection can be computed by local distance preservation E_i and local topology preservation q_{mi}
- These can be superimposed on the mapping:



- Local mapping faults can be disclosed
- $O(N^2)$ complexity of the quality measure !
- Consumes more time than the mapping itself
- **Remedy:** Fast interactive local alternative

Mapping Reliability

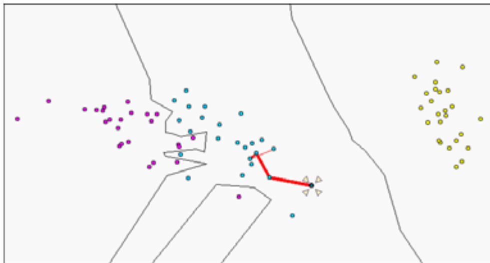
- Validation of mapping quality by actual local neighborhood display



- The actual nearest neighbors in **X** are computed for an interactively chosen pattern
- The projection points corresponding to these NN in **Y** are threaded by a red line in the visualization
- This discloses mapping errors at low computational cost for valid conclusion drawing and knowledge acquisition
- Similar to q_{mi}
- Compatible with navigation & hierarchical analysis

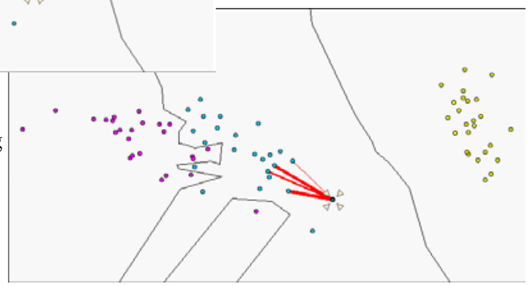
Mapping Reliability

- Alternative ALND implementations:



- Diminishing linewidth with growing distance from reference point improves the visualization

- Instead of point threading by a polyline, a pairwise connection can be beneficial & more lucid



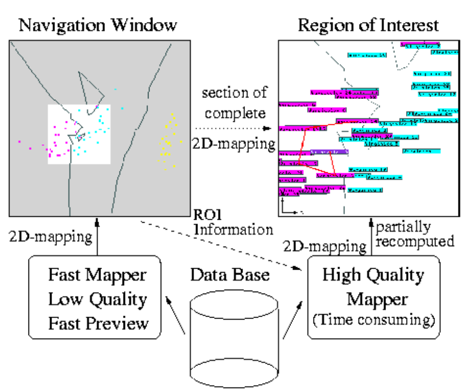
Hierarchical Visualization

- None of the presented methods will return in general a satisfying global and local mapping quality
- Three main strategies for improving mapping quality can be observed:

1) Hierachical clustering and mapping of the data

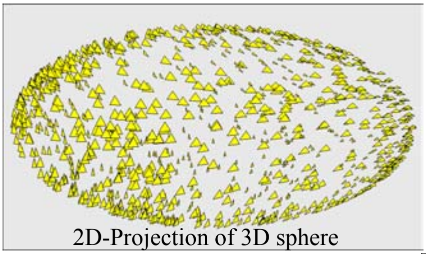
2) Hierarchical mapping & navigation approach:

3) Global to local shift of the mapping stress implemented by a limited or shrinking neighborhood during the mapping computation

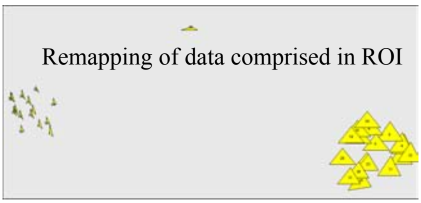


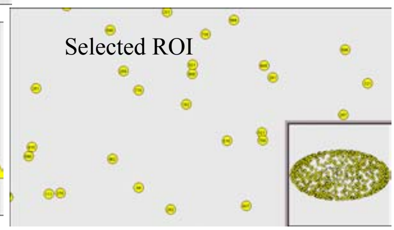
Hierarchical Visualization

- Advanced Projection and Visualization Techniques: Example of Case 2



- Artificial 3D sphere data
- Points distributed only on the sphere shell
- Intrinsically 3D data !
- Mapping error to 2D unavoidable
- Sphere is flattended: ROI contains vectors from front & rear
- Recomputation separates data !





Hierarchical Visualization

- Advanced Projection and Visualization Techniques: Example of Case 1
- Computational effort & mapping quality can be traded-off in a hierachical data clustering and mapping approach
- Data is first quantized by SOM, c-means, or other clustering method
- Then, the K centroids of the quantized data are mapped by NLM
- The complete data set is then mapped by NLM(R)
- Considerable speed-up can be achieved:

$$SMF = \frac{SA_{complete}}{SA_{hier}} \tag{6.3}$$

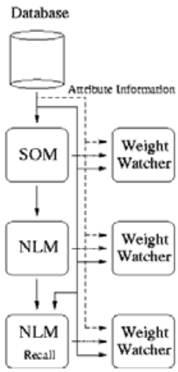
$$with \quad SA_{complete} = iterations * \frac{N*(N-1)}{2}$$

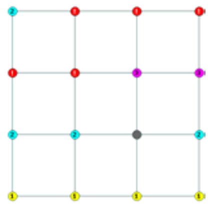
$$and \quad SA_{hier} = iter_t * \frac{K*(K-1)}{2} + iter_r * N * K$$

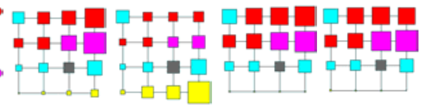
- Obviously, K controls computational effort & mapping quality
- The required computational effort for clustering must be included

Hierarchical Visualization

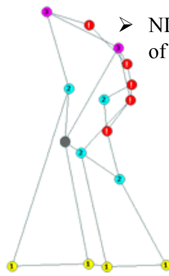
- Hierarchical Mapping of Iris Based on NLM and NLMR



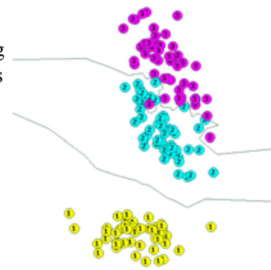




- Quantization of Iris data by SOM
- Optional component plane visualization



- NLM Mapping of 16 centroids

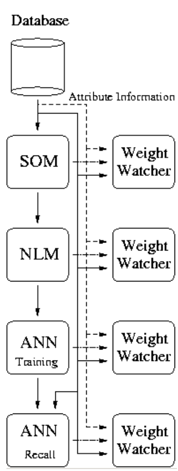


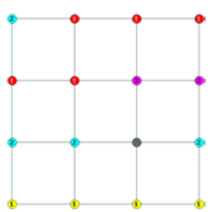
- Mapping of all 150 samples by NLM(R)

Sensor Signal Processing
Data Visualization

Hierarchical Visualization

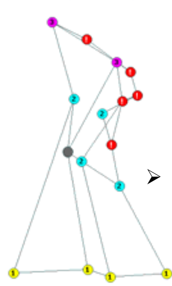
- Hierarchical Mapping of Iris Based on NLM and BP network



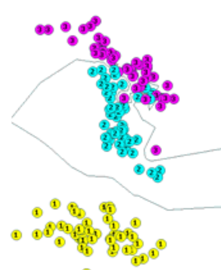


- Quantization of Iris data by SOM

- NLM Mapping of 16 centroids



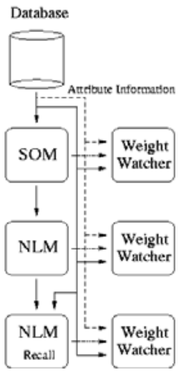
- Recall of BP with centroids

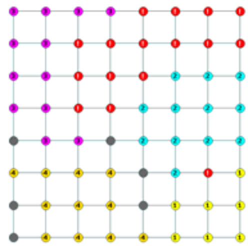


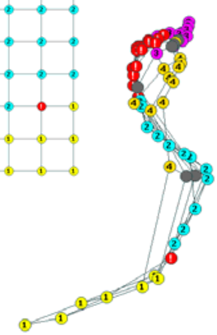
- Mapping of all 150 samples by BP recall

Hierarchical Visualization

➢ Hierarchical Mapping of Mechatronic Based on NLM and NLMR



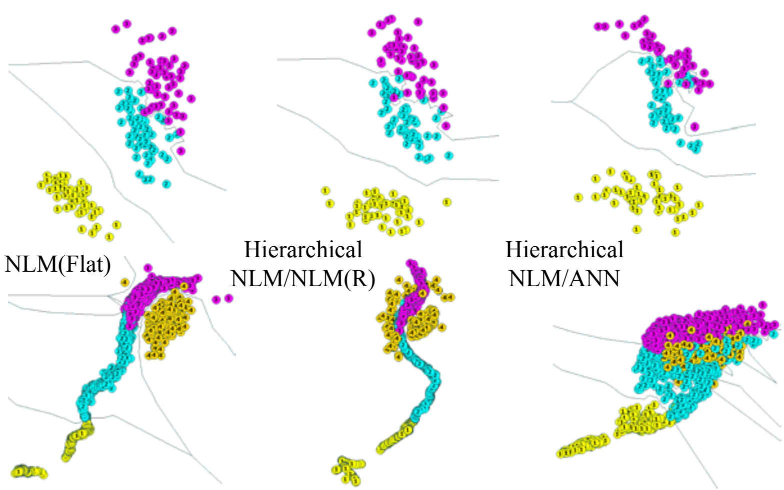




➢ Mechatronic (1775 samples, 24-D) data
➢ K=64 centroids, SMF=26.73 (K=3, SMF=591 !)

Hierarchical Visualization

➢ Visual Assessment of Hierarchical Mappings:



Hierarchical Visualization

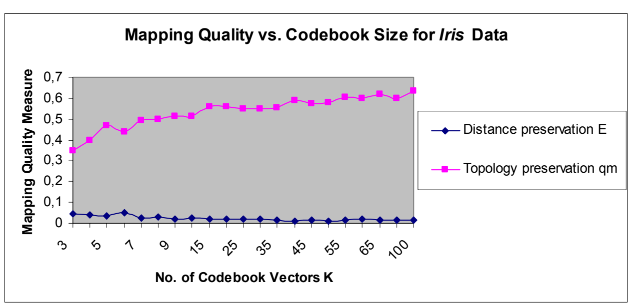
- Quantitative Assessment of Hierarchical Mapping
- The introduced distance & topology preservation measures are employed for quantitative comparison of three regarded data sets:

Data set	Mapping	E	q_m
Iris data	NLM	0.01439	0.6366
Iris data	NLM(R)	0.01791	0.5000
Iris data	BP	0.66365	0.4794
Mechatronic	NLM	0.02875	0.1671
Mechatronic	NLM(R)	0.05713	0.1130
Mechatronic	BP	0.75846	0.1059
RIAD	NLM	0.05116	0.2421
RIAD	NLM(R)	0.10400	0.1754
RIAD	BP	0.14935	0.1340

- The hierarchical mapping approach based on NLM & NLM(R) outperforms the NLM & ANN approach in terms of complexity, speed, user convenience, and results

Hierarchical Visualization

- Mapping Quality vs. Computational Expense in Hierarchical Mapping
- The introduced distance & topology preservation measures are employed for quantitative comparison of three regarded data sets:

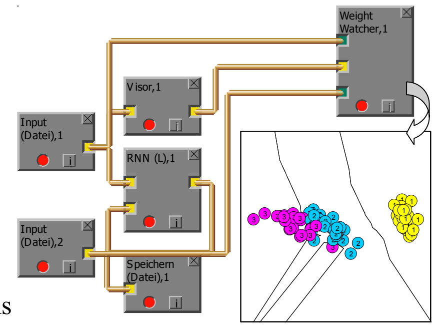


- Gradual increase of topoloy preservation & decrease of distance preservation error can be observed for increasing K. Nonmonotonicity is due to chosen clustering procedures (single runs)

Tools

- Tools for Interactive Multivariate Data Visualization: WeightWatcher
- **Research objective:** Achievement of a transparent & intuitive MMI to access and analyse multivariate data for information processing
- Employment of efficient projection techniques, enhanced user interaction, and CAD-inspired functionality
- Optimum exploitation of the remarkable human perceptive and associative capabilities

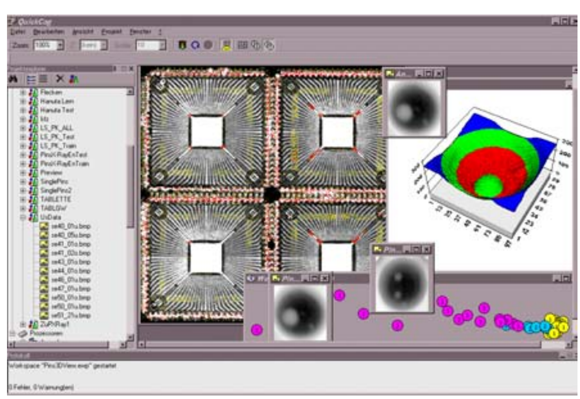
- Thus, the **WeightWatcher** (WW) was conceived for visual analysis of neural networks
- The concept was generalized to arbitrary multivariate data sets
- WW is part of QuickCog, a platform for fast & transparent design of vision & cognition systems



Tools

- WW in QuickCog serves for feature space visualization & assessment
- All described visualization techniques are implemented by WW
- The QuickMine-Toolbox comprises projection & clustering methods with WW

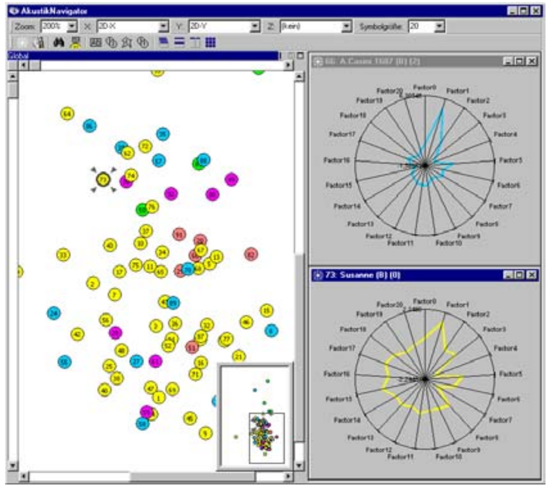
- An MS-Excel-Macro allows import of arbitrary application data
- QuickMine can be employed for general data visualization & analysis tasks



Tools

➢ Tools for Interactive Multivariate Data Visualization: Acoustic Navigator

- AN is dedicated to psychoacoustics & sound engineering
- Integrated in the processing hierarchy of PATS system
- Combines visualization & auralization in the analysis process
- Global & local sound data visualization
- Search functions & directed navigation
- Tentative edit functions for complex pattern synthesis [Sammon 69]

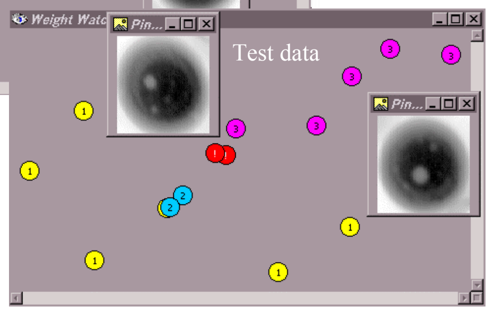


Applications

Benefits of Dimensionality Reduction and Visualization for the Rapid Design of Vision & Cognition Systems:

- ➢ Clustering & structure in the data can be observed
- ➢ Discriminance of the computed feature set can be assessed
- ➢ Assessment of different feature computation methods is feasible
- ➢ Open-loop parameter optimization for feature computation is feasible
- ➢ Interactive selection of significant features by WW component planes
- ➢ Outliers & identicals can be easily identified and eliminated from the data set
- ➢ Human errors during data acquisition, labeling & filing can be detected
- ➢ Large unlabeled or partially labeled data sets can be manually or semi-automatically according to human to the human observation of data structure and implied similarity in the projection
- ➢ Classification errors can be spatially located in the feature space projection
- ➢ Understanding of design problems & trouble shooting is alleviated

➢ Application Examples: BGA X-Ray Inspection in Electronics Manufacturing



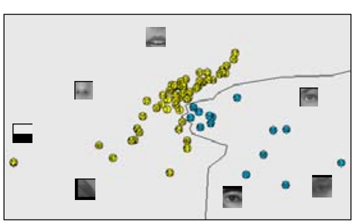
- Misclassifications & outliers can be tracked-back in the database by mouseclick in the WW display
- System design, validation, optimization, and trouble shooting is alleviated

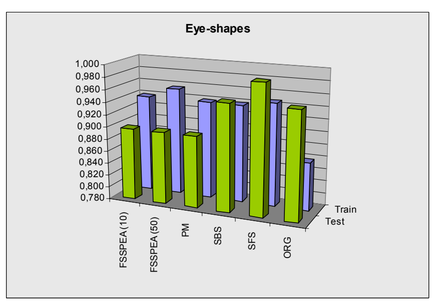
➢ Application Example: Eye-Tracking System (s. Introduction)

FS for Eye-Shape Classification:

- FSSPEA finds best solution
- Generalization is affected !
- SFS solution for HW-design

Feature space for FSSPEA





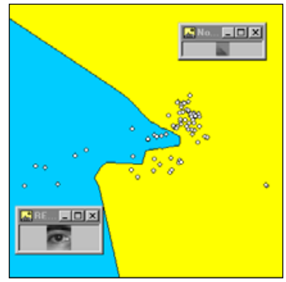
Method	Train	Test	Features
ORG	0,992500	0,981667	24
SFS	0,997500	0,928333	3
SBS	0,989333	0,981667	12
PM	0,997500	0,959167	6
SGA	0,997333	0,969167	12
FSSPEA	0,997500	0,928333	3

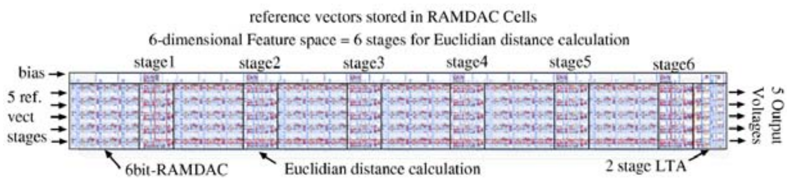
Applications

➢ Application Example: Eye-Tracking System <u>Hardware</u>

<u>RNN Low-Power Analog Eye-Shape Classifier:</u>

- 12-D Gabor jet reduced to 6 features by AFS
- RNN constrained to Manhattan distance, T_{RNN}=6
- 6 bit RAMDACs for weights
- AMS 0.6µm CUQ CMOS-technology
- Cell size 1595 um x 184 um (conservative)
- Power dissipation 348,9 nW, $3{,}5\times10^{-10}$ J
- Classification speed: 1ms/pattern, R=100%





Summary

➢ This chapter gave an introduction to and a focused survey of the principles and techniques for interactive visualization of multivariate data

➢ The strong relation with dimensionality reduction techniques, in particular distance preserving mappings, was pointed out

➢ Various auxiliary visualization modes for components, grids, attributes, hierarchical database access and outlier detection werde presented

➢ In particular, the issue of ubiquitous mapping errors was addressed and quantitative as well as fast visual assessment criteria were introduced

➢ Improvement of the concept by hierarchical approaches was introduced

➢ Two dedicated tools, the WeightWatcher & the Acoustic Navigator were presented

➢ Application examples were given, e.g., rapid prototyping and transparent design of recognition systems