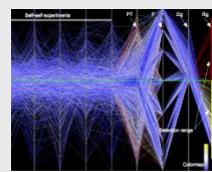


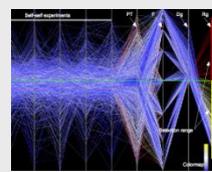
Informations- visualisierung

Thema:	8. Spezifische Verfahren - Parallele Koordinaten
Dozent:	Prof. Dr. Geric Scheuermann scheuermann@informatik.uni-leipzig.de
Sprechstunde:	nach Vereinbarung
Umfang:	2
Prüfungsfach:	Modul Fortgeschrittene Computergrafik Medizininformatik, Angewandte Informatik



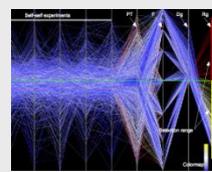
Wiederholung

- Was wird bei der SPM oberhalb und unterhalb der Diagonalen dargestellt?
- Was wird auf der Diagonalen dargestellt?
- Welche Ordnung ist in Zeile/Spalte definiert?
- Wie kann der Zusammenhang zwischen den einzelnen Scatterplots visuell analysiert werden?
- Wozu sind kategorische/diskrete Daten gut?
- Was ist das Problem bei hochdim' Daten?
- Wie können Dimensionen reduziert werden?
- Was ist der GrandTour-Ansatz?
- Wie unterscheidet er sich von Projection Pursuit?



Wiederholung

- Was ist das neue von ScatterDice?
- Wie sieht die Transition aus?
- Was unterscheidet kartesische und parallele Koordinatensysteme?
- Was sind die Probleme von Parallelen Koordinaten?



8. Spezifische Verfahren

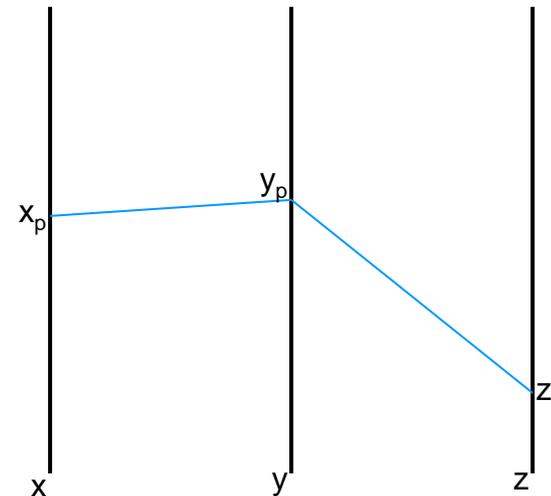
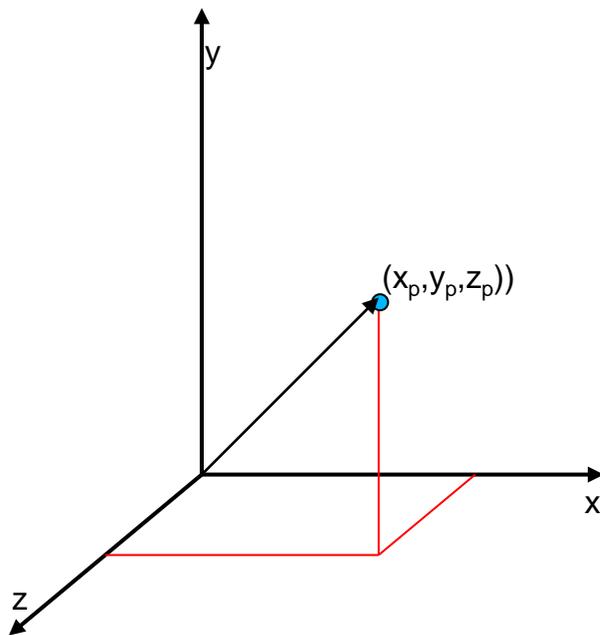
Verfahren für die Darstellung von Tabellen

- Histogramme
- Scatterplots
- Treemap & Mosaics
- Stacked Display / Dimensions
- Scatterplot Matrix
- **Parallele Koordinaten**

8.7 Parallele Koordinaten

Koordinatensysteme - Paralleles System [Inselberg 1985]

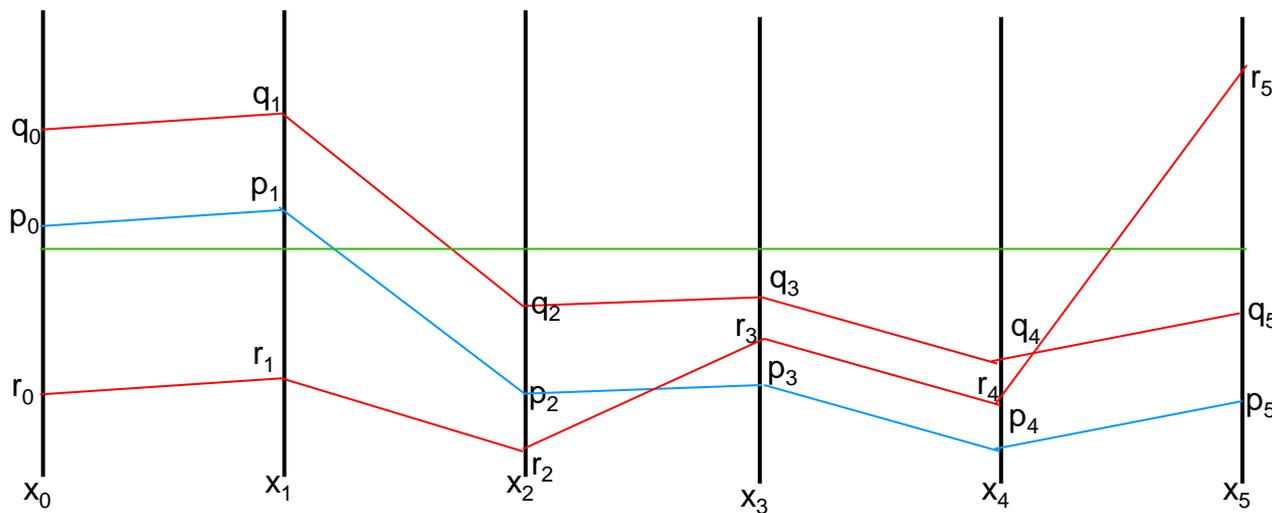
- Parallel Coordinates (Plot) - PC oder PCP
- Achsen liegen **parallel** zueinander
- Koordinaten sind **Tupel** der Dimensionswerte
- Werte im Koordinatensystem sind **Kantenzüge**

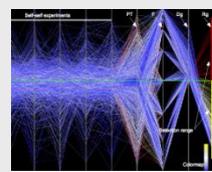


8.7 Parallele Koordinaten

Koordinatensysteme - Paralleles System [Inselberg 1985]

- Mehr Dimensionen, **mehr Achsen**
- **Reihenfolge** der Achsen liegt nicht unbedingt fest
- Skalierung der Achsen auf **Min/Max** der Dimension
- **Nulllinie/Mittellinie** kann angezeigt werden

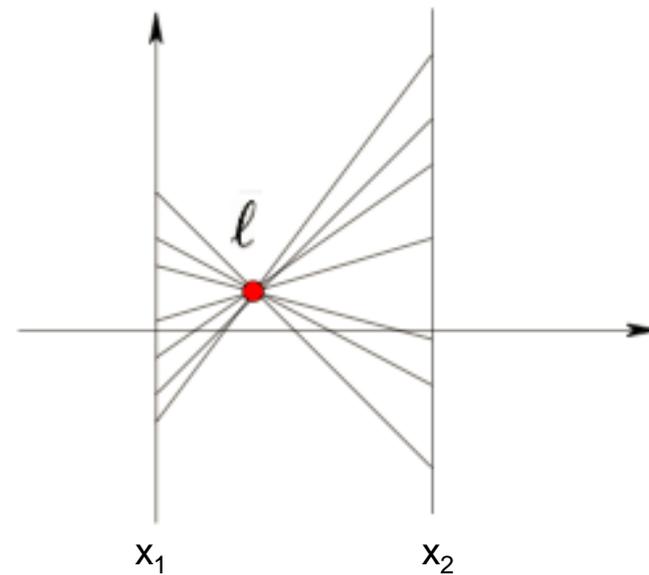
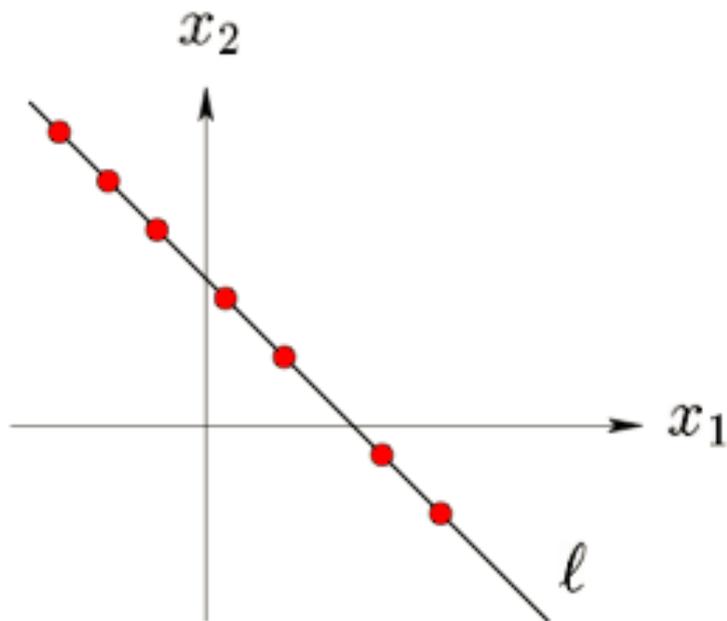




8.7 Parallele Koordinaten

Abbildungen von Geraden aus KKS ins PC

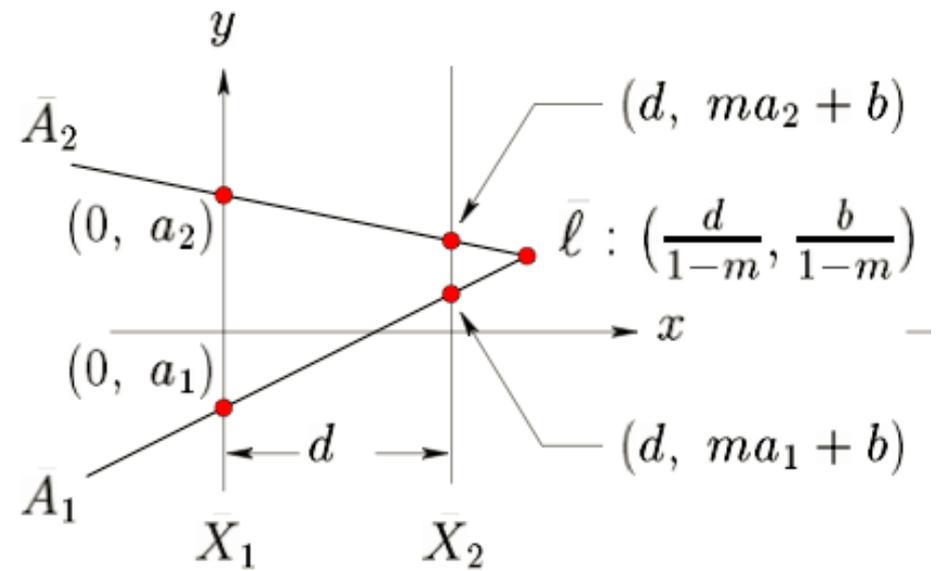
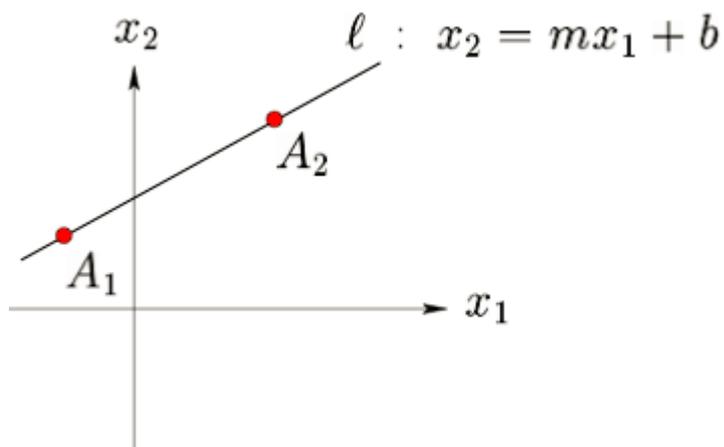
- Punkte auf einer Geraden l werden zu Kantenschar \bar{l}
- Linien- oder Kantendualität

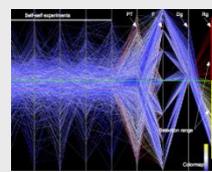


8.7 Parallele Koordinaten

Abbildungen von Geraden aus KKS ins PC

- Steigung $m > 0$, dann Koordinaten A_2 immer größer als von A_1
- Aus kartesischer Line entsteht dualer PC-Punkt \bar{l}
- Tendiert zu **positiver** Korrelation

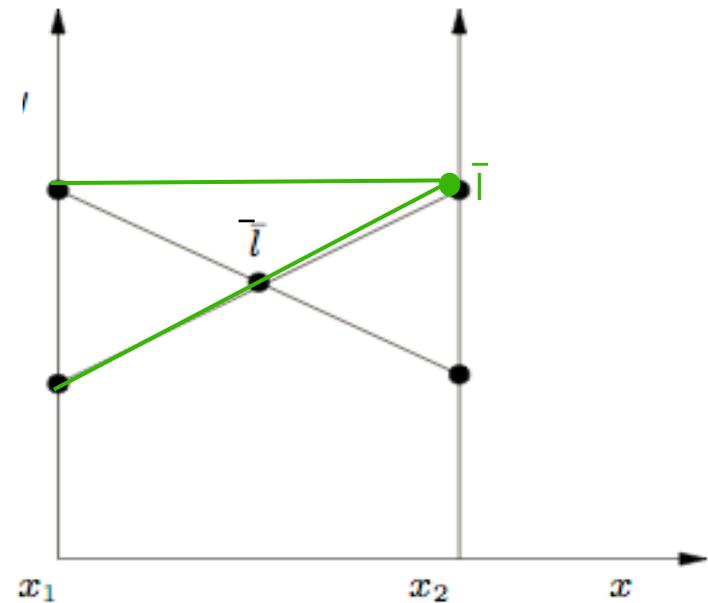
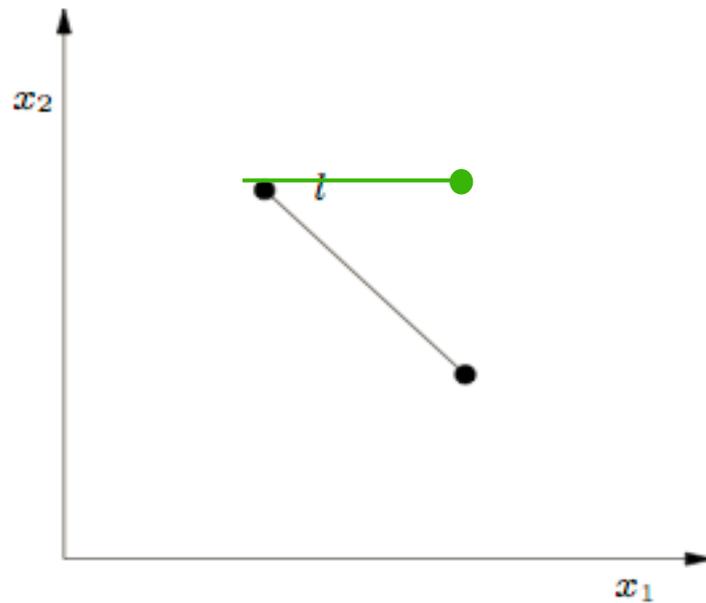


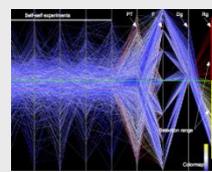


8.7 Parallele Koordinaten

Abbildungen von Geraden aus KKS ins PC

- Steigung $m < 0$, induzierter dualer Punkt \bar{l}
- Tendiert zu **negativer** Korrelation
- **$m=0$**

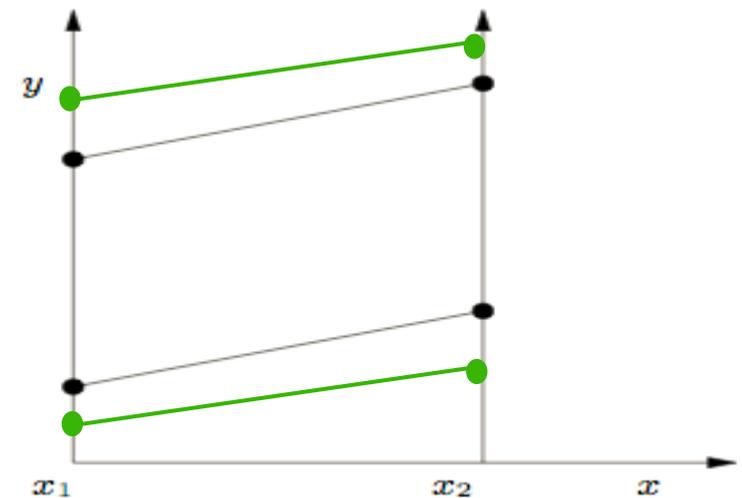
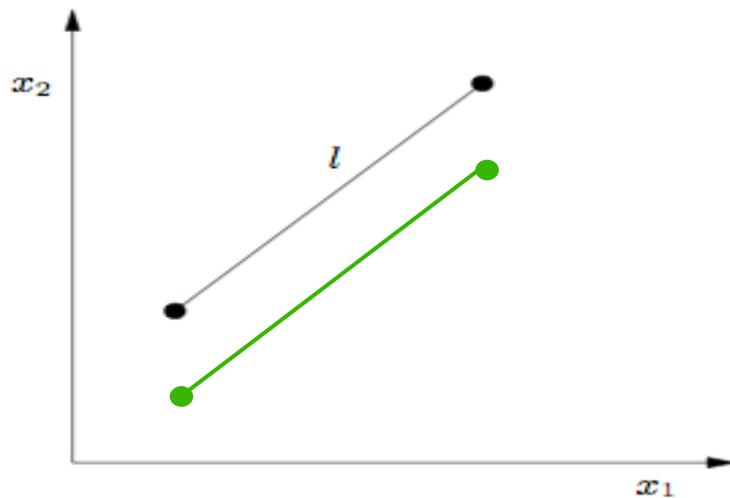


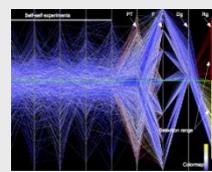


8.7 Parallele Koordinaten

Abbildungen von Geraden aus KKS ins PC

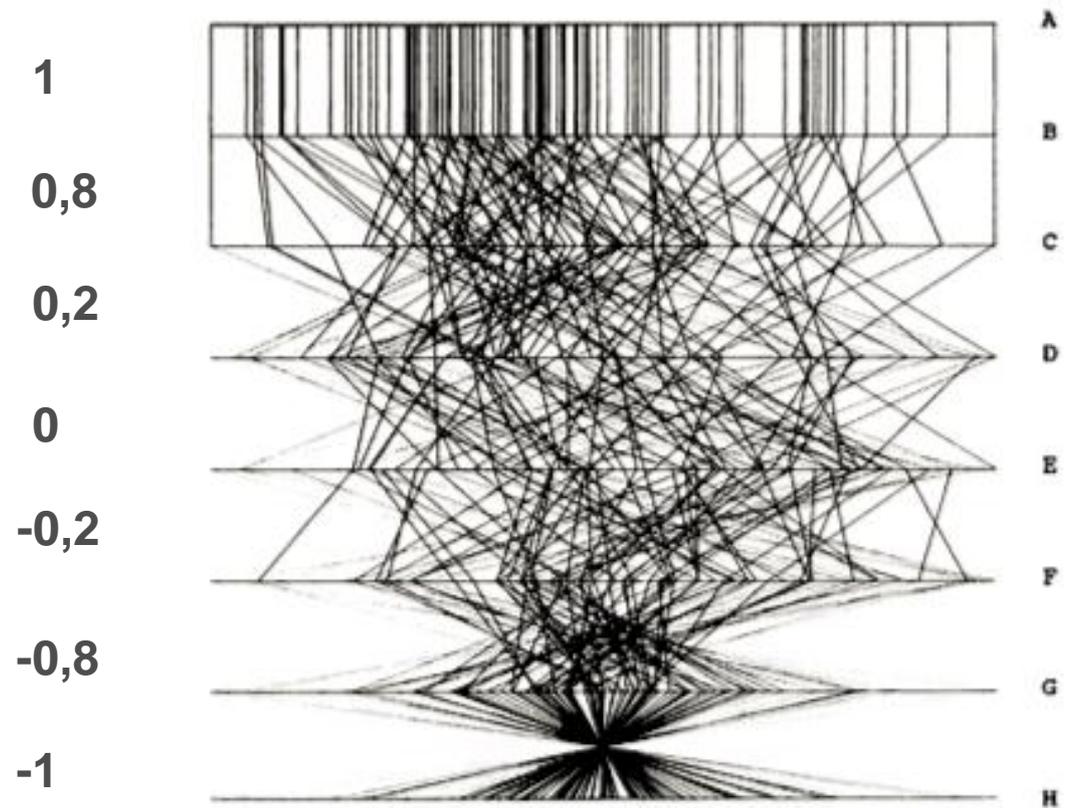
- Steigung $m=1$, induzierter dualer Punkt \bar{l} liegt im ∞ (Fernpunkt)
- Positive (starke)** Korrelation





8.7 Parallele Koordinaten

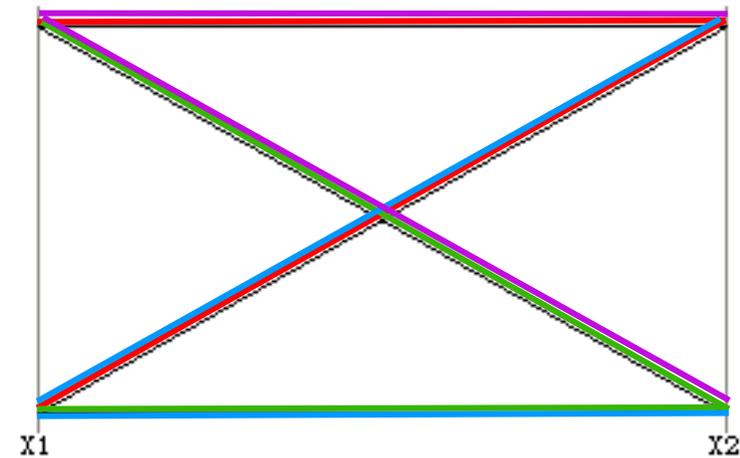
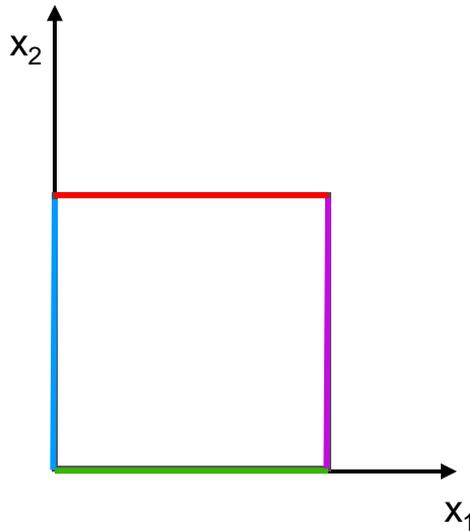
Korrelationen in 6D [Wegman 1990]



8.7 Parallele Koordinaten

Abbildungen von Objekten aus KKS ins PC

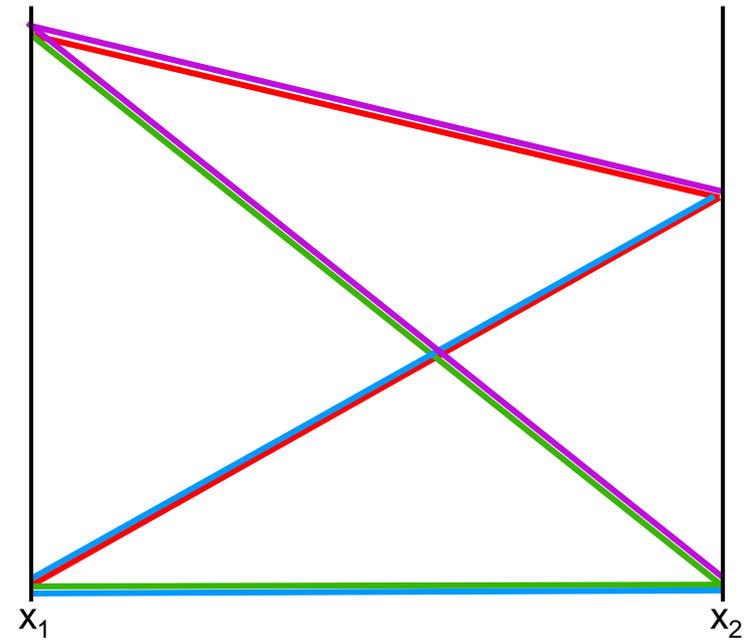
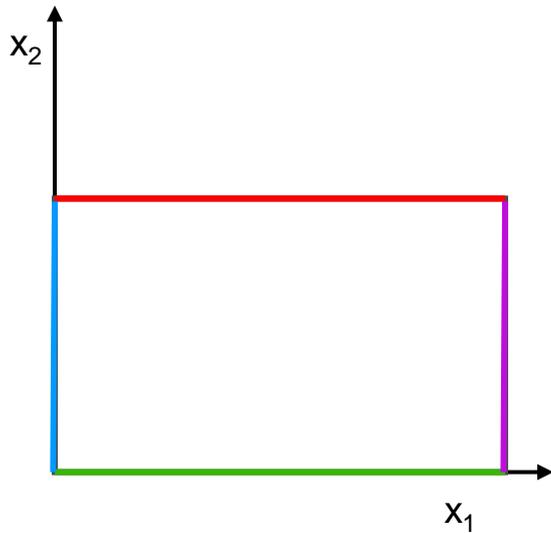
- **Quadrat, Projektion** in PC

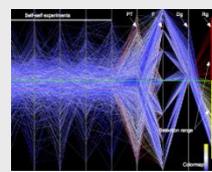


8.7 Parallele Koordinaten

Abbildungen von Objekten aus KKS ins PC

- **Rechteck, verzerrte** Projektion in PC

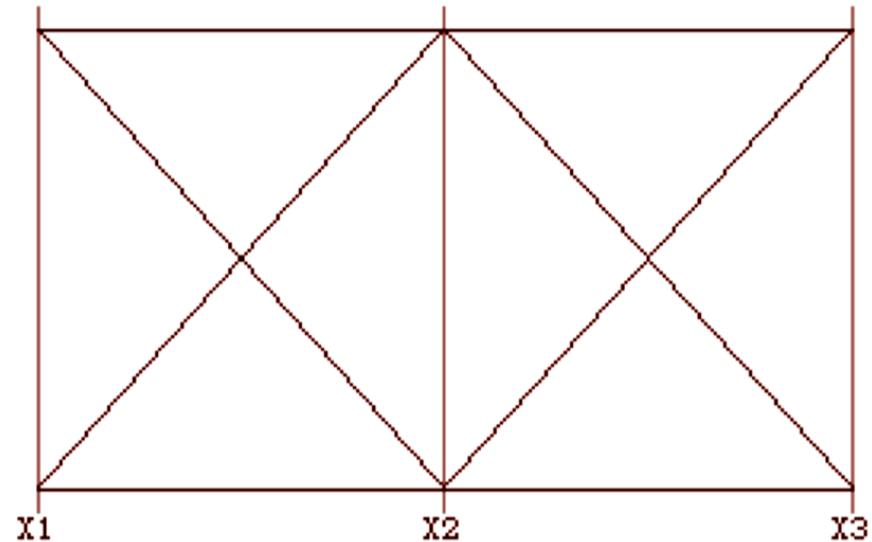
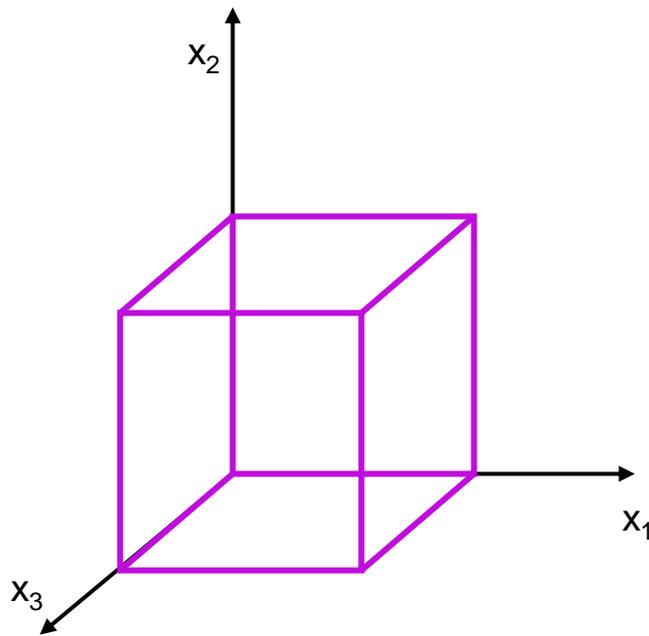


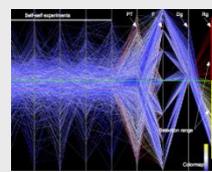


8.7 Parallele Koordinaten

Abbildungen von Objekten aus KKS ins PC

- Würfel (3D)

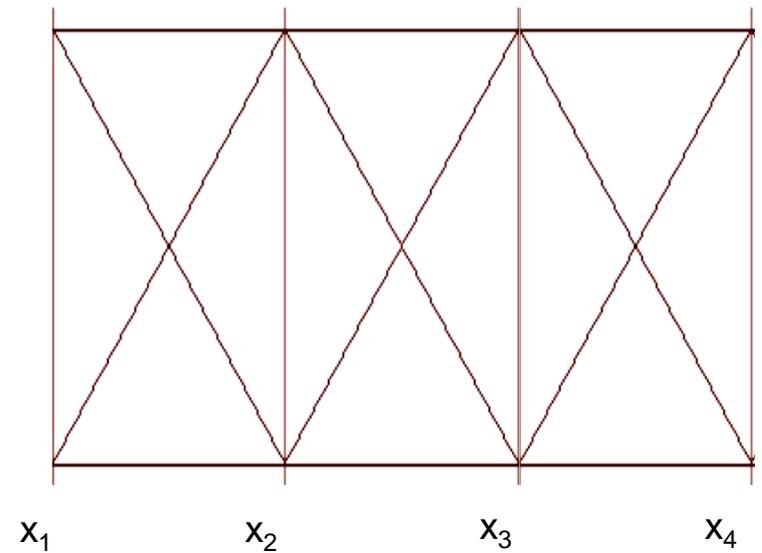
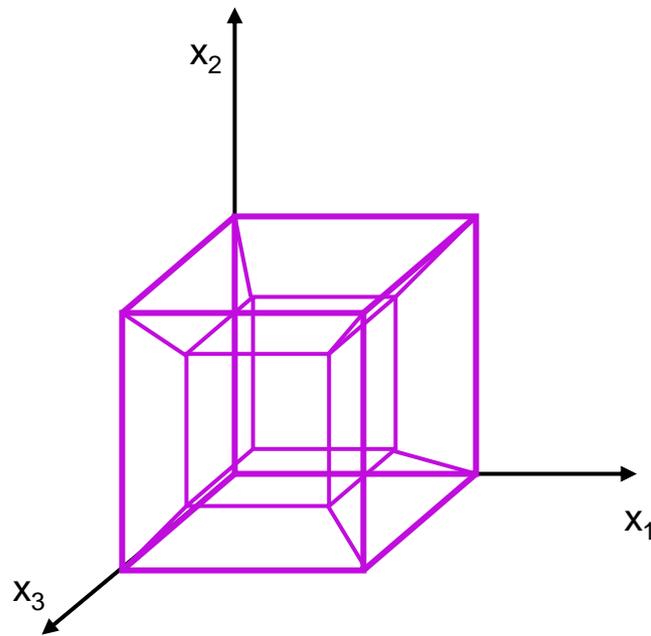


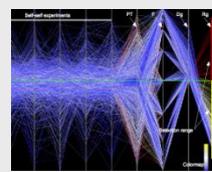


8.7 Parallele Koordinaten

Abbildungen von Objekten aus KKS ins PC

- Hyperwürfel (3D)

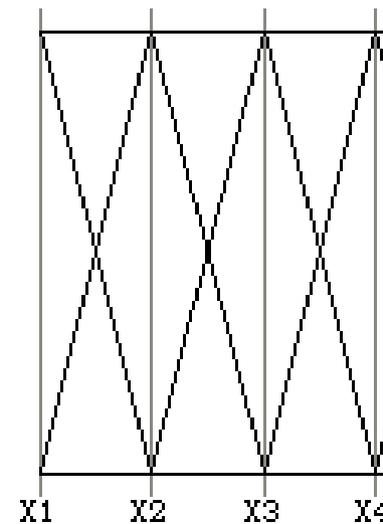
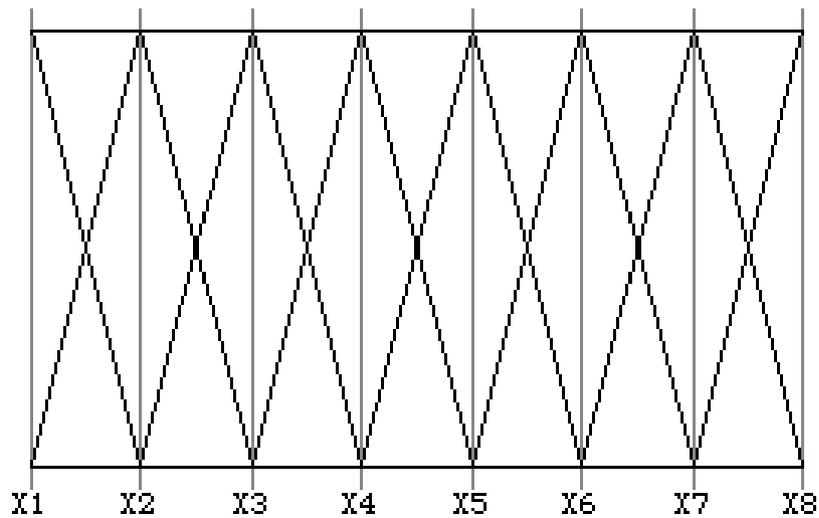




8.7 Parallele Koordinaten

Abbildungen von Objekten aus KKS ins PC

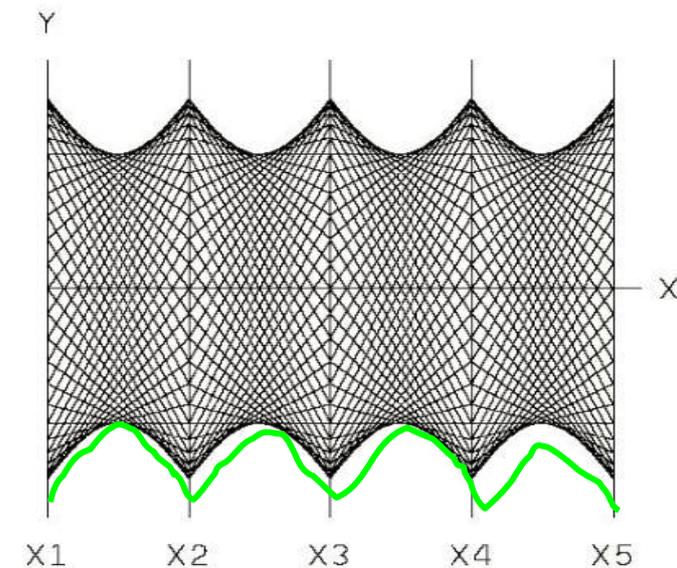
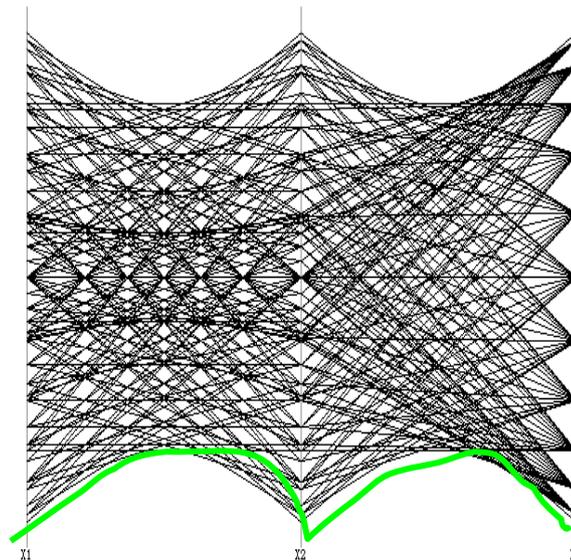
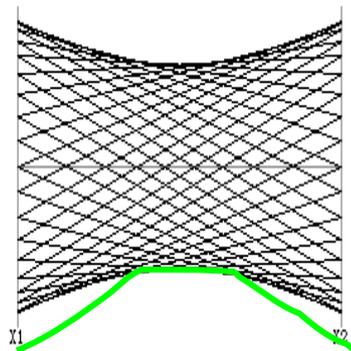
- Hyperwürfel (8D vs. 4D)

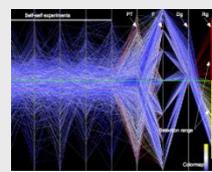


8.7 Parallele Koordinaten

Abbildungen von Objekten aus KKS ins PC

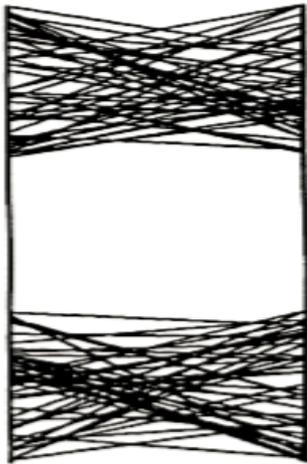
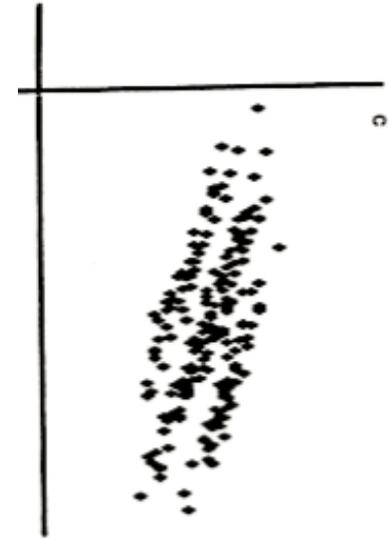
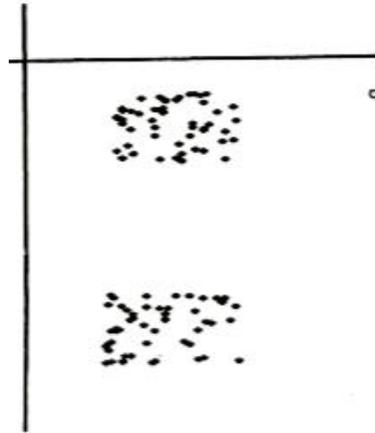
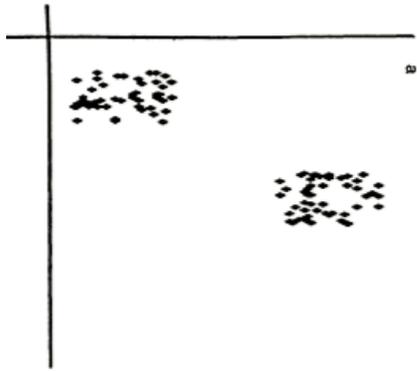
- Kreis (2D), Kugel (3D) und Hypersphäre (5D)
- #Kantenzüge hängen von **Tessellierung** ab
- **Unkorrelierte** Daten füllen hyperbolische **Hüllkurven**
- Korrelationen **vertiefen Hüllkurven** zu Ecken

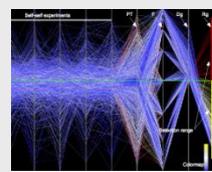




8.7 Parallele Koordinaten

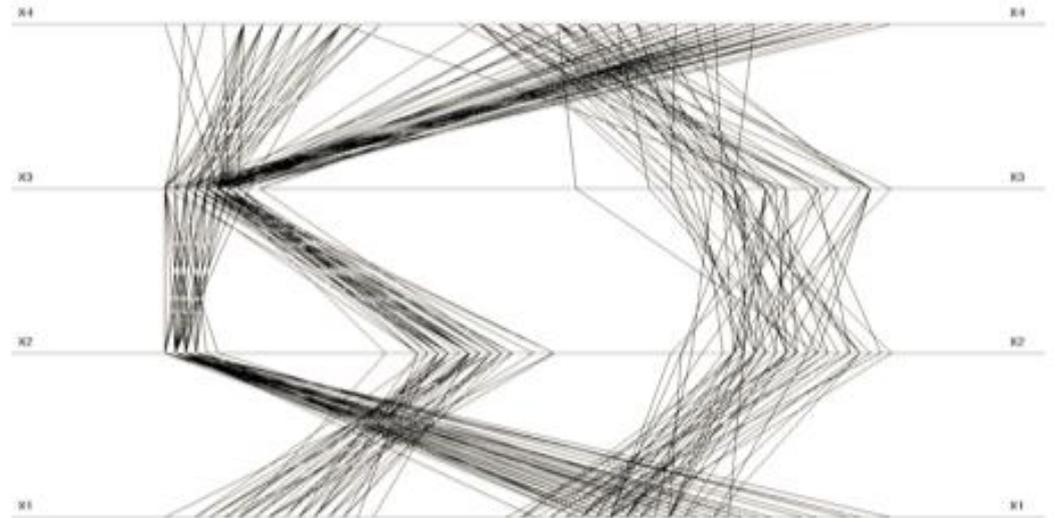
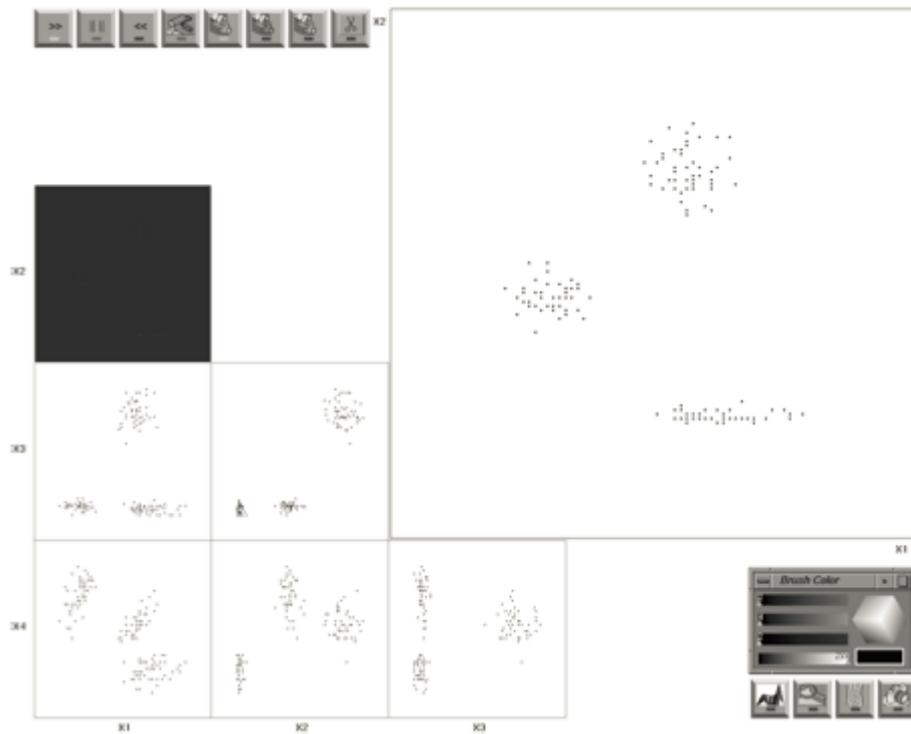
Cluster-Auftrennung [Wegman 1990]

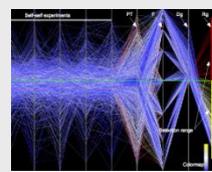




8.7 Parallele Koordinaten

Cluster-Auftrennung [Wegman, Luo 1996]

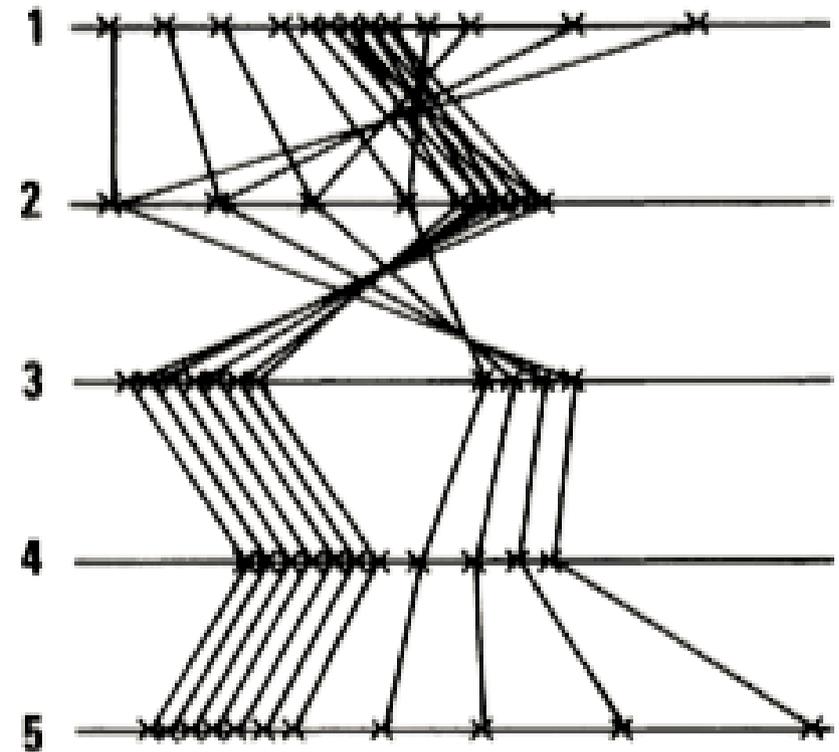


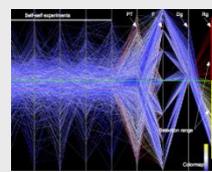


8.7 Parallele Koordinaten

Cluster-Auftrennung [Wegman 1990]

- **Verteilung** in 1,2 (Normal, χ^2)
- Korrelationen
- Cluster
- 3D **Hyperebene** in 5D



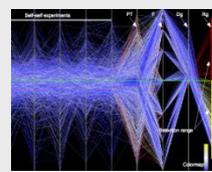


8.7 Parallele Koordinaten

Abbildungen von KKS ins PC

[Wegman 1990]

- Dualitäten (beide Richtungen) zwischen KKS und PC
 - **Punkte** werden zu **Linien**.
 - **Ellipsoide** werden zu hyperbolischen **Hüllkurven**.
 - **Rotationen** werden zu **Translationen**.
 - **Wendepunkte** werden zu **Scheitelpunkten**.



8.7 Parallele Koordinaten

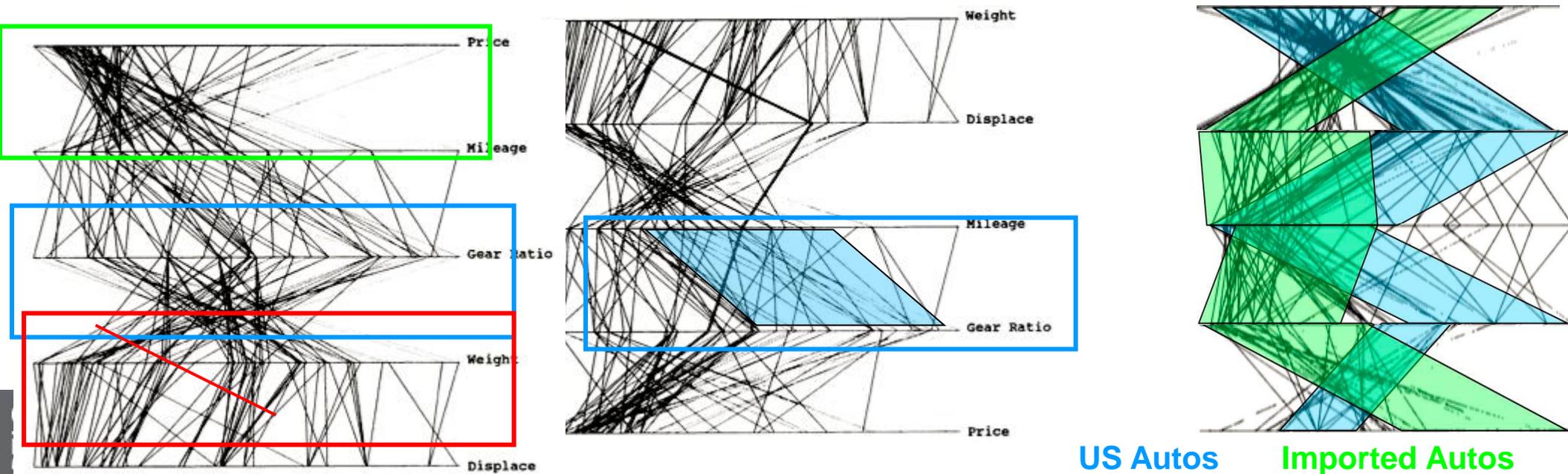
Interpretation [Wegman 1990]

- Interpretation von Korrelationen abhängig von **Achsenkonfiguration**
- **Permutationen** aller Dimensionsachsen:
(n+1)/2 für n Dimensionen
- Entsprechend erhöht sich der Interpretationsaufwand.

8.7 Parallele Koordinaten

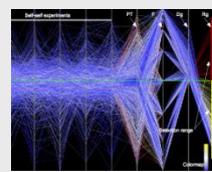
Interpretation [Wegman 1990]

- **Permutationen** der Dimensionsachsen
- Datensatz: 74 Autos des Modelljahres 1979



US Autos

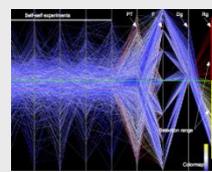
Imported Autos



8.7 Parallele Koordinaten

Interpretation [Wegman, Luo 1997]

- **Permutationen** der Dimensionsachsen
- Verwende **Grand-Tour Animationen**, um Permutationen „abzulaufen“



8.7 Parallele Koordinaten

Koordinatensysteme - Paralleles System

Lösungsansätze

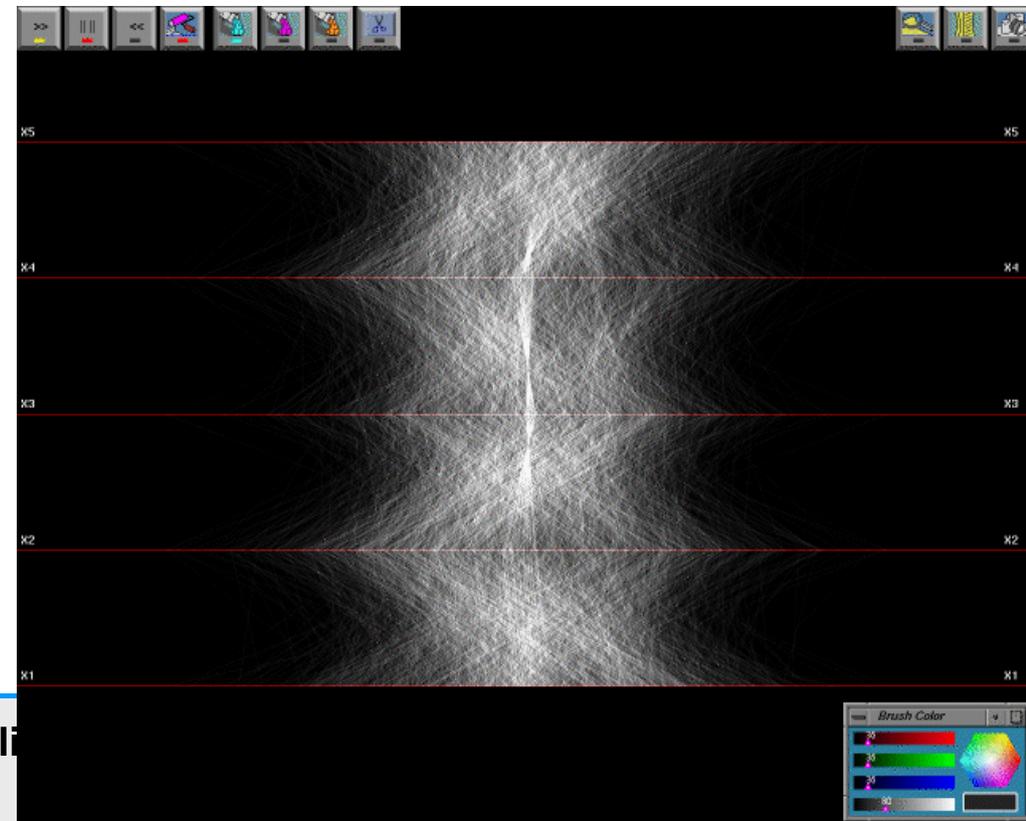
- Farb- und Opazitätsmodellierung
- Clustering
- Dimensionsreduktion
- Brushing-Techniken
- ...

8.7 Parallele Koordinaten

Visuelles Clustern [Wegman, Luo 1997]

Farbton/Sättigung Brushing

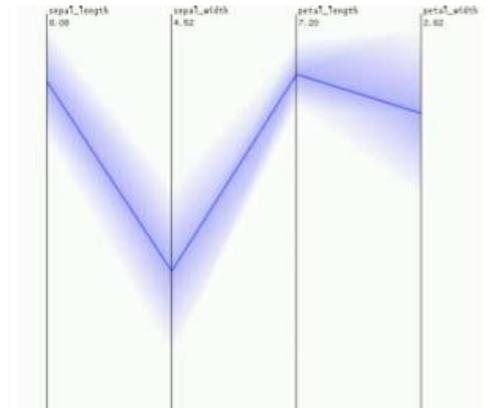
- Füge **leichten Blauton** hinzu: Bei **Overplotting** wird Darstellung blau
- „**Entsättigung**“ auf Schwarz: Je mehr **Overplotting**, desto heller (bei schwarzem Hintergrund)
- Alternativ auch als **Opazitätsmodellierung**



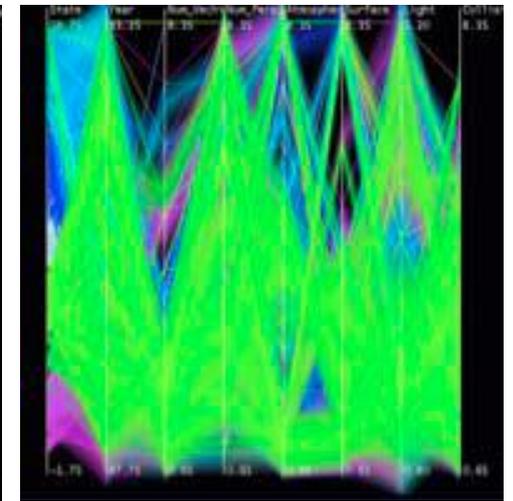
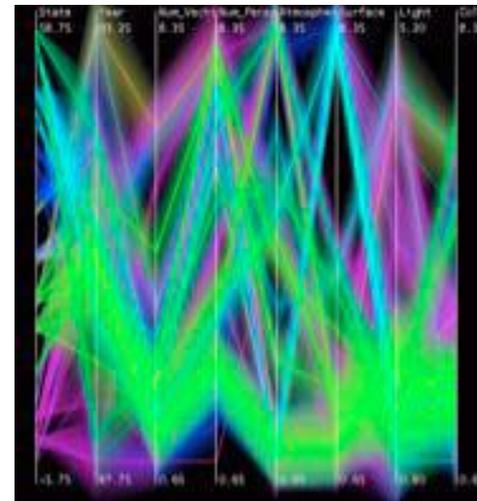
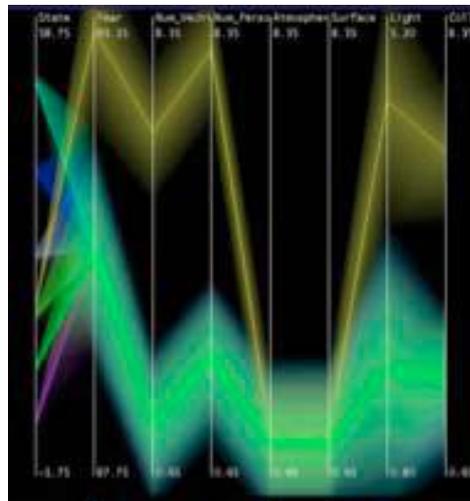
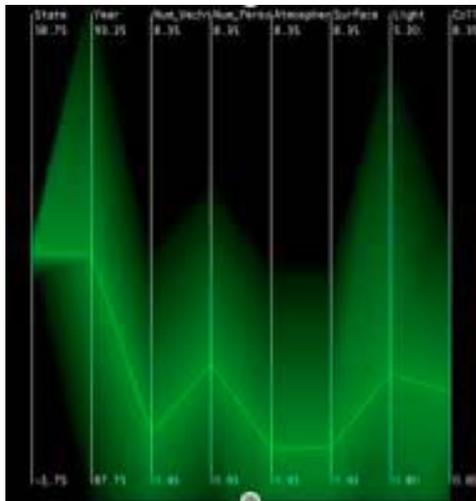
8.7 Parallele Koordinaten

Visuelles Clustern / Hierarchische PC [Fua et al. 1999]

- Hierarchisches Clustering \Rightarrow **Baumstruktur**
- Visualisierung von
 - Cluster-**Mitte**
 - Cluster-**Band**



Unterschiedliche **Detaillebenen** der hierarchischen Cluster

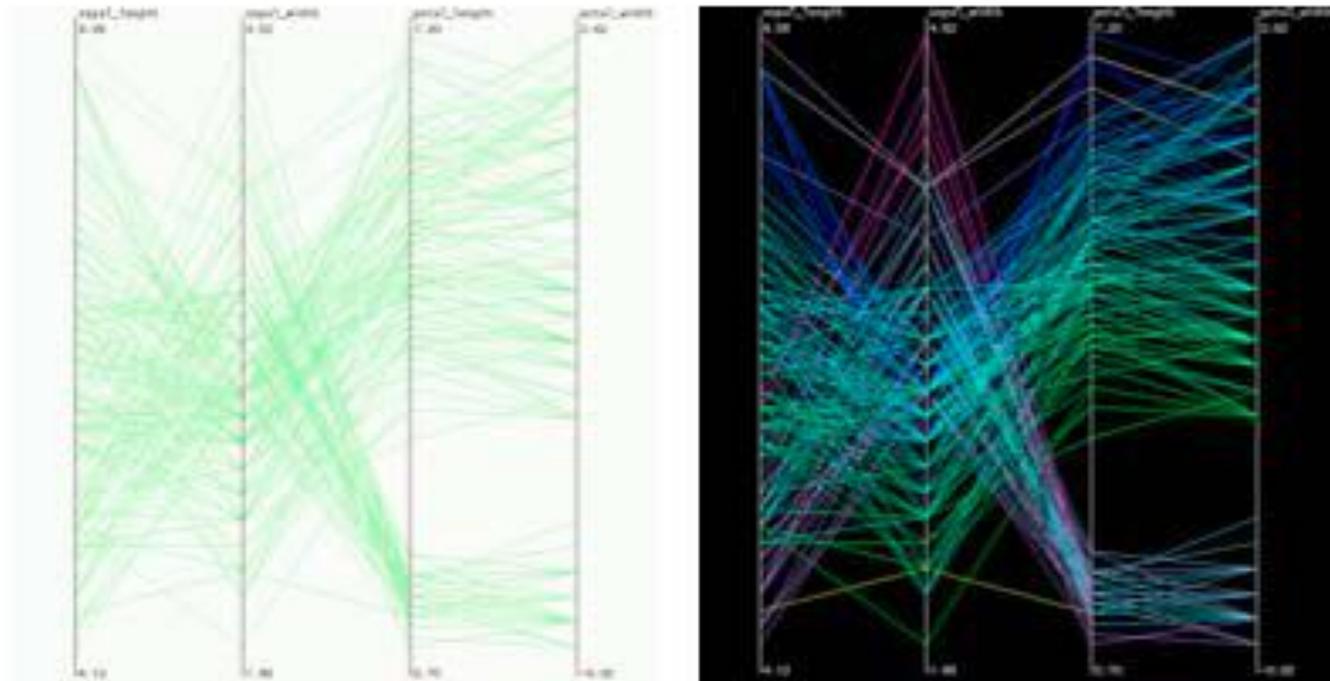


Informationsvisualisierung

8.7 Parallele Koordinaten

Visuelles Clustern / Hierarchische PC [Fua et al. 1999]

- Hierarchisches Clustering \Rightarrow **Baumstruktur**
- **Nähe**-basiertes Färben
 - Nutze Nachbarknoten zu einem Elternknoten
 - Nehme Distanz als Farbwahlkriterium - je näher, desto ähnlicher

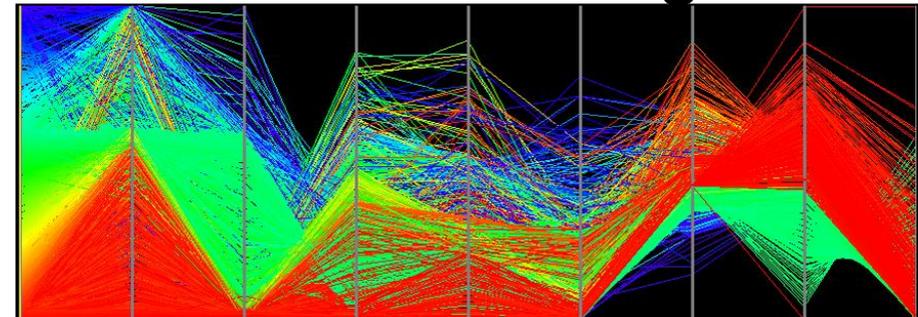
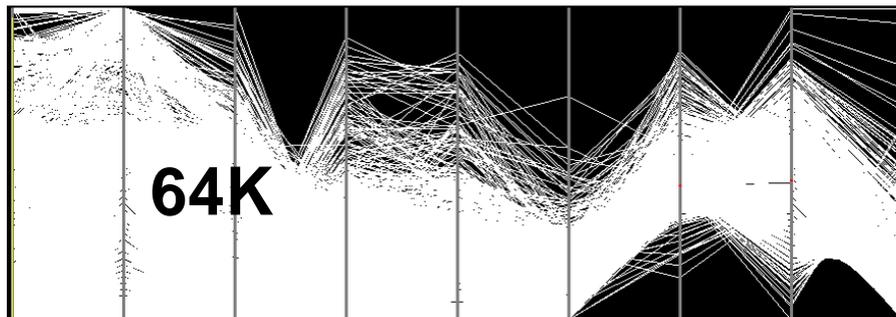


8.7 Parallele Koordinaten

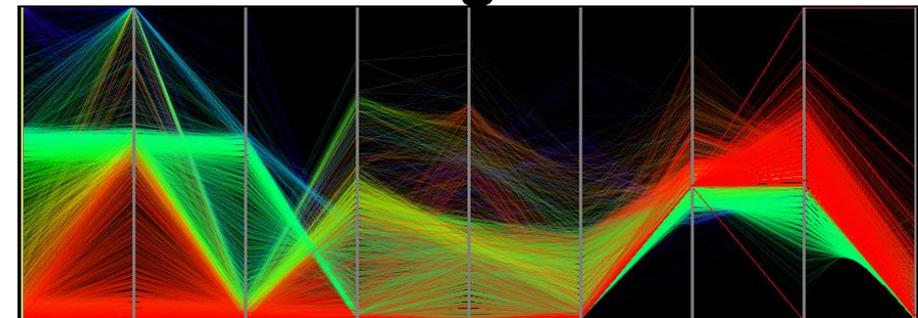
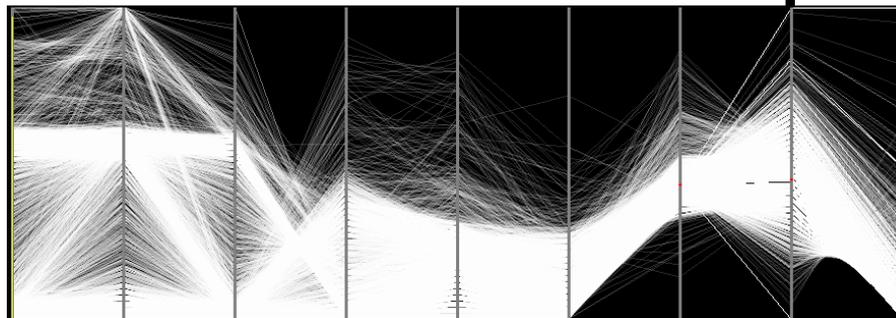
Visuelles Clustern [Pradhan et al. 2003/2005]

- ... über Opazität : **Gleiche Opazität** für alle Datensamples

Farbabbildung



Uniforme Opazitätsmodellierung

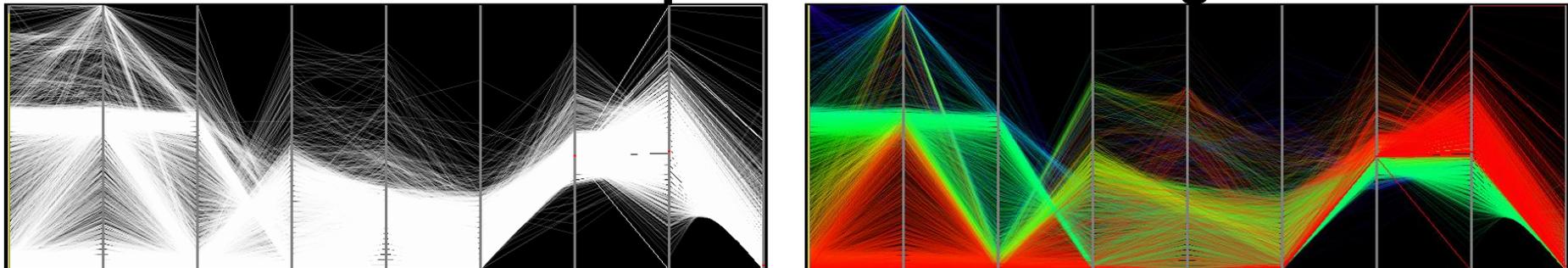


8.7 Parallele Koordinaten

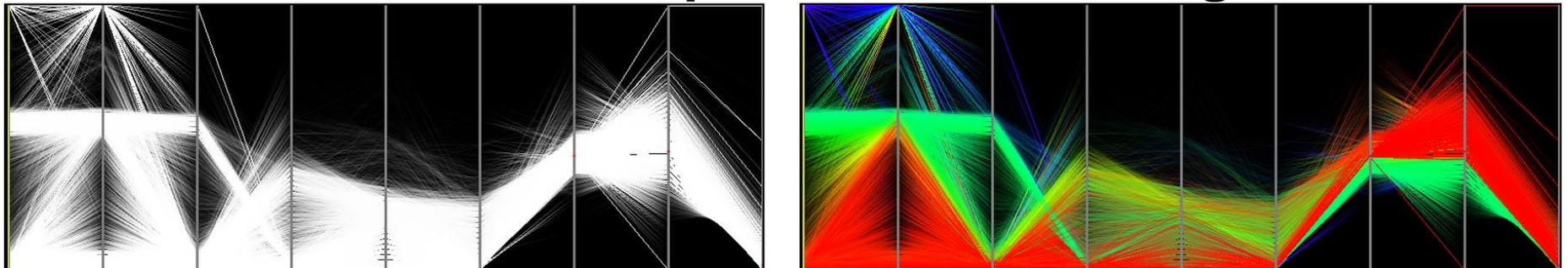
Visuelles Clustern [Pradhan et al. 2003/2005]

- ... über Opazität: **Gewichtung nach Anzahl** der Datensamples pro Klasse (pro Achse)

Uniforme Opazitätsmodellierung



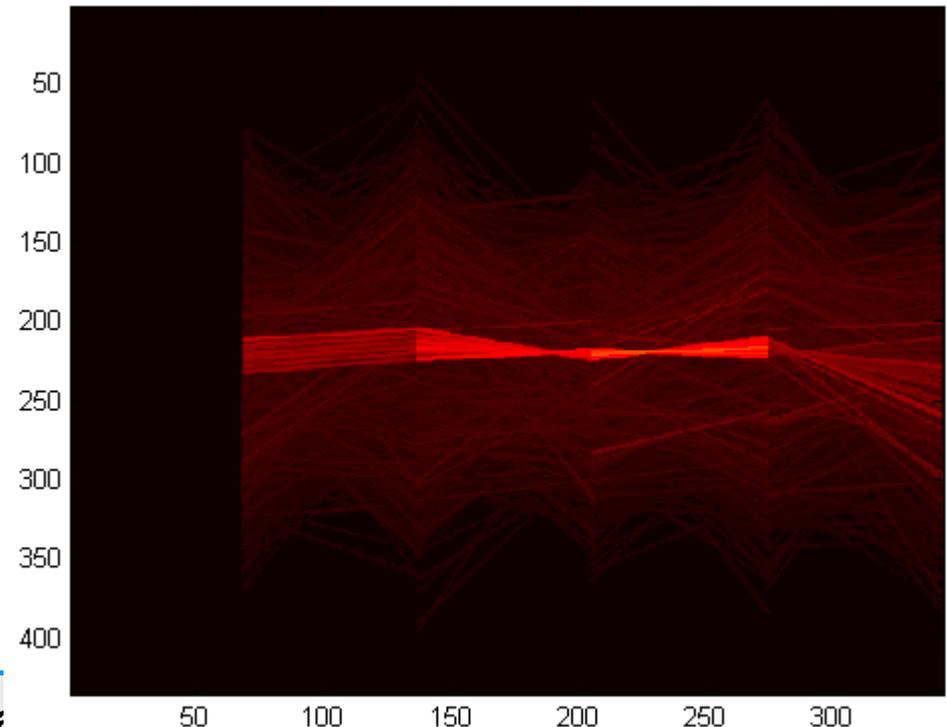
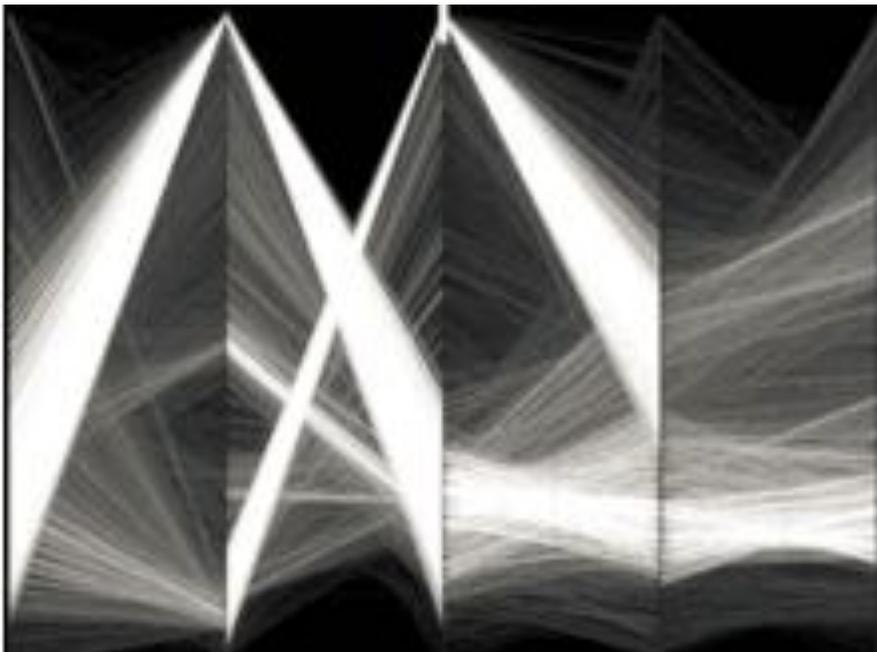
Gewichtete Opazitätsmodellierung



8.7 Parallele Koordinaten

Visuelles Clustern [Artero et al. 2004]

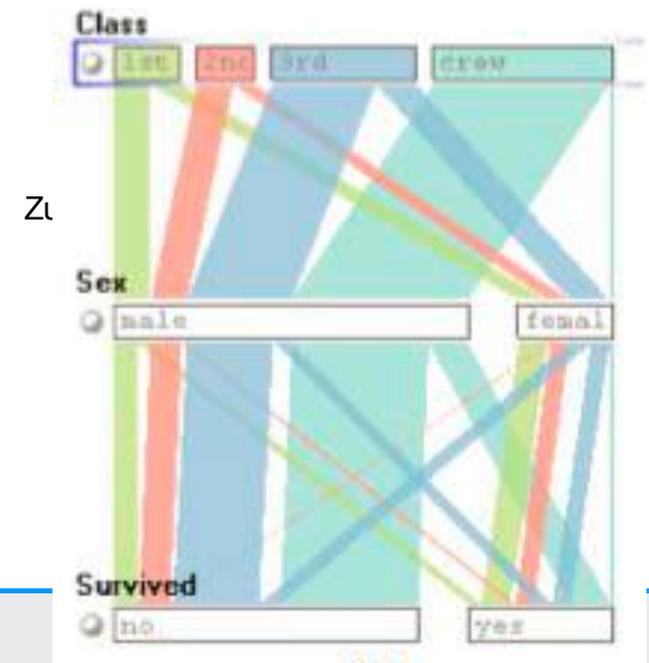
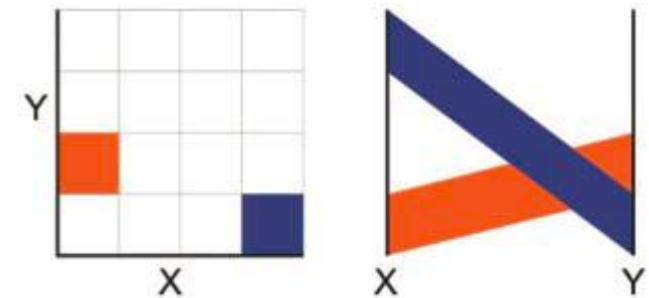
- ... über Intensität: **Gewichtung nach Anzahl** (Frequency) der Datensamples pro Klasse (pro Achse)
- Intensität **proportional** (linear) zur Anzahl; **Tiefe nach Intensität** sortiert
- Sammelklasse wird **gefiltert**



8.7 Parallele Koordinaten

Visuelles Clustern [Bendix et al. 2005]

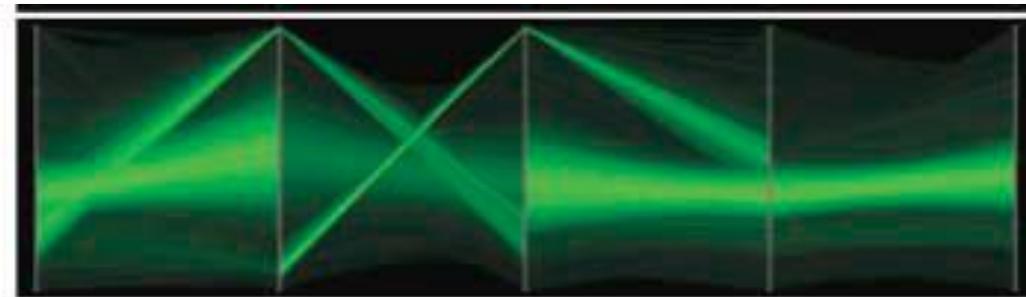
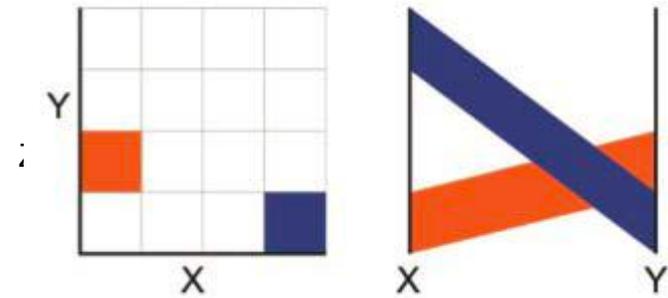
- ... über BinMap: **Aufteilung nach Häufigkeit** (Frequency) in Eimer zwischen **zwei Dimensionen**
- Für **kategorische** Daten
- Rendern als **Parallelogramm**
- Farbe entsprechend der ersten Dimensionsaufteilung
- Gruppieren, Umsortieren, Hervorheben, Filtern
- Beispiel: Titanic-Datensatz



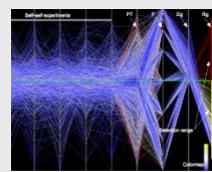
8.7 Parallele Koordinaten

Visuelles Clustern [Novotny, Hauser 2006]

- ... über BinMap: **Aufteilung nach Häufigkeit** (Frequency) in Eimer zwischen **zwei Dimensionen**
- Helligkeit entsprechend der **Samplendichte**
- **Outliers** mit anderer Farbe (rot)



128 x 128 BinMap



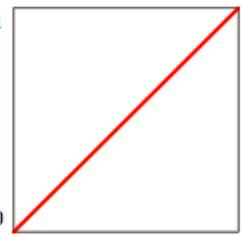
8.7 Parallele Koordinaten

Visuelles Clustern [Johansson et al. 2005]

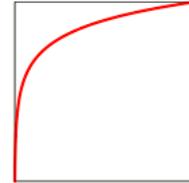
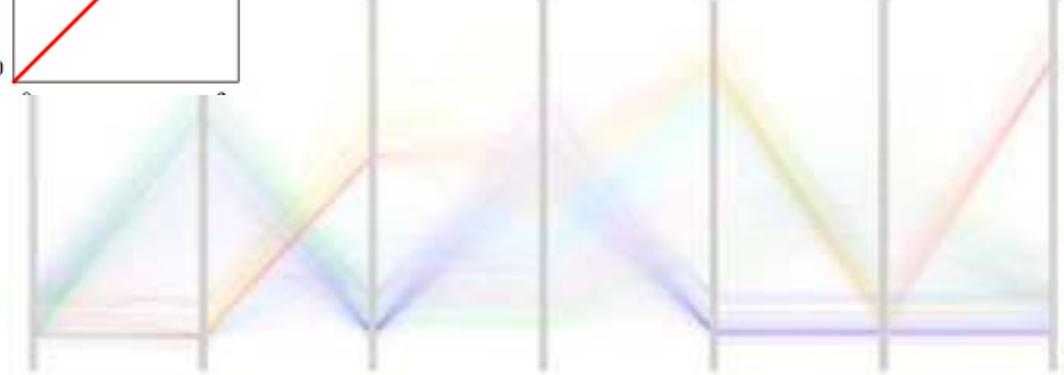
- Cluster-Berechnung in Vorverarbeitung
- Bestimmung der Cluster in Textur und **normalisiert nach Anzahl** (Frequency)
- Auf **ausreichend Datentiefe** achten
- Opazität des Cluster-Bands über **Transferfunktion** spezifizieren
- Cluster-Darstellung wird über **Varianz/Schiefe animiert**

8.7 Parallele Koordinaten

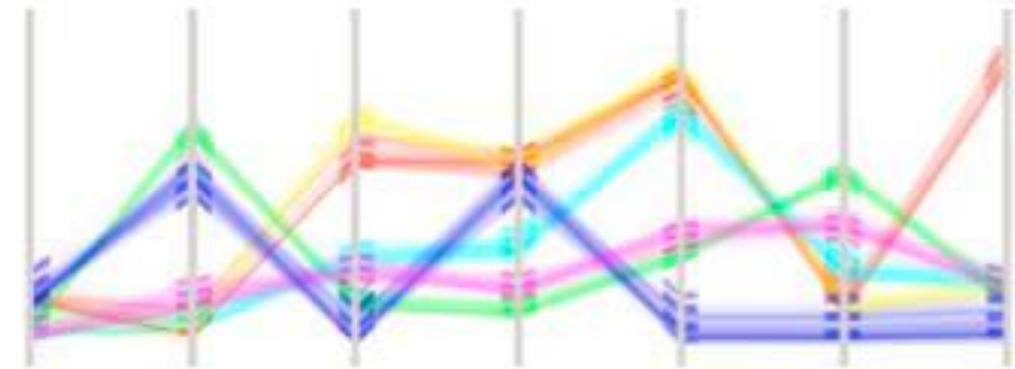
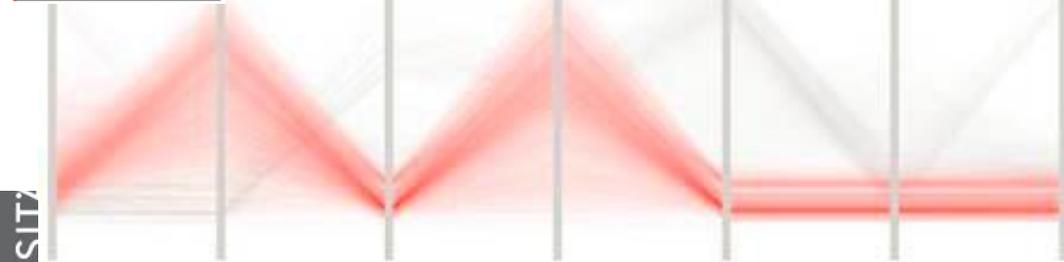
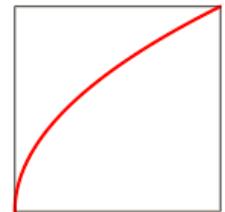
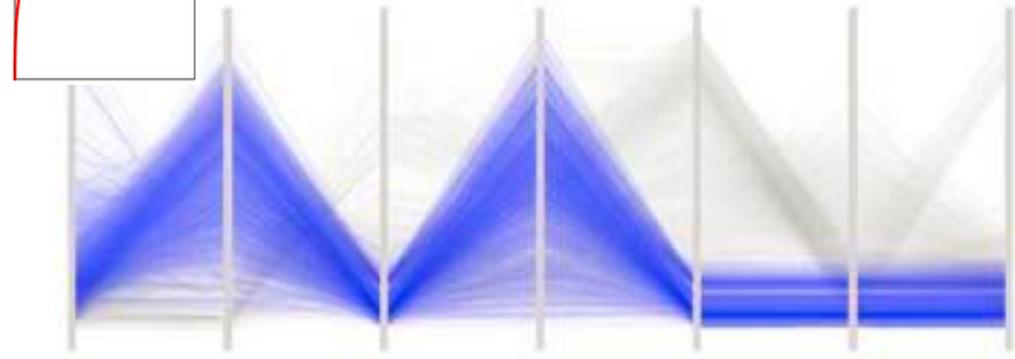
Visuelles Clustern [Johansson et al. 2005]



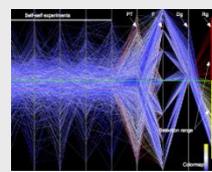
<= Linearer Verlauf der Transparenz
(Transferfunktion im Sinne des Volume Renderings)



<= Logarithmischer Verlauf der Transparenz



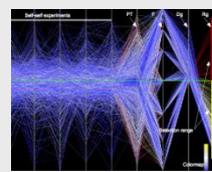
Feature Animation



8.7 Parallele Koordinaten

Dimensionsreduktion

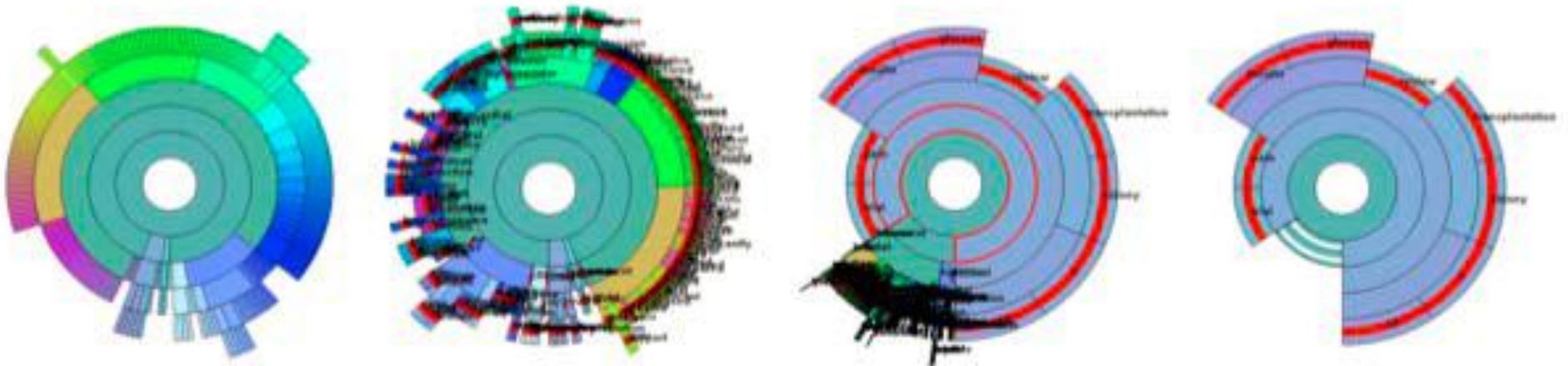
- Clusterverfahren bisher haben **Datensamples** zusammengefasst
- Hier werden nun **Dimensionen zusammengefasst**
- Viele verschiedene Techniken
 - Hauptkomponentenanalyse (PCA)
 - Multidimensional Scaling (MDS)
 - ...

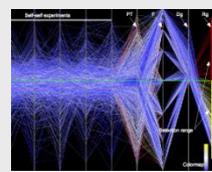


8.7 Parallele Koordinaten

Dimensionsreduktion [Yang et al. 2003]

- Bestimmt **Hierarchie von Dimensionen**
- Kontrolliert **Grad** von Dimensionen**details**
- Ermöglicht **Neuordnung** der Dimensionen
 - Mit Ähnlichkeitsmaß
 - Mit Relevanzmaß





8.7 Parallele Koordinaten

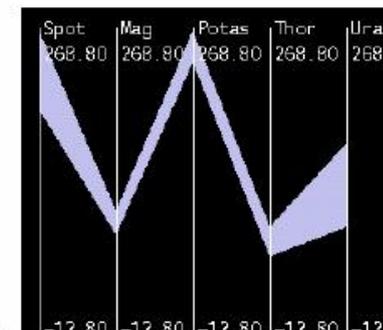
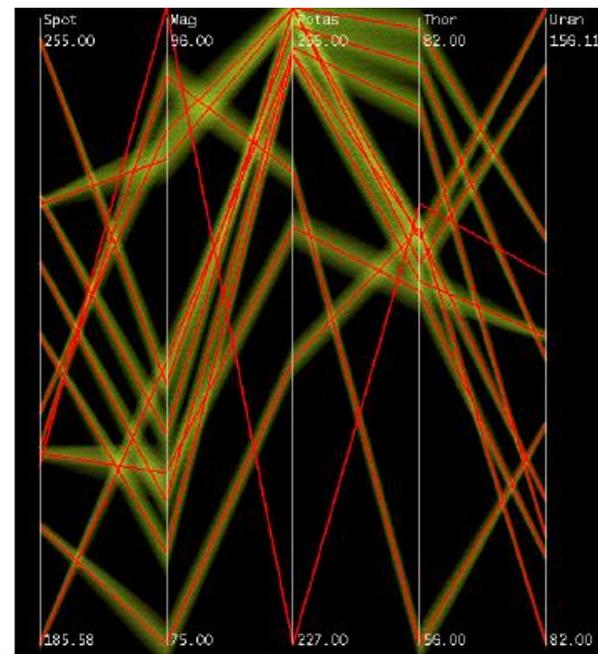
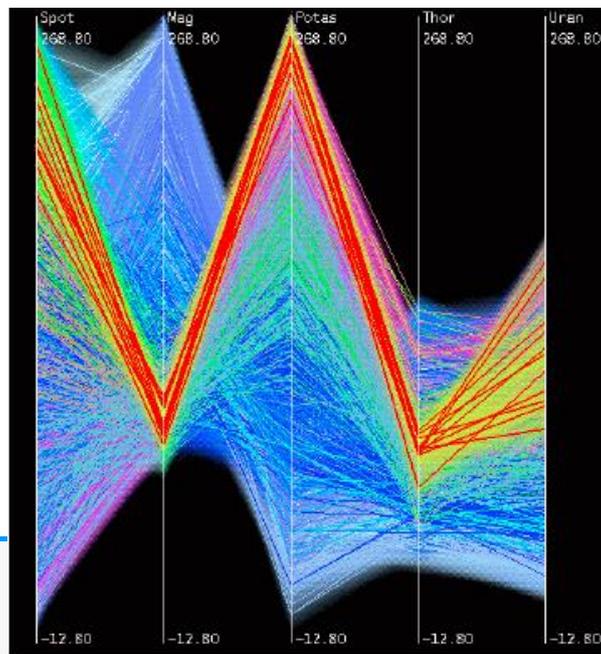
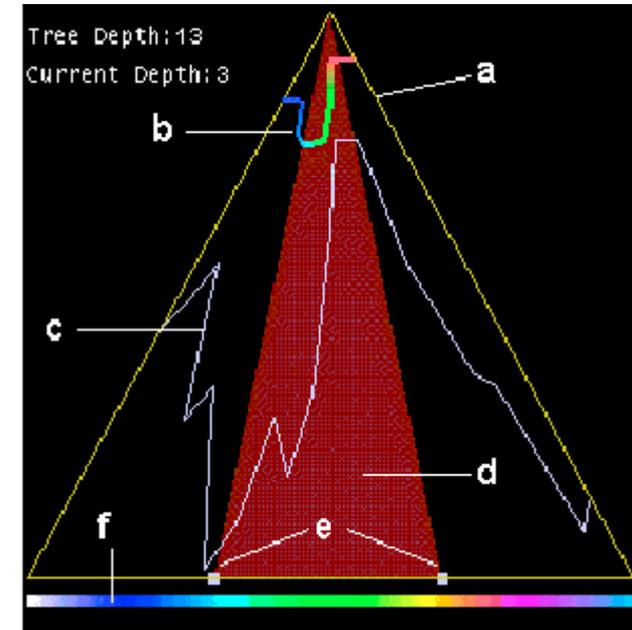
Brushing

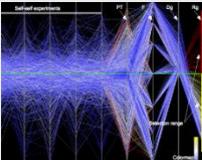
- **Markieren** von bestimmten Datensamples
- Spezifikation **bestimmter Eigenschaften** des Brushes

8.7 Parallele Koordinaten

Brushing [Fua et al. 1999]

- Structure-Based Brushing
 - Auswahl im **Hierarchie-Rahmen**
- Dimension-Zooming
 - Skalieren eines **Teilbereichs einer Dimension**
 - Übersicht mit **Mini-Map** des ausgewählten Bereichs





8.7 Parallele Koordinaten

Angular Brushing [Hauser et al. 2002]

- Standard-Brushing:
 - Auswahl eines **Dimensionsintervalls**
 - **Kombination** von Intervallen (Composite Brush)
- Hier: Auswahl von Dimension und **Steigungsintervall** (Winkel)
- Damit Auswahl der **Art (und Höhe)** der Korrelation
- Einfache Identifikation von
 - **Korrelationsclustern**
 - **Ausreißern** (extreme Winkel zu anderen Elementen)

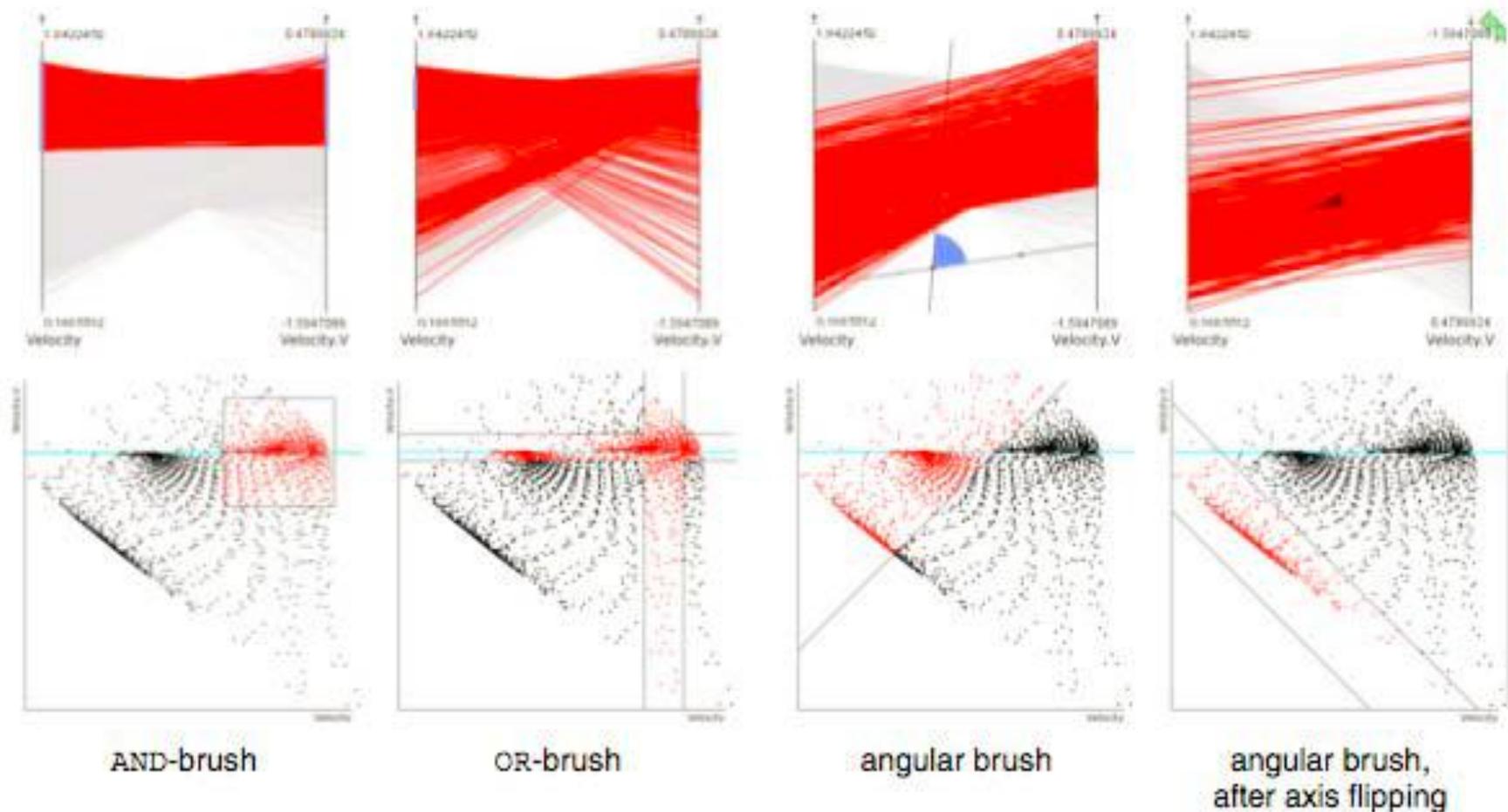
8.7 Parallele Koordinaten

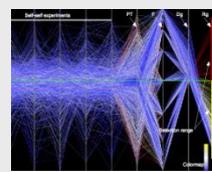
Angular Brushing [Hauser et al. 2002]

- Auswahl von Dimension und **Steigungsintervall** (Winkel)

Composite Brush

Angular Brush

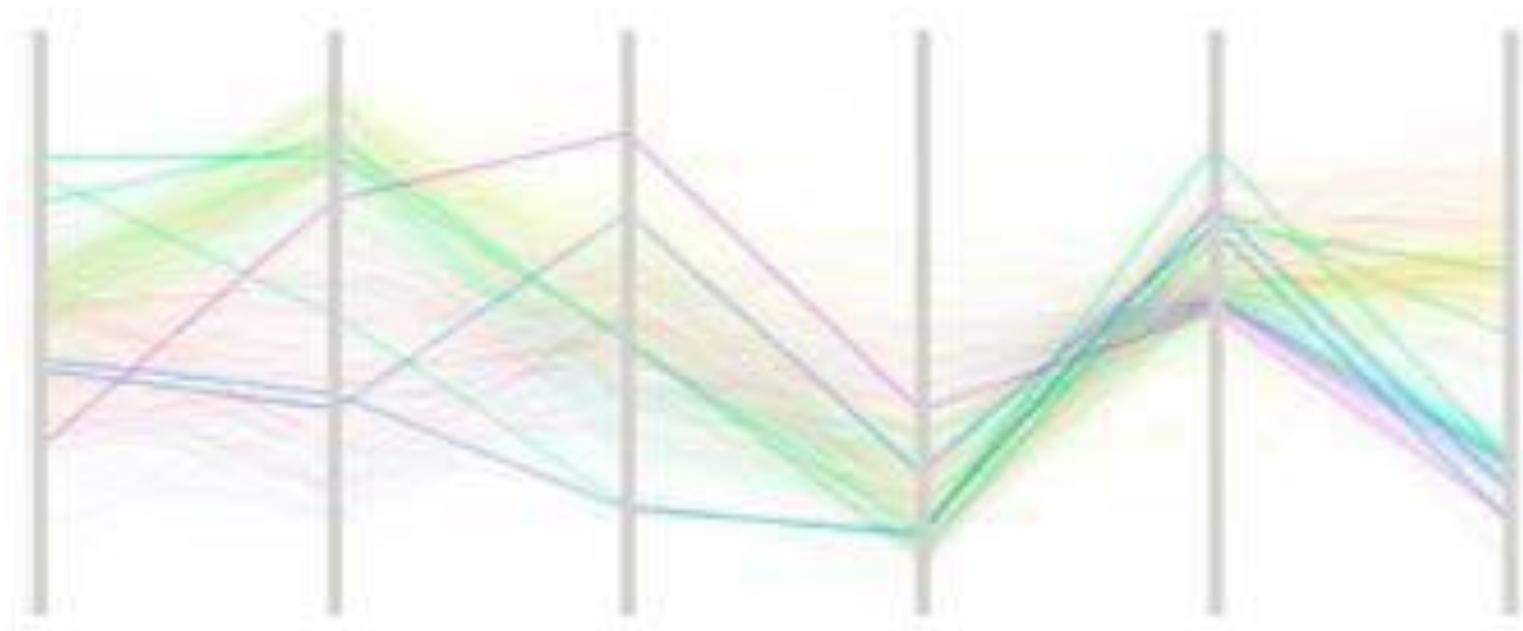




8.7 Parallele Koordinaten

Ausreißer-Behandlung

- Identifikation durch
 - Extreme/andere Korrelationen (Angular Brushing, ...)
 - Datenverteilung (Varianz, Interquartilenintervall, ...)

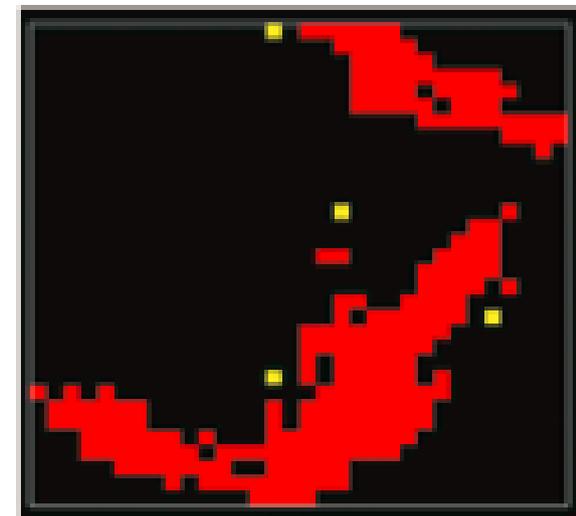
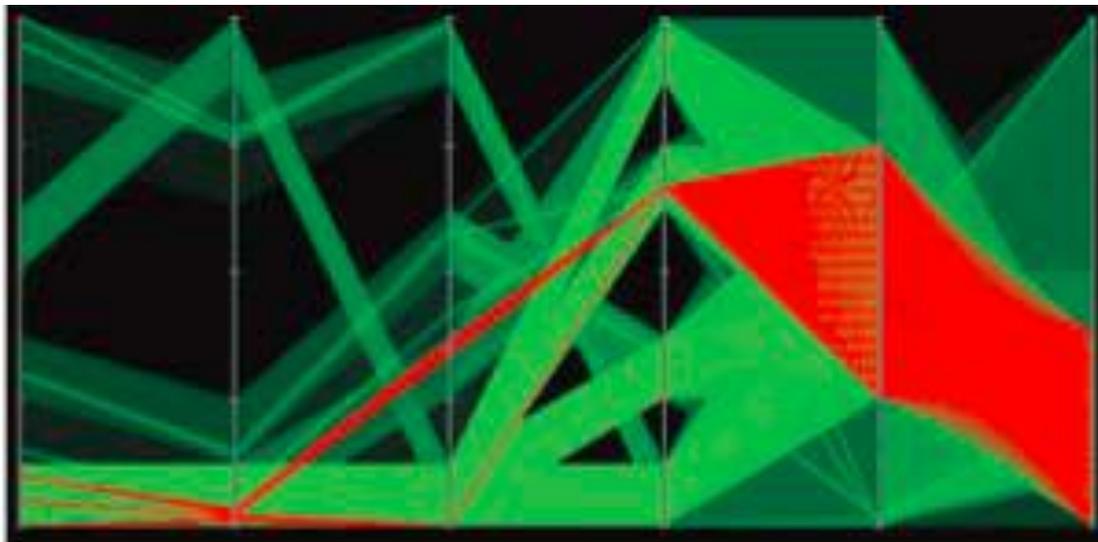


[Johansson et al. 2005]

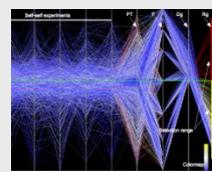
8.7 Parallele Koordinaten

Ausreißer-Behandlung

- Identifikation durch
 - Häufigkeitsklassen werden **LowPass** gefiltert und vom Original abgezogen
 - Dann **leere Klassen im LowPass** sind Ausreißerklassen



[Novotny, Hauser 2005]



Literatur

- Bendix, Kosara, Hauser: Parallel Sets: Visual Analysis of Categorical Data, In *Proc. IEEE Symposium on Information Visualization*, pp. 558-568, 2005.
- Fua, Ward, Rundensteiner: Hierarchical Parallel Coordinates for Exploration of Large Datasets, , In *Proc. IEEE Symposium on Information Visualization*, pp. 43-50, 2004.
- Inselberg: The Plane with Parallel Coordinates, *The Visual Computer*, 1(2):69-92, 1985.
- Inselberg: Parallel Coordinates: A Tool for Visualizing Multi-Dimensional Geometry, In *Proc. of IEEE Visualization*, pp. 361-376, 1990.
- Hauser, Ledermann, Doleisch: Angular Brushing of Extended Parallel Coordinates, In *Proc. IEEE Symposium on Information Visualization*, pp. 127-130, 2002.
- Heinrich: SpRay: Eine Visual-Analytics-Plattform für hochdimensionale Daten, Diplomarbeit Bioinformatik, Universität Tübingen, 2008.
- Johansson, Ljung, Jern, Cooper: Revealing Structure within Clustered Parallel Coordinates Display, In *Proc. IEEE Symposium on Information Visualization*, 2005.
- Novotny, Hauser: Outlier-Preserving Focus+Context Visualization in Parallel Coordinates, In *Proc. IEEE Symposium on Information Visualization*, pp. 893-900, 2006.
- Pradhan, Bartz, Mueller: SignatureSpace: A Multidimensional Exploratory Approach for the Analysis of Volme Data, In *WSI-TechReports*, WSI-2005-11, 2005.
- Wegman: Hyperdimensional Data Analysis Using Parallel Coordinates, *Journal of the American Statistical Association*, 85(411):664-675, 1990.
- Wegman, Luo: High Dimensional Clustering Using Parallel Coordinates and the Grand Tour, *Computing Science and Statistics*, 28:361-368, 1997.
- Yang, Peng, Ward, Rundensteiner: Interactive Hierarchical Dimension Ordering, Spacing and Filtering for Exploration of High Dimensional Datasets, In *Proc. IEEE Symposium on Information Visualization*, pp. 105-112, 2003.