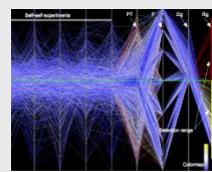


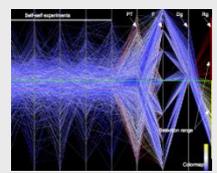
# Informations- visualisierung

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



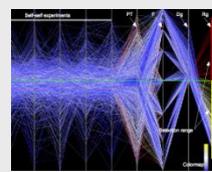
# Wiederholung

- Was sind TreeMaps und wie funktionieren sie?
- Wofür sind sie besonders gut geeignet?
- Was sind ihre Nachteile?
- Wie kann der Elementrand eingespart werden?
- Wie funktionieren Cushion-TreeMaps?
- Wie kann die Darstellung der Hierarchie verbessert werden?
- Was gibt es für Ansätze zur Verbesserung des Aspektratio?
- Was ist der wesentliche Unterschied zu Slice-and-Dice?
- Wie kann die Ordnung der Elemente erhalten bleiben?
- Wie funktionieren radiale Layouts?



# Wiederholung

- Was sind Mosaics?
- Wie teilen sie die Daten auf?
- Was ist Dimensional Stacking?



## 8. Spezifische Verfahren

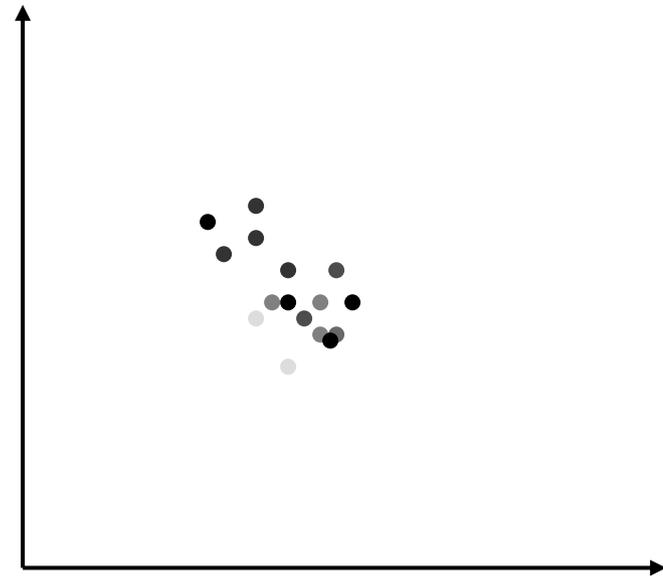
### Verfahren für die Darstellung von Tabellen

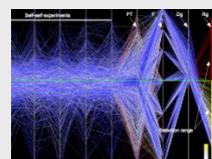
- Histogramme
- Scatterplots
- Treemap & Mosaics
- Stacked Display / Dimensions
- **Scatterplot Matrix**
- **Parallele Koordinaten**

## 8.6 Scatterplot-Matrizen

### Wiederholung Scatterplot

- Bildet **zwei Variablen / Merkmale** auf Abszisse und Ordinate
- Ist eine **Projektion**
- Häufigkeit wird als **Dichte, Opazität, oder Luminanz** abgebildet
- Pseudokontinuierlich
- Potentielle **Aliasing**-Probleme

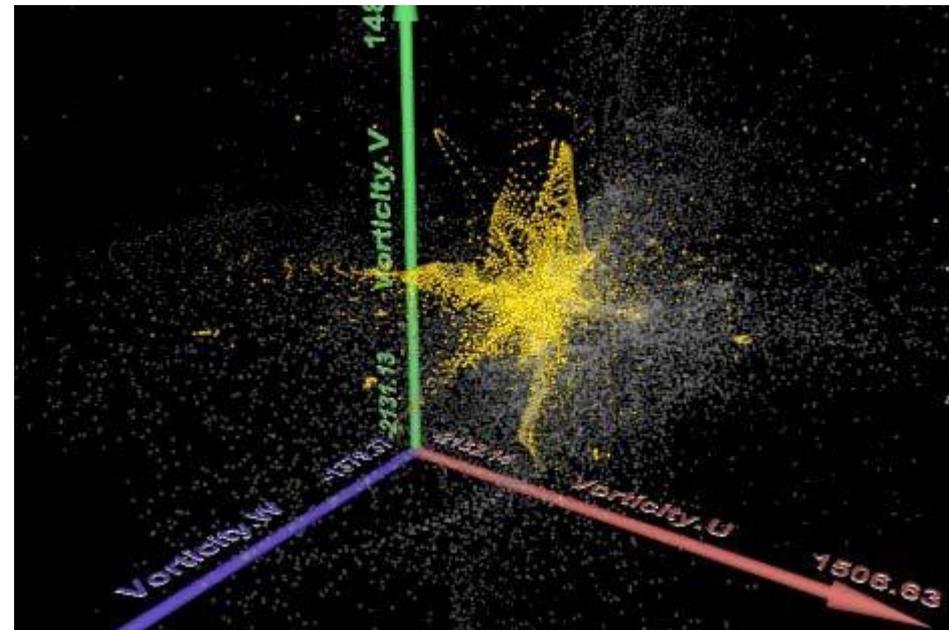




## 8.6 Scatterplot-Matrizen

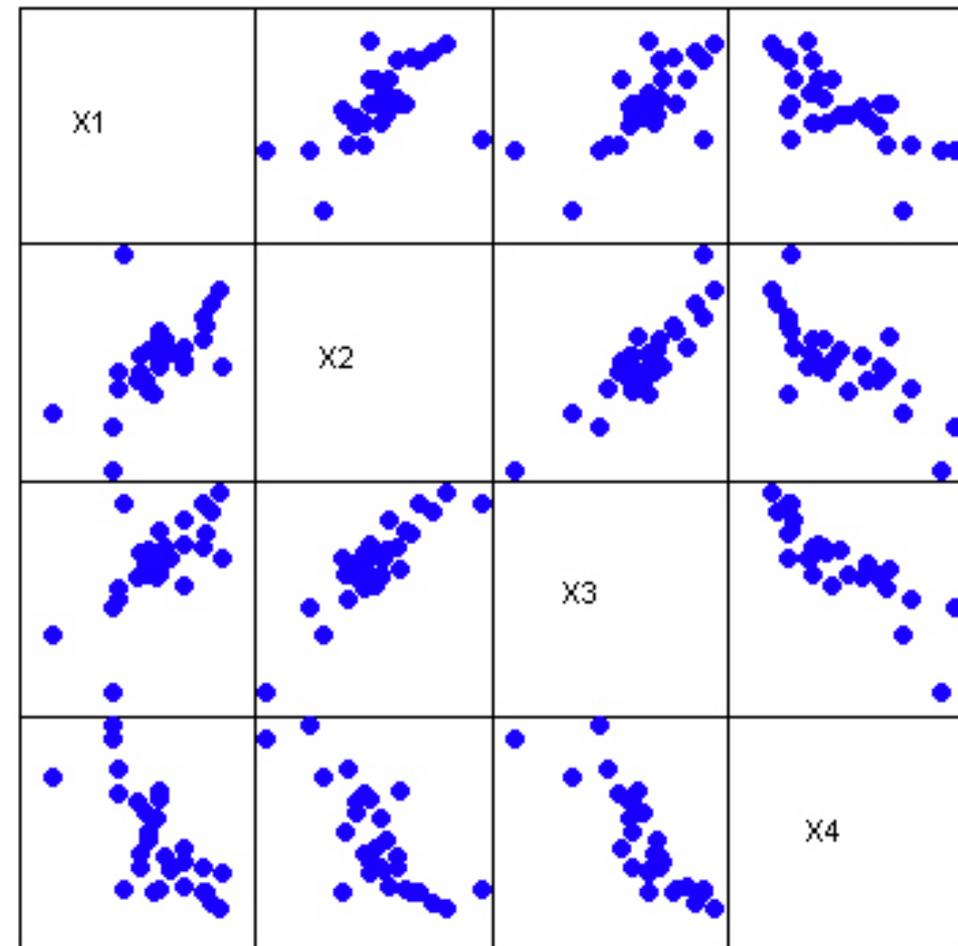
### Scatterplot: Integration von mehr Variablen

- 3D-Scatterplots
- Integration weiterer Variable in **3. Dimension**
- Erhebliches **Verdeckungsproblem**
- **Räumliches** Verständnis schwierig
- **Brushing schwierig**
- **Erfordert Tiefensortierung** für 3D und ggf. Alphablending



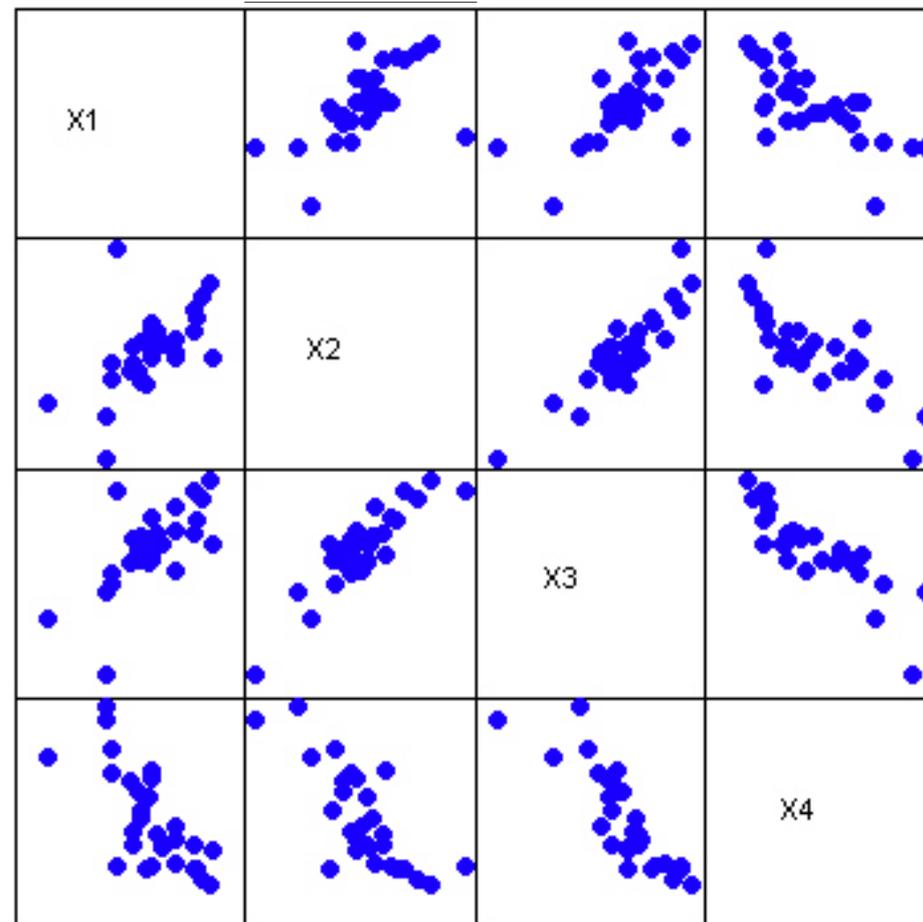
## 8.6 Scatterplot-Matrizen

- [Hartigan 1975, Chambers 1983]
- Auch Draftsman-Plot oder **Panel**darstellung/-matrix oder SPLOM
- Integration von mehr Variablen
- Ist eine **Projektion**
- **Paarweise Kombination** aller Parameter
- **Label** oder **Histogramme** (univariates Scatterplot) auf der Diagonalen
- Obere und untere **Diagonale** sind nur gespiegelt
- Legende oft nur **für jede zweite** Zeile/Spalte



## 8.6 Scatterplot-Matrizen

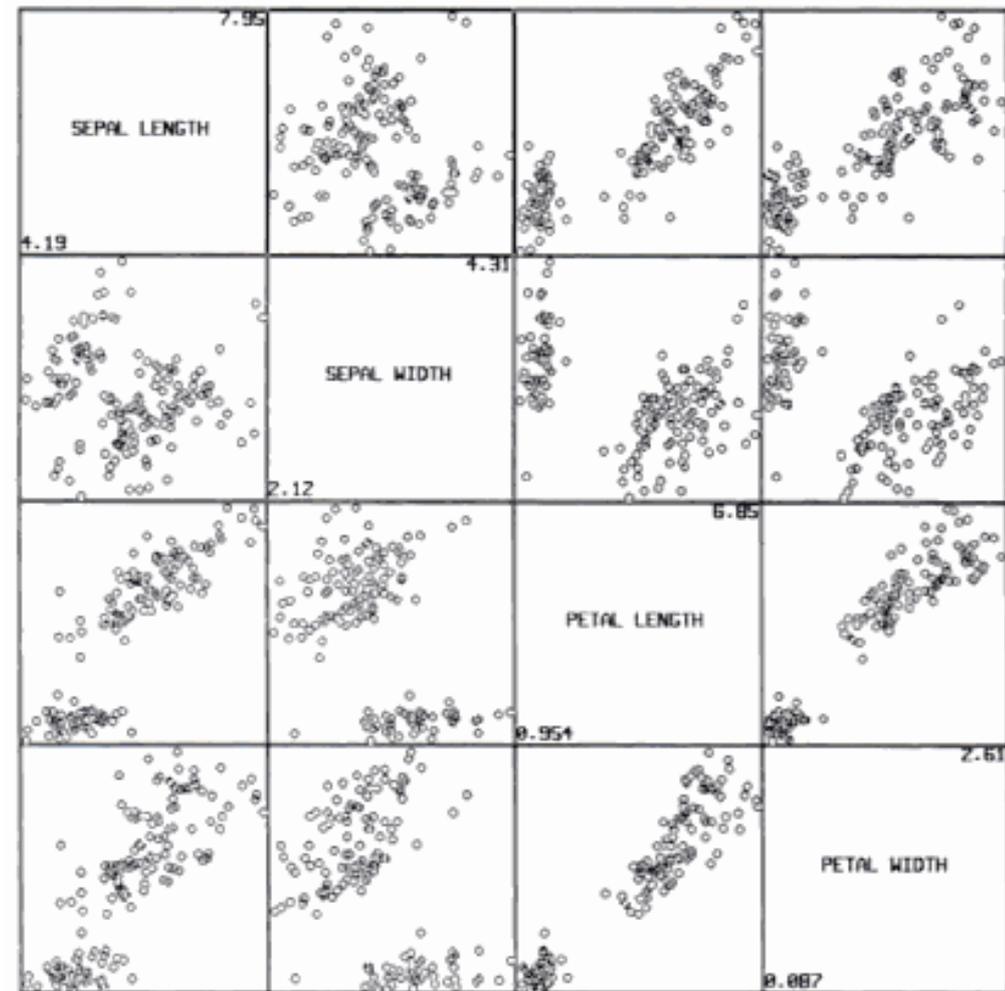
- Innerhalb einer Zeile/Spalte ändert sich **nur ein Parameter**
- **Keine** erzwungene **Ordnung** von Zeilen/Spalten
- **Schwierige** Interpretation von mehreren Parametern
- Problematisch bei **hoher Dimensionszahl**
- Varianten
  - Hyperslice
  - Prosection View



## 8.6 Scatterplot-Matrizen

### Interpretation der Scatterplot-Matrizen

- Schwertlilien: {Kelchkrone , Blütenblatt } x {Länge , Breite}
- Wie **entsprechen sich** Cluster/Daten?
- Einzelrotationen **erhalten Daten** teilweise

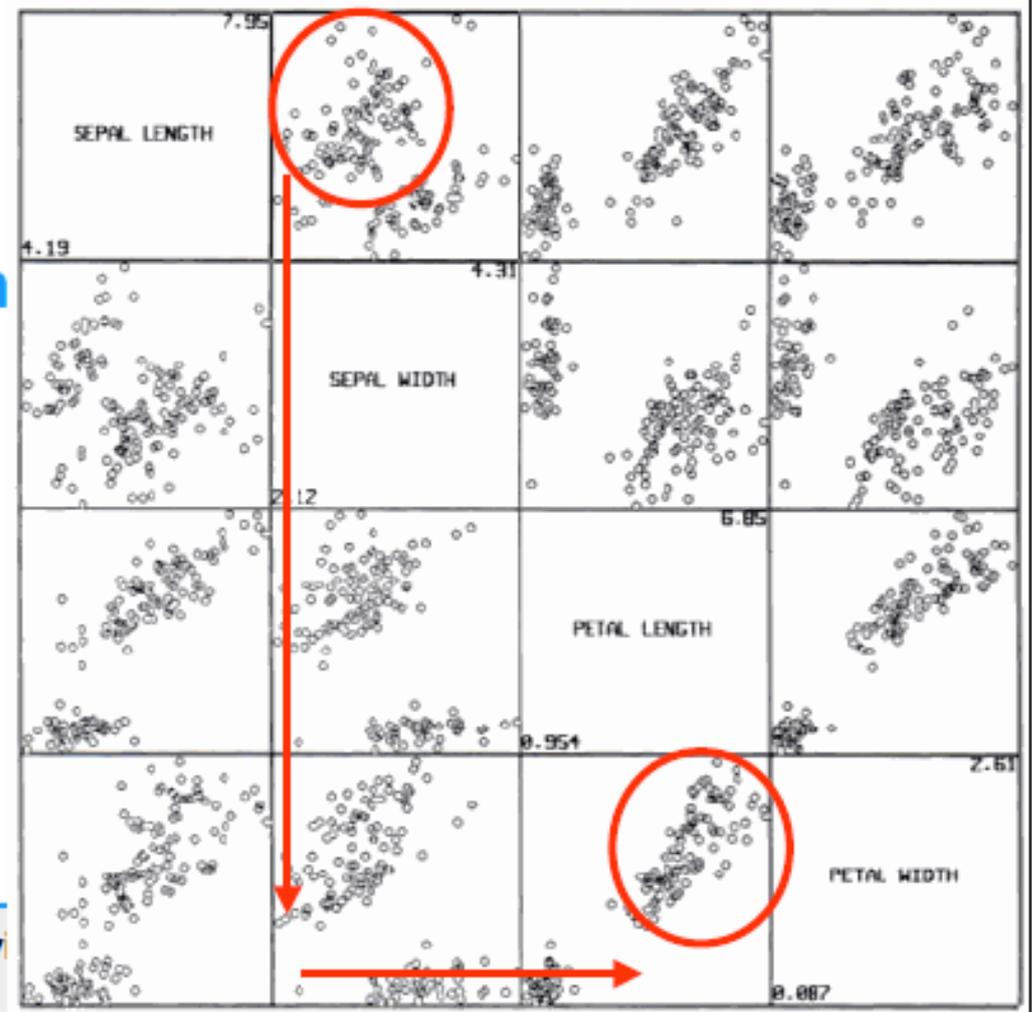


[Cleveland, McGill 1988]

## 8.6 Scatterplot-Matrizen

### Interpretation der Scatterplot-Matrizen

- Schwertlilien: {Kelchkrone , Blütenblatt } x {Länge , Breite}
- Wie **entsprechen sich** Cluster/Daten?
- Einzelrotationen **erhalten Daten** teilweise



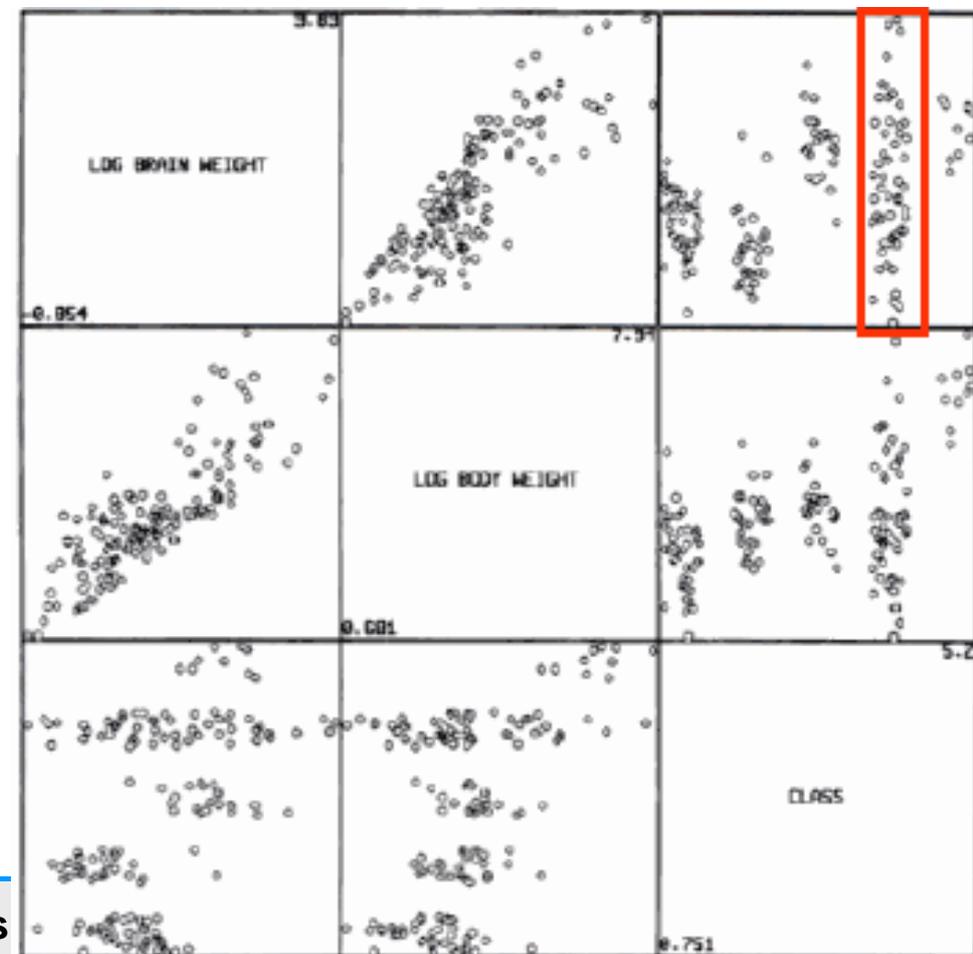
[Cleveland, McGill 1988]

Informationen

## 8.6 Scatterplot-Matrizen

### Interpretation der Scatterplot-Matrizen

- Gehirngewicht x Körpergewicht x Klasse (Vögel, Fische, Primaten, Nicht-Primat-Säuger, Dinosaurier)
- Brushing **nicht in allen Scatterplots** möglich
- **Kategorische** Daten können hilfreich sein
- Manchmal **Jitter** hinzufügen

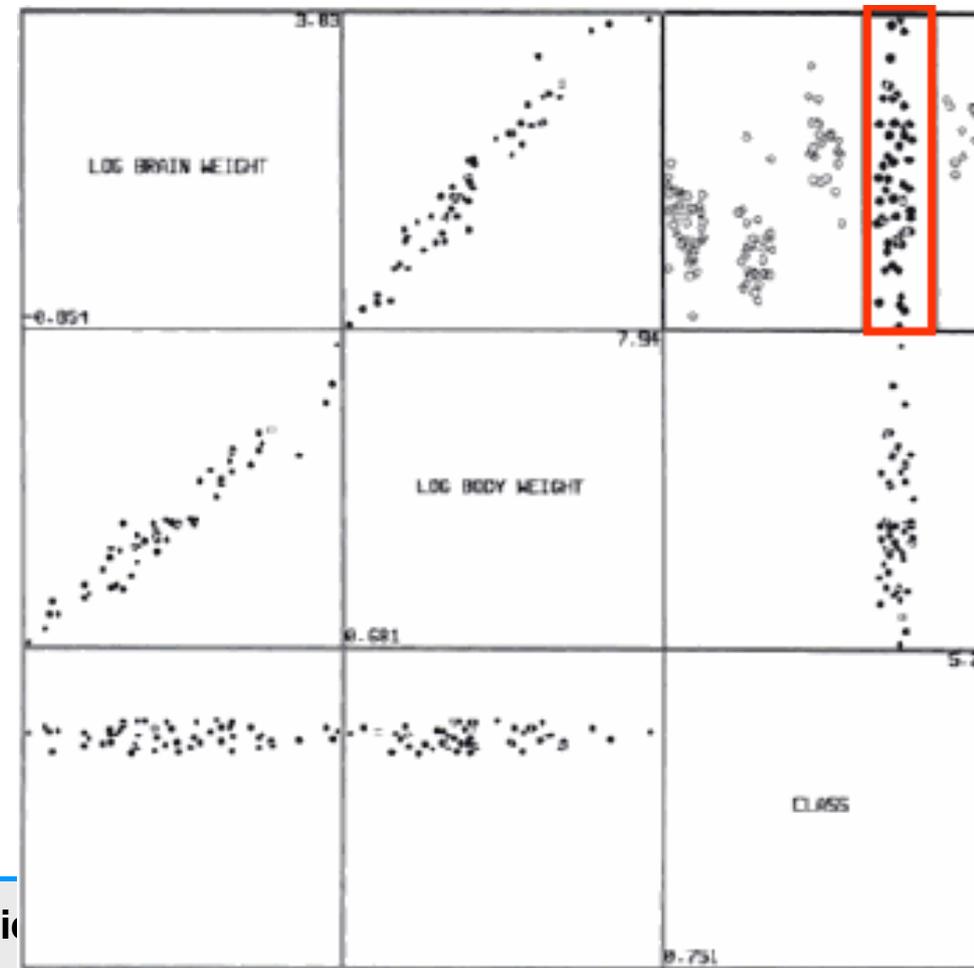


[Cleveland, McGill 1988]

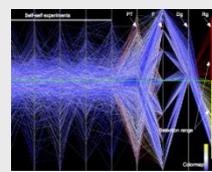
## 8.6 Scatterplot-Matrizen

### Interpretation der Scatterplot-Matrizen

- Gehirngewicht x Körpergewicht x Klasse (Vögel, Fische, Primaten, **Nicht-Primat-Säuger**, Dinosaurier)
- Brushing **nicht in allen Scatterplots** möglich
- **Kategorische** Daten können hilfreich sein



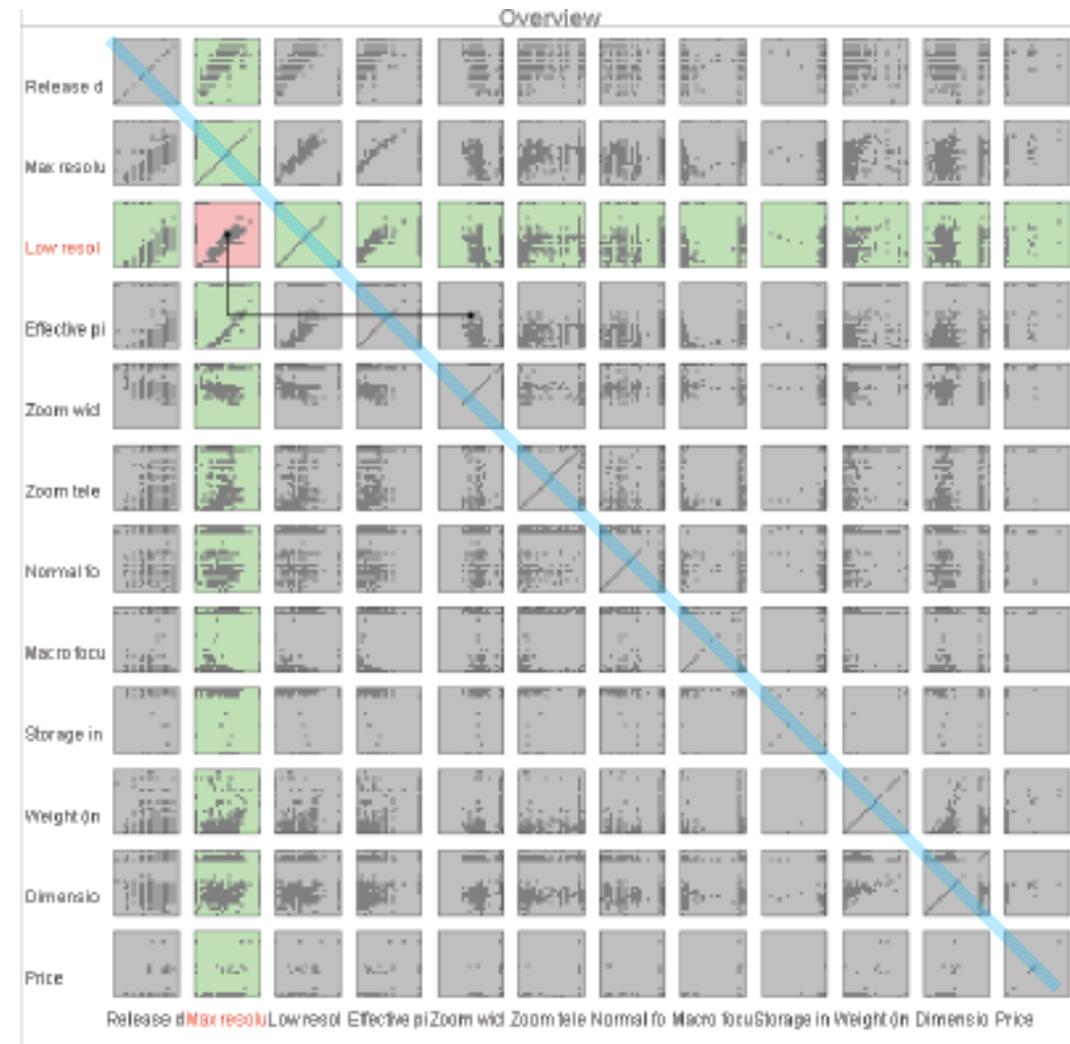
[Cleveland, McGill 1988]

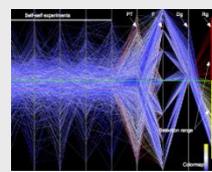


## 8.6 Scatterplot-Matrizen

### Probleme bei hoher Dimensionalität

- Kleiner **Plotbereich**
- Anzahl von  $p(p-1)/2$  Scatterplots **steigt zu schnell** mit  $p$
- Zusammenhang bleibt oft unklar

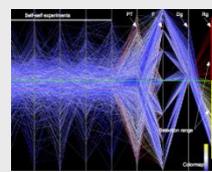




## 8.6 Scatterplot-Matrizen

### Dimensionsreduktion

- Reduktion der Dimensionalität der Scatterplot-Matrix durch **Bestimmung relevanter Dimensionen**
- Verwendung von „**Scagnostics**“  
[Tukey, Tukey 1985, Wilkinson et al. 2004]
- **Berechnung von k „geeigneten“** Scatterplots durch neue Kriterien
- **Neusortierung** / Beschränkung der Matrix entsprechend der Auswahl
- Reduziert  $O(p^2)$  zu  $O(k^2)$ , mit  $k \ll p$
- Beruht auf **Graphentheorie**
- **Vereinfachung** der z.T. komplexen Berechnungen

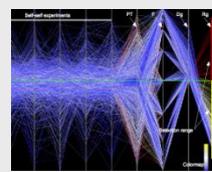


## 8.6 Scatterplot-Matrizen

### Scagnostics

Suche nach [Wilkinson et al. 2004]

- Ausreißer: Entfernung im Minimum Spanning Tree (MST)
- Form:
  - Konvexität (Alpha-Hülle zu Konvexe Hülle)
  - Magerkeit (Umfang zu Fläche)
  - Sehnigkeit (Pfadähnlichkeit: Breite/Länge des MST)
  - Direktheit (Pfadlänge zu Distanz)
- Trends: Monotonie (Korrelation)
- Dichte
  - Schrägheit
  - Klumpigkeit
- Kohärenz: Rifflichkeit (Glattheit der Pfade im MST)

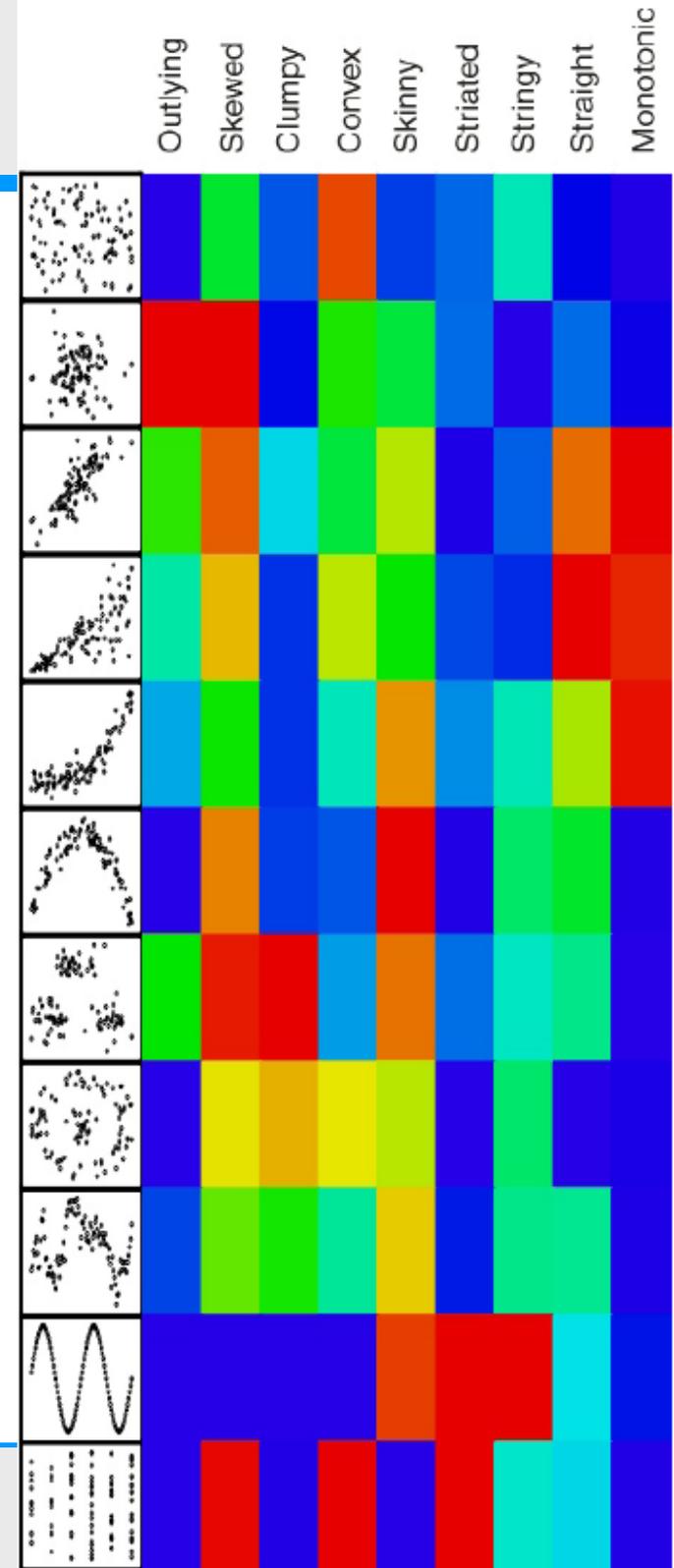


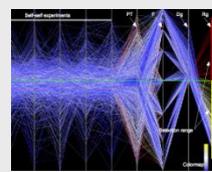
## 8.6 Scatterplot-Matrizen

### Scagnostics

[Wilkinson et al. 2004]

- Vergleich der Metriken
  - Blau: niedrig
  - Rot: hoch





# 8.6 Scatterplot-Matrizen

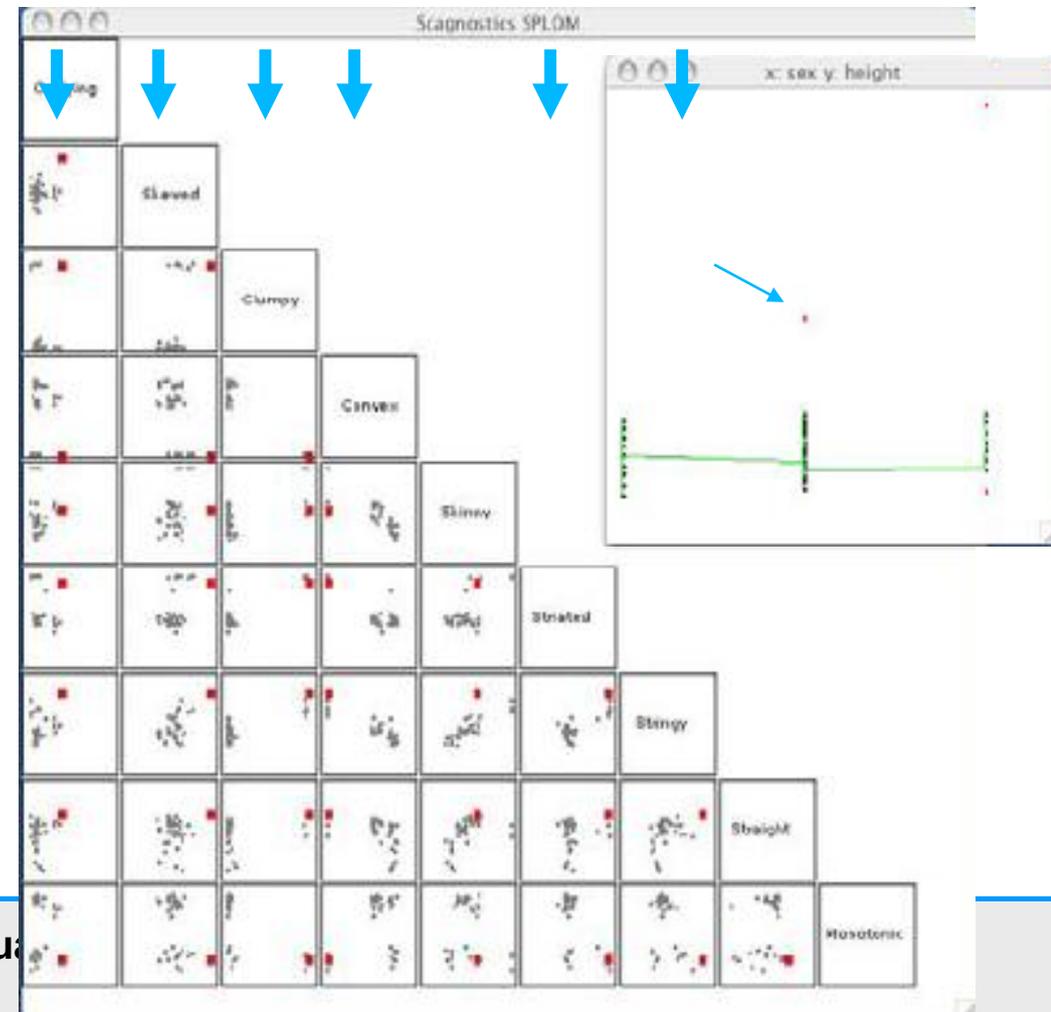
## Scagnostics

[Wilkinson et al. 2004]

- Vergleich mit Standard Scatterplot-Matrix
- Ohrschnecken-Datensatz
- **Ausreißer**-Erkennung



visua

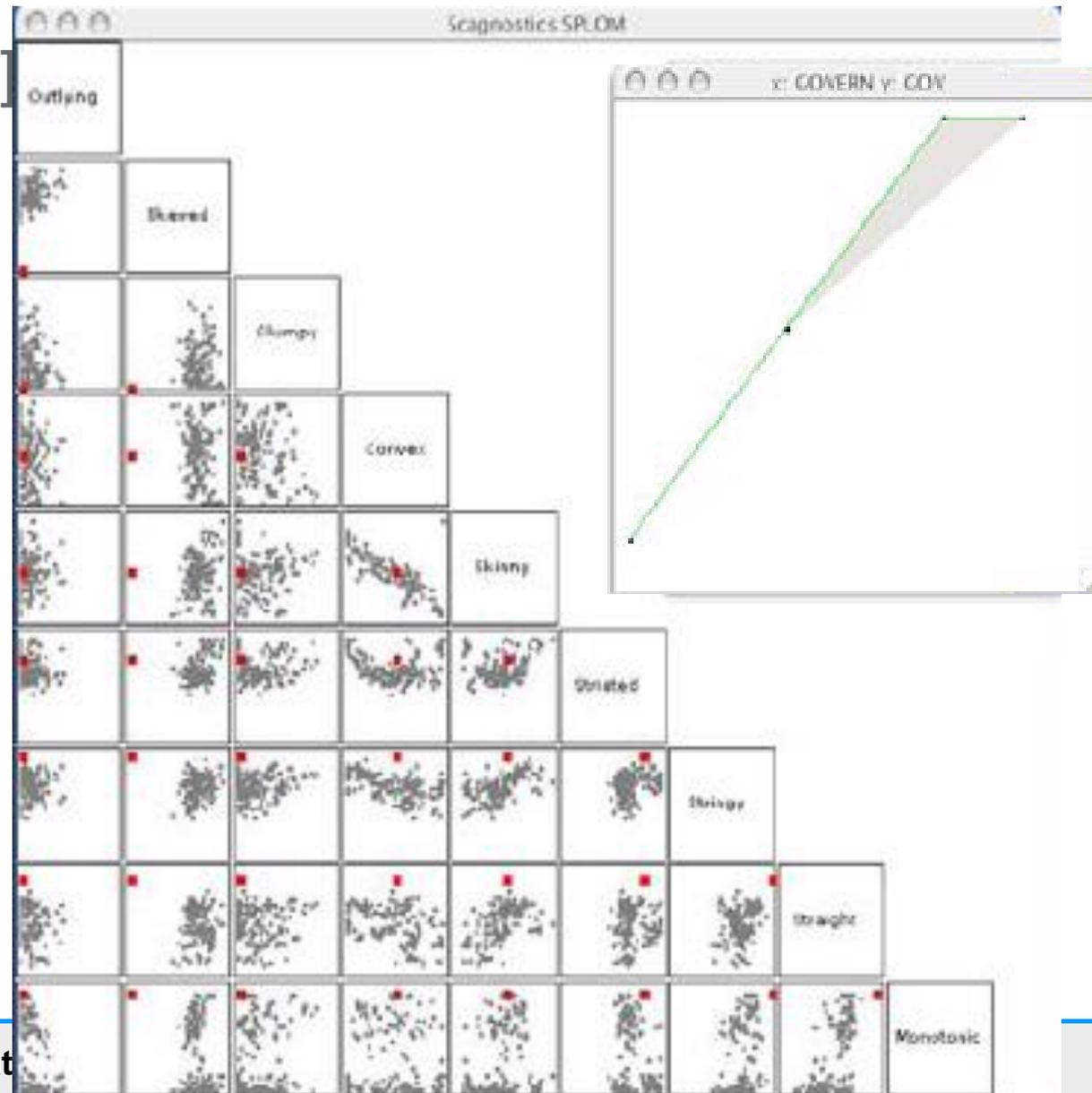


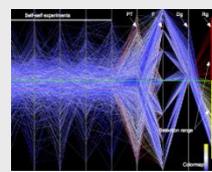
## 8.6 Scatterplot-Matrizen

### Scagnostics

[Wilkinson et al. 2004]

- WHO Länderliste
- Zwei Arten von Regierungsform (**Singularität**)





## 8.6 Scatterplot-Matrizen

### Scagnostics

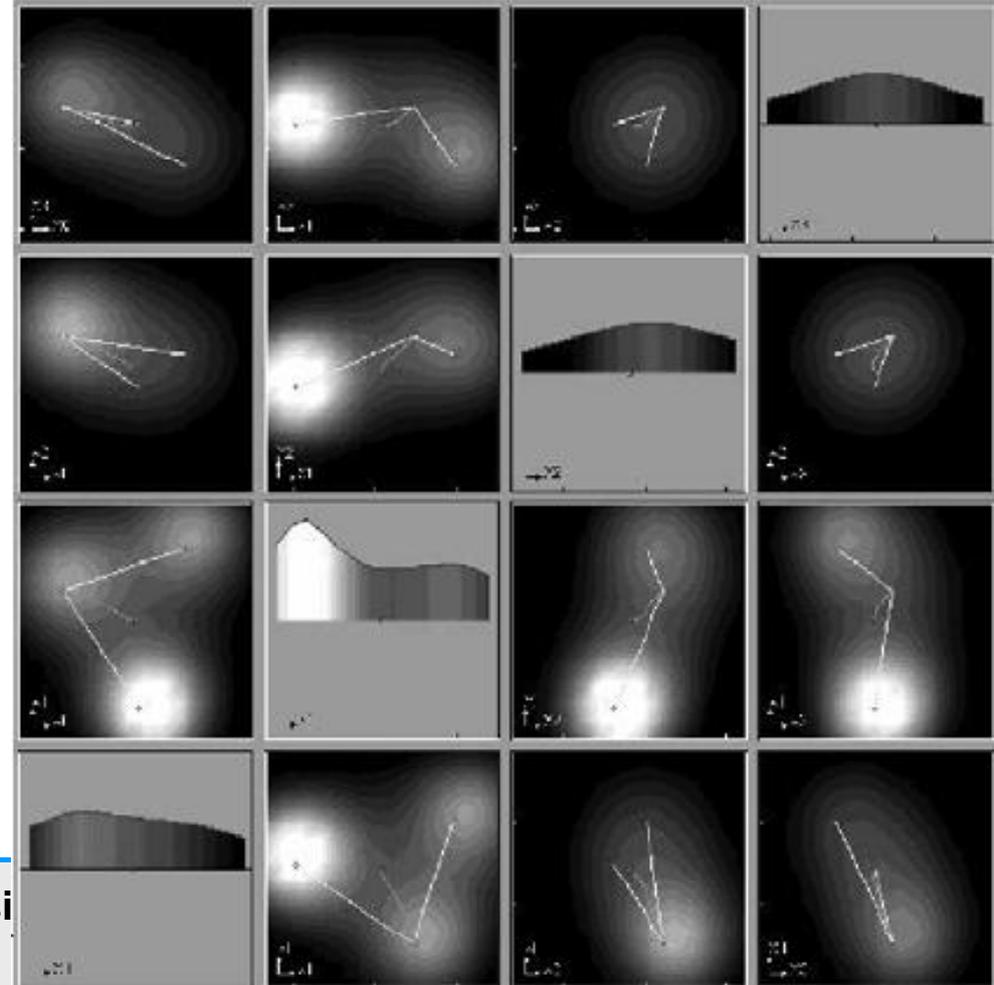
[Wilkinson et al. 2004]

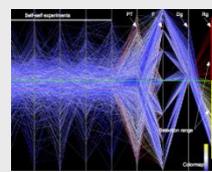
- Hängt stark von der **Qualität der Metriken** ab
- Keine Garantie, dass Metriken **relevante Werte gut abbilden**
- **Kein klarer Vorteil** gegenüber anderen dimensions-reduzierenden Verfahren auf Daten

## 8.6 Scatterplot-Matrizen

### Hyperslices [van Wijk, van Liere 1993]

- **Ähnlich** wie Scatterplot-Matrix
- **Beliebige Schnitte** durch Merkmalsraum
- **Selektion**, keine Projektion
- **Schnitte schneiden** sich im Focal Point/Current Point
- Hier 4D-Potentialfunktion

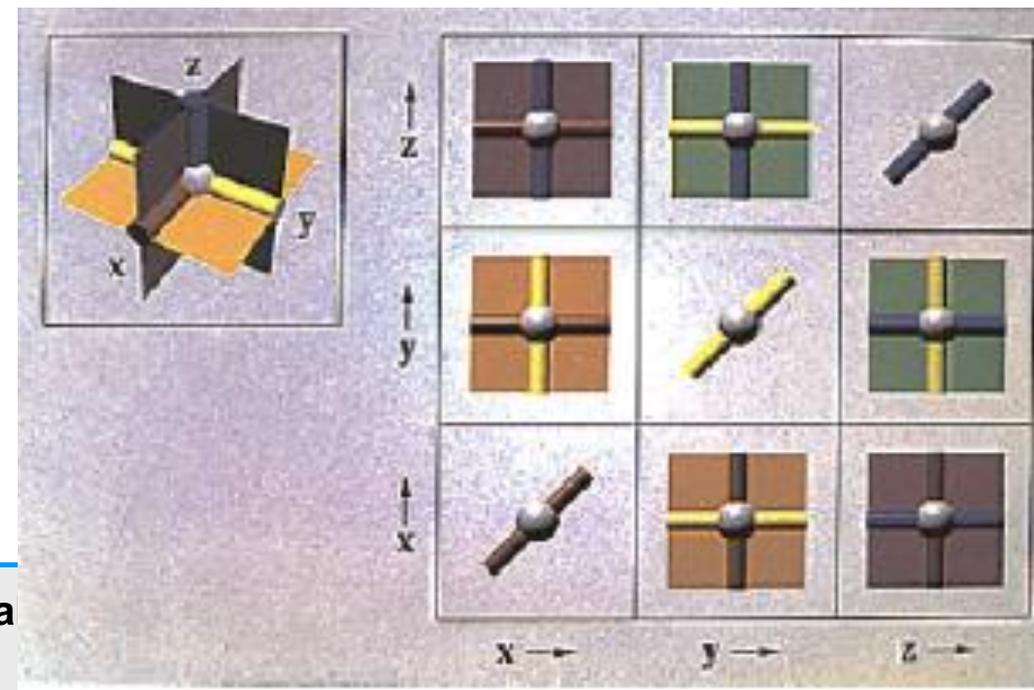


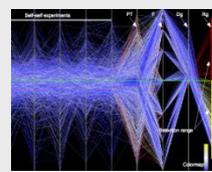


## 8.6 Scatterplot-Matrizen

### Hyperslices [van Wijk 1993]

- **Ähnlich** wie Scatterplot-Matrix
- **Oblique Schnitte** durch Merkmalsraum
- **Selektion**, keine Projektion
- **Schnitte schneiden** sich im Focal Point/Current Point





## 8.6 Scatterplot-Matrizen

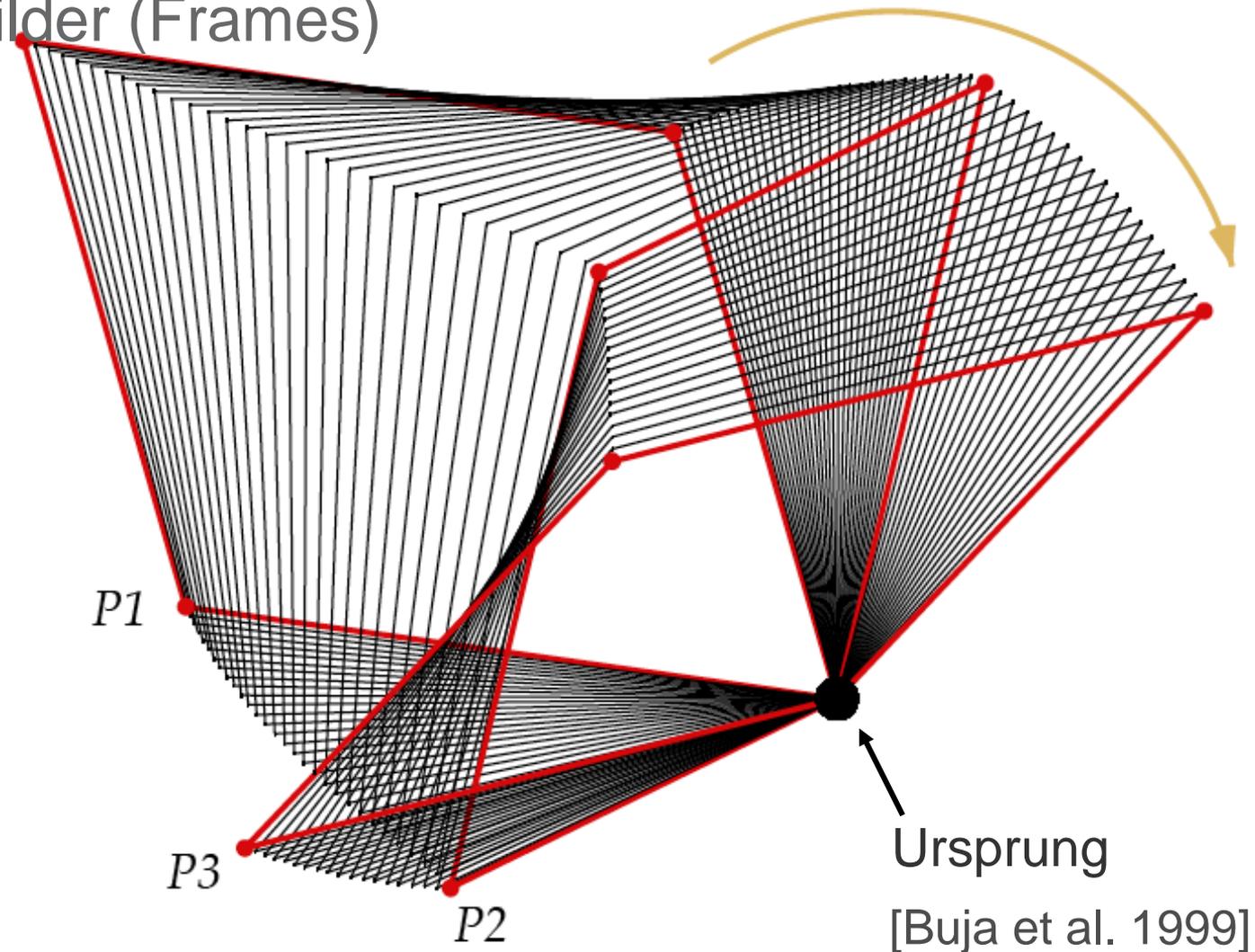
### Animationen der Scatterplot-Matrix

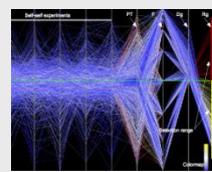
- GrandTour-Ansatz [Asimov 1985, Buja, Asimov 1986]
- Projektion der Scatterplot-Matrix
- Startet mit drei Dimensionen und **rotiert nacheinander** in neue Projektionsebenen
- Fügt weitere, **nicht verwendete Ebenen hinzu** bis alle Dimensionen durchlaufen sind
- Rotation ist **animierte Projektion auf Hyperebene** zwischen den Zieldimensionen
- Interpolationsparameter **als Zeitachse** der Animation

## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix

- GrandTour-Ansatz
- Drei Zwischenbilder (Frames) der Animation

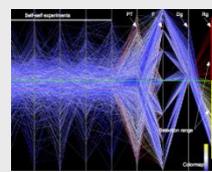




## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix

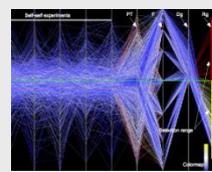
- GrandTour-Ansatz [Asimov 1985, Buja, Asimov 1986]
- Durch **Rotationen** wird räumliche Ausdehnung klarer
- **Veränderung** der Daten wird klarer
- **Werte** lassen sich jedoch **schlecht ablesen**
- **Zusammenhang zwischen Rotationen** schwierig erfassbar
- Bei großer Anzahl von Dimensionen wird Durchlaufen aller Möglichkeiten **langwierig**
- Rotation **eher zufällig** (kann aber kontrolliert werden)



## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix

- GrandTour-Ansatz [Asimov 1985, Buja, Asimov 1986]
- Kombiniert mit dem **Projection-Pursuit** [Huber 1985, Cook et al. 1995]
  - Bestimmt **PP Index**
  - Optimiert PPI entsprechend **verschiedener Metriken**
  - Animiert entsprechend des PPI



## 8.6 Scatterplot-Matrizen

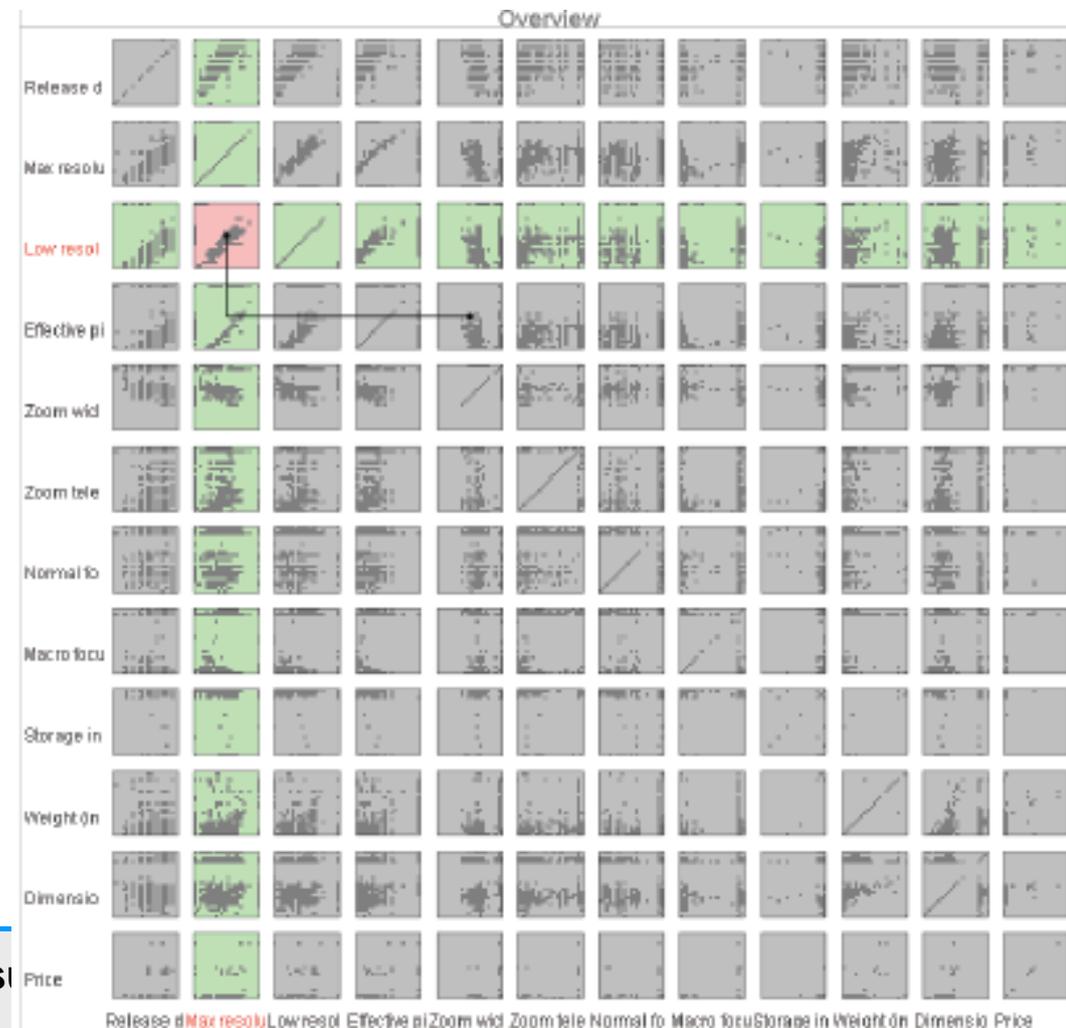
### Animationen der Scatterplot-Matrix

- **ScatterDice** [Elmqvist et al. 2008]
- **Nutzt Eigenschaften** der Scatterplot-Matrix aus
  - **Neuordnen** der Spalten und Zeilen möglich
  - **Veränderung** zu benachbarten Zellen überschaubar (eine Dimension)
- Nutzt Scatterplot-Matrix-Raum als **Navigationsraum**
- Überträgt Visualisierungsproblem auf ein Navigationsproblem

## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

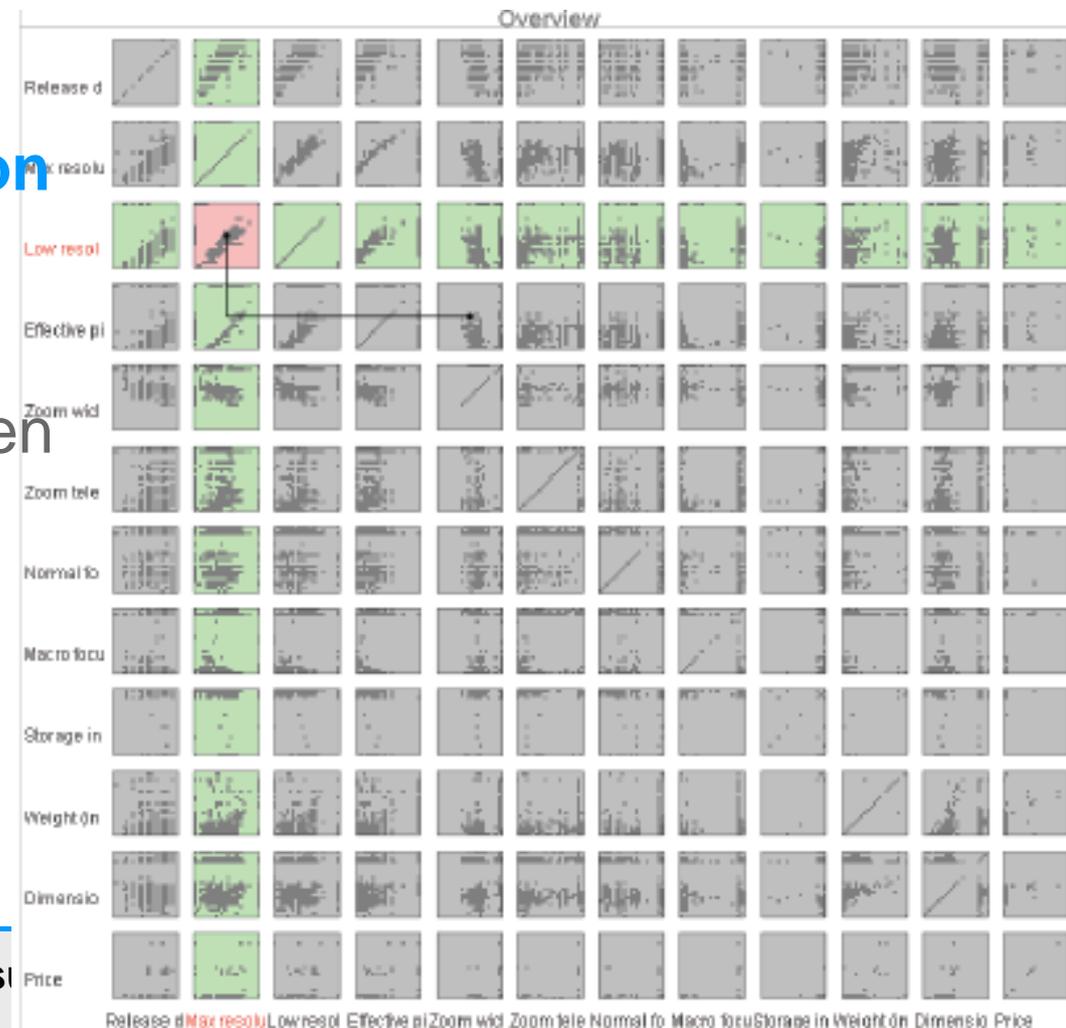
- **ScatterDice** [Elmqvist et al. 2008]
- Wie muss Transition aussehen?
  - Lineare Interpolation **schwierig zu interpretieren**
  - Keine klare Semantik

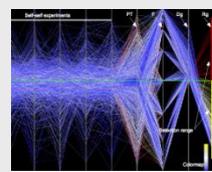


## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

- **ScatterDice** [Elmqvist et al. 2008]
- Wie muss Transition aussehen?
  - Nutzung der **3. Dimension**
  - **Metapher**: Rollender Würfel
  - Nur **Rechteck**stransitionen
  - **Schrittweise** Transition
- [Cleveland, McGill]

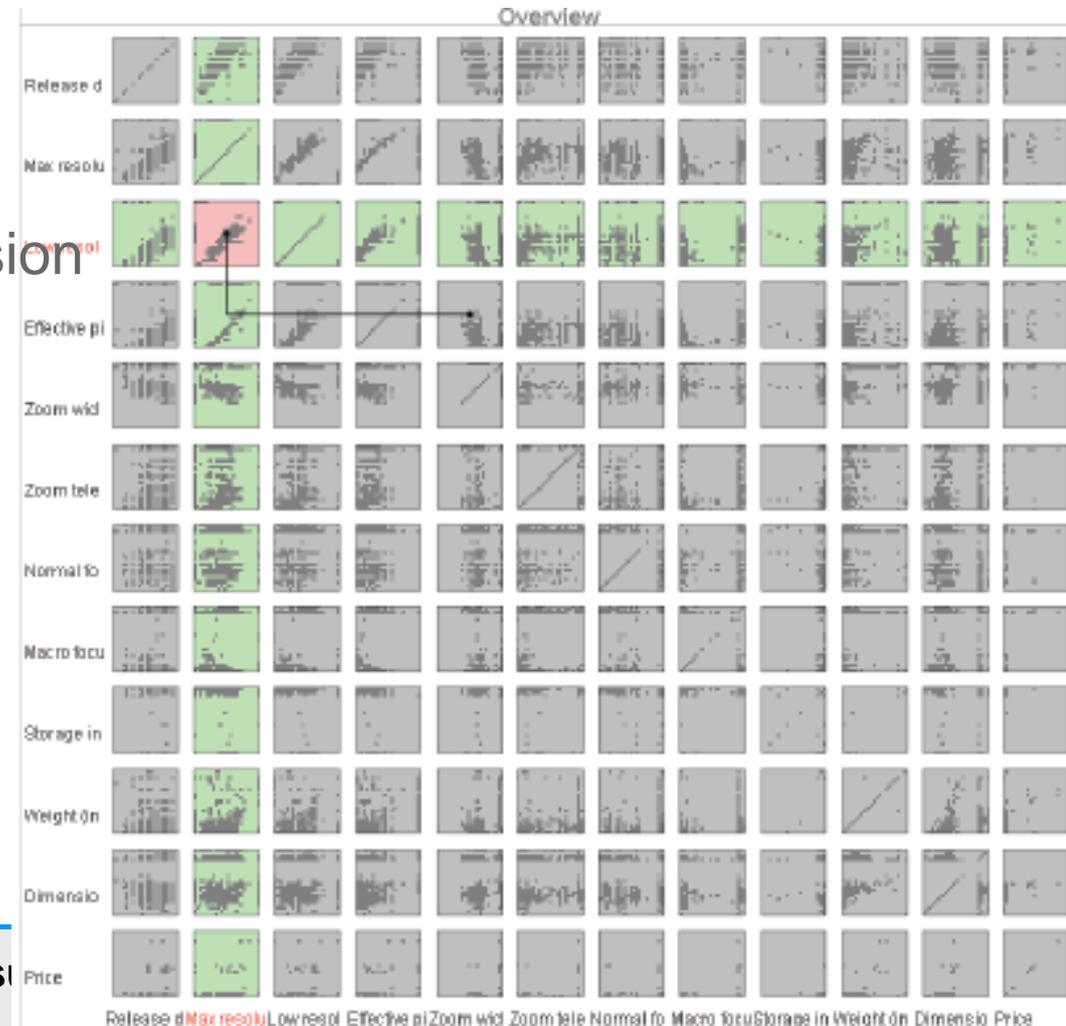


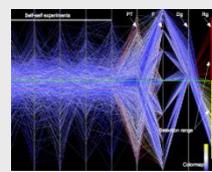


## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

- **ScatterDice** [Elmqvist et al. 2008]
- Wie muss Transition aussehen?
  - Schrittweise Transition
    - **Aufblasen** in 3. Dimension
    - **Transition** (Rotation)
    - **Projektion** in 2D
  - **Perspektive** Projektion (drei Schritte)
  - Orthogonale Projektion ist **schwieriger zu interpretieren**

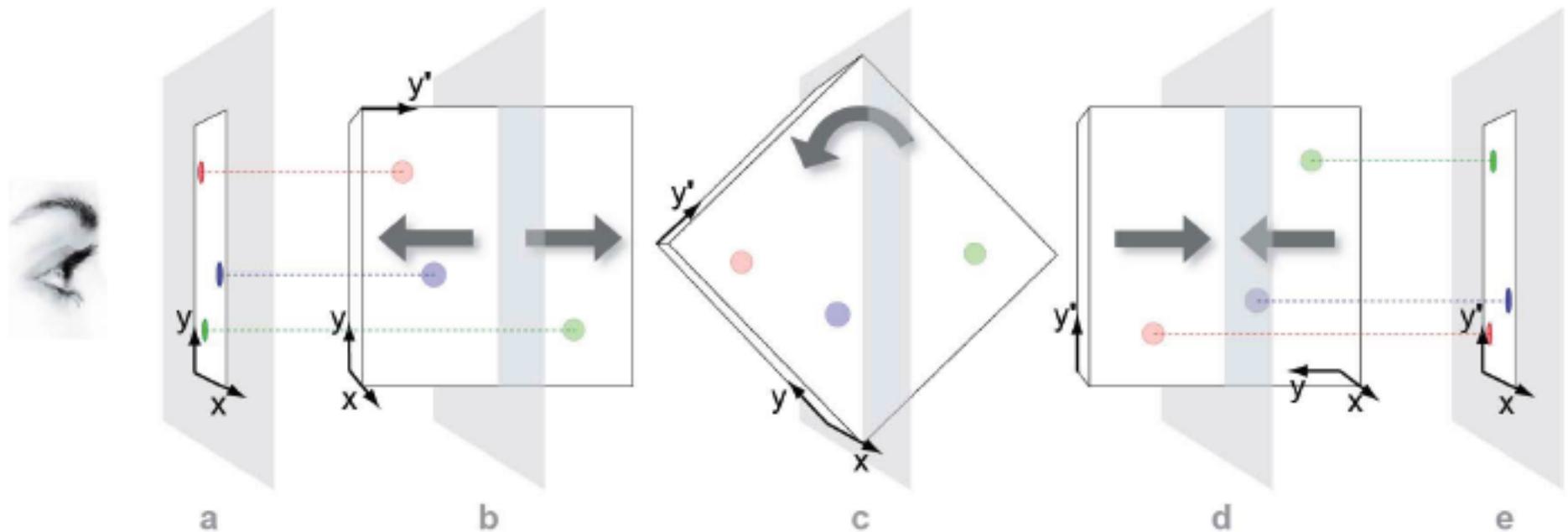


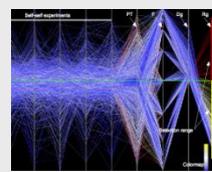


## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

- **ScatterDice** [Elmqvist et al. 2008]
- 3D **perspektivische** Transition





## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

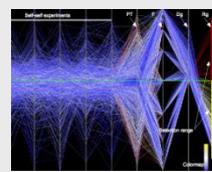
- **ScatterDice** [Elmqvist et al. 2008]

**Rolling the Dice**

Multidimensional Visual Exploration  
using Scatterplot Matrix Navigation

Niklas Elmqvist  
Pierre Dragicevic  
Jean-Daniel Fekete

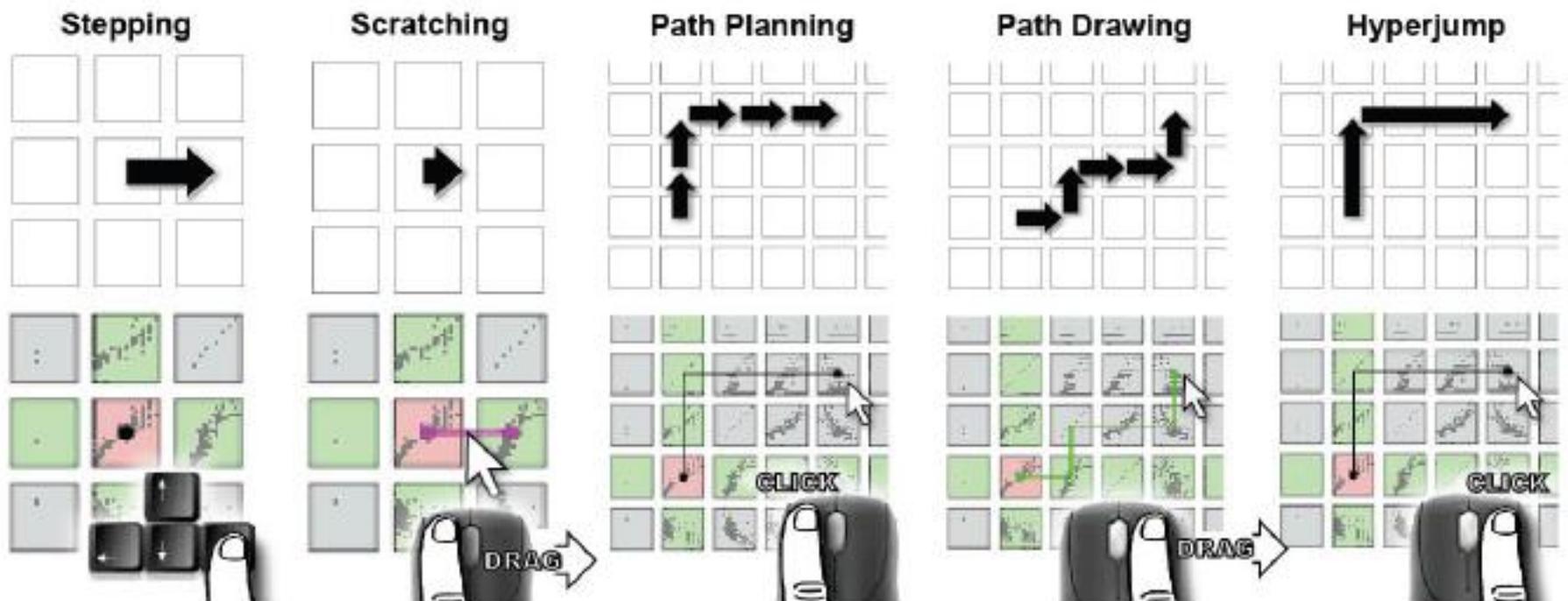
INRIA



## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

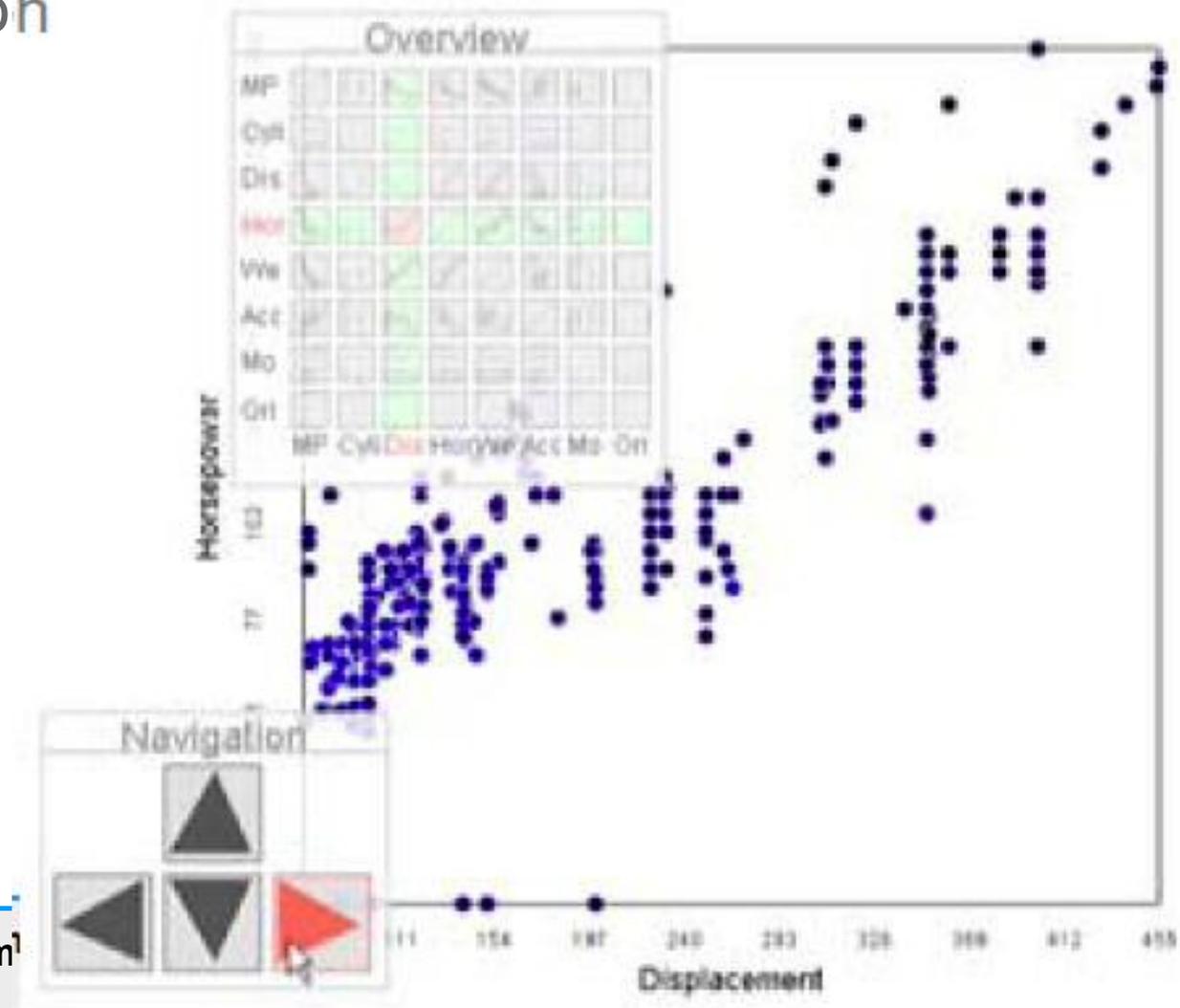
- **ScatterDice** [Elmqvist et al. 2008]
- Interaktion/Navigation
  - **Neuordnen** der Dimensionen
  - **Navigation**

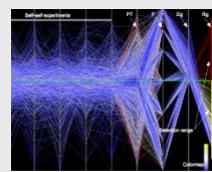


## 8.6 Scatterplot-Matrizen

### Animationen der Scatterplot-Matrix (SPM)

- **ScatterDice** [Elmqvist et al. 2008]
- Interaktion/Navigation
  - **Neuordnen** der Dimensionen
  - **Navigation**

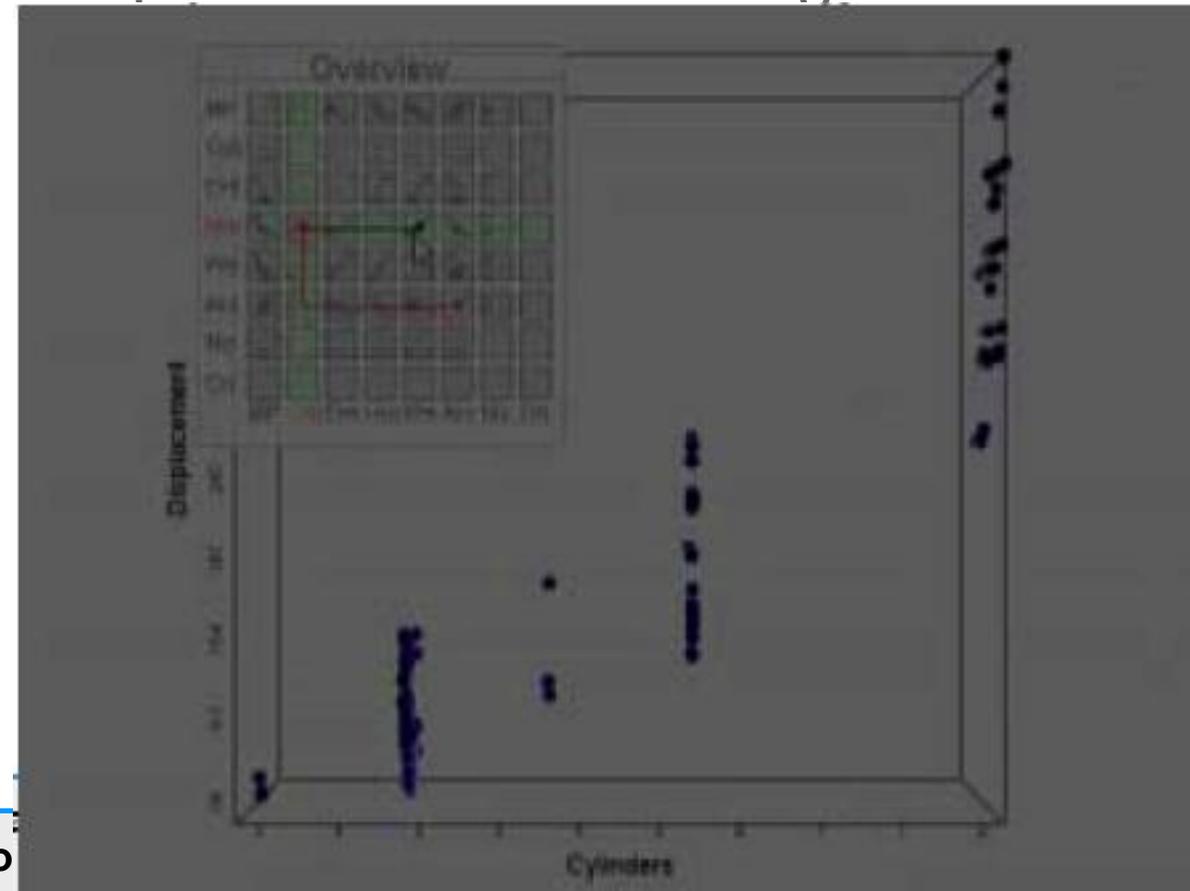




## 8.6 Scatterplot-Matrizen

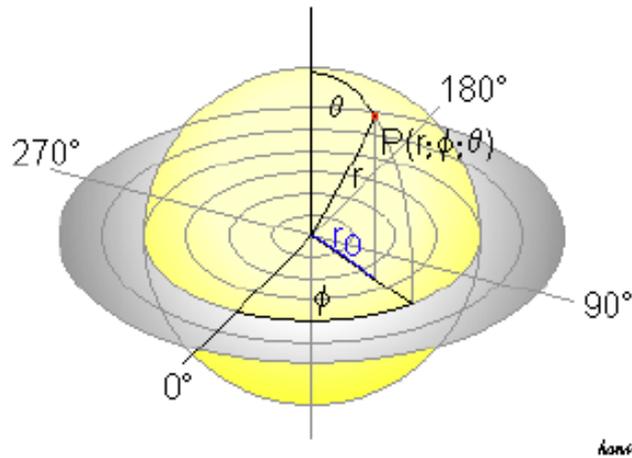
### Animationen der Scatterplot-Matrix (SPM)

- **ScatterDice** [Elmqvist et al. 2008]
- Analysen mit **Anfragen** (Queries)
- **Selektion** in einem Scatterplot und Visualisierung in weiteren Scatterplots

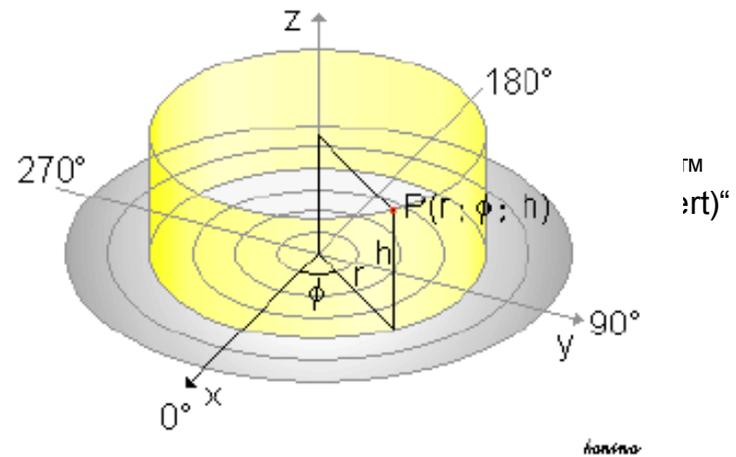


# 8.7 Parallelele Koordinaten

## Koordinatensysteme



Kugelkoordinaten



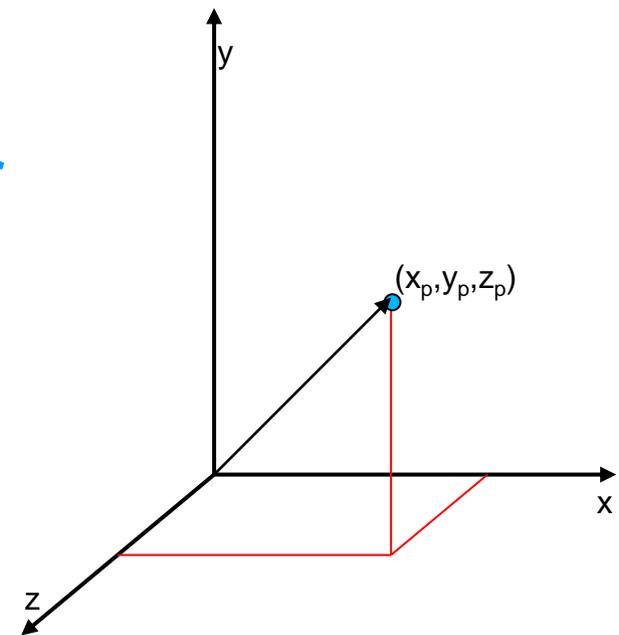
Zylinderkoordinaten

[Wikipedia.org]

## 8.7 Parallele Koordinaten

### Koordinatensysteme - Kartesisches System

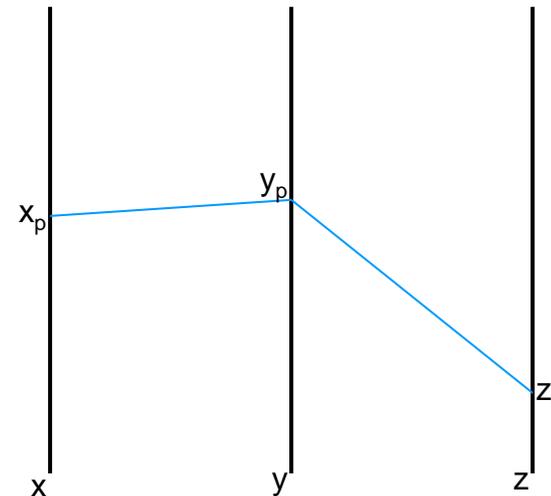
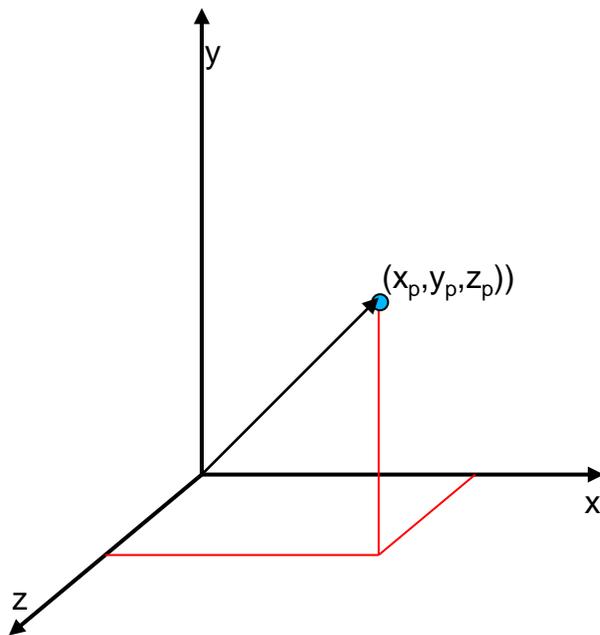
- Achsen sind **orthogonal** zueinander
- Koordinaten sind **Tupel** der Dimensionenwerte
- Werte im Koordinatensystem sind **Punkte**
- **1-2 Dimensionen gut leserlich**,  
3 Dimensionen schon schwieriger,  
4 schlecht möglich
- Mehr Dimensionen sind **kaum darstellbar**



## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System [Inselberg 1985]

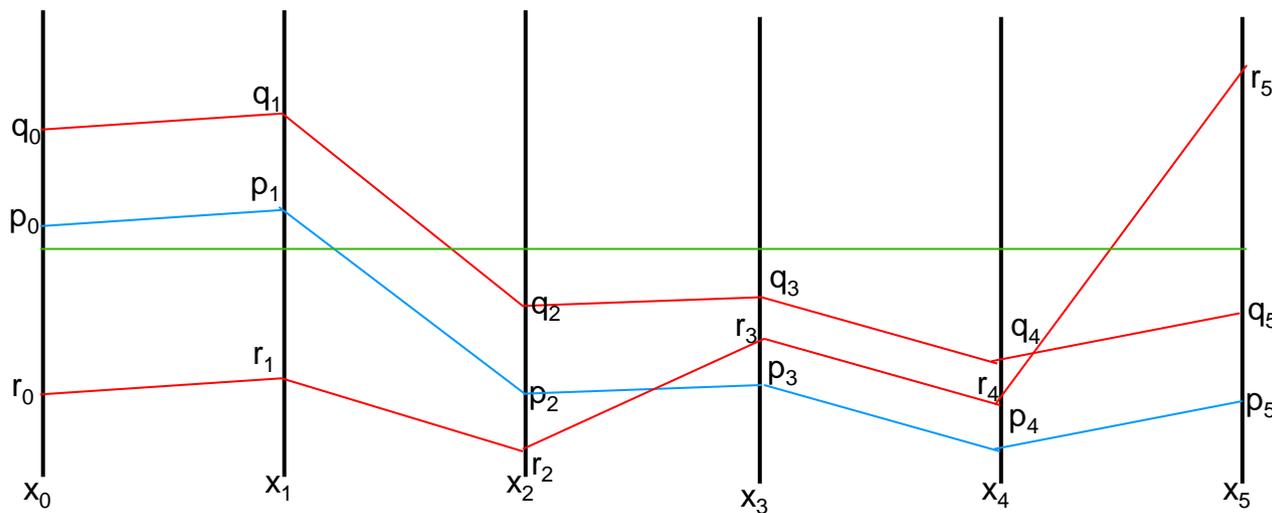
- Parallel Coordinates (Plot) - PC oder PCP
- Achsen liegen **parallel** zueinander
- Koordinaten sind **Tupel** der Dimensionswerte
- Werte im Koordinatensystem sind **Kantenzüge**



## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System [Inselberg 1985]

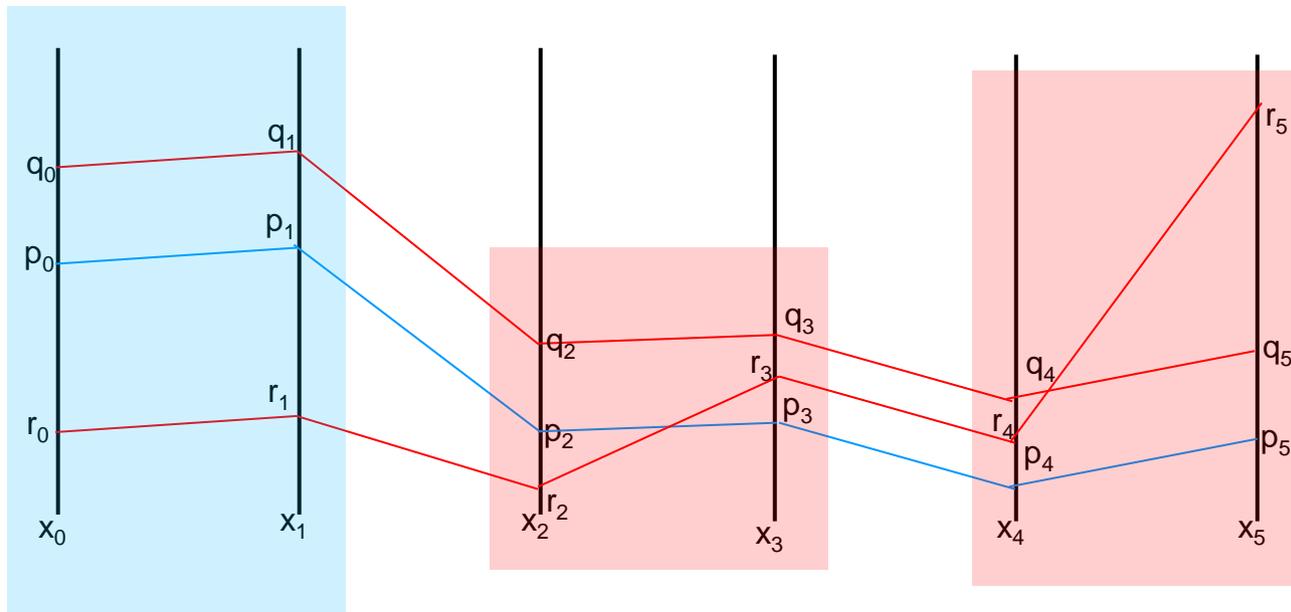
- Mehr Dimensionen, **mehr Achsen**
- **Reihenfolge** der Achsen liegt nicht unbedingt fest
- Skalierung der Achsen auf **Min/Max** der Dimension
- **Nulllinie/Mittellinie** kann angezeigt werden



## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System [Inselberg 1985]

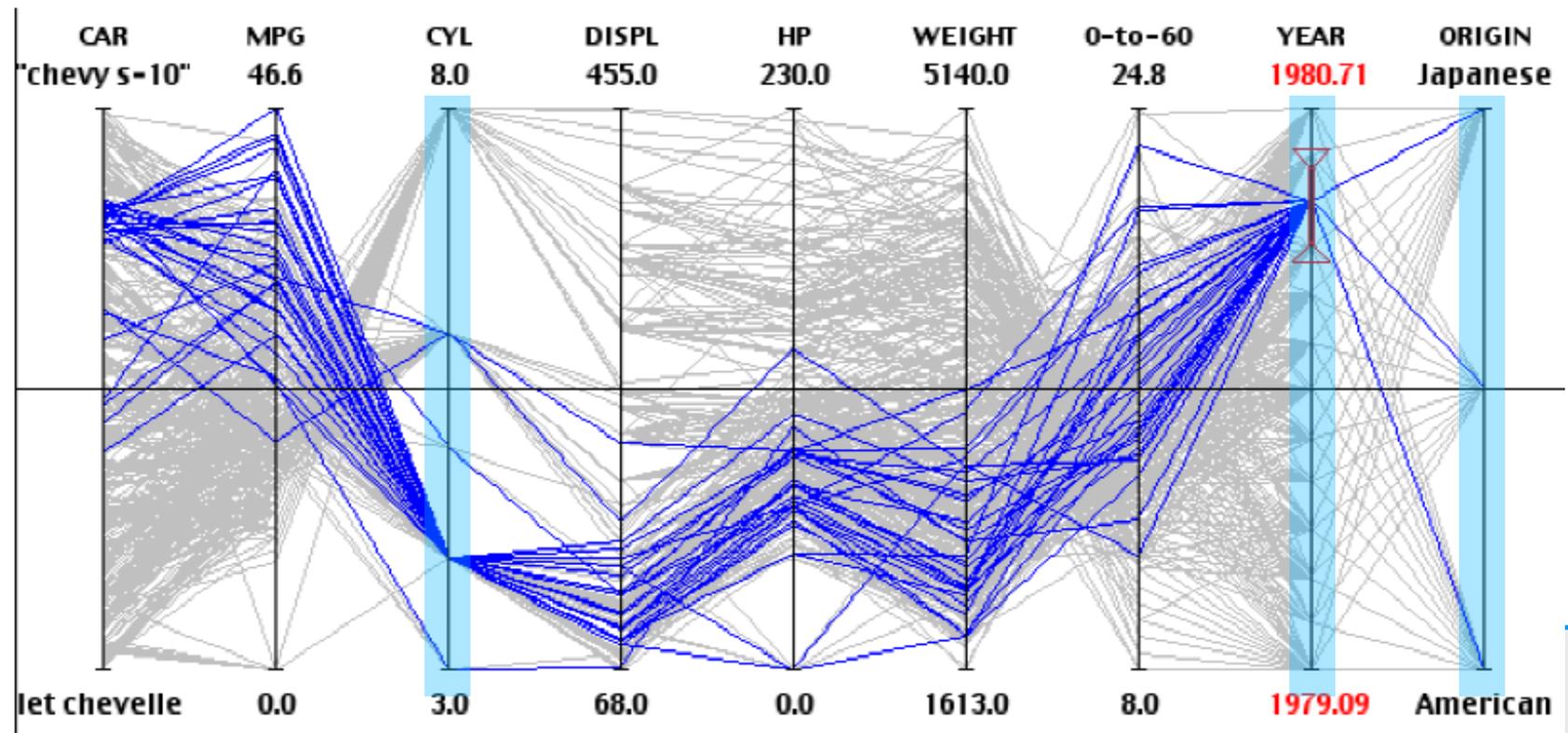
- Vergleiche der Datenwerte / Kantenzüge
- Korrelationen durch **gleiches** und **gegensätzliches** Verhalten
- **Skalierung** kann aber Erkennung wesentlich beeinträchtigen
- PC-Darstellung erlaubt aber nur **schlecht Rückschlüsse auf Häufigkeit**

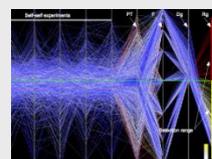


## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System

- PC-Darstellung erlaubt aber **nur schlecht Rückschlüsse** auf Häufigkeit
- Schlechte **Bildschirmeffizienz**
- **Diskrete/kategorische** Daten sind übersichtlicher

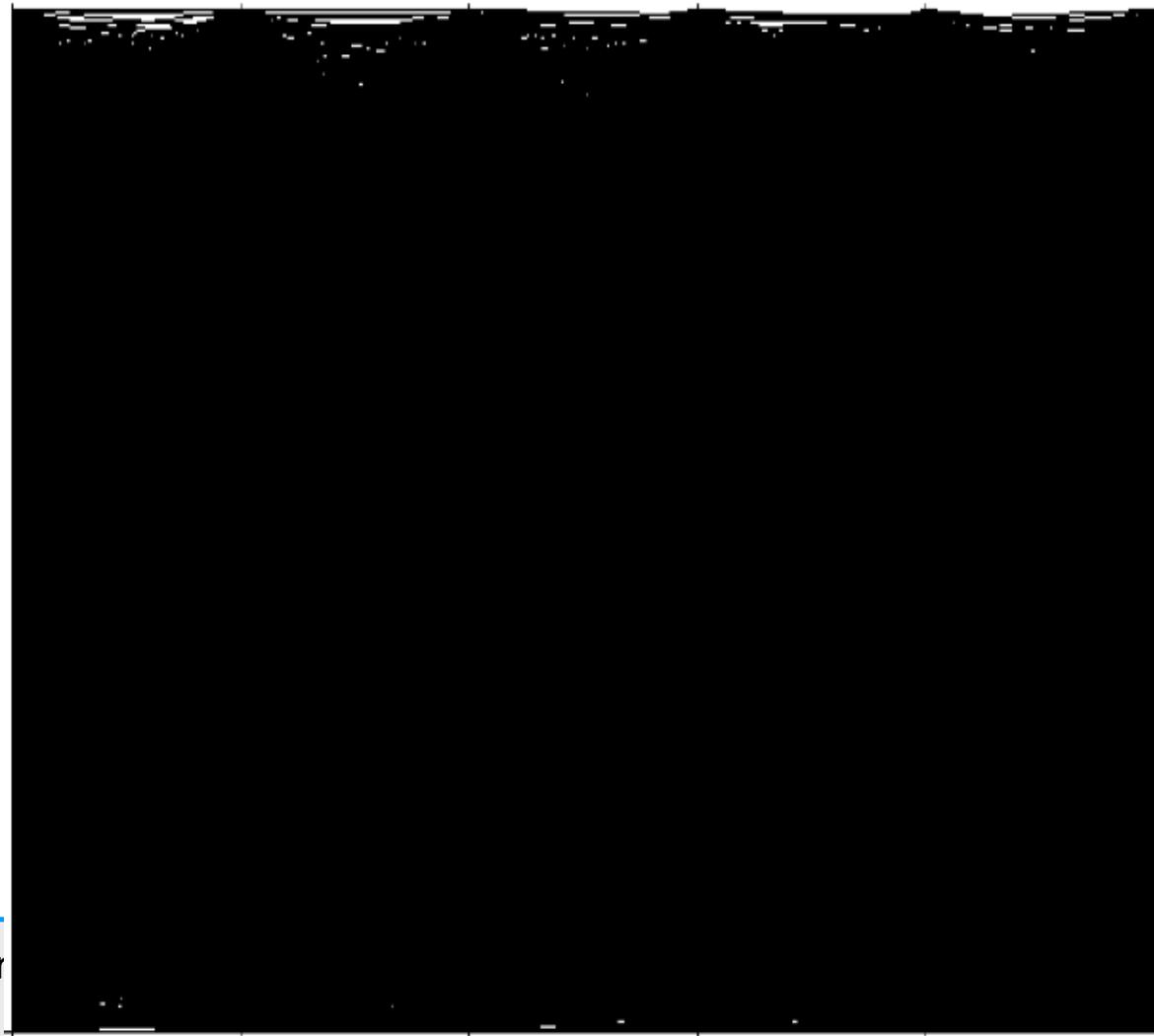




## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System

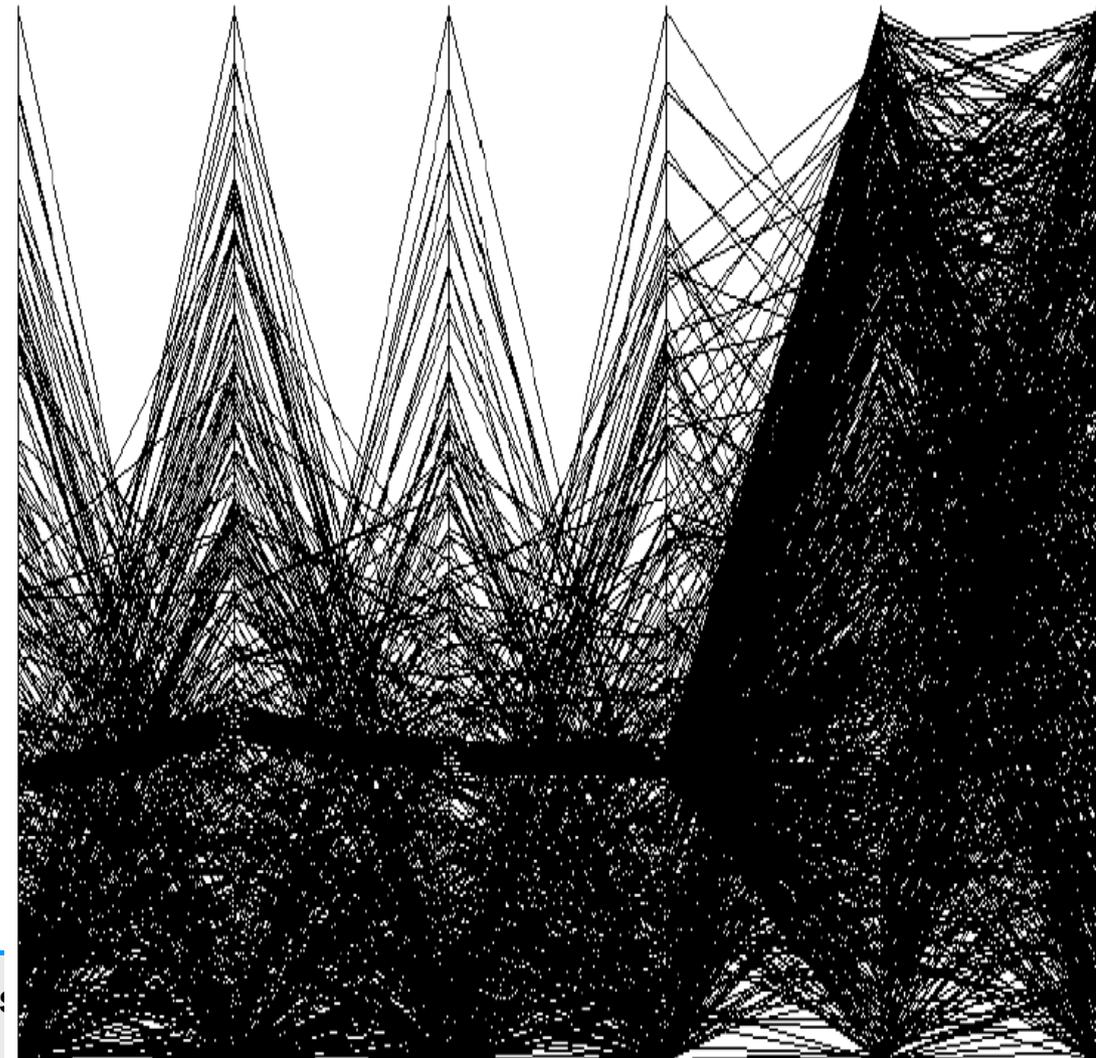
- Höhere Anzahlen von Datenwerten (Samples) **erhöhen Überdeckung**
- Schlechte **Bildschirmeffizienz**
- 15K Datenelemente, mit Rauschen



## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System

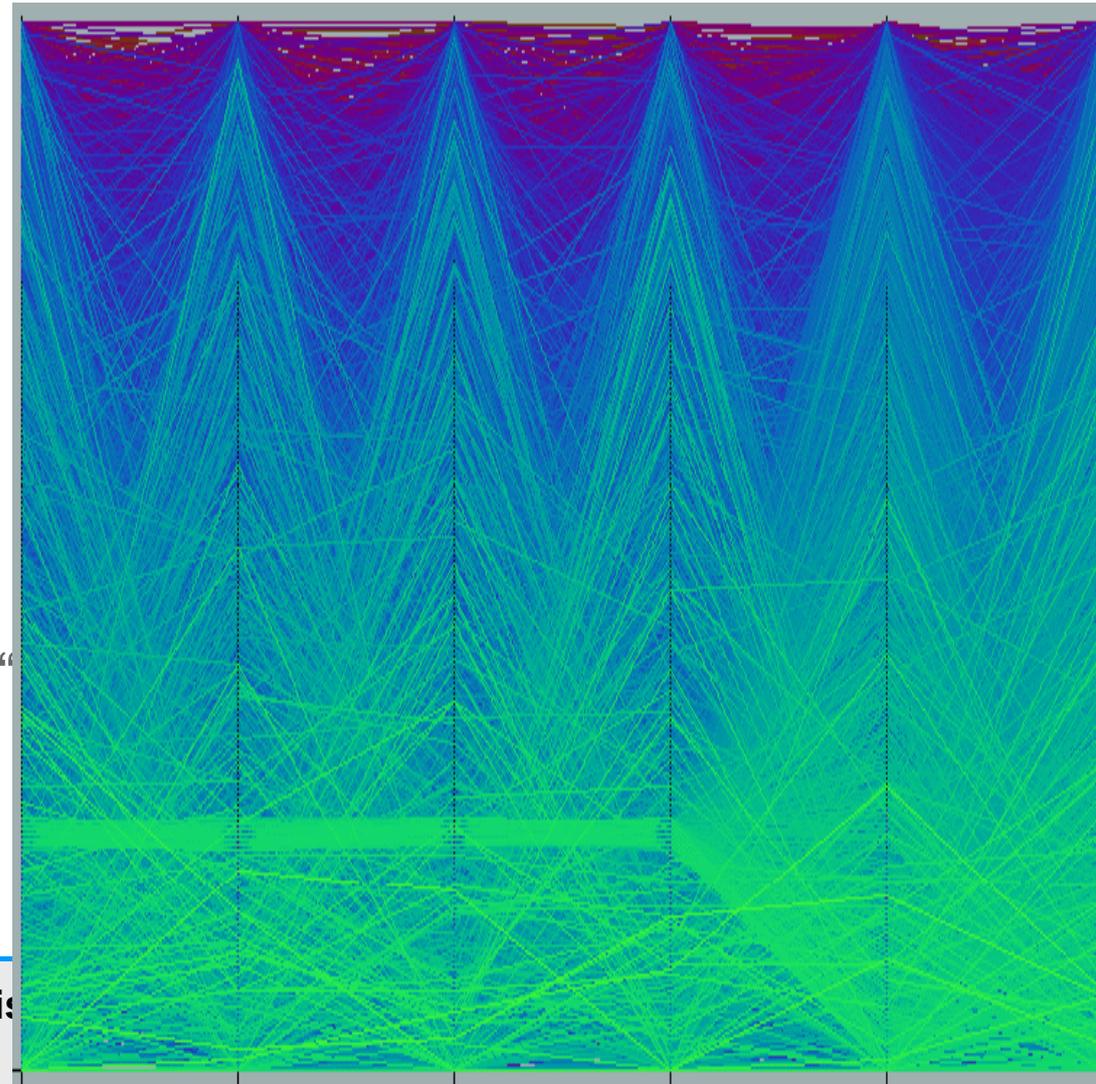
- Höhere Anzahlen von Datenwerten (Samples) **erhöhen Überdeckung**
- Schlechte **Bildschirmeffizienz**
- 15K Datenelemente, mit Rauschen
- Jetzt nur noch **5% der Daten** (750 Elemente)
- Auswahl**sampling** notwendig

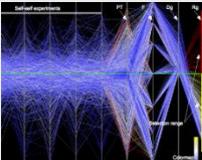


## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System

- Höhere Anzahlen von Datenwerten (Samples) **erhöhen Überdeckung**
- Schlechte **Bildschirmeffizienz**
- 15K Datenelemente, mit Rauschen
- Jetzt mit **Farbabbildung**
- **Überdeckungsproblem**, da keine Darstellungsreihenfolge
- **Unklar** ob „Zacken nach oben“ immer von gleichen Datenwerten erzeugt werden





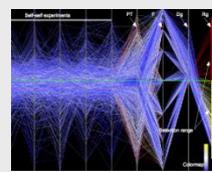
## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System

- Ermöglicht **hohe Anzahl** von Dimensionen
- Im Wesentlichen nur durch **Bildschirmplatz** beschränkt

### Aber

- Nicht **bildschirmeffizient**
- Standardmethode nicht günstig für **hohe Sample-Anzahl** (Überdeckung)
- Skalierung der Achsen **uneinheitlich**
- **Ordnung** der Achsen idR. beliebig
- **Korrelationserkennung** hängt an Skalierung und Ordnung der Achsen

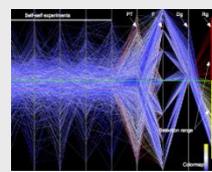


## 8.7 Parallele Koordinaten

### Koordinatensysteme - Paralleles System

#### Lösungsansätze

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



# Literatur

- Asimov: The Grand Tour. *SIAM Journal on Science and Statistical Computing IV*, (1):185ff 1985.
- Buja, Asimov: Grand Tour Methods: An Outline, In *Proc. of the 17th Symposium on the Interface of Computer Science and Statistics*, pp. 63-67, 1986.
- Buja, Cook, Asimov, Hurley: Theory and Computational Methods for Dynamic Projections in High-Dimensional Data Visualization, *Monograph*, 1999.
- Chambers, Cleveland, Kleiner, Tukey: Graphical Methods for Data Analysis, Wadsworth, Monterey, CA, 1983.
- Cleveland: The Elements of Graphing Data, Hobart Press, Summit, NJ, 1985.
- Cleveland, McGill: Dynamic Graphics for Statistics, CRC Press, Boca Raton, FL, 1998.
- Cook, Buja, Cabrera, Hurley: Grand Tour and Projection Pursuit, *Journal of Computational and Graphical Statistics*, 4(3):155-172, 1995.
- Elmqvist, Dragicevic, Fekete: Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation, In *Proc. of IEEE Symposium on Information Visualization*, pp. 1141-1148, 2008.
- Hartigan: Printer Graphics for Clustering, *Journal of Statistical Computation and Simulation*, 4(3):187-213, 1975.
- Huber: Projection Pursuit, *The Annals of Statistics*, 13(2):435-475, 1985.
- Inselberg: The Plane with Parallel Coordinates, *The Visual Computer*, 1(2):69-92, 1985.
- Tukey, Tukey: Computer Graphics and Exploratory Data Analysis: An Introduction. In *Proc. of Sixth Annual Conference and Exposition: Computer Graphics 1985*, National Computer Graphics Association, 1985
- van Wijk, van Liere: HyperSlice - Visualization of Scalar Functions of Many Variables. In *Proc. of IEEE Visualization*, 1993.
- Wilkinson, Anand, Grossman: Graph-Theoretic Scagnostics, In *Proc. of IEEE Symposium on Information Visualization*, pp. 157-164, 2004.