Multi-Dimensional Data Visualization (Parallel Coordinates)

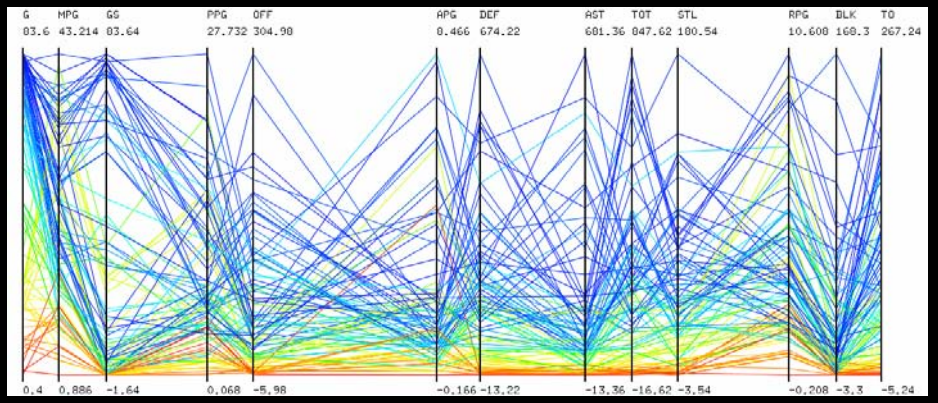
袁晓如

北京大学

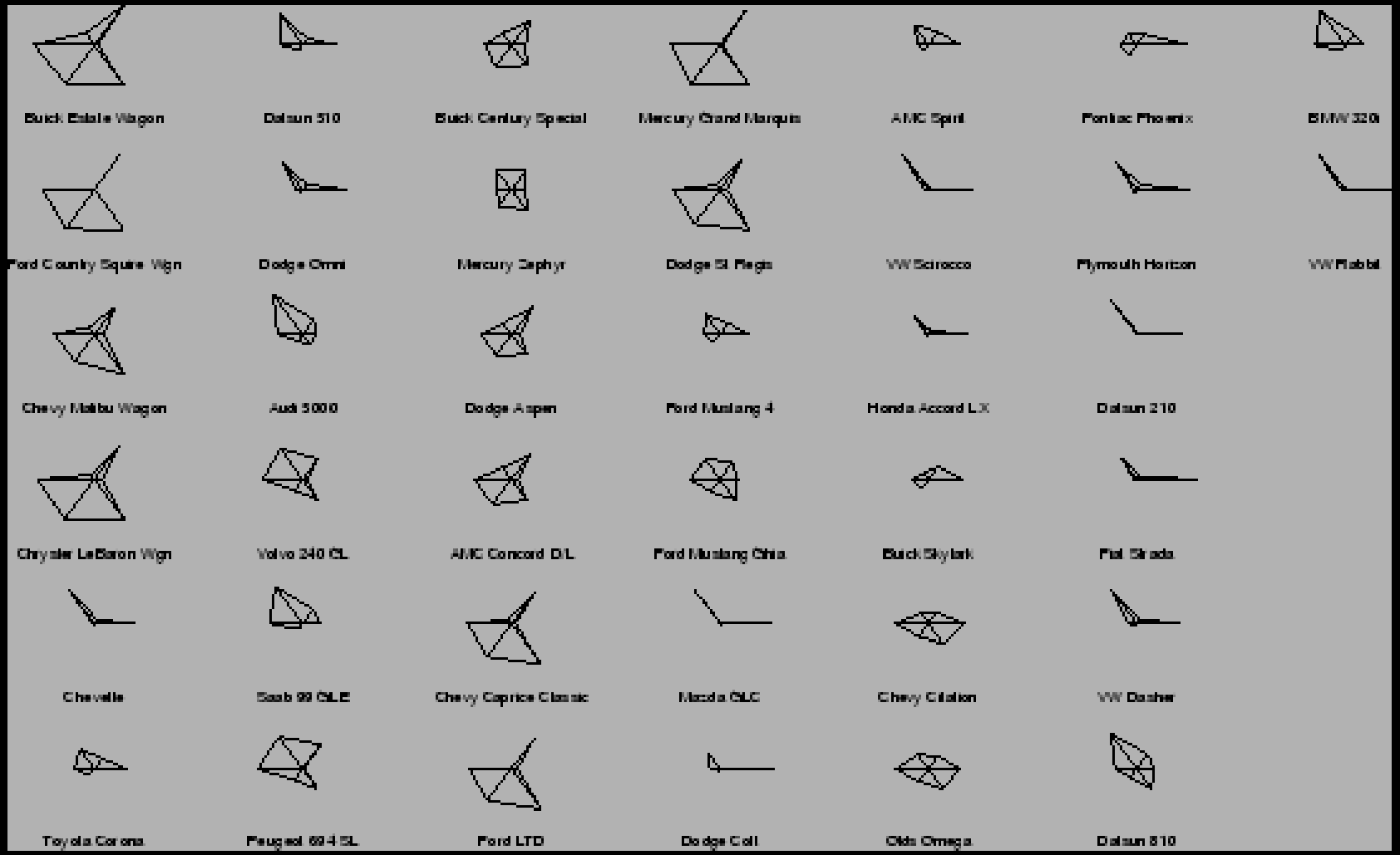
2009年12月25日

Parallel Coordinates History

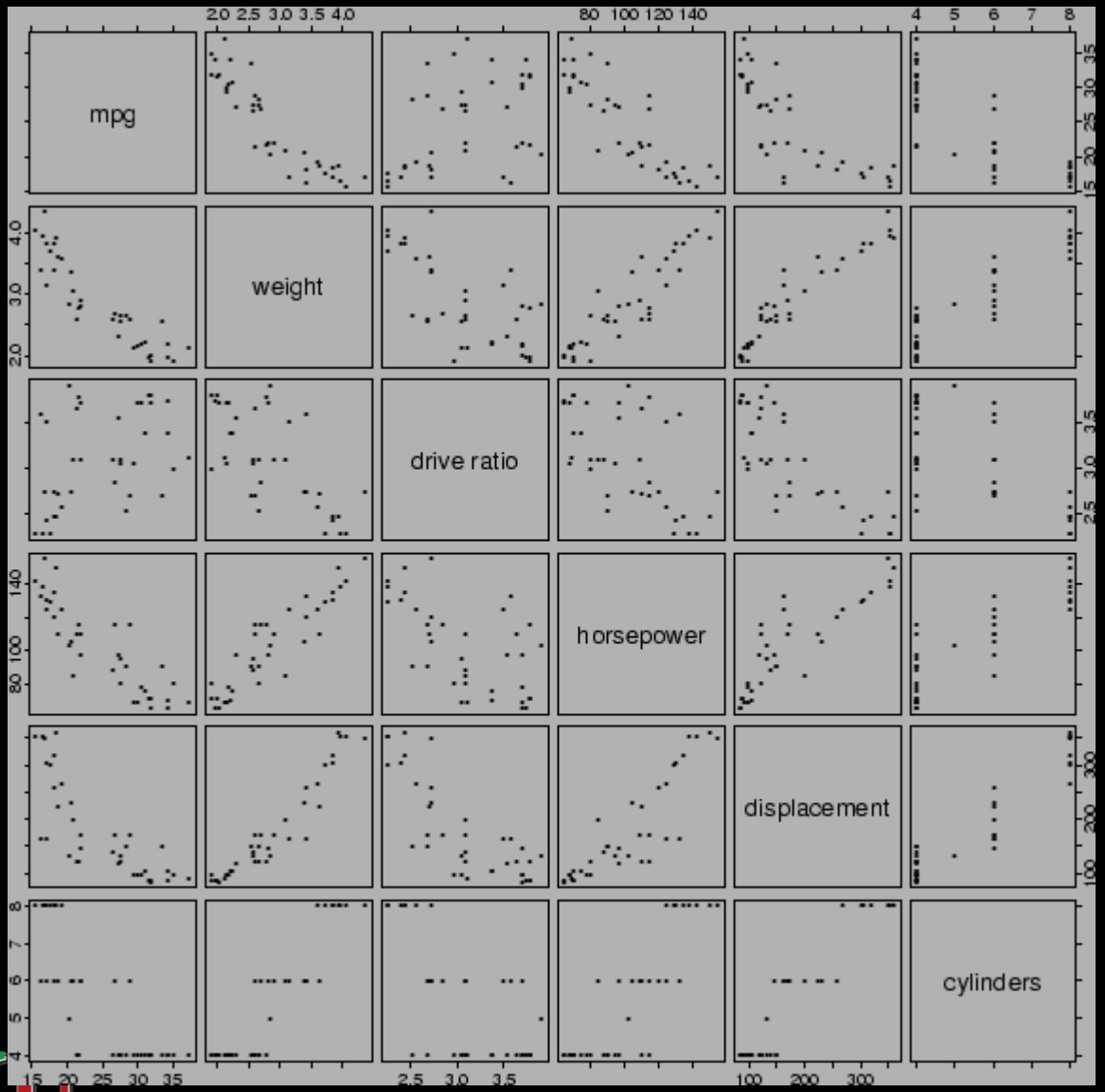
- Invented by Inselberg (1985) for computational geometry http://www.math.tau.ac.il/~aiisreal/
- Introduced by Wegman (1990) for hyper-dimensional data analysis
- Reveal data relationships, anomalies, structures & trends
- Widely used in the statistics, data analysis and visualization domain



Star Glyphs

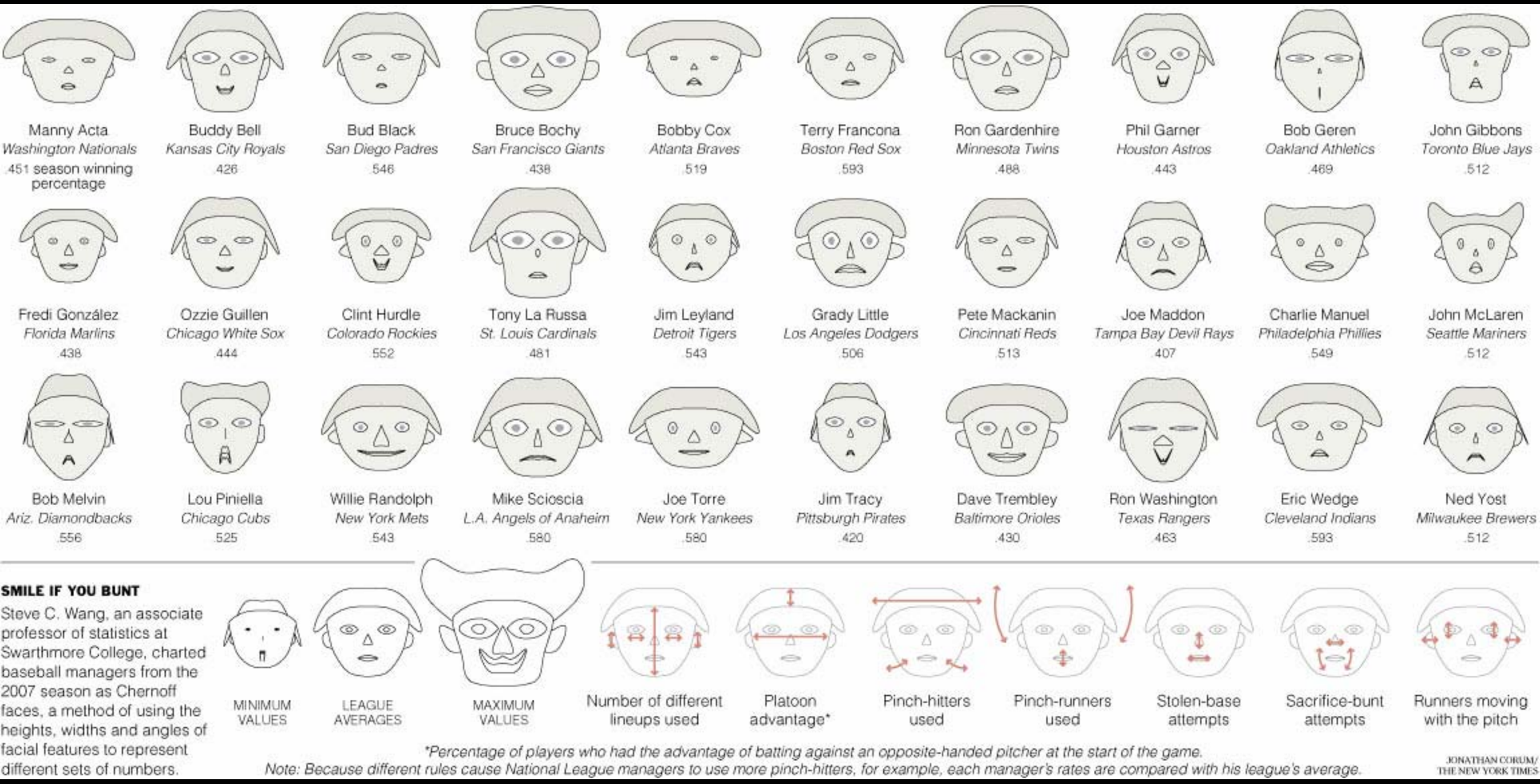

Buick Estate Wagon
Datsun 510
Buick Century Special
Mercury Grand Marquis
AMC Spirit
Pontiac Phoenix
BMW 320i
Ford Country Squire Wgn
Dodge Omni
Mercury Zephyr
Dodge St Regis
VW Scirocco
Plymouth Horizon
VW Rabbit
Chevy Malibu Wagon
Audi 5000
Dodge Aspen
Ford Mustang 4
Honda Accord LX
Datsun 210
Chrysler LeBaron Wgn
Volvo 240 GL
AMC Concord DL
Ford Mustang Ghia
Buick Skylark
Fiat Strada
Chevelle
Saab 99 GLE
Chevy Caprice Classic
Mazda GLC
Chevy Citation
VW Dasher
Toyota Corona
Peugeot 604 SL
Ford LTD
Dodge Colt
Olds Omega
Datsun 810

Scatterplot Matrix



VisPKU

Chernoff Faces



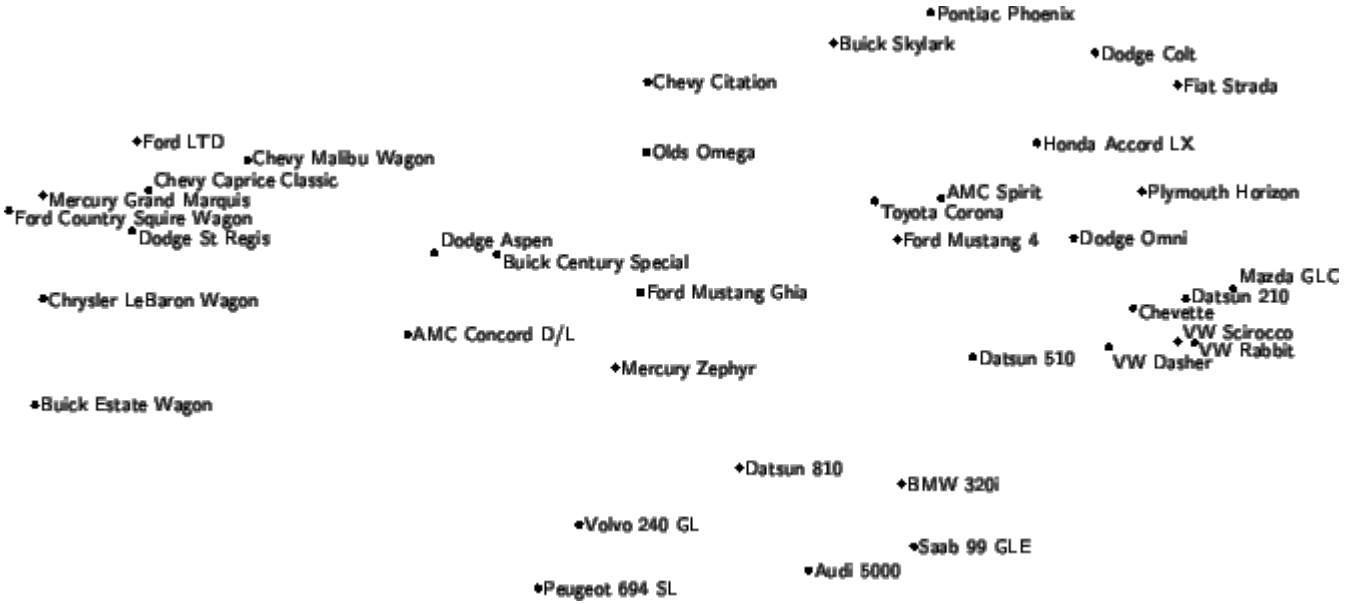
SMILE IF YOU BUNT

Steve C. Wang, an associate professor of statistics at Swarthmore College, charted baseball managers from the 2007 season as Chernoff faces, a method of using the heights, widths and angles of facial features to represent different sets of numbers.

*Percentage of players who had the advantage of batting against an opposite-handed pitcher at the start of the game.
Note: Because different rules cause National League managers to use more pinch-hitters, for example, each manager's rates are compared with his league's average.

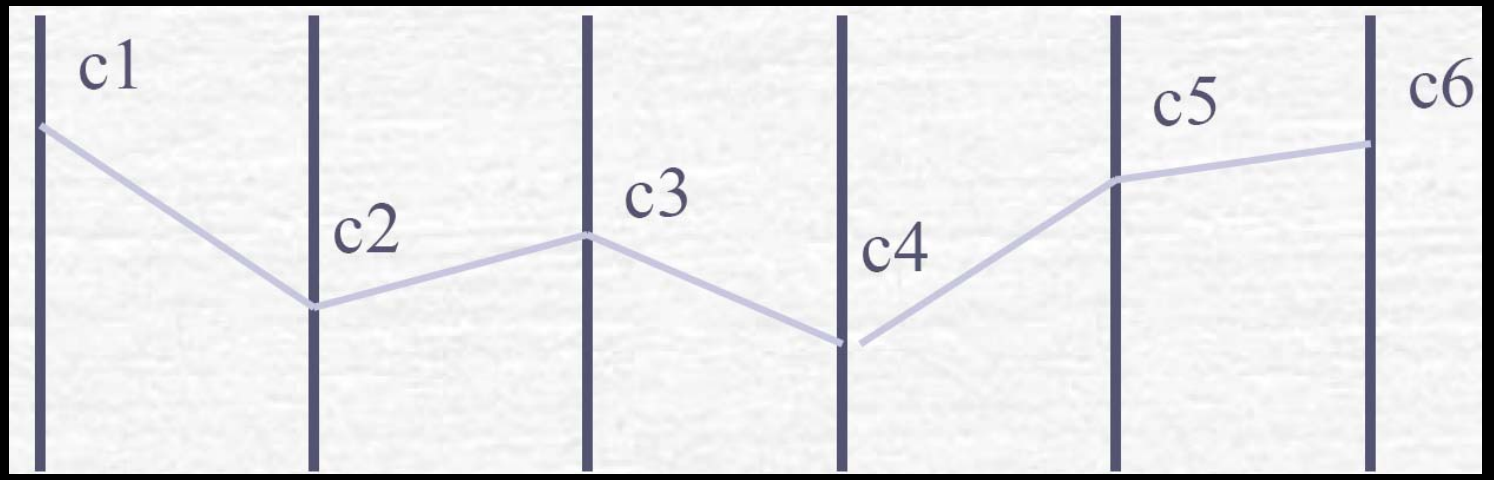
JONATHAN CORUM/THE NEW YORK TIMES

Multidimensional Scaling

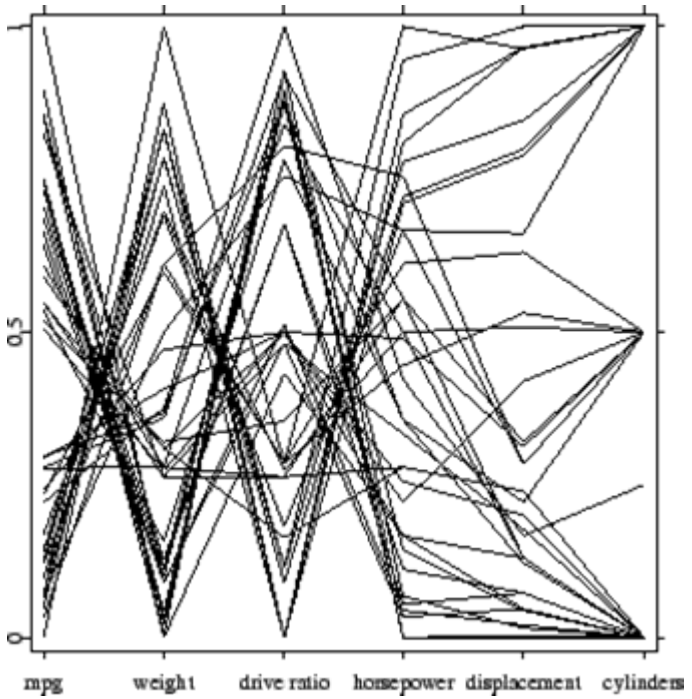


Parallel Coordinates Basics

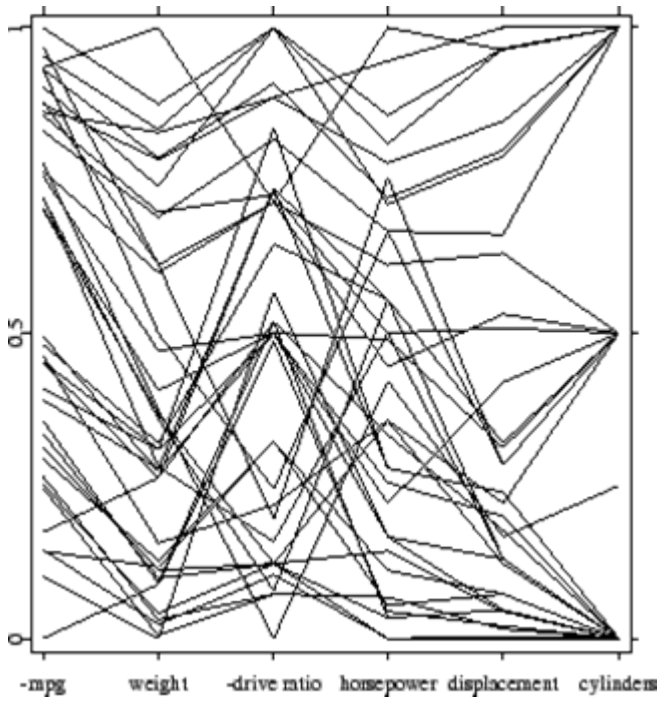
- To represent N dimensional data
  - Set N vertical axes in parallel
  - Put data to intersects on corresponding axes
  - Connect intersects



Reorder Parallel Coordinates

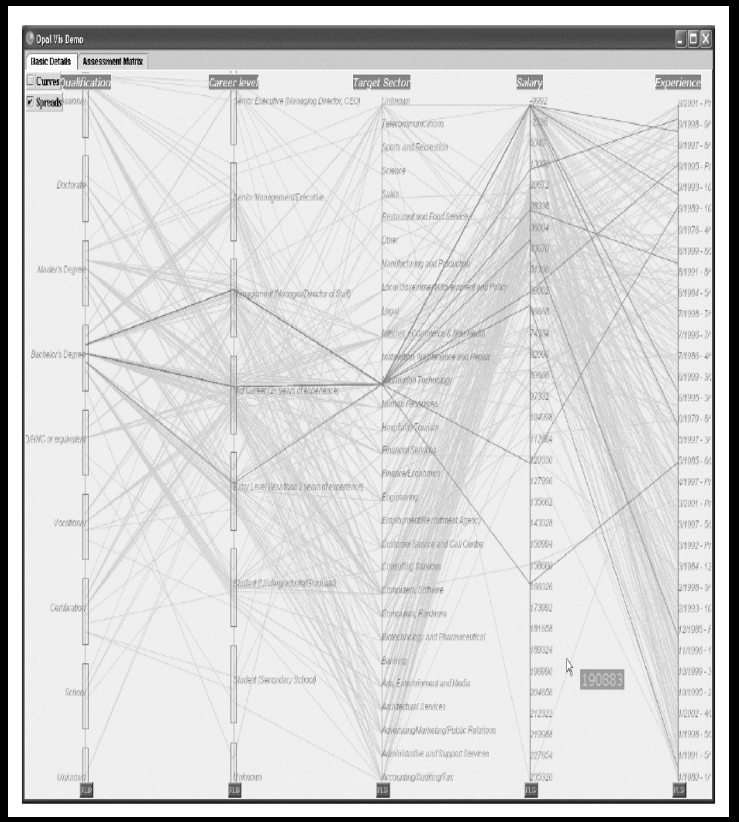


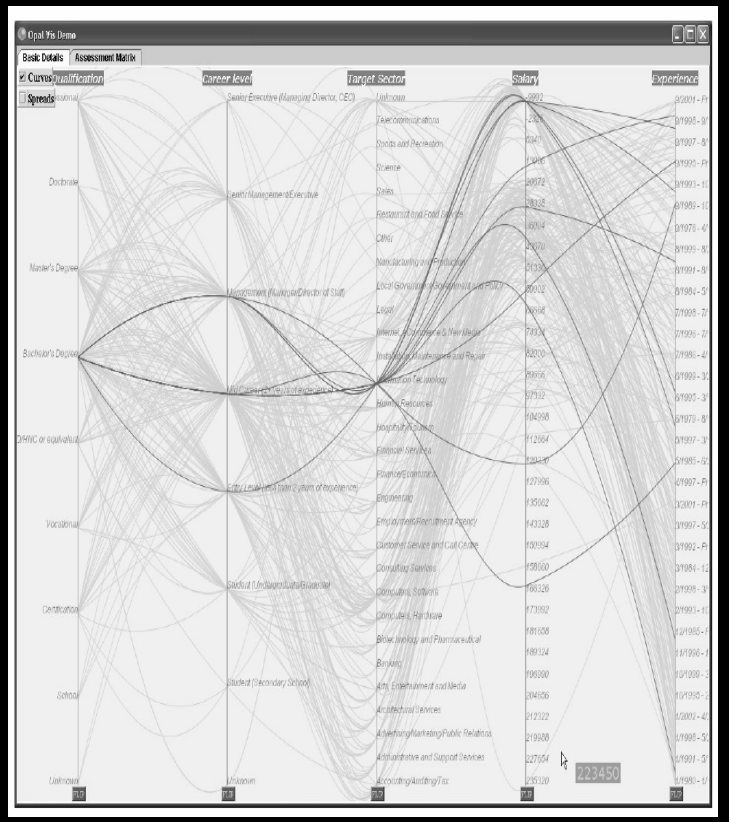
Original



Reordered

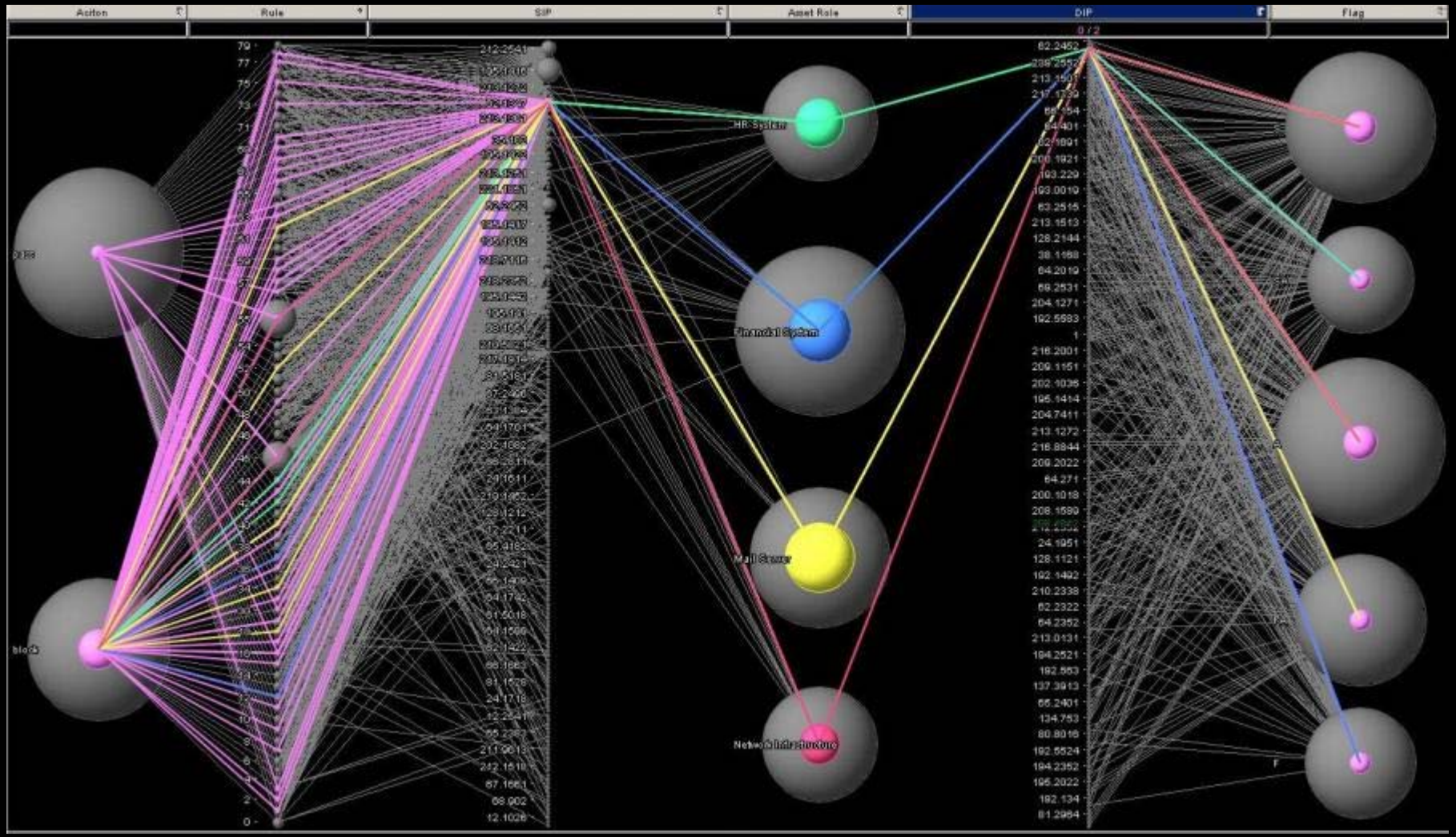
Curved Parallel Coordinates



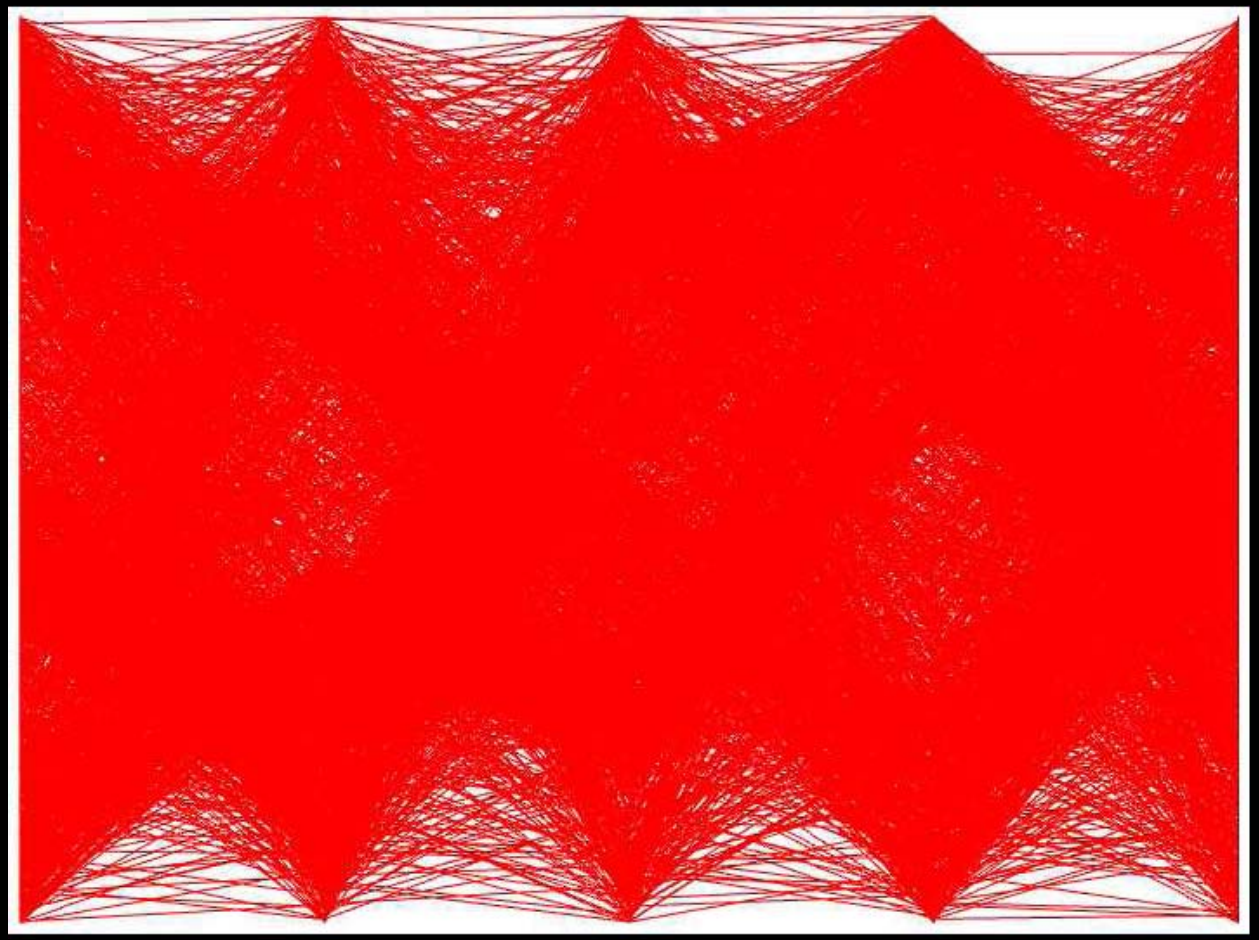


More fluently visibility but drawing curves takes more time and processing

Parallel Coordinate View of a Firewall Log File

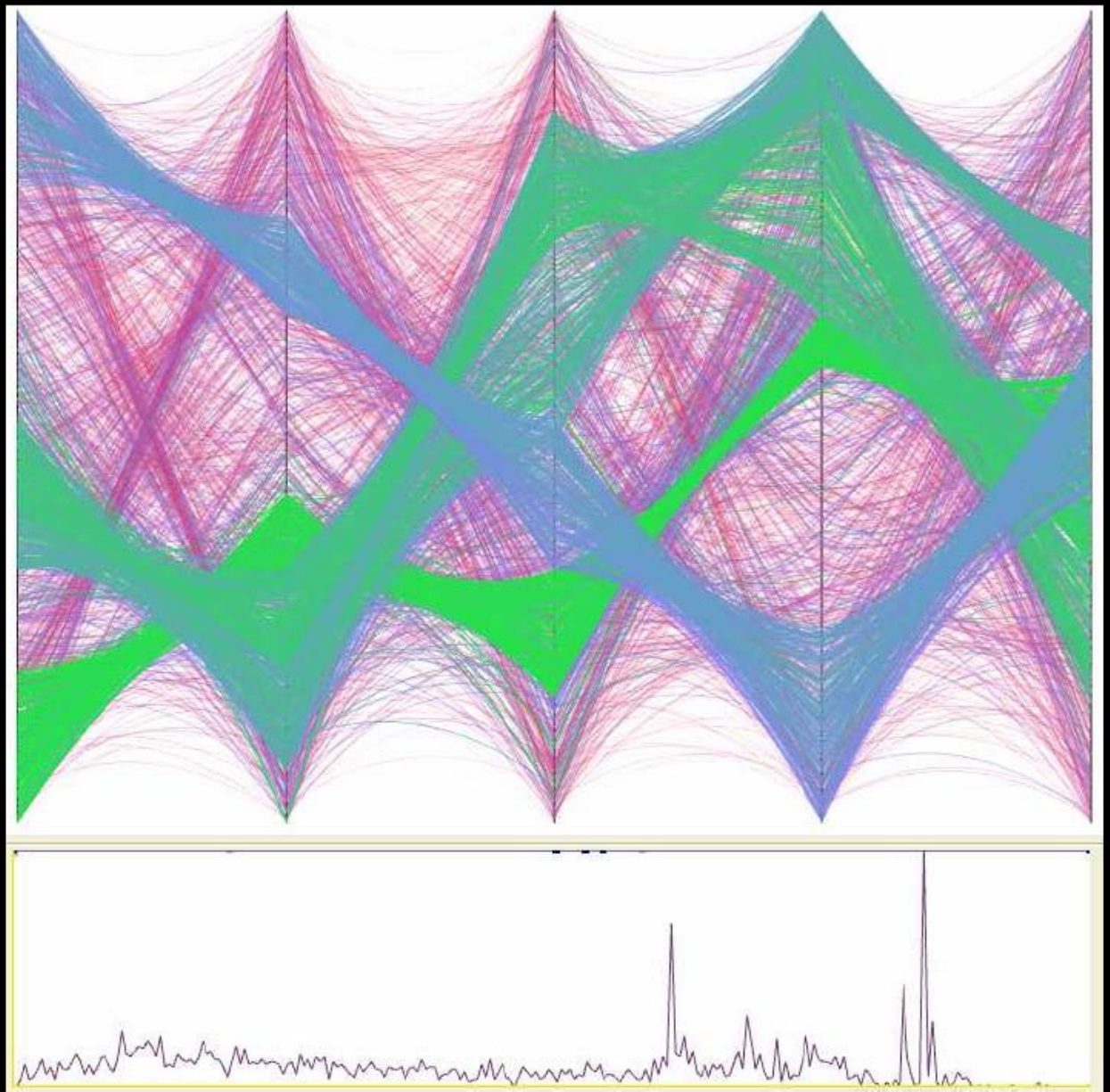


Visual Clustering in Parallel Coordinates



[EuroVis 2008]

Visual Clustering in Parallel Coordinates

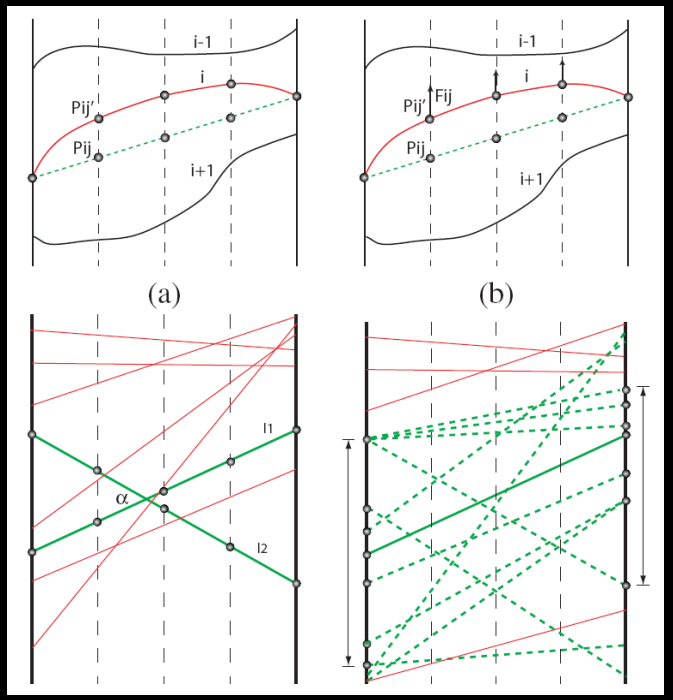


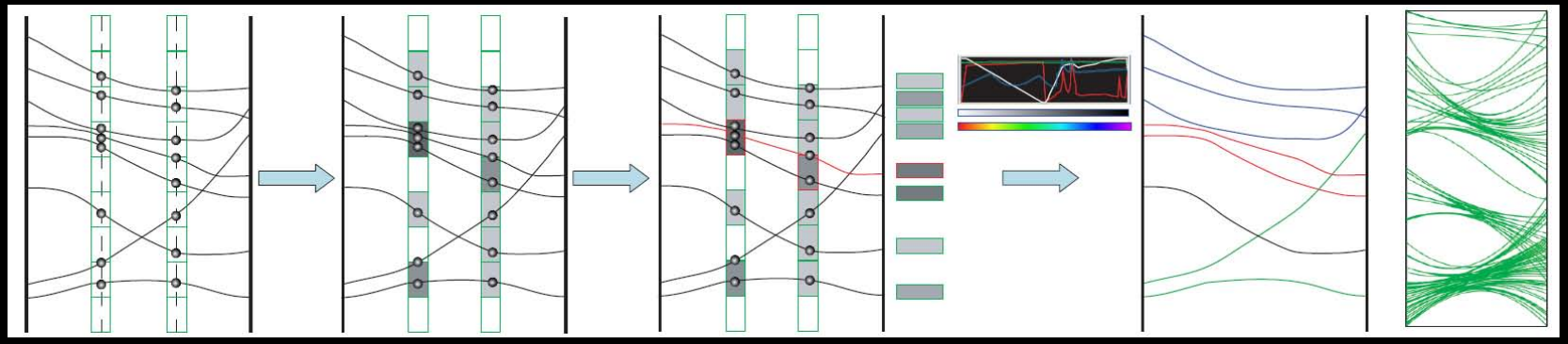
Modeling visual interaction between lines

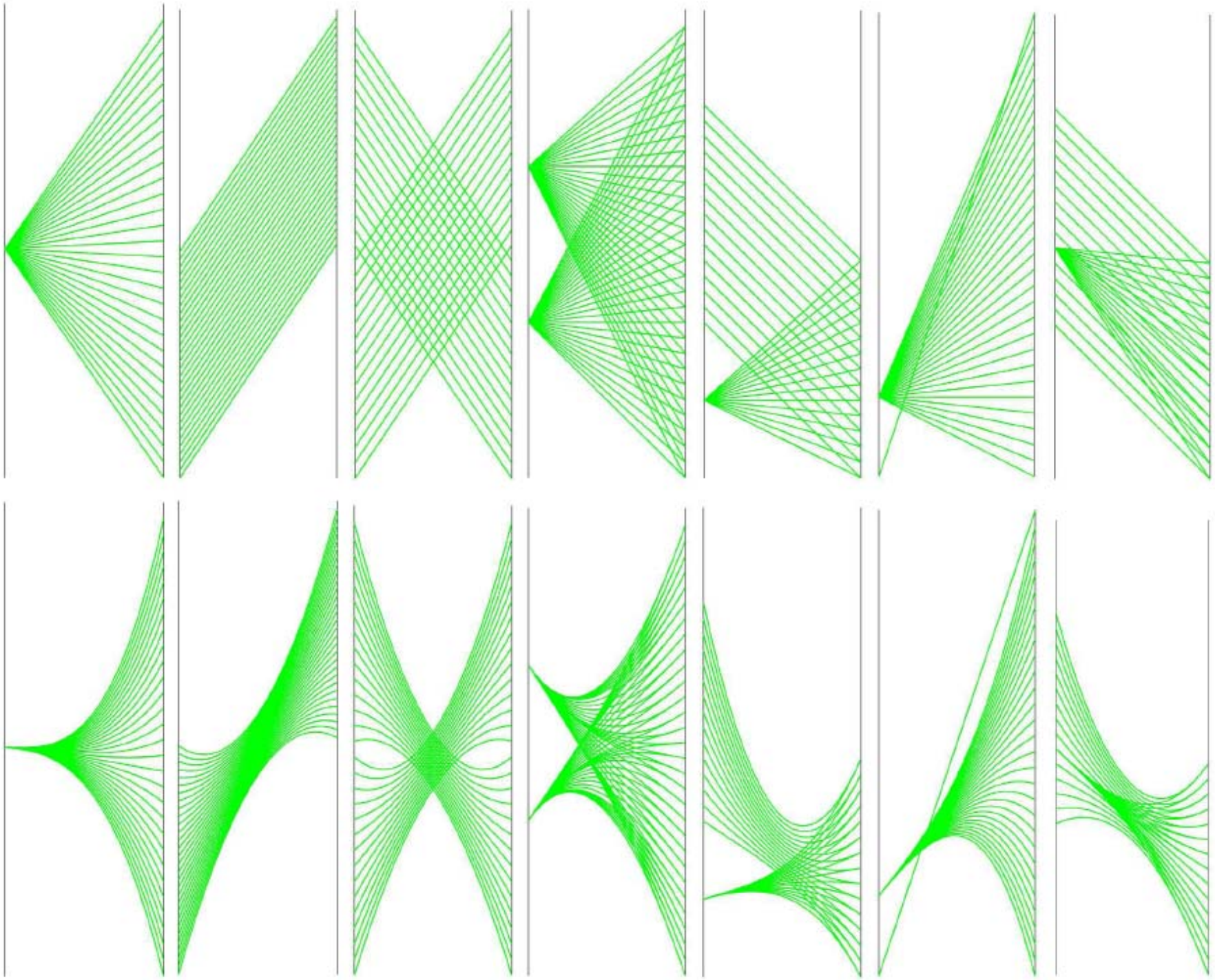
$$\mathbf{E} = \alpha_c \mathbf{E}_{curvature} + (1 - \alpha_c)\mathbf{E}_{gravitation}$$

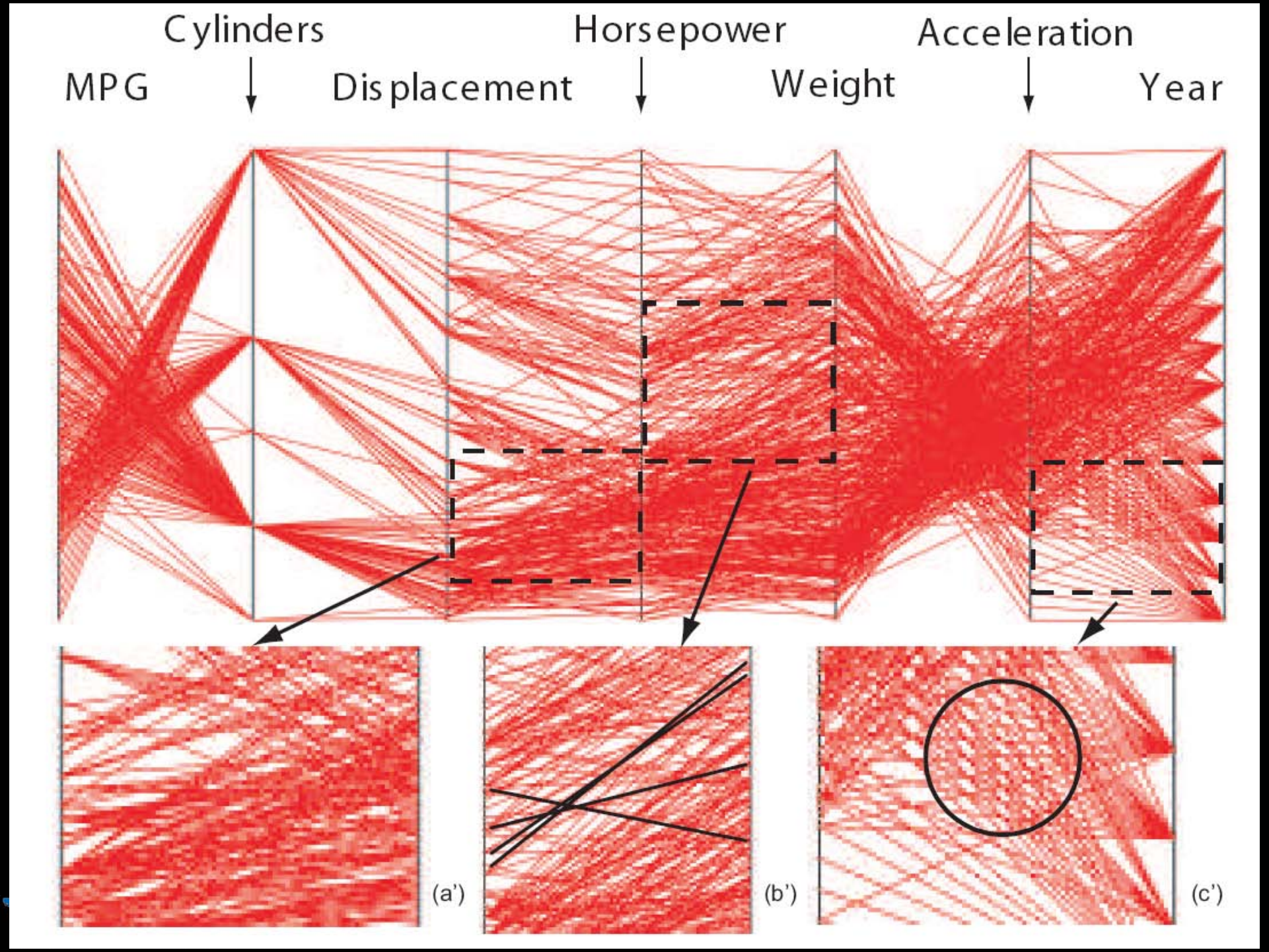
$$\mathbf{E}_{curvature} = \sum_{i=1}^{n}\sum_{j=1}^{m} |P'_{ij} - P_{ij}|$$

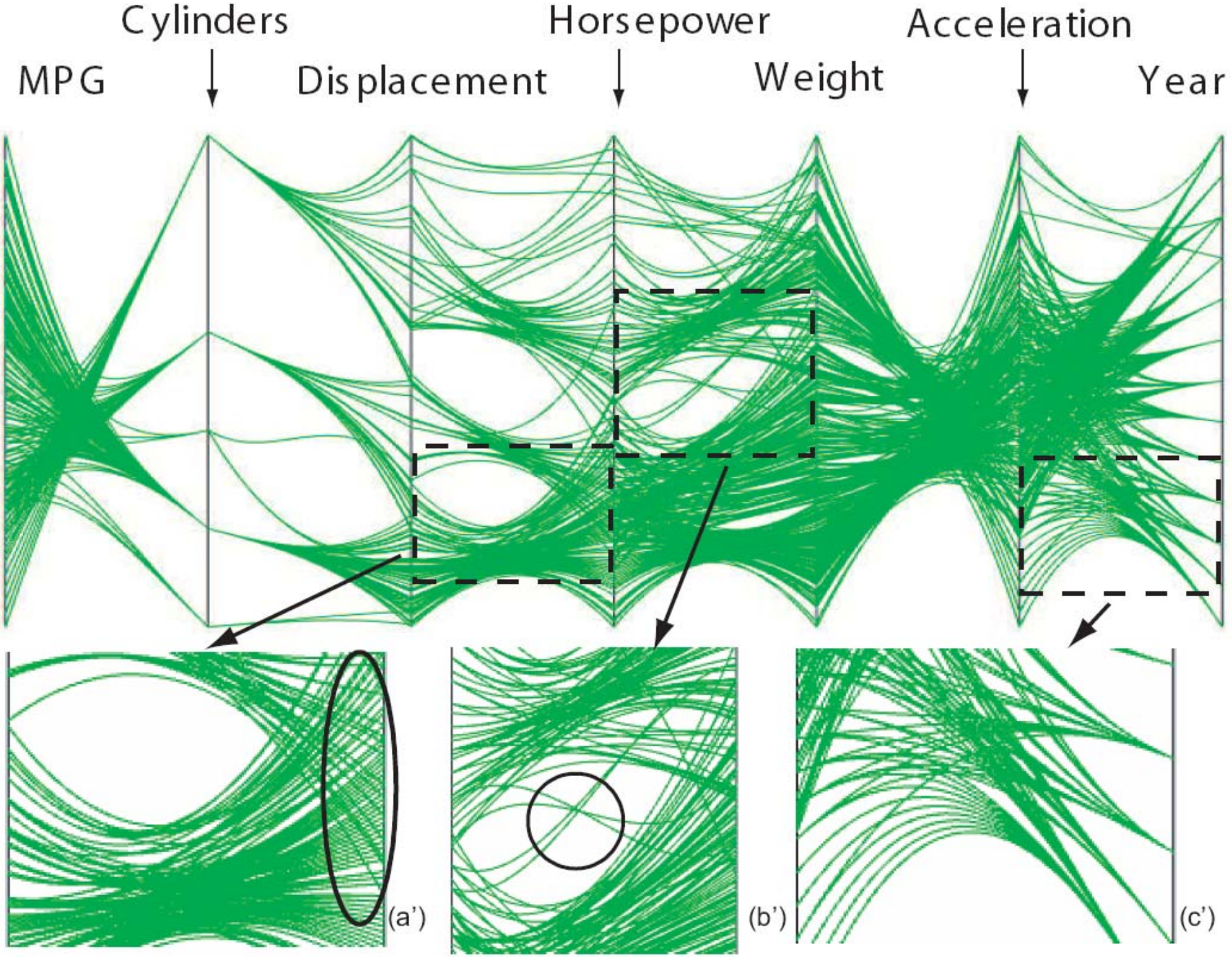
$$\mathbf{E}_{gravitation} = \sum_{i=1}^{n}\sum_{j=1}^{m} -F_{ij} \cdot (P'_{ij} - P_{ij}) + E_{ij}$$

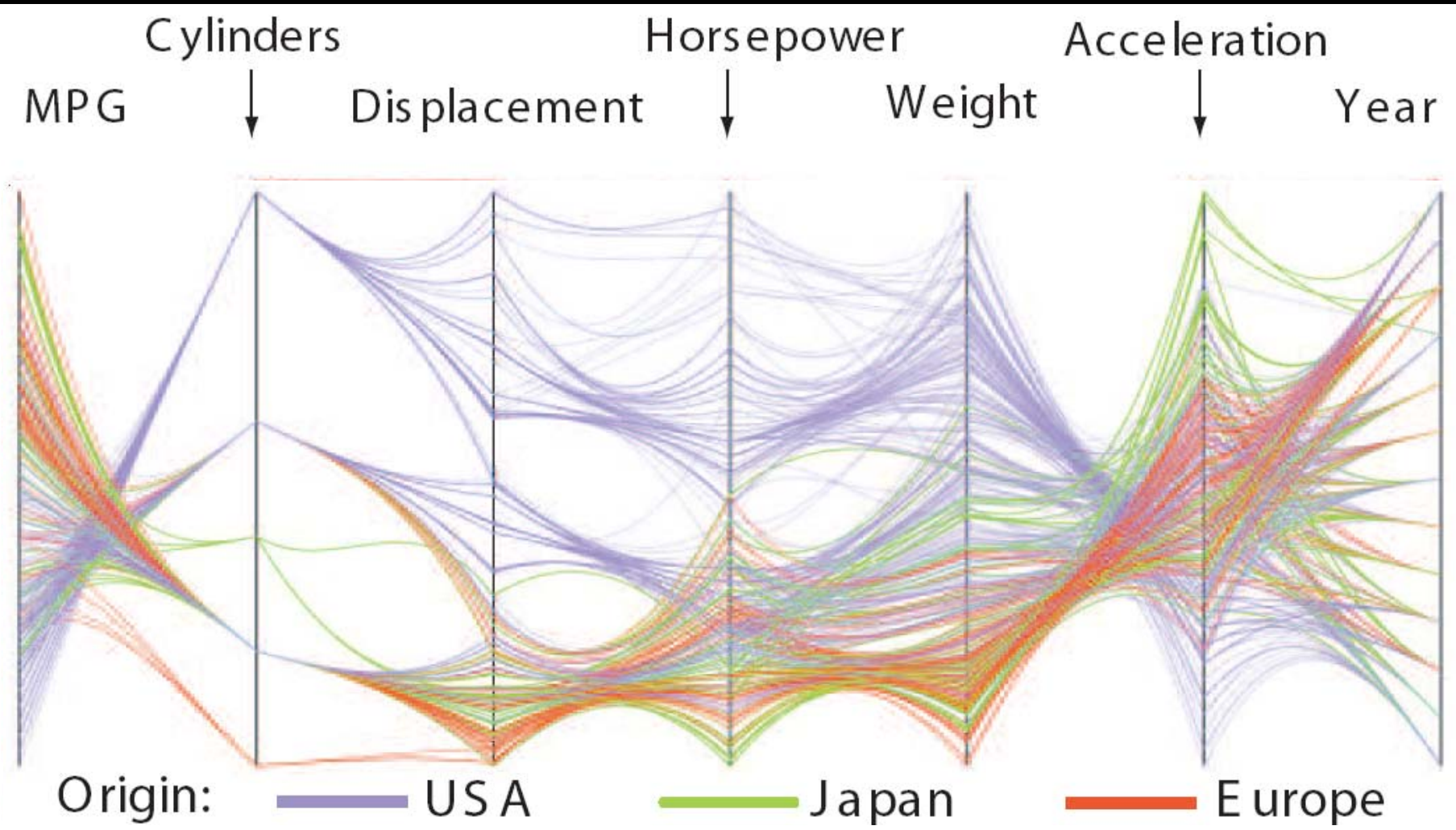


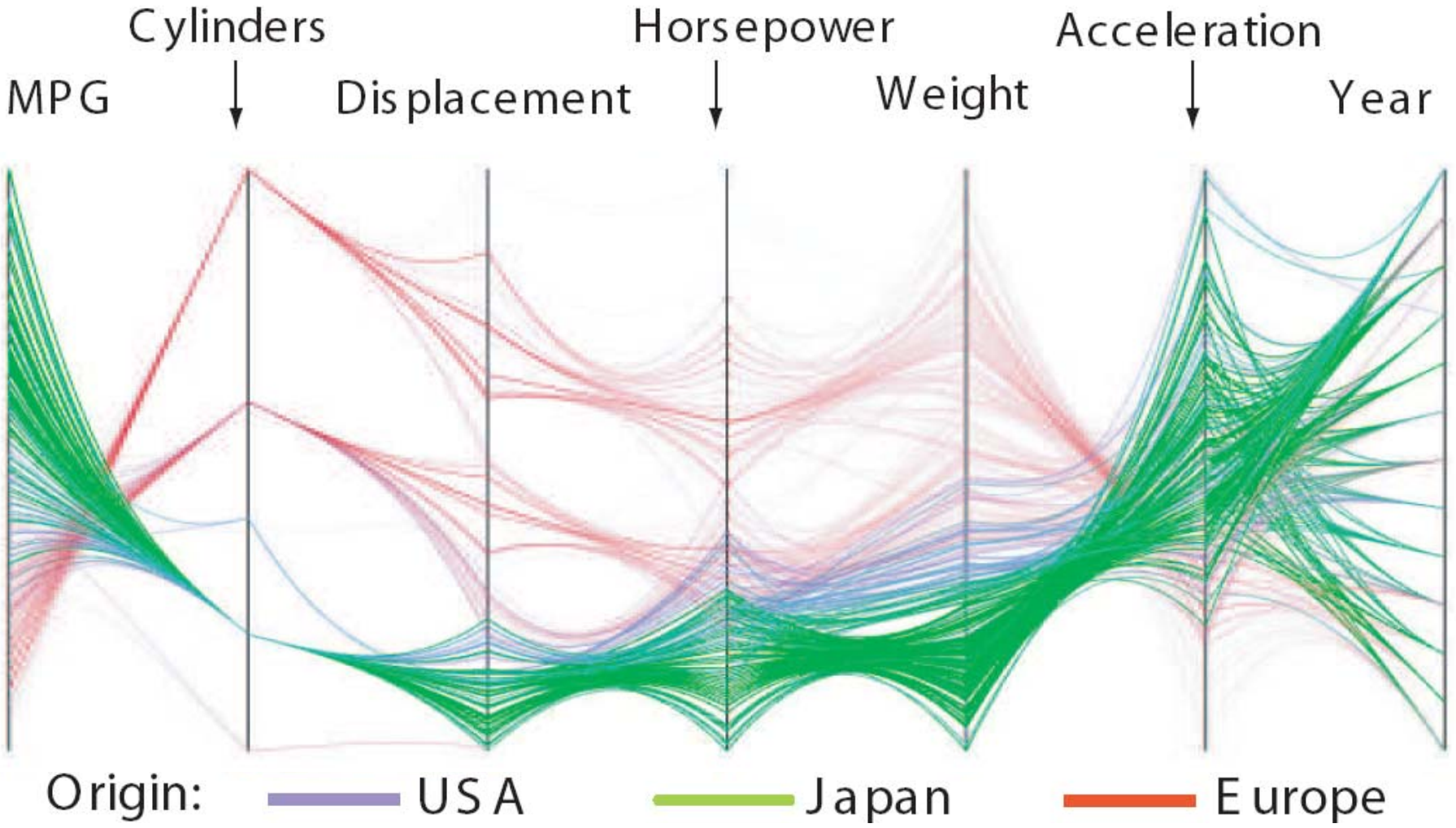


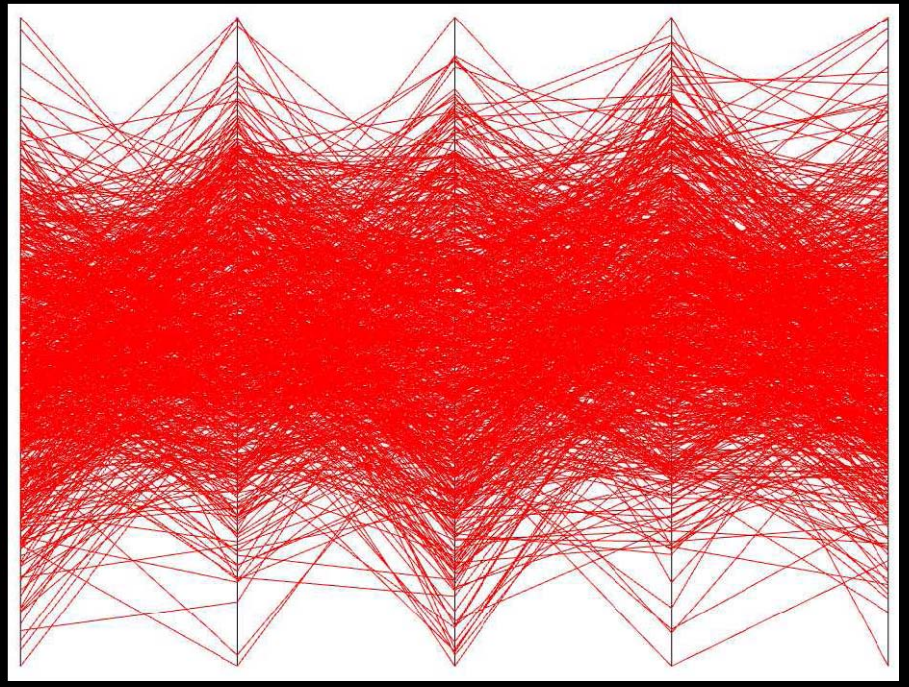


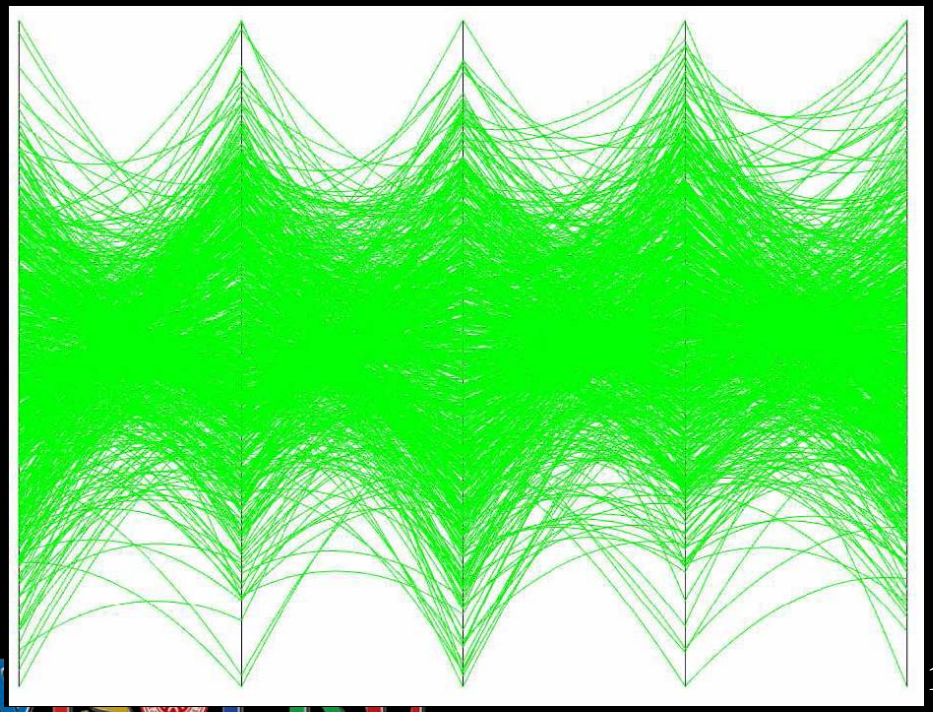

Cylinders
Horsepower
Acceleration
MPG
Displacement
Weight
Year
(a')
(b')
(c')

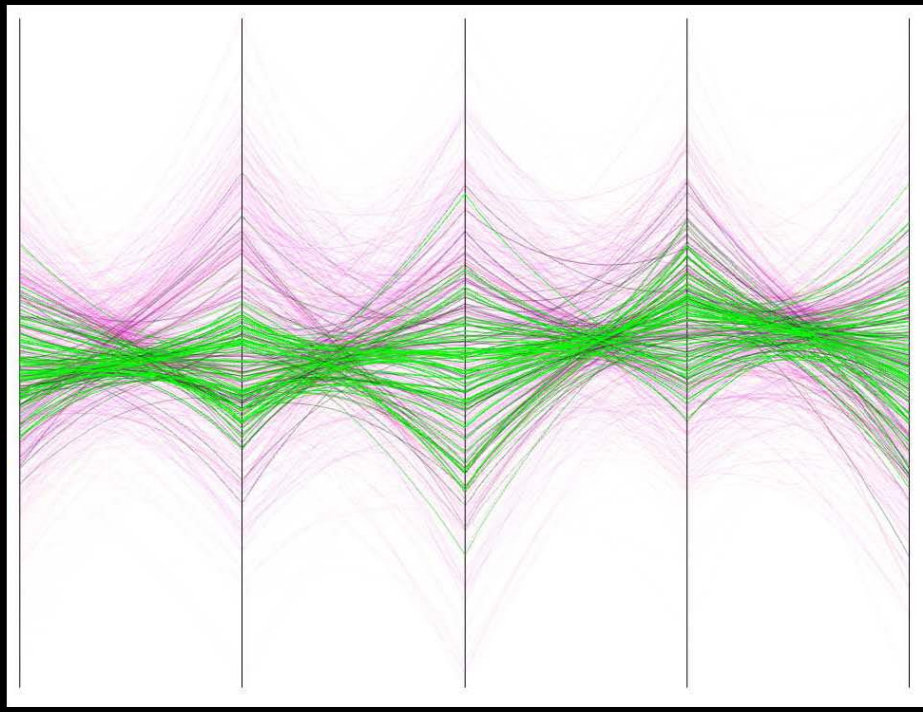

Cylinders
Horsepower
Acceleration
MPG
Displacement
Weight
Year
(a')
(b')
(c')


Cylinders
Horsepower
Acceleration
MPG
Displacement
Weight
Year
Origin:
USA
Japan
Europe

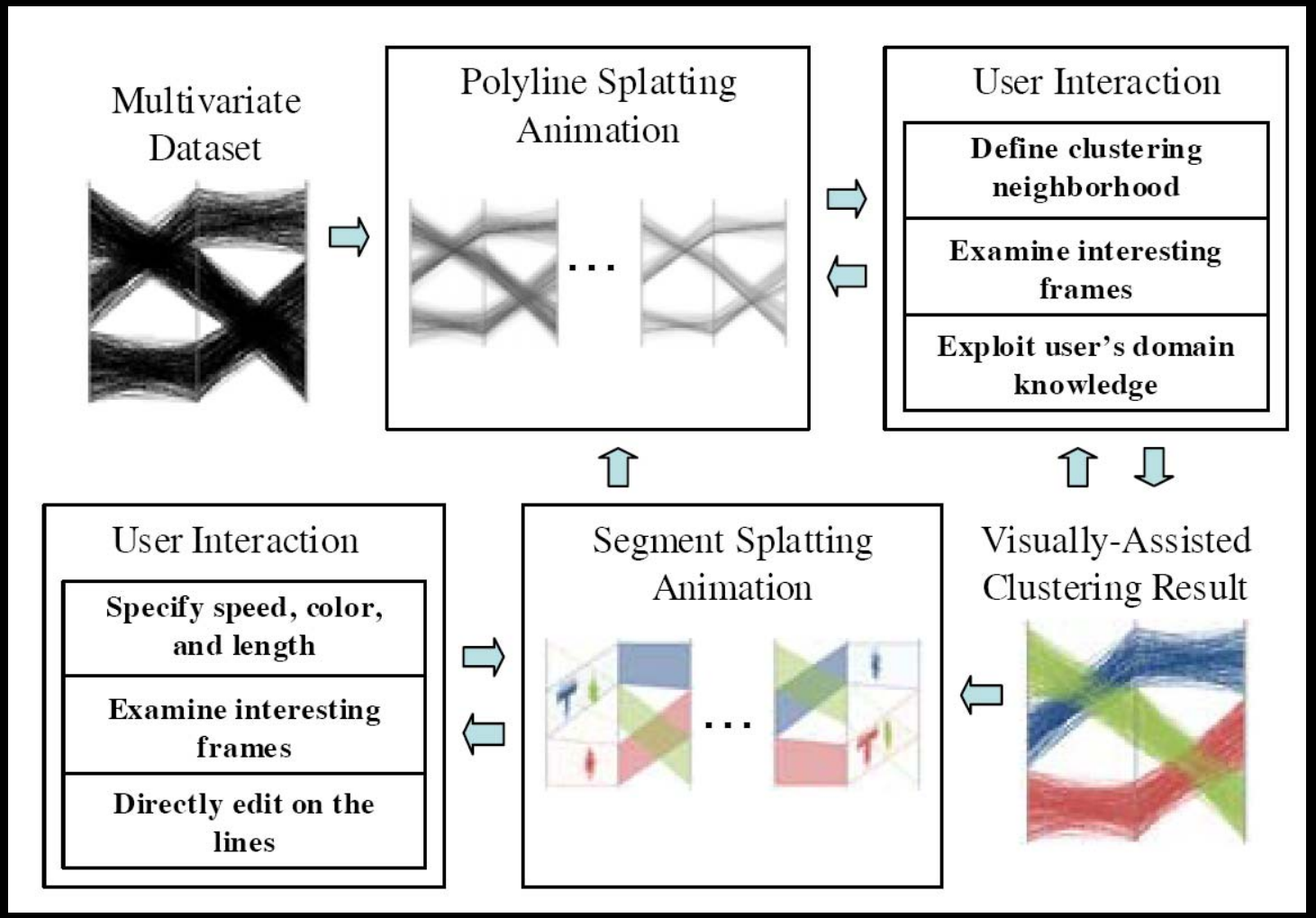

Cylinders
Horsepower
Acceleration
MPG
Displacement
Weight
Year
Origin:
USA
Japan
Europe







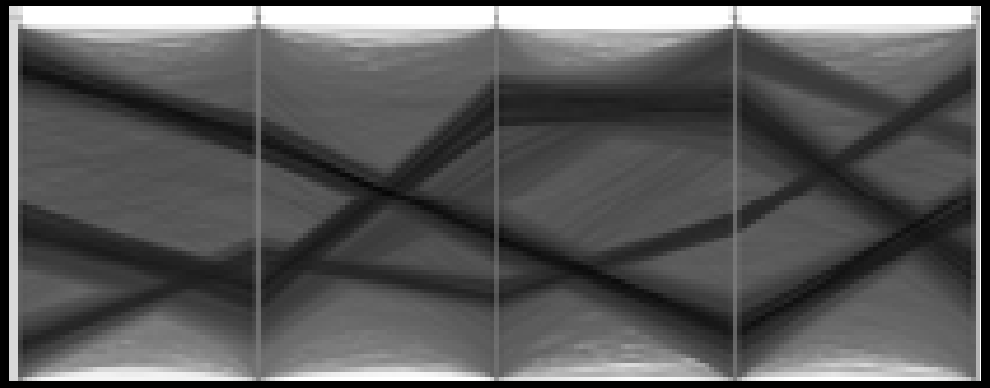
Splatting the Lines in Parallel Coordiantes

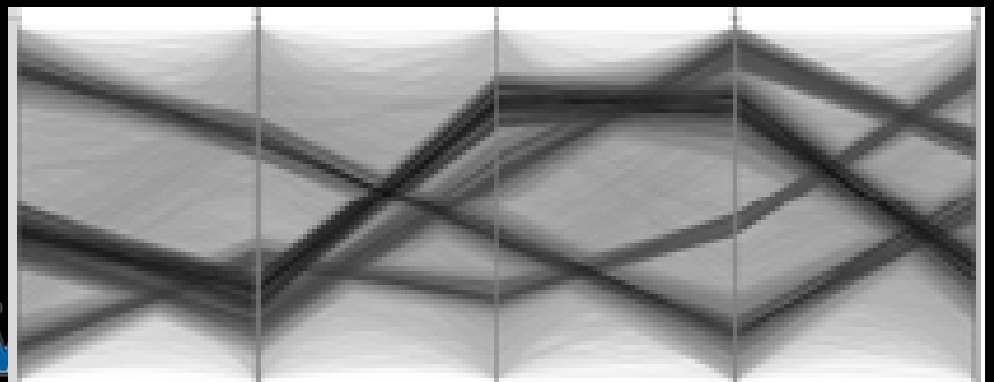


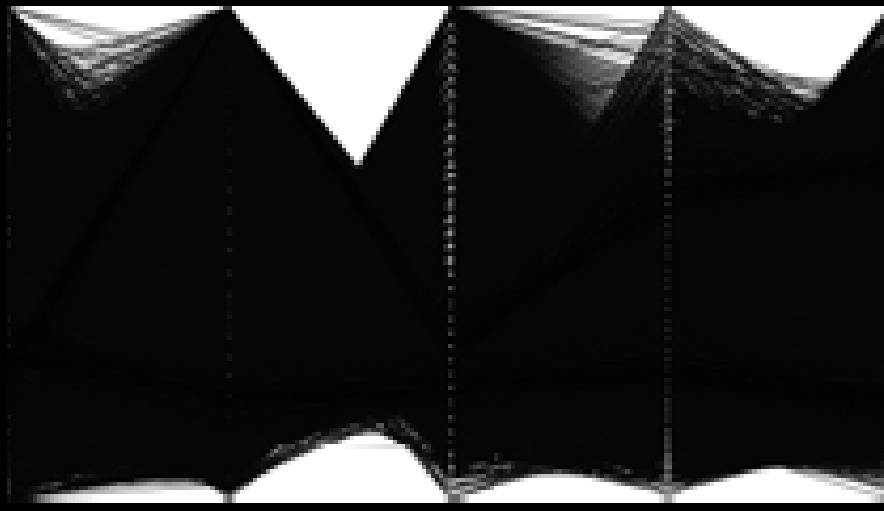
[EuroVis 2009]

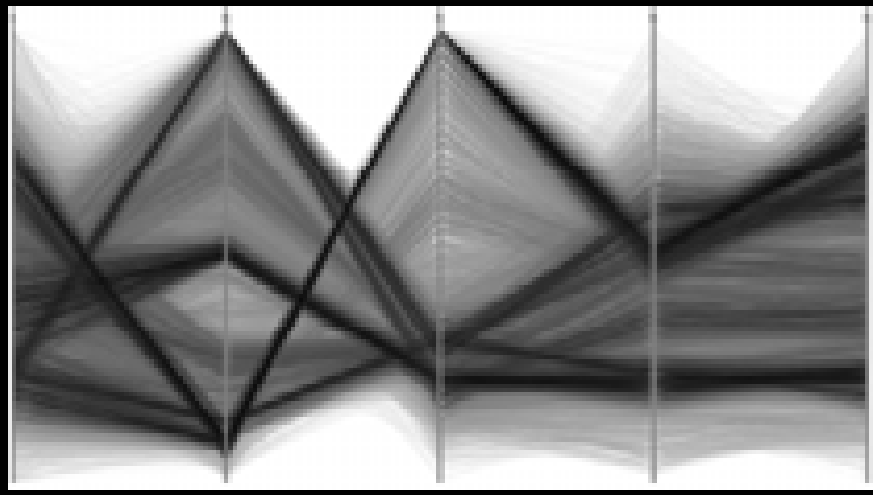
[EuroVis'09] (CGF'09)

Polyline Splatter

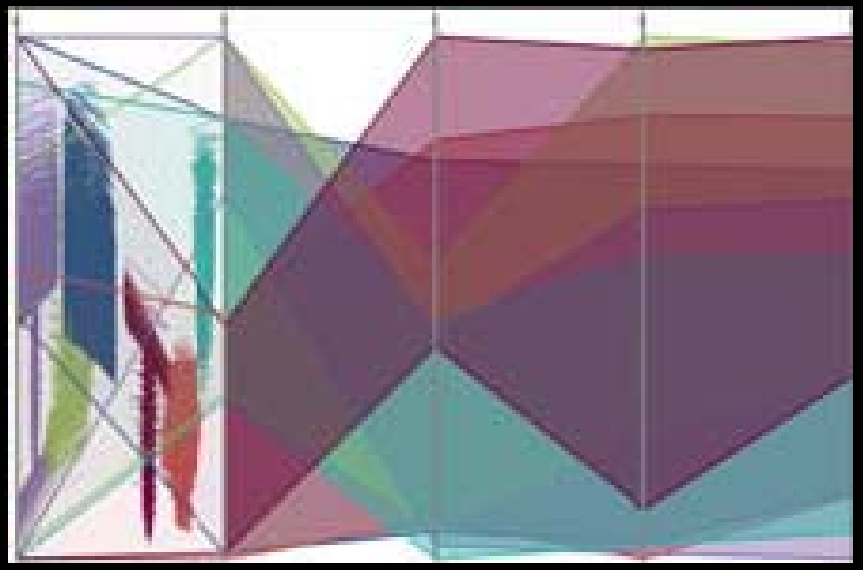


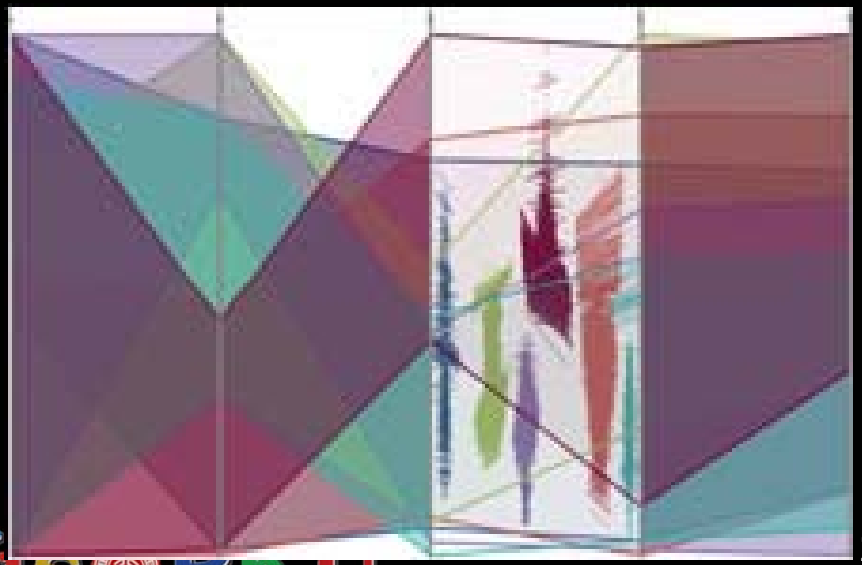


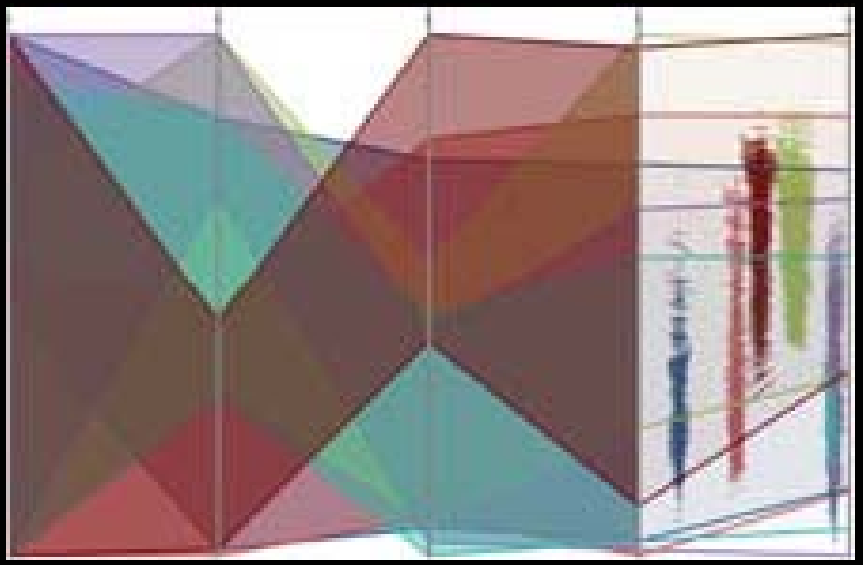




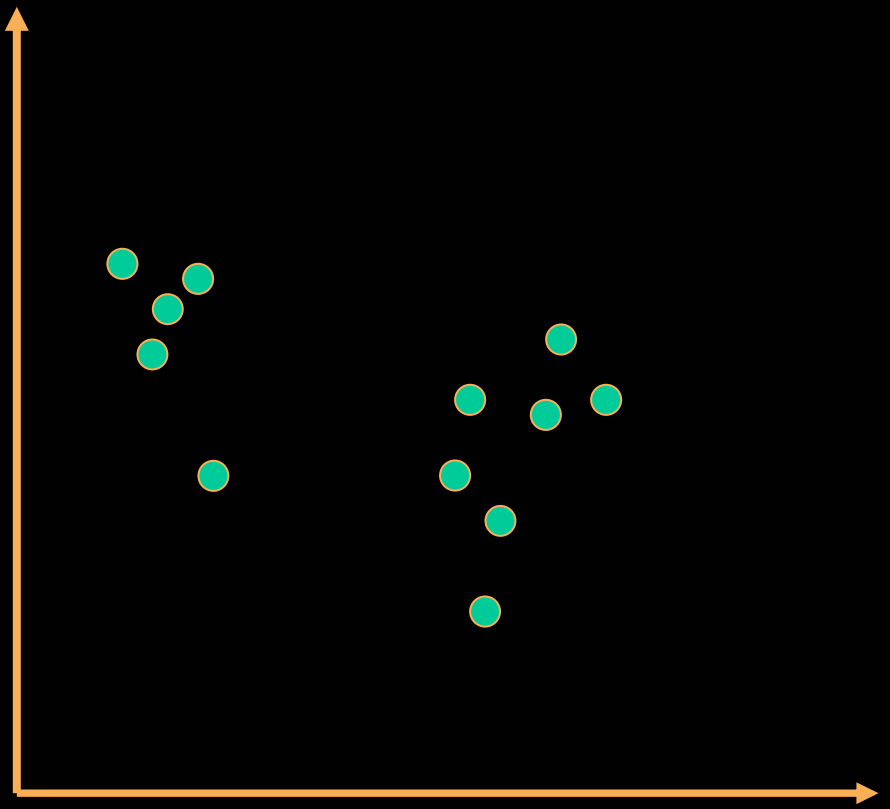
Segment Splatter



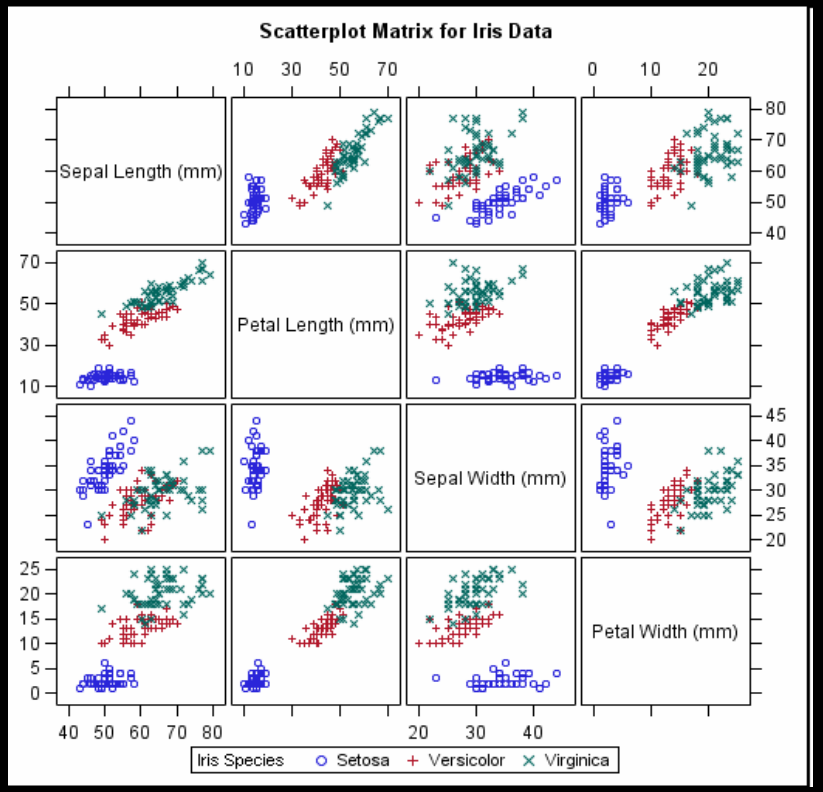




Scatter Plots



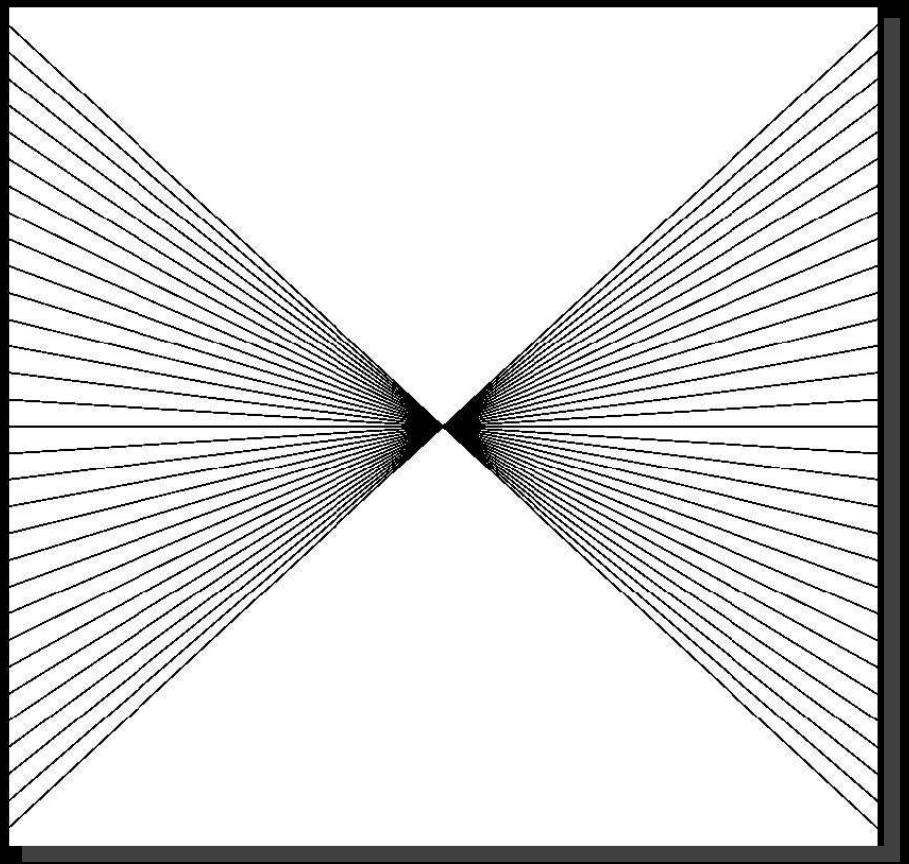
Scatter Plot

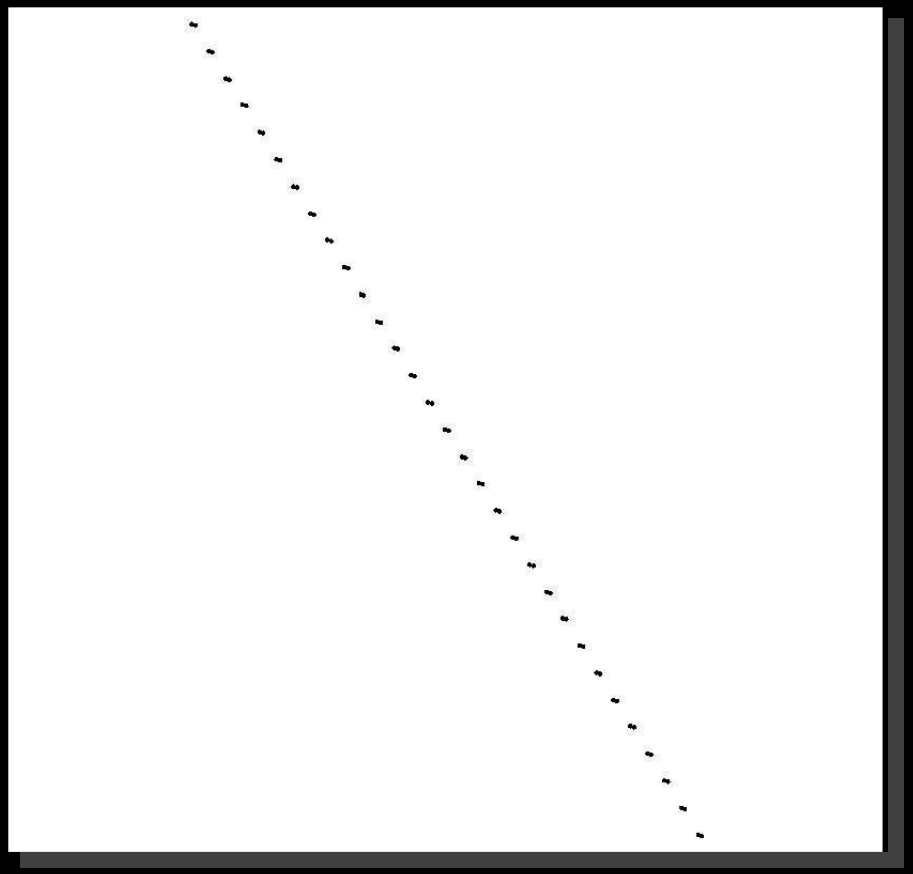


Scatter Plot Matrix

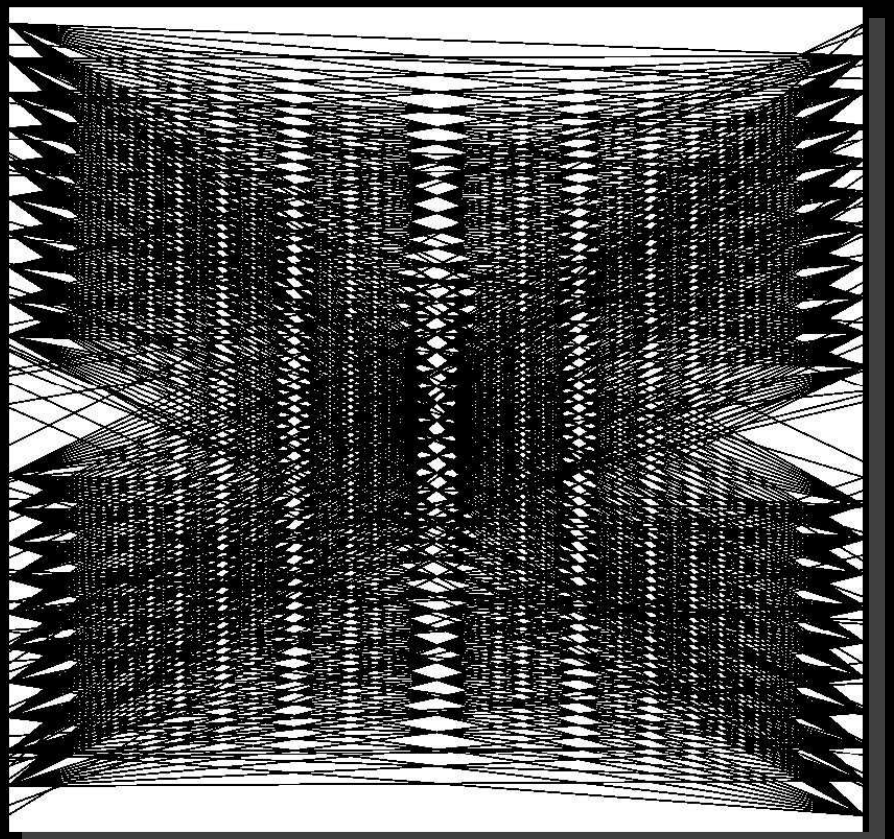
From sas.com

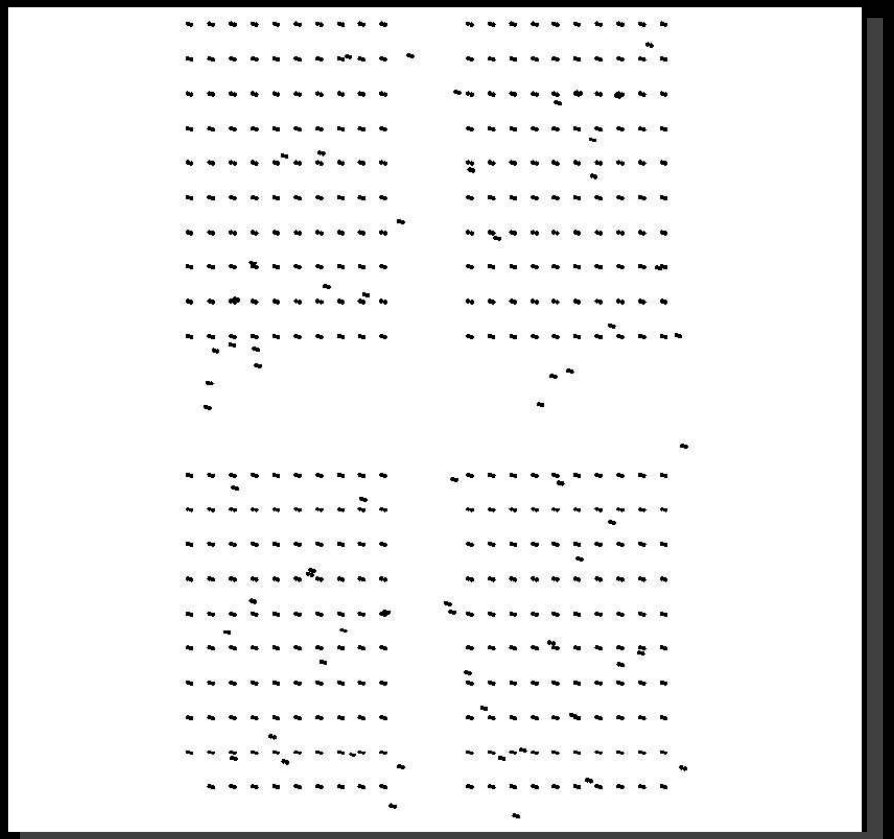
Line vs. Point Representation



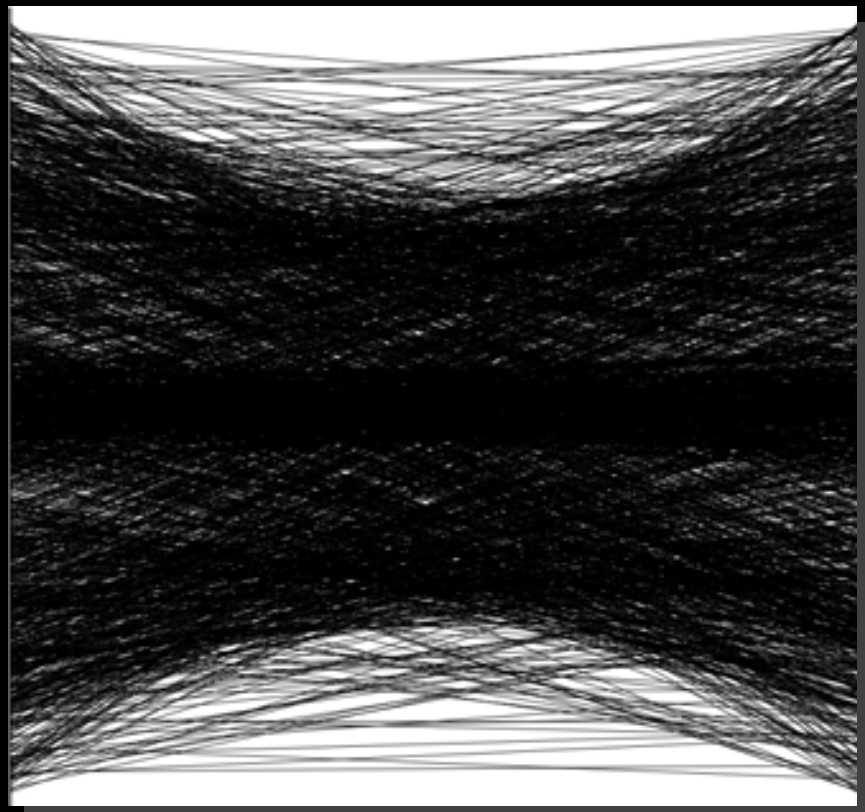


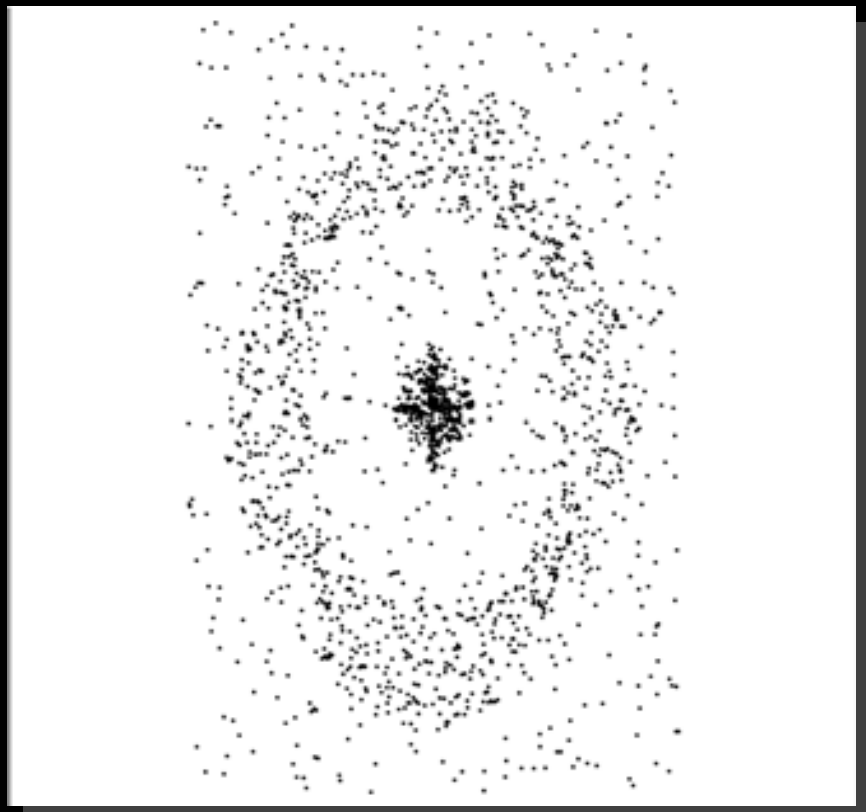
Line vs. Point Representation



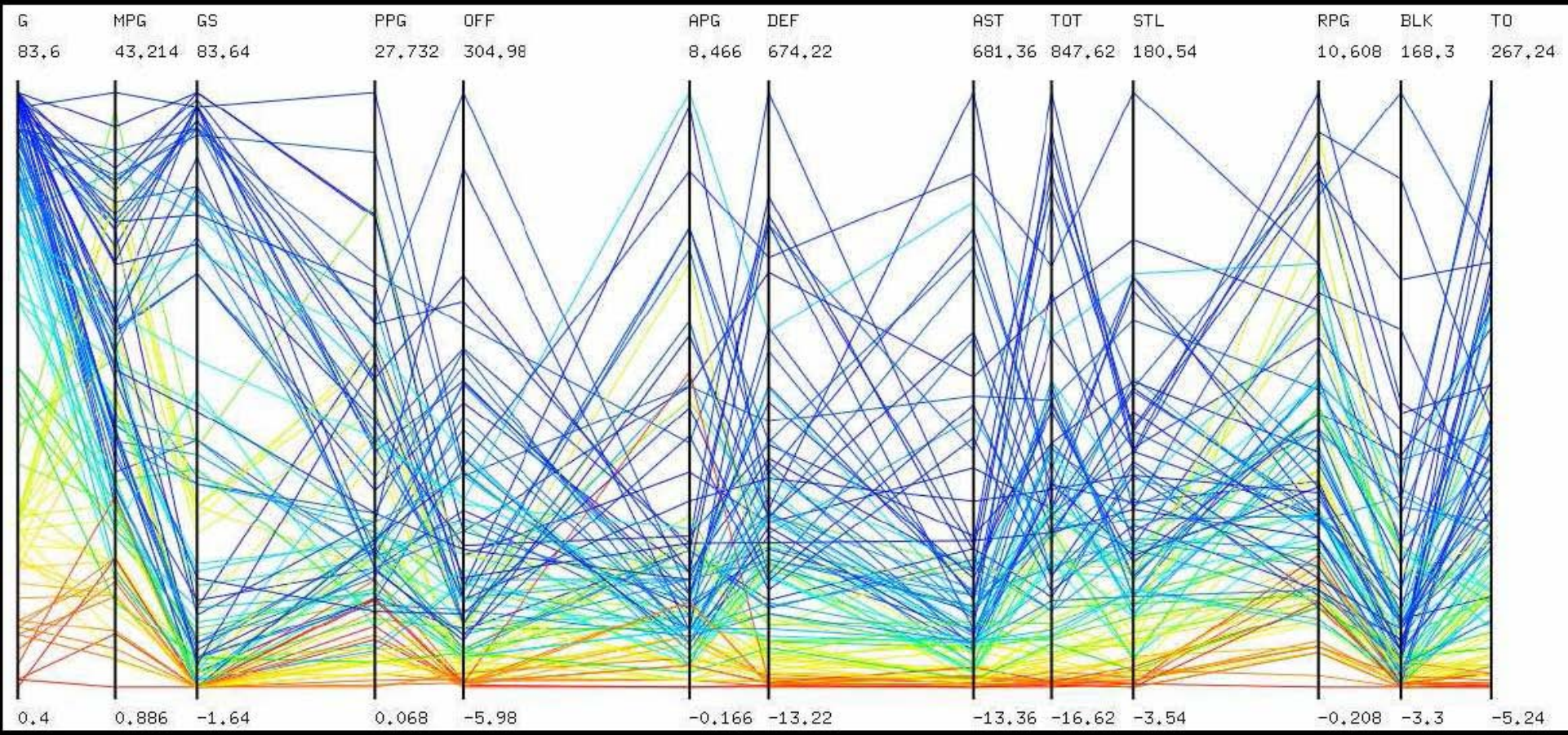


Line vs. Point Representation





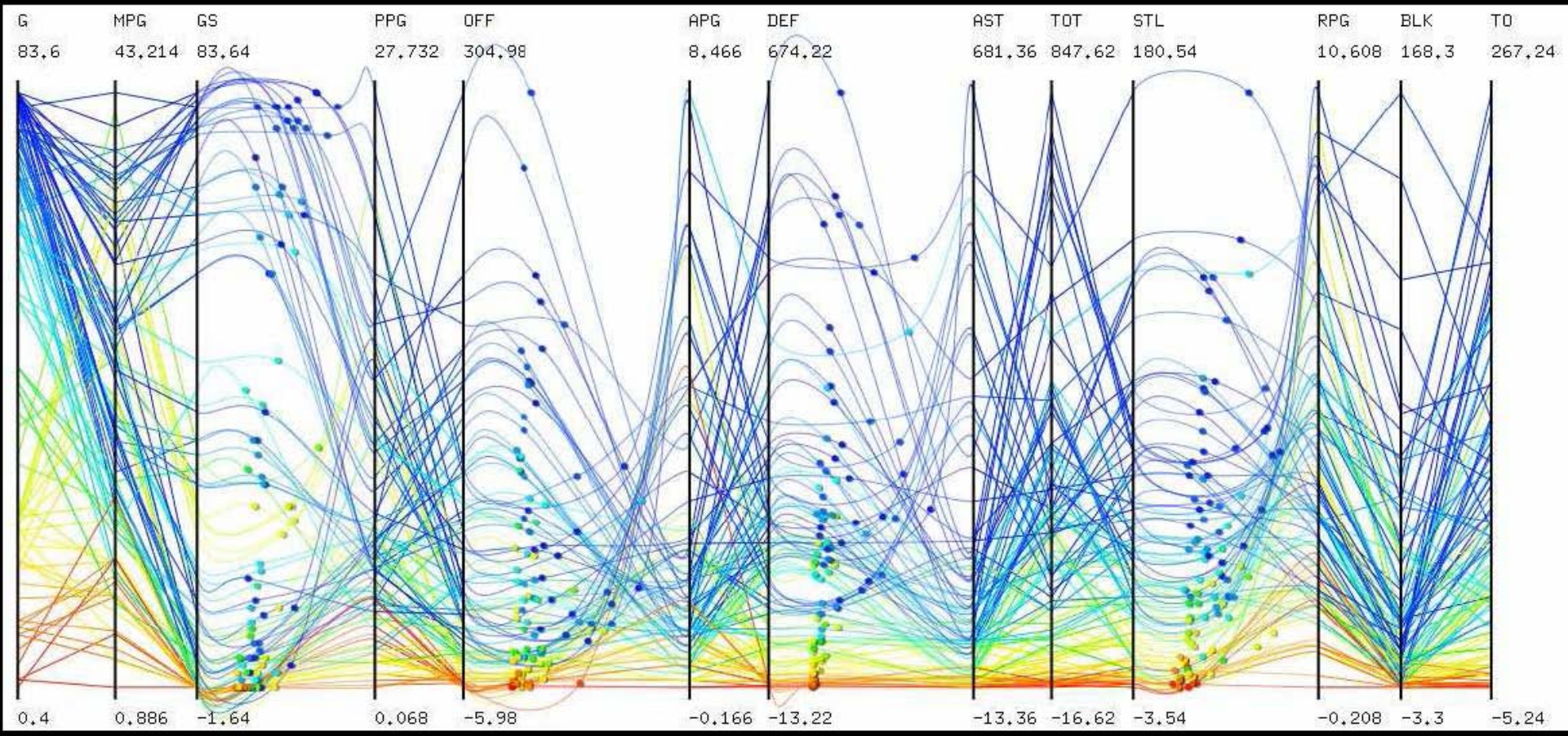
Data Exploration with PC



7 variables and 392 data items.

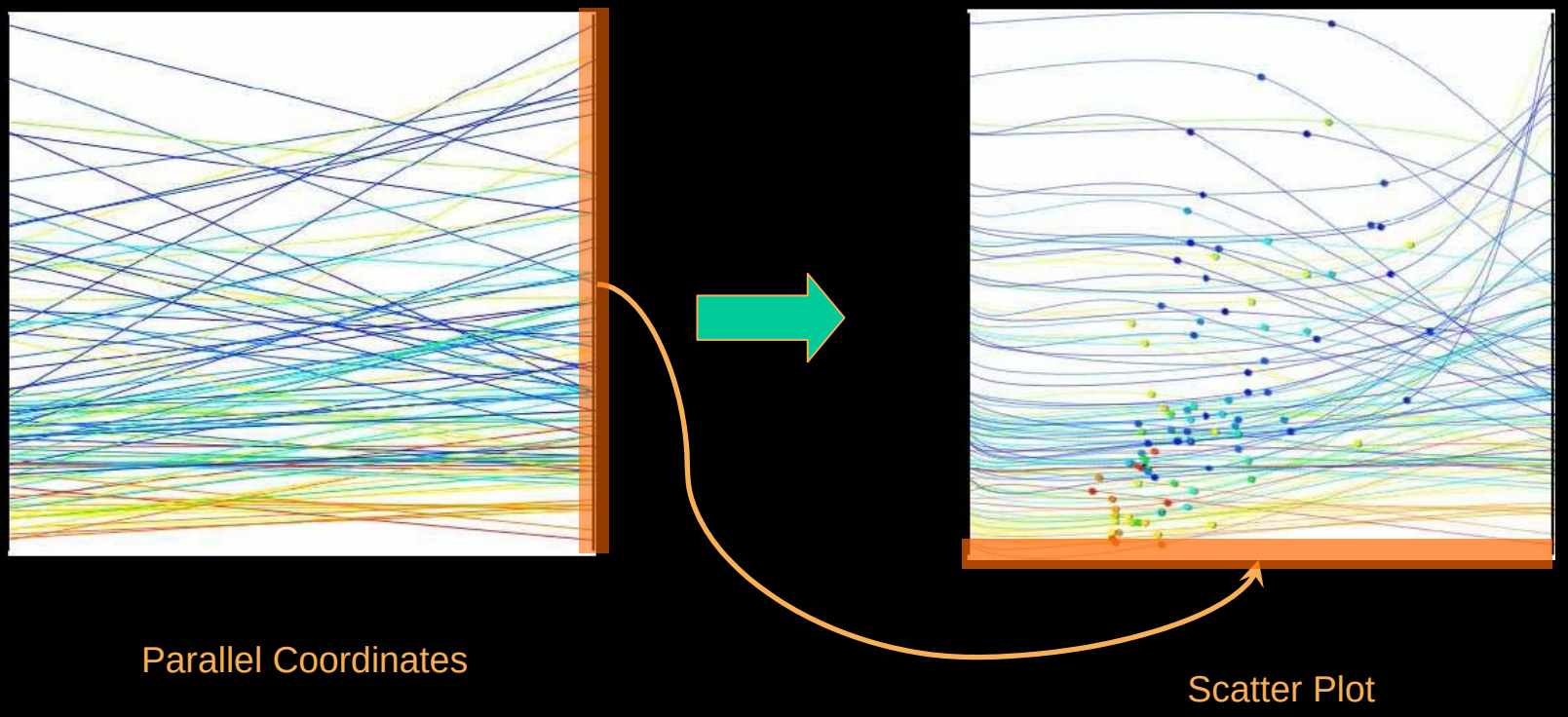
[InfoVis 2009] (Under review)

Data Exploration with SPPC



[InfoVis'09] (TVCG'09)

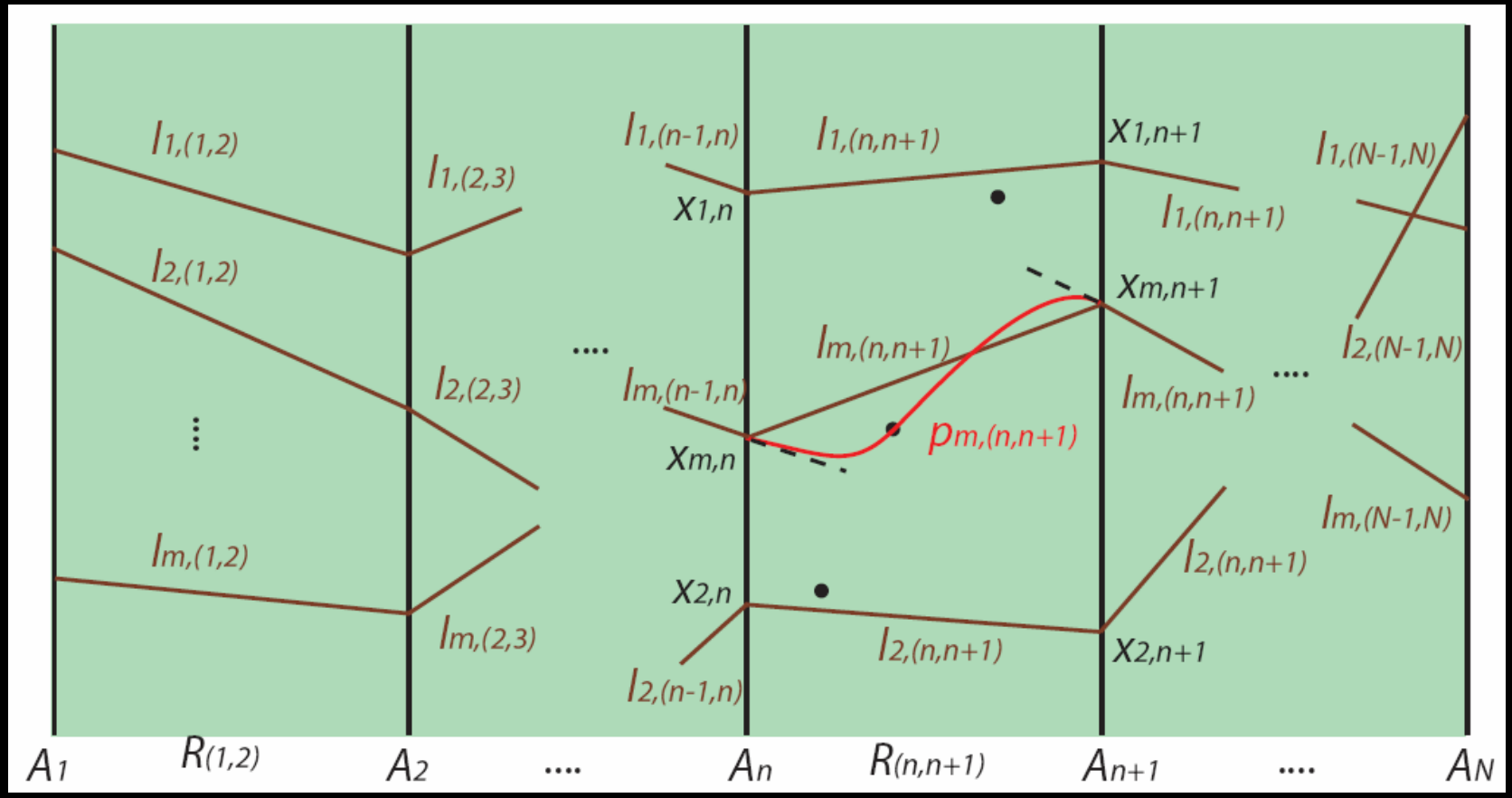
SPPC – Converting Lines into Points



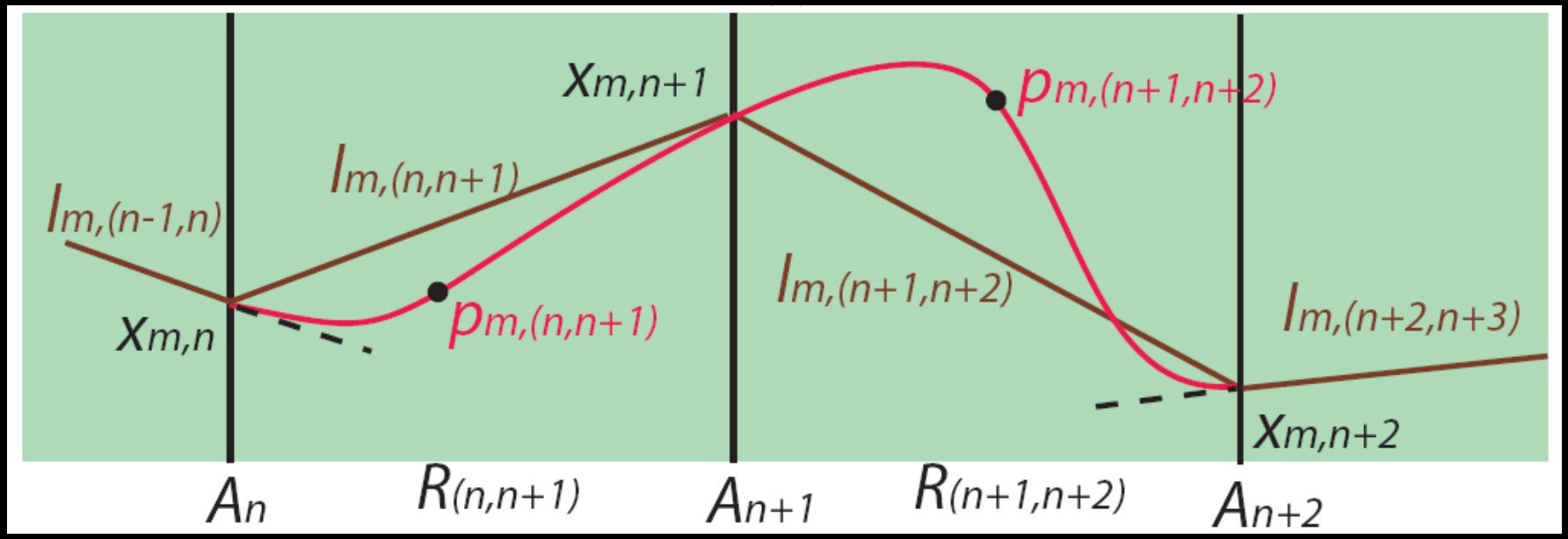
Parallel Coordinates

Scatter Plot

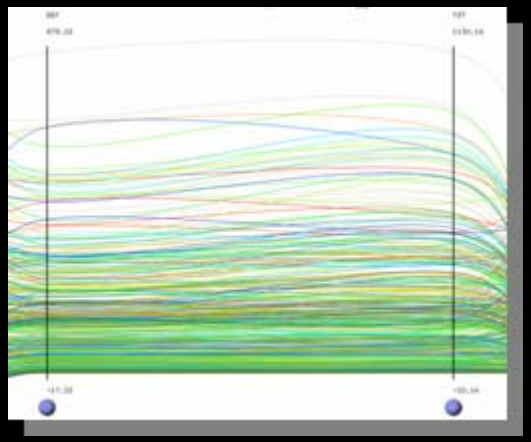
Scattering Points in Parallel Coordinates - SPPC



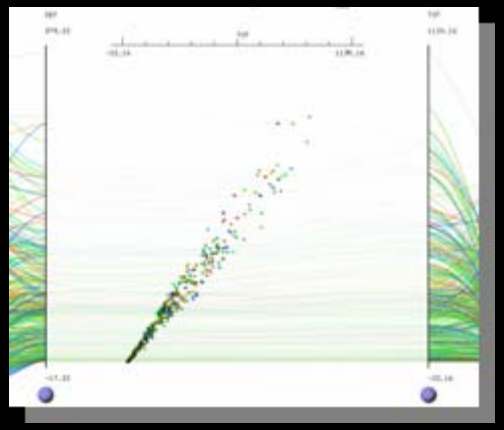
Scattering Points in Parallel Coordinates - SPPC



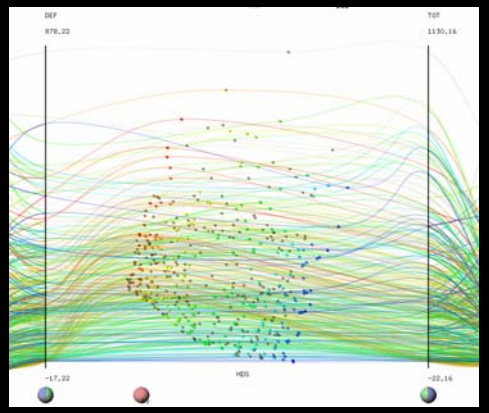
Extension in point distribution - DMS

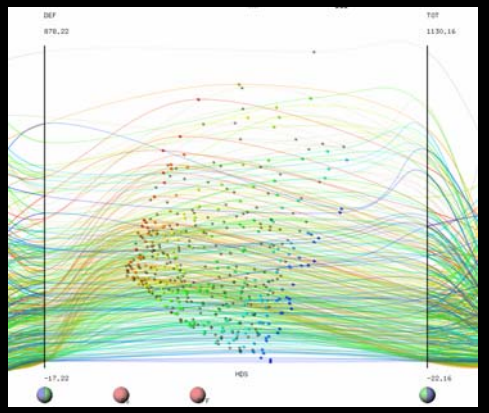


Parallel Coordinates



Scatter Plots

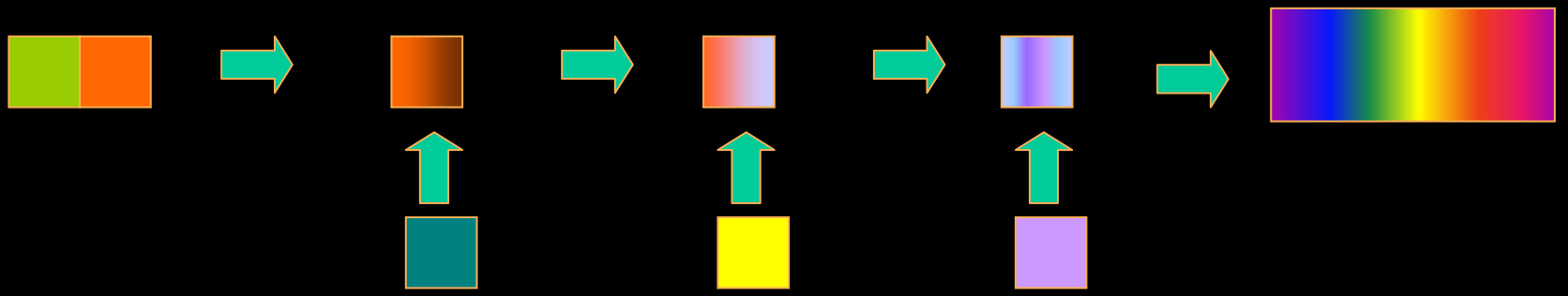




MDS

Dimensional Incremental Multidimensional Scaling (DIMDS)

MDS

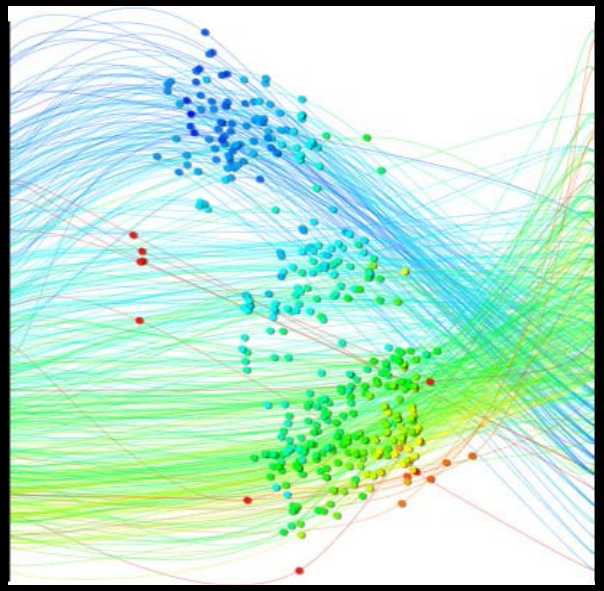


DIMDS

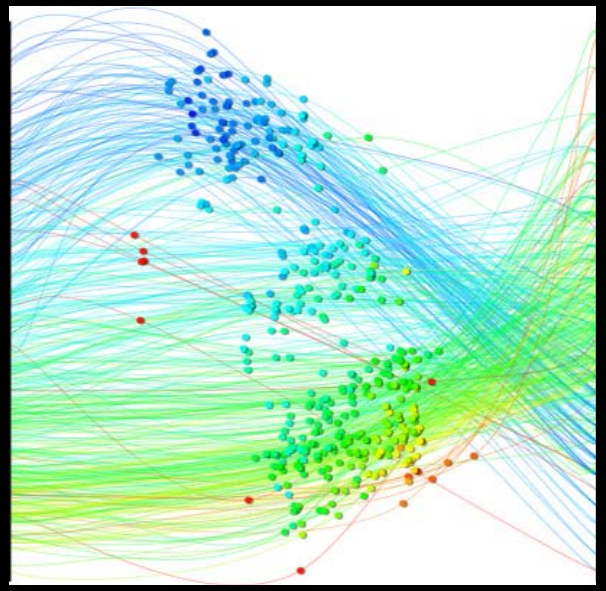
DIMDS Performance

Data	Dimensions	Data Size	CPU MDS	CPU DIMDS	GPU DIMDS	CPU DIMDS+	GPU DIMDS+
Car	8	406	5671ms	719ms (7.88)	503ms (11.3)	110ms (51.6)	92ms (61.6)
AAUP	14	1161	64312ms	1671ms (38.5)	862ms (74.6)	828ms (77.7)	593ms (108.5)
DNA	21378	68	1485ms	1453ms (1.02)	1043ms (1.42)	32ms (46.3)	81ms (18.3)

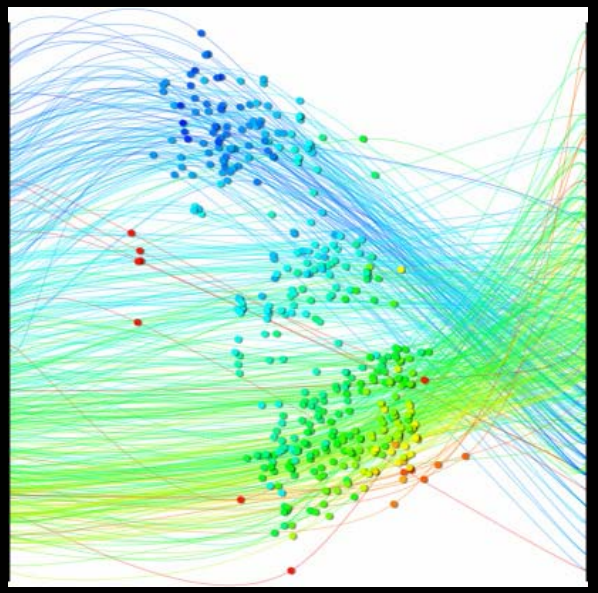
DIMDS Order



MDS

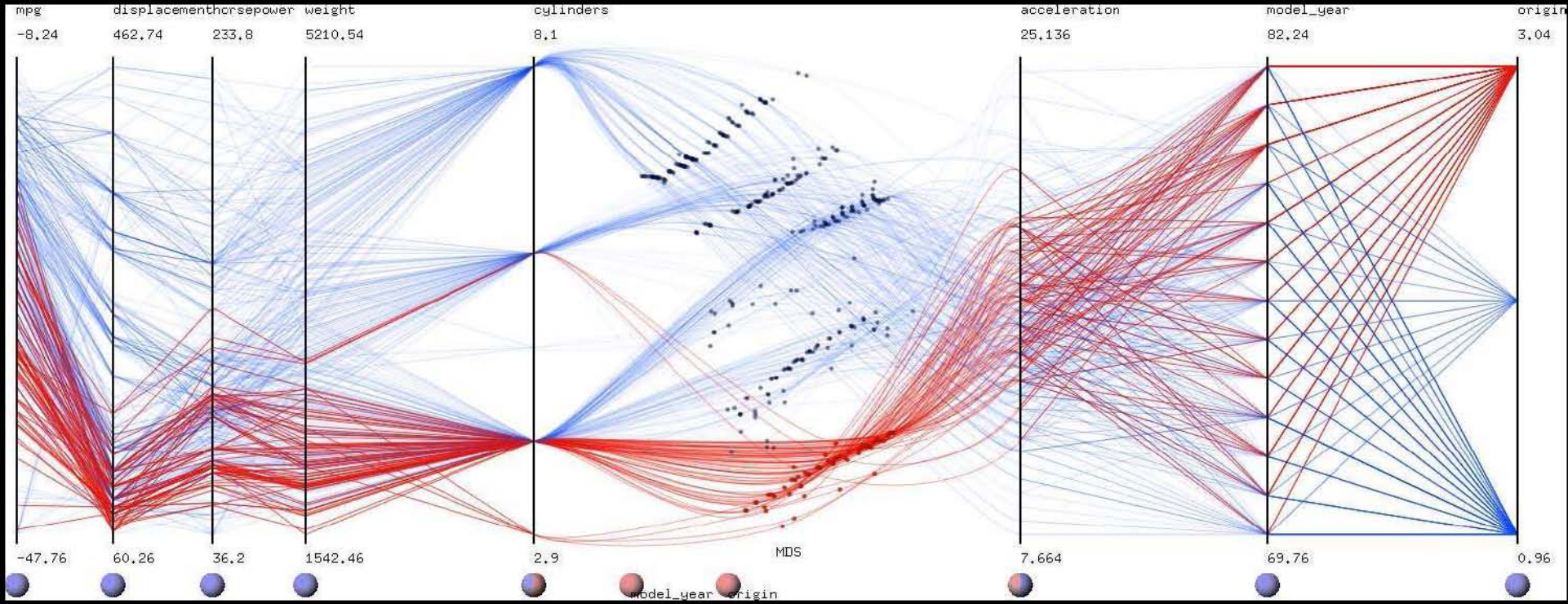


DIMDS - order 1



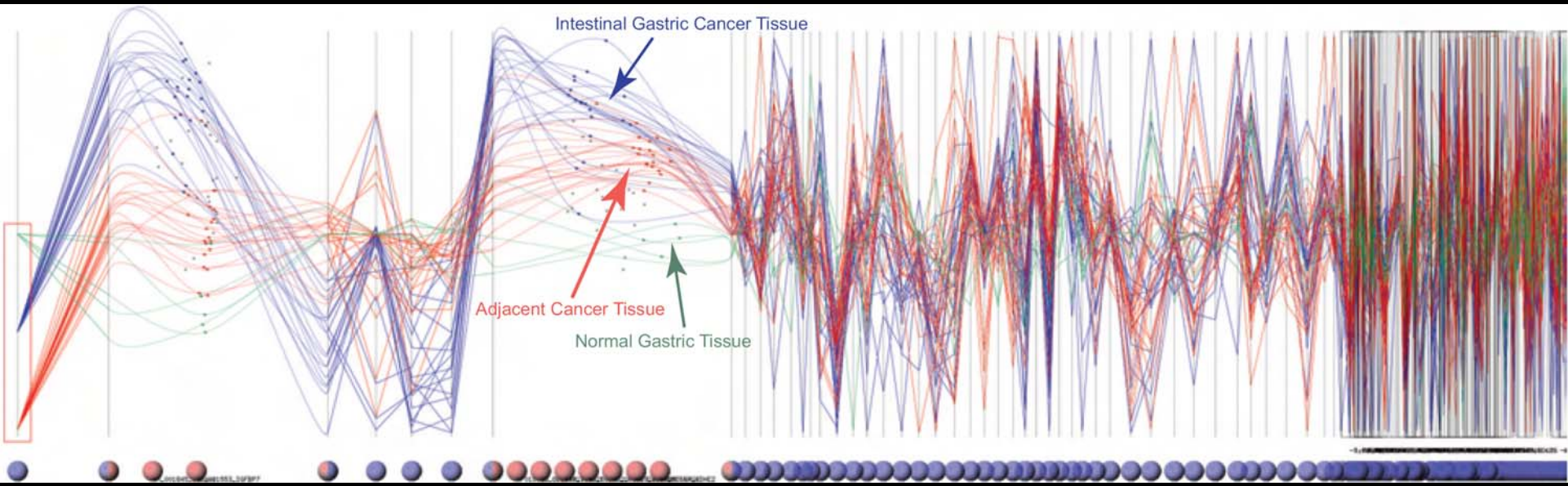
DIMDS - order 2

Data Exploration with SPPC

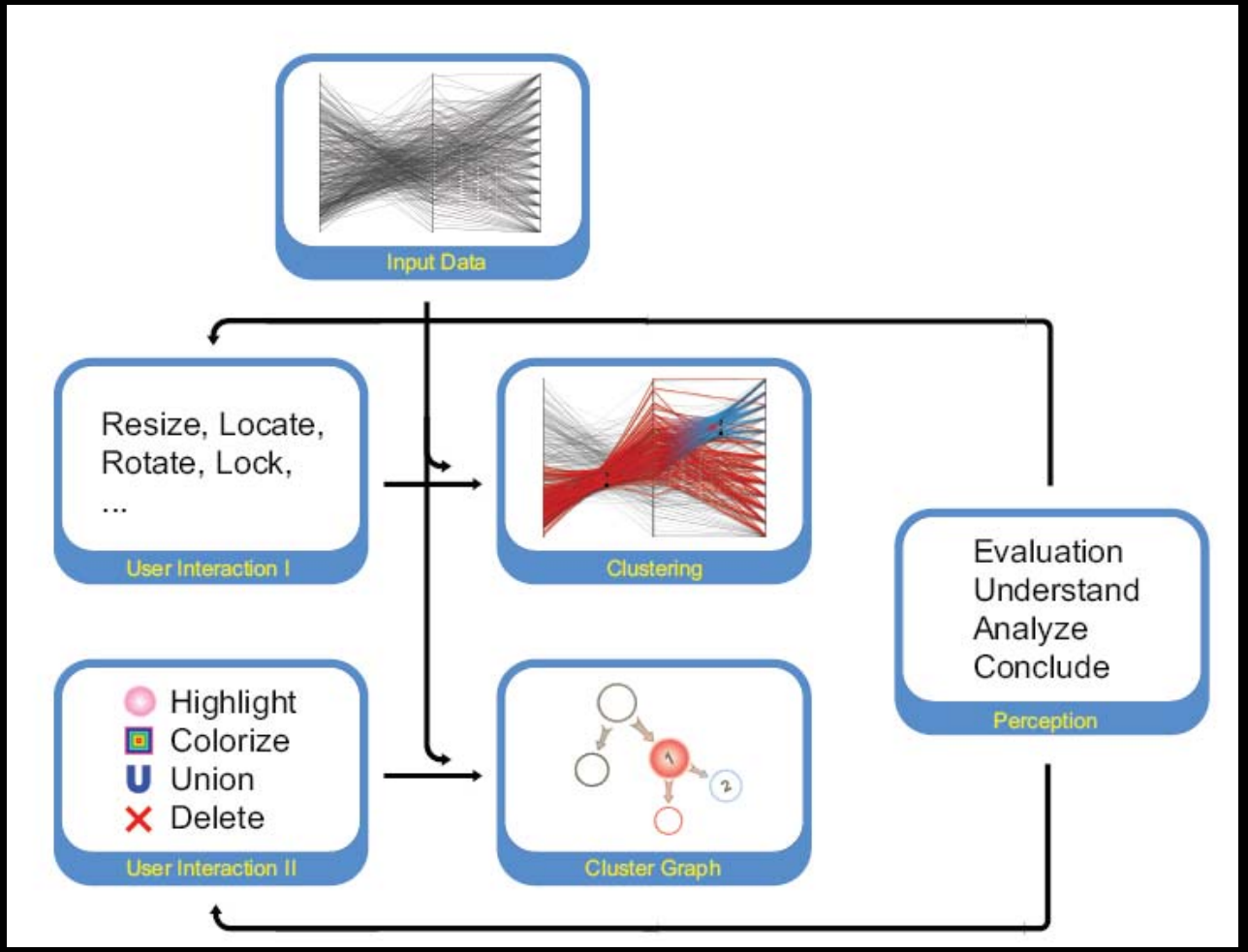


Application – Intestinal Gastric Cancer

- DNA Microarray data

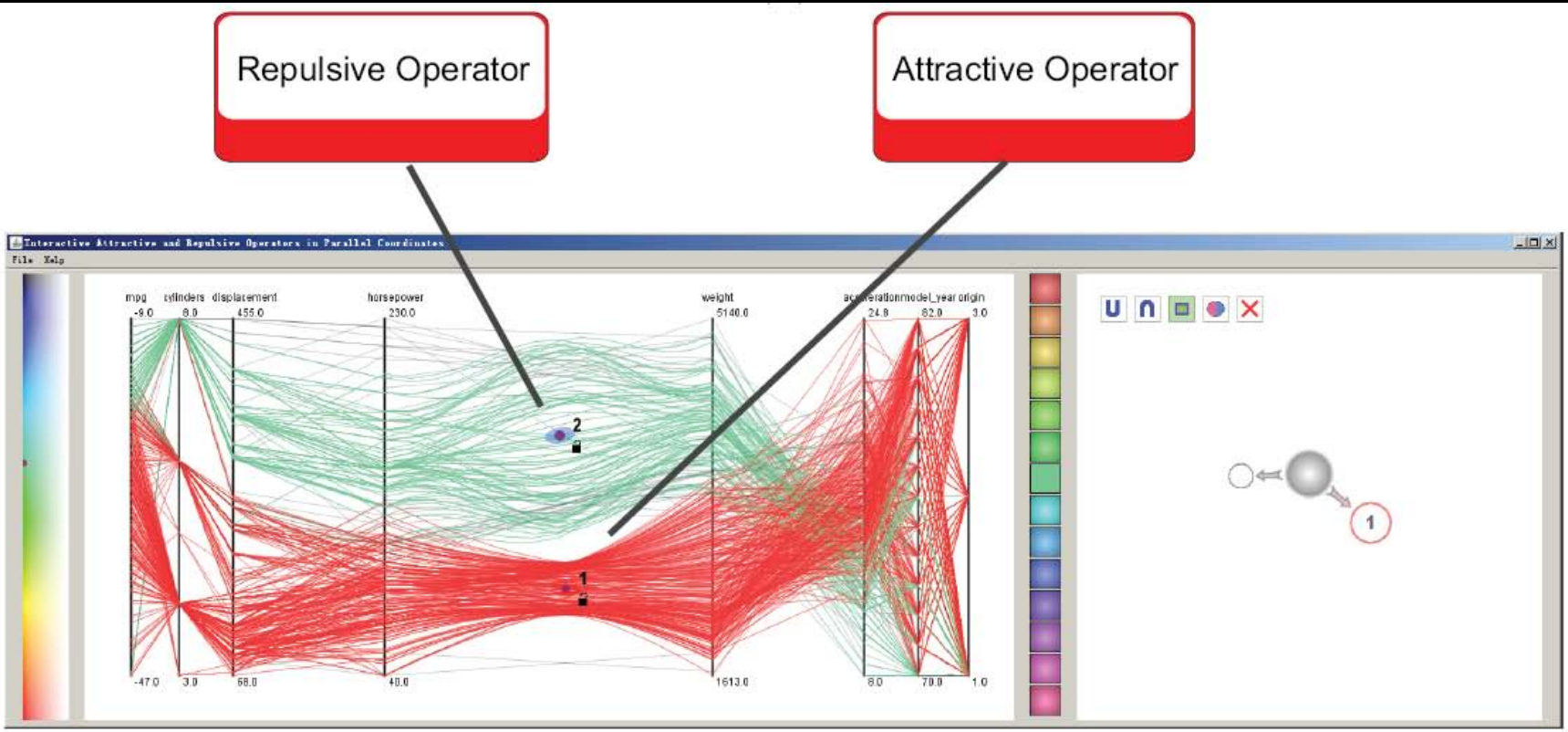


Interactive Local Clustering Operations

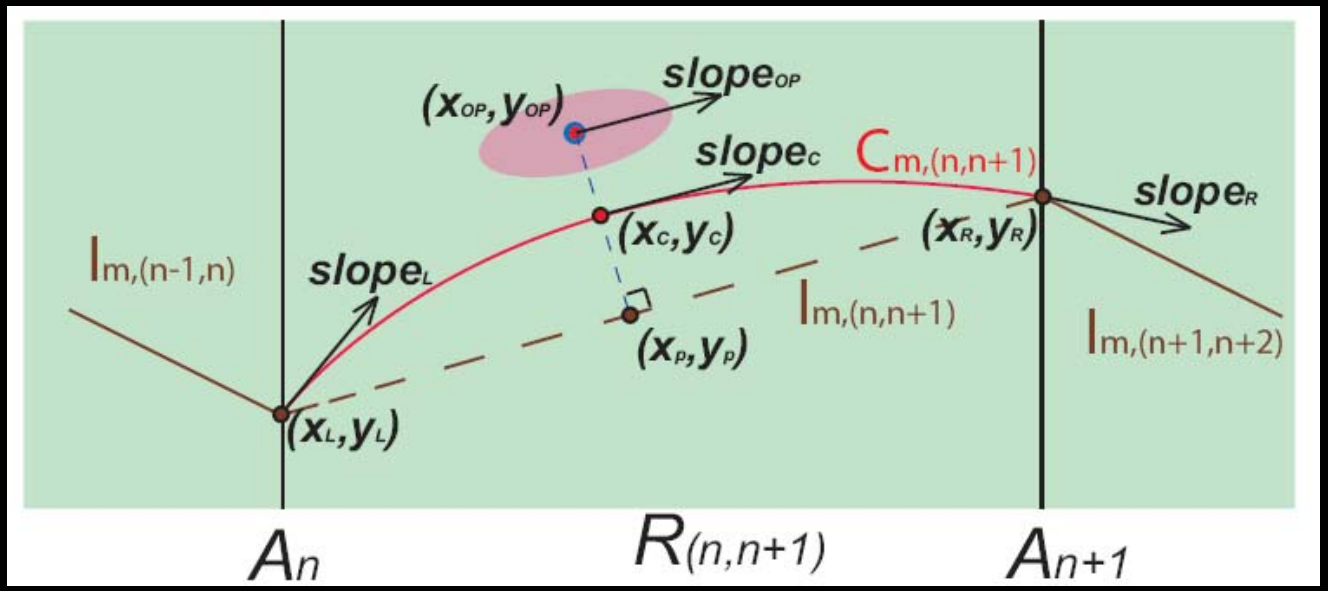


[PacificVis 2010]

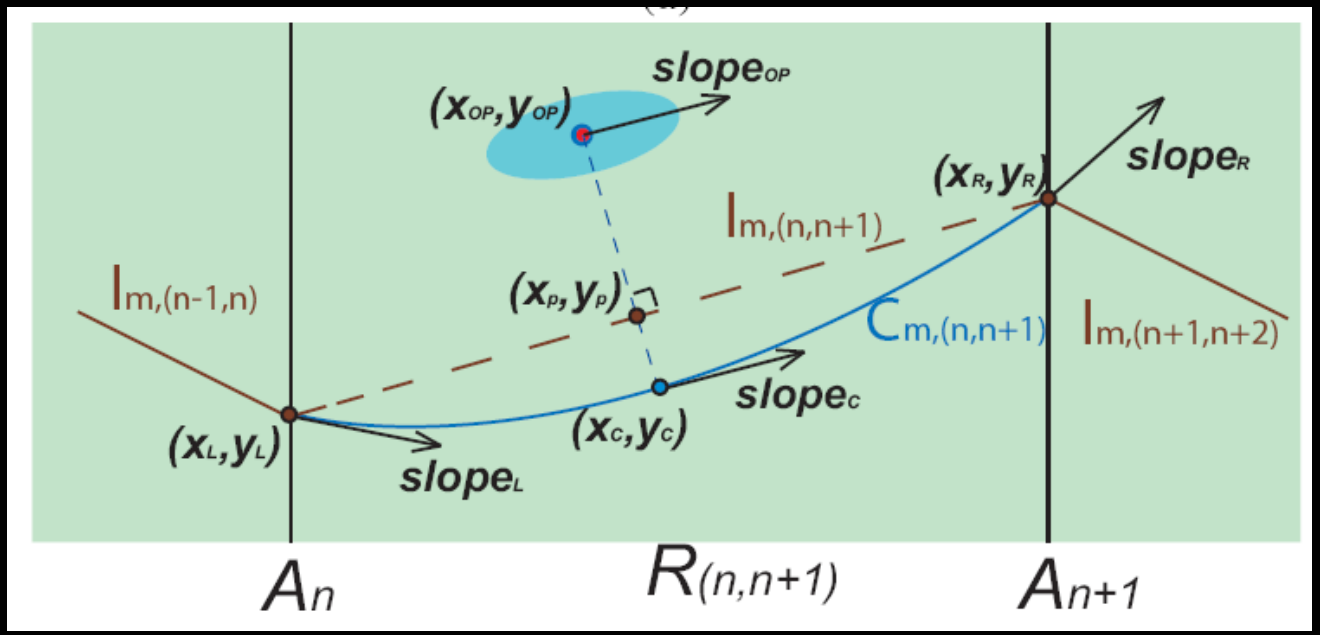
Interactive Local Clustering Operations



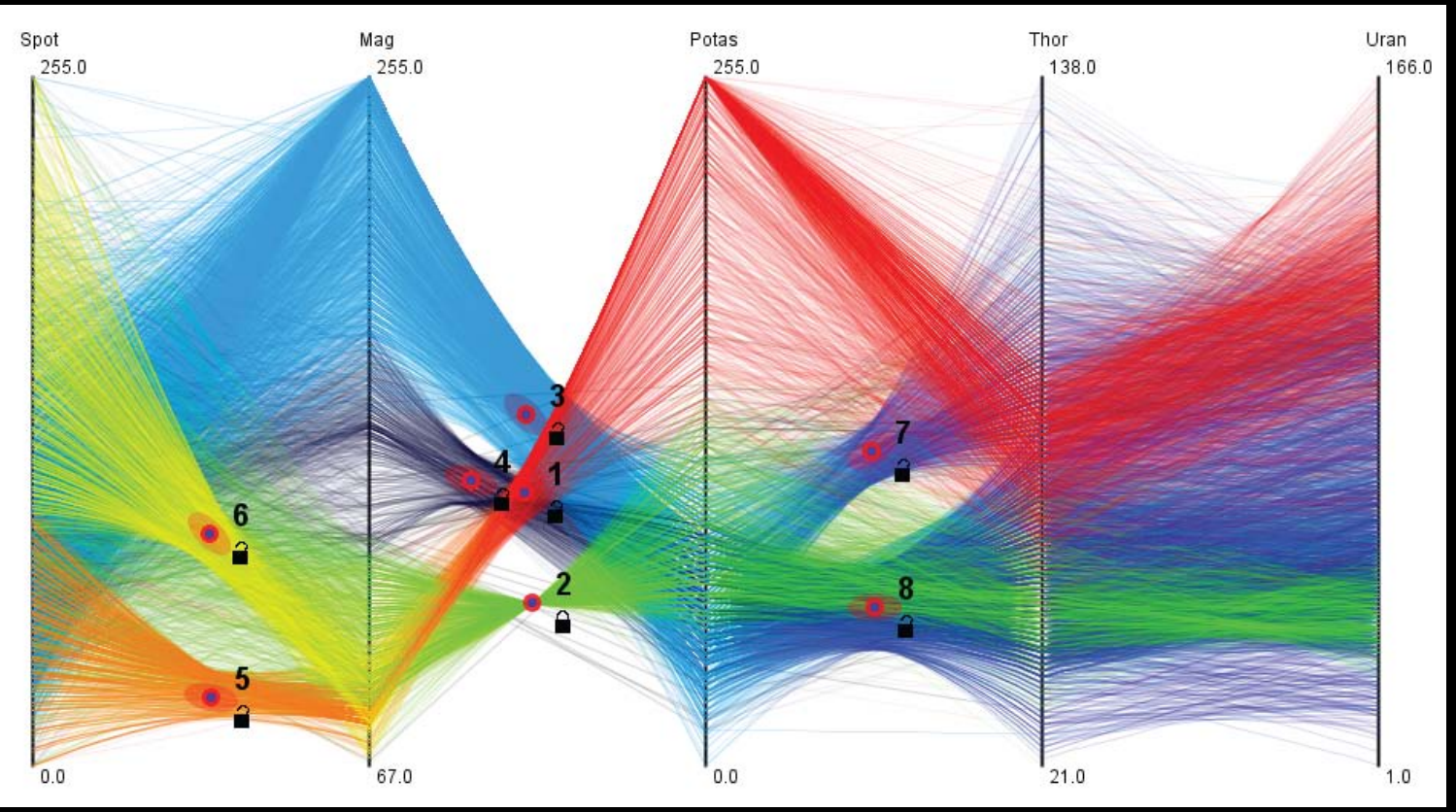
Local Clustering Operations

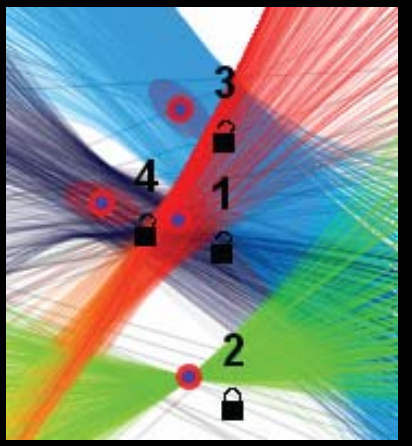


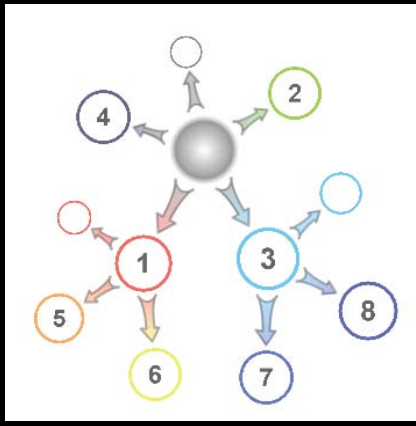
Local Clustering Operations



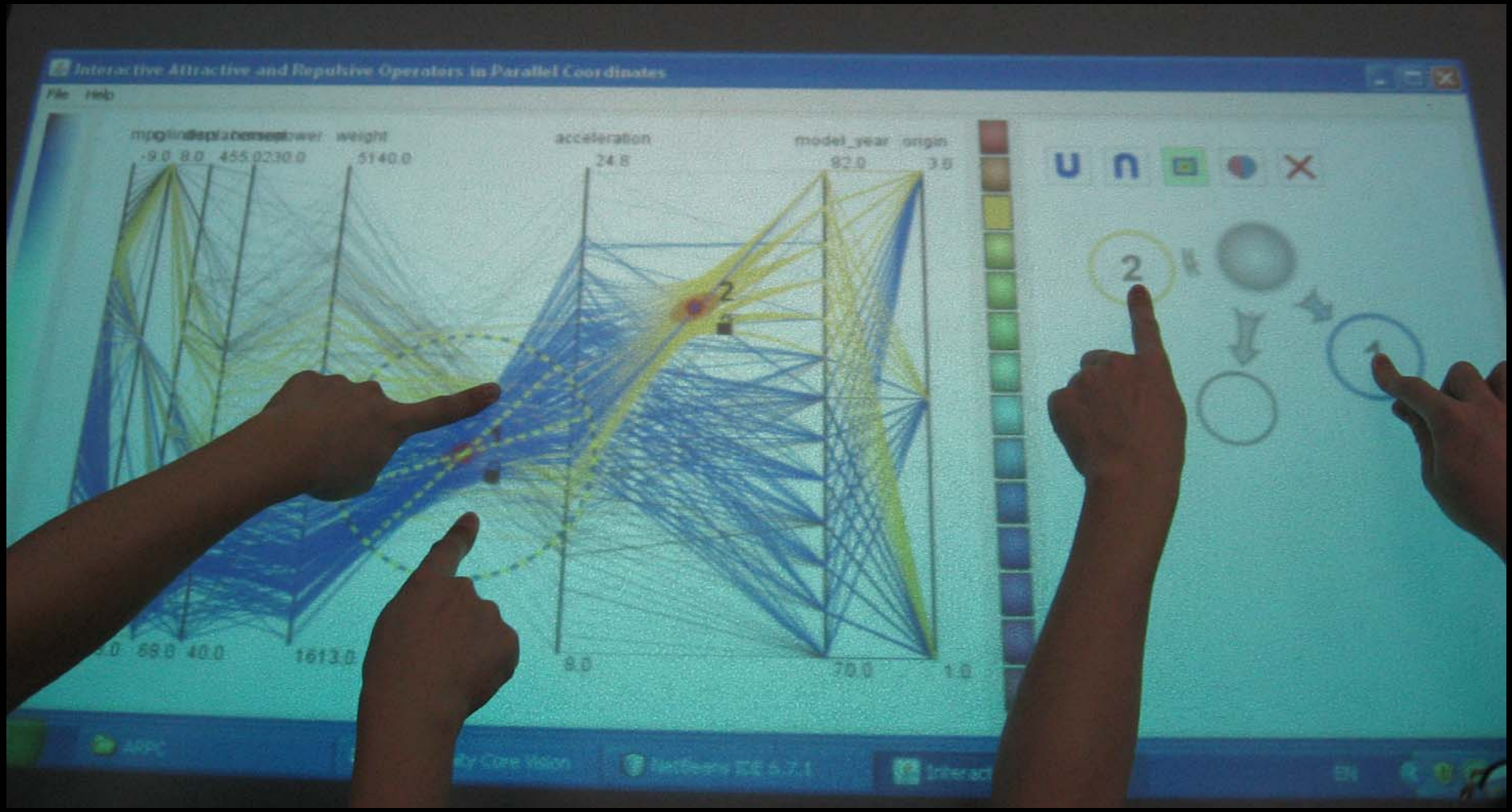
Interactive Local Clustering Operations



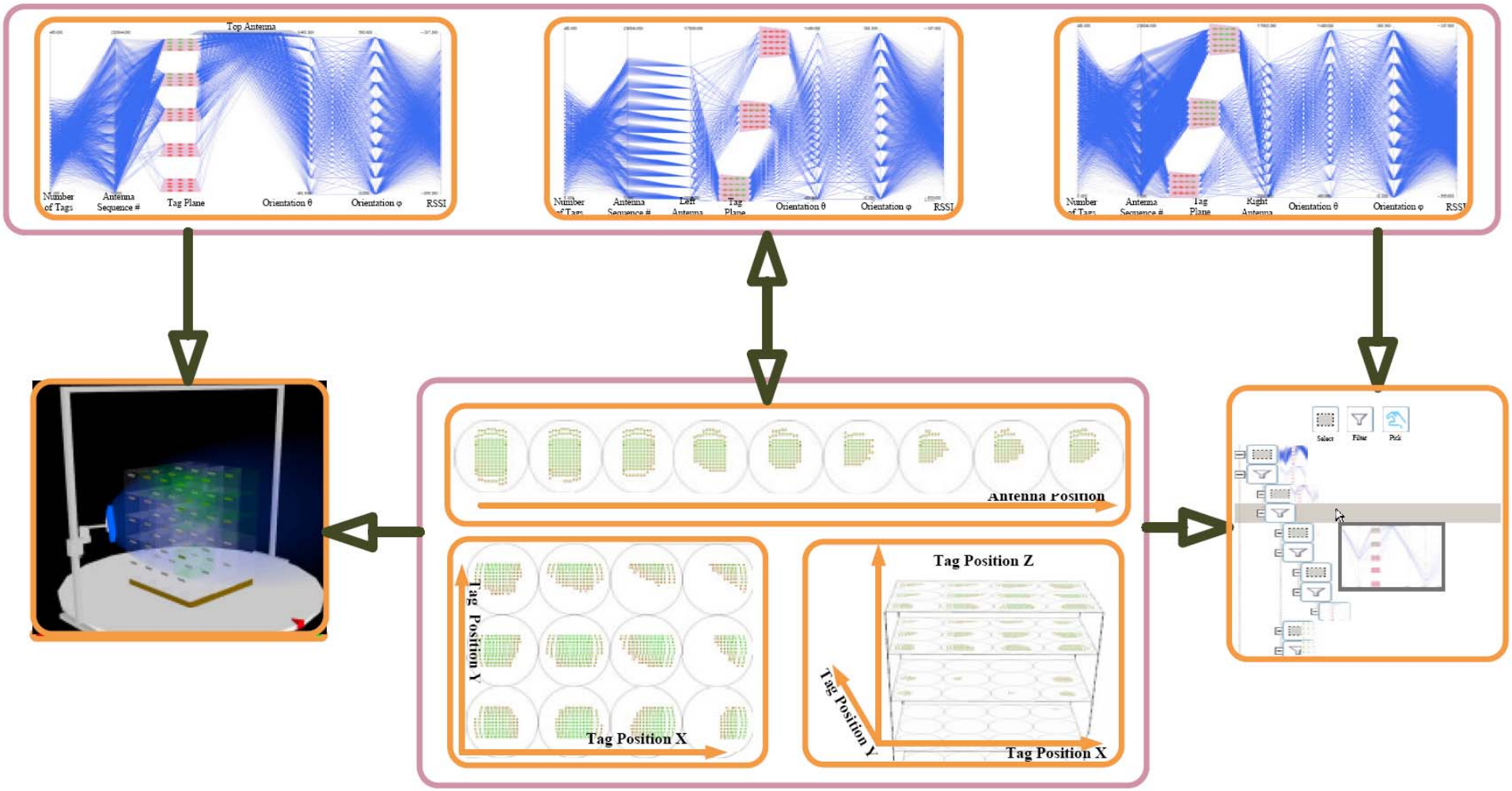




Interactive Local Clustering Operations



Parallel Coordinates Application in RFID



[Vis 2009] (TVCG'09)

Demo

- Video
- PC application (NBA data)
- Software download
  http://vis.pku.edu.cn/software/

Visualization Research at Peking University

http://vis.pku.edu.cn
http://vis.pku.edu.cn/wiki

Email: xiaoru.yuan@pku.edu.cn

VisPKU