

InfoVis – Parallel Coordinates

Sophie Steinparz, Richard Aßmair, Alois Bauer, and John Feiner
Graz University of Technology
Inffeldgasse 16c
A-8010 Graz, Austria

sophie.steinparz@student.tugraz.at
adamn@sbox.tugraz.at
alois.bauer@student.tugraz.at
j.feiner@student.tugraz.at

Contents

1	Introduction	2
2	Related Work	3
3	Data Sets and Data Preparation	4
3.1	Data Sets	4
3.2	Large Data Sets	5
3.3	Relation Between Data Points	5
3.4	Filtering	5
3.5	Metrics	5
3.6	Classification	7
4	Visualisation and Interaction	8
4.1	Visualisation	8
4.1.1	Clutter and Density Reduction	8
4.1.2	Clustering	9
4.1.3	Extent Scaling	9
4.1.4	Further visualization techniques	9
4.2	Interaction	10
4.2.1	Brushing	10
4.2.2	Axis	10
5	Interpretation	12
5.1	Notes	12
5.2	Analyse	13
5.3	Show important	13
5.4	Zoom	14
5.5	Filter and analyse further	14
5.6	Details on demand	15
6	Software and Tools	17
6.1	Open Source Tools	17
6.2	Xmdv	17
6.3	Parvis	17
6.4	PicViz	17
7	Outlook and Future	20
7.1	Possible Improvements	20
7.2	Concluding Remarks	20

8	Appendix	21
8.1	Well-Known Persons	21
8.2	Related Conferences	21
8.3	Visualisation Techniques – Advantages and Disadvantages	21

Abstract

The interpretation of data sets with many attributes – called multivariate data – is still an unsolved problem in many fields of data mining and information visualisation. The research focuses on different approaches to ease the way of displaying and interacting with large amount of data points. For example meteorological or geological data is used to analyse hurricanes and from the results of gene experiments we hope to deduct new findings about their relationship. One promising approach to handle multivariate data is the use of parallel coordinates where it is possible to display large data sets with many attributes. Each data point is drawn as a poly-line crossing several axis holding the corresponding attribute values. Fast display is granted by efficient algorithms stemming from geometry. Furthermore users are allowed to work with the data by rearranging axes, flipping axes, highlighting important lines via brushing, or filter data interactively. In this paper the steps for quick and successful interpretation of parallel coordinates are shown. By applying filtering, rearranging and brushing techniques the example cereal data set can be analysed and relation between cereals get obvious. Based on those findings the readers might interpret other data sets more easily in the future.

Chapter 1

Introduction

Constantly growing data requires more effective ways of presenting them without losing the overview. Information visualisation proceeds the assumption that a visual data representation allow us to understand, explore large data sets in an intuitive way. However, this is where techniques for visualising large data sets come into play, one of the general methods for multivariate data currently used is the *Parallel Coordinates*(PC) technique in Fua et al. [1999]. High dimensional data is being visualized and multivariate data is being analyzed for drawing conclusions. Such conclusions can implicate predicting seasonal hurricane activity in Chad A. Steed and Edward [2008], analyzing EEGs (electroencephalography) for finding relevant differences between healthy people and patients in ten Caat et al. [2007].

Multivariate data consists of more than one dimension/variable. Each axis represents a variable of the data set. The N-axis are drawn as vertical lines with equal spacing. Each data element displayed is a series of connected points along the dimensions. The higher the dimension, the higher the number of axis.

In this paper our primary goal is to survey the kinds of filter and displaying techniques for the PC technique, which are most useful. This paper is structured in the following way: First an overview about the history and concepts of PC is given in Chapter 2. Then the data acquisition is described in Chapter 3. Data visualisation and interpretation is discussed in Chapter 4. Some free available open source tools and software used for parallel coordinates are described in Chapter 6.

In this paper the usage of parallel coordinates is explained with many examples based on cereal data. The sample data is provided by statlib¹ or statcrunch² (The data set of BLS POS has been described by Wegman and Dorfman [2003] also). The import, the visualisation and different steps of modification (interaction) as well as the interpretation has been done by the authors themselves.

¹<http://lib.stat.cmu.edu/datasets/1993.expo/cereal>

²<http://www.statcrunch.com/5.0/index.php?dataid=285920>

Chapter 2

Related Work

In this chapter we provide an overview and background information for PC and related works. The general motivation for PC is finding some relevant information within multivariate data sets, e.g. hurricanes, Geo-Maps, EEGs, Computer Security.

PC were originally developed by Alfred Inselberg in 1985 in Inselberg [1985]. His motivation was due to the absence of visualization while studying multi-dimensional geometry. He experimented, put coordinate axes parallel to each other to finally get the representation seen in Figure 2.1. Each n-dimensional polyline is a set of connected points which intersects the y-axis ($x_1 \dots x_5$) at the value the data point has in the corresponding dimension. Figure 2.1 shows a five-dimensional point $C = (c_1, c_2, c_3, c_4, c_5)$ represented by polygonal line in a parallel coordinates plot.

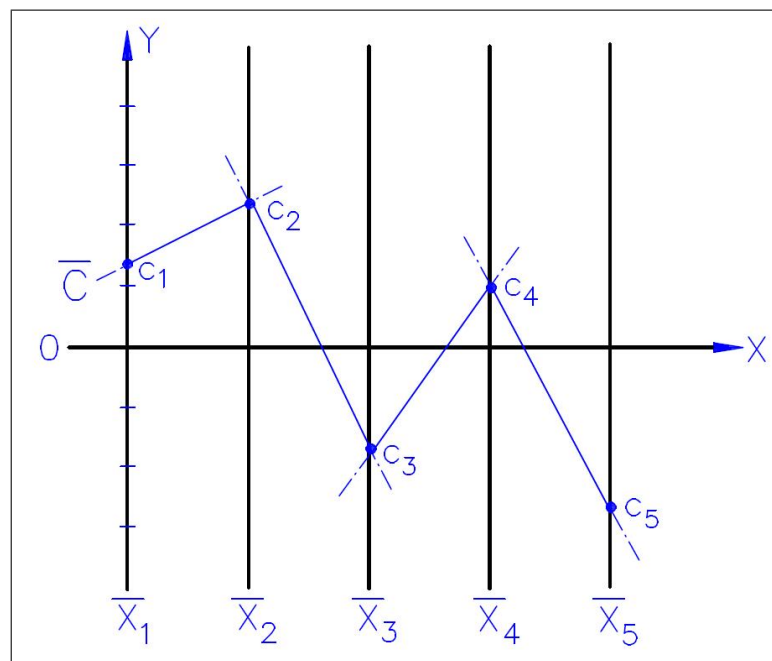


Figure 2.1: A polygonal line with 5 dimensions in a parallel coordinate system. (Source: http://www.math.tau.ac.il/~aiisreal/index_files/images.htm)

Many polylines can be drawn with several tools at once. But this also brings along some overview and significance problems. In the chapter 5 and especially in chapter 5 we present an example to pass over this problem, referring to methods in the literature.

Chapter 3

Data Sets and Data Preparation

In Chapter 2 the overall concepts for working with parallel coordinates have been laid out. In this chapter we focus on the raw data and the initial preparations necessary to work with (large) data sets. Hence examples for typical data sets are given and the (maximum) number of data entries discussed. Exploratory data analysis (EDA) – as stated in Inselberg [2009] – is the basis for 2D data displays which again is the basis for human pattern-recognition. Finding relations in complex data sets with hundreds of variables gets possible.

3.1 Data Sets

Real world objects (and subjects) are digitized and then stored and processed as bits and bytes in computers. Nowadays the main way of representing artefacts in digital storage is to persist data into relational databases in tables with fields (compare Chapter 3 "Representation" in Spence [2007]). In memory the data might be handled directly row by row, or converted into a tree-like structure in objects with attributes. For humans to understand information hiding in data it is necessary to prepare the data in some visual way. For example small sets of data can easily be displayed in tables. The raw data itself is hardly understandable if it is complex and/or large. See Figure 3.1 for an architecture necessary to process data for visualisation.

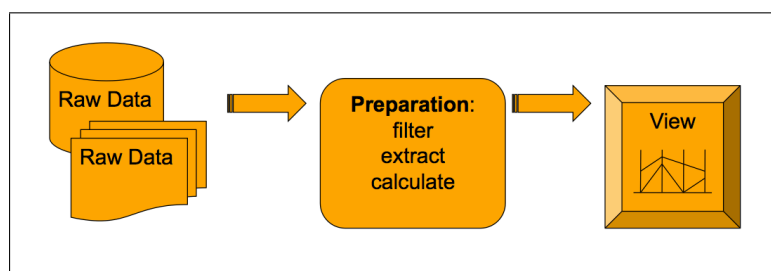


Figure 3.1: The architecture of many data acquisition and data processing frameworks.

Data sets with many attributes are called multivariate data sets. For example interpreting raw data of an EEG in medics, geological data about hurricanes, or geographical data like the size of countries. Analysing relationships between genes by showing data of many experiments with chemical reactions or flow simulation of liquids are other examples.

Especially that holds true for data with many features. Those "multivariate" data can have up to 50 different attributes such as a car could have attributes such as year, horsepower, top-speed, number of lethal accidents and so on. For a simple example see Figure 3.2 (image from http://florence.acadiau.ca/collab/hugh_public/index.php?title=GGobi_Intro).

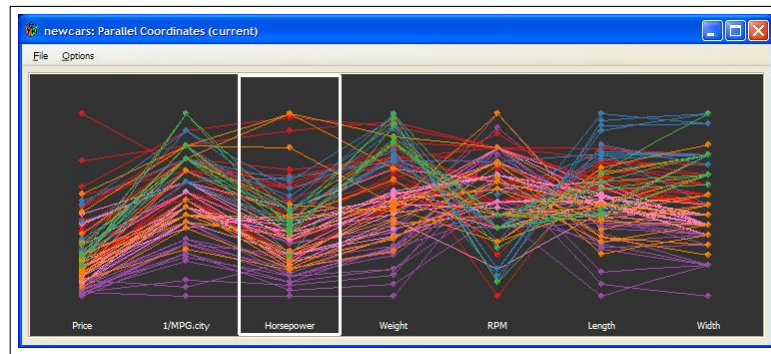


Figure 3.2: A simple car data set with attributes such as horsepower and top-speed

3.2 Large Data Sets

A special hard task is to handle large amount of data. Data sets larger than 10^3 to 10^6 data points might need sophisticated pre-preparation before visualisation. Hauser et al. [2002] describes how to prepare data of more than 10^5 data points for drawing (progressive rendering). In the paper of Novotny and Hauser [2006] data processing and visualisation with up to 3 millions data records and nearly 50 dimensions is explained. Furthermore a technique "data binning" is explained where equivalent polylines are counted (frequency based). Additionally, an overview of outlier detection, clustering, (density based representations) is available. The big advantage is that performance is not dependant on size of input data. [Inselberg, 2009, p400] experienced good results for data sets up to 800 vairables and 10.000 data entries.

3.3 Relation Between Data Points

As mentioned above the single data points consist of objects with attributes. If information in one data point refers to other data point it is even harder to make those relations visible. Not only the visualisation is tricky, also the efficient storing and retrieving of relations is not a trivial task. In the special case of paralell coordinates we assume no such (important) relation being available within the data directly. The variables are independent. Relations are not obvious at the beginning, nor stored in the data. Possible relations are the result of the data interpretation, which itself is easier when working interactively with parallel coordinate visualisations.

3.4 Filtering

To reduce the amount of data basic techniques such as filtering are common. The advantages are the reduced number of data points to visualise. Unfortunately it is sometimes hard to find out which data can be left off (filtered) without losing significant information. It is suggested to use advanced metrics including "smart" filtering and intermediate calculations.

3.5 Metrics

With mappings applied to the original data set it is possible to reduce the amount of data. Sometimes not the size of the data set is reduced, but the accessibility for computer to handle (i.e. visualise) data very fast is improved. Basic calculations could be metrics such as statistical operations max, min, median or the like. Additionally, those pre-calculated data is sometimes stored into meta-data databases for quick access. In Chapter 4 several options such as outlier detection, showing proximity or selecting sub-sets are discussed and therefore low level filter operations are necessary within the data.

To visualise data later on, we need to prepare data in such a way it can be fed into graphical libraries for drawing. Therefore Inselberg [2004] lays the foundation of geometric calculations for lines. He discusses multi-

dimensional lines, intersections, point curves. Then it is shown, how coplanarity can be detected. Hyperplanes are explained and necessary metrics such as neighborhoods or proximity are explained. Advanced preparations are discussed in Johansson et al. [2008], where methods to group data into patterns, applying threshold (received by an adaptive staircase procedure: positive, negative responses) are introduced. Find hierarchical clustering in Fua et al. [1999], where by means of a tree of nested clusters based on proximity information data is made up for interpretation. Very special metrics are described in Ellis and Dix [2006], where the idea of removing lines from "overcrowded" displays by simply showing randomly selected lines (instead of each and every line). Thereby the general trends and patterns keep visible. Technically speaking "binning" or statistical sampling is necessary and based on them different occlusion algorithms are introduced.

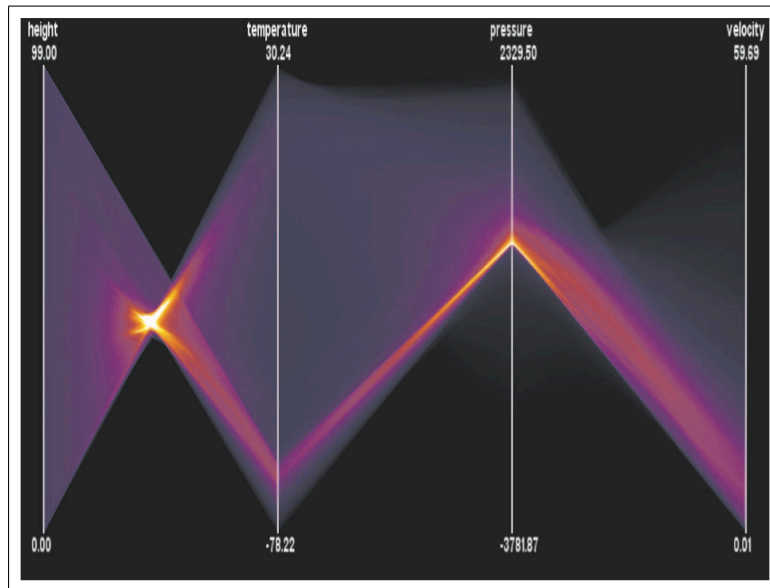


Figure 3.3: For generating density fields pre-preparation of data with algorithms of analytical computation is necessary [Heinrich and Weiskopf, 2009, pg.6].

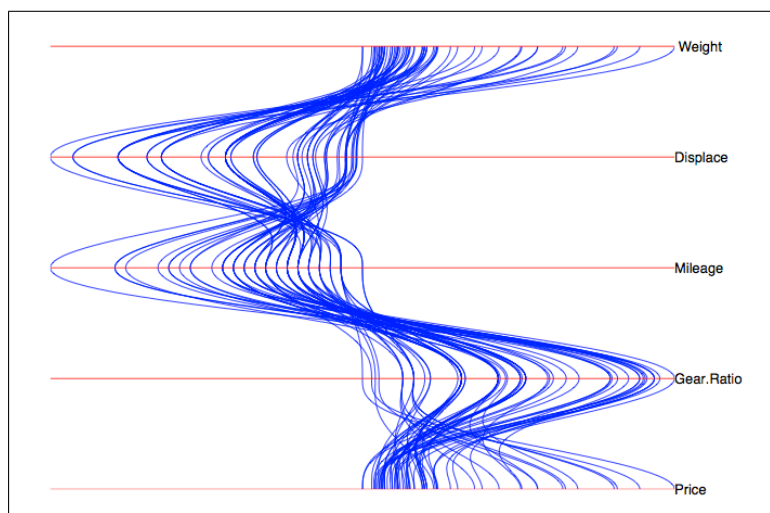


Figure 3.4: Smooth parallel coordinate plot by [Moustafa and Wegman, 2002, pg.14].

An interesting approach for improved visualisation states Heinrich and Weiskopf [2009] (see also the beautiful images such as Figure 3.3). For generating density fields within parallel coordinates - a technique coming from the "continuous scatterplots" different algorithms for numerical and analytical computation of those density fields are discussed.

To remove edge clutter Zhou et al. [2008] explains a method to employ visually clustering. This can be achieved by geometrically bundling lines and by respecting the local line density of curves. A similar problem is discussed in Lex [2008] where the crossing problem and viewing multiple relations are addressed. The most promising techniques mentioned are again cluttering and brushing. Moustafa and Wegman [2002] discusses parameter transformation for geometric structures (see smooth plots in Figure 3.4) as alternative to line interpolation. Smooth curves generated through clear axes quantization have some advantages for viewing and interpretation. Norm reducing properties and orthogonal crossing of axes are used. Furthermore the detection of second order structures is shown. Wegman [1990] explains hyper-dimensional data analysis for multivariate data representations emphasizing the tight relations to projective geometry.

For interaction such as setting thresholds for filtering we might also prepare the data accordingly. For example database "views" within the relational database might provide such low level data based filtering.

3.6 Classification

Another way to work on data is to classify input data. Tools accomplishing this tasks are called classifier. They create a characterisation of data points by distinguishing elements into a subset according given conditions and rules.

The next Chapter 4 discusses possible ways to visualise the pre-prepared data. Furthermore interaction techniques are shown, whereby one has to keep in mind, that the underlying data structure should be smart and powerful to support those (CPU consuming) operations in the best way possible.

Chapter 4

Visualisation and Interaction

Now that we know the data we want to work with, we have to decide how to work with that data. There are two steps needed. The first step is the visualization itself and the second the interaction with the visualized data to get the most information out of it.

4.1 Visualisation

Why is the question of how to visualize that important? We already know how a parallel coordinate diagram has to look like and we have the data to map them to this diagram. In Figure 4.1 we can see what could happen if we just map the data to the diagram. To overcome this problem there are several ways. Here is a brief overview about common ways to improve the visualization.

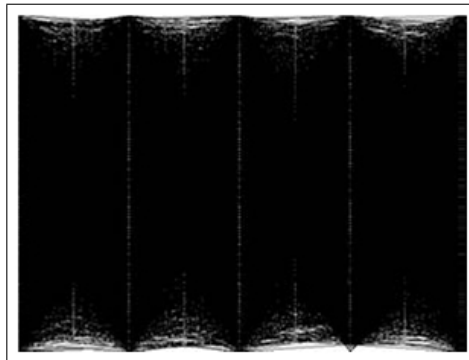


Figure 4.1: Overloaded parallel coordinate which needs better visualization Artero et al. [2004]

4.1.1 Clutter and Density Reduction

Clutter and high density are quite annoying in parallel coordinates. The analysts sitting in front of the parallel coordinates will not be able to see any information. Therefore there are techniques to reduce this disturbing factors:

- (Random) sampling and sample lens - see Figure 4.2, Dix and Ellis [2002]
- Continuous Parallel Coordinates - see Figure 4.3 and Figure 3.3
- Dynamic Masking: Fua et al. [1999] introduced this technique of dense reduction

This kind of techniques can also be called "Pruning". Yang [2005] use pruning and parallel coordinates to visualize generalized association rules. Dynamic masking concentrates on relative opacity between brushed and unbrushed areas.

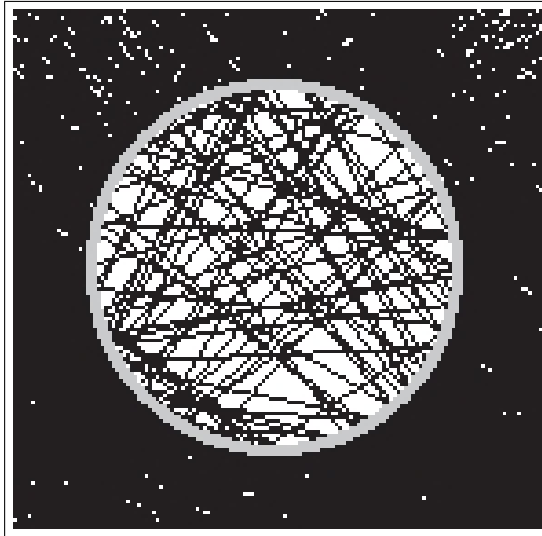


Figure 4.2: visualization of a sampling lens Ellis and Dix [2006]

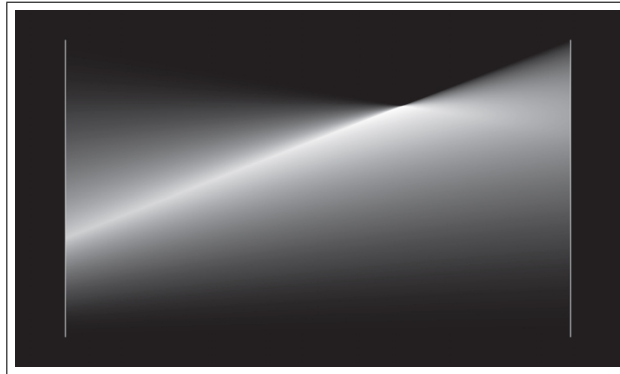


Figure 4.3: visualization of a continuous parallel coordinates Heinrich and Weiskopf [2009]

4.1.2 Clustering

Clustering is needed to recognize trends more easily in large data sets. Fua et al. [1999] defines large data sets with 10^6 to 10^9 data elements and looked for an efficient way to cluster them. Therefore they hierarchically cluster the data. In their context hierarchical means to structure and present data at different levels of abstraction. In 4.4 you can see the visual result.

Zhou et al. [2008] too use a method they called clustering. The difference is the non data-centric view. Hong et al. create clusters based on "geometrically deforming and grouping polylines while they are being plotted". In Figure 4.5 the difference can be seen.

4.1.3 Extent Scaling

In a parallel coordinate visualization there are often many overlapping bands. This makes it difficult to recognize each of the data set. To overcome this problem Fua et al. [1999] defined to scale the band thickness dynamically. For comparison see 4.6.

4.1.4 Further visualization techniques

- Dimension Zooming: mentioned in Fua et al. [1999]
- Drill-down and Roll-up: mentioned in Fua et al. [1999]

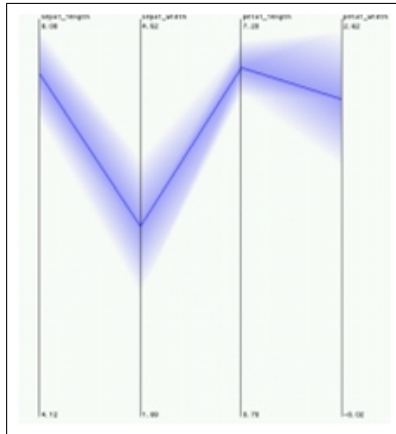


Figure 4.4: Clustering large data sets Fua et al. [1999]

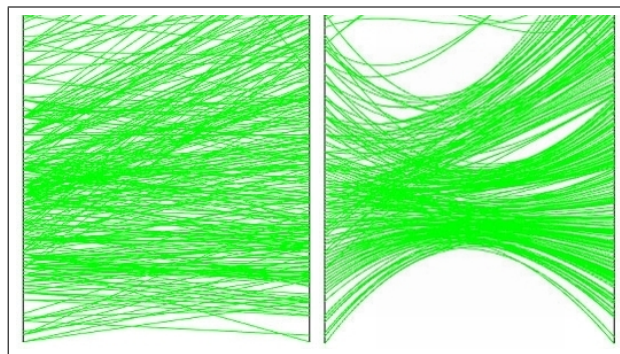


Figure 4.5: Data set without (left) and with (right) geometrical clustering Zhou et al. [2008]

4.2 Interaction

As important as a good visualization are the possibilities of interaction. Often it is hardly possible to see all the information at the first sight. To find correlations it helps to use the following interaction methods.

4.2.1 Brushing

As described in huey Fua et al. [1999] "Brushing is the process of interactively painting over a subregion of the data display using a mouse, stylus, or other input device that enables the specification of location attributes." Brushing may be direct or indirect which means that the selection is made by the mouse cursor or by additional tools (e.g. sliders). A brushed region can be seen as marked and the user can now interact with this specific subset of data Hauser et al. [2002].

There are existing lots of different kind of brushing methods. A brief overview can be found in huey Fua et al. [1999]. The first ones who described brushing were Becker and Cleveland in Becker and Cleveland [1987]. huey Fua et al. [1999] introduced their own variant of brushing called "Structure based Brushing".

4.2.2 Axis

There are many interaction possibilities related to the axes of the parallel coordinates. Following methods are taken from Hauser et al. [2002] which gives a kind overview of existing techniques.

- Axes reordering: Axes can be taken and rearranged. (e.g. via drag and drop)
- Flipping of axes: The value of axes can be inverted and top down ordering will be flipped to bottom up.

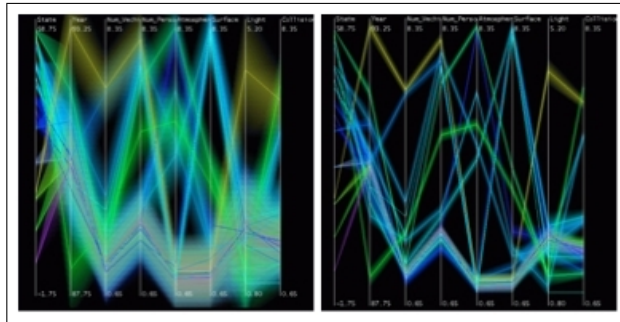


Figure 4.6: Before extent scaling, left. After extent scaling, right. Fua et al. [1999]

- Changing the mapping: The mapping of the min and max values can be changed to define an appropriate value scale. (e.g. on every axis zero should be on the same height)
- Deletion and addition of axes: information which doesn't seem useful can be taken out by deleting axes and vice versa.

Chapter 5

Interpretation

In this section steps for interpretation are suggested and a "how to" with examples is provided.

5.1 Notes

Analyse first, show the important, zoom, filter, and analyse further, details on demand by Keim [2005]

This statement from Daniel Keim, professor of University of Constance, created for a workshop with the topic: Workshop on Visual Analytics will be the mantra for this chapter.

Figure 5.1 shows the startingplot of the cereal dataset. This dataset will be used to demonstrate and explain the 5 Steps for the interpretation of a parallel coordinates dataset.

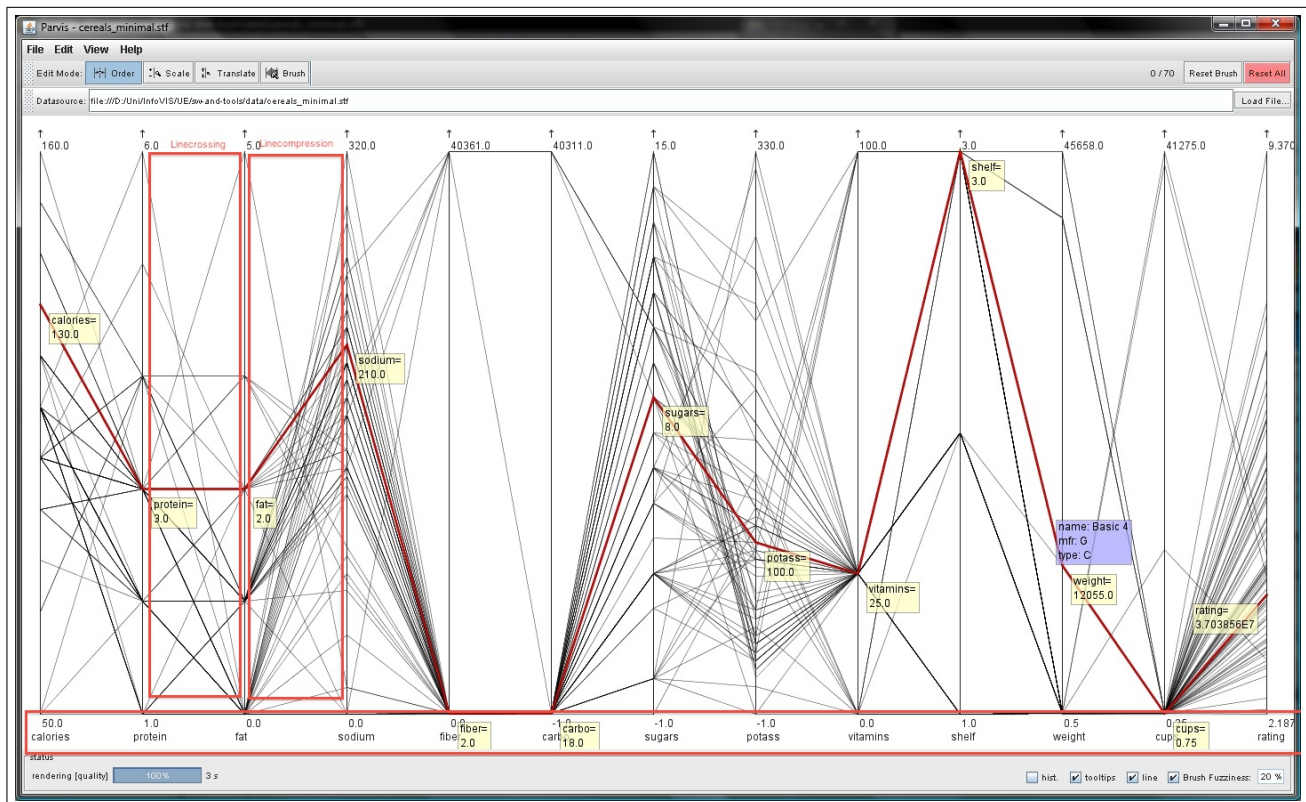


Figure 5.1: Shows the startplot of the cereal dataset

5.2 Analyse

To analyse the datasets, we have to look for a few indicators:

- Dimension of the dataset
- Line crossing
- Line compression
- Trends

What can we see in figure 5.1? At bottom are the dimensions listed. ParVis can change the order of the dimensions to get a better view for analysis.

The right marked rectangle shows a trend in this parallel coordinates. This means e.g. that the most lines (in this case one cereal) starts at one point or small section at the left dimension and ends in one point or small section at the other dimension. For e.g. cereals with low or zero fat are tend to high sodium values.

The left marked rectangle shows coherency between two dimensions. In special, this rectangle shows uncoordinated relations between two dimensions. To get some significant relations, you can change the order of the dimensions.

5.3 Show important

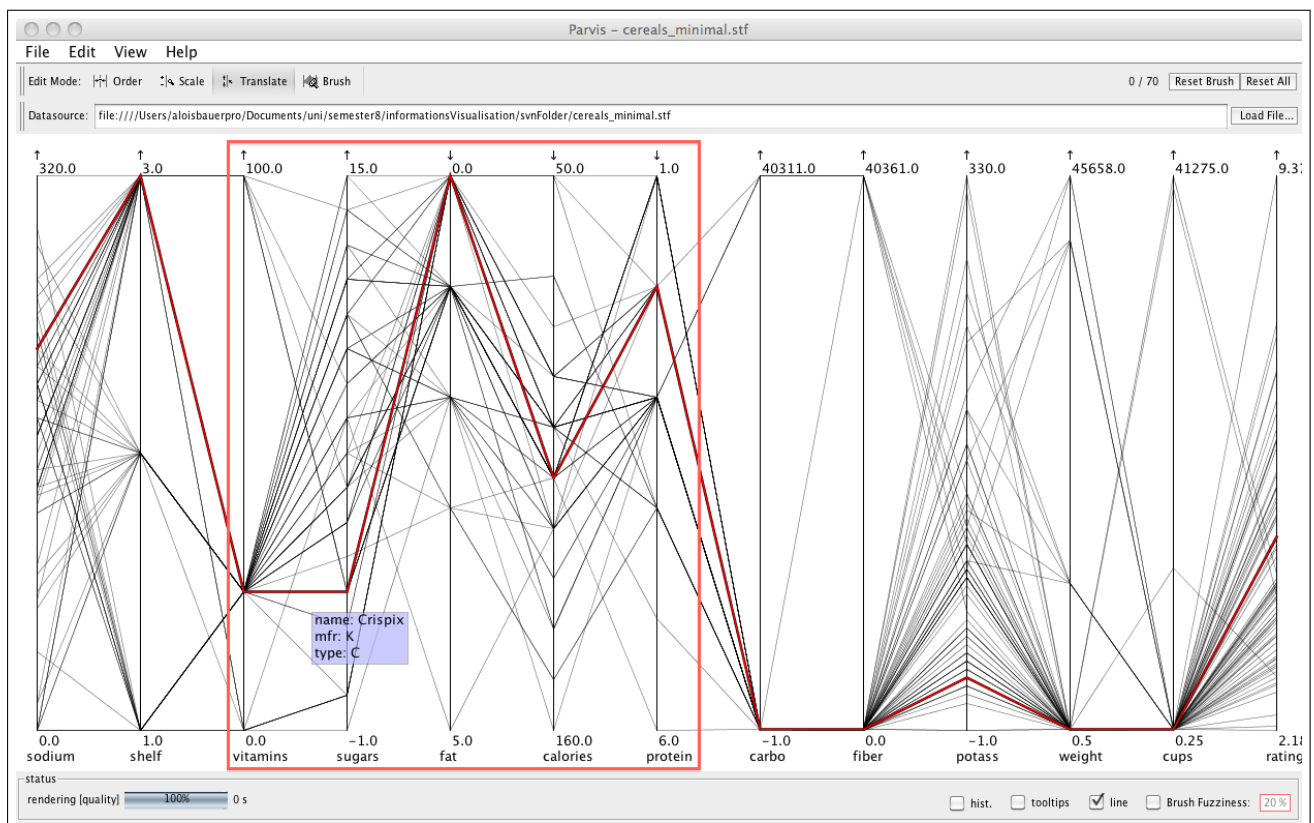


Figure 5.2: Shows a way to present important data - new order of dimensions

Figure 5.2 shows one way to present important data or dimensions in a parallel coordinates plot. If there are more or less important dimensions, you have to make a new order of the dimensions. Additionally you can collapse the grading (from maximal to minimal or minimal to maximal).

An other way to show important data is to brush some important lines. If you need data from a special dimension with a small section, this is the right way - like figure 5.2.

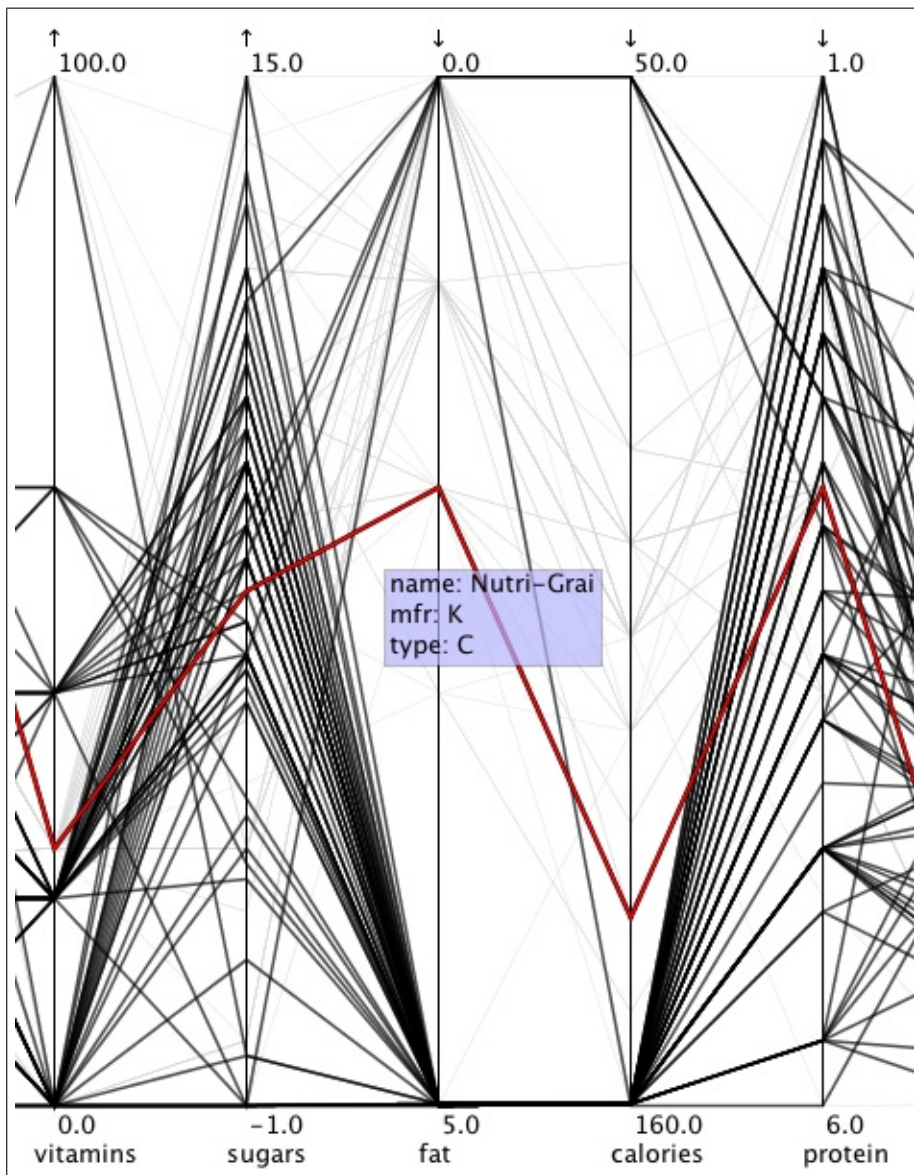


Figure 5.3: Shows a way to present important data - brush important data

5.4 Zoom

The next step for interpreting datasets is to zoom in relevant sections. ParVis supports this step. You have to change the scale in your dimensions. After this manipulation, you should get an view like figure 5.4.

5.5 Filter and analyse further

If you only want see a few dimensions, you can make a secondary click on an axis and remove (or add) all dimensions you do not need. Additionally you have to merge this view with the previous methods to zoom or filter the PCs. If you filter your datasets as per description, you should get a view like figure 5.5.

For the further analyze you have to merge all the techniques you need which are presented in this chapter.

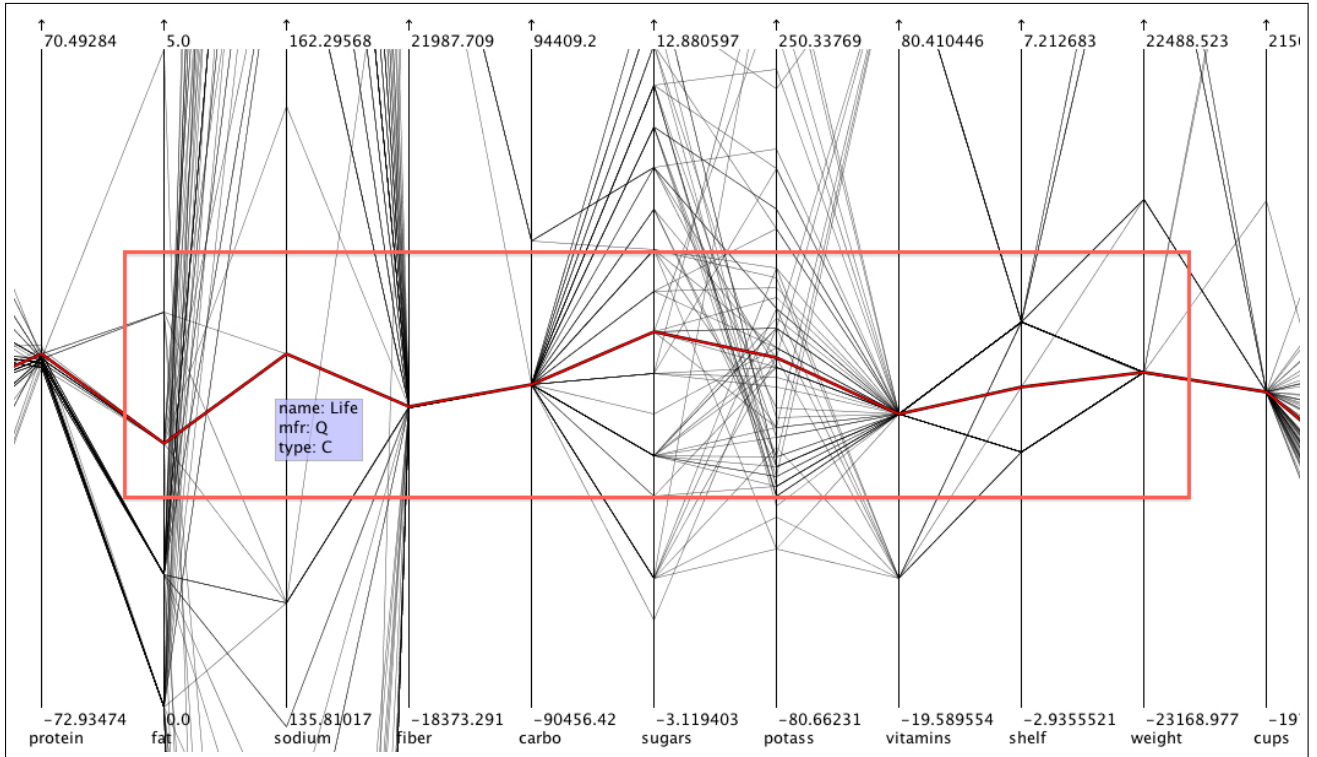


Figure 5.4: Shows how to fit in data

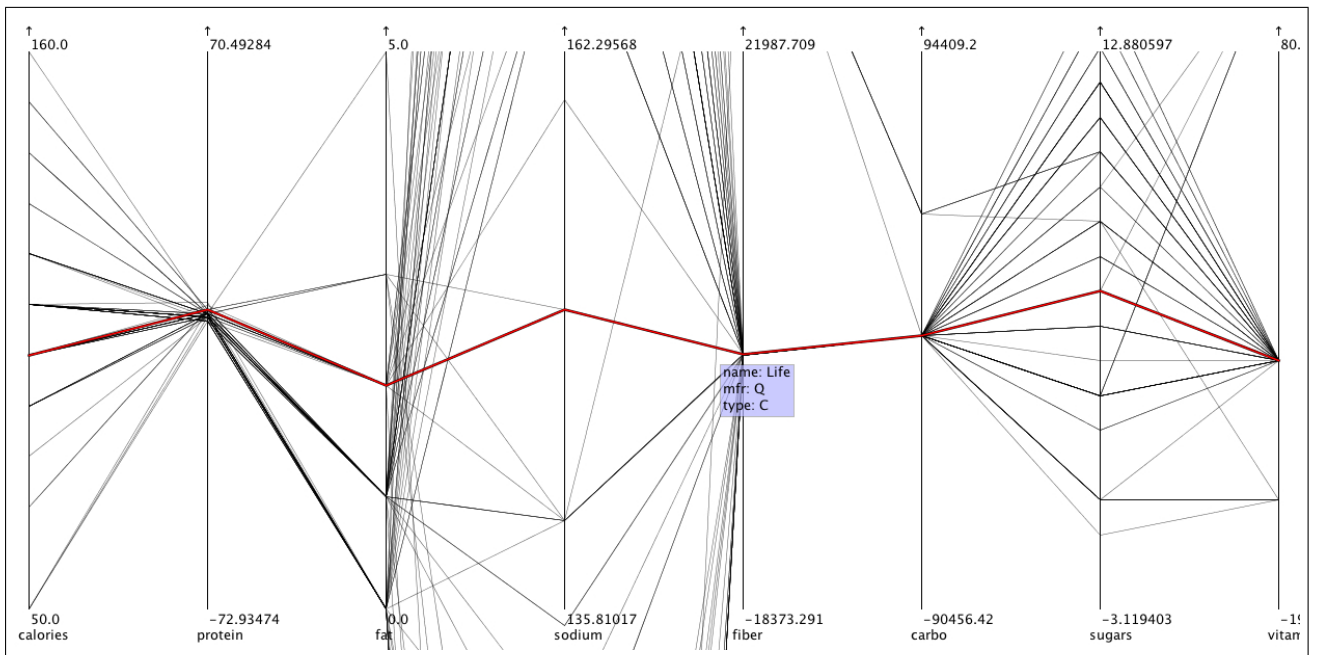


Figure 5.5: Shows a result of a filtered

5.6 Details on demand

The last section is to show details in a parallel coordinates plot. In ParVis there are some settings to view details, it is shown in figure 5.6.

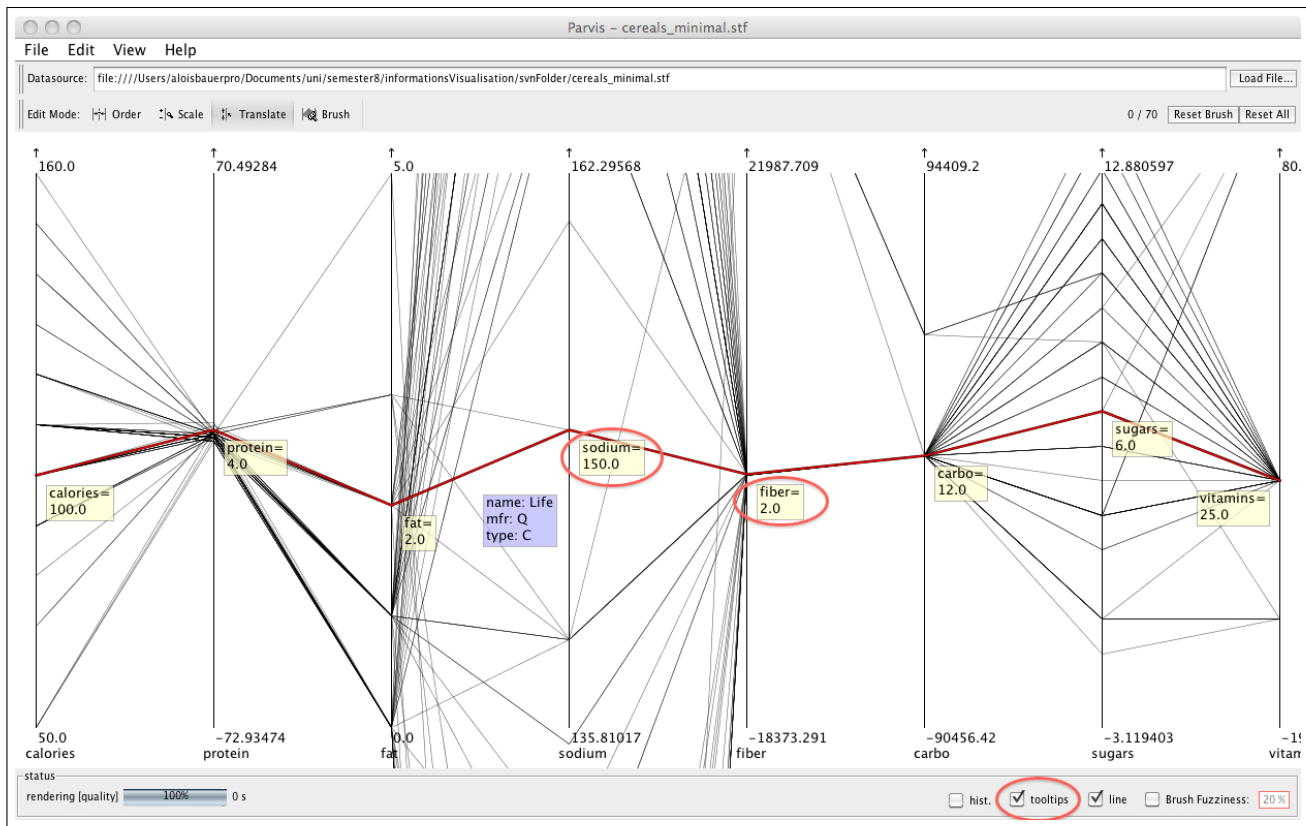


Figure 5.6: Shows a dataline with details

Chapter 6

Software and Tools

There are numerous tools for manipulating PC on the internet, but not all of them are freeware. Some of them are very basic and others are not only for displaying PCs. In this chapter we introduce only the most popular opensource tools.

6.1 Open Source Tools

Here is a list of prominent open source tools:

- Xmdv - The multivariate visualization tool
- Parvis - Parallel coordinate visualisation tool
- PicViz - Parallel coordinate plotter tool

6.2 Xmdv

Xmdv is a very basic tool originated in 1993, developed by Matt Ward. The tool is available for Unix and Windows platforms. The newest version supports HDF4 and HDF4 (a library and multi-object file format for storing and managing data between machines) as well as oracle databases. See their homepage¹ for further details. It allows visually exploration of multivariate data in a variety of methods, like hierarchical clustering - see 4.1.2, brushing - see 4.2.1, scaling, zooming and reordering. A screenshot of the program can be seen in Figure 6.1.

6.3 Parvis

Parvis is another open source tool for PCs, implemented in Java-Swing and Java-Beans. It provide the user with the necessary means of visual interaction with the data set, such as clustering, angular brushing, axis histograms, interactive rendering, scaling and translating of axes, easy drag and drop manipulation of the axis. See their homepage² for further details. A screenshot of the program can be seen in Figure 6.2.

6.4 PicViz

Picviz is an open source PCs tool with his own graph description language, implemented in C with Python bindings used for the GUI. The tool makes use of flex and bison (both are code generating tools). It is able to

¹<http://davis.wpi.edu/xmdv/>

²<http://www.mediavirus.org/parvis/index.html>

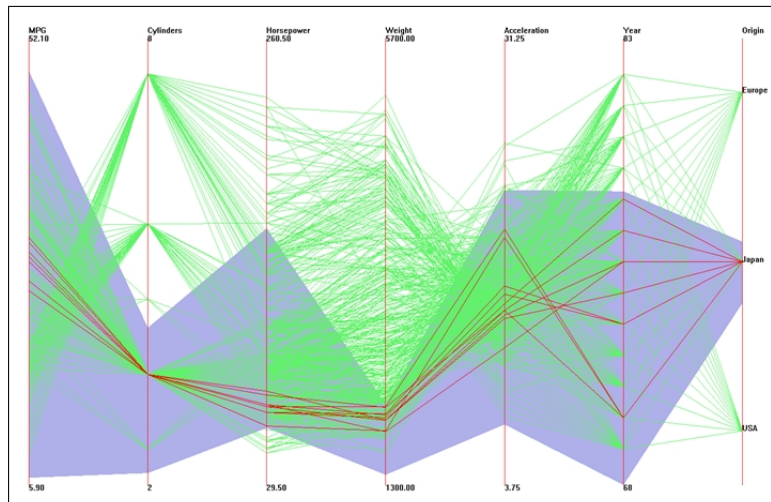


Figure 6.1: Xmdv. Polygonal lines with 7 dimensions in a parallel coordinate system representing the famous car data-set.

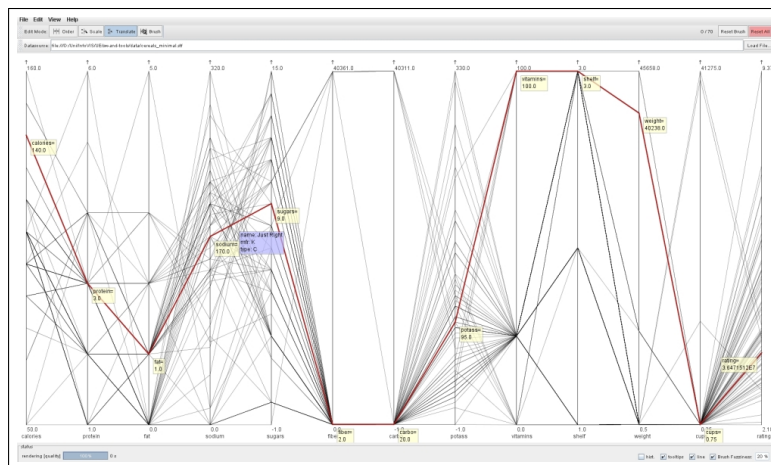


Figure 6.2: Parvis. The cereal data set with 13 dimensions. system representing the famous car data-set.

graph millions of events. Picviz is used to analyze SSH logs, Apache access logs and network traffic in a PC plotter. Picviz main goal is to find possible attacks that have been hidden before, quickly analyze problems and find correlations among variables. See their homepage³ for further details. A screenshot of the program can be seen in Figure 6.3.

³<http://www.wallfire.net/picviz/>

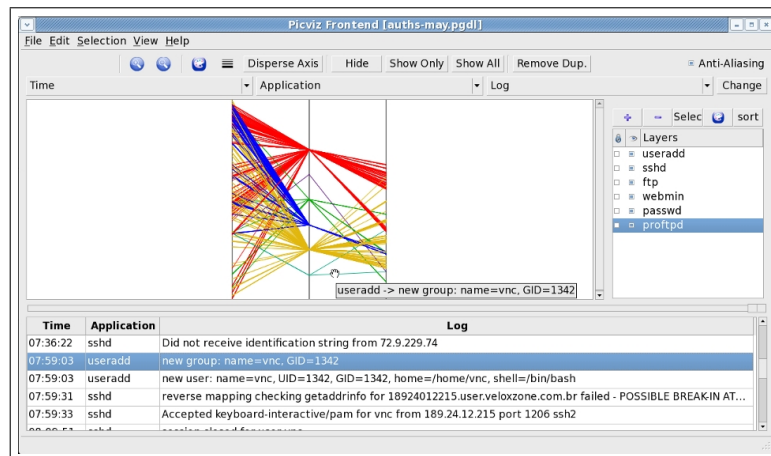


Figure 6.3: Picviz. The Picviz GUI for finding problems in different layers with a parallel coordinates plotter.

Chapter 7

Outlook and Future

This paper describes the way from raw data to interactive visualisation using parallel coordinates. The basis is laid in preparing data properly in Chapter 3, to several visualisation techniques in Chapter 4. Many ways to interact with the data are described. Then Chapter 5 discussed the problems and difficulties to interpret data. Especially, the advantage of using parallel coordinates for multivariate data is obvious as shown in section 4.2. Finally in Chapter 6 available open source tools have been compared. Every part of the visualisation process has been supported by examples of a given "cereal" data set.

The future might bring further improvements concerning the visualisation of parallel coordinates.

7.1 Possible Improvements

We expect specialised visualisations, which are adopted for mobile devices. That means the new interaction possibilities like gestures and sensors of smart phones and slate PCs with multi-touch features will become available. Fast graphic processing is available (e.g. OpenGL) and data transmission from central servers to mobile devices isn't a problem any more.

7.2 Concluding Remarks

Visualising information to comprehend the semantics of data is necessary for human beings. The way of visualizing, interacting and then analysing data to extract patterns and relationship between features of multivariate data sets will be necessary in the future even more. The amount of data is still growing exponentially and the need for data mining with drilling down into the data gets more important. For many companies e.g. in the field of e-commerce with customer behaviour analyzation working with parallel coordinate visualisation might be soon their first choice.

Chapter 8

Appendix

8.1 Well-Known Persons

Person	Description	Links
Alfred Inselberg	Many publications in the field of parallel coordinates. He claims to pursue the idea of parallel coordinates since about 1959.	http://www.math.tau.ac.il/~aiisreal/
Edward J. Wegman	More than 20 publications since 1985. in the field of parallel coordinates.	www.galaxy.gmu.edu/stats/faculty/wegman.html , www.galaxy.gmu.edu/stats/faculty/wegman.resume2.pdf , www.EdwardWegman.com/

8.2 Related Conferences

Conference	Description	Links
VisWeek2010: VIS, InfoVis, VAST	(IEEE) October 24 - 29, 2010 — Salt Lake City, Utah, USA.	http://vis.computer.org/VisWeek2010
IVAPP2010	International Conference on Information Visualization Theory and Applications, 17-21 May, 2010, Angers, France	http://ivapp.visigrapp.org/

8.3 Visualisation Techniques – Advantages and Disadvantages

Visualisation	Advantage	Disadvantage
Parallel Coordinates	many data points possible, multivariate data sets, high dimensional data	graphic power needed, occlusion, clutters, difficulties to comprehend in 3D
Scatterplots	very suitable for two dimensions, easy to comprehend	not suitable for high dimensional data

Bibliography

- Artero, Almir Olivette, Maria Cristina Ferreira de Oliveira, and Haim Levkowitz [2004]. *Uncovering Clusters in Crowded Parallel Coordinates Visualizations*. *Information Visualization, IEEE Symposium*, pages 81–88. doi:10.1109/INFVIS.2004.68.
- Becker, Richard A. and William S. Cleveland [1987]. *Brushing Scatterplots*. *Technometrics*, 29(2), pages 127–142. doi:10.2307/1269768.
- Chad A. Steed, T.J. Jankun-Kelly, Patrick J. Fitzpatrick and J. Edward [2008]. *North Atlantic Hurricane Trend Analysis using Parallel Coordinates and Statistical Techniques*. In *Workshop on Geo-Spatial Visual Analytics 2008, Fifth International Conference on Geographic Information Science*, pages 23–26.
- Dix, Alan and Geoff Ellis [2002]. *By Chance Enhancing Interaction with Large Data Sets through Statistical Sampling*. In *AVI '02: Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 167–176. ACM. 1-58113-537-8. doi:10.1145/1556262.1556289.
- Ellis, Geoffrey and Alan Dix [2006]. *Enabling Automatic Clutter Reduction in Parallel Coordinate Plots*. *IEEE Transactions on Visualization and Computer Graphics*, 12(5), pages 717–724. doi:10.1109/TVCG.2006.138.
- Fua, Ying-Huey, Matthew O. Ward, and Elka A. Rundensteiner [1999]. *Hierarchical Parallel Coordinates for Exploration of Large Datasets*. In *Proc. Conference on Visualization (Vis'99)*, pages 43–50. IEEE Computer Society Press. 0-7803-5897.
- Hauser, Helwig, Florian Ledermann, and Helmut Doleisch [2002]. *Angular Brushing of Extended Parallel Coordinates*. In *Proc. {IEEE Symposium on Information Visualization (InfoVis 2002)*, page 127. IEEE Computer Society. 076951751X. http://www.mediavirus.org/parvis/parvis_full.pdf.
- Heinrich, Julian and Daniel Weiskopf [2009]. *Continuous Parallel Coordinates*. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), pages 1531–1538. doi:10.1109/TVCG.2009.131.
- huey Fua, Ying, Matthew O. Ward, and Elke A. Rundensteiner [1999]. *Navigating Hierarchies with Structure-Based Brushes*. In *Proc. of Information Visualization 99*, pages 58–64. doi:10.1.1.38.9258.
- Inselberg, Alfred [1985]. *The Plane With Parallel Coordinates*. *The Visual Computer*, 1(2), pages 69 – 91. doi:10.1007/BF01898350.
- Inselberg, Alfred [2004]. *Lecture Notes PARALLEL COORDINATES (1999-2004)*. <http://www.stanford.edu/group/mmds/slides/inselberg-mmds.pdf>.
- Inselberg, Alfred [2009]. *Parallel Coordinates*. Springer. 978-0-387-21507-5. doi:10.1007/978-0-387-68628-8.
- Johansson, Jimmy, Camilla Forsell, Mats Lind, and Mattew Cooper [2008]. *Perceiving Patterns in Parallel Coordinates: Determining Thresholds for Identification of Relationships*. *Information Visualization*, 7(2), pages 152–162. doi:10.1145/1466620.1466624. <http://www.palgrave-journals.com/ivs/journal/v7/n2/full/9500166a.html>.

- Keim, Daniel [2005]. *Scaling Visual Analytics to Very Large Data Sets - Presentation*. <http://infovis.uni-konstanz.de/events/VisAnalyticsWs05/pdf/03DanielKeim.pdf>.
- Lex, Alexander [2008]. *Exploration of Gene Expression Data in a Visually Linked Environment*. Master's Thesis, TU Graz. http://www.icg.tu-graz.ac.at/Members/alex/lex_masters_thesis/at_download/file.
- Moustafa, Rida E. A. and Edward J. Wegman [2002]. *On Some Generalizations of Parallel Coordinate Plots*. Technical Report, Center for Computational Statistics, George Mason University, USA. <http://www.galaxy.gmu.edu/stats/syllabi/inft979/GeneralizedParallelCoordinates.pdf>.
- Novotny, Matej and Helwig Hauser [2006]. *Outlier-Preserving Focus+Context Visualization in Parallel Coordinates*. *IEEE Transactions on Visualization and Computer Graphics*, 12(5), pages 893–900. doi:10.1109/TVCG.2006.170.
- Spence, Robert [2007]. *Information Visualization – Design for Interaction*. Pearson/Prentice Hall. 9780132065504.
- ten Caat, Michael, Natasha M. Maurits, and Jos B. T. M. Roerdink [2007]. *Design and Evaluation of Tiled Parallel Coordinate Visualization of Multichannel EEG Data*. *IEEE Transactions on Visualization and Computer Graphics*, 13(1), pages 70–79. doi:10.1109/TVCG.2007.9.
- Wegman, Edward J. [1990]. *Hyperdimensional Data Analysis Using Parallel Coordinates*. *Journal of the American Statistical Association*, 85(411), pages 664–675. <http://www.jstor.org/stable/pdfplus/2290001.pdf>.
- Wegman, Edward J. and Alan Dorfman [2003]. *Visualizing Cereal World*. *Comput. Stat. Data Anal.*, 43(4), pages 633–649. doi:10.1016/S0167-9473(02)00297-9.
- Yang, Li [2005]. *Pruning and Visualizing Generalized Association Rules in Parallel Coordinates*. *IEEE Transactions on Knowledge and Data Engineering*, 17(1), pages 60–70. doi:10.1109/TKDE.2005.14.
- Zhou, Hong, Xiaoru Yuan, Huamin Qu, Weiwei Cui, and Baoquan Chen [2008]. *Visual Clustering in Parallel Coordinates*. In *Proc. Eurographics / IEEE-VGTC Symposium on Visualization (EuroVis2008)*, volume 27, page 8. The Eurographics Association and Blackwell Publishing Ltd. http://www.cs.ust.hk/~zhouhong/articles/Hong_EuroVis08.pdf.