

Survey of multidimensional visualization techniques

Abdelaziz Maalej, Nancy Rodriguez, Olivier Strauss

► **To cite this version:**

Abdelaziz Maalej, Nancy Rodriguez, Olivier Strauss. Survey of multidimensional visualization techniques. CGVCVIP'12: Computer Graphics, Visualization, Computer Vision and Image Processing Conference, Jul 2012, Lisbon, Portugal. pp.N/A. lirmm-00808108

HAL Id: lirmm-00808108

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00808108>

Submitted on 4 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SURVEY OF MULTIDIMENSIONAL VISUALIZATION TECHNIQUES

Maalej Abdelaziz

Rodriguez Nancy

Strauss Olivier

{abdelaziz.maalej, nancy.rodriguez, olivier.strauss}@lirmm.fr

ABSTRACT

In this survey we have chosen to present seven multivariate data visualization techniques. We describe each technique by specifying the means of interactions it can support and the most important developments which were brought to them. Our motivation for this work is to collect the most important techniques of multi-dimensional data visualization. Also this work is considered as a first step of our astrophysical data visualization project.

KEYWORDS

multi-dimensional data visualization, interaction, brushing

1. INTRODUCTION

A collaboration is maintained between LIRMM and LUPM, two French laboratories strongly involved in the project NUMEV. As part of a joint thesis supervision, this is the first development of a visualization tool for complex data from astrophysical models showing changes in the internal structure of stars with masses comparable to the mass of our Sun. This tool is intended to be used also by other users in other scientific domains like biology. Multivariate data visualization simply refers to the display of multidimensional data. A multidimensional data set consists of a collection of N-tuples, where each entry of an N-tuple is a nominal or ordinal value corresponding to an independent or dependent variable. In this paper we present five basic visualization techniques: Parallel Coordinates, Star Glyphs, Scatterplot Matrix, Table Lens and Dimensional Stacking and two hybrid techniques: Star Glyph integrated with Parallel Coordinates and Multiple Coordinated Views. We will not be interested in techniques like Curves, Ternary diagrams and Scatterplot due to their low number of dimensions.

Multidimensional visualization techniques are classified according to (Keim, D. A. and Kriegel, H. P., 1996) (Chan, W. W-Y., 2006) as geometric projection techniques (Scatterplot Matrix, Parallel coordinates, Table Lens), Pixel-oriented techniques (Space Filling curve), hierarchical display techniques (Dimensional stacking) or iconography techniques (Star Glyph).

2. INTERACTIONS: DEFINITIONS

The exploration of unknown data sets while making relations, detecting similarities, repetitions, differences and patterns can be reached through different interaction models. The interaction enables a real exploitation and analysis of the visualized data because humans are skillful to extract information from an environment on which they do interaction. According to (Gibson, J. J., 1979) user must act to perceive and must perceive to act:

Brushing is a process in which a user can highlight, select, or delete a subset of elements with a mouse or other suitable input device. In situations where multiple views of the data are being shown simultaneously, brushing in one view affects the same data in all other views (Ward, M. O., 1994).

Filtering is the method by which some data from the dataset is excluded from view by displaying only those records that meet specific criteria. Filtering permits the user to present varying views of the data stored in a dataset without actually affecting that data. Using filters narrows the scope of data that is presented to the user (Gajic, Z.).

Reconfiguring: Is the way how to rearrange pixels or to rearrange the dimensions called dimension reordering.

3. BASIC TECHNIQUES

3.1 Parallel coordinates

Parallel coordinates is a technique invented by Alfred Inselberg in the 70's which has been applied to a diverse set of multidimensional problems (Inselberg, A. and Dimsdale, B., 1990). In this method, each dimension is represented by an axis, and the N axes are organized as equidistant vertical lines labeled x_1, x_2, \dots, x_N , all having the same positive orientation as the y axis. An item C with coordinates (c_1, c_2, \dots, c_N) is represented by the polygonal line whose N vertices are at (c_i) on the x_i axis for $i=1, \dots, N$. In effect a 1-1 correspondence between points is established (see Figure 1). It is a system for doing and visualizing analytic and synthetic multidimensional visualization.

3.1.2 Brushing

The brushing interaction of this technique is a kind of a draggable and resizable slider on each dimension, these sliders cover the selected items which are coloured while the rest of the data is grayed.

3.1.3 Dimension reordering

Manual dimension ordering is available in some systems. For example (Stolte, C. and Hanrahan, P., 2000) allows users to manually select and order the dimensions to be mapped to the display. Similarly, in XmdvTool (Ward, M. O., 1994) users can manually change the order of dimensions from a reconfigurable list of dimensions.

3.1.4 Contributions

Too many developments have been brought to the parallel coordinates technique since its creation. We will cite two of them within clutter reduction context, the use of curves and the Hierarchical Dimension Clustering. In standard parallel coordinates, finding which entering line maps to which exiting line is very often a matter of chance. In (Graham, M. and Kennedy, J., 2003) the authors use curves solely to help differentiate lines that cross at axes, an occurrence that increases dramatically when using axes with a few discrete values. It was found that with many curves it became difficult to differentiate them if they were bunched close together along their paths. Methods of separating such curves were developed, enclosing boxes were proposed representing the single and same value.

3.2 Star Glyph

According to (Oellien, F., 2003) star glyphs are probably the most common technique to denote data points. A star glyph consists of a center point with equally angled rays. These branches correspond to the different dimensions and the length of the limbs marks the value of this particular dimension for the studied data point. A polygon line connects the outer ends of the spokes (see Figure 2). Every item is represented by a star Glyph.

3.2.2 Brushing

The XmdvTool 7.0 (Matthew, O.) includes linked brushing between all visualizations. It enables for the user a brush toolbox to resize the value of the dimensions. And according to the user's brushing it updates the whole star glyph visualization.

3.2.3 Dimension reordering (automatic and manual)

In order to counter the visual clutter, (Wei, P. et al, 2004) define a well structured star glyph if its rays are arranged so that they have similar length to their neighbors and are well balanced along some axis. And to reduce the clutter for the whole display, the authors propose to seek automatically for reordering the dimensions to minimize the total occurrence of unstructured rays in star glyphs considering these 3 criteria:

- Neighboring rays have similar lengths
- The lengths of rays are ordered in a monotonically increasing or decreasing manner on both sides of an axis
- Rays of similar lengths are positioned symmetrically along either a horizontal or vertical axis

In the XmdvTool 7.0 a tool for manual dimension reordering was proposed.

3.3 Scatterplot Matrix

Scatterplots are basic building blocks in statistical graphics and data visualization (Cleveland, W. S. et al., 1988). When researchers are interested in the relationships between pairs of several continuous variables, they often produce a series of scatterplots for each of the pairs. It can be convenient to view these graphs

together on a single screen or page using what is usually called a scatterplot Matrix (see Figure 3) (Brian S. E. and David C. H., 2005). With k variables, there are $k(k - 1)/2$ pairs, and therefore for even small numbers of variables the number of scatterplots can be large.

3.3.2 Brushing

(Becker, R. A. and Cleveland, W. S., 1987) define four brushing operations: highlight, shadow highlight, delete, and label. These operations are carried out by resizing a rectangle, called the brush, over one of the scatterplots. The effect of an operation appears simultaneously on all the scatterplots. The data analyst moves the brush to different positions on the scatterplot matrix by moving the mouse. At any time, the data analyst can stop the brushing and change one or more of the features the shape of the brush, the operation, or the paint mode and then resume the brushing.

3.3.3 Dimensions reordering

(Elmqvist, N. et al, 2008) adopt the systematic dimension reordering approach of (Ankerst, M. et al, 1998) where similar dimensions in a multidimensional dataset are placed next to each other. The use of a scatterplot matrix allows having independent row and column orderings. In their method, it makes sense to utilize this fact and display similarity on the column and dissimilarity on the row order.

3.3.4 Navigation

Instead of letting the user choose mappings for the axes of a single scatterplot, (Elmqvist, N. et al, 2008) create one scatterplot per every combination of dimensions and arrange them in a large scatterplot matrix (Cleveland, W. S. et al., 1988). The whole matrix serves as an overview of the dataset and also defines a visual space for navigation, turning the visual exploration process into a navigation task. Transitions from one scatterplot to another is performed using a 3D rotation that is consistent with the overall navigation metaphor and that provides more natural cues than standard interpolated animation (see Figure 4).

3.3.5 Contributions: Live Scatterplot

Scatterplots matrixes often expose interesting patterns of cloning between subsystems and point to opportunities for further analysis. However, the remaining question when such patterns are seen is always, "but what is that?". (Cordy, J.R., 2011) proposes the Live Scatterplots approach to aim at providing an immediate, intuitive answer that can help the analyst to quickly identify and access subsystems and clones involved in a pattern simply by directly pointing at it in the scatterplot. Live scatterplots approach exploits the table, title and hyperlink tags of standard HTML to provide this ability in any standard browser, without the need for custom frameworks.

3.4 Table lens

(Ramana, R. and Stuart C. K., 1994) present a new visualization, called the Table Lens, for visualizing and making sense of large tables.

The visualization uses a focus+context (fisheye) technique that works effectively on tabular information because it allows display of crucial label information and multiple distal focal areas. The Table Lens fuses symbolic and graphical representations into a single coherent view that can be fluidly adjusted by the user.

This fusion and interactivity enables an extremely rich and natural style of direct manipulation exploratory data analysis. The Table Lens uses a number of different types of graphical representations, called presentation types, to display the contents of the table cells. In particular, presentation types utilize graphical vocabulary including text, color, shading, length, and position to represent underlying cell values (see Figure 5).

3.4.1 Interaction: Distortion function

The table lens represents rows of a table as rows of pixels. The user can focus and stretch out rows or columns to see the data (see Figure 5), whilst maintaining surrounding context, it is based on the Degree Of Interest (DOI) notion (George, F. W., 1986) which shows if the user is interested in that item or not.

3.5 Pixel oriented

The basic idea of pixel-oriented technique is to map each data value to a colored pixel and present the data values belonging to one variable in separate windows. It was proposed by (Keim, D.A. 1996). Since in general this technique uses only one pixel per data value, the technique allows visualizing the largest amount of data possible on current displays (up to 1,000,000 data values). If each data value is represented by one pixel, the main question is how to arrange the pixels on the screen. (Keim, D.A. 1996) uses different arrangements for different purposes of the pixel-oriented technique. One straightforward possibility is to arrange the data items from left to right in a line-by-line fashion (see Figure 6). Another possibility is to arrange the data items top-down in a column-by-column fashion. More useful are techniques which provide a better clustering of closely related data items and allow the user to influence the arrangement of the data.

3.5.2 Pixel rearrangement

To allow the user to influence the arrangement of pixels that it becomes semantically meaningful, (Keim, D.,1996) proposes the Recursive Pattern Technique. Using the recursive pattern visualization technique the user is able to generate a wide range of different pixel-oriented visualizations. By structuring the visualization in such different ways.

3.5.3 Filtering

In (Keim, D.A. and Kriegel, H.P., 1998) the authors introduce the VisDB system which allows an exploration of large databases using visualization techniques. The VisDB system enables the user to modify his database query by using sliders for modifying the selection predicates and weighting factors.

3.5.4 Brushing

A rich set of navigation and selection tools has been developed in (Anilkumar, P., 2004) for the pixel oriented display in XmdvTool. Zooming and panning, brushing, distortion and pixel reordering help users to learn information about the dataset. Manual selection tools (Anilkumar P., 2004) allow users to perform human-driven dimension reduction by selecting subsets of dimensions for further exploration using the display as well as other multi-dimensional visualization techniques.

3.6 Dimensional Stacking

(LeBlanc, J. et al., 1990) created a tool that enables the user to project data of arbitrary dimensions onto 2-D image. Of equal importance to this, is the ability to control the viewing parameters, so that one can interactively adjust what ranges of values each dimension are displayed. This will allow an intuitive feel for the data to be developed as the database is explored.

Assume N dimensions $\langle D_1, D_2, \dots, D_n \rangle$, then consider a traditional 2-D display for the pair of dimensions $[D_1, D_2]$: each element of it represents a discrete value. If the cardinality of either dimension is less than the screen resolution, each element will map to a rectangular region of pixels. We can then use this area as a virtual screen for dimensions $[D_3, D_4]$, thus projecting 4-D data on a 2-D display. We can continue recursively stacking dimensions until the accumulated size exceeds the screen resolution (see Figure 7).

3.6.2 Dimension reordering

The clutter reduction technique proposed in (Wei, P. et al, 2004) using dimension reordering is applied on this technique in order to reduce the clutter (see Figure 7).

4. HYBRID TECHNIQUES

4.1 Star Glyph Integration with Parallel Coordinates

A disadvantage of a traditional 2D Star Glyph representation is that the loss of context can lead to difficulties when conducting comparisons across glyphs. In (Fanea, E. et al, 2005) the authors developed a visual representation called Parallel Glyphs to address these issues by interactions with the glyphs on the pivot axis and to address the problem of overlapping polylines in the parallel coordinates technique.

4.1.2 Rotation

In (Fanea, E. et al, 2005), users are able to rotate the visualization in all three directions around the pivot axis. A particularly interesting situation is when the glyphs are unfolded and the entire model is rotated such that the glyphs directly overlap each other which allows the user comparing the different values taken on the dimensions.

4.1.3 Item Rearrangement

The user has the capability to select one or several polylines (items) and then to rotate them around the pivot axis to a different position in the sequence. This allows to adapt the sequence of polylines in the data set to directly compare polylines that were previously far apart. Thus, the user can avoid one of the main issues of a Star Glyphs visualization, the difficult comparison of non-adjacent spikes.

4.2 Multiple Coordinated Views

Coordinated views provide several usually juxtaposed or tiled views of visualizations that are designed to be of use in relationship to each other (Collins, C. et al., 2010) (see Figure 8). The important factor for this visualization comparison discussion is that these coordinated views can be algorithmically linked such that actions and highlights in one view can be spread on other views. Coordinated views allow for reuse of the spatial visual variable, thus each relationship type is afforded spatial rights.

Alexander Lex and al. developed a system called Caleydo (Lex, A., et al., 2010) that provides powerful tools for gene expression analysis, as well as the Bucket, a tool for pathway analysis where multiple views are presented in a 2.5D arrangement.

4.2.2 Views reordering

To bring different views into focus, (Lex A., et al., 2010) enables the user to move them up and down. A very restrictive set of navigation operations turns out to be sufficient, providing the benefit of low cognitive load during navigation. Another system called VisLink (Collins, C. et al., 2010) addresses the issue of navigating in a 3D multi-view arrangement by providing hotkeys for predefined camera positions, while still allowing full 3D navigation. Views can simply be moved by drag and drop.

4.2.3 Linking between views

When working with data sources shown in separate applications, the information is contained in multiple application windows, spatially arranged within the display space, and potentially surrounded by other application windows not contributing to the analysis task. Without coordination of applications and visual linking, the user first has to gain an overview of available information in all application windows, followed by a closer investigation of potentially relevant items in these windows, to finally relate these items in order to identify trends, patterns, and correlations. Based on these considerations visual links (Waldner, M. et al., 2010) were created for the desktop, connecting related, but spatially distributed, data items of various sources and different representations (see Figure 8). With visual links we can explicitly and consistently show related items across independent applications.

5. DISCUSSION

We should attract the attention that a very important interaction is ignored in almost of all the works which is annotation. Annotation enables the users to note their remarks. The problem here is the way to save the annotations. Also we should focus on the integration of the time dimension which is almost represented by a slider tool (Yuichi, K. Y. et al., 1997) to go forward and backward in the time. Some contributions were brought to time-series visualization techniques i.e. Horizon Graph (Reijner, H.,2008) which enables the user to see a large number of time series on a single screen and to compare trends and spot patterns that would be very difficult or impossible to see in a standard report.

6. CONCLUSION

In this survey, we chose to present eight techniques of multidimensional data visualization. For each technique we described some interactions and some contributions which were developed to enable the user extracting much more information from the visualization. Each of these techniques has a certain limitation to avoid the clutter or to manage the huge amount of data to be visualized. In our astrophysical data visualization project we intend to implement a multiple coordinated views as a web application which enables simultaneous visualization of different visualization techniques in an interactive way such as the spread of the selection in the concerned views.

REFERENCES

- Andrews, K., 2002. *Information Visualization Tutorial Notes*. IICM Graz University of Technology.
- Anilkumar, P., 2004. *Pixel Oriented Visualization in XmdvTool*. Faculty of the Worcester Polytechnic Institute.
- Ankerst, M. et al, 1998. Similarity clustering of dimensions for an enhanced visualization of multidimensional data. *In Proceedings of the IEEE Symposium on Information Visualization*. pp. 52-62.
- Becker, R. A. and Cleveland, W. S., 1987. Brushing scatterplots. *Dynamic graphics for data analysis*, edited by W.S Bostock, M., 2010. <http://mbostock.github.com/protovis/ex/cars.html>.
- Brian, S. E. and David, C. H., 2005. *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Ltd, Chichester.
- Cleveland, W. S. et al., 1988. *Dynamic Graphics for Statistics*. Wadsworth and Brooks/Cole, Pacific Grove, CA.
- Chan, W. W.-Y., 2006. *A survey on multivariate data visualization in Science and technology*. Department of Computer Science and Engineering Hong Kong University of Science and Technology Clear Water Bay, Kowloon, Hong Kong.
- Collins, C. et al., 2010. Revealing relationships amongst visualizations. *IEEE Transactions on Visualization and Computer Graphics Proceedings of the IEEE Conference on Information Visualization*.
- Cordy, J.R., 2011. Live Scatterplots. *Proc. IWSC 2011, ICSE 5th International Workshop on Software Clones*. Waikiki, Hawaii. pp. 79-80.
- Elmqvist, N. et al, 2008. Rolling the Dice : Multidimensional Visual Exploration using Scatterplot Matrix Navigation. *IEEE Transactions on Visualization and Computer Graphics*. pp. 1539-1148.
- Fanea, E. et al, 2005. An Interactive 3D Integration of Parallel Coordinates and Star Glyphs. *In Proc. IEEE INFOVIS*.
- Gajic, Z., Chapter 8. *Delphi Database Course for beginners*.
- George, F. W., 1986. Generalized fisheye views. *In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. pp. 16-23.
- Gibson, J. J., 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin Boston.
- Graham, M. and Kennedy, J., 2003. Using curves to enhance parallel coordinate visualizations. *7th International Conference on Information Visualization, IV . IEEE International Conferences on Information Visualization. IEEE Computer Society Press*. London, UK. pp. 10-16. ISBN 0-7695-1988-1.
- Inselberg, A. and Dimsdale, B., 1990. Parallel coordinates :A tool for visualizing multi-dimensional geometry. *IEEE VISUALIZATION Conference*.
- Keim, D. A. and Kriegel, H. P., 1996. Visualization Techniques for Mining Large Databases: A Comparison. *IEEE Transactions on Knowledge and Data Engineering, vol.8, no.6*. pp. 923-938.

- Keim, D.A. and Kriegel, H.P., 1998. *VisDB: Database exploration using multi-dimensional visualization*. In *Computer Graphics & Applications*. pp 40-49.
- Keim, D., 1996. A Pixel-oriented Visualization Techniques for Exploring Very Large Databases. *Journal of Computational and Graphical Statistics*.
- LeBlanc, J. et al., 1990. Exploring N-dimensional databases. *Proceedings of Visualization*. pp. 230-237.
- Lex, A., et al., 2010. Caleydo: Design and Evaluation of a Visual Analysis Framework for Gene Expression Data in its Biological . *Context Proceedings of 2010 IEEE Pacific Visualization Symposium*. Taipei, Taiwan . pp. 57- 64.
- Matthew, O., 2006. Ward <http://davis.wpi.edu/xmdv/downloadxmdv.html>.
- North, C. and Shneiderman, B., 2000. Snap-together visualization: A user interface for coordinating visualizations via relational schemata. In *Proc. Of Advanced Visual Interfaces*. pp. 128-135.
- Oellien, F., 2003. *Algorithmen und Applikationen zur interaktiven Visualisierung und Analyse chemiespezifischer Datensätze*. Fakultäten der Universität Erlangen-Nürnberg.
- PetitpasP., 2011. <http://multimedia.fnac.com/multimedia/editorial/labo/reflex-3-2011.pdf>.
- Ramana, R. and Stuart Card, K., 1994. The table lens: merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. *CHI*. pp. 318-322 .
- Reijner H.,2008. The development of the horizon graph. In: *Proc. Vis. 2008 Workshop From Theory to Practice: Design*.
- Stolte C. and Hanrahan P., 2000. A system for query, analysis, and visualization of multidimensional relational databases. *Proc. IEEE Symposium on Information Visualization*. pp. 5-14.
- Waldner, M. et al., 2010. Visual Links across Applications - *Best student paper award Proceedings of the Conference on Graphics Interface*. Ontario, Canada.
- Ward, M. O., 1994. XmdvTool: Integrating Multiple Methods for Visualizing Multivariate data. *IEEE Conf. on Visualization*. pp. 326-333.
- Wei, P. et al, 2004. Clutter Reduction in Multi- Dimensional Data Visualization Using Dimension Reordering, *IEEE Symposium on Information Visualization InfoVis*. pp. 89-96.
- Yuichi, K. Y. et al., 1997. TimeSlider: An Interface to Specify Time Point. In *Proceedings of ACM Symposium on User Interface Software and Technology*. pp. 43-44.

Figure 1. Two brushes example (Bostock, M. 2010)

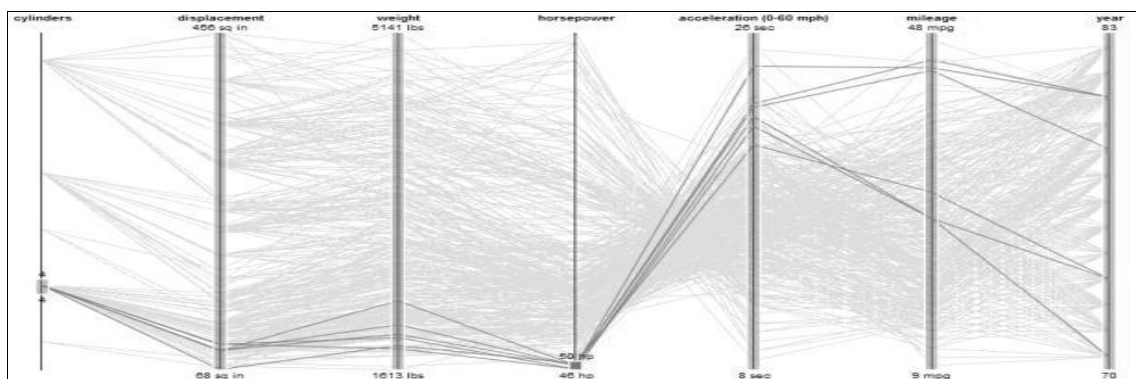


Figure 2. Star glyph for digital camera (Petitpas, P.)

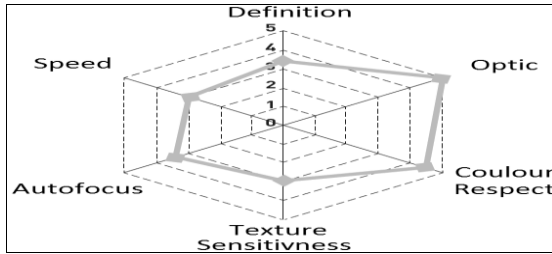


Figure 3. XmdvTool scatterplot matrix of cars dataset

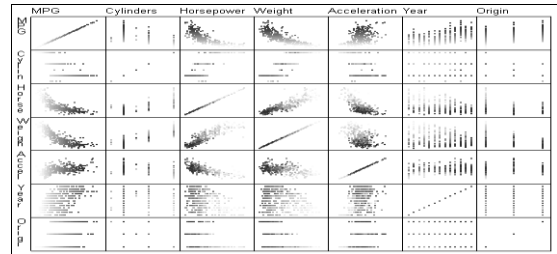


Figure 4. The overview matrix - 3D scatterplot

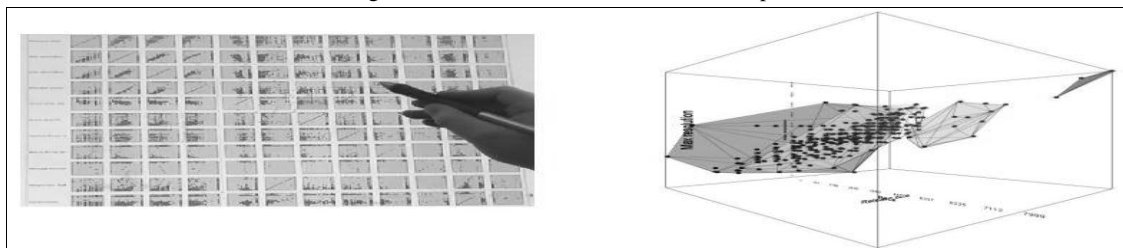


Figure 5. Table Lens player data set (Andrews, K.)

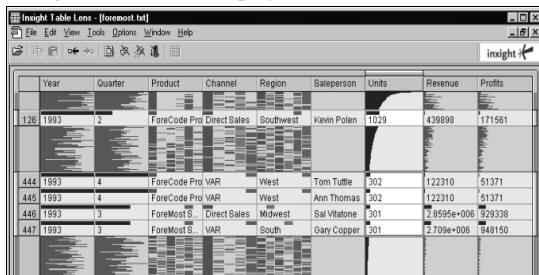


Figure 6. Brushing is accomplished by clicks on pixels

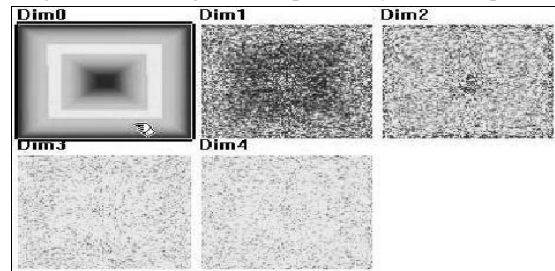


Figure 7. Dimensional Stacking cars dataset

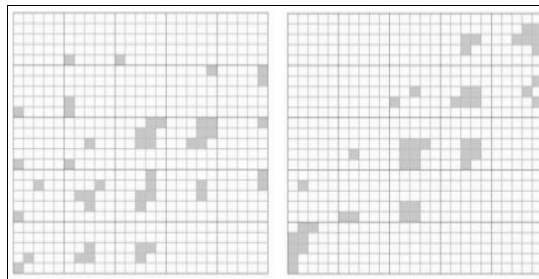


Figure 8. Visual links in Caleydo

