Project proposal: User-controlled construction of parallel coordinate plots

Lucas Rizoli
University of British Columbia
lrizoli [at] cs.ubc.ca

Keywords: Parallel coordinates, dimension ordering, interactive techniques, marking menus.

Index Terms: H.5.2 [Information interfaces and presentation]: User interfaces—Interaction styles; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

1 Introduction

Parallel coordinate plots [4] are useful for visualizing multidimensional data. They can also quickly become cluttered and difficult to read, especially when displaying a large number of dimensions. Parallel coordinate plots are useful with a reasonably small set of dimensions; the order of these axes is very important [5].

There have been a variety of proposed solutions to the number and order of axes in parallel coordinates: from dimension reduction [7] to computer-imposed axis ordering [3]. Both systems are complex, computationally expensive, and may not provide the best ordering in all cases. Ultimately, users find greater satisfaction with orders that they are able to manipulate [6].

2 Solution: User-controlled construction

To allow the user to discover the best order of axes in a parallel coordinates plot, the author proposes a system in which the user builds a plot axis by axis. This system aims to facilitate the choice of axes by showing the user what potential plots would look like, as well as providing an intuitive interface with which to build the plot. This would alleviate the burden of other parallel coordinate systems that require the user to find order by re-arranging axes in a plot, while still giving the user control over the final order.

The user begins with a single axis on the far left of the display, and adds axes from left to right. Before each addition, previews of the potential plots between the latest axis added to the plot and other axes are given. These are called plot thumbnails. Plot thumbnails can be scanned quickly, their gross patterns evaluated on sight without requiring the user to re-arrange axes directly on the plot. This is somewhat similar to the V-Miner system [8], which uses thumbnails of the trends between axes to automatically or semi-automatically determine their order. The V-Miner thumbnails are not thumbnails of the plots themselves, but rather the gross changes of the data in sequence. These describe another aspect of the data, one not commonly seen alongside a parallel coordinate plot, but do not directly aid in the ordering of the axes in plot.

The interface mechanism with which the user chooses an axis to add to the plot presents these thumbnails in a way that makes it easy to compare them and select the next axis: fan menus.

2.1 Fan menus

Fan menus are semicircular menus in which items are arranged like the sections in an oriental fan. There are many potential benefits to using fan menus for constructing parallel coordinate plots.

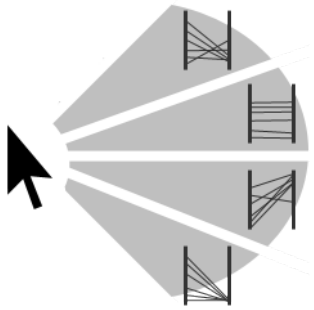

Figure 1: A mock-up of a fan menu with plot thumbnails.

- **Efficiency:** Fan menus are similar to pie or marking menus, which are known to be more efficient than traditional drop-down menus [1]. Although there appears to be little to no literature addressing fan menus explicitly, they share many of the same properties (short distance from the pointer to menu items) that have been credited for improved performance with marking menus.

- **Space for thumbnails:** The curved edge of a fan menu provides more space along which to arrange thumbnails. Since parallel coordinate plots deal with high-dimensional data, the number of thumbnails may well exceed the number of thumbnails that can be drawn along a single column (at least, at any useful size). By arranging them on a curve, taking advantage of horizontal space as well as vertical, it is possible to fit more thumbnails, to allow for more padding between them, or to render them larger than would be possible in a straight arrangement.

- **Ease of comparison:** Although somewhat dependent on their vertical order, the arrangement of the thumbnails along a single edge of the menu allows users to easily compare them. Were the thumbnails arranged along a circle, items opposite each other would be difficult to compare because they would be visually distant from one another. Although a conventional columnar arrangement may be best when comparing thumbnails, the finer distinctions they allow will likely not have any meaningful benefit over a curved arrangement, particularly when dealing with small thumbnails. Grid arrangements, where thumbnails appear in multiple columns, may crowd the display, or put some items at a disadvantage because they are in a strange local context (immediate neighbours in the grid alter their interpretation by the user) or because they are more difficult to select than others (smaller target area for the pointer, farther from the pointer's initial location).

- **Directionality:** The spread of the fan from the left to the right implies that there is a direction to the menu choice. This compliments the addition of axes from left to right in the plot, and ties the user's action with the menu to the final plot directly.

2.2 Thumbnail ordering and filtering

The order of the thumbnails in the fan menu is important to their potential for affecting the user's choice. There are many ways of sorting the thumbnails vertically in the fan menu, each with differing potential effects. Sorting thumbnails alphabetically by dimension label would allow users to quickly determine where a specific axes is in the menu display. This does not guarantee that the axes that result in the most dramatic descriptions of the data are easily found; sorting by correlation coefficient or by similarity could be more useful when exploring an unfamiliar data set. It may be best to provide control of the sorting method to the user through a combo box widget listing the methods, or providing additional information alongside the thumbnails (labels or correlation coefficients).

Another possibility is to implement hierarchical menus, in which the top most items are thumbnails of clusters of potential plots (similar to [2]) and deeper menu items are thumbnails of the plots summarized by their them. This would collapse large collections of thumbnails into smaller menus, but could also make the exploration of potential plots more difficult and increase the amount of computation required by the program (due to clustering).

2.3 Scenario of use

The proposed plot construction system would be useful in a variety of situations. Take for example, a user wanting to analyze a set of data concerning automobiles. The user would like to be able to visualize important relationships between certain vehicles, but is not certain what these relationships are. The given data is of reasonably high dimensionality (manufacturer, horsepower, fuel efficiency, time to accelerate from 0 to 60 mph, most common paint colour, cost, customer rating, etc.). In order to create a meaningful parallel coordinate plot, the user begins with the axis of manufacturers. Using the fan menu, the user can then see which of the other axes has the most interesting relationship with the manufacturer dimension. It could be that the thumbnail for manufacturer to paint colour is uninteresting, just near-straight lines without any major correlation; and that the thumbnail for manufacturer to customer rating shows some unexpected relationship, that customer ratings are far better for cars from one manufacturer than all the others. The user then selects that thumbnail and an axis for customer rating is added to the plot. The user is then presented with another fan menu of thumbnails and goes on to select the next axis (perhaps customer ratings and car colour correlate much better than ratings and fuel efficiency), and then the next (perhaps the user would like to compare fuel efficiency and manufacturer, regardless of the relationship between them, because the plot is going to be used in a presentation on car manufacturers' impact on the environment), until satisfied with the plot.

3 Project details

The proposed system can be investigated in any number of ways. The author plans to explore the interactive aspect of the system described, evaluating the usefulness and usability of the fan menus and thumbnails for building meaningful parallel coordinate plots. The aim of this investigation is to refine the interactive aspect of the proposed system.

3.1 Implementation

The plots and fan menu interface will be implemented in Processing[1], an open-source, Java-based environment. Processing allows interfaces and mid-range graphics to be coded quickly and simply (its official documentation refers to programs as "sketches"). This will allow for fast changes to the interface, as well as a prototype that can be demonstrated on almost any system equipped with a Java-enabled web-browser. The light-weight nature of Processing should not be an impediment since the rendering of parallel coordinate plots is not computationally expensive. Unfortunately, Processing does not include any parallel coordinate packages.

[1] See http://www.processing.org/

3.2 User studies

The prototype will undergo iterative design-test cycles during initial development: users will be observed and casually interviewed interacting with the program, which will affect the next stage of its development.

Formal user tests will be conducted once a the prototype has been through a few iterations of the early design cycle. In these tests, users will be given a data set and asked to perform certain tasks, from basic operations such as adding a specific axis to the plot, to more abstract and complex tasks such as building a plot that summarizes interesting aspects of the data. Experimenter's observations, user questionnaires, and time data will be collected and used to evaluate the efficacy of the interface for each task, as well as user satisfaction. These tests will involve volunteer users collected from within the department, likely other computer science students.

3.3 Milestones

- **Oct. 27:** Project proposal due
- **Nov. 3:** Gather data, refine design and testing procedure (tasks to be done, measurements to be taken)
- **Nov. 10:** First functional prototype
- **Nov. 10–20:** Iterative prototyping: repeated small, informal user tests, changes made to prototype
- **Nov. 14 or 16:** Project update presentation
- **Nov. 20–30:** Formal user testing
- **Dec. 1–13:** Preparation of final presentation and report
- **Dec. 6:** Complete analysis of testing results
- **Dec 14:** Final project presentation

3.4 Personal expertise

The difficulty of ordering axes in parallel coordinates seems like an interesting problem. The author has no previous experience in graphics programming, but has previously evaluated human-computer interfaces (task-analysis, questionnaires, interviews, quantitative studies).

References

[1] J. Callahan, D. Hopkins, M. Weiser, and B. Shneiderman. An empirical comparison of pie vs. linear menus. In *CHI '88: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 95–100, New York, NY, USA, 1988. ACM Press.

[2] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner. Hierarchical parallel coordinates for exploration of large datasets. In *VIS '99: Proceedings of the conference on Visualization '99*, pages 43–50, Los Alamitos, CA, USA, 1999. IEEE Computer Society Press.

[3] A. Inselberg and T. Avidan. The automated multidimensional detective. In *INFOVIS '99: Proceedings of the 1999 IEEE Symposium on Information Visualization*, page 112, Washington, DC, USA, 1999. IEEE Computer Society.

[4] A. Inselberg and B. Dimsdale. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *VIS '90: Proceedings of the 1st conference on Visualization '90*, pages 361–378, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.

[5] W. Peng, M. O. Ward, and E. A. Rundensteiner. Clutter reduction in multi-dimensional data visualization using dimension reordering. In *INFOVIS '04: Proceedings of the 2004 IEEE Symposium on Information Visualization*, pages 89–96, Washington, DC, USA, 2004. IEEE Computer Society.

[6] J. Yang, W. Peng, M. O. Ward, and E. A. Rundensteiner. Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. *INFOVIS '03: Proceedings of the Proceedings of the 2003 IEEE Symposium on Information Visualization*, page 14, 2003.

[7] J. Yang, M. O. Ward, E. A. Rundensteiner, and S. Huang. Visual hierarchical dimension reduction for exploration of high dimensional datasets. In *VISSYM '03: Proceedings of the symposium on Data visualisation 2003*, pages 19–28, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[8] K. Zhao, B. Liu, T. M. Tirpak, and A. Schaller. V-miner: using enhanced parallel coordinates to mine product design and test data. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 494–502, New York, NY, USA, 2004. ACM Press.