Published in Proceedings of AVI ‘2000: Advanced Visual Interfaces, Palermo, ITALY, May 23-26, 2000, pp. 173-180, ACM Press.

Multidimensional Information Visualization through Sliding Rods

Tom Lanning, Kent Wittenburg, Michael Heinrichs, Christina Fyock, Glenn Li
GTE Laboratories Incorporated
40 Sylvan Rd.
Waltham, MA 02254 USA
{tlanning, kentw, mheinrichs, cfyock, glenn.li}@gte.com

ABSTRACT

In this paper we propose new visual interface technology to address multidimensional data exploration and browsing tasks. MultiNav, a prototype from GTE Laboratories, is based upon a multidimensional information model that affords new data exploration and semantically structured browsing interactions. The primary visual metaphor is based on sliding rods, each of which is associated with an information dimension from the underlying model. Users can interactively select value ranges along the rods in order to reveal hidden relationships as well as query and restrict the set through direct manipulation. A novel focus+context view is afforded in which detail about individual items is revealed within the context of the global multidimensional attribute space. We propose a novel interaction technique to change focus, which is based on dragging rods from side to side. We relate this work on multidimensional information visualization to other research in the area, including Parallel Coordinates, Dynamic Histograms, Dynamic Queries, and focus+context tables.*

Keywords

Visual interface design, multidimensional information visualization, focus+context, shopping interfaces.

1. INTRODUCTION

There are two basic interaction paradigms for Web information finding for the vast majority of sites. In a standard browsing paradigm, users follow hyperlinks beginning with broad categories and terminating with detailed listings. False starts require users to back up and try again with page navigation links. In a query/response paradigm, users enter a query and then the system responds with an ordered list of results. Often users have to repeat this pattern several times over to refine the query. Page flipping, query refinement, and scanning through long lists of result items can be tedious and time-consuming. However, hyperlink or query/response techniques may be effective if the user knows what s/he is looking for and can recognize it when it is found. However, there are many browsing contexts when the user is not sure what s/he is looking for. Making a good choice may depend on what the total set of choices is and how individuals compare across a number of dimensions. Such tasks are common in the world of information access, particularly in the shopping domain. The current commercial approach is to rely on tables to present comparison information, but they are ineffective when there are more than a few items to compare.

While some data visualization tools address the need to understand global information dimensions and their relationships, most data visualization tools are not for the casual user nor are they targeted to tasks such as shopping on the Internet. Such tasks typically require not just understanding statistical correlations across multiple attributes, but also the need to jump back and forth from a global context to the details of individual items. The shopping task, along with information seeking dialogs in general, is characterized by the need to whittle down a large initial set of choices to just a few. Thus we recognize a need for developing techniques for multidimensional visualization that are easy to use and that integrate aggregate attribute information with browsing of individual items.

MultiNav, a prototype from GTE Laboratories, is based upon a multidimensional information model that affords new data exploration and semantically structured browsing interactions. The primary visual metaphor is based on sliding rods, each of which is associated with an information dimension from the underlying model. In what follows we discuss the information model, our interaction design based on sliding rods, and other aspects of our pro-

* GTE and the GTE logo are trademarks of GTE. All other marks are the property of their respective owners.

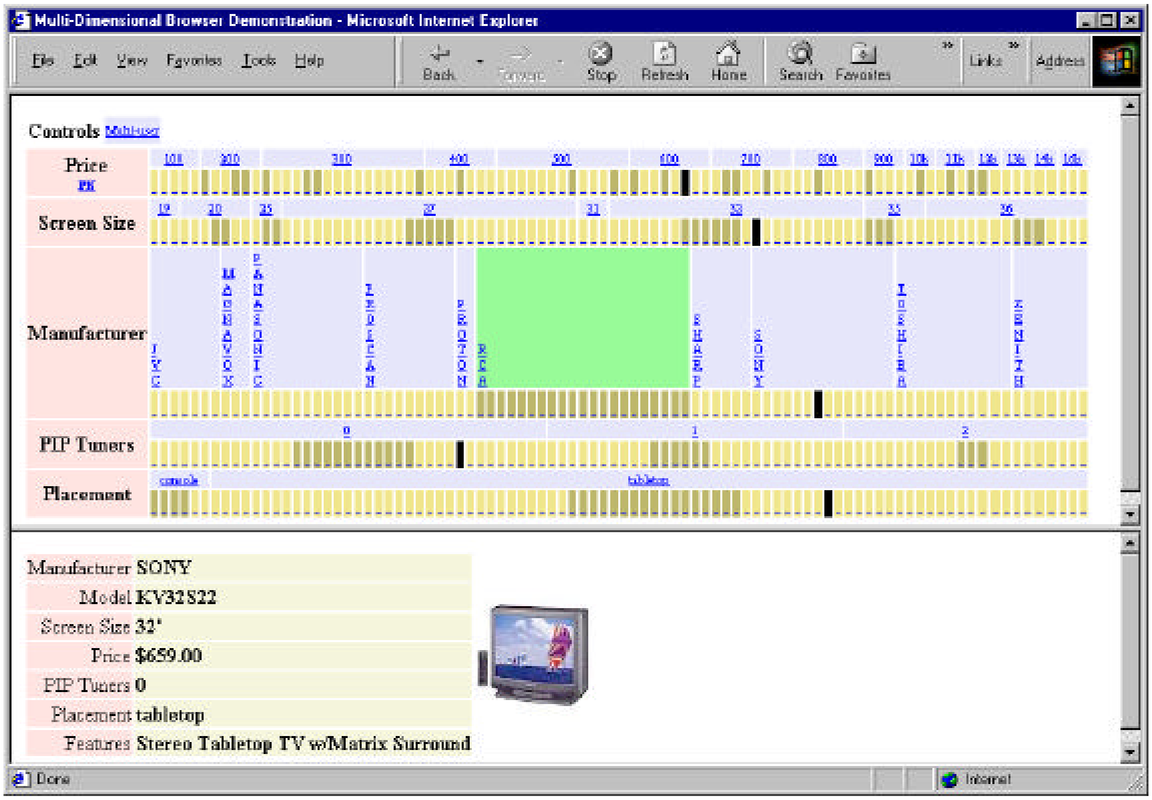

Figure 1: A screen shot of parallel attribute rows in a our first prototype (HTML/JavaScript)

totypes including the overall information-seeking dialog, data interfaces, and limits of the technique. We then relate this work on multidimensional information visualization to other research in the area, including Dynamic Querying, Dynamic Histograms, extended tables, and Parallel Coordinates.

2. MULTINAV

MultiNav is a tool that is designed to support data exploration and browsing in support of selection tasks. Instead of a two-stage query/response model that consists of forming queries and evaluating results in successive screens, MultiNav supports a tighter loop. Once a dataset is loaded to the client, users can explore aspects of the information space and navigate across individual items within the same visual and interaction context without return trips to the server and their associated delays.

2.1 Information model

MultiNav starts with an integration of two different but closely related perspectives on modeling information. The first, the traditional information theory model, represents the space (or domain) as a collection of entities, attributes of those entities, and the values associated with those attributes and entities. The second, the physics and computer graphics model, uses dimensions, coordinates along those dimensions, and the items residing (points plotted) at those coordinates. These two models are isomorphic. Entities map to items. Attributes map to dimensions. Values map to coordinates. By using both models, we can develop insights into the problem of modeling multidimensional spaces that we might have missed if we had only used a single model.

In fact, the primary insight to be gleaned from this perspective is that each item in the information space can have a location in a “physical” space. The problem, of course, is that the space is multidimensional. Most approaches to information visualization have assumed that the solution to visualizing a multidimensional space is to reduce the space to two or three dimensions and plot a location for all items within an absolute global projection. Our approach is, instead, to create visualizations of each dimension independently and synchronize them. We assume that every item has a linear position in every dimension through a sorting mechanism.

As we will see, these assumptions allow us to build user interface components to explore the space by visualizing the effect of value restrictions (queries) on the space in the aggregate as well as to navigate the space through a stepwise sequence of relative views.

2.2 Attributes as parallel rows

One prototype of a MultiNav interface component is shown in Figure 1. The content here consists of a set of product data for 100 televisions. Each of the rows in the visualization is labeled with an attribute in the data set, e.g., price on the top row. The fact that there are 100 items is evidenced by the fact that there are 100 bars positioned in each row. Each of the bars represents a different item, and they are all sorted. Thus we can see that each item has a "physical" position in each attribute row. (We acknowledge that not all attributes admit a natural sort order, but all attributes can be sorted, even if arbitrarily. The default text type uses alphanumeric sorting.)

Note next that each attribute row is labeled with values for that attribute. Price, for example, has been segmented into ascending groupings of $100, $200, etc. The length of the value cell grouping is an indication of the distribution of the items across value ranges. Thus we can see from the first row that the prices of televisions range from $100 to $1600, and that most fall in the range of $300-800.

At the bottom of the figure we see detailed information about one television set. The user clicked on one of the bars in a row, which was then highlighted (painted black) simultaneously in all rows. By observing the position of the black bar, we can get a quick sense of where this product falls in the space as a whole by seeing its position in each attribute sort. In this example, we can see that this Sony television falls in the lower-middle distribution of price, the lower end of screen size, and so on. Thus we get both a global view of the "context" of this item in the set as a whole as well as a detailed view. If the user chose to, s/he could browse individual items in this way. For instance, it might be interesting to see what that television is at the extreme high end of the distribution. What does one get by paying $1600 for a television? The user can click on the bar at the extreme right end of the first row to find the answer.

The interface in Figure 1 also allows interactive attribute-based data exploration. The large area highlighted in the middle of the figure is an indication that the user selected RCA in the manufacturer row. The effect is to make a restriction on the data set. The result is shown through changing the color of all item bars that match that restriction. The set of item bars with the value of RCA is then changed--the darker color is visible in the middle of the figure under the RCA area--but the same set is made visible in every other attribute row as well. Thus a user can see that RCA has products that range from $200 to $800 and from 20" screen size to 36" screen size. The user could continue to restrict the query further by clicking on other value cells. A click on another attribute row is treated as a logical AND. Thus a user could quickly discern what the price range of RCA televisions is in the 27" screen size by clicking on the value cell 27 in the attribute row screen size. A click in the same attribute row is considered a logical OR. Thus a user could ask the same questions for the set of Sonys and RCAs together by clicking on Sony.

2.3 Attributes as sliding rods

Figure 2 shows a second MultiNav prototype. The functionality is a superset of the earlier prototype. A major design difference is that the attribute rows, instead of remaining static during interactions, slide back and forth. The horizontal position of each of the attribute rods is determined by the item in focus. The rods are aligned such that the focused item is always in the center position.

Here is a scenario of how a user might interact with the sliding rods. First, the user moves the mouse cursor over the lower part of a rod until the cursor changes shape and a "Drag" tool tip appears. At this point, the user clicks down and drags the rod from side to side (or uses the keyboard arrow buttons). During the dragging of one rod, all the other rods move as well. Since the same item resides at some position on each of the other rods, the rods all move left or right as necessary to align that same item in the center. Users can see correlations by noting which rods tend to move in the same or different directions. For instance, one may notice in a product set that price generally goes up as some other attribute, such as screen size or resolution, goes up. In that case rods will move in the same direction and tend to be horizontally aligned. Reverse correlations will be exhibited by rods moving in different directions and "anti-alignment" in the horizontal plane. Attributes that don't correlate will be evident by rods that jump seemingly randomly while one of the rods is moved. While statistical analysis and other visualization methods are superior at revealing overall correlations, an advantage of this interactive method is that partial correlations over some restricted range of values might become evident.

Aside from sliding the rods, users can also glean other information about the data set. As in the earlier prototype, users can interactively select value ranges and see distributions in other dimensions of those items selected through value restrictions. In this implementation, users can drag

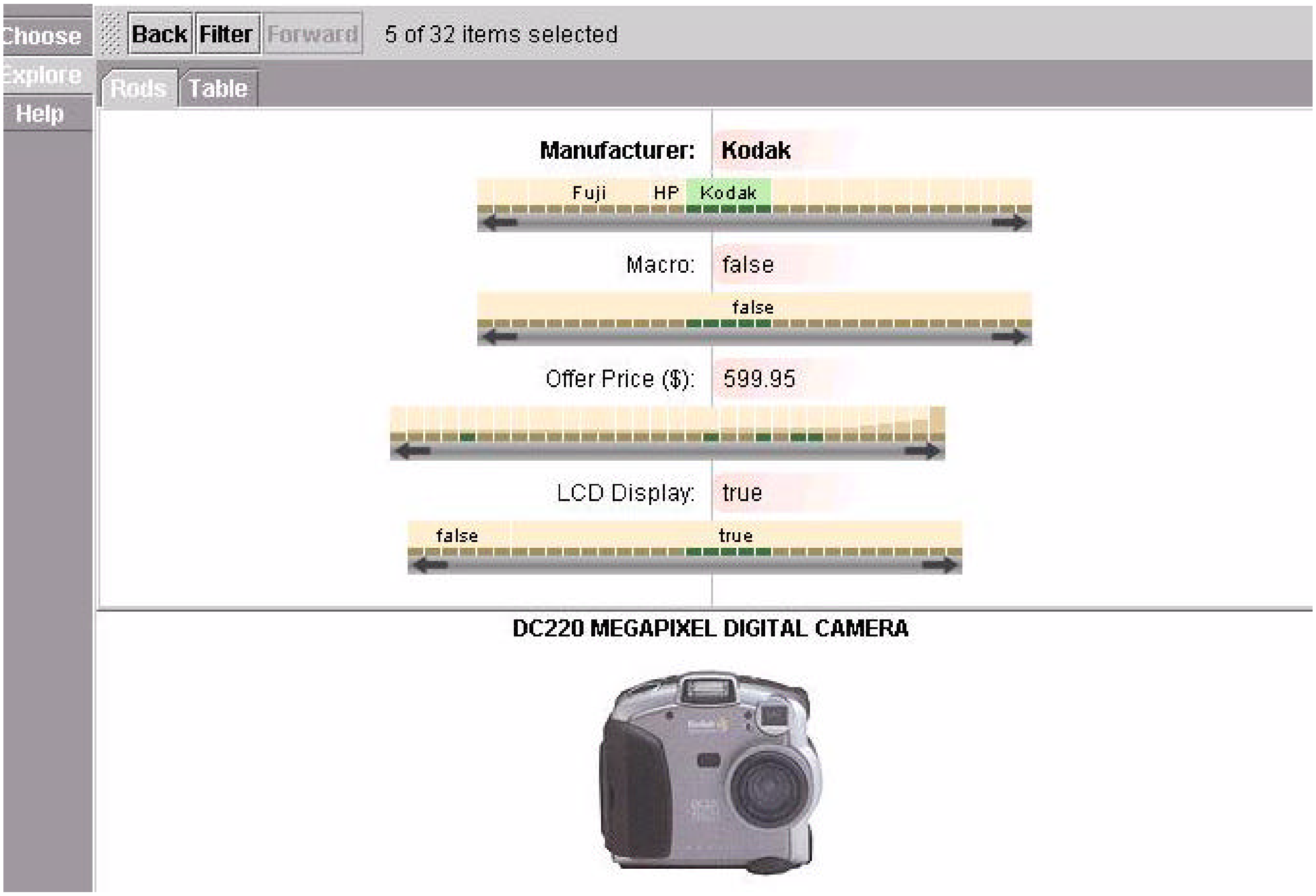

Figure 2: A screen shot of sliding rods in MultiNav (Java)

across the attribute value areas to select contiguous value ranges. Noncontiguous value ranges can be added as well. This interaction technique was inspired by ToggleMaps [2]. In Figure 2, the user has dragged across the area labeled "Kodak" to select products from just that manufacturer. The price ranges and so on of the restricted item set are colored on the other attribute rods. At any point a user may filter the set of items in view by clicking the "Filter" button. The previous set is then replaced with a MultiNav view of the restricted set. Users may navigate through a history stack of previous supersets or subsets through the "Back" and "Forward" buttons.

Users can also note outliers as well as min and max value ranges on the attribute rods. If the attribute is of numerical type, then histograms are shown for each value relative to the range in the current set. For example, in Figure 2 one can see that most of the digital cameras in this set are similarly priced in the low end of the distribution range. However, there are a few high-priced items at the right fringe. We use a similar technique for time/date value types.

Labeling values is a challenge for data sets of more than trivial size. In Figure 1, you can see one strategy in which all values are labeled, even if it takes up a lot of vertical screen real estate. In order to handle larger data sets with more attributes, we decided to opt for a labeling strategy that uses minimal space at the cost of requiring more interaction to extract the information. In the implementation shown in Figure 2, value labels are added to the attribute rods only if there is room. Otherwise, users may examine values through a tool tip mechanism. If they place the mouse over the upper half of any of the rods, then a screen appears with the value of that attribute of the item at that position.

3. INFORMATION-SEEKING DIALOG

MultiNav comes into play in a larger information-seeking dialog when a set of items of appropriate size and having a common attribute schema has been specified. Methods that might be used to select such a set include browsing over hierarchically organized categories or initiating a query. Such dialogs are evident on any of the comparison shopping sites on the Web today.

For such sites to utilize MultiNav, a server component

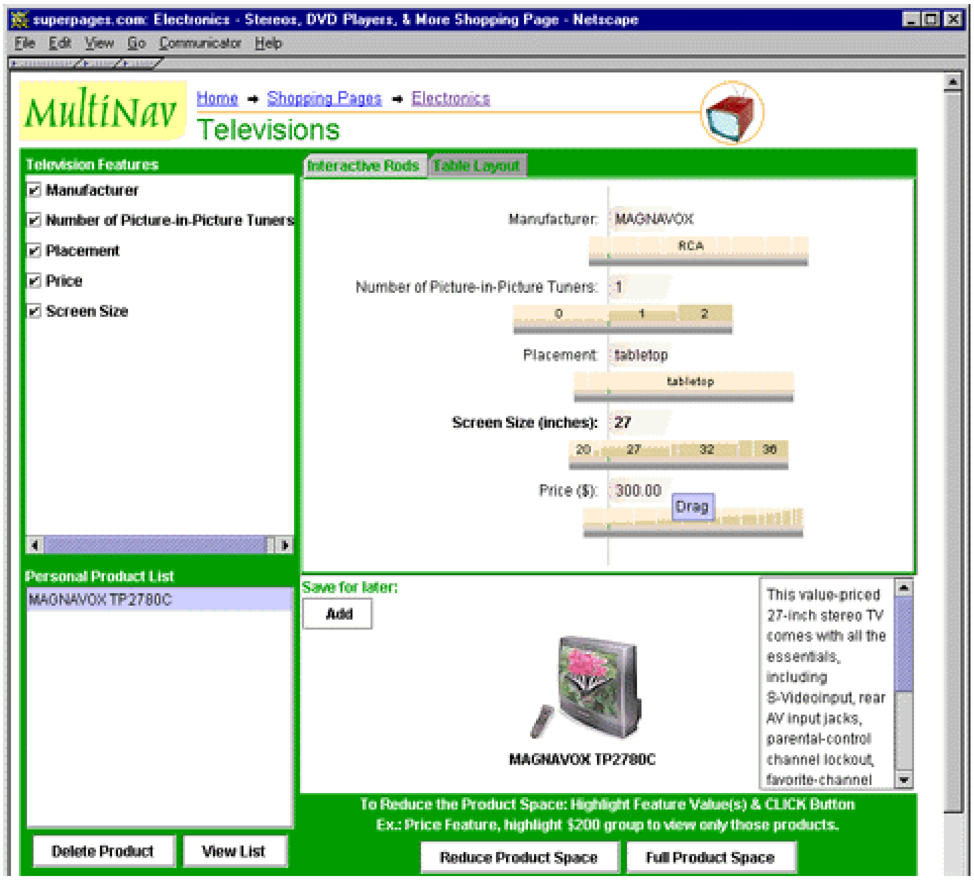

Figure 3: A MultiNav shopping applet prototype

would generate a MultiNav space given the specified set of items. An applet that reads in this dataset can then be inserted into the dialog. Figure 3 shows one example in the context of a comparison shopping prototype. A user has previously selected the category "televisions" for further investigation. This example site holds information on 100 television products, a number suitable for MultiNav.

Evident in this figure are two other functions available for users. First, a set of attributes is shown in the upper left pane. Users may interactively select those attributes that they wish to view in the MultiNav Rods component by selecting checkboxes. The data specification for MultiNav allows for any number of attributes to be made available for selection. Second, in the lower left area of the figure one can see a pane labeled "Personal Product List." Users can add any focused item from any of the views to the personal product list by clicking the "Add" button. This list itself may later be visualized as a table or within the sliding rods view. We believe that tables may be a preferable way of viewing small sets.

Although these additional panes are shown simultaneously in a "frame" layout in Figure 3, we recognize the fact that they take up precious screen real estate. We have also revised this applet to use popups and tabbed screen replacements to address this same functionality.

4. DATA INTERFACES

Multinav is designed to be an information visualization tool that can be linked with several data sources. Currently these data sources are XML documents, CSV files, and relational databases.

Our XML document specification is designed to allow any number of items along with any number of attributes. It includes meta-information regarding type and desired presentation. Presentation information is relevant to items (as exemplified in the detail view panes in Figures 1-3), but also to attributes and values. A value for an attribute such as camera resolution may be represented by a single integer for the purposes of sorting but be displayed as a text string such as "800 X 600" in the rod label. MultiNav spaces can also be exported as XML documents, which keeps the structure and presentation of the data model. (They can also read and write themselves to binary storage through a serialization mechanism.) The XML document is the preferred data format for Multinav.

In order for existing data to take advantage of Multinav, CSV import and relational database bridges are also provided. CSV is an interoperable exchange format supported by Microsoft Windows applications including Microsoft Excel. The fact that MultiNav can read CSV format means that any Excel spreadsheet can be exported and viewed with MultiNav. We have also built a relational database bridge using RDBC so that an SQL query can be sent by MultiNav to a networked database and its table of results interpreted and displayed.

We have used this I/O flexibility to explore the use of MultiNav for a number of datasets, including demographic statistics for towns, business investment information, and various types of products.

5. LIMITS OF THE TECHNIQUE

Because the sliding rods technique requires that any view state include a focused item, there are natural limits to the size of the set suitable for viewing. Our implementation makes the assumption, in fact, that each individual in the set is allotted a minimum of one pixel in each rod display. Since rods slide from side to side, the minimum horizontal screen real estate required to view the rods in their entirety is the cardinality of the item set times two. In effect, this bounds the maximum size of a set to the horizontal pixel resolution of a display area divided by two. Scrolling within the rods display area is in fact accommodated by our implementation, but visual context is significantly hampered when the ends of the rods are out of view. Given this metric, we suggest that a rough bound on the cardinality of a set for MultiNav sliding rods visualization is 500. This number also represents a reasonable maximum target considering data download times and animation performance during rod sliding.

Other factors come into play when set sizes approach this 500 item limit. The main one is that labeling becomes more difficult, and boundaries between value changes also might become invisible. When the attributes are numeric, histograms can convey information about value distribution, but this is not applicable when the values are textual (nominal or ordinal). We are still refining our presentation techniques in these cases.

6. RELATED WORK

In the space of information visualization research, the sliding rods technique we have introduced here is related to a body of work that utilizes parallel dimensions as visual structures [4]. Parallel Coordinates [7] is one such technique that also associates the items in a data set with a visual presentation of separate but parallel dimensions in which the total set of items is each ordered. With this technique, relationships across value distributions are revealed through connecting individuals across parallel coordinates with lines. The goal is for such lines to reveal patterns in the data through a multidimensional geometry. As is evident in the discussion of how to use this technique [6], it is suitable for sophisticated users to probe relationships in high-dimensional data through a kind of detective work. It is not suited for relatively naive users who faced with the comparatively simple task of selecting products.

Where the Parallel Coordinates technique utilizes static visual patterns to uncover relationships, the MultiNav sliding rod technique relies on interactions. One type of interaction, selecting a range of objects in one display and having those same objects colored on other displays, is a well-recognized technique in high-dimensional information visualization. It is sometimes referred to in the general sense as brushing [3].

Others who have utilized brushing techniques in the context of parallel dimensions include Attribute Explorer [12] and Influence Explorer [11]. These systems use interactive histograms as the display of visual structure. Sliders are deployed to select value ranges. The display is by value rather than by item. That is, a value range is plotted on an axis and the number of items that fall into that value is plotted on the second axis to yield a histogram. This is done for each attribute dimension. In MultiNav, each horizontal position on a dimension rod is associated with an individual item rather than a value. Value distributions can be plotted on the vertical dimension within the rod. This design decision is related to the fact that the sliding rods technique requires each individual to be associated with a unique horizontal position. As far as we can judge, the trade-offs of these two design alternatives are as follows. The plotting by value is good at reflecting aggregate distributions. It is also likely to scale better than sliding rods. However, the plotting by individual has the advantage of revealing a stronger visual signature for individual items as well as reflecting correlations through animation of sliding rods. Overall, we sense that sliding rods are better at browsing since a natural gesture is afforded to view nearby items in any dimension.

The selection mechanism employed by our recent MultiNav implementations allows for selection of discontinuous values a la ToggleMaps [2]. It is more akin to interactions with paint programs than typical interface widgets such as sliders or buttons. Eick earlier proposed a similar interaction technique for what he called Data Visualization Sliders [5]. Data Visualization Sliders incorporate the “painting” selection technique across a range of visual information displays including color scales, barplots, and density plots. The proposal combined controls for restricting values with the display of the values themselves. Data Visualization Sliders exhibit a richer set of techniques for data visualization than fit the sliding rods model, again, because of the constraint that each individual be associated with a unique horizontal position.

Dynamic Querying [1][9] is an influential paradigm that relies on a tight feedback loop between attribute restrictions (queries) and results. It inspired MultiNav as well as a large body of interactive visualization designs. A difference between MultiNav and Dynamic Query applications is the inclusion in Dynamic Querying of a visualization of result sets that is independent from the controls. MultiNav, along with Data Visualization Sliders and interactive histograms, integrates the two. MultiNav sliding rods is a more generic solution for information visualization problems since Dynamic Querying involves customizing the visualization (but not the controls) to the particulars of specific domains. MultiNav can be employed to visualize any spreadsheet. However, its usefulness will depend on whether attributes (columns) admit a natural sort order. Presentation information also needs to be customized in many cases for optimal results.

Another set of related work is based on extending conventional table layouts with focusing interaction techniques. One such system is TableLens [8], a general data analysis tool. Another is FOCUS, a tool for product comparison and selection [10]. Both incorporate focus+context visualization techniques within table layouts so that areas of interest can be given relatively more visual real estate. FOCUS, later renamed InfoZoom, also allows users to form boolean queries through selection and direct manipulation. Users can thus begin with a fairly large data set and gradually restrict it to items of most interest. MultiNav offers this same functionality (although without the animated transitions that InfoZoom offers). As with conventional tables, FOCUS and TableLens allow for sorts based on attributes. Users can focus on attribute value distributions in order to explore the data as well as restrict it through selection of value cells. The selection operations for querying are greatly aided by sorting since neighboring values are spatially congruent. What table-based layouts do not offer is the ability to view multiple attributes simultaneously within their natural sorts. This would be impossible since table layout conventions require that every column aligns the values of a every item across every row. MultiNav breaks that mold by aligning the attribute values of just one item at a time, allowing for remaining items in the attribute rows to maintain their natural order. Users can then easily select value ranges for any attribute in every view state, and we believe this offers a more simple and straightforward interaction model.

The focusing mechanism for MultiNav thus contrasts with focus+context tables as follows. The focus for MultiNav is an individual item; its context is a global multidimensional attribute space. The focus for focus+context tables is a range over a single attribute's value distribution; its context is the remaining attributes and values, but value ranges are not easily revealed in attributes other than the focused one since in general only one attribute at a time can be used to sort. We should add that TableLens, in particular, offers many other options for focus+context views within the scope of a single attribute dimension that may be possible to incorporate into a system that incorporates a sliding rod layout and interaction model.

Despite the fact that Dynamic Querying and many other information visualization techniques have been in circulation for years, it is notable that none have succeeded in making their way into standard Web information-seeking dialogs. Shneiderman, in 1994, mentioned challenges related to performance and scale [9]. Perhaps another challenge is that most Web service providers seem compelled to provide their interface at the lowest common denominator, namely HTML. The MultiNav prototype shown in Figure 1 is in fact implemented largely in HTML, using table markup tags. This may suggest an avenue for some information visualization techniques to win market acceptance. However, sliding rods require richer implementation languages. We can only hope that eventually the Web will settle on standard environments that will enable large scale deployments of dynamic information visualization applications. The sliding rods version of MultiNav is written in Java 2 and is feasible for deployment in that environment when/if it becomes standard. We believe that other implementations are feasible as well, including Java 1, JavaScript, and Macromedia programming environments.

7. CONCLUSION AND FURTHER WORK

The sliding-rods technique for visualizing multidimensional information offers a clear contrast to traditional information-seeking interaction paradigms. At any point a user can browse items that are "near" a focused item in any of the dimensions that matter to him or her. Detail is revealed of the focused item at the same time that its broader context in a multidimensional space is shown. Ideally, at a glance, users can extract a visual signature of the item they are focusing on that is based on its relationship to other items in the space. They can also uncover hidden relationships in the data set through interactively sliding dimension rods and restricting value ranges. With the direct manipulation querying and filtering features of this interface, users should have confidence that they haven't missed a better choice that they wouldn't otherwise have known about. They are always presented with a global context when looking at the next level of detail for individual items.

We have discussed how this work relates to Parallel Coordinates, Dynamic Histograms, Dynamic Queries, and focus+context table layouts. MultiNav sliding rods integrate browsing, which focuses on individual items, with a context consisting of global, always-sorted attribute information. The requirement that it always renders a focus item is also a limitation on its scale. Compared to static tables, however, it represents a significant scaling advantage. Compared to dynamic focus+context tables, it offers a more simple interaction model for selection.

As for further work, the fundamental hypothesis of this paper, that MultiNav sliding rods improve the user experience as well as performance on certain information-seeking tasks for casual users, has yet to be proven. Our sense from informal observation is that users need some initial help at understanding how to use the controls but that the initial learning curve is quickly mastered. We are exploring other visualization presentation interaction techniques that use the same underlying model. We also are working on extending the set of generic types for attribute dimensions.

8. ACKNOWLEDGMENTS

We thank John Vittal for his support of this work under the Advanced Interactive Internet Technologies project in the Advanced Systems Lab of GTE Laboratories.

9. REFERENCES

[1] Ahlberg, C., and Shneiderman, B. (1994) Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In Proceedings of CHI '94 (Boston MA, April 1994), ACM Press, 313-317.

[2] Baudisch, P. (1998) Don't Click, Paint! Using Toggle Maps to Manipulate Sets of Toggle Switches, In Proceedings of UIST ‘98 (Madison WI, November 1998), ACM Press, 65-66.

[3] Buja, A., D. Cook, and D. F. Swayne (1996) Interactive High-Dimensional Data Visualization. Journal of Computational and Graphical Statistics, 5(1), 78-99.

[4] Card, S. K., Mackinlay, J.D., and Shneiderman, B. (eds.) (1999) Readings in Information Visualization: Using Vision to Think, Morgan Kaufmann, 1999.

[5] Eick, S. G. (1994) Data Visualization Sliders. In Proceedings of UIST ‘94 (Marina del Rey CA, November 1994), ACM Press, 119-120.

[6] Inselberg, A. (1997) Multidimensional Detective. In Proceedings of IEEE Information Visualization ‘97 (Phoenix AZ, October 1997), IEEE Press, 100-107.

[7] Inselberg, A., and Dimsdale, B. (1990) Parallel Coordinates: A Tool For Visualizing Multi-dimensional Geometry. Visualization '90 (San Francisco CA, 1990), IEEE Press, 361-370.

[8] Rao, R., Card, S.K. (1994) The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information. In Proceedings of CHI ‘94 (Boston MA, April 1994), ACM Press, 318-322.

[9] Shneiderman, B. (1994) Dynamic Queries for Visual Information Seeking. IEEE Software, 11(6), 70-77.

[10] Spenke, M., Beilken, C., Berlage, T. (1996) FOCUS: The interactive table for product comparison and selection. In Proceedings of UIST '96 (Seattle WA, November 1996), ACM Press, 41-50.

[11] Tweedie L.A., Spence, R., Dawkes, H., and Su, H. (1996) Externalizing Abstract Mathematical Models. In Proceedings of CHI '96 (Vancouver Canada, April 1996), ACM Press, 406-412.

[12] Tweedie, L.A., Spence, R., Williams, D., and Bhogal, R. (1994) The Attribute Explorer. In Video Proceedings and Conference Companion of CHI ‘94 (Boston MA, April 1994), ACM Press, 435-436.