MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Rapid Serial Visual Presentation Techniques for Visualizing a 3rd Data Dimension

Wittenburg, K.; Lanning, T.; Forlines, C.; Esenther, A.

TR2003-20 April 2003

Abstract

In order to circumvent the challenges of conventional 3D interfaces in providing a third dimension to abstract data, we propose a set of techniques from rapid serial visual presentation(RSVP) interfaces. The basic idea is to explode a sequence of 2D presentations into a 3D trail. Interactions afford integration over the sequence in time, but also take advantage of human visual perception to incorporate peripheral information as the sequence approaches or recedes from the focal point. By performing analysis on the sequence analogous to clustering and segmentation of a video stream, patterns and points of interest can be revealed in spatial layout of the third dimensional trail. Users can also navigate quickly to points of interest.

To be presented at HCI International Conference, Crete, Greece, June 2003.

MERL – A MITSUBISHI ELECTRIC RESEARCH LABORATORY
http://www.merl.com

Rapid Serial Visual Presentation Techniques for Visualizing a 3rd Data Dimension

Kent Wittenburg, Tom Lanning, Cliff Forlines, Alan Esenther

TR-2003-20 June 2003

Abstract

In order to circumvent the challenges of conventional 3D interfaces in providing a third dimension to abstract data, we propose a set of techniques from rapid serial visual presentation (RSVP) interfaces. The basic idea is to explode a sequence of 2D presentations into a 3D trail. Interactions afford integration over the sequence in time, but also take advantage of human visual perception to incorporate peripheral information as the sequence approaches or recedes from the focal point. By performing analysis on the sequence analogous to clustering and segmentation of a video stream, patterns and points of interest can be revealed in spatial layout of the third dimensional trail. Users can also navigate quickly to points of interest.

To be presented at HCI International Conference, Crete, Greece, June 2003.

Rapid Serial Visual Presentation Techniques for Visualizing a 3rd Data Dimension

Kent Wittenburg, Tom Lanning, Cliff Forlines, Alan Esenther

Mitsubishi Electric Research Laboratories
201 Broadway
Cambridge, MA 02139
{wittenburg, lanning, forlines, esenther}@merl.com

Abstract

In order to circumvent the challenges of conventional 3D interfaces in providing a third dimension to abstract data, we propose a set of techniques from rapid serial visual presentation (RSVP) interfaces. The basic idea is to "explode" a sequence of 2D presentations into a 3D trail. Interactions afford integration over the sequence in time, but also take advantage of human visual perception to incorporate peripheral information as the sequence approaches or recedes from the focal point. By performing analysis on the sequence analogous to clustering and segmentation of a video stream, patterns and points of interest can be revealed in spatial layout of the third dimensional trail. Users can also navigate quickly to points of interest.

1 Introduction

Visual data mining is a process that emphasizes the human in the loop (Keim & Kriegel, 1996). Although the field of data mining has focused to date largely on automatic extraction of patterns and implicit information in large multidimensional datasets, the introduction of visualization methods affords the opportunity to use the extraordinary capabilities of human visual perception. Humans can process information rapidly in visual form and respond opportunistically in ways that computers cannot. Visual data mining is one of many complex tasks that require careful system design in order to know when and how to take advantage of human capabilities to direct the course of the analysis in an iterative process (Spence & Apperley, 1977) (Anderson et al., 2000).

Despite the advanced techniques for information visualization that have been proposed by the research community in recent years (Card, Mackinlay, & Shneiderman, 1999) (Spence, 2001), the most common practice in the industry for visualizing abstract data still consists of generating 2-dimensional charts and graphs. In fact, almost all of the visual data mining techniques surveyed by Keim (2002) are based on 2D presentations. Taking the next obvious step, namely, adding a third spatial dimension to plot the data, has met with limited success. There are inherent challenges in 3D information visualization (Card et al., 1999, p. 61), including graphic design complexity, legibility of text, and excessive freedom of movement for navigation controls. Information is often occluded from certain viewpoints, necessitating complex manipulations on the part of users to reveal hidden information. With so many degrees of freedom typical in manipulating the viewpoint in a three-dimensional display, it is easy for the complexity of navigation and other UI operations to overwhelm users. Nevertheless, adding a third dimension is an attractive proposition in order to show data relationships that are otherwise awkward or impossible to reveal in 2D.

In the field of scientific visualization, a commonly used technique for adding a third dimension is the use of time, i.e., the production of "data movies." However, this is often a labor-intensive task whose main purpose is the presentation of findings rather than exploratory analysis, which requires a quick turn-around from specification to visualization. Also, the standard method for presenting data as a movie allows only transitory integration of the data presented in the temporal dimension. In other words, a user's ability to gain insight into data along the time dimension will be limited by the fact that the information is continuously streaming. Memory resources would have to be stretched in order to recognize a pattern that was discontinuous or evident over anything other than a short time window.

In this paper we propose a new approach to adding a third dimension to data that is derived from work in Rapid Serial Visual Presentation (RSVP) interfaces. As with the related work by Mackay and Beaudouin-Lafon (1998), it borrows insight from the field of multimedia interfaces, adding a spatial dimension to the familiar temporal presentation methods of image sequences.

2 RSVP Interfaces

Spence (2002) provides an overview of work to date in the area of RSVP as it applies to human-computer interfaces. The most basic of the RSVP methods involves temporal sequencing of single images where each successive image displaces the previous one much as in a typical video player. Spence calls this keyhole mode, emphasizing the constricted view of the presentation in time. He also discusses four other variants of RSVP interfaces: carousel mode, collage mode, floating mode, and shelf mode. These all incorporate some form of spatial layout of the image frames that add additional movement or displacement of the image content. A related presentation method in the area of document interfaces may be found in Freeman & Gelernter (1996).

Much is known about human perception and cognition of imagery presented in keyhole mode since it has formed the basis of many psychology experiments. (See, e.g., Coltheart 1999.) There have been a few experiments also in applications of these techniques in human-computer interfaces, but it is largely unexplored territory. However, in one experiment DeBruijn and Spence (2002) make the observation that different subjects used different eye-tracking strategies. They noted that one user seemed to focus only on the screen area in which new images would appear (before they moved off to the background). Another seemed to track the images as they were moving. The work we present here further explores this territory, allowing for a design for RSVP interfaces that explicitly courts both these strategies. One of our hypotheses is that a 3D "distance" model rather than a 2D "scrolling" model is easier to process.

3 Shift, Analyze, and Collect

Our model of spatial/temporal layout incorporates traversal through a sequence of 2D image frames that takes advantage of depth perception integral to the human visual system. This 3D layout distinguishes it from most of the other RSVP interfaces. Our spatial/temporal layout is integrated with the keyhole mode of RSVP and enhanced with a number of techniques for analysis and grouping. We will refer to our method as *Shift, Analyze, and Collect*.

Visual objects that are relevant to a *Shift, Analyze, and Collect* presentation may be 2D images generated from any variety of sources (maps, video frames, charts, or graphs). Analyses of the image sequence are designed to provide an "interest" function so that certain images drop out from image trail. In all cases we combine the spatial/temporal layouts with a common set of controls for continuous forward/back and adjustable rates of speed (Wittenburg, Chiyoda, Heinrichs, & Lanning, 2000). The visual effect of such controls is to advance the linear sequence of objects towards or away from the main focus position close to the viewer.

3.1 Spatial layout

We propose an spatial layout of a sequence of visual objects in a 3D space such that the trajectory formed as a line from center point to center point in the visible object sequence comprises a straight line or curve in which the position of the center point increases in the z dimension. (In two-dimensional systems, the scale of the object can be used to approximate the depth dimension z.) The effect is analogous to a roadway that may be straight or curved along which a series of signs appear in a spaced sequence. The viewer (driver) sees only a certain subsequence of the signs at any given point in time through the display (windshield).

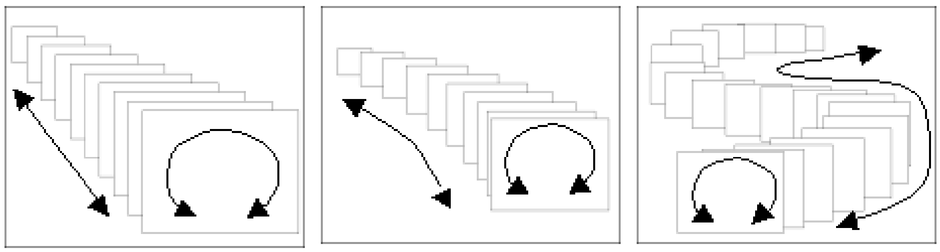

Figure 1: Three layouts

In the most basic case, all visual objects along a trajectory in a 3D space represent merely a strict ordering and are spaced evenly. However, the method provides for any spacing along a scale of choice. A spatial layout may reflect, say, temporal placement as well as any other semantically coherent relation that can be mapped onto values in a one-dimensional linear or nonlinear scale.

3.2 Traversing the sequence

The user is given controls governing the traversal of the sequence by manipulating forward/back direction and speed. As the user advances the sequence, the objects appear to move closer (or further away) along a stable trajectory path. The analogy with signs along a highway is that the user can control the speed and direction of a car. Looking out the windshield, the signs will appear to move past at regular temporal and spatial intervals, appearing initially from far away and then moving closer. Or, if the car is moving in reverse and the driver remains facing forward, the signs will appear in view from over the shoulder and then regress into the distance. Whether the perceptual effect is that the viewpoint moves or that the trail of objects moves is an interesting question discussed elsewhere (Wittenburg, Ali-Ahmad, LaLiberte, & Lanning, 1998a,b).

3.3 Collector frame

In the mode of operation we have described above, as an object gets closer, it will eventually move past the field of view and disappear. This is the standard metaphor of what would happen if a driver in a vehicle passes signs along a roadway, looking straight ahead. We propose a variant on this basic mode that has the advantage of maintaining continuity across related visual objects (say, frames within a video) and providing a focal point for the users so that they do not have to shift their eyes to process a sequence at high speed. We use a collector frame, which is a visual container that exists at a z position closest to the viewer along the main trajectory. As the visual objects move closer in the sequence they eventually approach the position of the collector frame. Instead of moving past, they replace the last object that was formerly in the view. Our method thus combines the advantages of conventional video players (with fast forward and reverse) with a layout that affords being able to look ahead (or behind) the focal point in the sequence of image frames or visual objects. The frames in the frontmost position in Figure 1 are collector frames, the circular arrow representing replacement rather than movement during advancement.

3.4 Analysis (clustering/segmentation) of the sequence

Having a sequence of images "exploded" into a third dimensional trail has advantages in providing a spatial context for a temporal presentation. It can provide an aid to visual memory unavailable in keyhole mode and can help users find and target a frame of interest. However, the advantages for visual data mining applications are more evident when the sequence is analyzed. For example, if the sequence consists of a sequence of 2D line graphs, then it may be relevant to know when the slope of the line changes from a bell-shaped curve to a U-shaped curve. If the images are 2D scatter plots, perhaps there are outliers that should be singled out. There are a limitless number of techniques that may be relevant to finding the interesting positions in the sequence, including clustering analyses. Our method provides a uniform method for viewing the output of such analyses. Our informal experience so far suggests that the most effective method for viewing an analyzed sequence is to drop out the "less interesting" frames from the trail while maintaining the original relative positions of the "interesting" ones. However, as the user advances the sequence, all the frames are played in the capture frame to aid continuity.

4 Examples

Figure 2 shows example screenshots from our prototype that we have built to explore design variations on this type of interface. In both cases the user can use the controls to move the whole sequence forwards or backwards at varying rates of speed. Focusing on just the capture frame in the frontmost position is like watching a video. In the example on the left, each image shows a graph of annual precipitation in Boston. The z dimension is ordered by year. One can see immediately that data does not exist for all years. (Data collection was interrupted as a result of World War II.) The user can filter by adjusting a slider representing precipitation totals.

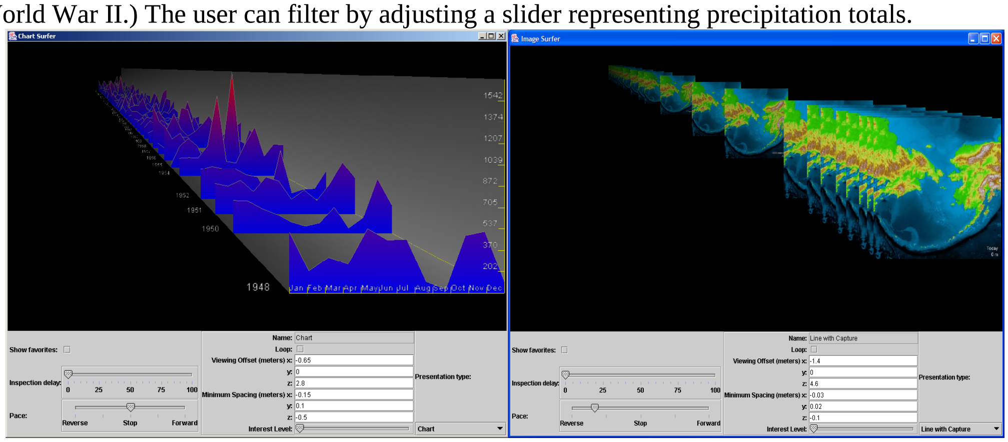

Figure 2: Screen snapshots from two datasets

On the right is data showing a reconstruction of the geological history of the flooding of the Bering land bridge (Manley, 2002). Each frame represents a view of the geology at intervals of 1000 years. However, the Z layout position is based on sea level, with the present day level in the foreground. From this view one can discern that the pace of sea level change was not as great in recent geological time as it was 8000 years ago.

5 Conclusion

This paper is about RSVP techniques we call *Shift, Analyze, and Collect* that have a place in larger systems for visual data mining. Our proposal is generic in that it can be combined with any two-dimensional presentation that is then ordered and analyzed in a third dimension. We hypothesize that an advantage to the user is a simplification of the navigational controls compared to usual approaches in 3D modeling. The research is at an early stage, having evolved from earlier work exploring the potential for rapid serial visual presentation in interfaces for video and other applications (Wittenburg et al., 2000). Future work will involve evaluation as well as incorporation of these techniques into a more complete system for visual data mining that includes multidimensional interactive bargrams (Wittenburg, Lanning, Heinrichs, & Stanton, 2001).

References

Anderson, D., Anderson, E., Lesh, N.B., Marks, J.W, Mirtich, B., Ratajczack, D.; Ryall, K. (2000) "Human-Guided Simple Search", National Conference on Artificial Intelligence (AAAI), August 2000 (AAAI Press).

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.) (1999) *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA USA: Morgan Kaufmann.

Coltheart, V. (Ed.) (1999) Fleeting Memories: Cognition of Brief Visual Stimuli. Cambridge, MA USA: MIT Press.

De Bruijn, O., & Spence, R. (2002) Patterns of Eye Gaze during Rapid Serial Visual Presentation. In Proceedings of Advanced Visual Interfaces (AVI 2002), (Trento, Italy, May 2002), 209-217.

Freeman, E., & Gelernter, D. (1996) Lifestreams: a storage model for personal data. *ACM SIGMOD Record*, 25(1), 80-86.

Keim, D. A. (2002) Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8 (1), 1-8.

Keim, D. A., & Kriegel, H.P. (1996) Visualization Techniques for Mining Large Databases: A Comparison. *IEEE Transactions on Knowledge and Data Engineering*, 8 (6), 923-938.

Mackay, W. E., & Beaudouin-Lafon, M. (1998) DIVA: exploratory data analysis with multimedia streams. *Proceedings of the ACM CHI'98 Conference on Human Factors in Computing Systems*, (Los Angeles, CA, April 1998), 416-423.

Manley, W.F. (2002) Postglacial Flooding of the Bering Land Bridge: A Geospatial Animation: INSTAAR, Univ of Colorado, v. 1, http://instaar.colorado.edu/QGISL/bering_land_bridge.

Spence, R. (2001) *Information Visualization*. Harlow, England: ACM Press & Addison-Wesley.

Spence R. (2002) Rapid, serial and visual: a presentation technique with potential. *Information Visualization*, 1 (1), 13-19.

Spence, R. & Apperley, M. D. (1977) The interactive-graphic man-computer dialogue in computer-aided circuit design. *IEEE Trans. on Circuits and Systems*, CAS-24 (2), 49-61.

Wittenburg, K., Ali-Ahmad, W., LaLiberte, D., & Lanning, T. (1998a) Polynesian Navigation in Information Spaces. In *Proceedings of CHI '98 Summary*, Los Angeles, CA USA April 1998, 317-318.

Wittenburg, K., Ali-Ahmad, W., LaLiberte, D., & Lanning, T. (1998b) Rapid-Fire Image *Previews for Information Navigation.* In *Proceedings of AVI '98: Advanced Visual Interfaces*, L'Aquila, ITALY, May 1998, 76-82.

Wittenburg, K., Chiyoda, C., Heinrichs, M., & Lanning, T. (2000) Browsing Through Rapid-Fire Imaging: Requirements and Industry Initiatives. In *Proceedings of Electronic Imaging '2000: Internet Imaging*, San Jose, CA, USA, January 26-28, 2000, 48-56.

Wittenburg, K. Lanning, T., Heinrichs, M., & Stanton, M. (2001) Parallel bargrams for Consumer-based information exploration and choice. In *Proceedings of 14th Annual Symposium of User Interface Software and Technology (UIST '01)*, Orlando, FL USA, 51-60.