Dynamic Statistical Graphics in the C2 Virtual Reality Environment

J. Symanzik

Department of Statistics
Iowa State University
Ames, IA 50011
symanzik@iastate.edu

D. Cook

Department of Statistics
Iowa State University
Ames, IA 50011
dicook@iastate.edu

B. D. Kohlmeyer
ICEMT
Iowa State University
Ames, IA 50011
mink@iastate.edu

U. Lechner
ICEMT
Iowa State University
Ames, IA 50011
uli@iastate.edu

C. Cruz–Neira
ICEMT
Iowa State University
Ames, IA 50011
carolina@icemt.iastate.edu

Abstract

The C2 is a high–end immersive virtual reality environment that allows 3–dimensional projections of higher dimensional objects. It is our intention to make use of the C2 to explore new statistical graphics applications based on single, multiple, or sequences of 3–dimensional projections. We also plan to explore navigation tools for steering through high–dimensional space, and interaction tools such as brushing and identification of points.

1 Introduction

This paper contains work related to dynamic statistical graphics and the C2 virtual reality environment. So far, most applications of dynamic statistical graphics have involved working with 2–dimensional display devices. The C2 is a high–end virtual reality environment that allows direct viewing of 3–dimensional projections of higher dimensional objects.

It is our intention to make use of the C2 to explore new statistical graphics applications based on single, multiple, or sequences of 3–dimensional projections. We also plan to coordinate the use of 3–dimensional navigation tools and other interactive graphical methodology such as interactive brushing, painting, and identification of observations. These methods have been finessed for the 2–dimensional screen but their realization in a virtual reality environment may require different approaches.

In Section 2 we will provide an overview on virtual reality and the C2 in particular. The background to dynamic statistical graphics methods will be discussed in Section 3. In Section 4 we will describe our approach to dynamic statistical graphics in the C2. We finish this paper with a discussion of what we have learned about the environment so far in Section 5 and an outline for near future work in Section 6.

2 Virtual Reality Overview and the C2

In this section we will define virtual reality (VR) in general and highlight a few aspects of its history. We will also introduce a particular virtual environment, the C2.

2.1 Definition of VR

Within the literature, there does not exist a unique definition of the term *Virtual Reality*, but almost every person has a different understanding. In Cruz–Neira (1993) the following definitions are given: "VR is the body of techniques that apply computation to the generation of experientially valid realities." (William Bricken), "Virtual reality is the place where humans and computers make contact." (Ken Pimentel, Kevin Teixeira), "VR has to do with the simulation of environments." (Gregory Newby), "VR provides real–time viewer–centered head–tracking perspective with a large angle of view, interactive control and binocular display." (Daniel Sandin), "Virtual reality refers to immersive, interactive, multi–sensory, viewer–centered, three–dimensional computer generated environments and the combination of technologies required to build these environments." (Carolina Cruz–Neira), "An experience in which a person is surrounded by a three–dimensional computer–generated representation and is able to move around in the virtual world and see it from different angles, to reach into it, grab it, and reshape it." (Howard Rheingold), "Virtual reality is a media to recreate the world in which we live and to create illusions of new and yet unknown worlds." (Anonymous).

In addition, terms such as *Artificial Reality*, *Virtual Environments*, and *Cyberspace* sometimes are used interchangeably for VR, while other people make clear distinctions between each of them.

2.2 A few Aspects of the History of VR

Even though there exist different definitions of VR, there is little doubt that the origin of VR dates back to 1965 when Ivan Sutherland proposed the *Ultimate Display* (Sutherland, 1965). In 1968 he also built the *Sword of Damocles* which is considered to be the first *Head Mounted Display* (HMD). It consists of two cathode ray tubes that are mounted alongside each of the user's ears and additional hardware that is suspended from the ceiling by a mechanical arm to measure the user's head position and orientation. In 1971 Frederick Brooks developed the *GROPE–II System*, which used the *ARM*, one of the first force–feedback devices. It was not before 1985 that Thomas Zimmerman designed the *Data Glove*, a device that is capable of measuring the degree to which each of the user's fingers is bent. Another VR device, the *BOOM* was commercialized in 1989 by Fake Space Labs. It is a small box that contains two cathode ray tubes which can be viewed through two eye holes. The box is attached to a mechanical arm that measures its position and orientation while users move it around to explore the virtual world.

A brief chronology of further events that influenced the development of VR can be found in Cruz–Neira (1993), a more complete overview can be found in Pimentel and Teixeira (1995).

2.3 The C2

The C2 is a device where the correct projection of the imagery on large screens creates a VR experience. The C2 is available at Iowa State University (ISU) at Ames. Technical details and applications of a VR environment similar to the C2 can be found in Cruz–Neira et al. (1992), Cruz–Neira et al. (1993a), Cruz–Neira et al. (1993b), Cruz–Neira (1995), Roy et al. (1995).

The C2 is a projection–based virtual reality environment, which uses 3D computer graphics, position tracking, and auditory feedback to immerse users in a 3D environment. It has a floor print of 12×12 feet and a height of 9 feet. Stereographic images are rear projected onto three side walls and front projected onto the floor (see Figure 1). The illusion of 3D is created through the use of CrystalEyes Stereographics' LCD shutter glasses and two high–performance Silicon Graphics graphics computers. The glasses are synchronized to the computer display through infrared emitters alternating the left and right eye viewpoints at 96 hz. The user's brain, as it does in the real world, combines the two views into a 3D stereoscopic image. The position and orientation of the user's hands and head are determined through the use of a magnetic based tracker, a cyberglove, and a hand–held wand. Audio feedback is transmitted to the user through multiple speakers.

multiple papers of VR.

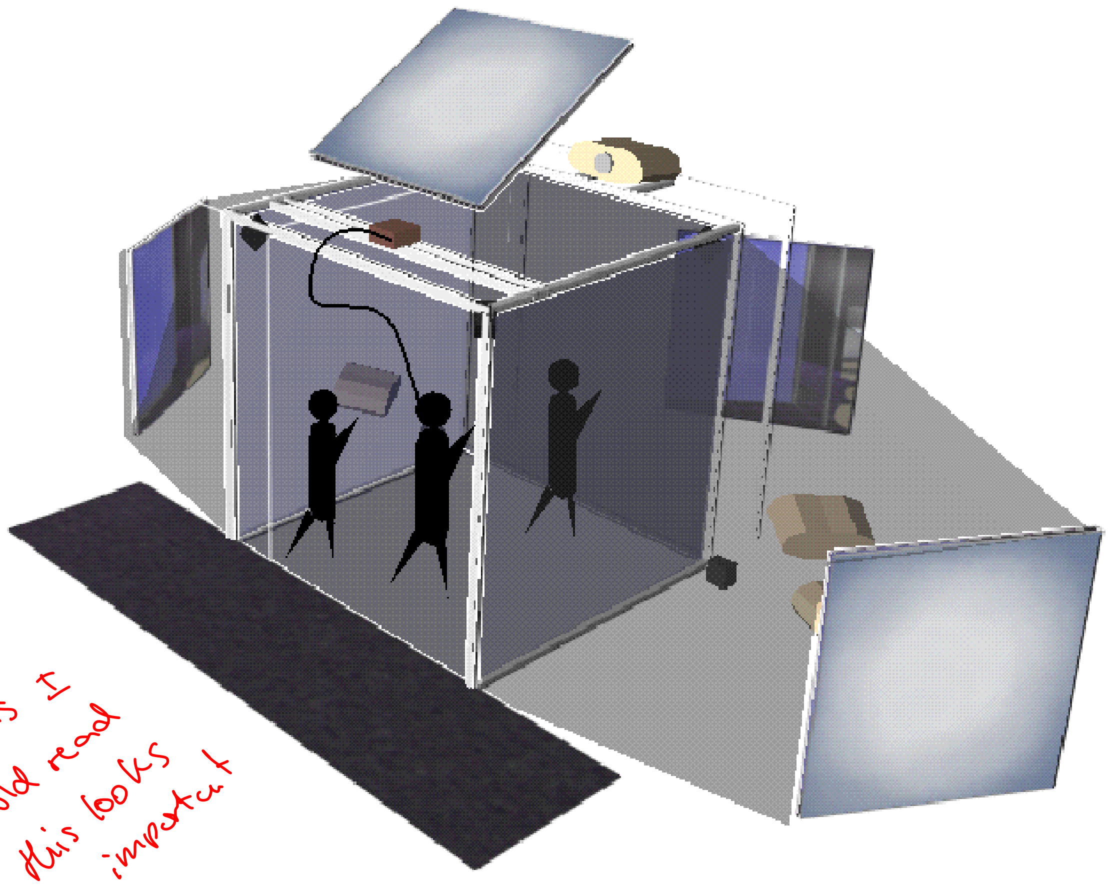
A detailed technical comparison of a VR environment similar to the C2 with other display devices for VR such as *Cathode Ray Tube* (CRT), *Binocular–Omni–Orientation–Monitor* (BOOM), and *Head Mounted Display* (HMD) can be found in Cruz–Neira et al. (1992). The C2 is an easy–to–learn, high–resolution VR interface that is superior to these devices in particular because of its full field–of–view, its visual acuity, and the lack of intrusion. It requires only very lightweight, unrestrictive equipment to be worn that does not make the user feel uncomfortable. Moreover, the C2 allows multiple viewers to enter the C2 and share the same virtual environment at the same time to benefit from the visual experience. Thus, it is a very helpful tool for collaborative work. It might also be helpful for a new user to join a guide, i. e., an expert navigator, in the C2 and get introduced to the particular problem before exploring the virtual environment him/herself.

3 Dynamic Statistical Graphics

Dynamic statistical graphics enables data analysts in all fields to carry out visual investigations leading to insights into relationships in complex data. Dynamic statistical graphics involves methods for viewing data in the form of point clouds or modeled surfaces. Higher dimensional data can be projected into 1–, 2– or 3–dimensional planes in a set of multiple views or as a continuous sequence of views which constitutes motion through the higher dimensional space containing the data.

There is a strong history of statistical graphics research on developing tools for visualizing relationships between many variables. Much of this work is documented in videos available from the American Statistical Association Statistical Graphics Section Video Lending Library (`http://orion.oac.uci.edu/~rnewcomb/`

look at this.



It seems I should read this looks important

Figure 1: *The C2.*

`statistics/graphics/graphics.html`).

A video clip of the successive stages in a multidimensional scaling algorithm (made by Kruskal, 1970) is one of the first examples of dynamic statistical graphics. A second example by Chang (1970) shows an interactive search for a structured 2–dimensional projection in 5–dimensions where 3 of the 5 dimensions are noise. PRIM–9 (Picturing, Rotation, Isolation and Masking in up to 9–dimensions; Fisherkeller et al. (1974a), Fisherkeller et al. (1974b)) is the landmark example of early dynamic statistical graphics. Projections formed the fundamental part of the visualization system, and were complemented with isolation and masking.

We mention these works to place our work in the C2 in the context of the earliest underpinnings of visualization of high–dimensional data. A good explanation of the importance of projection as a tool for visualizing structure in high–dimensional data can be found in Furnas and Buja (1994).

One major breakthrough in using projections for visualizing higher dimensions was made by Asimov (1985) in his work on the grand tour. The grand tour in an abstract sense shows a viewer all possible projections in a continuous stream (which could be considered to be moving planes through p–space). Several possibilities for "showing all possible projections" were explored in the original work, but the most successful method to arise from it is based on interpolating between random planes. Another common approach to displaying high–dimensional data can be found in (Becker and Cleveland, 1987) where data is plotted in a matrix of pairwise scatterplots (and users can do linked brushing between the plots which is very powerful). This is also an example of using projections to display multiple variables: it is a special set of projections along the coordinate axes.

Since the introduction of PRIM–9 most interactive and dynamic statistical graphics have been restricted to display at most 2 dimensions at a time. However, there have been some approaches to display statistical data in 3 or more dimensions. Stereo plots and anaglyphs have been used within statistics by Carr, Littlefield, and Nicholson (Carr et al., 1983; Carr and Littlefield, 1983; Carr and Nicholson, 1985; Carr et al., 1986). In particular, anaglyphs can be considered as an important means to represent 3–dimensional pictures on flat surfaces. They have been used in a variety of applithis can be done now

cation areas such as geometry, chemistry, architecture, and mining, but found only little use as yet in statistics. Interactive statistical anaglyph programs have been developed by Symanzik, Hering, and von der Weydt and obvious advantages of anaglyphs over 2–dimensional projections relevant for statisticians have been pointed out in Hering and von der Weydt (1989), Hering and Symanzik (1992), Symanzik (1992), Symanzik (1993a) and Symanzik (1993b). Other applications include cartographics mappings, spatial statistics, time series, and growth curves.

To our knowledge, there is only little research done that combines dynamic statistical graphics and VR, by researchers at George Mason University. Some basic ideas have been discussed in Section 6.9 of Wegman and Carr (1993).

4 Dynamic Statistical Graphics in the C2

We have begun the work from within the framework of 3–dimensional projections of p–dimensional data, using as a basis the methods developed and available in XGobi (Swayne et al., 1991). XGobi uses multiple linked windows to display scatterplots of multiple views of high–dimensional data. It has interactive and linked brushing, identification and scaling of the point scatter, and univariate, bivariate, trivariate and tour plot modes. This approach has a history of development throughout dynamic statistical graphics research history, as discussed in the previous section. The main difference is that with *XGobi the user interface is rather like a desktop with pages of paper whilst the C2 environment is more like having the whole room at our disposal for the data analysis.* So clearly there will be some applications which will benefit greatly from the C2 environment.

which?

The initial implementation contains a 3–dimensional grand tour (Asimov, 1985; Buja and Asimov, 1986; Buja et al., 1989; Buja et al., 1996). The basic idea of a grand tour is that a continuous sequence of projections is shown to the user. In XGobi the sequence contains 2–dimensional projections of the data, and in the C2 the projections are 3–dimensional. Although we will see in the example, that separate views of the data may be 1–, 2– or 3–dimensional. Continuity of motion allows the user to mentally make connections between different views of the data, and taking arbitrary 3–dimensional projections can expose features of the data not visible in marginal plots.

In the C2 there are a number of controls (see Figure 2). There is one viewing box, delimited by a wireframe cube with spheres at the vertices. At some future time we hope to have multiple viewing boxes but which will involve invoking more than one data drawing process. The viewing box can be re–shaped or rotated within the C2 environment. At the right hand front edge there is a speed pole so that the user can manually control the speed of motion during a tour with a gloved hand, visible as a floating hand and forearm in this image.

In the viewing box is a projection of a data set generated by placing points at the vertices of a 6–dimensional cube. The points are painted different colors and glyph types. The palette of colors available for painting is shown to the right of the viewing box. There are 4 different symbol types — sphere, cube, pyramid and star — and the size of the symbol can be changed by using the resizing pole next to the color palette. There are two different shapes of brush (not shown) that can be used — a sphere or a rectangular prism — which can be resized easily.

Not visible in the picture are variable spheres which show the contribution to the projection of each variable and also allow the user to select and deselect variables from being included in the tour.

Interaction with the control tools is reinforced by sound feedback.

5 Discussion

The C2 virtual reality environment is remarkably different to the display devices that are commonly available for data analysis. It extends beyond the small gain of one more dimension of viewing space, to being a completely defined "real" world space. The temptation is to grab the objects, and to flinch as a point flies into your nose. The objects surround the viewer so that the environment encompasses the viewer, inviting interaction with the data. A large amount of our current work has been on developing the user interface for dynamic statistical graphics and this is likely to continue.

One type of data and analysis which may benefit substantially from the environment is geographically referenced data: surfaces in 3–dimensions genuinely can be seen as such and in relation to other surfaces in the C2 environment in a manner not possible on a 2–dimensional computer display.

We need to emphasize that the C2 is a very expensive conglomeration of equipment, and we are lucky to have one available to us at Iowa State University, as well as very experienced personnel to introduce us to its use. We are taking advantage of this opportunity to explore what we, as statisticians, might gain from VR for a data analytic environment. Ultimately VR may become fairly

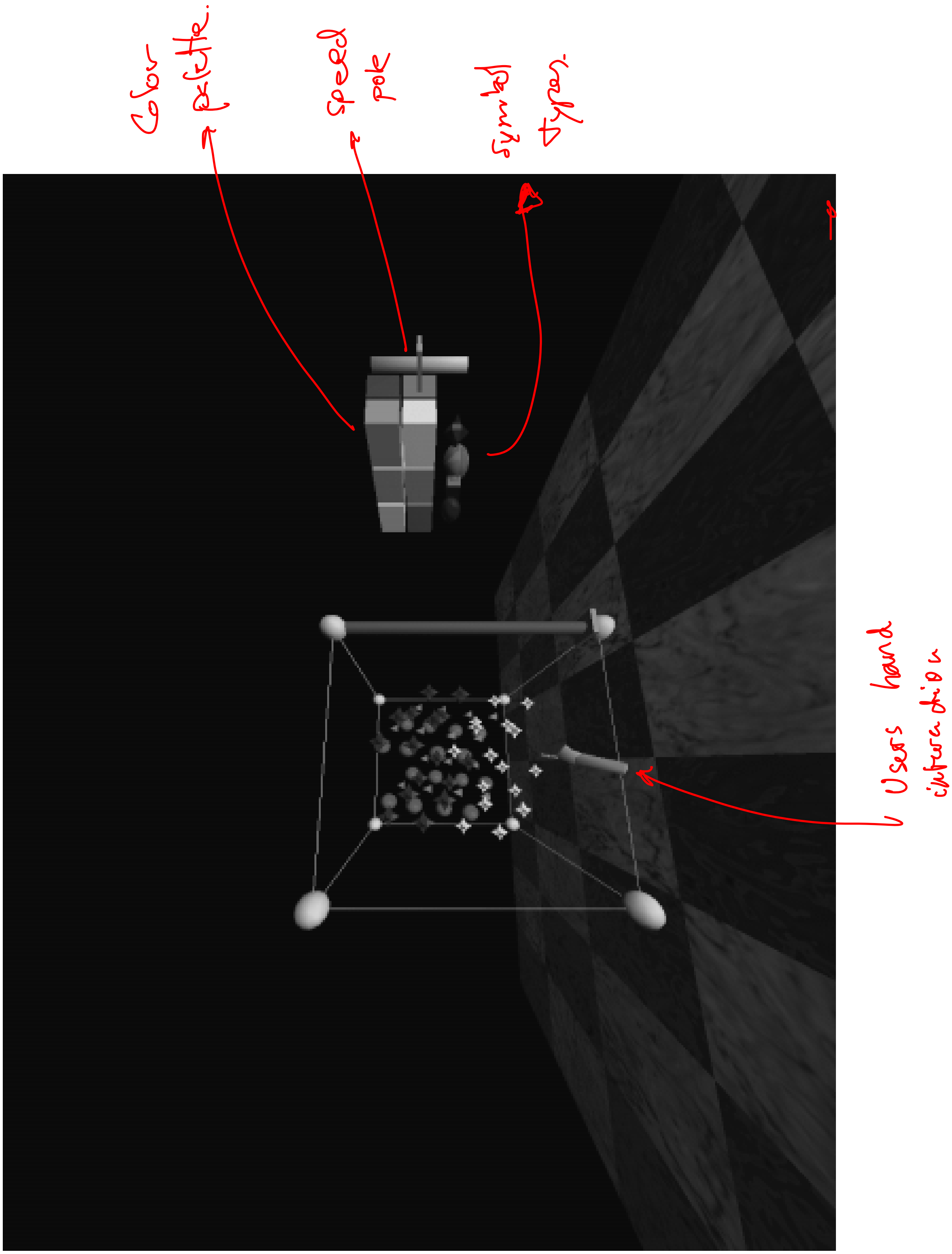


Figure 2: *Example view of the C2 statistical graphics application.*

commonly available, so we are hoping to develop tools for the C2 environment and help to shape the future of VR to suit our needs as data analysts.

6 Future Work

In the future the following directions of the use of the C2 should be explored for statistical graphics applications:

1. Development of navigation tools that allow us to keep track of the location of the current projection in the high–dimensional space, and relative positions of projections.

2. How much gain in intuition and understanding of high-dimensional structures do we get by using 3–dimensional projections over 2–dimensional projections?

3. Coordinating the use of other tools such as interactive brushing, painting, isolation, identification of data.

4. Displaying several 3–dimensional projections simultaneously.

5. Surface visualization, (using slicing or transparency) for density estimation, and curve estimation.

6. User interface issues: connecting interaction devices, and integrating sound.

7. Connecting to high–speed networks and connections to supercomputing for processing of massive data.

Acknowledgements

The research reported in this article has been supported by National Science Foundation Grant DMS–96–32662 and the U.S. Environmental Protection Agency through Cooperative Agreement CR822919 with Iowa State University. This paper has not been subjected to the Environmental Protection Agency's peer and administrative review.

References

Asimov, D. (1985). The Grand Tour: A Tool for Viewing Multidimensional Data. *SIAM Journal of Scientific and Statistical Computing*, 6(1):128–143.

Becker, R. A. and Cleveland, W. S. (1987). Brushing Scatterplots. In Cleveland, W. S. and McGill, M. E., editors, *Dynamic Graphics for Statistics*, pages 201–224. Wadsworth, Monterey, CA.

Buja, A. and Asimov, D. (1986). Grand Tour Methods: An Outline. In Allen, D. M., editor, *Proceedings of the 17th Symposium on the Interface between Computing Science and Statistics*, pages 63–67, Lexington, KY. Elsevier.

Buja, A., Asimov, D., and Hurley, C. (1989). Methods for Subspace Interpolation in Dynamic Graphics. Technical Memorandum, Bellcore.

Buja, A., Cook, D., Asimov, D., and Hurley, C. (1996). Dynamic Projections in High-Dimensional Visualization: Theory and Computational Methods. *Journal of Computational and Graphical Statistics*. Submitted.

Carr, D. B. and Littlefield, R. J. (1983). Color Anaglyph Stereo Scatterplots — Construction Details. In Gentle, J. E., editor, *Proceedings of the 15th Symposium on the Interface between Computer Science and Statistics*, pages 295–299, New York, NY. North–Holland Publishing Company.

Carr, D. B., Littlefield, R. J., and Nicholson, W. L. (1983). Color Anaglyph Stereo Scatterplots — Construction and Application. In *ASA Proceedings of the Section on Statistical Computing*, pages 255–257, Alexandria, VA. American Statistical Association.

Carr, D. B. and Nicholson, W. L. (1985). Evaluation of Graphical Techniques for Data in Dimensions 3 to 5: Scatterplot Matrix, Glyph, and Stereo Examples. In *ASA Proceedings of the Section on Statistical Computing*, pages 229–235, Alexandria, VA. American Statistical Association.

Carr, D. B., Nicholson, W. L., Littlefield, R. J., and Hall, D. L. (1986). Interactive Color Display Methods for Multivariate Data. In *Statistical Image Processing and Graphics*, pages 215–250. Marcel Dekker, New York, NY.

Chang, J. (1970). Real-time Rotation. ASA Statistical Graphics Video Lending Library (contact: dfs@bellcore.com).

Cruz-Neira, C. (1993). Virtual Reality Overview. SIGGRAPH '93 Course Notes #23. pp. 1–18.

Cruz-Neira, C. (1995). *Projection–based Virtual Reality: The CAVE and its Applications to Computational Science.* PhD thesis, University of Illinois at Chicago.

Cruz-Neira, C., Leigh, J., Papka, M., Barnes, C., Cohen, S. M., Das, S., Engelmann, R., Hudson, R., Roy, T., Siegel, L., Vasilakis, C., DeFanti, T., and Sandin, D. J. (1993a). Scientists in Wonderland: A Report on Visualization Applications in the CAVE Virtual Reality Environment. In *IEEE 1993 Symposium on Research Frontiers in Virtual Reality*, pages 59–66.

Cruz-Neira, C., Sandin, D. J., DeFanti, T., Kenyon, R., and Hart, J. C. (1992). The CAVE, AudioVisual Experience Automatic Virtual Environment. *Communications of the ACM*, 35:64–72.

Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993b). Surround–Screen Projection–Based Virtual Reality: The design and Implementation of the CAVE. In *ACM SIGGRAPH '93 Proceedings*, pages 135–142, Anaheim, CA.

Fisherkeller, M., Friedman, J. H., and Tukey, J. (1974a). PRIM-9: An Interactive Multidimensional Data Display and Analysis System. ASA Statistical Graphics Video Lending Library (contact: dfs@bellcore.com).

Fisherkeller, M., Friedman, J. H., and Tukey, J. (1974b). PRIM-9: An Interactive Multidimensional Data Display and Analysis System. Technical Report SLAC-PUB-1408, Stanford Linear Accelerator Center, Stanford, Ca.

Furnas, G. and Buja, A. (1994). Prosection Views: Dimensional Inference Through Sections and Projections. *Journal of Computational and Graphical Statistics*, 3(4):323–385.

Hering, F. and Symanzik, J. (1992). Anaglyphen 3D — Ein Programm zur interaktiven Anaglyphendarstellung. Forschungsbericht 92/1, Fachbereich Statistik, Universität Dortmund, (in German).

Hering, F. and von der Weydt, S. (1989). Interaktive Anaglyphendarstellungen als Hilfsmittel zur Analyse mehrdimensionaler Daten. Forschungsbericht 89/7, Fachbereich Statistik, Universität Dortmund, (in German).

Kruskal, J. B. (1970). Multidimensional Scaling. ASA Statistical Graphics Video Lending Library (contact: dfs@bellcore.com).

??

Pimentel, K. and Teixeira, K. (1995). *Virtual Reality through the new Looking Glass (Second Edition).* McGraw–Hill, New York, NY.

Roy, T., Cruz-Neira, C., and DeFanti, T. A. (1995). Cosmic Worm in the CAVE: Steering a High Performance Computing Application from a Virtual Environment. *Presence: Teleoperators and Virtual Environments*, 4(2):121–129.

Sutherland, I. E. (1965). The Ultimate Display. In *Proc. IFIP 65, 2*, pages 506–508, 582–583.

Swayne, D. F., Cook, D., and Buja, A. (1991). XGobi: Interactive Dynamic Graphics in the X Window System with a Link to S. In *ASA Proceedings of the Section on Statistical Graphics*, pages 1–8, Alexandria, VA. American Statistical Association.

Symanzik, J. (1992). Computerdarstellungen von Anaglyphen — Ein Hilfsmittel der multivariaten Statistik. Diplomarbeit, Fachbereich Informatik, Universität Dortmund, (in German).

Symanzik, J. (1993a). Anaglyphen 3D — A Program for the Interactive Representation of Three–Dimensional Perspective Plots of Statistical Data. In Opitz, O., Lausen, B., and Klar, R., editors, *Information and Classification. Concepts, Methods and Applications. Proceedings of the 16th Annual Conference of the "Gesellschaft für Klassifikation e. V."*, pages 384–389, Berlin, Heidelberg. Springer.

Symanzik, J. (1993b). Three–Dimensional Statistical Graphics based on Interactively Animated Anaglyphs. In *ASA Proceedings of the Section on Statistical Graphics*, pages 71–76, Alexandria, VA. American Statistical Association.

Wegman, E. J. and Carr, D. B. (1993). Statistical Graphics and Visualization. In Rao, C. R., editor, *Handbook of Statistics, Vol. 9*, pages 857–958. Elsevier Science Publishers, Amsterdam.