Dynamic Statistical Graphics in the CAVE Virtual Reality Environment

Jürgen Symanzik, Dianne Cook,
Department of Statistics, Iowa State University, Ames, IA, 50011
Bradley D. Kohlmeyer, Carolina Cruz–Neira
Iowa Center for Emerging Manufacturing Technology, Iowa State University, Ames, IA, 50011
e–mail: symanzik@iastate.edu

Abstract

The CAVE 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 CAVE 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.

1 Introduction

This paper contains work related to dynamic statistical graphics (DSG) and the CAVE virtual reality environment. So far, most applications of dynamic statistical graphics have involved working with 2–dimensional display devices. The CAVE is a hig–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 CAVE 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 CAVE in particular. The background to dynamic statistical graphics methods will be discussed in Section 3. In Section 4 we will describe our approach to DSG in the CAVE. 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 CAVE

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 CAVE.

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 & Teixeira (1995).

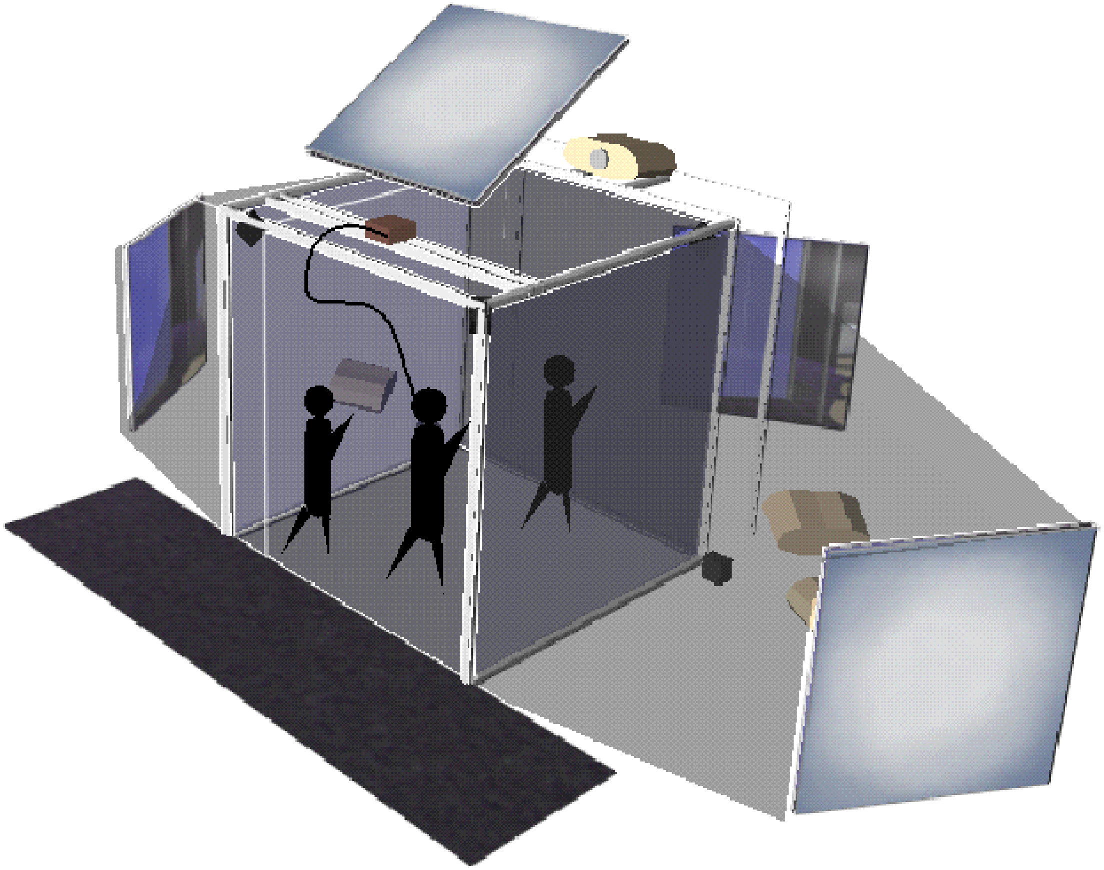

Figure 1: *The CAVE.*

2.3 The CAVE

The CAVE is a device where the correct projection of the imagery on large screens creates a VR experience. The CAVE was designed and implemented in 1991 at the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago. An extended version, called CAVE^{++}, is available at Iowa State University (ISU) at Ames since May 1996. The name, "CAVE", has two meanings: it is a recursive acronym (CAVE Automatic Virtual Environment) and it is also a reference to "The Alegory of the Cave" found in Plato's Republic (Plato c.375BC), in which the philosopher discusses inferring reality (ideal forms) from projections (shadows) on the cave wall. Technical details and applications of the CAVE can be found in Cruz-Neira, Sandin, DeFanti, Kenyon & Hart (1992), Cruz-Neira, Sandin & DeFanti (1993), Cruz-Neira, Leigh, Papka, Barnes, Cohen, Das, Engelmann, Hudson, Roy, Siegel, Vasilakis, DeFanti & Sandin (1993), Cruz-Neira (1995), and Roy, Cruz-Neira & DeFanti (1995).

The CAVE is a projection–based virtual reality environment, which uses 3D computer graphics, position tracking, and auditory feedback to immerse users in a 3D environment. The CAVE^{++} at ISU 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.

A detailed technical comparison of the CAVE 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 CAVE 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 CAVE allows multiple viewers to enter the CAVE 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 CAVE and get introduced to the particular problem before exploring the virtual environment him/herself.

3 Dynamic Statistical Graphics

Dynamic statistical graphics (DSG) enables data analysts in all fields to carry out visual investigations leading to insights into relationships in complex data. DSG 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 (contact: `dfs@bellcore.com`).

A video clip of the successive stages in a multidimensional scaling algorithm (made by Kruskal (1970)) is one of the first examples of DSG. 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, Friedman & Tukey (1974*a*), Fisherkeller, Friedman & Tukey (1974*b*)) is the landmark example of early DSG. 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 CAVE 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 & 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 & Cleveland (1987) where data is plotted in a matrix of pairwise scatterplots (and users can do linked brushing between the plots which is quite 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, Littlefield & Nicholson 1983, Carr & Littlefield 1983, Carr & Nicholson 1985, Carr, Nicholson, Littlefield & Hall 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 application 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 & von der Weydt (1989), Hering & Symanzik (1992), Symanzik (1992), Symanzik (1993*a*), and Symanzik (1993*b*). Other applications include cartographics mappings, spatial statistics, time series, and growth curves.

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

4 Dynamic Statistical Graphics in the CAVE

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, Cook & Buja 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 CAVE 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 CAVE environment.

The initial implementation contains a 3–dimensional grand tour (Asimov 1985, Buja & Asimov 1986, Buja, Asimov & Hurley 1989, Buja, Cook, Asimov & Hurley 1996). The basic idea to 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 CAVE 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 CAVE there are a number of "goody boxes" and controls (see Figure 2. At present there are three goody boxes, one for a point cloud view, another for a convex hull, and

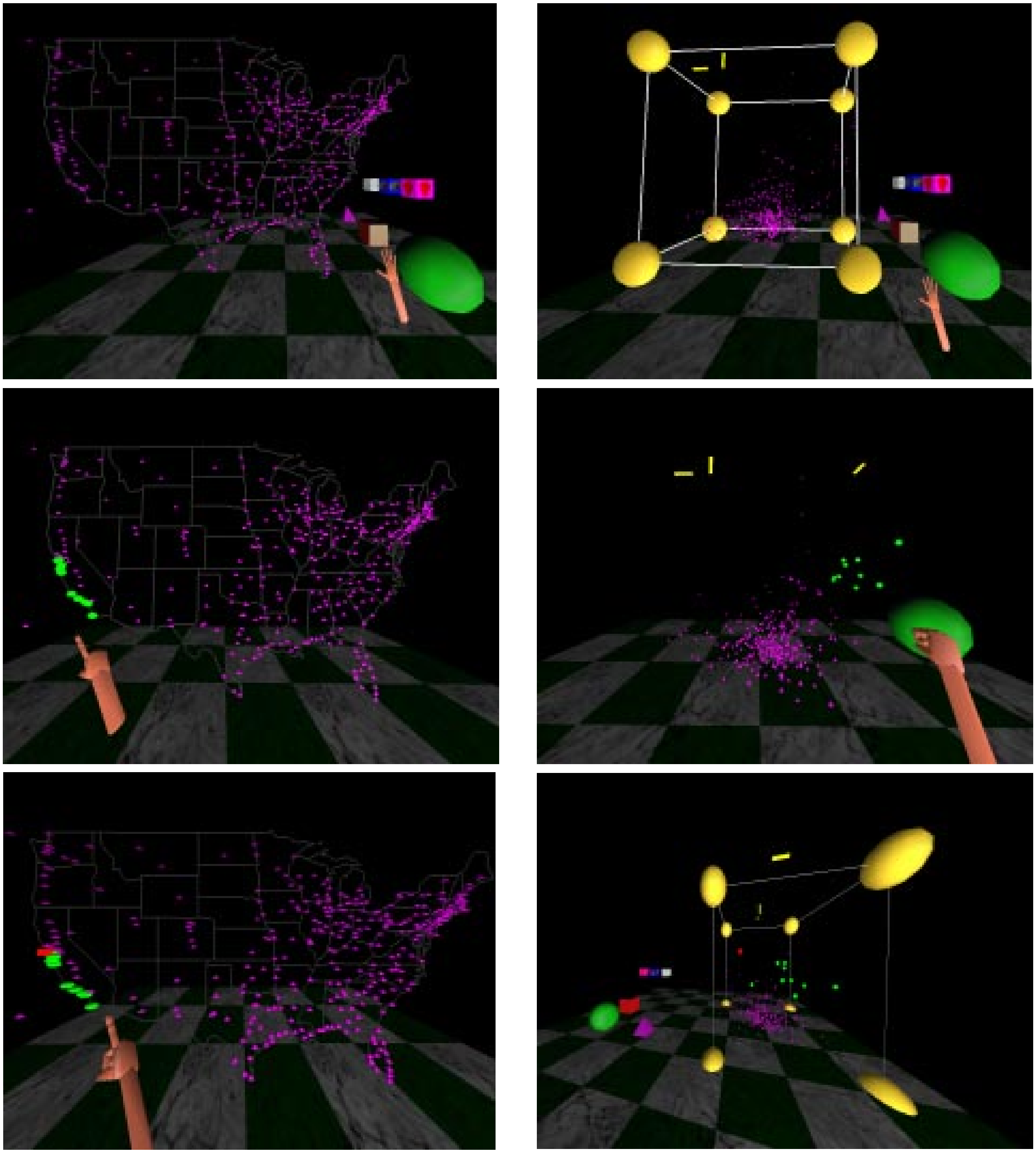

Figure 2: *Places Data in the CAVE: Mapview (Left) and 3–dimensional point cloud displaying housing cost, climate & terrain, and education (Right). "Goody boxes", "glyph" types, and boundary box are visible (Top Row). Cities with nice climate & terrain and high housing cost are brushed and happen to fall into California (Middle Row). Among the brushed points is one city (San Francisco) with an outstanding value for education (Bottom Row).*

another (as yet incomplete) box for a density estimate. The user can grab the goody box to bring another viewing box of either a point cloud, convex hull, or density estimate into the

CAVE. Each viewing box has its own controls: a speed "pole" for adjusting the velocity of the touring, a set of variable "spheres" to allow variables to be added or removed from the view. The spheres also contain a radial line indicating the direction and magnitude of the variable's contribution to the 3–dimensional projection.

Near the goody boxes are a selection of colors and "glyph" types (cube, sphere, pyramid). The user can select a color and glyph to paint points in the scatter cloud. Corresponding points in different views will simultaneously be painted. Changing the glove action to identification results in the points being verbally identified as the gloved hand moves around the point cloud.

In Figure 2 we see an old and familiar (to statisticians) data set being viewed in the CAVE. This is the "places data" (Boyer & Savageau 1981), which was distributed to ASA members some years ago so that they could apply contemporary data analytic methods to describe the data and then present results in a poster session at the ASA annual conference. It contains nine measures of livability at 329 cities in the United States: climate & terrain, housing costs, health care & environment, crime rate, transportation, education, arts, recreation, and economics. The latitude and longtitude of the cities have been added as additional variables to the data (by Paul Tukey).

We have used "one virtual wall" of the CAVE as a pin board on which to tack a map of the United States and place a point at the latitude and longtitude of each city in the data set. In the center of the room is a 3–dimensional point cloud displaying the variables housing cost, climate & terrain, and education. We can brush an unusual group of points in a different color and symbol by grabbing a new glyph and color with the glove and moving the gloved hand through the point cloud. Corresponding points in the map will be painted accordingly. Pointing at a location in the map generates a verbal identification of the name of the city.

5 What We Have Learned So Far

The CAVE virtual reality environment is remarkably different to the display devices that are commonly available for data analysis. Even a skeptical mind should be convinced from experiencing the environment that it has huge potential for data analysis. It extends way 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 future work will be on developing the user interface for DSG.

One area which stands to 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 CAVE environment in a manner not possible on a 2–dimensional computer display. One question that immediately arose is whether the map should be "pinned to the wall" or underneath our feet. Pinned to the wall is feasible in the real world so it may be more natural to look at geographic data this way, but a surface plot of the elevation over a geographic domain may be better viewed from a bird's eye perspective and thus located on the floor.

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

6 Future Work

So far, we have only looked at two examples of statistical data in the CAVE (i. e., the 3–dimensional grand tour and the linked brushing between a map and a 3–dimensional point cloud). In the future the following directions of the use of the CAVE should be explored for statistical graphics applications:

1. Adaptation/Implementation of methods for 3–dimensional projections.

2. Adapting the CAVE interface to allow the different methods for generating sequences to be applied at will.

3. 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.

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

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

6. Displaying several 3–dimensional projections simultaneously.

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

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

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

The inclusion of spatial data, in particular, satellite imagery, makes the intended research computationally heavy. Small satellite images consist of several million points — usually multivariate data of the spectral bands measured at each pixel. Spatial analysis will involve spatial covariance modelling which can potentially square the number of points to be processed. Therefore, we need high performance software and hardware that allows us to process these massive data sets in real time.

We plan to link software that provides spatial data at a given site (e. g., US agencies) without permanently copying this data with software that provides new applications for DSG in the CAVE environment. We expect the exchange of data and the communication of the different programs to happen in real time which requires the use of modern communication networks.

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 Scientific and Statistical Computing* **6**(1), 128–143.

Becker, R. A. & Cleveland, W. S. (1987), Brushing Scatterplots, *in* W. S. Cleveland & M. E. McGill, eds, 'Dynamic Graphics for Statistics', Wadsworth, Monterey, CA, pp. 201–224.

Boyer, R. & Savageau, D. (1981), *Places Rated Almanac*, Rand McNally, Chicago, IL.

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

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

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

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

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

Carr, D. B. & 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', American Statistical Association, Alexandria, VA, pp. 229–235.

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

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. & Sandin, D. J. (1993), Scientists in Wonderland: A Report on Visualization Applications in the CAVE Virtual Reality Environment, *in* 'IEEE 1993 Symposium on Research Frontiers in Virtual Reality', pp. 59–66.

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

Cruz-Neira, C., Sandin, D. J., DeFanti, T., Kenyon, R. & Hart, J. C. (1992), 'The CAVE, AudioVisual Experience Automatic Virtual Environment', *Communications of the ACM* **35**, 64–72.

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

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

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

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

Plato (c.375BC), *The Republic*, The Academy, Athens.

Roy, T., Cruz-Neira, C. & 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', pp. 506–508, 582–583.

Swayne, D. F., Cook, D. & 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', American Statistical Association, Alexandria, VA, pp. 1–8.

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

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

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

Wegman, E. J. & Carr, D. B. (1993), Statistical Graphics and Visualization, *in* C. R. Rao, ed., 'Handbook of Statistics, Vol. 9', Elsevier Science Publishers, Amsterdam, pp. 857–958.