EDITORIAL

"Three in 93 and four in 94" was proclaimed as the rallying call of the new editors of the joint *Statistical Computing and Statistical Graphics* newsletter. If we have succeeded then this first issue of 1993 should be in your hands before the April Interface meeting. We are planning to mail the second issue before the Joint Statistical Meetings in August, and the third sometime around Thanksgiving. More issues mean the newsletter can be more timely, provide announcements and encourage dialogue. Please use your newsletter to communicate with the membership of the two largest sections of the ASA. Our deadlines for the remaining issues in 1993 are the last day of June and October.

Many regular columns will continue, but we solicit your help with new ideas and offers to write columns or once-only pieces. Please keep those e-cards and e-letters coming!

This issue has two feature articles. The first feature describes the availability and discusses the design of public domain matrix and linear algebra routines. Many academic computer installations can make available high quality subroutines of the standard algorithms without the need to purchase or lease commercial software packages, e.g. IMSL or NAG. Algorithms are available, as described by Colin Goodall, to anyone with FTP software and access to the internet.

The second feature, about graphics and stereoscopic displays, is like the Sunday night mini-movie on network television. It looks like a feature article, but is really the first episode of a new column. Dan Carr has stepped down as editor but couldn't resist the challenge of a regular column. The graphic images which accompany this article should also be seen as the beginning of more innovative graphical material which we would like to print.

The Newsletter is now being set/typeset in LaTeX. After three years of colorful newsletters we have tried to restrain ourselves to typographic spice and a change in format. Neither of us has any sense of color, so we will take the conservative (and probably boring) tack and use black text on a neutral background. We want to credit Kevin Fox, the design artist from Penn State, for the new masthead and for keeping our link to the past with the intersecting circles punctuating the articles.

Submissions should be sent by email to either of the editors. If you can prepare your article in TeX or LaTeX that will make our lives just a little easier. Otherwise plain old ASCII format is fine.

In preparing this issue we have realized yet again the herculian efforts of the newsletter's founding editors, Sallie Keller-McNulty and Dan Carr. We hope we can live up to the standards they have set for us.

James L. Rosenberger
Editor Statistical Computing Section
JLR@stat.psu.edu

Mike Meyer
Editor, Statistical Graphics Section
mikem@stat.cmu.edu

SECOND FEATURE

Production of Stereoscopic Displays for Data Analysis

Daniel B. Carr *George Mason University*

Dedicated to David L. Hall, colleague and friend.

Stereoscopic displays help the analyst escape from the limited domain of 2-D visualization into the natural domain of 3-D visualization. The goal of producing 3-D scatterplots motivates much of the following discussion. The goal has strong implications in terms of selecting a stereo projection. In the every day world, familiarity with objects and many depth cues facilitates fusion of left and right retinal images into a stereo image. Monocular depth cues include linear perspective (objects of equal size transect areas inversely proportional to their distance), interposition or occlusion (when one object is in front of another it obscures the more distant object), shadows (we generally assume light comes from above), detail perspective (no fine detail appears in distant objects due to limited visual acuity), and aerial perspective (greater optical depth through the air leads to a blue shift). For most elementary 3-D scatterplots prior knowledge about the form to be perceived is limited and monocular cues are restricted, so care must be taken in the selection of a stereo projection.

Two Infinite Families of Reasonable Stereo Projections

Different geometric models lead to different stereo projections. (Geometric models are idealized in that each eye has a blind spot, a region of high resolution, and various imperfections.) A simple model (Newman and Sproull 1979) presumes that the eyes converge on a single focal point and constructs left and right images by projecting onto left and right projection planes. The projection planes contain the focal point and are orthogonal to the respective lines of site. This fixed-focal-point model is appropriate for advanced dynamic systems that update immediately as the eyes change their focal point. However, the fixed-focal-point projection does not correspond to the data analyst's typical stereo-viewing scenario. Valyus (1962) states, "it has been shown experimentally that eye movements performed when stereoscopic pictures are viewed are similar to those performed in observing a real object. As the gaze is transferred from one object to another the eyes perform conjugate movements directed to the subjectively most important regions, and at the same time co-ordinated convergence movements take place." While the fixed-focal-point projection has proved passable for showing non-updated images of familiar scenes, image fusion problems result when looking at points in the corners of the plot. Thus fixed-focal-point projection is inadequate for 3-D scatterplots.

When the eyes have multiple focal points within the same stereo image, a reasonable compromise uses a single common projection plane that is parallel to a frontal view of the face. In a multiple-focal-point model, a data point projected into the projection plane has the same y coordinate for both left and right images. This is a fundamental requirement for 3-D scatterplot projections.

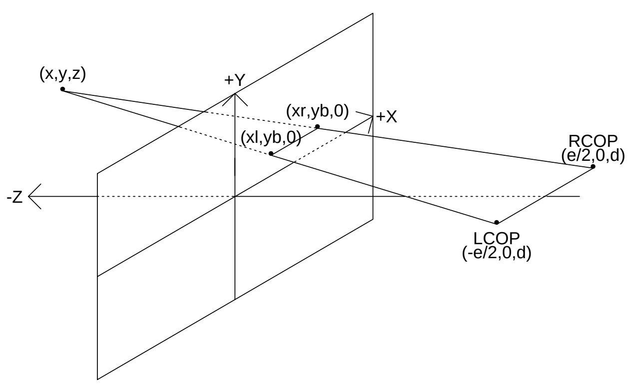

Figure 1

Two classes of projections satisfy the multiple-focal-point (single projective plane) constraint. The first class uses standard projective methods with separate centers of projection for the left and right eyes, denoted LCOP and RCOP respectively. Assume a right handed coordinate system with positive z toward the viewer. Then the LCOP and RCOP coordinates relative to the center of the workstation screen are $(-e/2, 0, d)$ and $(e/2, 0, d)$ where e is the eye separation (see Figure 1). The projected coordinates can then be found by scaling the vector from the eye to the data point by a constant, s, so that the scaled vector touches the screen (the z coordinate is 0). For the right eye this yields

$$s \times [(x, y, z) - (e/2, 0, d)] + (e/2, 0, d) = (x_r, y_r, 0) \quad (1)$$

Solving both right-eye and left-eye equations for s based on the z coordinate and substituting yields:

$$\begin{aligned} x_r &= (dx - ez/2)/(d - z) \\ x_l &= (dx + ez/2)/(d - z) \\ y &= dy/(d - z) \end{aligned} \quad (2)$$

The fact the both left and right images have the same y coordinate is evident from geometric considerations. Consider a data point appearing behind the projection plane (or viewing screen). The two eyes and this data point form a triangle that intersects the screen. If the frontal view of the face is parallel to the screen and the eyes are level, then the twin projected points must also be level.

The second class of multiple-focal-point stereo projections may be called a depth-cued orthogonal projection. In this projection, the LCOP and RCOP are shifted for each data point so that the midpoint between the eyes has the same x and y coordinates as the display-scaled data point to be projected. The depth-cued points are then

$$\begin{aligned} x_r &= x - ez/[2 \times (d - z)] \\ x_l &= x + ez/[2 \times (d - z)] \\ y &= y \end{aligned} \quad (3)$$

The two projection classes have continuous projection and viewing parameters. Projection parameters include eye separation (e), projection distance (d) and the size of the viewing cube into which we translate and scale the data. For convenience consider the workstation screen as a viewing cube that is 20 centimeters on a side with its front face centered and aligned with the screen surface. Viewing conditions may differ from the projection model and that introduces magnification (m) and viewing distance (d') parameters. These parameters can be varied over an interval and still lead to comfortable image fusion of a 3-D point cloud. Consequently both stereo projection classes are infinite.

Projection Parameter Bounds and Parallax

While an infinite number of projections are effective, some parameter bounds should not be violated and some projections are more desirable than others. The following provides some background concerning parameter constraints. The curious reader is referred to Valyus (1962) for additional detail.

For most people, eye separation falls in the interval between 5.2 and 7.4 centimeters. Fortunately, using the exact eye separation for each individual is not crucial as evidenced by stereo publications that are enjoyed by diverse audiences.

Parallax is a key concept for understanding the projection parameter bounds. The horizontal parallax, p, of a point refers to the distance between the projected coordinates on the screen, x_r and x_l. Then for our multiple-focal-point stereo projection

$$p = x_r - x_l = -ez/(d - z) \tag{4}$$

A point appearing in front of the screen will have a positive z, $z < d$, and the parallax will be negative. Similarly, a point appearing behind the screen will have positive parallax. Parallax must be limited to provide acceptable stereo fusion. Maintaining image focus constrains the amount of acceptable parallax. Suppose the eyes converge on the twin images of a point as if the point were real. The apparent location of the point is the focal point. Eye convergence is usually coupled with accommodation (lens focusing) so that a region in front and in back of the focal point is in focus. If this region includes the workstation screen the image of the point will be clear. While those experienced in stereo viewing often learn to decouple the convergence and accommodation of their eyes, a mismatch can lead to either fusion or focus problems. Constraining the parallax to keep the perceived image depth close to the screen avoids such problems. The horizontal parallax is related to angular parallax on the retina. Studies (Valyus 1962, Yeh and Silverstein 1990) have related the speed of image fusion to angular parallax and provide guidelines. In short, restricting the viewing cube to 23 centimeters on a side or smaller will generally satisfy the less restrictive (but slower fusion) Valyus bound. Lipton (1982) recommends the equivalent of centering the viewing cube depthwise on the screen to control the parallax. In this case the viewing cube can be made substantially larger before the parallax becomes excessive.

Magnification and Viewing Distance

View related parameters include the viewing distance d' and the image magnification m. The equation of parallax can be solved for the apparent distance from a point to the screen, z', in terms of the actual viewing distance, d'.

$$z' = d'/(1 - e/p) \tag{5}$$

If the parallax is fixed then doubling the viewing distance doubles the apparent distance of the point to the screen. Thus the viewing cube frame can be made to appear squashed or elongated by selecting a different viewing distance, d', than actual projection distance, d.

Often a stereo image is shown to a large audience on a big screen. If the image is not designed for the room, magnification may cause problems. Magnifying the image magnifies the parallax. A Taylor series expansion of the depth equation about $p = 0$ yields

$$z' = -d' \times (p/e + (p/e)^2 + (p/e)^3 \cdots) \tag{6}$$

When the parallax is small relative to eye separation, magnifying the image increases the apparent depth almost linearly. As the parallax to eye-separation ratio approaches one-half, the nonlinear terms contribute equally and magnification distorts the image. As magnified parallax approaches eye separation, the depth of image theoretically approaches minus infinity. This causes both depth distortion and fusion problems before the apparent depth reaches infinity.

Control of Perspective Differences

Under natural conditions the left and right eyes have different views of the world. The field of view for a single human eye is about 150° horizontally and 135° vertically. In binocular vision the field of view covered by both eyes is 120° so 30° or $1/5$ of each eye's field of view is unique to the eye. For 3-D scatterplots all points must be seen unless hidden by occlusion, so images must be restricted to the shared field of view. If the viewing cube is sufficiently smaller than available display space on the workstation screen this is not a problem.

Hodges (1992) provides an astute comment on stereo production for workstations. He notes that some hardware systems provide only a single center of projection. Then the standard trick for producing stereo projections is to use the "off-axis" projection. For each eye's view this shifts the data toward midpoint between the eyes, projects from the midpoint and then shifts the result back. While the projections produce coordinates identical to the LCOP and RCOP projections, the fields of view differ. This provides another reason to constrain the size of the viewing cube. For scatterplots all corresponding left and right points must appear on the screen.

Perspective differences within the shared field of view can cause fusion and interpretation problems. Perspective induced fusion problems may be evident when showing the viewing-cube frame in small side-by-side plots. If, for example, the left cube face has the x coordinate at $-e/2$, the left cube face will project as a line in the left eye image but as a trapezoid in the right eye image. This radical perspective-induced discrepancy complicates image fusion. Since the viewing cube frame is simple, avoiding radical differences in perspective is straightforward. For example rotating the viewing cube often suffices. In fact some analysts prefer perspective views of the cube that have two or

Figure 2 - Random Points on a Mobius Strip. The structure is apparent with any 3-D viewing approach. In flatland recognizing the structure is much harder. Piecing scatterplot matrix views together is not so easy without brushing or conditioning. Conditioning or slicing is helpful in moving to higher dimensions. For real data, overplotting is more of a problem and density- based presentations become advantageous.

Figure 3 - Selected Stereo Contours for a 3-D Density Estimate.

Figure 4 - Steepest Ascent Ridge Traces for a 2-D Density Surface. Ridges and their projections in domain space provide another way of viewing density estimates. This example is provided courtesy of Qiang Luo.

Figure 5 - Random Dot Stereogram. This contour completion illusion at different depths takes a while to fuse. This image is provided courtesy of Nathan Carr, who used it in a high school science project that repeated experiments pioneered by Bela Julesz (1971).

three infinity points rather than the common head-on view that has only one infinity point. Another trick bases the projection on a distance substantially larger than the actual viewing conditions. This reduces perspective differences.

The foreshortening (interpoint distances appear smaller as a function of depth) of perspective views complicates the reading of coordinates based on given axes. Equations (3) provide an orthographic stereo projection that removes all perspective differences and allows the x and y coordinates to be interpreted directly. Carr and Littlefield (1983) describe a simple implementation that exploits the scaling in standard statistical graphics packages. Plot the left eye and right eye coordinates of a point using $(X - X_p, Y)$ and $(X + X_p, Y)$ respectively where the values are in data units. The expression for the half parallax, X_p, in X data units, is

$$X_p = k \times \text{range}(X) \times (Z - Z_{\text{min}})/\text{range}(Z) \quad (7)$$

The constant k is chosen as a conveniently small value such as .026.

Littlefield provided a color-anaglyph implementation in a statistical package over a decade ago by modifying Minitab (remember when source code was available). This modification allowed drawing of arbitrary glyphs and included color table control to handle light mixing. Red and green points overplotted as yellow as they should and color control allowed subtle depth-based shading. Color polaroids demonstating that a third variable helps little in group discrimination were shown at the first ASA Graphics Exposition in 1982 and additional examples were published (Carr, Nicholson, Littlefield, and Hall 1986). While color anaglyph stereo is not the most desirable form of stereo, the anaglyph work demonstrated the simplicity of stereo plot production as long as statistical packages provide for control of color mixing (hint).

Side by Side Stereo Examples

An article on graphics should have some graphics. Stereo production methods are incredibly diverse. For workstations, time-multiplexed methods that alternately route images to the left and right eyes are gaining in popularity. Side-by-side methods are common in non-electronic publications and so will be used here. Note that other disciplines often publish stereo images in color (*Editors' note: Maybe someday the newsletter will be able to afford to do this*). For example see the ray-traced image in Hodges (1992). The additional color-based depth cues are especially helpful when showing surfaces. The images here are monochrome point and line drawings and for brevity focus on geometry rather than data analysis. The legends provide a brief description of the simple examples.

The examples are designed for parallel fusion. To fuse the images look at the left image with the left eye and the right image with the right eye from a distance of about a 50 centimeters. It may be helpful to separate the images with a card or to use an inexpensive magnifying stereopticon. These parallel fusion figures were produced using an S function (stereo.pairs) which is available via anonymous ftp from `galaxy.gmu.edu` (in subdirectory /submissions/eda).

Future Articles

Separate articles will give background for more advanced examples and provide some discussion of statistical interpretation. The next article will focus on alpha blending and with luck will include a translucent stereo rendering of overlapping contours of a density estimate for 3-D data. While many practicing statisticians do not yet have the requisite hardware, computing environments are changing very rapidly. Soon a high percentage of statisticians will be able to study 3-D structure in the calm of non-rotating images. Future articles will discuss enhancement tools for the representation and study of data.

Additional References

Many additional details are available in the literature, concerning stereoscopic resolution, hypersteroscopy, and enhancements for statistical images, common production difficulties, etc. Some useful starting references are Wegman and Carr (1992), Hodges (1992), Carr and Nicholson (1988), Papathomas and Julesz (1988), Huber (1987) and Lipton (1982). Valyus (1962) provides one of the most detailed expositions and the fascinating work of Julesz (1971) produces considerable insight into visual processing via study of optical illusions encoded in random dot stereograms.

References

Carr. D. B. and W. L. Nicholson. (1988), "EXPLOR4: A Program for Exploring Four-dimensional Data."*Dynamic Graphics for Statistics*, eds. W. S. Cleveland and M. E. McGill, pp. 309-329. Belmont, California: Wadsworth.

Carr, D. B., W. L. Nicholson, R. J. Littlefield, and D. L. Hall. (1986), "Interactive Color Display Methods for Multivariate Data." *Statistical Image Processing and Graphics*, eds. E. J. Wegman and D. J. Depriest, pp. 215-250, New York: Marcel Decker.

Carr, D. B. and R. J. Littlefield. (1983), "Color Anaglyph Stereo Scatterplots – Construction Details." *Computer Science and Statistics: Proceedings* *of the 15th Symposium on the Interface.* New York: North Holland Publishing Company.

Hodges, L. R. (1992), "Tutorial: Time-Multiplexed Stereoscopic Computer Graphics." *IEEE Computer Graphics & Applications*, pp. 20-30.

Huber, P. J. (1987), "Experiences With Three-Dimensional Scatterplots." *Journal of the American Statistical Association*, **82**(398), 448-453.

Julesz, B. (1971), *The Foundations of Cyclopean Perception.* Chicago: University of Chicago Press.

Lipton, L. (1982), *Foundations of the Stereo-Scopic Cinema, A Study in Depth.* New York: Van Nostrand Reinhold.

Newman, W. M. and R. F. Sproull. (1979), *Principles of Interactive Computer Graphics, Second Edition.* New York: McGraw-Hill.

Papathomas, T. V. and B. Julesz. (1988), "The Application of Depth Separation to the Display of Large Data Sets." *Dynamic Graphics for Statistics*, eds. W. S. Cleveland and M. E. McGill, pp. 353-377. Belmont, California: Wadsworth.

Valyus, N. A. (1962), *Stereoscopy.* New York: The Focal Press.

Wegman, E. J. and D. B. Carr. (1992), "Statistical Graphics and Visualization." Technical Report No. 84. Center for Computational Statistics, George Mason University, Fairfax, VA 22030.

Yeh, Y. Y. and Silverstein, L. D. (1990), "Limits of Fusion and Depth Judgement in Stereoscopic Color Displays." *Human Factors*, **32**, 45-60.

*Minitab is a trademark of Minitab,Inc.
S is a trademark of AT&T.

Daniel B. Carr
George Mason University
dcarr@galaxy.gmu.edu

DEPARTMENTAL COMPUTING

Not just hardware and software

A computer system has four major components—hardware, software, communications and people. The hardware and software angles of computing get discussed at length. In this column I will try to focus on issues in computing that have to do with getting a coordinated computing system going and maintained. The issues cut across types of hardware and software, operating systems and organizations. While my personal experience has been in developing a large (100+ computers) UNIX-based environment in academia, the issues surrounding departmental computing appear in business and government and across platforms. My department has networked PCs, Macs and multiple Vaxes in addition to the UNIX network. I'll try to keep the presentations as independent as I can of any particular computing platform. We'll discuss platform specifics when necessary for examples of general principles.

I confess to having several basic premises that guide my thinking about departmental systems.

- Departmental computing is the right level for support of statistical activities. Reliance on University or corporate level computing systems naturally leads to inadequate resources.
- A coordinated system in which users can send and receive mail, share files, share printers and other peripherals, share software installations, share documentation and share user support is a reasonable goal.
- Planning is possible. Despite budget uncertainties and power politics, planning can and should be done.
- Diversity happens. We can't stop it and we have to work to accept it. The best system for a fiscal person may not be the best system for a statistical scientist. Should all the statistical scientists in a department have the same platform? Let's discuss this. Let me know what you think. See the end of this column for my e-mail address.
- A network connection to the Internet is essential for proper access to the information needed to do one's job. Again, let me know what you think.
- Larger systems require full-time systems management. We'll start here (see the next section).

In future columns I hope to cover issues that hinder acceptance and development of departmental computing resources. Here are some of the issues I see as important (in no particular order):

- Acquisition of systems. Choosing platform(s), fund raising, planning, purchasing. Breaking in—how does one get started? What about solo machines?
- Administration. System management, user support.
- Maintenance (hardware, software, network gear) of modern departmental computing networks.
- Resource sharing. Coordinating funds as well as equipment.