"reasonable." The point to be made here is that each alternative eats up some menu space and potentially confuses the user by mode trapping (you never are in the mode you think you are). I am not aware of *any* data analytic problem where the first or third of the modes offers a substantial advantage over the second. What sometimes is required in addition is a *variant* of the first method, for moving one object (point cloud) as a rigid body relative to some other. In order to do this, one would have to control all six degrees of freedom of rigid body motion simultaneously (for which, incidentally, the mouse is a woefully inadequate interaction device).

As the authors stress, scatterplot matrices are nice because they provide a single integrated view of the data (Section 2.6). But I believe they understate the resolution problems, and that for all but the very smallest values of p an alternagraphic solution is preferable. For example, show only two scatterplots at one time, keeping one fixed, and use the other space to flip through all the plots in one row or column of the scatterplot matrix.

Some shorter comments.

End of Section 3.1. The distributed processor model is not only at a design disadvantage, but it also creates a software maintenance nightmare.

Section 3.3: *Integer Arithmetic*. I believe these considerations are no longer relevant after the advent of coprocessors (8087, 68881, etc.).

Section 3.4. These sad comments on graphics standards, unfortunately, are even true for semistatic graphics (where the only interaction is the all important identification of labeled observations).

Section 3.5. Windows are great if there is enough screen resolution (800 by 1000 or better) and we immediately got hooked on them with our first Apollo in 1982. But just like the proverbial *goto* in programming, extensive use of windows may actually be harmful in data analysis. Data analysis is an experimental science, and a "laboratory journal" metaphor is more appropriate than a messy "desktop," especially since the electronic version can be messed up much more thoroughly than a real one, and in much less time!

Of course, no survey can cover everything in depth. Still, because of their importance, I believe the following topics would have deserved a more thorough treatment: techniques for identifying and isolating clusters (the letter I of the original PRIM-9) and the role and use of colors.

ADDITIONAL REFERENCES

HUBER, P. J. (1987). Experiences with three-dimensional scatterplots. *J. Amer. Statist. Assoc.* **82** 448–453.

TUKEY, P. A. and TUKEY, J. W. (1981). Graphical display of data in three and higher dimensions. In *Interpreting Multivariate Data* (V. Barnett, ed.). Wiley, New York.

Comment

William F. Eddy

The authors are to be thanked for their thorough review of the current state of interactive graphics (although I think I detected a certain bias in favor of methods they and their colleagues have developed). As I started careful reading of this paper I found myself repeatedly asking: What if Option 2 instead of Option 1? It seems obvious that there is considerable work yet to be done in deciding which choices should be made. I will resist the temptation to produce a long list of such questions but rather point in other directions.

William F. Eddy is Professor of Statistics, Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213. This report was prepared while the author was a Resident Visitor at Bell Communications Research, Morristown, New Jersey.

1. STATISTICAL ROOTS

Graphical techniques have always been a part of statistics. Nevertheless, I was very struck on reading this paper that graphical statistics is currently in very much the same state that mathematical statistics was about 100 years ago. In fact I went back and reread parts of some of Karl Pearson's long series of papers on the mathematical theory of evolution that was published in the *Philosophical Transactions of the Royal Society of London* between 1894 and 1916. In the very first paper, Pearson gives a graphical method for calculating the first five moments of a probability density function.

Reading Pearson's papers reveals interesting parallels between the state of mathematical statistics then and the state of graphical statistics now:

1. There is a strong emphasis on description rather than analysis.
2. There is a tendency to modify existing tools rather than develop totally new ones.
3. There is a tremendous joy in new insights developed from the use of the techniques.
4. There is a very limited number of researchers in the area.

The first of these features is common to any very young science; we can only hope that graphical statistics will become more scientific. The second of these features is also indicative of infancy of development in the field; we can only hope that graphical statistics will grow as did mathematical statistics. The third of these features is one of the real attractions of working in a developing field; we can only hope that all statisticians will have the opportunity to use interactive graphical techniques. The fourth of these features means there is a tremendous opportunity to engage in useful research in the area, particularly the kind that can have a lasting influence; we can only hope that many people will be attracted to research in graphical statistics.

I interpret all these similarities in a very hopeful way. The heyday of mathematical statistics (that is, the time in which techniques were developed that actually affected statistical practice) was, say, the first half of this century. Given the accelerated pace of developments now I guess we should look forward to a ten year "Golden Age of Graphical Statistics."

2. DYNAGRAFIX: WHERE DO WE GO FROM HERE?

It is unreasonable to suppose that in the next five or ten years we will have enough computational power for whatever interactive graphical tasks we wish to perform. Our expectations will increase at least as fast as hardware capabilities. This means that at the heart of dynamic graphical techniques there will always be the core problem of how to get the computations done in real time. The evolution of graphic display systems obviously follows the evolution of other computer hardware. First, there were dedicated mainframes, then we had time-shared systems. Now the workstation model described in this paper is becoming ubiquitous. It seems clear that the next step is a distributed approach to interactive graphics. Anyone with a workstation that supports multiple windows and a local area network can use the poor man's distributed approach: open a window on a remote machine for the purposes of computation and display the results on the local workstation. A more serious problem occurs when the remote machine is not able to supply the needed "horsepower". This is all too common in statistics research groups where the remote machine may only be an unused workstation. In this case, it is necessary to use multiple machines and the decomposition alluded to at the end of Section 3.1 of this paper becomes a research dilemma. Decomposition of statistical calculations (not only graphical) for distributed computation is a problem that merits much wider attention.

One of the great failures of statistical research in recent years has been our inability to capture the attention of the larger scientific community. In part this is simply because most science feels it can proceed with little or no statistical input. In part also this is because much statistical research is not relevant to the larger scientific community. And in part also this is because we do not use modern "marketing" techniques for the ideas we have which are generally useful to this community. (For proof by a not-exactly-statistical example, consider the fast Fourier transform with Richard Garwin behind it. See Garwin, 1969.) Dynamic graphical techniques for the analysis of high dimensional data sets provides statistics with the opportunity to transcend this past.

The increasingly widespread use of supercomputers has created a greatly increased need for techniques of data analysis that can rapidly pick through huge quantities of output and find the "interesting" stuff. Dynamic graphical techniques obviously have great potential to assist in this task and help overcome the first two hurdles in the previous paragraph. With respect to the third (marketing), I propose, with due apologies to the late L. Ron Hubbard, that we name these techniques, collectively, DYNAGRAFIX, and we start advertising them as a replacement for AI. A few carefully designed and properly planted success stories for DYNAGRAFIX will either set AI back 50 years or trigger widespread interest in statistics, especially graphical.

ACKNOWLEDGMENT

This research was partially supported by the Office of Naval Research under Contract N00014-84-K-0588.

ADDITIONAL REFERENCE

GARWIN, R. L. (1969). The fast Fourier transform as an example of the difficulty in gaining wide use for a new technique. *IEEE Trans. Audio Electroacoust.* **AU-17** 68–72.