MEMORANDUM 28 October 1992

To: Attendees of "Data Visualization Summit", 19-20 October
From: Tim Bastian

In this memo I attempt to summarize the ideas and discussion presented at a meeting on software development for data visualization which took place in Charlottesville on 19-20 October. I have also included some followup discussion which occurred at a meeting in Socorro on 23 October. The latter meeting was attended by T. Bastian, S. Beland, E. Brinks, T. Cornwell, C. Flatters, R. Hjellming, and D. Wood. Its purpose was to discuss local development of visualization software.

The goals of the Charlottesville meeting were: 1) to introduce scientists and programmers to the hardware and software purchased or otherwise acquired by the NRAO for the purpose of visualizing and recording representations of astronomical data sets and; 2) to discuss near-term and medium-term development of data visualization software in the NRAO using available means. With regard to the first point, the meeting was successful; with regard to the second, less so.

It became clear during the course of the meeting that discussion was often hampered by the fact that most of those present did not have broad experience in the general problem of data visualization beyond what is routinely encountered in AIPS (except for E. Greisen, who had gained some experience in the rendering of 3D data through use of Sun's, now obsolete, TAAC board). More importantly, many of those present had had very little exposure to the various software packages available at the NRAO (e.g., *AVS, PV Wave, Miriad X Visualizer*). As a consequence, the issue of which of the commercial or public domain software packages, if any, would adequately serve as an environment for further development in the area of data visualization could not be addressed effectively. Obviously, a good deal more experience with these, and perhaps other, packages is required before such judgements can be made, a point I return to below.

Issues related to the question of what software package(s) should serve as the basis of development in data visualization are those of the *scope* of the development effort and the degree to which the effort should be *integrated* with *AIPS* and/or *AIPS*$^{++}$.. These issues were not discussed in any detail. From a practical standpoint it would seem reasonable to assume that development in data visualization will make contact with both *AIPS* and *AIPS*$^{++}$, the former by reason of necessity since that is where most interferometric data is currently reduced and analyzed. On the other hand, development efforts should be aware of developments in *AIPS*$^{++}$ and the development effort in data visualization should be carried forward with the idea that parts of it may eventually be integrated into *AIPS*$^{++}$. That is, development in data visualization may also serve to prototype parts of the *AIPS*$^{++}$ system.

As mentioned above, however, little progress can be made in sharpening our approach to any of these issues until both scientists and programmers become far better acquainted with the IBM RS 6000s and with the various software options available. In order to do so, several "zeroeth-order" things must be done. First, we must ensure that the relevant software is installed and that documentation is available, as pointed out by H. Liszt. Second, a small-scale effort needs to be carried out to enable users to get astronomical data into the various software environments (e.g., reliable FITS readers or other - see below) Third, scientists need to work with data that they are familiar with in one or more of the less familiar software environments in order to

assess the capabilities of a given package and to determine whether or not a large-scale development effort is warranted.

D. Wood, A. Bridle, T. Bastian, and H. Liszt each gave short presentations describing the kinds of tools they would like to see developed on the IBM RS 6000s for data visualization and analysis in the near-term. While each presentation was colored to some degree by the type of science the particular individual was involved with, there was an encouraging degree of overlap. A "wish list" has been distilled from these presentations and is presented in an appendix. I'll give the gist of its content below.

The near-term functionality desired by scientists can be loosely broken up in to four areas: data import/export, data display, data analysis, and hard copy. Several of us emphasized that the *process* of data reduction and analysis is tightly coupled to data visualization and that, ultimately, the tools we use for data visualization should be fully integrated with whatever software environment is used for data reduction and analysis. Given that the observatory is in the midst of a difficult transition from *AIPS* to AIPS^{++}, a transition which may extend over a considerable period of time, it is unrealistic to impose the demand that development efforts in the area of data visualization should be fully integrated with *AIPS* and/or *AIPS*$^{++}$. From a practical standpoint, it seems far more reasonable to require that, regardless of the commercial software package(s) we choose to exploit, data import/export should be extremely easy and flexible. Using FITS as the only medium of exchange is cumbersome and is, in any case, inadequate. It should be a trivial matter for a user to import an ascii file of data in a tabular format, an unformatted binary file, or a file with random access organization. Given the fact that in-house workstations are networked, the fact that data reduction, analysis, and visualization will occur on different machines and in different software environments need not be a substantial barrier.

Since data display and data analysis are so closely intertwined it was also generally agreed that interaction with the data had to be supported from the beginning. That is, it is rarely useful to simply view data, regardless of how elegant the rendering is – several means of interacting with the data must be available to users. These include generalized cursor readback functions, mapping coordinates onto the data, and flexible display of both image and graphical data in one or more windows. True, development of some of these tools is nontrivial. On the other hand, full support is not needed, at least in the early stages of development. For example, general support of coordinate mapping is difficult – attaching coordinates to orthogonal 2D slices in a 3D data cube is not (e.g., α and δ, α and v, or δ and v).

Finally, there is the question of hard copy. The NRAO has purchased monochrome and color postscript printers for hard copy on paper or transparencies. It has also purchased Solitaire film recorders. Support of both the postscript printers and the Solitaire seems to be progressing quite reasonably in Charlottesville. Once development in these areas has progressed further, Socorro will follow. Remaining tasks include calibrating the Solitaire (e.g., the gamma correction factor) and support of 24 bit displays (e.g. true color, RGB, or HSV displays). Support of these devices in AIPS is also intended. An area of "hard copy" which has not been developed is that of video recording, editing, and display. It appears that this can be remedied in the near-term to the extent that the necessary hardware can be purchased in fairly short order. It was left unclear, though, who would be responsible for the specifications and acquisition of a video system.

To sum up, the Charlottesville meeting drew attention to the fact that a good deal more exposure to the IBM RS 6000s and the various software options is necessary

before directed development in the area of data visualization can occur. Hence, in the very near term, we must gain the necessary exposure. It was agreed that the we should attempt to install the *MXV* package on the IBMs since it supports many of the tools which appear on the wish list. It is not clear that it would serve as a useful platform for further software development, however. It is also necessary to determine whether or not *IBM Explorer* can or will support 24 bit display. If it cannot, it is not worth looking at further. In any case, it would appear that in-house efforts in data visualization require that we first assess the capabilities of *AVS, PV Wave, MXV,* and possibly *IBM Exporer.* While contact with *AIPS* and *AIPS*$^{++}$ is necessary in the former case and desirable in the latter, the general effort in data visualization should be regarded as experimental for the time being and as being in some sense "standalone".

On 23 October there was a followup meeting in Socorro to decide how best to proceed locally. S. Beland and C. Flatters reckoned that they could devote of order 50% and 15% of their time, respectively, to the problem of data visualization. In view of the above remarks, it was decided that D. Wood and E. Brinks would work closely with S. Beland and, to a lesser degree with C. Flatters, for a period of 2–3 weeks on spectral line data in *AVS.* In addition to familiarizing themselves with general aspects of *AVS,* the intent is for Eli, Doug, and Stephane to exploit the existing capabilities of *AVS* as much as possible to analyze line data, i.e., determine which items on the "wish list" are already supported or could be by means of suitably designed networks using existing modules. Those items on the wishlist which are not supported by existing *AVS* modules and networks need to be identified.

While *PV Wave* is not available on the IBMs, it is a possible candidate for development. Before asking whether it would be worthwhile to obtain a license for *PV Wave* on the IBMs, "zeroeth order" development of the kind described above can proceed on a Sparc. T. Bastian will devote some effort to writing a convenient interface to enable users to pull AIPS MA or UV data directly into *PV Wave* data structures so that others begin to experiment with the manipulation and display of familiar data types using *PV Wave's* extensive software toolbox. It is also likely that it will soon be possible to load CL, SN, and/or BP tables into *PV Wave* for display and manipulation.

The Socorro group will meet every few weeks in the near-term in order to review progress in assessing the various software packages and, eventually, to sharpen development efforts with an eye toward fulfilling the initial goals specified on the wish list. It would also be highly desirable if interested parties in both Charlottesville and Socorro discussed these issues periodically, but perhaps less frequently than the local meetings.

APPENDIX

WISH LIST: DATA VISUALIZATION SOFTWARE DEVELOPMENT
(in no particular order)

- Support of flexible data import/export
- Support of data coordinates and values
- Support of data structures (*á la PV Wave*)
- Generalized data display
  - Display of image and graphical data in multiple windows
  - Point & click plotting capability for extracting spectra or light curves
  - Calibrated control knobs and sliders with "save" capability
  - data spreadsheet
- 24 bit display
  - True color
  - RGB & HSV displays
  - Windows with independent color palettes and TRFs
- Generalized TV animation
  - Very high rates
  - Support of "deep" cubes (e.g., 1000 planes)
  - side-by-side animation
  - animation of image plus graphical data (e.g., overlaid contours)
  - animation of 24 bit data
- Generalized cursor readback
  - readback of data coordinates & values from 2D and 3D representations
  - "hot" cursor for data tagging and flagging
  - generalized (3D) blotch and TVBOX functions
- Support of very large images
  - Fast, interactive blinking
  - Roam with full pixel resolution
- Generalized slice & dice
  - 1D slice in 2D and 3D data sets
  - 2D slice in 3D data sets
  - curvilinear slice or brick excavation
- Rendering of 3D data
  - "true" ray-tracing, rendering of geometrical data
  - rubber sheets, isosurfaces, nested isosurfaces
  - support of annotation
- General support of data geometries
  - geometry transformations (translate, rotate, non-linear stretch)
  - data coregistration (via transfer of geometry, data cross-correlation, etc.)
- Generalized statistics
- Hardcopy of screens, user-specified areas, or windows
  - color postscript printer
  - solitaire film recorder
  - video

To: Attendees of Visualization Workshop

From: R. Burns

Date: Nov. 4, 1992

Subject: Additions to Tim Bastian's Memorandum Summarizing the Visualization Workshop

Tim's note summarizes the workshop very well. The workshop made clear that we do not have adequate experience with the various packages available to make informed recommendations. Perhaps the most useful near term (6 months) activity that we can recommend is that we increase our experience on a broad front. It is important that we involve more observers in investigating these packages and that we involve more software personnel in development activities with the packages.

Although we need to get more people involved, we should do so without encouraging negative rivalries, intersite or other, and without excessive duplication of efforts. After the six month period we should review our overall direction to reach some level of consensus. In the interim it is important to keep our colleagues informed of our activities and progress.

In Charlottesville, we have had no formal meetings since the workshop; however a number of activities are taking place. We will probably use our Computer Lunch format to exchange information on the subject.

Paul Shannon has installed PV-Wave Point & Click, Version 1.61, on a trial basis for Harvey Liszt on Harvey's workstation, and has requested the upgrade to Point & Click, Version 2.0, from PVI. He plans to install the upgrade, PV-Wave Command Language, GT-Grid and NAG on the NSF server at the end of November.

During the workshop there was a question concerning the possible implementation of PV-Wave on the IBM workstations. At present the Command Language version is available for the IBM (the price is the same as for SUNs). The Point & Click version is not yet available.

Glen Langston has worked with Alan Bridle to make available MXV. Alan will investigate MXV and also give a lunch talk after he has a reasonable level of experience.

Eric and I plan to discuss the IBM Explorer package with Loyd Treinish, its designer. We hope to determine quickly whether the package supports 24 bit display and whether it makes use of our special hardware graphics board. If it doesn't, there is no need to pursue it further. If there is merit to further investigation, I will obtain from IBM a second 60 day trial/demo license.

Eric has been using AVS and intends to set up a user friendly environment for it. He also plans to work on the gamma correction problem using AIPS as an environment. Eric is going to have to spend a substantial amount of time on SUN Solaris related implications for AIPS. We are not sure how difficult this will be.

Paul Shannon is working with Geoff on AIPS++ graphics development using about 15% of his time. He is studying and experimenting with Interviews, a C++ user interface toolkit, and Unidraw, a library built upon Interviews which allows for the creation of domain-specific graphical editors. There are two goals: to acquire the skills for the eventual construction of an AIPS++ graphical user interface, (GUI), and to build a graphical editor for the manipulation and display of 2D radio astronomical data, suitable for inclusion in journal articles and other publications.