

VERY LARGE ARRAY COMPUTER MEMORANDUM NO. 175

DATA DISPLAY:

SEARCHING FOR NEW AVENUES IN IMAGE ANALYSIS

Arnold Rots

30 July 1985

DATA DISPLAY :

SEARCHING FOR NEW AVENUES IN IMAGE ANALYSIS

Arnold Rots
30 July 1985

TABLE OF CONTENTS

1.	INTRODUCTION	1
2.	OBJECTIVES FOR DATA DISPLAY	2
3.	IMAGE DISPLAY	3
3.1.	Visibility Data	3
3.2.	2-D Maps	4
3.2.1.	Zoom and Pan	4
3.2.2.	Color Schemes	4
3.2.3.	Transfer Functions	5
3.3.	Image Analysis	6
3.3.1.	Image Segments	6
3.3.2.	Map Statistics	7
3.3.3.	Comparison of Maps	7
3.3.4.	Arithmetic Operations	7
3.3.5.	Intensity-Hue Display	8
3.3.6.	2-D Convolutions	8
3.4.	3-D Maps	10
3.4.1.	2-D Representations	10
3.4.2.	3-D Solids	11
3.4.3.	Stereoscopic Images	12
3.4.4.	Animation	13
3.4.5.	Holograms	14
3.5.	Specialized Hardware	15
3.5.1.	Control Panel	15
3.5.2.	Hardcopy	16
3.5.3.	Digital Image Storage	16
3.5.4.	Video Recording	17
3.5.5.	3-D Solids Machine	17
4.	GRAPHICS DISPLAY	18
4.1.	2-D Displays	19
4.2.	3-D Displays	20
4.3.	Device Independence	21
5.	WORKING ENVIRONMENT AND SUPPORT FUNCTIONS	23
5.1.	Work Stations	23
5.2.	Image Recording Facilities	24
5.3.	Photo/Graphics Facilities	24
6.	CONCLUSION AND RECOMMENDATIONS	25

1. INTRODUCTION

Display of data is an immensely important area in data processing since it functions as the prime interface between the user and his data as the latter are being processed. Yet, most people are relatively unfamiliar with its potential, usually because the necessary manpower is not available to implement sophisticated software for display devices so that most of the capabilities remain hidden for those working on data analysis software as well as for the users. For instance, the IIS image display systems used in GIPSY and AIPS are quite powerful, not only in rendering very flexible image display, but also in performing certain simple image analysis functions. Full use of those capabilities would make data processing with these systems easier for the users, while it also could reduce part of the burden on the host computer. In neither system, however, are those capabilities fully realized; the IIS systems are basically used as frame buffers for the TV monitors with rudimentary slope-and-intercept transfer function control. Also, from users' responses to more advanced features available in only one of these systems, it has become clear that users deem such features essential only after they have gained experience with them. This is true for sophisticated display techniques as well as for image analysis functions which provide interactive "quick-and-dirty" preview of equivalent functions available in the host computer; when done in the host they take longer and require more resources, thereby often precluding extensive experimentation.

Of particular concern is the processing of three-dimensional data, where users have so far been forced to interpret their observations from displays that are essentially mere modifications of designs made for the two-dimensional case. Some research has been done in astronomy on display of three-dimensional data. This has only scratched the surface so far, although it has revealed some enticing vistas.

These are the matters I would like to address in this memo in the hope to raise people's awareness of the opportunities that are available right now and the possibilities that may become feasible through a concerted effort in research, design, and hard coding work. One should not imagine that the display tools come for free. It will need more manpower than traditionally has been invested in this area; but it will be warranted by the enhanced power of the data analysis system as a whole.

In terms of hardware and applications there are two types of data display: graphics (line drawing) and image (grey-scale) display. Both may be in monochrome or in color. In the next section I will try to define the objectives for the design of display systems. Then, in the following two sections I shall deal with the two display device types. Finally, there will be a discussion of off-line support functions and a concluding section.

2. OBJECTIVES FOR DATA DISPLAY

The data reduction and analysis process consists of a string of reduction and analysis functions applied to the data. At each point in this sequence the astronomer has to decide on subsequent processing, based on his assessment of the data and his own objectives. He also wants to understand his data. The main objectives of data display are to facilitate this by giving the astronomer an instant overview of the contents of his data, as well as enabling him to lift out details. These functions should be performed fast, interactively, and under easy control of the user.

This may not be the most complete definition of data display objectives, but it is certainly a valid one and, in my opinion, quite adequate in the context of this memo.

The objectives may also sound rather trivial. One should be aware of the implications, however. "Instant overview of the contents of the data" means considerably more than just "a picture of the data"; it requires precisely such a display that the full contents can be grasped instantaneously, which is by no means a trivial requirement. Also implied is a good deal of capability to "play" with the data. Ideally, only "good" intermediate results should be calculated and stored by the host computer. Tools to determine the correct parameters and processes to achieve this should be provided, at least in part, by the display system. In some cases flexible display is sufficient to make a sound decision; in other cases one needs interactive "quick-and-dirty" versions of analysis functions present in the display device for preview.

Easy control of display functions similarly has far-reaching consequences: language - and thus interaction through a character keyboard - is a poor medium to control displays. Like theatre lighting systems, displays are much more easily controlled by analog devices such as switches, knobs, slides, and buttons.

3. IMAGE DISPLAY

For aperture synthesis observations there are three types of data to be displayed, each with their own specific requirements: visibility data, two-dimensional maps, and three-dimensional maps (predominantly spectral line data). There are two elements in the application of display devices to the user's data: the pure display function which tries to translate as many bits as possible into a comprehensible image, and image analysis which allows the user to manipulate the bits interactively in order to gain an even better understanding. It will be clear that the distinction between the two is not always sharp. Finally there is some specialized hardware to be considered that is extremely helpful for the enhancement, the efficiency, and the ease of use of display systems.

3.1. Visibility Data

Aperture synthesis visibility data consist of complex numbers as a function of two coordinates: base line and time, or u and v . Image display of these data forms a powerful tool in editing and calibration of the observations since discordant data values are easily recognizable when arranged in such a way that large-scale patterns are to be expected. Depending on the nature of the observations and the deficiencies to be detected, either (base line, time) or (u,v) display may be preferable.

For detection of bad receiver behavior, for instance, a (base line, time) display is the most revealing. Care should be taken that the base lines are arranged such that patterns caused by source structure can easily be recognized. This probably means for the VLA that intra-arm base lines for each arm should be grouped together, while inter-arm base lines need careful sequencing.

The most generally useful way to represent complex quantities is through an intensity-hue display (see also section 3.3.5.); here amplitude is represented by the intensity of the pixel, phase by the color (hue). A cyclic color scheme (blue-cyan-green-yellow-red-magenta-blue) is most appropriate for this purpose since then there will be no discontinuity at phases of $+180/-180$ or $0/360$. The display system should allow the user to independently vary the transfer functions for amplitude and phase in order to suit the particular range of values he is interested in at any one moment. This is an important requirement that not only determines the versatility and usefulness of this display, but also restricts the choice of acceptable models of image display machines since most models are incapable of supporting this feature.

Interactive support functions should include: dynamic amplitude, phase and antenna pair display for pixels selected under cursor control; and flagging (editing) of pixels under cursor control.

3.2. 2-D Maps

Two-dimensional maps are the traditional realm of image display systems. We assume that we are dealing with astronomical maps containing some variable (e.g., brightness distribution) as a function of two coordinates - either two spatial coordinates, or one spatial coordinate and velocity. In this section we shall restrict ourselves to the techniques that optimize the display of such a single image.

Map sizes will vary considerably, most commonly from 64x64 to 2048x2048. Image storage for at least 1024x1024 should be provided. TV monitors that can display a pixel grid of this size are available, although they are not (yet) common. One also ought to consider whether anything is to be gained by switching to this size monitor from the standard 512x512 displays, considering the amount of detail the eye can take in.

Some display systems (e.g., Vicom) will allow optimal use of memory space because they are capable of handling arbitrary image sizes which are not tied to fixed positions in memory.

3.2.1. Zoom and Pan

Zoom and pan are very essential features, both for lifting out details and for accurate positioning of the cursor (down to the pixel level). Most systems provide zoom factors of 1, 2, 4, and 8; zooming is usually accomplished by pixel replication. Ideally, one would want any integer zoom factor, which some systems provide, a negative zoom (to display 1024x1024 images on a 512x512 screen), which is extremely rare, and some sort of interpolation instead of pixel replication, which is not available in hardware. In general, however, these wish list items either are non-essential niceties or can be emulated in other ways.

If the monitor screen is no bigger than 512x512 one wants at least a roam feature that allows interactive roaming through a 1024x1024 image.

Some systems have nasty definitions built into them. The IIS, for instance, has a very inconvenient definition of the zoom center.

3.2.2. Color Schemes

Although a gray scale representation is often quite adequate and usually provides more dynamic range than most color schemes, pseudo color is a useful option to have. Schemes presently in use include spectral colors and a variety of discrete (contrasting) color palettes. The latter options are often referred to as color contours and include the capability to compress or stretch (recycle) the colors which is especially useful for optical data.

There definitely is a need for multiple color schemes; choice of a particular scheme depends on the data, the user's taste, and the user's physiology (such as partial color-blindness).

It is highly desirable to be able to decouple the definition of the color scheme from the transfer function (see section 3.2.3.). This requires at least two stages of look-up tables as IIS provides. Some other manufacturers provide this capability with a special color table, which is less flexible.

3.2.3. Transfer Functions

Transfer functions form the heart of image display, in a sense. The question is how to transform the range of intensities present in an image - or, the part of that range that the user is currently interested in - in such a way that the maximum amount of information can be perceived by the human eye. There are two parts to solving this problem. First, a transformation has to be performed on the image intensities as they are kept on a storage device to compress them into 8 bits; this is usually not a problem when the dynamic range of the image is less than about a factor of 100. Then, the 8 bit values in the display device have to be transformed into intensities on the monitor. On the whole, the approaches to both transformations are similar, although the requirements at any given moment may be different; e.g., the user may wish to load as much of the image intensity range into the display memory as possible in order to look at a number of sub-ranges in detail.

There are two classes of requirements for the transfer functions. One is to use the available viewing (or display memory) dynamic range as efficiently as possible for the entire range of intensities. How this is done depends on the data. For images that are buried in the noise, for instance, a straight linear transfer function (1:1) may be used, while high-dynamic range images (with relatively few points near the high end of the intensity range) usually profit from a logarithmic or a histogram-equalized transfer function. Especially histogram equalization is a powerful method that allocates the available output intensities as efficiently as possible by distributing them over the input intensities on the basis of the distribution of those input intensities. If the image device is equipped with a 16 bit mode and a histogram generation option, the bulk of the work for this technique can be done on the display device which significantly unburdens the host by halving the number of required I/O operations on the image. Whether used in loading the image or in displaying it, histogram equalization has the effect of maximizing the amount of detail one can see in a high dynamic range image, thus enabling the user to grasp as much of the contents of an image as possible from one picture. It is useful to have several options available for the equalizing algorithm since not all images can be handled the same way. Another important feature is the possibility to interactively vary the range in the histogram to which the equalization algorithm is being applied.

The second class is to lift out part of the input intensity range to look at it in great detail; it means concentrating the bulk of the output intensity range on a more or less small part of the input range. Traditionally, this often has been done by slope-and-intercept transfer functions. They have the advantage that they are very easy to control interactively since there are only two input parameters with functions that are easily understood by the user. However, they are also rather crude and do not allow designating a small portion of the output range to the remainder of the input range. Two-kink / three-segment transfer functions are far superior in that respect, but not perfect yet and control is more involved. It should be possible, though, to develop a family of transfer functions that is better than both of the previous ones in either respect.

Often, the quality of a picture can be improved even more by combining the use of sophisticated transfer functions with two-dimensional filtering functions. These will be discussed in section 3.3.6.

3.3. Image Analysis

Modern image display devices usually incorporate a certain amount of processing capability that allows the user to perform a number of different basic operations on entire, or parts of, images in a very short time (typically one refresh cycle per operation). By stringing several basic operations together more sophisticated processes can be performed. There are two advantages to making use of these facilities: doing them in the image display device relieves the host computer from these tasks and the display machines can perform them much faster (essentially in an array processor fashion) with instant display, making it feasible to do it interactively. The disadvantage is that all operations are performed with limited accuracy (8 bits), although in certain cases one may be able to use 16 bit options, provided there is enough refresh memory. There are five areas of application:

- The case where a quick-and-dirty operation is sufficient for the user's purpose. An example is when the user wants to see the difference between two maps and a limited accuracy subtraction in 8 bits is enough to show what he wants to see. Intensity-hue display essentially also falls in this category.

- Enhanced image display such as histogram equalization, intensity-hue display and the use of two-dimensional filtering for display purposes.

- The case where the user wants to handle the data interactively in order to make decisions on further processing. An example in this category is using a two-dimensional smoothing implemented in the display device (fast, but with limited accuracy) in order to determine the optimal parameters for a proper smoothing on the original data.

- Image statistics. Image display devices that have blotch and histogram functions offer the capability of performing image statistics (average, sum, r.m.s., median, histogram, etc.) of selected areas of an image almost instantaneously and usually with enough accuracy. The retrieval of single pixel intensity values is a special case of this.

- Definition of image segments, regions of interest, or blotches. These can be used for blanking or differential treatment either in further processing or for display functions.

In the following I shall briefly describe some analysis tools, not according to the applications outlined above, but rather according to the techniques used.

3.3.1. Image Segments

Defining different segments in an image is often needed for a variety of applications: outlining the area in which one believes the source is contained or, instead, defining an area in which one believes there to be no radiation, for such purposes as "Clean boxes", spectral line windowing, image statistics, and different application of transfer functions. The most natural way to define these areas is by drawing polygons in the image itself as it appears on the TV monitor, with the aid of a pointing device (light pen, tablet, track ball) and outlining them in a graphics overlay plane. This is especially powerful if the image display device offers the capability of distinguishing between the "blanked" and "non-blanked" regions in performing statistics and arithmetic and logical operations.

3.3.2. Map Statistics

Map statistics of selected areas in an image provide essential information on the quality of the data and form an indispensable supplement to the image itself. The statistics include: average, median, sum, standard deviation, and histograms of intensity distributions. If the image display device is equipped with a histogram generator which can operate on the entire image, in the region of interest (see section 3.3.1.) or outside that region, the statistics can be obtained very fast and usually with enough accuracy. Increased accuracy can be obtained by reading the map again or if 16 bit mode is available.

Histograms are also important for fast processing of histogram equalization algorithms (see section 3.2.3.).

3.3.3. Comparison of Maps

In judging the results of image processing techniques the user often wants to compare two maps. The technique that has a long tradition in astronomy for this purpose is the blinking of the two images. This can easily be accomplished in a display device by loading the maps into different image planes and then switching the monitor at some rate, interactively controlled by the user, between the two images. This requires the two images to have very similar appearance which can be done by fiddling with the transfer functions while both are shown in a split screen mode.

Another technique is to put both images on in split screen mode and give the user interactive control over the position of the split, so that he can move it back and forth over the area of interest. Although it depends on the personal preference of the user, this technique is on the whole more effective than blinking.

Finally, subtracting the two maps may also provide a very effective means of comparison. This can be done by subtracting the original maps in the host computer and displaying the difference in the display device, or by performing the subtraction in the display device itself, which is infinitely faster but has limited accuracy (see section 3.3.4.).

3.3.4. Arithmetic Operations

Arithmetic operations between maps are very common. Adding (or averaging) two maps and subtracting two maps from each other (either for comparison or subtracting the continuum from spectral line data) are obvious, but multiplication and division also occur. Image display devices can perform such operations very fast, by using an Arithmetic and Logical Unit in conjunction with a Feedback Unit, and/or the Look Up Tables. When two image planes are switched on their contents (as modified by the Look Up Table) are added. Addition can be achieved by applying two linear transfer functions. Subtraction by a positive and a negative linear transfer function. Multiplication and division can be accomplished by applying logarithmic transfer functions to the individual images and an exponential transfer function to the sum; this, by the way, requires separate Look Up Tables for each image plane, as well as for the sum (see also section 3.3.5.).

Having these functions available in the display device is useful for a quick-and-dirty preview and for cases where the limited accuracy is acceptable and the result does not have to be kept.

3.3.5. Intensity-Hue Display

Intensity-hue display forms an excellent tool for situations where the user wants to view two parameters simultaneously in two-dimensional space, especially when one of them carries intensity information and implies a credibility criterion for the other. Examples are: column density and velocity (spectral line), percentage polarization and polarization angle, amplitude and phase (visibility data), total intensity and percentage polarization, continuum flux density and optical depth, and flux density and spectral index. The former of the parameters in these pairs controls the intensity of the image, the latter the color (hue).

It will be clear that human beings in general think in terms of intensity, hue, and saturation (often the last two are combined into the household term "color"), rather than prime color mixes. Although there is nothing particularly difficult in making image display devices follow this natural trend, oddly enough, the industry is still in the Middle Ages where the architecture of the devices (with one exception) is dictated by the (TV monitor) hardware; henceforth, the programmer has to think in Red, Green, and Blue. Unfortunately, since image display machine designers seem to be so far removed from reality, in most of these machines it turns out to be impossible to emulate intensity and hue display. IIS is one of the very few that can do it, for the same reason why it can multiply two maps: it has three separate Look Up Tables for each image plane as well as for the red, green, and blue sums.

One should be aware that the range of fully saturated colors generally used in intensity-hue display does not offer a large dynamic range to the eye. It is therefore mandatory that the transfer functions mapping the pixel intensities into screen intensity and color can be controlled interactively and independently by the user. The choice of colors deserves some consideration and depends on the data being displayed. Angular units (phase, polarization angle) usually benefit from a cyclic color scheme, since they are cyclic themselves (see section 3.1.). For velocity the jargon (redshift-blueshift) suggests a spectral color sequence, while for the other parameters either a spectral sequence, or a sequence going from "dark" to "light" colors is the most appropriate.

It should be emphasized that intensity-hue display is primarily an analysis tool which enables the user to recognize or discount certain features in his data quickly. When looking at the two parameter maps separately, he might have missed this interpretation, or only arrived at it with considerably more effort.

3.3.6. 2-D Convolutions

Two-dimensional convolutions (or filtering) have two basic applications in image analysis: for changing the image resolution (smoothing) and for image enhancement. The user may want to smooth (degrade the resolution of) his map in order to increase the signal-to-noise ratio for weak extended features, to make a strongly elliptical "beam" rounder, or to bring his map at the same resolution as other data for comparison; this last case may involve observations at different frequencies or observations made with another instrument. Two-dimensional filtering for image enhancement purposes usually involves combining original and/or smoothed and/or edge-enhanced images. It is

interesting to see that these two applications can functionally be lumped together. They are both essentially image enhancement operations: one enhances the image through a low-pass filter bringing out extended features buried in the noise, the other through a high-pass filter accentuating the small scale details.

Although convolutions are conceptually simple operations, they do require a fair amount of resources when done in the host computer and can, in that mode, not be done at interactive speeds. Image display devices with an ALU/Feedback option can perform the basic operations needed for convolutions (multiplication, translation, and addition of entire images) at very high speed (burst speeds of 15 million operations per second) and have the added advantage that the result is immediately available in the display device for viewing. Admittedly, the precision is limited since one usually deals with 8-bit images, but in general 16 bits can be used for accumulation and possibly even for the entire operation if enough memory is available.

Even so, the result of such a limited precision convolution may be able to tell the user what he wants to know, or at least give him the information needed on optimal parameter choice to initiate a "proper" convolution operation in the host. Experience has shown that users, due to the lack of interactive capability of this sort, tend to experiment very little with convolution parameters, even though obvious improvements could be made. Rough timing estimates provide the reason for this: starting a convolution operation in the host and displaying the image on the monitor will typically take at least 3 to 5 minutes of real time, while the image display device can perform the same task in less than 15 seconds.

The field of two-dimensional filtering for image enhancement has hardly been explored in radio astronomy, although users are increasingly becoming aware of the usefulness of high quality image display. It is not quite clear at the moment whether final quality display involving filtering can be handled with the limited precision of image display devices - although it may -, but it certainly would be useful to provide the option as part of the standard interactive image enhancement tools available during the data analysis process. An experimental program, MFILTR, doing this kind of two-dimensional filtering has been implemented on the PDP-11/44 computer DISPLY and its IIS image display device at the VLA site.

3.4. 3-D Maps

An especially massive display and data interpretation problem in radio astronomy is posed by aperture synthesis spectral line observations. In this case the user ends up with a three-dimensional image where the coordinate axes are formed by two spatial coordinates on the sky and frequency (or Doppler velocity). For the interpretation of such data it is imperative that the user has a means of forming at least a mental picture of the brightness distribution in all three dimensions simultaneously; a clear understanding of the three-dimensional structure and continuity of the object(s) is necessary for intelligent analysis of the data. So far, this mental picture has usually been built up from a large number of displays of two-dimensional cross-cuts through the data cube, but obviously that assembly process could be made much faster and much more efficient if the data could be displayed in three dimensions directly. The following sections will deal with some attempts at implementing such display tools. From this introduction it will be clear that the main objective is to provide tools for use during the analysis process, not the generation of fancy final presentation displays.

3.4.1. 2-D Representations

Because of the nature and geometry of common display materials (paper, TV monitor screens) two-dimensional representations will remain important. But in addition there is the fact that the three-dimensional displays that we can presently conceive of are not easily quantifiable. Hence, there is a need for two-dimensional displays, both as a substitute and for work copies to look at quantitative detail after the user has gained a three-dimensional understanding of the total contents of the data.

A mosaic of two-dimensional cross cuts is very useful for the latter purpose. One may think here of contour maps in the spatial coordinates, one at each observed velocity, or in velocity and one spatial coordinate along parallel lines. It is in general useful if the display program can automatically compose such displays in a mosaic directly (like, for instance, the AIPS task KONTR).

Another application of a stack of two-dimensional (gray-scale) maps, covering the entire cube, is to put them into a time sequence (animation) and to "travel" through the cube in time. An example of such a program is the AIPS verb TVMOVIE. This technique is capable of aiding in the detection of three-dimensional continuity in a rather powerful way. Its draw-backs are that it requires the user to mentally accumulate what he sees over time in order to build up the three-dimensional picture in his mind, and that the display cannot be captured in a hardcopy to be studied at leisure. However, the latter disadvantage may be overcome by recording the scene on video tape and playing it over and over again on a cheap viewing station. The main problem in creating these displays is fast loading of the individual frames so that a realistic animation effect can be achieved; normal loading of the data from disk into the image display device is far too slow. There are three possible solutions. First, one can load each frame and transfer it to video tape individually. This requires a video recorder with editing capability (in order not to lose the sync) and is rather tedious since the whole process has to be repeated for every loop

one wants to record. Secondly, each frame can be loaded and transferred to a specialized image storage device, from where the sequence can be retrieved at high speed (and, if so desired, recorded on video tape in real time); examples of such a device are the video disk used with GIPSY in Groningen and the image storage device being developed by NRAO. Finally, one can use the image display device itself, preloading all frames in sections of the refresh memories and displaying them in rapid sequence. An image display device like VICOM is completely flexible in its allocation of refresh memory and allows this type of operation easily. In a device like an IIS one can display a zoomed image and rapidly change the zoom center; naturally, the individual frames are then restricted to 512x512, 256x256, 128x128, or 64x64. In using this technique one obviously has to contend with a trade-off between the size of the individual frames and the number of frames in the time sequence, the product being limited by the amount of available refresh memory. But on the other hand, it provides much faster and easier access to time sequences than the other two methods and is therefore a useful (quick-and-dirty) option, even when any of the others are available.

There is one more application of two-dimensional displays to three-dimensional data: the use of false color. If one has a series of two-dimensional maps at different velocities, each velocity could be assigned a slightly different color in, say, a spectral sequence. Although powerful for some data, there are severe draw-backs; it is especially tedious to change the color transfer function (requiring a complete rebuild of the image, reading through all the data again). In practice, the display is usually limited to an image built from a two-parameter representation of the profile (the intensity as a function of velocity at each spatial pixel): zeroth and first moment. One can then let the former map control the intensity, the latter the hue, while it is possible to vary the transfer functions of both maps (like in the AIPS verb TVHUEINT, developed by Jim Torson). See section 3.3.5. for details. A disadvantage is obviously that one has lost all information on the exact shape of the profiles and that the human eye does not have a tremendous resolution along the saturated colors of the spectral sequence.

3.4.2. 3-D Solids

What one ideally would like to do is put the contents of a data cube in space and view it from various angles. To do this one has to make a three-dimensional object out of the data, or a 3-D solid surface. Such an object can be constructed by wrapping a surface through the cube at a particular threshold intensity: everything inside the surface has a higher intensity than the threshold, everything outside a lower intensity. Basically this is the true three-dimensional equivalent of a two-dimensional contour map with a single contour level. This three-dimensional solid surface object can then be painted on the TV monitor screen as seen from any viewing position, including ambient light, direct lighting, perspective, and depth cueing (diminishing intensity at larger distances). In principle one could also add other features, like surface texture, reflectivity, true shading, and even transparency, but such sophistications do in general add so much overhead that they are not really warranted.

Increasing the threshold value would have the effect of peeling the onion skins off, getting closer and closer to the heart of the object. Objects can be cut open in which case it may be advantageous to color-

code the intensities seen in its bowels. It is to be recommended to add a colored back drop in order to be able to discriminate between dark (shaded) parts on the object and holes in it.

Even though this is one of the better techniques to show three-dimensional structure, there are draw-backs inherent in these displays. The problem is that when we see a (two-dimensional) picture of a familiar three-dimensional object, such as a house, a chair, an animal, we can in our minds reconstruct the three-dimensional structure because of the familiarity: we have walked around these objects and know how they are put together. This is not true for most astronomical objects, especially not in the strange (right ascension, declination, line-of-sight velocity) phase space. The images discussed in this section do therefore not convey enough information to understand the exact three-dimensional shape of the objects or to even decide (in the case of several unconnected structures) what is in front and what is in the back. Solutions to this problem are discussed in the next sections.

3.4.3. Stereoscopic Images

An obvious technique to turn one's attention to for the display of three-dimensional data is that of stereoscopic images. There are two ways of displaying the data and two techniques for viewing them.

The first mode of display is just to stack the images along the third axis in two pictures, with an offset depending on the position along that axis. This essentially produces a transparent object. More experimentation is needed with this form of display, but one of the disadvantages is confusion between foreground and background features.

The second mode is to produce a stereoscopic display of the three-dimensional solids discussed in the previous section. This is done by generating two views with viewing angles slightly different in azimuth; this difference should typically be in the range 5 to 15 degrees.

The best way to view a stereo-pair is through a stereoscope. For this purpose the two images can be put on the TV monitor and hard copies can be made which are then placed in the stereoscope; if necessary, the images may be reduced on a copying machine.

A fast way of viewing stereo-pairs is the use of anaglyphs: one image is put on the TV monitor in red, the other in green and the result is viewed through red-and-green "stereo-glasses". In positive, the right hand image is red, in negative it is green. Care has to be taken that the images do not "cross" too much in the back; they may have to be slightly offset in horizontal direction. In general, the TV screen phosphors are well matched to the lenses in these glasses. This mode of viewing is also useful for audience presentations where the use of stereoscopes is technically impossible. A more satisfactory technique is to use polarizers instead of red-and-green, but this involves more sophisticated machinery. For audience presentations, two projectors (each with the image for one eye) have to be used with orthogonal polarizers in front of the lens. One has to project onto a non-depolarizing screen and the audience has to be equipped with corresponding polarizing glasses; usually, linear polarizers (at 45 and 135 degrees) are being used, but circular polarization has great advantages because it makes the mounting of the polarizers less critical and allows more tilting of the viewers' heads. On a TV monitor something similar can be done. Stereographics Corporation will sell a complete system (including the monitor) for about \$20,000, but it may be possible to do something simpler: one fills the odd lines of the image

with one view and the even ones with the other; a circular polarizer and a variable retarder are placed in front of the screen; the variable retarder either lets light through unchanged or changes the sense of the polarization and is triggered by the vertical retrace; the user, finally, is equipped with circularly polarized glasses.

One thing one has to be aware of, however, is that the effectiveness of these techniques is highly subjective: some people simply do not have stereopsis - possibly because of some physiological defect -, some people are very good at it, and most people can "sort of see it", but could be good at it with some practice. Similarly, positive stereo-pairs work better for some people, while others prefer negatives; this may also depend on the objects, and may in particular be different for gray-scale and graphics images.

Finally, there are some three-dimensional imaging devices on the market now; one of these works with a vibrating mirror. However, it appears that the amount of data that such devices can handle is, at least at the moment, insufficient for our purposes.

3.4.4. Animation

By far the most effective way to show the three-dimensional structure of objects is to move them; for instance, by rotating the 3-D solids described in section 3.4.2. Unfortunately, this is not always easy. For one thing, one has to construct a fairly large number of frames; one needs intervals no larger than 5 degrees and preferably half that. Then, the animation has to be effected; this can be done in four ways: using the zoom-and-pan of the image display device, using an image storage device with a high data rate to the display device, or making an actual animation movie, on film or video tape.

The zoom-and-pan option has been described extensively in section 3.4.1. One should keep in mind, though, that the limited number of frames that can be displayed is probably a rather severe restriction for this application, since one is likely to deal with frames no smaller than 256x256. Nevertheless, it may be sufficient for a back and forth "rocking" of the object which may be satisfactory.

The image storage device will be dealt with in section 3.5.3. It may be useful to point out here that to obtain a smooth animation (which this application requires) one needs to run at frame speeds of at least 10 Hz which may not be achievable with digital image storage devices.

If one rejects the options above, one is left with making an animation movie, for which, incidentally, an image storage device still comes in handy. The problems involved in recording on video tape will be discussed in section 3.5.4. The recording of animation movies on film is something we do have some experience with. Ideally, one would like to generate these on a high quality film recorder, such as a Dicomed film recorder. At the same time, though, one would like to use 16-mm film because processing, editing, and projection facilities are usually readily available. This can be done adequately on a Dicomed Model D-48 recorder; on a Model D-47, such as we have, it is not feasible. Alternatively, one can shoot a movie directly off a TV monitor screen. This requires some control software and equipment and one has to be aware that it must be done in single frame mode with exposure times of at least 1 second. This means that movie production, although not difficult, is a lengthy process - of the order of half a minute to one minute of recording time per second screen time. Although the medium itself (film) is obviously more expensive than video - at

about \$10 per minute, original or copy - the equipment is cheap now that everybody is switching to video: used 16-mm cameras can be had for a few hundred dollars. Nevertheless, because of the hassle of set-up and processing, it does not lend itself easily for quick interactive work. At the VLA we have just been lucky to have 1-3 day turn-around available.

A new possibility that has recently emerged is the development at the University of Pennsylvania Hospital of a 3-D Solids machine that will provide the user with interactive control over the viewing of a three-dimensional solid surface (see section 3.5.5.), thus enabling real time animation.

3.4.5. Holograms

While thinking about display of three-dimensional data, one inevitably turns one's thoughts to holograms. It would especially be useful if production of white light holograms were easy and could be automated. However, the state of the art is currently not such that holograms are a practical medium for display in astronomy.

3.5. Specialized Hardware

There are some pieces of specialized hardware that are eminently suited to make life easier in interactive image display. They should be considered integral parts of the image display stations since they allow the user to realize the full capabilities of image display devices.

3.5.1. Control Panel

Controlling the functions of an image display is in the mind of the user essentially an analog operation. The use of a keyboard for interaction with the image is experienced as an extremely inconvenient intrusion of this cumbersome device. Manufacturers of display devices have realized this - to a certain extent - and provided trackballs, mice, light pens, tablets, push buttons, etc. They may be used to move cursors around and control various functions, like zoom, split screens, transfer functions, etc. Nevertheless, one has to resort to menus and multiple button definitions (not to mention multiple cursor definitions) in order to cover all control functions. And even then the system will at the very least be cumbersome to use, and probably confusing.

It therefore makes more sense to put the control of the image display device almost entirely on a hardware panel that looks analog to the user. The following sections could be envisaged to be part of this panel. A module that controls which image planes are switched on. A number of transfer function modules that control transfer functions through rotating knobs or slides with push button selection of the image planes they control. A zoom module. An image plane cum split screen control (including positioning of the split). A blink module with control of the planes involved and of the blink rate. A graphics plane control module. A color scheme control module. And any others one may fancy. Not only gives such a panel easy - and easily understandable - control of the display, but it also provides an instantaneous status display. Basically, all this falls in the realm of ergonomics and care should be taken that things are "right": the organization of the panel, the "feel" of the knobs, the slant of the panel, as well as placement of the monitor and lighting.

In principle the control panel could communicate with the image display device through the host, but it would make life easier if the panel could communicate with the device directly; this would also take a burden off the host. It would be necessary to build a small processor into the control panel.

If the display device cursor is used for any functions it is helpful for the user to adopt different cursor colors to distinguish between these functions and to store text and/or numeric information in the cursor array (assuming that the display device has such a feature). The advantage of the latter technique is that this information is not zoomed with the rest of the image (as it would be when displayed through the graphics planes of IIS Model 70 machines) and that the information is displayed right at the center of the user's focus of attention. These techniques have been employed in GIPSY and in the program MFILTR (see section 3.3.6.).

3.5.2. Hardcopy

The capability of obtaining hard copies of the screen is a tool the user cannot do without. There are two types of hard copies, each with its own purpose and requirements.

The first will mainly be used for reference and as work copies. The main requirements here are: fast generation (less than a minute), reasonable size (at least 8x8 inches), and a material that can be written on. Black-and-white is sufficient for these purposes and one does not need a tremendous dynamic range in gray scale intensity. The Honeywell hardcopy devices used at the VLA site are an excellent illustration of the kind of thing needed in this respect.

The second need for hardcopy is for pictures that can be shown to others to present the data. High quality in dynamic range, color, accuracy, and geometry is required here. Two forms of hardcopy are needed: slides (transparencies) and prints. Movie capability (film and/or video) could be added. Accurate film recorders and easy-to-use software are needed, as well as photographic facilities. The latter can in principle be fairly simple but it probably would make sense to combine them with other, more sophisticated photographic needs such as producing overlays. Turn-around time can be longer (up to a few days).

3.5.3. Digital Image Storage

Experience has shown that there is a very important need for a capability for storage of images. There are two categories: just storing an image that has been fine-tuned for later retrieval (for comparison, display, or recording) and storing a sequence of images that can be run off as a time sequence. This latter function comes about because conventional loading of images is not fast enough to produce a realistic time sequence. The requirements for an image storage device are: fast loading (in both directions), flexible control, capacity of several hundred images, archival facility, and, preferably, a digital format.

Fast loading enables realistic time sequences and avoids irritation at having to wait for the device to do its job. Flexible control (of especially the time sequences) can be effected through the control panel described in section 3.5.1. Digital format will allow the user to fiddle with the image after retrieval and makes the images compatible with any other images loaded from the host; this includes the capability (if adequate header information is stored in the device) of retrieving individual pixel intensity information. It also should produce a higher quality image. Archival storage, for instance by being able to dump or load the image storage device to or from cassette tape, enables users to pick up again at a later session without being bothered by intervening users; it also allows the user to build up a library of time sequences.

Such a device does assume, of course, that the image display device can be "dual ported", i.e., that it can, in some way, be loaded from the host as well as from an external device. This requirement is not strictly necessary for the control panel, but does make things easier and faster for it; if the capability is there, it does make sense to integrate the storage device and the control panel.

3.5.4. Video Recording

The image display device puts images on a TV monitor, so what would be more logical than to record these images on video tape for later display and/or the generation of time sequences? Unfortunately, things are not that simple.

To generate time sequences, frame by frame, requires a video tape format with frame encoding. The most likely candidate is U-matic 0.75 inch tape, not the most popular VHS 0.5 inch tape format. This is true for any sequence which is not recorded in one shot and where one does want to avoid irritating flicker. To be able to view such recordings on the garden variety VCRs then requires a copying capability to VHS tape.

The signal coming out of display devices is R-G-B-sync (four cables). In order to enable the user to feed this into a VCR one has to turn it into composite video, which is not a big deal, but still a \$3500 box.

Our TV monitors display 512 lines. Standard (American - NTSC) TV only displays 480 lines. So one will lose the bottom 32 lines, or one-sixteenth of the image. Our European colleagues are better off: they only get a black band underneath the image.

Our TV monitors run in "underscan" mode: to achieve a 1:1 aspect ratio, the TV beam only makes a partial horizontal scan. When displayed on a regular TV set, the image will be distorted: stretched in horizontal direction by a factor 1.33. Our European colleagues are even worse off. Of course, it should be possible to squeeze the image horizontally to achieve a satisfactory aspect ratio, but to my knowledge boxes performing just that service are not available commercially; they are available as part of sophisticated (and expensive) studio equipment that will do much more than we will ever need. Alternatively, the squeezing could be done in software, before the images are loaded, but that is not really satisfactory either, since it violates the rule "what you see is what you get".

As one can see, there are some problems with video recording. They can be solved, but it costs money (about \$10,000 to \$20,000). Movie making (on 16 mm film) is actually cheaper as far as equipment is concerned. Nevertheless, I think the capability should be developed. In the first place, it is in principle capable of producing canned displays with a very short turn-around time. Secondly, if we can record display sequences on video tape, the user can be provided with a cheap (\$1000) work station to view his data ad nauseam without having to hog the image display device itself, its TV monitor, a host terminal, and, possibly, the image storage device and the control panel.

3.5.5. 3-D Solids Machine

The Medical Imaging Section of the Department of Radiology at the University of Pennsylvania Hospital has recently developed a Voxel Processor. Basically, it is a memory that can hold a cube of intensities with specially configured processors that calculate and load an image of a three-dimensional solid surface through that cube. The user has dynamic control over the viewing parameters (angle, perspective, clipping), as well as the threshold intensity. The current prototype can handle 64x64x64 cubes and generate frames at 15 Hz. A 256x256x256 Voxel Processor is under development. Such a device would be extremely helpful for the display and interpretation of three-dimensional data.

4. GRAPHICS DISPLAY

The realm of graphics display is vastly different from that of image display. This is not only expressed in the definition, but also, more importantly, in the commercial availability of products. In graphics we are dealing with geometrical objects (points, line segments, polygons, etc., most often translated into vectors), which may or may not have attributes such as color and intensity. Notwithstanding the vectorite nature of graphics display, much of it is displayed on raster devices today. In image display the data consist of actual intensities, measured on some regular sampling grid in space. Image display is more demanding in sophistication, I/O rates, and memory, while its application is confined to a relatively small number of specialized fields. Graphics, on the other hand, can be dealt with at a much lower level of sophistication and resources (although in the upper ranges it can certainly compete with image display in these respects), and has permeated virtually every facet of society, most importantly the business community. Since business has perceived graphics display as an important tool for increased productivity and effectiveness in nearly every type of work (public relations, promotion, management, production, process control, etc.) there has been a great incentive (as well as the financial resources) to develop an enormous spectrum of commercially available graphics hardware and software, at a reasonable price. This does, of course, mean that much of what is available - actually much of what gets most of the attention - is of little use in astronomy: pie charts, bar charts, oversophisticated lay-outs, which all seem to have become the trade mark of successful business.

The use of commercial software is very rare in astronomy for various reasons. The astronomical community has traditionally had needs that either exceeded the capabilities of such software or were not covered by it; capital funds were often limited, whereas manpower was available; and the tradition of doing things ourselves has proven very strong. As a result, a situation has developed where it is accepted to spend money on hardware and even on software where it concerns things like operating systems and compilers, but where it is considered a waste to buy other commercial software because "we can do it much cheaper ourselves". However, it may be wise to consider the practice of using bought software by the business community a little more serious than just to assume that they lack the expertise to do it themselves. The reasons are three-fold: commercial software usually provides very flexible software tools at a price that is lower than the manpower cost of in-house production; the (often considerable) burden of maintenance is put on the software supplier; and commercial packages provide hardware independence, with regard to the host as well as the peripherals used.

Graphics software is probably the field where astronomy can benefit most from the use of commercial software packages, for the reasons outlined above. The only experience we have is with DI-3000, a product of Precision Visuals, Inc. This experience shows that commercial packages can provide very powerful, flexible software tools that greatly increase programmer production, are easy to use, have excellent documentation, and provide device independence. Usually with

no extra effort a higher quality display is produced than one would otherwise bother to go to, while also flexibility for the user is increased. As far as cost is concerned, it should be kept in mind that software providers often offer significant educational discounts.

4.1. 2-D Displays

There are three types of two-dimensional graphics displays in use in astronomy: contour maps, graphs of various kinds, and interactive graphics in conjunction with (mostly) image display. This last type of display (drawing lines, boxes, polygons, or other types of regions of interest) is never used on its own, but only as part of another graphics or image display.

Similarly, there are three forms of output: non-permanent (on CRT or TV monitor), disposable permanent (i.e., work copies), and publication quality. The purposes and applications are self-evident, although it may be useful to mention the possibility of overlaying contour maps on images. Also, the three output media show a trial-and-error path toward final data display in the order given above. Ideally, the user should be able to get the same piece of graphics on different media by the flip of a switch.

CRT graphics display is rather poorly developed in astronomy. It has never really got beyond the Tektronics 4012. Attention should be given to the capabilities of more modern graphics terminals and the use of color.

In the field of work-copy devices for a long time the choice has been Versatec-like devices - which are extremely messy. Laser printers are beginning to replace them; however, they have one disadvantage over the Versatecs: the size of the plot has a firm and rather small upper limit.

Publication quality, although attainable on a Versatec, is still most easily produced on a good old-fashioned pen plotter. With the advent of plotters with built-in micro-processor these devices have become faster and more flexible, taking much of a burden off the host. The improvements include publication quality character fonts. Great assets are the availability of different color pens (indispensable for the contouring of velocity fields; facilitating contouring of complicated regions; allowing overlaying of different plots) and pens with different thicknesses (for publication). One has to be aware, though, that pen plotters, by their nature, require a slightly different style of programming (see section 4.3).

There are a few special needs for the contour maps that we usually deal with. One is that the contouring algorithm has to be able to deal with undefined points in an otherwise regular grid (for velocity fields, optical depths, dispersions, spectral index, etc.). Many algorithms either do not allow for this or are based on randomly sampled data; neither is acceptable. Another is that arbitrary sizes must be allowed; if necessary, it should be possible to assemble large plots from long strips. Yet another need is the production of contour map mosaics for spectral line observations. Finally, we need the capability to overlay other information on the plots: star positions, polarization vectors.

Ruled surface plots deserve special mention. There are two applications: to display two-dimensional data (e.g., an image) as a stack of one-dimensional cross-sectional profiles, which is especially

sensitive to gradients, instead of or in addition to a contour map; and to display a collection of truly one-dimensional profiles (e.g., amplitude as a function of time for a number of base lines), which is very sensitive to "odd" points.

There is one more type of graphics that needs attention: plotting a number of related parameters on a two-dimensional grid. An example of this is the display of the parameters of multiple Gaussian components to spectral profiles. Ulrich Schwarz has developed a satisfactory scheme for this case.

4.2. 3-D Displays

There are two applications for three-dimensional graphics displays in radio astronomy: three-dimensional contour maps of spectral line data and graphic display of visibility data. The former constitutes a substitute as well as an extension for the three-dimensional solids discussed in section 3.4.2. and needs no further clarification. The latter forms an extension of the image displays discussed in section 3.1., as well as of the graphs mentioned in the previous section and will be expanded upon below.

Three-dimensional graphics devices have commercially been available for a long time; Evans and Sutherland are probably the best known manufacturers. These devices allow real-time rotation, translation, hither- and yon-clipping, depth cueing, perspective, and zooming. Especially the motion cues introduced by real-time rotation are very powerful in conveying three-dimensional structural information. In addition, they will allow some form of stereoscopic display in conjunction with the other operations. Polygon fill has become available since manufacturers started switching to raster monitors and may, to some extent, emulate 3-D solid surfaces. Because of all these features and capabilities, the three-dimensional graphics devices are very powerful tools in exploring multi-dimensional data. Traditionally, these machines have been very expensive (well above \$50,000), but in recent years much cheaper devices doing essentially the same thing have come on the market, like the IBM 5080.

For the graphics display of visibility data, the 3-D graphics devices facilitate displaying amplitude or phase along the z-axis, while u and v lie along the x- and y-axes, as well as representing complex visibility as a true vector as a function of either base line and time or u and v. It is true, of course, that any view of such a representation can be displayed on a 2-D device, but the dynamic changing of the viewing angle greatly facilitates quick understanding of trends and patterns (e.g., the twisting of phase). Jim Torson made a demonstration program that simulates the effect on the VLA PDP-11/40 with the aid of an array processor. A similar display mode, lacking the dynamic capabilities, is in use for MERLIN data and supports the usefulness of these representations.

4.3. Device Independence

The great variety of output devices has been a blessing and a curse in the graphics world. It has led to a large degree of flexibility in display choices, but at the same time in an even greater inflexibility in application of software due to large amounts of device dependent code. With the proliferation of graphics and graphics software into every aspect of modern life (including those areas where there are large amounts of money), however, the impetus to standardize and introduce a large degree of device independence has become strong enough to achieve some results. The standards battle between CORE and GKS has been decided in favor of the latter. Three-dimensionality is still lacking in GKS, but will probably be brought in through the PHIGS extension. GKS also has, in principle, some image display elements, but it would, in my opinion, be unwise to rely on those; they are currently inadequate and it is doubtful whether they would ever become satisfactory.

The aim of the graphics standard is to achieve device independence. By defining the standard the interface with the output graphics device has been defined; therefore, once a driver has been written for a particular device that adheres to this standard, the device is available to all software that complies with the same standard. The use of such a graphics standard per se is not that easy. However, by using a graphics package that interfaces with the standard one can achieve considerable gains:

- the same code can be used for a number of output devices, i.e., the user can, in real time, choose the device that is most appropriate for his current needs (e.g., CRT or hard-copy).
- the quality of the display can be matched to the device and purpose of the output (e.g., quick-and-dirty or publication quality).
- increased programmer productivity (this depends on the quality of the package used).
- device independence from a management point of view: decisions on acquisition or replacement of graphics display devices do not have to be made anymore taking into consideration the match between a particular device and existing code. Also, mixing smart and dumb devices is not a consideration anymore since good device drivers will take advantage of all the capabilities of particular devices and simulate the rest in software.

The use of a commercial graphics package (discussed in the introduction to section 4) facilitates the ease of programming and ensures the availability of device drivers adhering to the standard; for the time being, device manufacturers cannot be relied on to provide the latter.

All in all, device independence is a wonderful thing. However, having said that, we also have to sound a warning. Device independence does introduce inherent inefficiencies: in writing code for a particular device one can take the peculiarities of the device into account and tailor the software to run as efficiently as possible. In using device independent code one has to pay in decreased performance as far as efficiency is concerned. In general, the advantages will outweigh the disadvantages, but in certain cases the result will be unsatisfactory. For instance, in writing a contouring program one has the choice between scanning the image line-by-line and following the contours. The former requires less memory, is simpler, and will run faster. In addition, it is the natural way to do it for dot-printing

devices. But it is totally unacceptable for pen plotters: not only will it result in very slow plotting with a lot of unnecessary pen movements, but the appearance of the plot will be very bad as well, because of all the little discontinuities where individual contour segments join. Use of the latter technique is mandatory for pen plotters. Therefore, some considerations of the characteristics of the device actually used will still have to enter into the code.

5. WORKING ENVIRONMENT AND SUPPORT FUNCTIONS

It goes without saying that in order to take full advantage of the potential that data display techniques hold, the work environment and support functions offered to the users must be thought out rather carefully. One can have the most sophisticated display software, but it is not going to profit the user if the TV monitors are mediocre, or if they have to be viewed under cramped or otherwise uncomfortable conditions, or if there is no hard copy available, or if the most beautiful images can be recorded but there is no way to process the film; one can add to this list ad libitum.

5.1. Work Stations

There are three requirements for organizing work stations: they should be pleasant to work at, there should be a spectrum of different types of work stations tailored to users' needs, and there should be enough of them.

The first requirement involves such things as: enough space to move around, pleasant temperature, low noise level, enough desk space, good lighting, and general ergonomic considerations. All controls should be in easy reach and placed at the right height and angle. Chairs should be comfortable and not interfere with the work. TV monitors should be placed at the right distance, height, and angle. Lighting should be adequate for desk and monitor viewing, and be adjustable.

Not all work stations have to be the same. Efficient use of resources dictates that the user can choose a work station that fits his needs for a particular session, but no more than those. A proper mix should be established of very sophisticated image display work stations with adequate graphics facilities, sophisticated graphics display work stations with simple image display facilities, and simple image display work stations (on-line as well as off-line, possibly just viewing stations). In addition, regular alphanumeric terminals (e.g., for preparation of batch oriented jobs) should be available. In general, rather simple hard copy output will be adequate for the work stations. High quality hard copy devices (plotter, image recorder) should be present, but at a central location; not only does one need only a few (or just one) of each, operations usually run smoother when they are under the care of a central operator.

Finally, there should be enough of these work stations. This implies that a proper mix, as outlined in the previous paragraph, is maintained. It also means minimizing the frustration level of the users. Hence, there should be enough of them to keep all users at a given time busy at an acceptable level, but not so many that the systems become clogged and responses sluggish.

5.2. Image Recording Facilities

There are probably three types of needs for high quality recording of images: slides for presentations, pictures for publications, and public relations work (posters, etc.). All three warrant the use of a high precision image recorder. The requirements are: recording areas of at least 4096 x 4096 pixels, high geometric accuracy, high linearity in intensity for at least 8 bits, good color rendition, and flexibility in film size uses. Experience has shown that the different needs have slightly different requirements for the type of film used and that the usage of the recording facility depends on how easy it is for a given user to get his slide or print made and how quickly he can get it. For these reasons the facility should be able to routinely handle 35 mm roll film and 4 x 5 inch sheet film, in color slide film, B/W negative, B/W reversal, and color negative film with no more than 24 hours turn-around time; in addition, a 16 mm film capability would be very desirable.

Obviously, there are other applications for image recorders. One might conceive of creating an alternative output path for graphics display through them. As a matter of fact, some graphics packages can provide a driver for a device like the Dicomed film recorder.

5.3. Photo/Graphics Facilities

As indicated in the previous section already, fast turn-around for images recorded on film is indispensable. This not only applies to the processing of the film but also to the capability of producing high quality prints. The same is true for the handling of high quality graphics work. The use of a film recorder for graphics display has been mentioned. Also needed are a high precision pen plotter and drafting services, as well as the photographic facilities to process their results.

In addition, astronomy requires some specialized photographic facilities: reproduction of optical material and the production of overlays.

All together, this calls for a well-equipped and -manned photolab that can take care of all photo processing needs on time scales of one day to one week and a drafting/graphics department. Part of the photo work would be the routine processing of computer-generated outputs.

6. CONCLUSION AND RECOMMENDATIONS

We conclude that judicious use of sophisticated data display hardware and techniques can make life a lot easier for the user and at the same time unburden the computing resources. This is achieved by providing interactive display and analysis functions that are aimed at providing the user with powerful tools for a better understanding of the contents of his data by moving the execution of those functions to the display devices. By doing this one not only shifts the compute power needed for the experimentation out of the host computers, but also enables the user to make well founded decisions on how to proceed with the data processing, resulting in less unnecessary processing and a faster reduction process for the user. It does require, however, an investment in display hardware, software, and research.

As for the actual image display devices, we recommend that machines be used with capabilities comparable with the IIS Models 70 or 75, but preferably with 1024 x 1024 pixel resolution, and with adequate amounts of refresh memory. The capabilities should at least include multiple image refresh memories with 8 bits each, the double lookup table architecture, 16 bit mode, and ALU/Feedback unit.

A variety of graphics devices should be available. Each work station should have a graphics CRT and access to a high quality hard copy device. One work station with a sophisticated three-dimensional graphics device would be extremely useful. To make usage of the various devices transparent to the user and the programmer, as well as to provide flexibility in management, it is strongly recommended that a good graphics package is adopted that adheres to an international graphics standard and that provides drivers for as large a number of devices as possible.

Display devices can very profitably be supplemented by special purpose hardware, but only in those cases where the desired features cannot be achieved in any other way.

Hard copy capability is very important. Each work station should have access to instant hard copy devices, for images and graphics. A central facility should provide high quality hard copy for both, with good procedures and support to make these available to the user in a reasonable time.

Appropriate attention should be paid to the working environment and the ergonomics of the work stations. This includes designing controls that make using the displays natural and easy for the user. There should be a proper variety of types of work stations, reflecting the varying needs of different users as well as of the various stages of data reduction and analysis.

Display of three-dimensional images can be made far more effective than the current ones are. A strong research effort in this direction could yield considerable gains.

Finally, no system is better than its support. One is not going to improve matters without an adequate level of and balance between the three support branches: software, hardware, and services. The first two have been discussed above. The third determines whether the user can actually achieve lasting use of the results of his toils; it is hard to overspecify photo/graphics services for a facility where image display plays a central role.