Design of and image editing with a space-filling three-dimensional display based on a standard raster graphics system

Henry Fuchs, Stephen M. Pizer
Department of Computer Science, University of North Carolina at North Chapel
New West Hall 035A, Chapel Hill, North Carolina 27514

E. Ralph Heinz
Department of Radiology, Duke University, Durham, North Carolina 27706

Sandra H. Bloomberg, Li-Ching Tsai, Dorothy C. Strickland
Department of Computer Science, University of North Carolina at North Chapel
New West Hall 035A, Chapel Hill, North Carolina 27514

We are developing graphics systems, image preprocessing methods, and interactive manipulation techniques for a space-filling 3D display using a varifocal mirror principle. Our driving problem is a medical imaging need for presentation of three-dimensional intensity information.++

The major goal of both the image preprocessing and the interactive manipulation has been to overcome obscuration, which we feel is coming to be recognized as the central problem in any space-filling display. In our system, the preprocessing step highlights important image features such as surfaces. At display time, the object can be dynamically edited and rotated for convenient viewing from various directions.

Our particular hardware design allows the 3D display to be constructed as an inexpensive add-on to a standard video graphics system. The interactive rotation and other manipulations are achieved by the standard built-in graphics processor.

1. 3D Display of Space-Filling Images

The majority of scenes displayed on graphics systems have been either two-dimensional or made up of three-dimensional lines and surfaces. An entirely different class of images is described by a grey (or color) level at each position (volume element) in three-space. Such data is produced in medicine by imaging modalities such as computed tomography (CT), ultrasound echography, and nuclear magnetic resonance, and in the earth sciences by satellite measurements and earth soundings in the exploration for oil. This type of three-dimensional data also arises in scientific areas such as physics and biochemistry, where, for example, one might have data on the electron density distribution in three-space. In all of these areas the intensity at each point specifies the value of some physical parameter there.

The objective of the display of such objects may be not only to visualize the objects from the physical measurements that produced the data, but also to manipulate objects to somehow match the space-filling image data. For example, one might wish to specify a radiation dose distribution that is high in the region of a tumor and low in the regions of sensitive organs, or one might wish to specify the 3D spatial conformation of a molecule so that the distribution of its atoms' electrons match the electron density distribution that has been measured. These matching objects normally consist of points, lines, or surfaces, and they often need to be displayed superimposed in 3D on the image of the space-filling data.

While projection together with certain depth cues such as stereo or the kinetic depth effect can be used to present the required 3D images on normal 2D display devices, more effective display can be obtained by space-filling 3D display devices that actually present an image distributed in three-space. The major difficulties with such display, whether on a 2D or a 3D device, are the high data rates required for fusion with the very large data bases that three dimensions implies and the obscuration of image portions by other portions in front of or behind a region of interest. The high data rates are a greater problem for devices with a sequential final display unit than for more parallel units, but it is a real problem with all display systems for space-filling 3D data. The obscuration problem, that simple grey-scale display in three-space produces a muddy image, is inherent to the display of space-filling data independent of the method or device. Consider, for example, the problem of appreciating the information in a stack of slides, even if all of the slides passed light as if they were transparent.

++This work is supported by NIH Grant #NS/HL 16759-01.

2. Approaches to Effective Space-Filling 3D Display

We have seen that the large number of points to be displayed causes difficulties of both data rate and obscuration. We are thus led to an attempt to limit the number of points without losing the important information in the data. Two approaches occur: an importance measure can be applied automatically to the image data and the display limited to the points with an appropriately high value of this measure, or the user can interactively specify objects or image regions of interest or manipulate the display so that these objects or regions are well visualized.

2.1. Measuring and thresholding on importance

One automatically determinable importance measure comes from noting that in many applications important information is located where the intensity is changing quickly, that is, near the edges of objects. Thus we have preprocessed the data to measure the edginess (edge strength) of each point, specified by the three-dimensional intensity gradient. Note that this determination is straightforward, e.g. by a 3D version of the Sobel operator that we have used, as compared to edge detection, where finding the edge strength must be followed by the much more difficult edge following step. With a 3D display device the edge following can be done by the visual system of the viewer.

We are often not interested in all edges. For example, in our major application of visualizing the carotid artery in the neck and the plaque on its wall from CT scan data, we may be interested in only edges of the blood and the plaque. The importance measure used is edge strength as weighted by a quickly falling function of spatial distance from a point with intensity corresponding to the contrast material in the blood. In our application we have found that limiting display to points with an appropriately high value of this importance measure produces a large reduction of the number of points to be displayed, making the data rate reasonable on our varifocal mirror display system, discussed below. It also produces quite a useful image with greatly lessened problems of obscuration.

2.2. Interaction

We noted above that interaction can be useful to specify objects or regions of interest and to manipulate the display so that these objects or regions are well visualized. Furthermore, in matching objects to space-filling data, as discussed in section 1, interaction is necessary to modify the matching object. For example, one might wish to stretch or rotate a bond to a residue in matching a molecular structure to an electron density distribution.

We have found one manipulation to be quite important for visualization of space-filling 3D objects in displays that can be viewed from a limited range of orientations: changing the point of view on the image (rotating the image). These rotations can be implemented by means that have been well-developed for graphics on 2D display devices.

We have found a variety of methods that are useful for interactively altering the image to make it more understandable. Among these are three methods based on location or intensity:

1) limiting display to (or highlighting) certain spatial regions, usually with simple shapes such as three-dimensional rectangles or ellipses,
2) limiting display to points with intensities in a particular range (or with importance measures in a particular range), or
3) superimposing a grey-scale image limited to a particular spatial region on objects selected by an importance measure. For example, we display a portion of a single CT slice superimposed on artery surfaces selected by the edginess measure that serve as a roadmap for the slice, and then allow the user interactively to change the location at which a slice is displayed.

Another category of methods depends on subdivision of the image into objects. Once this subdivision has been accomplished, one can blink or highlight certain objects, or one can choose to display only certain objects. Note that in this latter case, not only the obscuration problem but also the data rate problem is ameliorated.

To provide the second category of methods, there must be a means to subdivide the image data into objects. This is often straightforward if the objects are being defined interactively, as in producing a matching structure. It may also be possible by a preprocessing pattern recognition step, though we have not yet attempted such a step. Finally, a way may be provided by combining interactive pointing and region of interest specification with some pattern recognition.

As we examine the display and interaction objectives discussed above, we note the standard objectives of a general graphics system.

1) Objects need to be described.
2) These objects need to be transformed by rotation, blinking, highlighting, etc.
3) The resulting transformed objects need to be displayed - normally this involves refreshing a display device.
4) The objects need to be able to be selected and modified.
5) New objects need to be able to be drawn on-line or formed from other objects.
6) The illumination or selection for display of image portions needs to be determinable by their spatial or intensity value.

The only difference between these needs and those of a standard graphics system is item 6, which limits obscuration. Inhibiting objects or parts of objects for display is done not only because we do not wish to see these objects, since they obscure others, but also because with a limited data rate such inhibition allows increased quality of the display of the objects or regions of interest.

Since the objectives are only slightly extended from those of a general graphics system, it is reasonable to implement them using the more or less standard approach. A hierarchical object description in terms of primitives is created, this description in terms of primitives is organized to be stored in the memory of a graphics system, the display processor transforms these object primitives into a refresh buffer, and the display generation hardware causes the information in the refresh buffer to be presented, possibly after some final transformations. Our implementation of these ideas will be described in section 3.

Before describing the software implementation of these ideas, we must consider the selection of hardware for 3D display that is consistent with these ideas, has good display properties, and is not too expensive.

3. Realization of solutions

3.1. Possible 3d display systems

Of the many methods for displaying three-dimensional objects that have been developed, most exhibit only a few of the depth cues experienced by people in the "real world." Realistic rendered images, whether painted by hand or by computer, make use of obscuration, lighting of object surfaces, and occasionally, shadows. Real-time line-drawing systems use the kinetic depth effect. Neither of these systems exhibits head-motion parallax, so each has a rather limited set of depth cues. Other more esoteric methods include the display of stereo pairs [e.g. Valyus, 1966; Lipscomb, 1981], holograms [e.g. Lesem & Hirsch, 1968; Huang, 1971; Benton, 1977], integral photography using arrays of lenses [Lippman,1908; de Montebello,1977a], and vibrating or rotating mirrors [e.g. Withey, 1958; de Montebello, 1977b; Muirhead, 1961]. Holograms and integral photography generally require long times for image production and thus do not allow natural interaction. Many systems using vibrating or rotating mirrors have mechanical properties that make them inappropriate for routine use. The varifocal mirror [Traub, 1967; Hobgood, 1969; Fuchs, Pizer, et al, 1979; Baxter, 1981] is an exception.

3.2. Varifocal Mirror Display Principles

The term varifocal mirror display could as well be called a "vibrating mirror display." In this kind of system, 3D display is accomplished by the viewer observing the screen not directly, but by looking at a reflection of the screen in a flexible mirror (Figure 1). The mirror consists of flexible material such as aluminized mylar which is stretched taught over a rigid ring, like a drumhead. A loudspeaker mounted closely behind the ring is driven by a smooth sine wave of approximately 30Hz. The viewer's image of the screen appears extended in depth due to the mirror's vibrations. (A square image on the screen is extended into roughly a rectangular parallelepiped behind the mirror.) Now the major step! The screen is chosen to be a point-plotting CRT with a fast phosphor, and the list of displayed points is refreshed onto the screen at the same rate the mirror is vibrating (30Hz.). Then any particular point will always appear on the screen at the same time in the refresh cycle and thus when the mirror is in the same position, so the point will appear to the viewer to be at a particular depth in the parallelepiped. So, any collection of 3D points within the parallelepiped can be displayed simply by placing each 3D point at the place in the refresh buffer determined by its depth (Z) component. (Of course, since the mirror is always moving, only a single point can appear at any exact depth, but human depth resolution is sufficiently low that points close enough to their

Figure 1. A Viewer at the UNC Varifocal Mirror Display

desired location in the refresh list to be displayed within 0.1 inches of their desired depth appear to be at the same depth to the viewer.) With a 30Hz refresh and vibration rate, all the points are perceived simultaneously, so a true 3D image is seen with virtually all the depth cues, including stereopsis (different views from right and left eyes) and head-motion parallax (the object is perceived from different points of view as the viewer moves about the mirror). However, the hiding cue is not present, as the images are translucent. [The main difficulty, in fact, is that one cannot communicate the richness of the image via pictures in papers, videotape, etc. Of course one can take pictures of the images, as well as movies and videotape, but all these are mere flat 2D versions of a rich 3D environment! See Figure 2.]

As illustrated in Figure 2, a system for producing a 3D display using this method then consists of a 2D image presentation component made from a fast-phosphor point-plotting CRT; a vibrating mirror component made from a mirror, speaker, and stand; a component for producing one cycle of a 30Hz sine wave at a specified time; a refresh memory for storing points to be displayed; and a module for reading the refresh memory and displaying the points onto the CRT. There must also be a way to synchronize the point-plotting operation and the 30Hz sine wave. The system may also include a processor to transform the memory information in response to user interaction. Further details of general system design issues may be found in [Fuchs, Pizer, et al, 1982].

3.3. Previous work

A 3D display using the varifocal mirror display principle was first developed by A. Traub at MITRE based on work of Muirhead [Muirhead, 1961]. Later work included that of Rawson at Bell Labs and Hobgood at the University of North Carolina [Rawson, 1969; Hobgood, 1969]. In 1978 L. Sher of BBN developed an improved mirror for varifocal mirror displays. Rather than using the original taut membrane, his mirror is constructed of rigid fiberglass precisely machined and weighted at the perimeter so that it resonates at 30Hz. A system based on this mirror is starting to be marketed by Genisco Computers as SpaceGraph Terminal. Another group, at the University of Utah, has also developed a system [Baxter, 1981]. This one is being commercialized as part of a radiotherapy dose treatment planning system by Capintec, Inc. of Ramsey, N.J.

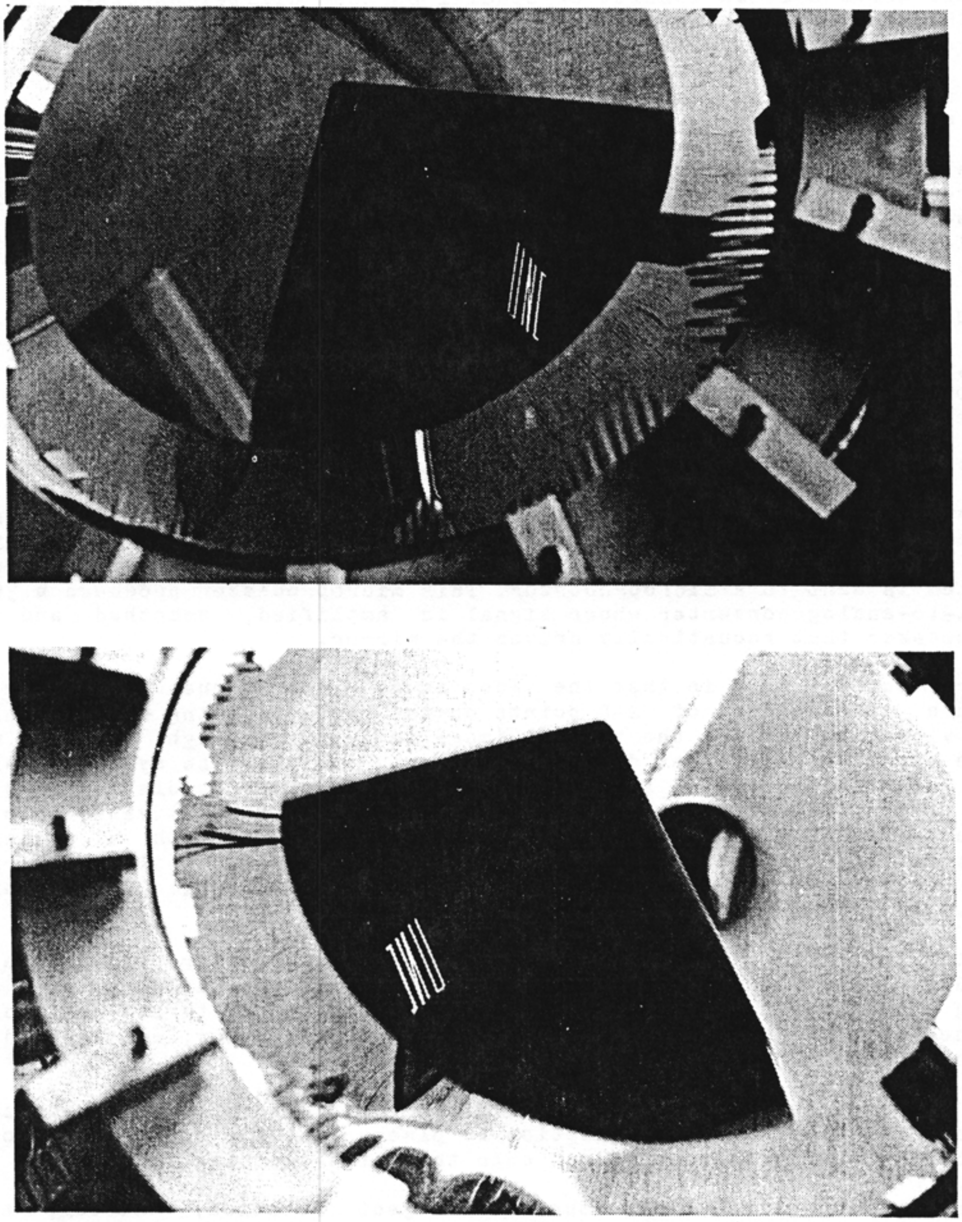

Figure 2. The Same Image Seen from Two Viewer Positions
The letters 'UNC' are drawn along the Z axis.

Both of these systems maximize the number of points that they can display with a "raster" mode in addition to the simple point plotting mode described above. Remember that in point mode, the refresh list consists of a sequence of points that are displayed onto a point-plotting CRT in synchrony with the mirror vibrations. The principal limitation to the number of points that can be displayed with this technique is the CRT deflection electronics themselves. Since the points can be anywhere on the screen, the CRT beam may have to be deflected from one edge of the screen to the other between plotting one point and the next. (The fastest generally available CRT's take about 100-200ns to do this.) If the consecutive points are close together on the screen surface, however, only a fraction of this deflection time is necessary, so more points can be displayed. To make the speedup significant, many points have to be near each other. This happens most frequently when entire slices of a 3D volume are to be displayed, such as a single CT image, and when such a slice lies at a particular depth of the mirror's virtual image. (Remember, the entire slice will be displayed while the mirror has moved very little.) In view of the discussion earlier on the difficulty of comprehending 3d images in which most of the volume has been filled with points, it appears to us that such raster modes are of limited utility.

Both of the above systems consist of special purpose digital hardware to display the data on a CRT and control the mirror. Due to lack of mass production, at least at this point, they are quite expensive. (The Genisco system sells for between $100,000 and $120,000.)

3.4. The UNC system

Whereas all other systems for varifocal mirror display that we know of have special-purpose electronics for memory and image point generation and thus are quite expensive, our system is based on a general purpose color raster graphics system. The result is that any raster graphics system with 16 or more bits per pixel can be enhanced with a varifocal mirror display for well under $10,000 in parts cost.

It should be noted that our design does not easily allow realization of the "raster" mode of the two other systems described above. Even though our system is based on a 2D raster system, we treat each pixel of the 2D refresh buffer as a distinct point in the x,y plane of the point-plotting CRT. From the above discussion of the difficulty of visualizing solid 3d images, we believe the lack of the raster mode is not a serious limitation.

3.4.1. System organization

The major components of our system are illustrated in Figure 3. The outputs of the video graphic system, the red, green, and blue signals, are connected directly to the x,y, and intensity inputs of the point-plotting CRT. The frame synchronization signal from the graphics system is sent to a microprocessor. This microprocessor produces a 30Hz sine wave onto a digital-to-analog converter whose signal is amplified, smoothed and sent to an ordinary loudspeaker that acoustically drives the mirror.

The central concept here is that the video image -- the sequence of color pixel values -- now becomes a sequence of 2-D points on the point-plotting CRT. By observing these points in the vibrating mirror, the virtual image is swept through some volume and the various points appear at different depths -- the ones early in the cycle appear closer, the ones somewhat later in the cycle, farther away. Thus, in order to display a 3D point having a particular (x,y,z,intensity) value, one puts the (x,y,intensity) value into the graphics system's image buffer at a place determined by the z value. (With a system generating interlaced video, all the points in the odd field are displayed while the mirror is moving from front to back; the points in the even field are displayed while the mirror is moving back to front.)

In addition to the above-mentioned components, our particular video graphics system (an Ikonas RDS 3000) contains a programmable graphics processor, so even excluding the mirror-driving microprocessor, our overall system consists of three distinct processing components: the VAX11/780 host, the Ikonas bipolar microprocessor, and the video generator. We utilize each of these units for a variety of tasks.

Figure 4 illustrates the simplest model of the system's internal organization. Each of the 3D points in the object description is placed into the appropriate location in the (2D) refresh buffer, which is scanned out onto the point-plotting CRT.

For highly interactive tasks such as object rotation, however, the available processing capacity of the main processor and its connection to the image buffer are inadequate. To overcome these limitations, we have implemented much of the current software system within the graphics system's internal programmable processor. In order for this processor to operate efficiently, we maintain a display file of 3D object descriptions within the graphics system. This enhanced conceptual organization is illustrated in Figure 5.

Before describing the organization further, it is helpful to note the simplification of two major otherwise-complicating factors about the mirror motion. During each 1/30 second cycle the mirror actually sweeps past each point in the 3D working volume twice -- once during its front-to-back travel, then again in its back-to-front travel. Thus there are two locations in the refresh buffer for each point; for a common interlaced display one of these two locations will be in the odd field, the other in the even field. Precise calibration between all such pairs of points, however, is somewhat complex, so for now we utilize only one of these two phases, effectively painting null intensities during the other phase.

The other complication about mirror motion is that since the mirror is always moving there is only a single point at any given depth; we increase this number by noting that nearby positions in the refresh buffer appear at about the same depth to the user. We therefore divide the refresh buffer into a fixed number of pieces ("slabs") and consider every position in each slab as being at the same Z. This allows a simple bucket sort

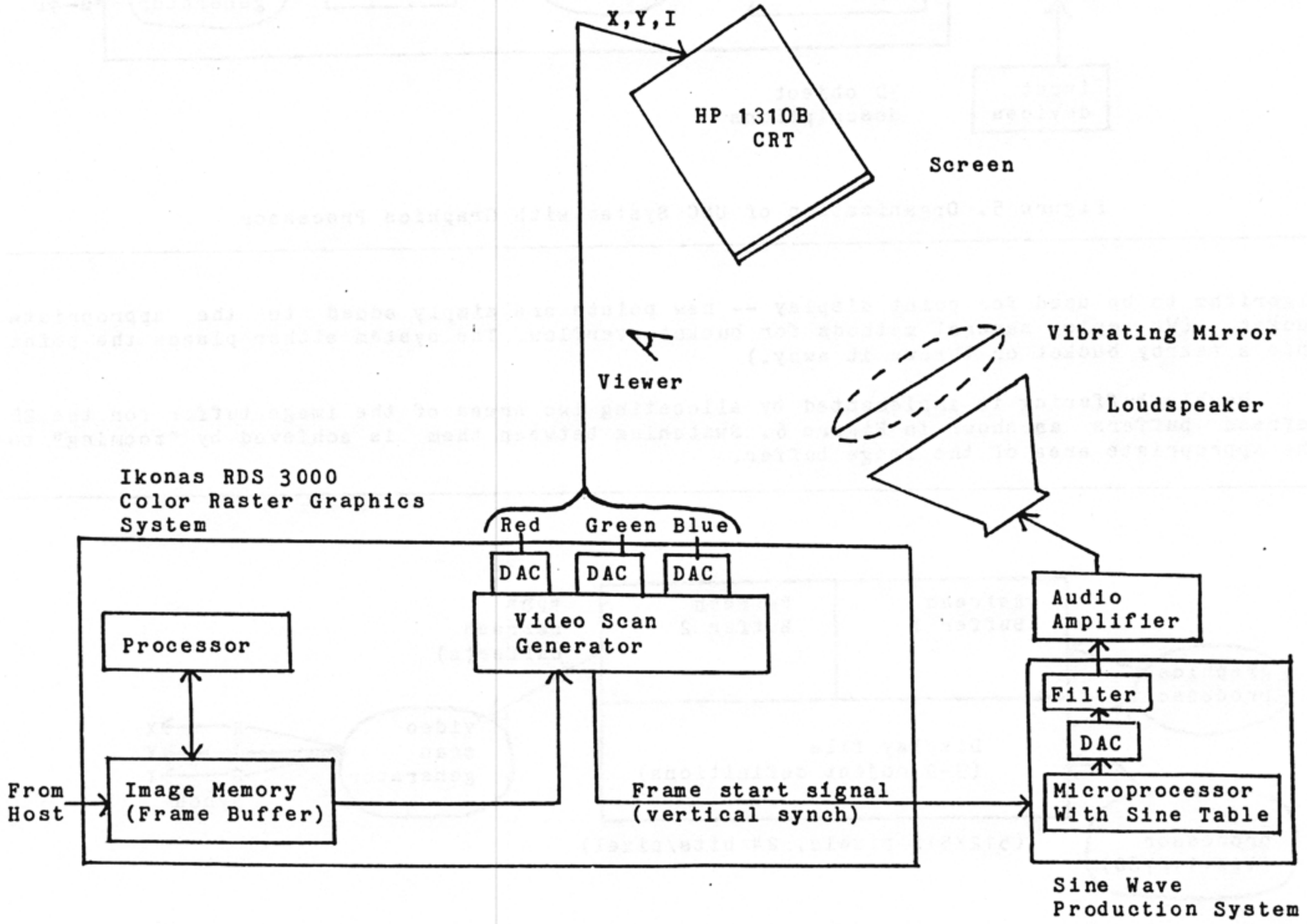

Figure 3. Major components of UNC Varifocal Mirror Display System

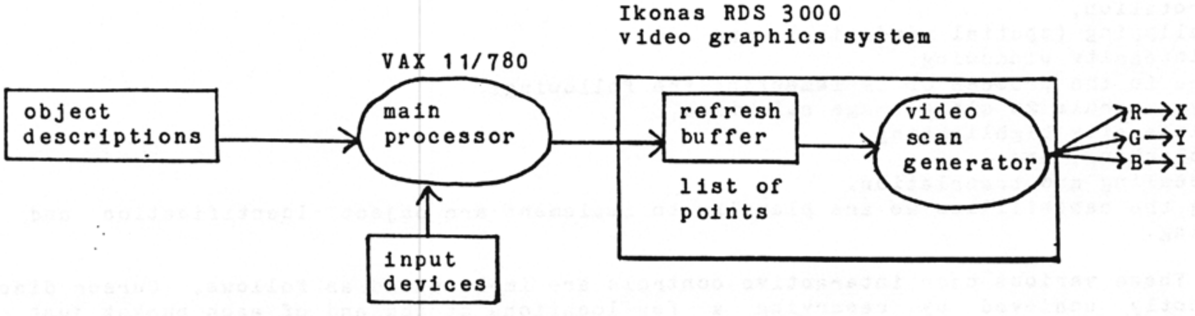

Figure 4. Simple Model of System Organization

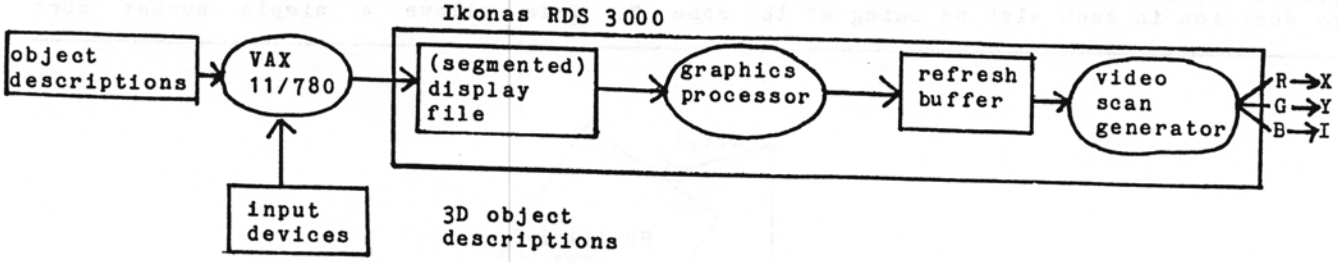

Figure 5. Organization of UNC System with Graphics Processor

algorithm to be used for point display -- new points are simply added to the appropriate bucket. (We employ several methods for bucket overflow. The system either places the point into a nearby bucket or throws it away.)

Double buffering is implemented by allocating two areas of the image buffer for the 2D refresh buffers as shown in Figure 6. Switching between them is achieved by "zooming" to the appropriate area of the image buffer.

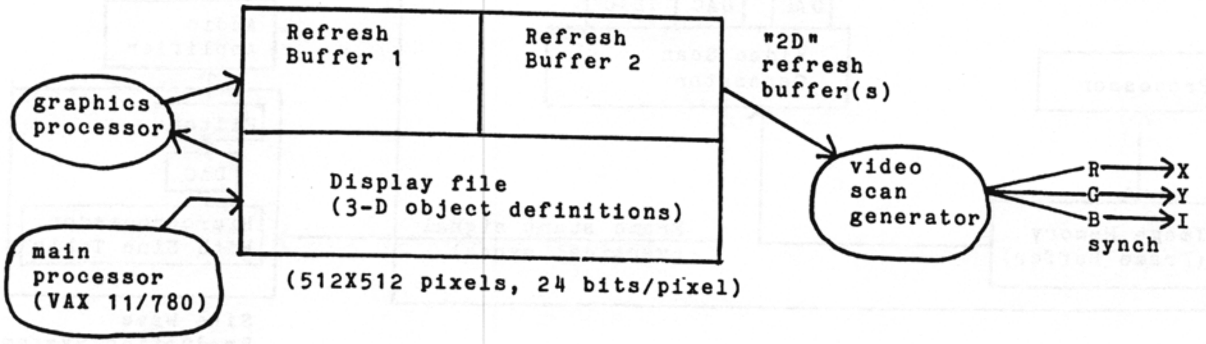

Figure 6. Image Buffer Organization

3.4.2. Available interactive operations

Section 2.2 described a set of interactive operations thought to be useful for any 3D display. Of that set, we now have the following operational:

1) cursor movement,
2) cursor-driven inking,
3) rotation,
4) clipping (spatial windowing), and
5) intensity windowing.

We are in the process of implementing the following:

1) grey-scale 2D slice image selection,
2) intensity highlighting,
3) blinking, and
4) scaling and translation.

Among the capabilities we are planning to implement are object identification and object editing.

These various user interactive controls are implemented as follows. Cursor display is currently achieved by reserving a few locations at the end of each bucket just for the cursor. The cursor is moved by erasing its current position and writing it into its new one. Inking is done by putting the cursor's point into a regular location in the bucket whenever the cursor is moved.

To achieve rotation, the user manipulates a 3D joystick from which the VAX host formulates the appropriate 3D transformation matrix and transmits this matrix to the graphics processor. The graphics processor transforms each of the points in the 3D display list by this matrix to obtain the XYZ location of the point in the working volume. The point is then inserted into the appropriate slab ("bucket") based on its new Z value (Figure 7). The mapping from Z to a slab address is a table lookup.

Clipping is achieved in different ways for the X and Y axes than for the Z axis. For the X and Y axes, the video lookup table is changed to reflect the mapping from the original value (the input to the table) to the value that should go to the CRT. All those entries higher than the right cut-off location are set to the maximum value. This appears on the screens as all points past the cut-off value are pushed aside (to better reveal the structure of the part of the object that remains). Z clipping is achieved by changing the viewport and window registers to avoid refreshing onto the screen certain portions of the image. (If this were viewed on a standard video monitor, it would appear as part of the image, say the top 30%, going black.) Intensity windowing is achieved in identical fashion to the X,Y windowing.

3.4.3. User interaction facilities to come

We are in the process of implementing interactive display of the original grey-scale slices. Display of these is complicated by the difficulty of refreshing so many points within each cycle time. In addition, when displaying more than a few slices, the obscuration problem becomes severe. The user, however, often may wish to refer back to the original slice, especially how it relates to the extracted information being displayed in the working volume. Thus fast response is important. To enable this feature (as well as some others), we are segmenting the display file into numerous pieces, each associated with a slice or some other object. Each object is marked by the user as either "active" (to be displayed), or "inactive" (not to be displayed). The user's latest selected "active" objects will be displayed at the very next cycle.

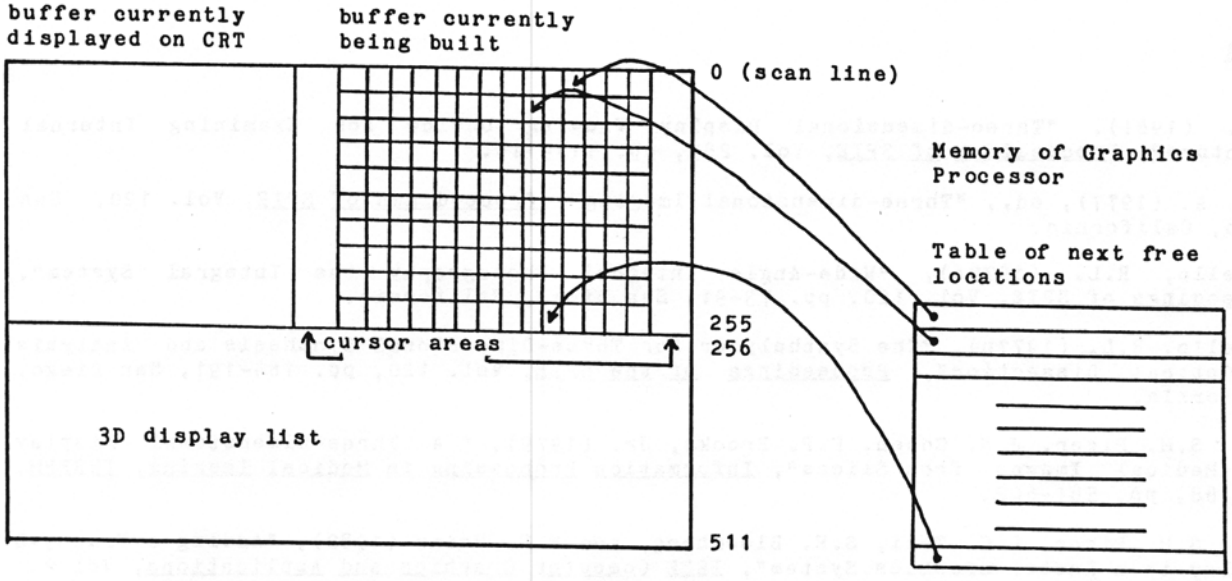

Figure 7. Organization of Buckets in Refresh Buffer

Intensity highlighting can be achieved to a limited degree by intensity modification of the Blue video lookup table, the one controlling the intensity output to the CRT. Higher intensities will be achieved by repeating points of highlighted objects into the refresh buffers so the beam will spend significantly more time at the points of highlighted objects than at points of unhighlighted ones.

We expect to achieve blinking capability by allocating one bit of the 24 bits per pixel/point specification. The intensity ("Blue") look up table will be switched between two different sets of values at the blinking rate to force blink-marked points to zero intensity during half of each blink cycle.

Scaling and translation capability will require only a small enhancement to the current rotation matrix, just as in any conventional "3D" graphics systems.

We are just beginning to plan for object identification and object editing capabilities. We want the system to be able to assist the user in identifying the extent of objects -- an artery, for instance. Object editing should enable the user to erase parts of objects and split objects into multiple objects. These capabilities will utilize heuristic techniques probably running in the VAX host computer.

4. Summary

We have implemented a true 3-D display based on a standard raster-graphics system. Our system is most frequently used to display 3-D density distributions of anatomical data. We have discovered that new display and interaction techniques are needed to understand these complex space-filling images. Some of these techniques include display of points only within a certain selected intensity range and display of edge-strength data rather than original density information. Real-time intensity windowing allows the user to explore details of small objects and subtle boundaries between objects. Similar interactive windowing within X,Y, and Z dimensions allows the user to cut away unwanted obstructions. Physicians viewing the display report readily noting information not apparent in the original sequential 2D images. Our experience with this system indicates that it has a bright future. Its general capabilities should be useful for other applications; its low add-on cost will make it accessible to many users.

REFERENCES

Baxter, B. (1981), "Three-dimensional Display Viewing Device for Examining Internal Structure", Proceedings of SPIE, Vol. 283, pp. 111-115.

Benton, S. A. (1977), ed., "Three-dimensional Imaging", Proceedings of SPIE, Vol. 120, San Diego, California.

de Montebello, R.L. (1977a), "Wide-Angle Integral Photography the Integral System", Proceedings of SPIE, Vol. 120, pp. 73-91, San Diego, California.

de Montebello, R.L. (1977b), "The Synthalyzer for Three-Dimensional Synthesis and Analysis by Optical Dissection", Proceedings of the SPIE, Vol. 120, pp. 184-191, San Diego, California.

Fuchs, H., S.M. Pizer, J.S. Cohen, F.P. Brooks, Jr. (1979), " A Three-Dimensional Display for Medical Images from Slices", Information Processing in Medical Imaging, INSERM, Vol. 88, pp. 581-602.

Fuchs, H., S.M. Pizer, L.C. Tsai, S.H. Bloomberg, and E.R. Heinz (1982), "Adding a True 3-D Display to a Raster Graphics System", IEEE Computer Graphics and Applications, Vol 2.

Hobgood, W.S. (1969), A Three-Dimensional Computer Graphics Display Using a Varifocal Mirror, Master's Thesis, University of North Carolina, Chapel Hill.

Huang, T.S. (1971), "Digital Holography", Proceedings of the IEEE, Vol. 59, No. 9, pp. 1335-1346.

Lesem, L.B. and P.M. Hirsch (1968), "Computer Synthesis of Holograms for 3-D Display", Communications of the ACM, Vol. 11, No. 10, pp. 661-674.

Lippmann, G. (1908), Comptes-Rendus, Vol. 146, pp. 446-451.

Lipscomb, J. (1981), "Three dimensional display of molecular models", Ph.D. Dissertation, University of North Carolina at Chapel Hill.

Muirhead, J.C. (1961), "Variable Focal Length Mirrors", Review of Scientific Instrumentation, Vol. 32, pp. 210-211.

Pizer, S.M., H. Fuchs, E.R. Heinz, S.H. Bloomberg, and L.C. Tsai (1982), "Varifocal Mirror Display of Organ Surfaces from CT Scans", Proceedings of the World Federation of Nuclear Medicine and Biology, Paris, France, August 1982.

Rawson, E.G. (1969), "Vibrating Mirrors for 3-D Imaging", IEEE Spectrum, Vol. 6, pp. 37-43.

Sher, L. (1978), Personal communications, Bolt Beranek and Newman, Boston, Mass.

SpaceGraph Terminal (1981), Promotional literature, Genisco Computers Cooperation, Costa Mesa, California.

Traub, A.C. (1967), "Stereoscopic Display Using Rapid Varifocal Mirror Oscillations", Applied Optics, Vol. 6, No. 6, pp. 1085-1087.

Valyus, N.A. (1966), Stereoscopy, New York: Focal Press.

Withey, E.L. (1958), "Cathode-Ray Tube Adds Third Dimension", Electronics, Vol. 31, No. 21, pp. 81-83.