4/30/89
→ IEEE
computer

Interactive Visualization of 3D Medical Data

Henry Fuchs, Marc Levoy, and Stephen M. Pizer

January, 1989
(revised April, 1989)

Computer Science Department
University of North Carolina
Chapel Hill, NC 27599

ABSTRACT

The increasing use of CT, MRI, and other multi-slice imaging modalities in medicine has inspired the development of new techniques for visualizing multi-dimensional data. This paper compares the three currently dominant techniques for displaying three-dimensional medical data: surface-based rendering, binary voxel rendering, and volume rendering. It also describes some of the display hardware being developed to help visualize these datasets: stereo viewers, varifocal mirrors, real-time image generation systems, and head-mounted displays. Particular emphasis is given to real-time volume rendering because current research suggests that it may become the method of choice for many clinical applications.

Keywords: visualization, medical imaging, voxel, volume rendering.

INTRODUCTION

This past decade has seen a revolution in techniques for noninvasively imaging the interior of the human body. New data acquisition modalities include computed tomography (CT), single photon emission computed tomography (SPECT), positron emission tomography (PET), magnetic resonance imaging (MRI), and ultrasound. All of these modalities have the potential for producing three-dimensional arrays of intensity values.

Unfortunately for the clinician, no fully satisfactory method for viewing this data yet exists. The currently dominant method consists of printing slices of the data onto transparent films for viewing on a back-lit screen. This approach makes detection of small or faint structures difficult. It also hampers comprehension of complex three-dimensional structures such as branching arterial systems. Without such an understanding, the planning of surgery, radiation therapy, and other invasive procedures is difficult and error-prone.

It has long been recognized that computer-generated imagery might be an effective means for presenting three-dimensional medical data to the clinician. Researchers have over a period of fifteen years built up a large repertoire of techniques for visualizing multi-dimensional information. The application of these techniques in clinical settings has until recently been limited due in part to low image quality and in part to the large quantity of data involved. The recent advent of larger computer memories and faster processors has spawned a period of rapid growth in software and hardware techniques for data visualization. Widespread clinical use of these techniques can be expected to follow within the next five or ten years.

RENDERING TECHNIQUES

Three-dimensional medical data has, by itself, no visible manifestation. Implicit in the visualization process is the creation of an intermediate representation - some visible object or phenomenon - that can be rendered to produce an image. This intermediate representation can be almost anything: dots, lines, surfaces, clouds, gels, etc. Since the human perceptual system expects sensory input to arise from physically plausible phenomena and forms interpretations on that basis, the representation should be of a physically plausible object. To promote easier interpretation, it should be of an intuitively familiar one. The majority of techniques for displaying three-dimensional medical data fall into three broad categories depending on the intermediate representation they employ.

Surface-based techniques first apply a surface detector to the sample array, then fit geometric primitives to the detected surfaces, and finally render the resulting geometric representation. The techniques differ from one another mainly in the choice of primitives and the scale at which they are defined. A common technique is to apply edge tracking on each data slice to yield a set of contours defining features of interest. A mesh of polygons can then be constructed connecting the contours on adjacent slices [Fuchs77]. Figure 1 was generated using this technique. Alternatively, polygons can be fit to an approximation of the original anatomy within each voxel, yielding a large set of voxel-sized polygons [Lorensen87].

These techniques have several desirable characteristics. Geometric primitives are compact, making storage and transmission inexpensive. They also have a high degree of spatial coherence, making rendering efficient. Unfortunately, automatic fitting of geometric primitives to sample data is seldom entirely successful. Contour following algorithms, for example, occasionally go astray when processing complicated scenes, forcing the user to intervene. These techniques also require binary classification of the incoming data; either a surface passes through the current voxel or it does not. In the presence of small or poorly defined features, error-free binary classification is often impossible. Errors in classification manifest themselves as visual artifacts in the generated image, specifically spurious surfaces (false positives) or erroneous holes in surfaces (false negatives).

Binary voxel techniques begin by thresholding the volume data to produce a three-dimensional binary array. The *cuberille* algorithm renders this array by treating 1's as opaque cubes having six polygonal faces [Herman79]. Alternatively, voxels can be painted directly onto the screen in back-to-front (or front-to-back) order [Frieder85], or rays can be traced from an observer position through the data, stopping when an opaque object is encountered [Schlusselberg86].

Because voxels, unlike geometric primitives, have no defined extent, resampling becomes an important issue. To avoid a "sugar cube" appearance, some sort of interpolation is necessary. These techniques also require binary classification of the sample data, and thus suffer from many of the artifacts that plague surface-based techniques.

Volume rendering techniques are a variant of the binary voxel techniques in which a color and a partial opacity is assigned to each voxel. Images are formed from the resulting colored semi-transparent volume by blending together voxels projecting to the same pixel on the picture plane. Quantization and aliasing artifacts are reduced by avoiding thresholding during data classification and by carefully resampling the data during projection.

Volume rendering offers the important advantage over surface-based or binary voxel techniques of eliminating the need for making a binary classification of the data. This provides a mechanism for displaying small or poorly defined features. Researchers at Pixar, Inc. of San Rafael, California first demonstrated volume rendering in 1985. Their technique consists of estimating occupancy fractions for each of a set of materials (air, muscle, fat, bone) that might be present in a voxel, computing from these fractions a color and a partial opacity for each voxel, geometrically transforming each slice of voxels from object-space to image-space, projecting it onto the image plane, and blending it together with the projection formed by previous slices [Drebin88]. Figures 2, 4, and 5 in this paper were generated using an algorithm developed at the University of North Carolina. It is similar in approach to the Pixar technique, but computes colors and opacities directly from the scalar value of each voxel and renders the resulting volume by tracing viewing rays from an observer position through the dataset

[Levoy88].

Despite its advantages, volume rendering suffers from a number of problems. High on this list is the technique's computational expense. Since all voxels participate in the generation of each image, rendering time grows linearly with the size of the dataset. Published techniques take minutes or even hours to generate a single view of a large dataset using currently available workstation technology. To reduce image generation time in our rendering algorithm, we employ several techniques that take advantage of spatial coherence present in the data and its projections. The first technique consists of constructing a pyramid of binary volumes to speed up the search for non-empty (non-transparent) voxels along viewing rays. The second technique uses an opacity threshold to adaptively terminate ray tracing. The third technique consists of casting a sparse grid of rays, less than one per pixel, and adaptively increasing the number of rays in regions of high image complexity. Combining all three optimizations, savings of more than two orders of magnitude over brute-force rendering methods have been obtained for many datasets [Levoy89].

Another drawback of volume rendering is its lack of versatility. Many clinical problems require that sampled data and analytically defined geometry appear in a single visualization. Examples include superimposition of radiation treatment beams over patient anatomy for the oncologist and display of medical prostheses for the orthopedist. We have developed two techniques for rendering mixtures of volume data and polygonally defined objects. The first employs a hybrid ray tracer capable of handling both polygons and sample arrays. Figure 2 was generated using this algorithm. The second consists of 3D scan-converting the polygons into the sample array with anti-aliasing, then rendering the ensemble. Figure 5 includes polygons rendered using this algorithm [Levoy89].

The computation of voxel opacity from voxel value in volume rendering algorithms performs the dual tasks of classifying the data into objects and selecting a subset of these objects for display. The classification procedure commonly used to display medical data consists of a simple one-for-one mapping from voxel value to opacity. In many cases, it is impossible to obtain the desired visualization using such a mapping. For example, in the MR study shown in figure 4, the scalp and cortical surface have roughly the same voxel value. Any mapping that renders the scalp transparently will also render the cortical surface transparently. How then to see the cortical surface? One possible strategy is to manually edit the dataset to remove the scalp as shown in figure 3. In this case, the boundary between erased and unerased voxels falls within regions that are rendered transparently. The boundary is therefore not seen in the volume rendering and need not be specified precisely. An alternative solution would be to combine volume rendering with high-level object definition methods in an interactive setting. Initial visualizations made without the benefit of object definition would be used to guide scene analysis and segmentation algorithms, which would in turn be used to isolate regions of interest, producing a better visualization. If the output of such segmentation algorithms included confidence levels or probabilities, they could be mapped to opacity and thus modulate the appearance of the image.

DISPLAY HARDWARE

The generation of a two-dimensional image from a three-dimensional dataset necessarily involves a reduction in information content. In particular, images do not utilize many faculties of the human visual system specifically adapted for perceiving depth: stereopsis, head motion parallax, and the kinetic depth effect. To convey as much three-dimensional information as possible to the observer, a number of specialized display devices have been developed.

Stereo viewers offer a simple, inexpensive means for widening the pathway from the computer to the observer. Our current favorite is a polarizing liquid crystal plate manufactured by Tektronix. The plate is mounted on the front of a video display, left and right eye images are displayed on alternate video fields, and the plate's direction of polarization is electronically switched in synchrony with the alternation of images. The user wears inexpensive passive glasses with polarizing material of different direction in each lens. Using this system, multiple users can see the same stereo image, and each user can look at multiple stereo displays.

Varifocal mirrors provide "true" three-dimensional perception of a distribution of glowing points of light. Our version consists of a vibrating aluminized mylar membrane stretched on a drumhead-like structure and acoustically vibrated from behind by a large conventional speaker driven by a low frequency (30Hz) sine wave. Users looking into this vibrating (and "varifocal") mirror see the image of a simple point-plotting CRT on which a sequence of dots is rapidly displayed. The displayed list of dots is repeated at the 30Hz rate, synchronized to the mirror vibration. Each dot is thus perceived at a particular depth, depending on the position of the mirror when the dot is displayed. As users move their head relative to the display, they see different views of the data.

In addition to providing both stereopsis and head-motion parallax, this display device can be viewed by several users simultaneously without their needing to wear any special apparatus. A major drawback of the system is the limited size of the displayed volume, which is constrained by the size of the CRT and the deflection of the mirror membrane. Another drawback is the lack of obscuration by the displayed dots; features in the foreground do not hide features in the background.

Cine sequences consist of sets of images that have been precalculated, stored in a frame buffer of sufficient memory capacity, and played back in real-time. The most common type of sequence is one in which the light source and observer are fixed and the object rotates. The resulting play of reflections across surfaces in the data enhances comprehension of the scene. A type of sequence frequently used in volume rendering is one in which the object, the light source, and the observer are all fixed and only the opacities of voxels change. For example, our physician colleagues find it useful to interactively vary the position of a clipping plane, allowing them to study in detail the tiny complex structures of the middle and inner ear. We provide this capability by computing a sequence of images in which the opacities of all voxels lying between the observer and the clipping plane are attenuated to nearly zero, and each image in the sequence has the clipping plane in a slightly different position relative to the data.

When calculating a sequence of images, we usually vary only a single parameter: observer position, light source position, cutting plane position, opacity mapping, etc. Users would often prefer to vary multiple parameters. Unfortunately, the number of images that must be precalculated to provide such multi-axis cine sequences is the product of the numbers of steps in the variation of each parameter; a modest 20 steps for each of two variables requires the calculation and storage of 400 images. The largest frame buffer currently in our laboratory is the Pixar Image Computer, which has a capacity of 64 512 x 512 12-bit images. This seriously limits the extent to which we can independently vary multiple parameters during playback of cine sequences.

Real-time image generation systems obviously offer more flexibility than cine sequences for the presentation of moving imagery. For surface-based rendering techniques, high-performance graphics workstations are available from Ardent, Silicon Graphics, and Stellar. At the University of North Carolina, we have been using our locally developed Pixel-Planes 4 system to provide clinicians with real-time display of contour-based polygon renderings (see figure 1).

For binary voxel representations, Phoenix Data Systems, Dimensional Medicine, and Dynamic Digital Displays have fielded systems capable of real-time or near real-time image generation. In addition, Kaufman is currently prototyping a system (and associated algorithms for 3D scan-conversion of surfaces and polyhedra) that will be capable of displaying a 512 x 512 x 512 voxel binary volume in real-time [Kaufman88].

At this writing, no system exists that can perform volume rendering in real-time. The most widely used system is the Pixar Image Computer. Using a four channel SIMD processor and Pixar's ChapVolumes software package, the Image Computer can generate high-quality images of 256 x 256 x 256 voxel datasets in about one minute. At the University of North Carolina, we are currently developing Pixel-Planes 5, a massively parallel raster graphics engine consisting of 32 20-MFLOP graphics processors, 1/4 million pixel processors, a 512 x 512 pixel color frame buffer, and a 640 Mb/sec ring network [Fuchs89]. In addition to enhancing our surface-based rendering capabilities, we intend to integrate Pixel-Planes 5 into a real-time volume rendering workstation. The frame rate we expect from this workstation depends on what parameters change from frame to frame. Preliminary estimates suggest that for changes in observer position alone, we will be able to generate a sequence of slightly coarse images at between 10 and 20 frames per second and a sequence of images of the quality of figure 4 in this paper at between 1 and 2 frames per second.

Head-mounted displays, pioneered by Ivan Sutherland at the University of Utah in 1968, incorporate stereopsis, head-motion parallax, and realistic images (including obscuration). These devices present stereo images to the user through displays mounted in front of each eye, track the motion of the user's head, and generate new images in real-time in response to changes in the user's head position and direction of view. This provides the illusion of walking about a simulated object or even of walking inside a simulated environment. Most of the development in this area has been for military and aerospace applications at such places as Wright Patterson Air Force Base and NASA Ames Research Center. We have also been developing head-mounted displays for several years.

Aside from the need for very fast update rates, these systems currently suffer from several technological weaknesses: poor resolution of the small video displays, high latency in the tracker-to-image feedback loop, and limited-range tracking systems. Nonetheless, they continue to hold great promise, and we hope to eventually incorporate a head-mounted display, Pixel-Planes 5, and our volume rendering algorithms into an interactive radiotherapy treatment planning system.

TOPICS FOR FUTURE RESEARCH

An important component of any computer-assisted medical image display system is a satisfactory user interface. In addition to the firms and research groups already mentioned, interactive medical image display systems have been developed by CEMAX, the Mayo Clinic, Sun Microsystems, and Lee Westover of our own department. User interfaces for many existing systems are constrained by the inability to generate images in real-time. For volume rendering systems in particular, feedback during selection of rendering parameters is usually provided by *meta-visualizations* such as 2D plots of color and opacity versus input value, wire-frame representations of viewing frustrums and motion paths, etc. If future systems succeed in generating volume rendered images in real-time, these meta-visualizations can be omitted or relegated to a supporting role. In such systems, sequences of rendered images would serve as feedback to the user of changes made in rendering parameters.

An important problem that has only begun to be addressed is how to visualize multiple datasets. Examples include superimposition of anatomic information from CT data and metabolic information from PET data, or simultaneous visualization of anatomy and radiation dose for cancer therapy planning. Realistic shading models have been used to visualize multiple 2D scalar fields by assigning one dataset to surface relief and another to surface albedo or color. For three-dimensional data, the second dataset can be used to modulate the color or opacity computed for voxels in the first dataset. Alternatively, the second dataset could be used to perturb the normals, shininess, or other properties of the surfaces displayed from the first dataset. Recent investigations in this area by Hoehne et al. [Hoehne88] and workers at the University of Chicago [Levin88] are very promising.

Current perspective suggests that the applicability of volume rendering to certain disciplines - diagnostic radiology in particular - hinges on answering the question: how correct is a volume rendering? Volume rendering has the disadvantage relative to surface-based techniques that it renders fuzzy surfaces represented by a bandlimited sample array rather than sharply defined surfaces represented by geometric primitives. Fuzzy surfaces do not occur in daily life; they are therefore not intuitively familiar. Moreover, volume rendered surfaces do not look fuzzy; they look precise. The danger therefore exists of interpreting these visualizations incorrectly. Since the radiologist is often required to make subtle judgements of feature size and shape, the potential exists of making erroneous diagnoses based on these visualizations. It is clear that volume rendering is useful for navigating through complex 3D data and for identifying gross anatomic abnormalities. It remains to be seen how useful it is for perceiving fine features, and the extent to which it can actually replace slice-by-slice presentation of medical data in clinical practice.

ACKNOWLEDGEMENTS

The CT study was provided by NC Memorial Hospital, the MR study was provided by Siemens AG and edited by Juiqi Tang, and the electron density map was provided by Chris Hill of the University of York. We thank our collaborators in the Radiation Oncology Department at UNC School of Medicine and NC Memorial Hospital, especially Julian Roseman, Ed Chaney, and George Sherouse. We also thank Andrew Skinner and Richard David for assistance with clinical studies, Phil Stancil for hardware systems support, and Rick Pillsbury for observations about the utility of cine sequences. This research is partially supported by NIH Grant #1 PO1 CA47982- 01, NSF Grant #CDR-86-22201, DARPA ISTO Order #6090, ONR Contract #N00014-86-K- 0680, and IBM.

REFERENCES

[Drebin88] Drebin, R.A., Carpenter, L., and Hanrahan, P., "Volume Rendering," *Computer Graphics,* Vol. 22, No. 4, August 1988, pp. 65-74.

[Frieder85] Frieder, G., Gordon, D., and Reynolds, R.A., "Back-to-Front Display of Voxel-Based Objects," *IEEE Computer Graphics and Applications,* Vol. 5, No. 1, January, 1985, pp. 52-59.

[Fuchs77] Fuchs, H., Kedem, Z.M. and Uselton, S.P., "Optimal Surface Reconstruction from Planar Contours," *Communications of the ACM,* Vol. 20, No. 10, 1977, pp. 693-702.

[Fuchs89] Fuchs, H., Poulton, J., Eyles, J., Greer, T., Goldfeather, J., Ellsworth, D., Molnar, S., Turk, G., Tebbs, B., and Israel, L., "Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor-Enhanced Memories," *Computer Graphics,* Vol. 23, No. 4, July, 1989.

[Herman79] Herman, G.T. and Liu, H.K., "Three-Dimensional Display of Human Organs from Computer Tomograms," *Computer Graphics and Image Processing,* Vol. 9, No. 1, January, 1979, pp. 1-21.

[Hoehne88] Hoehne, K.-H., Bomans, M., Tiede, U., and Riemer, M., "Display of Multiple 3D-Objects Using the Generalized Voxel-Model," *Proc. SPIE,* Vol. 914, 1988, pp. 850-854.

[Kaufman88] Kaufman, A. and Bakalash, R., "Memory and Processing Architecture for 3D Voxel-Based Imagery," *IEEE Computer Graphics and Applications,* Vol. 8, No. 6, November, 1988, pp. 10-23.

[Levin88] Levin, D.N., Hu, X., Tan, K.K., Galhotra, S., Herrmann, A., Chen, G.T.Y., Pelizzari, A., Balter, J., Beck, R.N., Chen, C.T., and Cooper, M.D., "Integrated 3-D Display of MR, CT, and PET Images of the Brain," *Proc. NCGA '89,* Philadelphia, PA, April, 1989, Vol. I, pp. 179-186.

[Levoy88] Levoy, M., "Display of Surfaces from Volume Data," *IEEE Computer Graphics and Applications,* Vol. 8, No. 3, May, 1988, pp. 29-37.

[Levoy89] Levoy, M., "Display of Surfaces from Volume Data," Ph.D. dissertation, University of North Carolina at Chapel Hill, May, 1989.

[Lorensen87] Lorensen, W.E. and Cline, H.E., "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," *Computer Graphics,* Vol. 21, No. 4, July, 1987, pp. 163-169.

[Schlusselberg86] Schlusselberg, D.S. and Smith, W.K., "Three-Dimensional Display of Medical Image Volumes," *Proc. NCGA '86,* Anaheim, CA, May, 1986, Vol. III, pp. 114-123.

Figure 1: Interactive surface-based rendering on Pixel-Planes 4 of a female pelvis with vaginal inserts for radiation sources. The long bulbous object surrounding the tip of the three vertical shafts is a polygonally defined radiation isodose surface.

Figure 2: Volume rendering of 256 x 256 x 113 voxel CT study of a human head. A polygonally defined tumor (in purple) and radiation treatment beam (in blue) have been added using our hybrid ray tracer. A portion of the CT data has been clipped away to show the 3D relationships between the various objects.

Figure 3: Slices from 256 x 256 x 156 voxel MR study of human head being edited to remove the scalp and thereby reveal the underlying cortical surface. Depicted is a typical view of the workstation screen in our interactive IMEX (IMage EXecutive) medical image editing program written by Peter Mills.

Figure 4: Volume rendering of edited MR dataset. The apparent mottling of the facial surface in the volume rendering is due to noise in the acquired data. Note that the user was not called upon to define surface geometry, but merely to isolate a region of interest.

Figure 5: Volume rendering of isovalue contour surface from 24 x 20 x 11 voxel electron density map of Staphylococcus Aureus ribonuclease. 3D scan-converted backdrop planes, cast shadows, and a color-coded stick representation of the molecular structure have been added to enhance comprehension. Similar techniques can be used to render isodose contour surfaces for radiation treatment planning.