

Interactive Visualization of 3D Medical Data

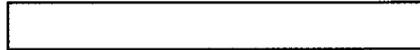
Henry Fuchs, Marc Levoy, and Stephen M. Pizer

University of North Carolina

Techniques for noninvasively imaging the interior of the human body have undergone a revolution in the last decade. 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 backlit 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.

Computer-generated imagery offers an effective means for presenting three-dimensional medical data to the clinician. Researchers have, over a period of 15 years, built up a large repertoire of techniques for visualizing multidimensional information. Applying these techniques in



Though several techniques for displaying multidimensional data are being developed, research suggests that real-time volume rendering will be preferred for many clinical applications.

clinical settings has until recently been limited, partly because of low image quality and partly because of 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 to 10 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. Most techniques for displaying three-dimensional medical data fall into the three broad categories of surface-based rendering, binary voxel rendering, and volume rendering. The characteristics of each category depend on the intermediate representation employed.

Surface-based techniques. The first category, surface-based techniques, applies a surface detector to the sample array, then fits geometric primitives to the detected surfaces, and finally renders 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 applies 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.¹ Figure 1 was generated using this technique. Alternatively, polygons can be fit to an approximation of the original anatomy within each voxel (a pixel in 3D), yielding a large set of voxel-sized polygons.²

These techniques have several desirable characteristics. Geometric primitives are compact, making storage and transmission inexpensive. They also exhibit 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. These 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.³ Alternatively, voxels can be painted directly onto the screen in back-to-front (or front-to-back) order,⁴ or rays can be traced from an observer position through the data, stopping when an opaque object is encountered.⁵

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.

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. (CT data courtesy of NC Memorial Hospital)



Volume-rendering techniques. These 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, semitransparent 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 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.⁶ Figures 2-4 were generated using an algorithm developed at the University of North Carolina. Though similar in approach to the Pixar technique, the algorithm 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 data set.⁷

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 data set. Published techniques take minutes or even hours to generate a single view of a large data set 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 nonempty (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 data sets.⁸

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



Figure 2. Volume rendering of $256 \times 256 \times 113$ -voxel computed tomography study of a human head. A polygoneal 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. (CT data courtesy of NC Memorial Hospital)

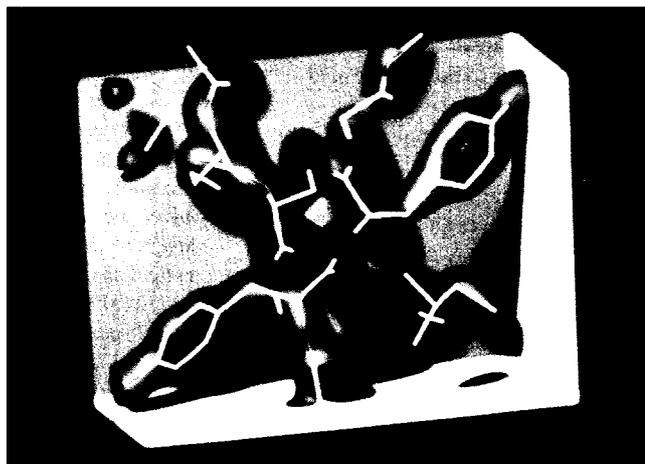


Figure 3. Volume rendering of isovalue contour surface from $24 \times 20 \times 11$ -voxel electron density map of staphylococcus aureus ribonuclease. 3D scan-converted backdrop planes, cast shadows, and a color-coded stick representation of the molecular structure have been added to enhance comprehension. Similar techniques can be used to render isodose contour surfaces for radiation treatment planning. (Electron density map courtesy of Chris Hill, University of York)

arrays. Figure 2 was generated using this algorithm. The second consists of a 3D scan conversion of the polygons into the sample array with antialiasing, then rendering the ensemble. Figure 3 includes polygons rendered using this algorithm.⁸

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 magnetic resonance 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 data set to remove the scalp as shown in Figure 5. 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

Generating a two-dimensional image from a three-dimensional data set necessarily reduces information content. In particular, images do not use many faculties of the human visual system specifically adapted for perceiving depth: stereopsis, head motion parallax, and the kinetic depth effect. Such specialized display devices as stereo viewers, varifocal mirrors, cine sequences, real-time image-generation systems, and head-mounted displays have been developed to convey as

much three-dimensional information as possible to the observer.

Stereo viewers. Relatively simple and inexpensive, stereo viewers offer a 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. By combining a vibrating reflective surface with a point-plotting CRT, the varifocal mirror provides "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



Figure 4. Volume rendering of edited magnetic resonance data set. 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. (MR data courtesy of Siemens AG, edited by Juiqi Tang)

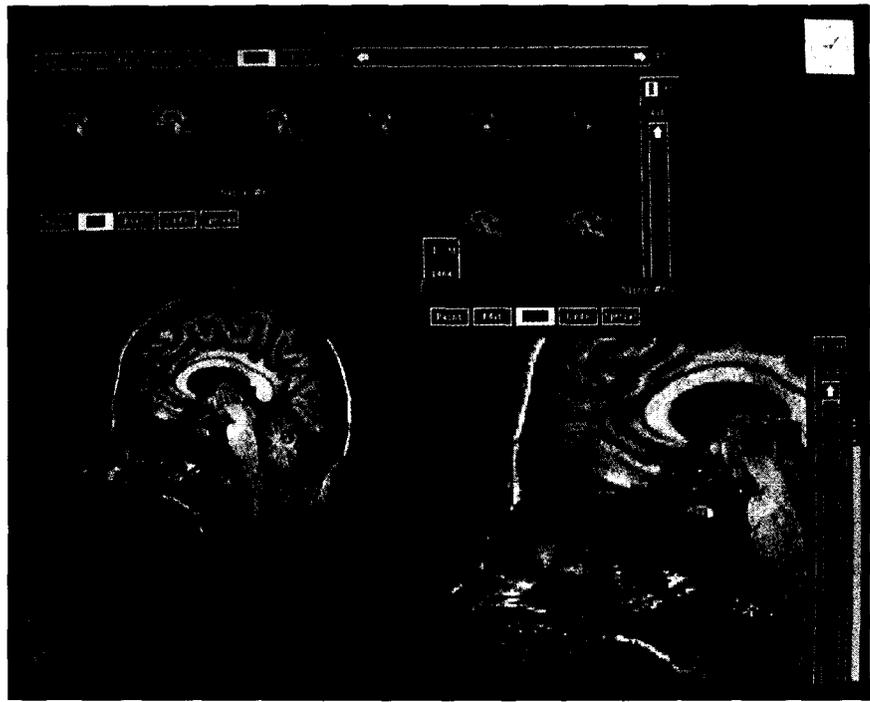


Figure 5. Slices from a $256 \times 256 \times 156$ -voxel magnetic resonance study of a human head being edited to remove the scalp and thereby reveal the underlying cortical surface. Depicted is a typical view of the workstation screen in our interactive Imex (image executive) medical-image-editing program, written by Peter Mills.

vibrated from behind by a large conventional speaker driven by a low-frequency (30 Hz) sine wave. A user looking into this vibrating (and "varifocal") mirror sees 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 30 Hz 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 a user moves his head relative to the display, he sees different views of the data.

This display device provides both stereopsis and head-motion parallax. In addition, it can be viewed simultaneously by several users without their wearing any special apparatus. The limited size of the displayed volume, which is constrained by the size of the CRT and the deflection of the mirror membrane, presents a major drawback. Another disadvantage is the lack of obscuration by the displayed dots; features in the foreground do not hide features in the background.

Cine sequences. Sets of images that have been precalculated, stored in a frame

buffer of sufficient memory capacity, and played back in real-time are called cine sequences. 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 such as observer position, light source position, cutting plane position, or opacity mapping. Users would often prefer to vary multiple parameters. Unfortunately, the

number of images that must be precalculated to provide such multiaxis 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 require 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 $64\,512 \times 512$ 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. Offering more flexibility than cine sequences for the presentation of moving imagery, real-time image-generation systems have found broad acceptance. 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 $512 \times 512 \times 512$ voxel binary volume in real time.⁹

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, single-instruction, multiple-data processor and Pixar's ChapVolumes software package, the Image Computer can generate high-quality images of $256 \times 256 \times 256$ -voxel data sets 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 16 40-Mflop graphics processors, 250,000 pixel processors, a 512×512 pixel color frame buffer, and a 640 megabits-per-second ring network.¹⁰ 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 solely in observer position, we will be able to generate a sequence of slightly coarse images at between 10 and 20 frames per second. A sequence of images comparable in quality to Figure 4 will be generated at between one and two frames per second.

Head-mounted displays. Pioneered by Ivan Sutherland at the University of Utah in 1968, head-mounted displays 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.

These systems demand very fast update rates and 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

A vital 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 metavisualizations such as 2D plots of color and opacity versus input value and wire-frame representations of viewing frustums and motion paths. If future systems succeed in generating volume-rendered images in real time, these metavisualizations 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 data sets. 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 data set to surface relief and another to surface albedo or color. For three-dimensional data, the second data set can be used to modulate the color or opacity computed for voxels in the first data set. Alternatively, the second data set could be used to perturb the normals, shininess, or other properties of the surfaces displayed from the first data set. Recent investigations in this area by Hoehne et al.¹¹ and workers at the University of Chicago¹² show promise.

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, relative to surface-based techniques, has the disadvantage of rendering fuzzy surfaces represented by a band-limited sample array rather than sharply defined surfaces represented by geometric primitives. Since fuzzy surfaces do not occur in daily life, they are 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 judgments of feature size and shape, the potential exists for making erroneous diagnoses based on these visualizations. Clearly, volume rendering is useful for navigating through complex 3D data and for identifying gross anatomic abnormalities. It remains to be seen how useful it can be for perceiving fine features and to what extent it can replace slice-by-slice presentation of medical data in clinical practice. □

Acknowledgments

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 No. 1 PO1 CA47982-01, NSF Grant No. CDR-86-22201, DARPA ISTO Order No. 6090, ONR Contract No. N00014-86-K-0680, and IBM.

References

1. H. Fuchs, Z.M. Kedem, and S.P. Uselton, "Optimal Surface Reconstruction from Planar Contours," *Comm. ACM*, Vol. 20, No. 10, Oct. 1977, pp. 693-702.
2. W.E. Lorensen and H.E. Cline, "Marching Cubes: A High-Resolution 3D Surface Construction Algorithm," *Computer Graphics (Proc. SIGGRAPH)*, Vol. 21, No. 4, July 1987, pp. 163-169.
3. G.T. Herman and H.K. Liu, "Three-Dimensional Display of Human Organs from Computer Tomograms," *Computer Graphics and Image Processing*, Vol. 9, No. 1, Jan. 1979, pp. 1-21.
4. G. Frieder, D. Gordon, and R.A. Reynolds, "Back-to-Front Display of Voxel-Based Objects," *IEEE Computer Graphics and Applications*, Vol. 5, No. 1, Jan. 1985, pp. 52-59.

5. D.S. Schlusserberg and W.K. Smith, "Three-Dimensional Display of Medical Image Volumes," *Proc. NCGA 86*, Nat'l Computer Graphics Assn., Fairfax, Va., Vol. III, May 1986, pp. 114-123.
6. R.A. Drebin, L. Carpenter, and P. Hanrahan, "Volume Rendering," *Computer Graphics* (Proc. SIGGraph), Vol. 22, No. 4, Aug. 1988, pp. 65-74.
7. M. Levoy, "Display of Surfaces from Volume Data," *IEEE Computer Graphics and Applications*, Vol. 8, No. 3, May 1988, pp. 29-37.
8. M. Levoy, *Display of Surfaces from Volume Data*, doctoral dissertation, University of North Carolina at Chapel Hill, May 1989.
9. A. Kaufman and R. Bakalash, "Memory and Processing Architecture for 3D Voxel-Based Imagery," *IEEE Computer Graphics and Applications*, Vol. 8, No. 6, Nov. 1988, pp. 10-23.
10. H. Fuchs et al., "Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor-Enhanced Memories," *Computer Graphics* (Proc. SIGGraph), Vol. 23, No. 4, July 1989.
11. K.H. Hoehne et al., "Display of Multiple 3D-Objects Using the Generalized Voxel-Model," *Proc. SPIE*, Vol. 914, 1988, pp. 850-854.
12. D.N. Levin, et al. "Integrated 3D Display of MR, CT, and PET Images of the Brain," *Proc. NCGA 89*, Nat'l Computer Graphics Assn., Fairfax, Va., Vol. I, pp. 179-186.



Marc Levoy is a research assistant professor of computer science and radiation oncology at the University of North Carolina at Chapel Hill. He is the principal developer of the Hanna-Barbera Computer Animation System and served as its director from 1980 through 1982. His research interests include scientific visualization, texture mapping, volume rendering, medical imaging, and molecular graphics. He received his bachelor's and master's degrees from Cornell University in 1976 and 1978 and his PhD in computer science from the University of North Carolina at Chapel Hill in 1989.



Stephen M. Pizer is a professor of computer science, radiation oncology, and radiology at the University of North Carolina at Chapel Hill, where he heads the multidisciplinary Medical Image Display Research Group. His research covers human and computer vision, interactive 3D graphics, contrast enhancement, and workstation development. He conducts active collaborations with laboratories in the Netherlands, Germany, and the US, and is associate editor for display of *IEEE Transactions on Medical Imaging*. He received a PhD in computer science from Harvard in 1968.



Henry Fuchs is the Federico Gil Professor of Computer Science and an adjunct professor of radiation oncology at the University of North Carolina at Chapel Hill. His research interests include interactive 3D graphics, VLSI architectures, and 3D medical imaging. The principal investigator for several research projects funded by the Air Force, DARPA/ISTO, NIH, and NSF, he also served as chairman of the 1985 Chapel Hill Conference on VLSI and the 1986 Workshop on Interactive 3D Graphics held at UNC-Chapel Hill. He is on several advisory committees, including NSF Division of Microelectronics Information Processing Systems and Stellar Computer's Technical Board. He received a BA from the University of California at Santa Cruz in 1970 and a PhD from the University of Utah in 1975.

All three authors can be contacted at the Department of Computer Science, Sitterson Hall, University of North Carolina, Chapel Hill, NC 27599-3175

August 1989



Universität
Gesamthochschule Paderborn

On the basis of agreements between the Westphalia Foundation, the State of Northrhine-Westphalia and the University of Paderborn.

The Heinz-Nixdorf-Institute

Interdisciplinary Research Centre for Information Technology (CIT)
is being created within the University.

The task of **CIT** is to strengthen teaching and research at the university in application-oriented areas of informatics and technology including scientific foundations and topics from the fields of the arts and social sciences. **CIT** is a central academic institute of the University. The position is a full professorship integrated into the Department involved and included in the curriculum of relevant courses of study. It is the aim of **CIT** to stimulate innovation in research by means of joint projects between researchers from different fields. The affiliation of a centre for postgraduate research is one of the components of the agreement between the Westphalia Foundation and the State of Northrhine-Westphalia.

The selection of applicants will be undertaken by the Department in cooperation with the Westphalia Foundation and the Board of Founding Directors of **CIT**. Candidates are expected to show particular willingness to cooperate in joint projects between different fields.

The foundation period as planned requires the following position to be filled at the next possible date:

Department 17 - Mathematics/Informatics

A University Chair (Grade C4) in Practical Informatics,
particularly Practical Studies in Distributed Systems

Applications should be qualified in one of these fields: "Computer Networks/Operating Systems", "Software Engineering for Distributed Systems" or "Distributed Databases". Within the framework of the research topics of **CIT**, the successful candidate will be expected to pay particular attention to problems involved in the development of distributed systems. Practical experience in the implementation of relevant projects is an advantage.

Required qualifications:

Section 49 para. 1 no. 4a, WissHG NW (University Law of Northrhine-Westphalia): Professorial thesis or equivalent academic/scientific qualification inside or outside university.

Application deadline: 6 weeks after publication.

The University of Paderborn is seeking to increase the proportion of women in university teaching positions and expressly requests women to apply.

Severely handicapped applications with equivalent suitability to others will be given priority.

Applications are invited to be presented with the usual documents (curriculum vitae, certificates, list of publications in major scientific periodicals) stating the code number 403 to the dean of the Department 17 - Mathematics/Informatics, University of Paderborn, Postfach 1621, D-4790 Paderborn.