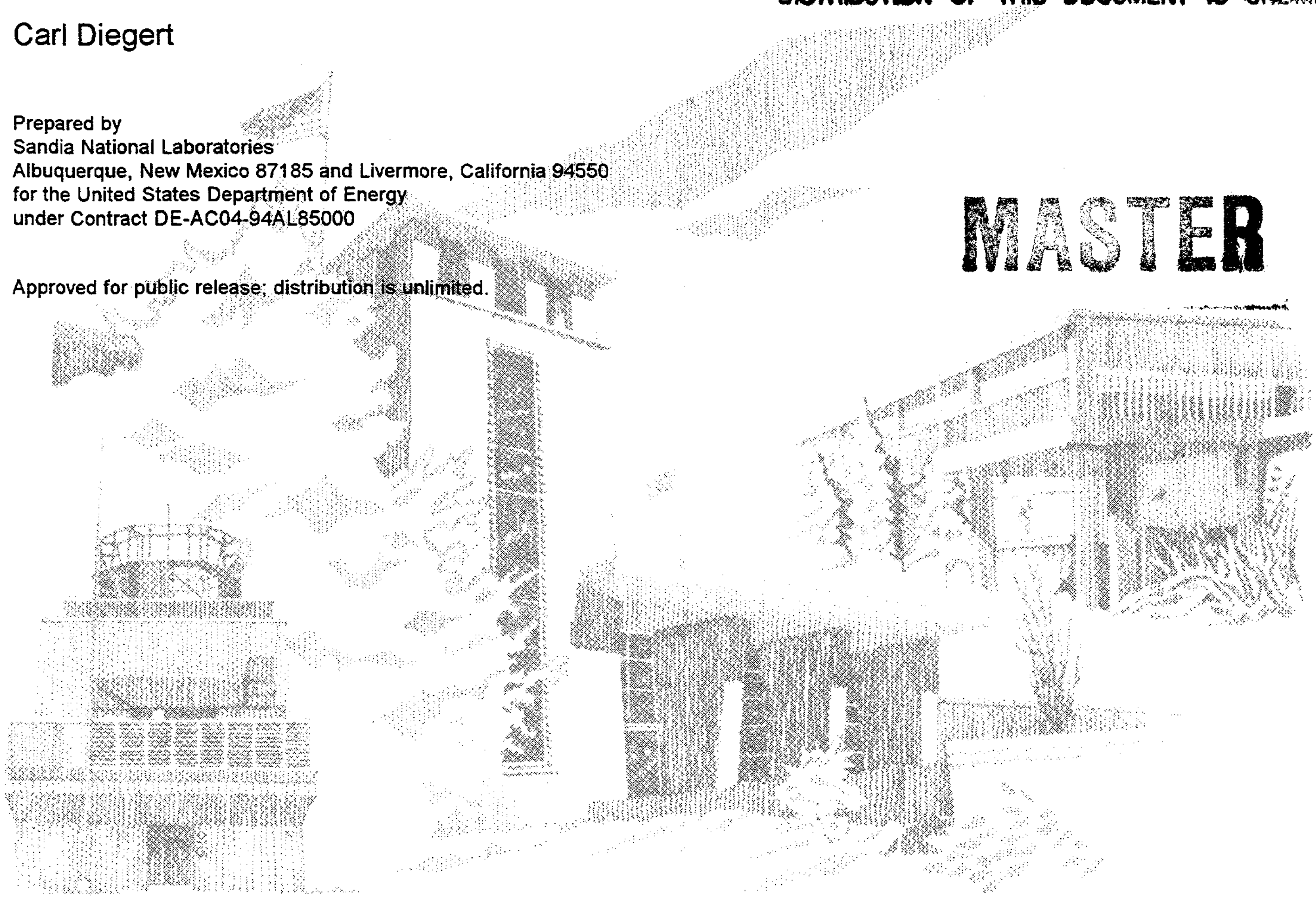
SANDIA REPORT

SAND97–0251 • UC–405
Unlimited Release
Printed January 1997

Innovative Computing for Diagnoses from Medical, Magnetic-Resonance Imaging

Carl Diegert

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-94AL85000



SF2900Q(8-81)

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A03
Microfiche copy: A01

DISCLAIMER

Distribution
Category UC-405

Sand97-0251
Unlimited Release
Printed January 1997

Innovative Computing for Diagnoses from Medical, Magnetic-Resonance Imaging

Carl Diegert
Computer Architectures Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-0318

Abstract

We present a final report on a Laboratory-Directed Research and Development (LDRD) project, *Innovative Computing for Diagnoses from Medical, Magnetic-Resonance Imaging*, performed during fiscal years 1992 and 1993. The project defined a role for high-performance computing in surgery: the supercomputer can automatically summarize the three-dimensional extents of lesions and other clinically-relevant structures, and can deliver these summaries to workstation-based, augmented-reality environments at the clinical site. We developed methods and software to make these summaries from the digital data already acquired using clinical, magnetic-resonance machines. In joint work with Albuquerque's Department of Veterans Affairs Hospital, we applied this work, and obtained a basis for planning, for rehearsal, and for guidance during surgery

Introduction

New capacities for acquiring, processing and communicating digital data have already begun a revolution in surgery. The new focus is on a broad, and ever widening set of digital data to more completely describe the patient, the disease, the machines in the operating room, the surgical team, and other hospital processes. The result can be more effective intervention, and reduced net cost of intervention. Already, effective intervention often cannot be completed without first at least obtaining digital information from automated "scans" of a patient before surgery. Successful outcomes increasingly depend on managing and exploiting an exploding base of digital data.

With the new focus on large-scale, digital datasets, come abundant roles for automatic computation. Here we describe one role: how automatic computation can extract summary information that is both manageable in the operating room, and can improve a surgeon's speed and accuracy. More specifically, we outline a role where we apply a high-performance computer to obtain summaries of the three-dimensional extent of lesions, and other clinically-relevant structures. We then deliver these summaries to workstations at the clinical site in time to provide a basis for planning, for rehearsal, and for guidance during surgery.

Today, the computational role we describe demands high-performance computer hardware. We chose a role that does not demand that the supercomputer be online during surgery. The role's process flow demands only that the supercomputer receive, process, and return data, all within a few hours. Present, sub-gigabit computer networks are enough to allow a supercomputer remote to the clinical site to fill the role.

The marginal cost of using today's supercomuters and networks to enhance treatment of at least some patients may already be justified by lowered costs of surgery, and of better outcomes. But there is a more compelling reason to begin experimental clinical use now, despite the trouble and expense of using remote supercomputers. Within a year or two, technological advance and volume production will drive the cost of computing hardware for this role down to levels that allow economical deployment of the hardware at, say, the thousands of sites that already do clinical imaging. By getting through development and clinical trials now, the enhanced treatment can be ready for wide deployment as soon as decreased computing hardware costs make it cost effective. From a business perspective, projecting the cost of future computing hardware is low risk compared to the risk in predicting clinical utility. Clinical experience now can reduce the risk in business plans, allowing industry to have product ready to exploit new computing hardware soon after it becomes available.

The conclusion of this paper shows that an effective way to implement and deliver the computation is to code using explicit message-passing, and to execute the code on a distributed-memory computer.

Improving the surgeon's accuracy

We will present the computational role by beginning with the problem, and concluding with the computation and its implementation.

A case study illustrates the problem. A central surgical goal is to quickly and completely resect certain brain tissue, while minimally disturbing other tissue. Information observed by prior magnetic-resonance (MR) acquisitions is indispensable in planning and guiding the intervention. The problem is that the surgeon and radiologist work from a less-than-ideal presentation of a small subset of the MR data. The problem is even greater in a typical surgery because information from multiple MR acquisitions, as well as information from new types of functional imaging all should help plan and complete the intervention.

About two hours into the five-hour surgery, the surgeon has exposed the brain, and must decide where to make an initial incision with his scalpel. Figure 1a shows how the patient presents for this surgery. Figure 1b is the surgeon's view of the patient just before making an initial incision.

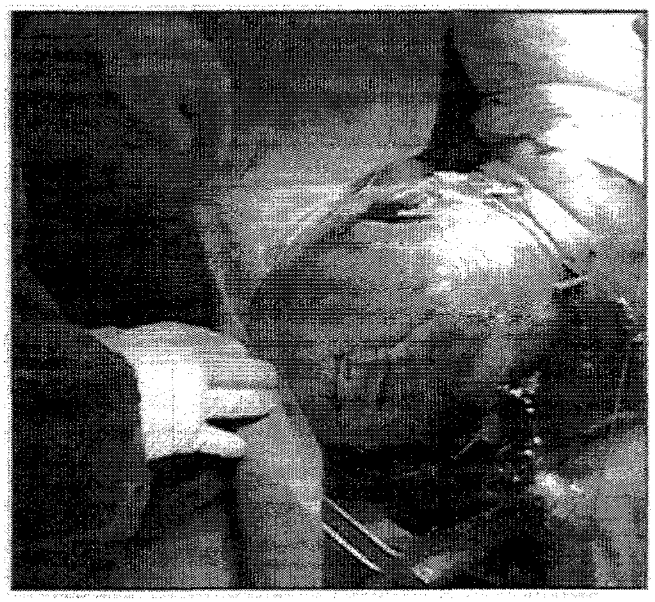


Figure 1a. The patient at the start of a five-hour surgery to remove a brain lesion. An electronic version of this paper, including all figures in full color, is at http:www.cs.sandia.gov/~diegert.

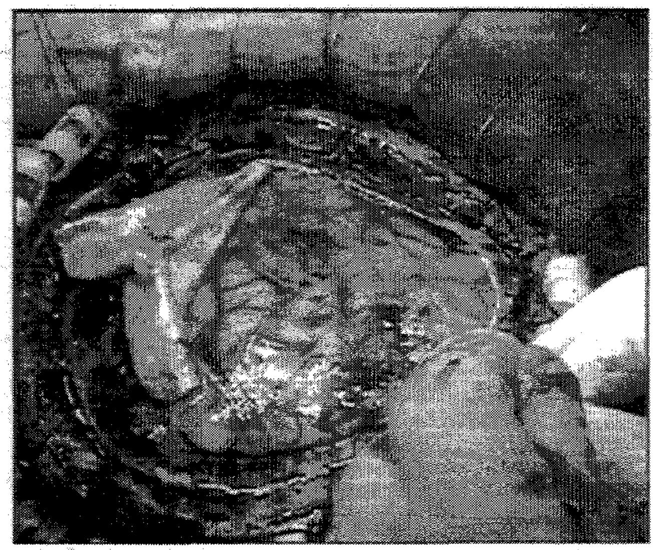


Figure 1b. About two hours into the surgery, the surgeon has exposed the brain, and must decide where to make an initial incision.

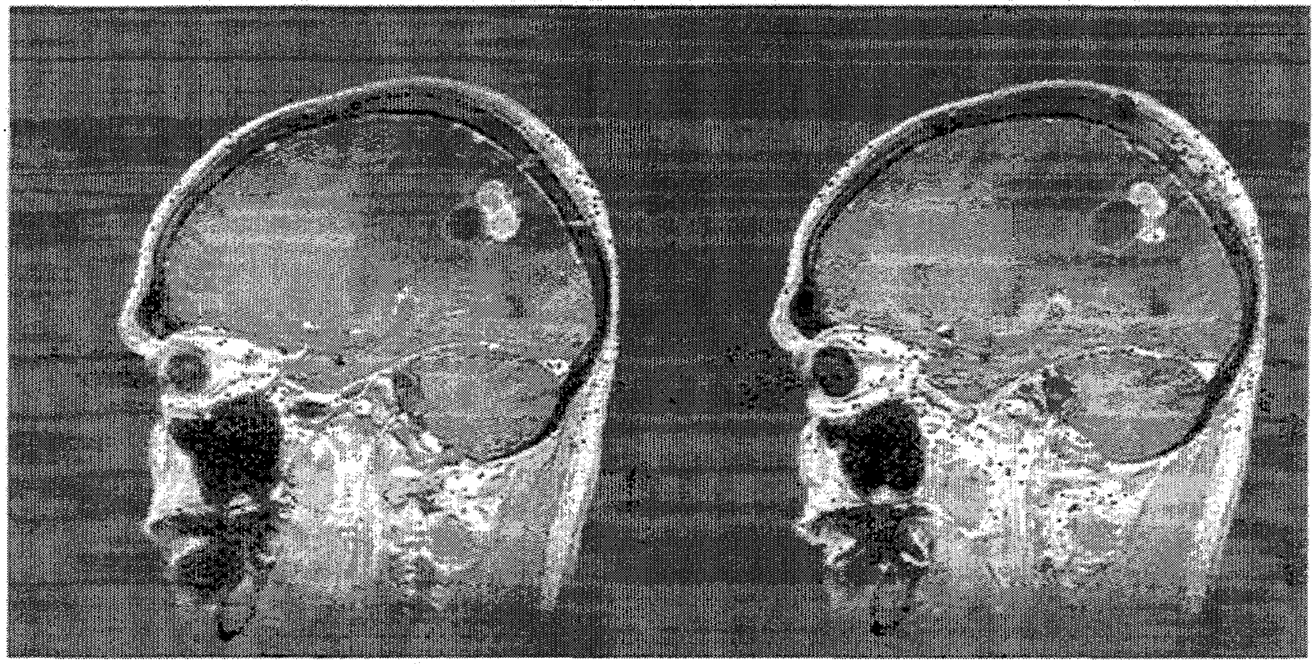


Figure 2a. *Figure 2b.*

Before making the incision, the surgeon completed a final review of the patient's MR data by stepping away from the patient, to stand beside five light boxes on the operating room (OR) wall. Each box displayed a black-and-white film mosaic of about 24 images. Figure 2a, a somewhat enhanced version of one of these images, shows a two-dimensional cross-section through the head in a plane perpendicular to the ear-to-ear axis. At the back of the head (at about 2:00 o'clock) are well-defined, adjacent bright and dark regions, together comprising a lesion, the tissue to be removed. Figure 2b is another image from the same volume acquisition, sliced perpendicular to a point a few millimeters closer to the patient's right ear. Other images (not presented here) show location of critical brain function, tissue to leave undisturbed. All of the images on the light boxes are two-dimensional sections. Some images show the same volume observation as figure one, but sliced at additional points along the ear-to-ear axis. Others are from acquisitions tuned to better show other features, acquisitions by other modalities (computed tomography X-ray and magnetic source imaging), and acquisitions taken at prior dates. The multiple acquisitions are taken at different sampling resolutions, and taken with the sampling lattice aligned with different axes. Understanding changes and other relations among these data demands skill and careful study.

Only a tiny fraction of the imaging acquisitions entered into the planning and completion of this surgery. In addition to the images on the five light boxes, hundreds more are on films in a stack of envelopes. Hundreds more were never printed to film -- medical technicians chose only a few "best" images from each volume acquisition to print to film. However, the entire digital dataset is archived on computer media in the hospital's radiology department. The radiologist and surgeon did make at least some use of these additional films in planning the surgery, but not during surgery. Access to the patient's complete digital database was impractical while planning the surgery, and was impossible during surgery.

Relevant, usable summary information

A role for high-performance computing is to give the surgeon a more suitable presentation of the digital database. The computation provides two opportunities for improvement. First, by aggressively summarizing the whole digital dataset, we seek to provide summaries that include more relevant information, dramatically more than can be presented in a small collection of "best" image slices. Second, our presentation must actually allow a human surgeon to improve his speed and accuracy.

Applying high-performance computing, as described below, we can deliver a compact representation of the three-dimensional extent of a lesion to a workstation in the operating room. In this computation, we load all of the digital information from a MR acquisition into the memories of a high performance computer, then find the extent of a lesion (or other structure of interest) in three dimensions. The computation and result are at the full resolution of the MR acquisition. The summary product includes only the geometric extent of the lesion. Substituting the summary product for the entire MR acquisition discards most of the information in the acquisition, and retains just a compact summary of the relevant information. With the reduced dataset size, a workstation is all that is needed to manage the summaries in the OR.

The form of the data is also changed when the summary is substituted for the MR acquisition. When the summary correctly defines the clinically-relevant extent of a lesion, we are adding value (for use in surgery) to the acquisition data. The segmentation algorithms that make this summary must be continually tuned and checked. Therefore, the computation cannot add value without demanding some continued clinical expertise. We have been fortunate to collaborate with Dr. Sanders (MR specialist), Dr. Orrision (radiologist), and Dr. Baldwin (surgeon), and have added our expertise in statistics and high-performance computing, to experiment with this new computational role.

A workstation can easily load our summary information, and can

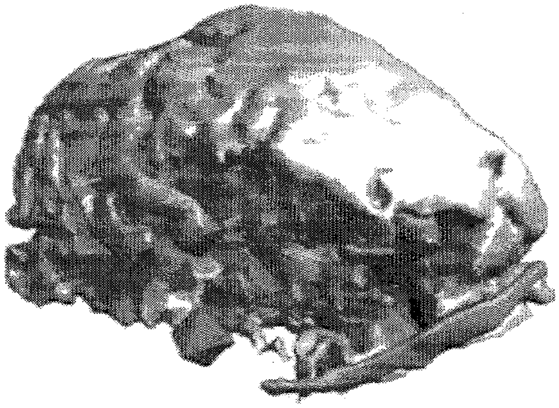


Figure 3a.

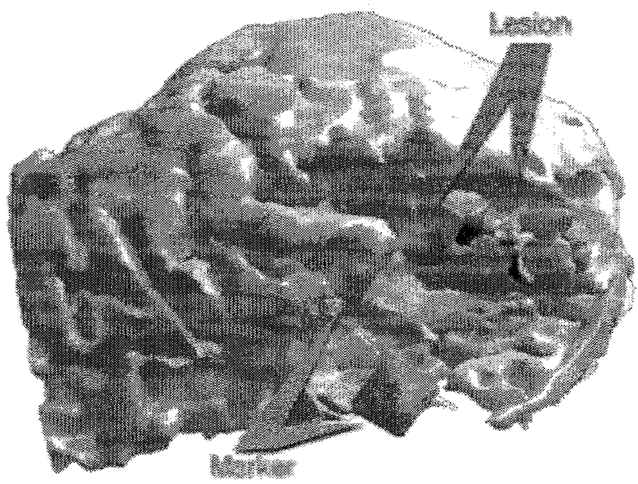


Figure 3b.

Two views, rendered on a workstation from the supercomputer's summary information. The summary estimates the extent of a lesion, the extent of healthy brain tissue (this tissue is severely compressed by swelling at the lesion site), and the extent of the deep vein. The two octagon shapes mark the location of brain function.

generate realistic images of the lesion and related structures from any viewing position. Two views are shown in Figures 3a and 3b. These views may better convey the shape of the lesion to the surgeon, but are probably not sufficient to dramatically improve the accuracy of locating his incision.

The final step to clinical impact

The second opportunity for high-performance computing is to provide the reduced data that can actually improve the speed and accuracy of a human surgeon. The groundwork for taking this step has been realized in a research protocol now in use at the Albuquerque Department of Veterans Affairs Medical Center. The groundwork comprises two components. First, a 3D tracking device in the OR makes the measurements that can align prior datasets with the patient, in the position he assumes during surgery. Second a workstation in the OR completes intraoperative computation.

The workstation is loaded with programs and data, then disconnected from the Internet, and rolled into the OR before surgery (Figure 4). The tracking system (Pixsys, Boulder CO) is also moved to the OR before surgery, and is connected to the workstation. With this innovative setup, the surgeon uses a sterile probe equipped with infrared marker lights (part of the tracking system) to point to a few fiducial points on the patient, sending three-dimensional location of these points to the workstation. Workstation software, then, uses the fiducial information to establish a map between the coordinate system used by the surgeon's

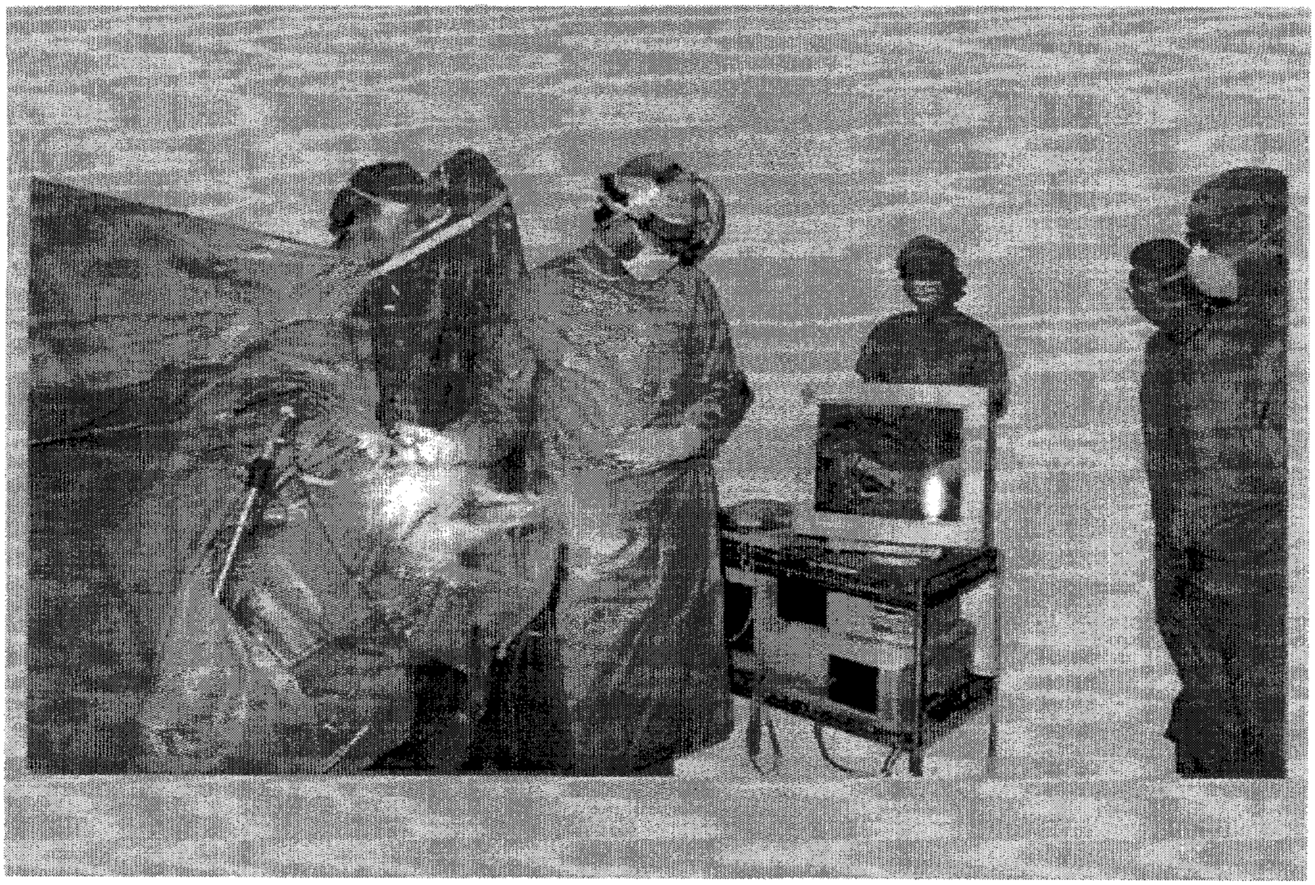


Figure 4. Dr. Baldwin consults a workstation display during a surgery at the Albuquerque Department of Veterans Affairs hospital. The cart also carries a tracking system. We altered this image to show the display discussed below, an image Dr. Baldwin actually reviewed after completing the surgery.

pointing device, and the coordinate system used in the prior MR acquisition. Using the map, the surgeon can point to any location, and automatically call up an annotated display of the three orthogonal "slices" of the MR acquisition that intersect this point.

By adding a video camera and workstation video interface to the equipment already in the OR, the surgical team can obtain a promising display from our three-dimensional summary information, as shown in Figure 5b.

When the surgeon can obtain enhanced reality displays like Figure 5b on a workstation display screen in the OR (Figure 4), his accuracy and speed should show a dramatic improvement over the present standard of care, where he must rely on the films (Figure 2). Once the utility of these displays is established, the incremental improvements of substituting a heads-up display for the workstation screen may be worth some additional disruption to the clinical procedures. We anticipate using a scanned, visible laser mounted on the OR ceiling (together with the 3D tracking cameras) to "write" guiding marks directly on the patient.

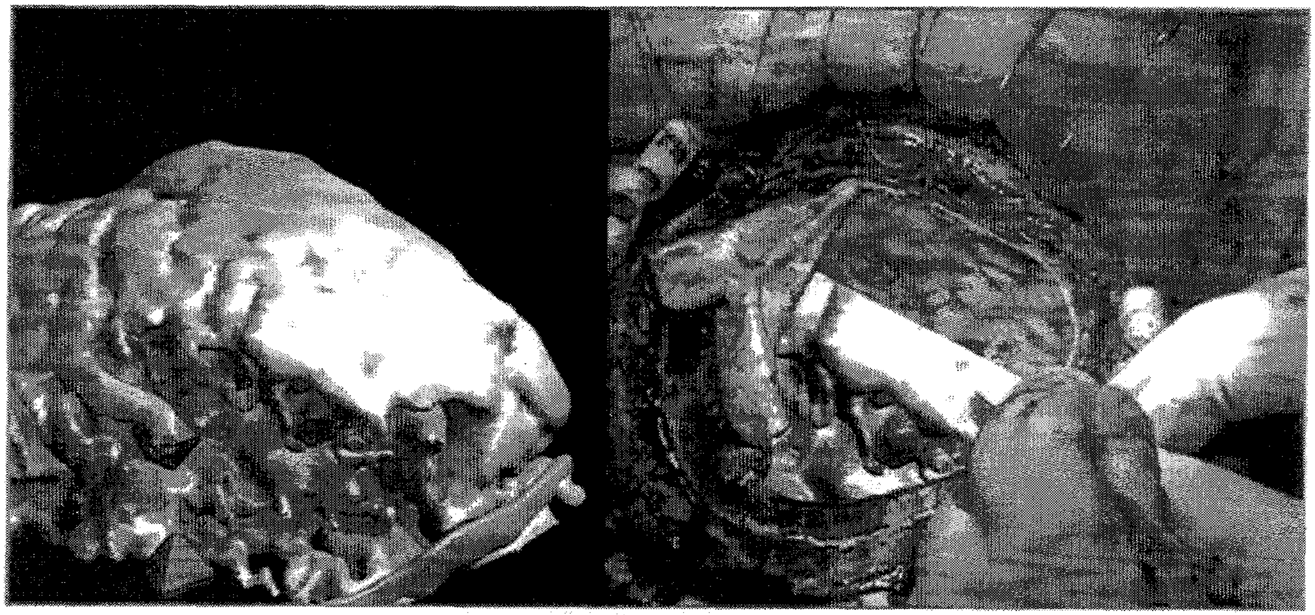


Figure 5. A tracking system in the OR can observe the particular viewpoint (and other parameters) of an intraoperative camera. The left picture (5a) is a rendering of the summary surfaces from the camera's viewpoint. The right picture (5b)is a composite of the camera's optical image (shown in Figure 1b), and the rendered image 5a. In the composite, some of the optical image has been painted away, revealing the location of the lesion underneath. In this case, the lesion's location relative to the bone cut could help locate the surgeon's incision.

Summary models are key to other roles

In the high-performance computing role just described, we use the summary models to improve the surgeon's speed and accuracy, as the models allow a workstation in the OR to guide the surgeon with enhanced-reality images. The key technology behind making the enhanced-reality pictures is an ability to segment and summarize the clinical MR data, producing compact, clinically-relevant summary models.

The same technology for making summary models can allow surgical rehearsal, training, and numerical simulations. To deliver a synthetic rehearsal environment using both a head-tracking visual display and a haptic manipulation device, we must generate both images and forces at interactive update rates. Without compact models, updates rates will be too slow to achieve useful fidelity. For rehearsal, we must extract the models after the patient's last MR acquisition, and in time to build the rehearsal environment before his surgery. High-performance computing offers speed and a potential for automation to meet this demanding schedule. Finally, numerical simulation involving patient anatomy depends not just on the 3D extents of the anatomy, but also on computational meshes on these regions. Our work on extracting summaries suitable for automatic mesh generation and subsequent numerical simulation are well underway.

Parallel implementation

The image segmentation and summary for the surgical role is still best accomplished on high-performance, general-purpose computers, where flexible segmentation and image generation are completed at interactive speeds. In contrast, it takes tens to hundreds of minutes for high-end workstations, equipped with about one-half gigabyte of main memory, to complete most segmentation and summary computations on single acquisitions. This is too slow to support the interactive data exploration that we still need to refine and tune the algorithms.

Our implementation delivers the required interactive performance by using explicit message passing software, executing on a scalable, distributed-memory computer. In fact, by executing on computers with more processors, the timings scale to execute much faster than demanded in this role.

For the specific computation timed in Table 1, we have a single MR acquisition from the case just presented loaded into the distributed memories of an Intel Paragon XPS system. The data are represented as a regularly-sampled lattice of 32-bit floating-point numbers with 256 by 256 by 128 entries. (Figure 1 shows two slices from this acquisition.) Our code, VENGINE, builds on a portable, massively-parallel software library MPSLIB, written by Mark Sears (Mail Stop 1111, Sandia National Laboratories, Albuquerque NM). Mark wrote MPSLIB principally for completing electronic structure calculations. A secondary goal was to make

TABLE 1.
Mean Milliseconds Execution (+/- Gives Range of Two Measurements)
Speedup Percent (Relative to 32-Node Timing)
Efficiency Percent (Relative to 32-Node Timing)

Nodes	32	64	128	256
Trimmed Filter	2,710 +/- 3	1,365 +/- 3	694 +/- 1	357 +/- 1
	100	199	390	759
	100	99	98	95
Extract Geometry	2,242 +/- 1	1,217 +/- 0	659 +/- 0	361 +/- 0
	100	184	340	621
	100	92	85	78
Generate Image	2,074 +/- 1	1,428 +/- 1	979 +/- 1	697 +/- 1
	100	145	212	298
	100	73	53	37

MPSLIB useful in other applications, a goal he demonstrated both by broadening the library to better support VENGINE, and by

An early step in exploring noisy MR data is to tune the parameters of a 3D, nonlinear filter, apply the filter, and observe the effect on a subsequent extraction of 3D surfaces. The time to a new image in this interactive exploration is the sum of three times: filtering, surface extraction, and image generation. Table 1 shows a range in time-to-image from seven seconds down to 1.4 seconds, as executed by our code VENGINE (version of Jan-30-1995) on Paragon systems with between 32 and 256 processing nodes.

For the first step, input parameters for the trimmed filter are a threshold scalar value and a self-weight value. For each voxel, if the voxel value is below threshold, its value is unchanged by the filter. Otherwise, the filter considers the voxel and its six neighbors. If three or fewer neighbors are over threshold, the voxel value is unchanged by the filter. Otherwise, the filter replaces the voxel value with a weighted mean of itself (with weight self-weight), and the mean of a subset of the neighbor voxel values that are over threshold (with the complimentary weight). The subset is formed by discarding a highest and a lowest value.

The efficiency of the parallel filter implementation decreases slightly as more processors are used because the amount of information we must share between processors to allow access to neighbor voxel values becomes relatively more burdensome. The table shows that filtering of just a single MR acquisition is a large enough computation relative to this burden to allow extremely high efficiency on Paragon systems as large as 256 processing nodes.

The surface extraction step timed in Table 1 is a table-driven marching cubes algorithm, with two input parameters. The first input parameter is a fuzz value, to suppress generation of triangles too small to be useful. The second is the value of the isosurface to be approximated by the output triangle list. The execution timed in the table produces a list of 1,142,256 triangles.

A more sophisticated extraction algorithm, or further reduction of marching-cubes triangle list is needed before the summary is useful on the workstation in the OR. However, for exploration using a high-performance computer, working directly with the large triangle lists is often more effective than complicating the interaction by demanding compact models. Thus for exploration, the final step to an image is to just shade all of the triangles. The timings are for computing simple, flat-shaded triangles. They do not even benefit from an additional step of balancing the computation: the processor that generates the triangle in the last, surface-extraction step just shades it. The execution time for image generation also includes forming a composite image from all of the triangles, where the image is itself distributed across many processing nodes. The efficiency of image generation could be improved on the larger Paragon systems by adding code to dynamically balance the computation.

To best show how a current-generation Paragon system treats the real, clinical problem, Table 1 times processing of the fixed-sized problem. Efficiencies, of course, would be much higher if we scaled up the size of MR acquisitions to better fill the systems with larger numbers of processing nodes. This scalability of our implementation will be important in the future, as clinical instruments acquire data with more resolution, and with more attributes.

The next step

To date, we have developed the high-performance computing role without influencing patient outcomes. We did work with real patient cases, and with the clinical people treating the patient. We did not, however, make our results available in time to influence patient outcomes. Working in this mode was effective in focusing our development, and in demonstrating promising results. While we can continue to refine the computing role without influencing patient outcomes, we cannot directly measure the effectiveness of the computational role in this mode. Effective computation will not deliver benefits to health care unless it is used to treat patients.

DISTRIBUTION:

5 MS 1109 Carl Diegert, 9215

1 MS 0318 George S. Davidson, 9215

1 MS 0188 C. E. Meyers, 4523

1 MS 0188 Donna Chavez, 4523

1 MS 9018 Central Technical Files, 8940-2

5 MS 0899 Technical Library, 4414

2 MS 0619 Review and Approval Desk, 12690
For DOE/OSTI