A Collaborative Virtual Environment for the Simulation of Temporal Bone Surgery

Dan Morris[1], Christopher Sewell[1], Nikolas Blevins[2], Federico Barbagli[1],
Kenneth Salisbury[1]

[1] Stanford University, Department of Computer Science
[2] Stanford University, Department of Otolaryngology

{dmorris, csewell, barbagli, jks}@robotics.stanford.edu,
nblevins@stanford.edu

Robotics Laboratory
Gates Building 1A
Stanford CA 94305-9010, USA

Abstract. We describe a framework for training-oriented simulation of temporal bone surgery. Bone dissection is simulated visually and haptically, using a hybrid data representation that allows smooth surfaces to be maintained for graphic rendering while volumetric data is used for haptic feedback. Novel sources of feedback are incorporated into the simulation platform, including synthetic drill sounds based on experimental data and simulated monitoring of virtual nerve bundles. Realistic behavior is modeled for a variety of surgical drill burrs, rendering the environment suitable for training low-level drilling skills. The system allows two users to independently observe and manipulate a common model, and allows one user to experience the forces generated by the other's contact with the bone surface. This permits an instructor to remotely observe a trainee and provide real-time feedback and demonstration.

1 Introduction

1.1 Temporal Bone Surgery

Several common otologic surgical procedures – including mastoidectomy, acoustic neuroma resection, and cochlear implantation – involve drilling within the temporal bone to access critical anatomy within the middle ear, inner ear, and skull base. As computer simulation is becoming a more frequently used technique in surgical training and planning, this class of procedures has emerged as a strong candidate for simulation-based learning.

The time spent on a procedure in this area is typically dominated by bone removal, which is performed with a series of burrs (rotary drill heads) of varying sizes and surface properties. Larger burrs are generally used for gross bone removal in the early part of a procedure, while smaller burrs are used for finer work in the vicinity of target anatomy. Surgeons employ a variety of strokes and contact techniques to precisely control bone removal while minimizing the risk of vibration and uncontrolled drill motion that could jeopardize critical structures. Drills are generally driven by pneumatic pressure, which is regulated by a floor-mounted footswitch. A combined

irrigation/suction instrument is typically used to keep the contact area moist and to remove accumulated blood and bone dust.

The primary risks to patient safety emerge from drilling in close proximity to the facial and vestibulocochlear nerves, which can be damaged as a result of vibration, heat, and direct contact with instruments. Electrodes are typically placed on the face and neck to monitor nerve activity; these electrodes are critical in allowing surgeons to locate and avoid injury to nerve bundles.

Surgery is typically performed using a binocular microscope. Some soft tissue work is required, including skin incision, muscle retraction, elevation of the tympanic membrane (eardrum), displacement of brain and nerve, and removal of lesions.

1.2 Current Training Techniques

Resident surgical training typically includes dissection of preserved human temporal bones. This allows residents to become acquainted with the mechanical aspects of drilling, but does not incorporate physiological information, continuous feedback for hazard avoidance, or soft tissue work. Temporal bone labs are also costly to maintain, and cadaver specimens can be difficult to obtain in sufficient quantity. This approach also limits the precision with which an instructor can monitor a trainee's drilling performance, as the instructor cannot feel the fine details of the trainee's interaction with the bone surface, and cannot easily share the drill and bone surface for demonstration. A further limitation of cadaver-based training is that instructors have little or no mechanism for controlling anatomic variations or the presence of specific pathology that can lead to challenging training scenarios.

Interactive atlases such as [6] are available for training regional anatomy. Two-dimensional simulations [3] are available for high-level procedure training.

2 Previous work

Previous work in interactive simulation of temporal bone surgery [2,4,9] has focused primarily on haptic rendering of volumetric data. Agus et al [2] have developed an analytical model of bone erosion as a function of applied drilling force and rotational velocity, which they have verified with experimental data. Petersik et al [9] model their drilling instrument as a point cloud, and use a modified version of the Voxmap-Pointshell algorithm [10] to sample the surface of the drill and generate appropriate forces at each sampled point. Each of these projects has incorporated haptic feedback into volumetric simulation environments that make use of CT and MR data and use volume-rendering techniques for graphical display.

Agus et al [1] describe several enhancements to their simulation environment that incorporate additional skills, including the use of irrigation and suction; and additional sources of intraoperative feedback, including real-time rendering of bone dust.

The system described in this paper draws on the haptic rendering algorithms presented in the above work. In order to extend training efforts beyond physical skills and to focus on clinical decision-making, we extend this work to allow tool-specific haptic rendering, networked interaction among users, realistic auditory feedback based on experimental recordings, and physiological monitoring of simulated nerves. Each of these features represents a key skill or source of feedback required for effective temporal bone surgery training and/or rehearsal. We also incorporate a representation of soft tissues that is not available in purely volumetric simulators.

3 Methods

3.1 Haptic rendering of volumetric data

We use a volumetric representation of bone for haptic rendering, which presents several advantages for drilling simulation. Volumetric data can be derived either from patient-specific CT data or from manually-assembled surface models [7], so this approach generalizes to a variety of data sources. A voxel representation also allows computationally inexpensive collision-detection for points (via direct indexing) and removal or density reduction of solid material. In our system, voxels are stored in a compact hash table; each voxel is associated with color, surface normal, and density information.

Virtual instruments are controlled using a SensAble Phantom [8] haptic feedback device, which provides three-degree-of-freedom force-feedback and six-degree-of-freedom positional input. Users can select from a variety of drills, including diamond and cutting burrs ranging from three to seven millimeters in diameter.

We adopt a haptic feedback approach similar to [9], in which the drill is represented as a cloud of sample points, distributed approximately uniformly around the surface of a spherical burr (we use 82 sample points). At each time step, each sample point is tested for contact with bone tissue. By tracing a ray from each immersed sample point toward the center of the tool, the system can generate a contact force that acts to move that sample point out of the bone volume. The net effect is a realistic sense of mechanical contact between drill and bone [9].

Our system associates a "drilling power" with each sample point; as each traced ray passes through a voxel, it removes an amount of bone density that depends on the drilling power of the sample point and the distance of the voxel from the instrument's main axis. These parameters allow us to simulate key aspects of drill/bone contact, particularly the fact that the equatorial surface of the burr carries a larger linear velocity than the polar surface and thus removes more bone per unit of applied force. The precise distribution of drilling power and the dependence of bone removal on axial distance are varied with the type of drill being used, which allows us to model critical differences among burr types. For example, our model captures the fact that cutting burrs typically show more dependence on drilling angle than diamond burrs do.

Voxels are stored in an in-memory hash table. When a voxel's density reaches zero, it is removed and is no longer accessed for haptic or graphic rendering.

A key aspect of the haptic sensation associated with drilling is the vibration of the instrument, which varies with applied force and with burr type. In order to generate realistic drill vibration frequencies, we outfitted a physical drill with an accelerometer and collected vibration data at a variety of applied drilling forces. These data are summarized in FIGURE 1. The key spectral peaks were identified for each burr type and used to synthesize vibrations during the simulation. Since we are driving our haptic feedback device at only 1 kHz, we are able to preserve only the lower-frequency vibrations identified from these experimental recordings. However, we are able to preserve the variations in vibration associated with changes in burr type and/or changes in applied drilling force.

Our goal in drill modeling is not necessarily to simulate every detail of physical reality, but to allow enough system parameterization for an experienced surgeon to tune the experimentally-based parameters. With this in mind, we have parameterized all of

the values discussed in this section – bone removal rate, dependence of bone removal on axial position, and vibration amplitude and fundamental frequency – and made them available via on-screen sliders in a "tuning" version of our system. Initial work with an otologist within our group has allowed us to determine effective values for each of these parameters.
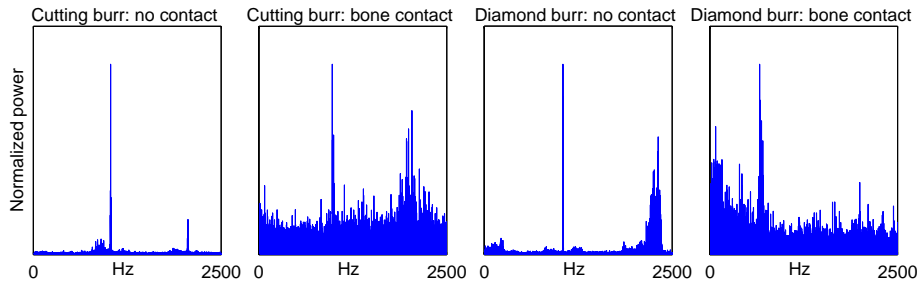

FIGURE 1: A spectral representation of drill vibration, collected from cutting and diamond drilling burrs, when in contact with bone and when powered but held away from the bone surface. Only the low spectral peaks have been plotted; those are the vibrations that can be captured with a typical haptic rendering device. The sharp spectral peaks make this data suitable for real-time vibration synthesis.

3.2 A hybrid data structure for graphic rendering

Volumetric data is well-suited for haptic simulation of bone removal, as it allows rapid collision detection and efficient removal of material units. However, direct rendering of volumetric data – while appropriate for visualization of transparent volumes – limits the degree to which primarily opaque surfaces (such as bone) can be realistically displayed, due to resolution limitations. Furthermore, a purely volumetric system will be unable to leverage the current trend toward increasingly powerful surface-rendering capabilities in graphics hardware.

With that in mind, our system maintains a hybrid data structure in which voxel data is used for haptic rendering and bone manipulation (as is described above), and triangulated surfaces are used for graphic rendering. Voxel locations are used directly as triangle vertices. When a voxel model is loaded into our system, the following algorithm is performed to isolate a polygonal surface that bounds the underlying voxels:

```
for each voxel v1 that is on the bone surface
  for each of v1's neighbors v2 that is on the bone surface
    for each of v2's neighbors v3 that is on the bone surface
      generate a triangle (v1,v2,v3) oriented away from the bone surface
```

Normals and texture coordinates for each voxel are derived from information embedded in the voxel data file, which is generated by flood-filling a temporal bone surface model [7]. Redundant triangles are eliminated by assigning a unique integer ID to all voxels and rejecting all triangles whose vertex ID's do not appear in sorted order. Further culling of subsurface (and thus invisible) triangles is performed according to [5]. To permit backface culling, each triangle is oriented to face outward by comparing its surface normal with the average of the normals at each vertex.

Each time a voxel is removed, all triangles containing that voxel are removed from the rendering list, and the above algorithm is repeated locally, for each neighbor of

the removed voxel that is now part of the bone surface. The local density gradient is also computed at each voxel bordering the removed voxel, to allow normal re-calculation, and texture coordinates are propagated to newly-revealed voxels. In this manner, a consistent bounding mesh is maintained, with limited (constant-time) computation performed each time the surface is modified. FIGURE 2 shows a bone surface that has been modified by a drill, along with the underlying volumetric representation.
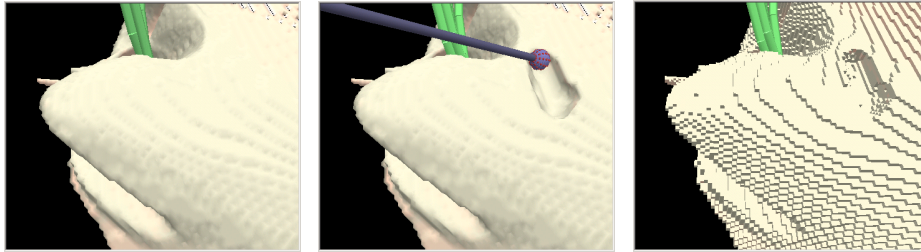

FIGURE 2: Interactive bone-drilling simulation. 2a (left): a smooth surface representation of a temporal bone model. 2b (center): virtual drilling of this surface; the surface is re-meshed as voxels are removed. 2c (right): an opaque rendering of the underlying volumetric representation of the surface, which is used for haptic feedback.
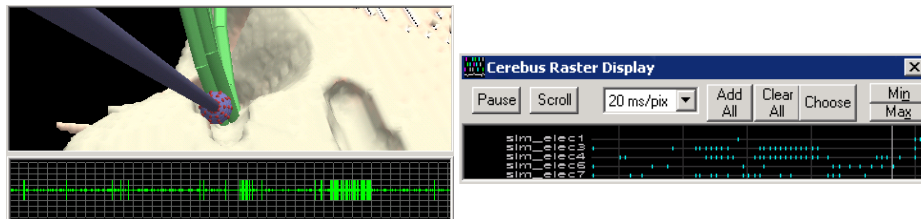

FIGURE 3: (a) Left: a virtual drilling procedure that requires drilling in the vicinity of simulated nerves. The bottom of the display represents a virtual neurophysiology monitor; note the bursts of activity resulting from close contact with the drill. (b) Right: A remote system monitors the activity of multiple simulated neurons.

3.3 Neurophysiology monitoring

Neural electrodes are often used during temporal bone surgery to monitor the facial and/or auditory nerves. The output from these electrodes – typically monitored by a dedicated individual or team – is used to guide navigational decisions. Our system thus incorporates a set of virtual nerves (displayed in FIGURE 3a), whose firing rate increases with proximity of an active drilling tool, but decreases with prolonged drill proximity or direct drill contact. The user can also safely stimulate virtual nerves using an electrical stimulator tool; this is analogous to the direct stimulation performed in the OR in conjunction with electrode monitoring, used to locate critical nerves that are not immediately visible.

An activity monitor can be displayed on the surgeon's display, as is shown in FIGURE 3, and sound can be presented to represent a neural spike train. Since nerve monitoring in certain procedures involves a dedicated neurophysiology team, our system also exports neural activity over a local network for monitoring on another computer. More specifically, our simulator emulates the Cerebus (CyberKinetics, Inc.) neural recording system, a recording amplifier and digitizer used in human and animal neurophysiology research. This allows us to use a variety of visualization

tools to present neural activity to users or instructors, who may be monitoring a trainee's performance. Furthermore, the flexible client architecture will eventually allow us to develop a more precise simulator of the neural monitoring equipment used intraoperatively. FIGURE 3b shows a screenshot from the Cerebus client software, demonstrating the activity of several simulated nerves.

3.4 Sound

Sound is a key source of intraoperative feedback, as it provides information about drill contact and about the nature of the underlying bone. We simulate the sound of the virtual burr as a series of noisy harmonics, whose frequency modulates with applied drilling force. Building upon the harmonic-based synthesis approach presented in [4], we have recorded audio data from cutting and diamond drill burrs under a series of drilling forces in order to determine the appropriate frequencies for synthesized sound, as well as the dependence of this data on drill type and applied drilling force. FIGURE 4 contains example spectral information collected from diamond and cutting burrs.

Sound can also be a key indicator of bone thickness intraoperatively; sound quality and frequency change significantly as the drill contacts a thin layer of bone, providing a warning that the surgeon is approaching sensitive tissue. In our simulator, the pitch of the synthesized sound increases when the drilled area becomes thin. In order to estimate the thickness of bone regions, we used a raytracing algorithm similar to that used for haptic rendering. At each voxel that is determined to be on the surface of the bone, the surface gradient is used to approximate the surface normal, and a ray is cast into the bone along this normal. The ray is traced until it emerges from the bone volume, and the thickness is estimated as the distance from the ray's entry point to its exit point. For sound synthesis, this thickness is averaged over all surface voxels with which the drill is in contact. The computed thickness values are also used to visually shade thin bone, which tends to appear red and translucent as the drill approaches soft tissue.
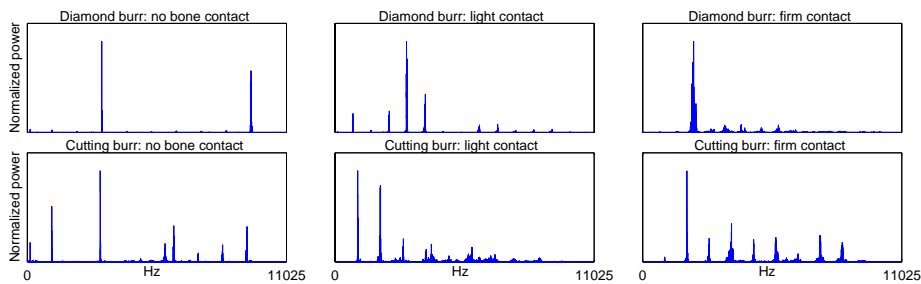

FIGURE 4: A spectral representation of audio data collected from diamond (top) and cutting (bottom) drilling burrs. Columns represent no bone contact, bone contact without significant pressure, and bone contact with a typical drilling pressure (applied by an experienced surgeon). The sharp spectral peaks and distinct variation among drill types and contact forces make this data suitable for real-time synthesis.

3.5 Collaborative networking

A key advantage of simulation-based training is the ability for multiple users to observe and manipulate a common model without interfering with each other's actions.

For example, it is useful for an instructor to be able to (possibly remotely) monitor a trainee's performance, and interactively provide feedback and demonstrations.

With this in mind, our system allows a second user to log in via Ethernet and interact with the bone model. We assume that the model is available at both stations, so only model modifications and drill positions/forces need to be sent over the network. We use a private gigabit Ethernet intranet, which allows us to send a continuous stream of position and force data with minimal latency (approximately 5ms).

Additionally, rather than interacting directly with the local model, a user can choose to drive his haptic device using the forces that are generated by the remote user's drill. This allows a surgeon to demonstrate proper drilling technique while a trainee actually experiences the corresponding contact forces, something that is not possible with traditional training techniques. The ability to send forces at haptic update rates allows even high-frequency vibrations to be rendered remotely.

3.6 Soft Tissue Simulation

A training surgeon using a temporal bone simulator will benefit from a logical representation of initial soft tissue incisions and – where appropriate – lesion resection. Our environment uses a system of masses and springs to simulate layers of skin and muscle that can be cut and displaced interactively (see FIGURE 5).
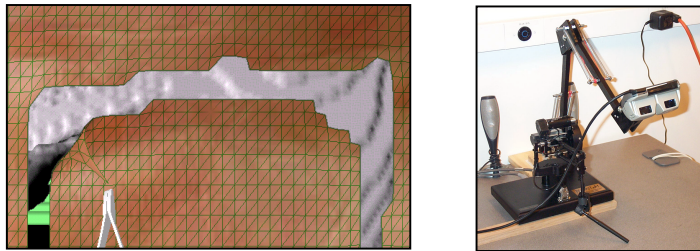

FIGURE 5 (left): Skin, modeled as a 3-d mesh of masses and springs, is interactively cut and retracted to reveal the underlying bone volume.
FIGURE 6 (right): The binocular stereo display and haptic rendering device used for simulation.

3.7 System Architecture

Our simulation environment runs in Windows XP on a 2-CPU 3GHz Xeon with 1GB RAM, using a SensAble Phantom for haptic feedback. OpenGL is used for graphic rendering; each bone surface mesh is rendered as a single vertex array, embedded in a display list. STL hash tables are used to represent voxel information and the mapping from voxels to vertex arrays. A high-priority thread handles all bone interaction, haptic rendering, and networking; reorganizations of the polygon mesh are queued by this thread and performed at graphic update rates by the rendering thread.

The scene is rendered in stereo using either shutter glasses or a binocular display consisting of two small LCD's mounted on an adjustable arm, depicted in FIGURE 6. The latter is analogous to the binocular microscope used intraoperatively.

The virtual drill is enabled and disabled using a floor-mounted switch, which simulates the foot-controlled pneumatic valve that is typically used for regulating drill power. The switch is interfaced via the workstation's parallel port.

4 Conclusion and Future Work

We have collaborated with a Stanford otologist to iteratively refine our simulation. Force and sound data are based on experimental results, but wherever possible, we allow physical parameters to be tuned online. The result is an effective simulation of the haptic and auditory components of drilling.

This paper focuses on novel sources of feedback that are not available in other temporal bone simulation environments; in the next iteration of our simulator, we plan to include additional features that have been previously described in [2], including a particle-based simulation of blood, dust, irrigation, and suction.

Our current source of data is a series of surface models of temporal bone structures, manually created by an otologist in our group. The use of surface models allows us high rendering resolution, provides texture-mapping and normal information for each surface voxel, and will allow us to leverage existing modeling packages as we move toward "scenario training", in which an instructor will be able to modify patient anatomy directly to prepare specific challenges and complications. However, we also plan to incorporate direct use of patient-specific CT data in the next revision of our simulator, a natural extension of our underlying volumetric representation. Also, we are currently applying this work to related procedures and data sets in dental and craniofacial surgery.

References

1. Agus, M,. Giachetti, A., Gobbetti, E., Zanetti, G., Zorcolo, A.: A multiprocessor decoupled system for the simulation of temporal bone surgery. Comp Vis in Science, 5(1): 35-43, 2002
2. Agus, M,. Giachetti, A., Gobbetti, E., Zanetti, G., John, N.W., Stone, R.J.: Mastoidectomy simulation with combined visual and haptic feedback. Proc of the Medicine Meets Virtual Reality 2002 Conference, Newport Beach, CA, Jan. 23-26, 2002
3. Blevins, N.H., Jackler, R.K., Gralapp, C: Temporal Bone Dissector. Mosby, January 1998.
4. Bryan, J., Stredney, D., Wiet, G., Sessanna, D.: Virtual Temporal Bone Dissection: A Case Study. Proc. of IEEE Visualization 2001, Ertl et. Al., (Eds): 497-500, October 2001.
5. Bouvier, D.J.: Double-Time Cubes: A Fast 3D Surface Construction Algorithm for Volume Visualization. Int'l Conf on Imaging Science, Systems, and Technology, June 1997.
6. Hohne, K.H., Bomans, M., Riemer, M., Schubert, R., Tiede, U., Lierse, W.: A 3D anatomical atlas based on a volume model. IEEE Visualization 1992, 12 (1992) 72-78
7. Kaufman, A., Shimony, E.: 3D Scan-Conversion Algorithms for Voxel-Based Graphics. Proc ACM Workshop on Interactive 3D Graphics. Chapel Hill, NC, Oct 1986, 45-76.
8. Massie, T.H., Salisbury, J.K.: The PHANTOM Haptic Interface: A Device for Probing Virtual Objects. Symp. on Haptic Interfaces for Virtual Environments. Chicago, IL, Nov. 1994.
9. Petersik, A., Pflesser, B., Tiede, U., Hohne K.H., Leuwer, R.: Haptic Volume Interaction with Anatomic Models at Sub-Voxel Resolution. Proc IEEE VR, Orlando, FL, Mar 2002.
10. Renz, M,. Preusche, C., Potke, M., Kriegel, H.P., Hirzinger, G.: Stable haptic interaction with virtual environments using an adapted voxmap-pointshell algorithm. Proc Eurohaptics, p149-154, 2001.