Quantifying Risky Behavior in Surgical Simulation

Christopher Sewell[1] Dan Morris[1] Nikolas Blevins[2] Federico Barbagli[1] Kenneth Salisbury[1]
Departments of [1]Computer Science and [2]Otolaryngology, Stanford University

Abstract. Evaluating a trainee's performance on a simulated procedure involves determining whether a specified objective was met while avoiding certain "injurous" actions that damage vulnerable structures. However, it is also important to teach the stylistic behaviors that minimize overall risk to the patient, even though these criteria may be more difficult to explicitly specify and detect. In this paper, we address the development of metrics that evaluate the risk in a trainee's behavior while performing a simulated mastoidectomy. Specifically, we measure the trainee's ability to maintain an appropriate field of view so as to avoid drilling bone that is hidden from view, as well as to consistently apply appropriate forces and velocities. Models of the maximum safe force and velocity magnitudes as functions of distances from key vulnerable structures are learned from model procedures performed by an expert surgeon on the simulator. In addition to quantitatively scoring the trainee's performance, these metrics allow for interactive 3D visualization of the performance by distinctive coloring of regions in which excessive forces or velocities were applied or insufficient visibility was maintained, enabling the trainee to pinpoint his/her mistakes and how to correct them. Although these risky behaviors relate to a mastoidectomy simulator, the objectives of maintaining visibility and applying safe forces and velocities are common in surgery, so it may be possible to extend much of this methodology to other procedures.

1. Introduction

The education of a surgeon-in-training involves the acquisition of the sensorimotor skills necessary for performing surgical tasks as well as the refinement of the cognitive processes involved in performing a full procedure. While a number of existing surgical simulators have been developed to train specific skills, there is also substantial benefit to providing trainees with increased experience through simulation in dealing with the wide range of potential scenarios that can arise in the course of performing a full procedure. Ideally, such a simulator should allow the trainee to interact with the virtual environment in a free-form manner, while evaluating his/her performance according to criteria devised and tuned by the instructing surgeon. It should also provide the user with feedback, both in the form of quantitative metrics and constructive criticism, detailing the trainee's weaknesses and how they can be improved.

At a basic level, the trainee's performance can be critiqued according to whether he/she achieved an objective (such as exposing a lesion; see Figure 1A) while avoiding "injurous" actions (such as cutting a nerve; see Figure 1B). We have previously proposed an event based framework that allows for the development of such simulations, and have illustrated the feasibility of the methodology in a simulation of a mastoidectomy procedure [1]. However, a more thorough simulator should also be able to assess the trainee's adherence to stylistic guidelines specified by the instructing surgeon. While there may be multiple techniques that a trainee, with a little luck, may be able to use to perform a procedure while avoiding injurous actions, it would be far better for him/her to learn the

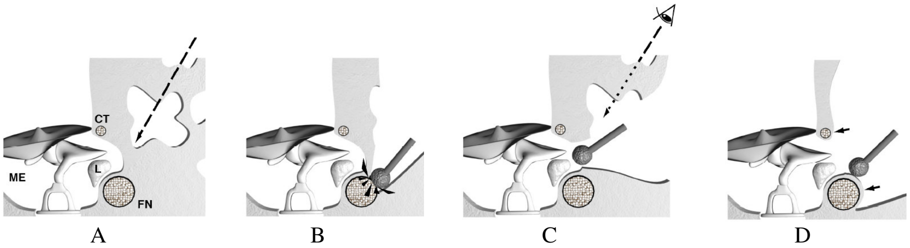


Figure 1. A) Schematic cross-sectional view of the temporal bone illustrating a lesion (L) within the middle ear (ME) cavity. The goal of this simulation is to access this lesion, in the direction shown by the arrow. The chorda tympani (CT) and the facial nerve (FN) are shown. B) An "injurious" action has occurred, with the burr contacting the facial nerve. C) A "dangerous" action has occurred, when the surgeon has drilled away bone without first establishing clear exposure of the region. D) Correct bone removal has occurred. The nerves have been avoided, and adequate exposure has been established. When the drill is in close proximity to a vulnerable structure, as shown here, applied force and velocity magnitudes should be sufficiently small.

specific technique developed over many years by expert surgeons that minimizes the overall risk to the patient. Nevertheless, such criteria are significantly more difficult to specify and to quantify.

2. Types of Risky Surgical Behaviors in a Mastoidectomy

In this paper, we consider three types of risky behavior in a mastoidectomy procedure: removing bone outside the current field of view (see Figure 1C); moving the drill too quickly, especially when in close proximity to vulnerable structures (such as the facial nerve or the sigmoid sinus); and applying excessive force, again especially when operating near vulnerable structures (see Figure 1D).

Recognizing the need for customizable surgical simulators, we have previously developed a graphical scripting environment in which the instructing surgeon can design specific training scenarios using finite state machines. However, it would be much more difficult to attempt to encode stylistic guidelines in this way. Also, explicitly segmenting the bone and assigning the appropriate velocities and forces in each region would be very tedious, error-prone, subjective, and applicable only to one specific model. Substantial research has been conducted both in the fields of non-explicit encoding of procedures [2] and of applying probabilistic and machine learning techniques to the evaluation of surgical skill [3, 4]. The paradigm of "programming by demonstration" can enable the simulator to develop an internal model of "good style" for a procedure by learning from exemplary runs of the simulation by expert surgeons.

2.1 Drill Force

One of the primary components of good technique in a mastoidectomy is applying appropriate forces with the drill when removing bone. The maximum "safe" force is some function primarily dependent on the distance of the current drilling location from key vulnerable anatomic structures. In our simulator, the bone is represented haptically as a collection of voxels (although a hybrid data structure is used to graphically render a triangulated mesh [5]), and the drill is modeled haptically as a cloud of points, as in the method described by Petersik et. al. [6]. Forces are generated by testing the drill points for intersection with bone voxels (via direct indexing), and summing the vectors from each intersected voxel to the drill center (see Figure 2A).

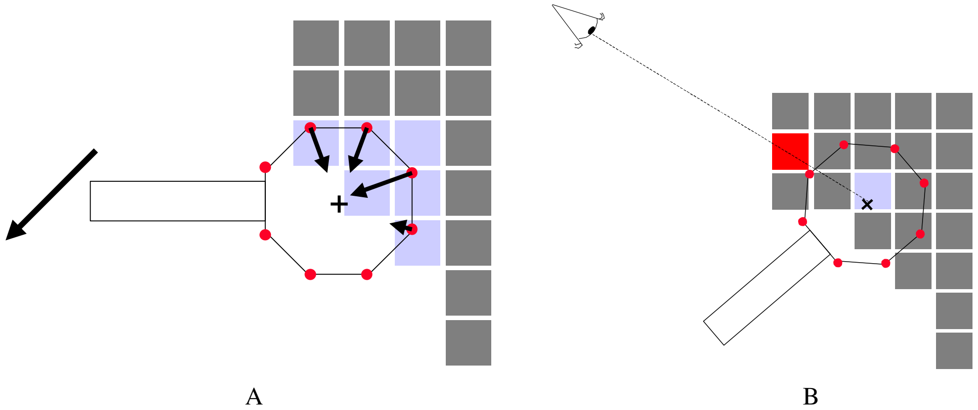


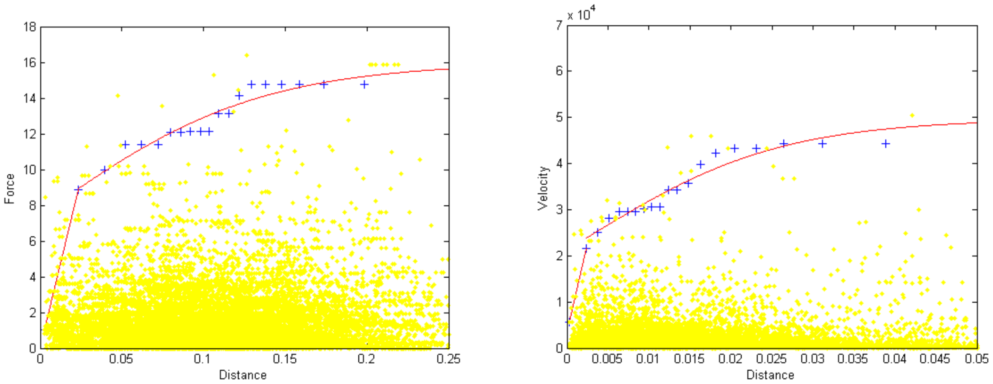
Figure 2. A) Algorithm for generating forces when the drill contacts bone. The bone is represented as voxels (shown here in 2D as squares), and the drill as a cloud of points (the dots along the perimeter of the circular drilling surface). Each drill point is tested for intersection with the voxel mesh (using direct indexing based on location). The generated force vector is the sum of the vectors from each intersected voxel to the drill center. The density of each voxel along these vectors (here shown with light-colored squares) is decreased, eventually causing them to be removed and turn into bone dust (rendered graphically with a simple particle simulator). B) Algorithm for visibility testing. In order to determine whether bone being removed is within the current field of view, a line is simply traced from the voxel (here shown as the light-colored square) to the view point (the OpenGL camera position). Points at discrete intervals along this line are tested for intersection with the voxel mesh. If any voxels (besides those covered by the drill) are intersected (such as the one on the middle far left highlighed here), they obstruct the view and the removed voxel is determined to not be visible. Otherwise, the voxel is visible.

We want to develop a function that, given a trainee's drill location, returns the maximum force magnitude that can safely be applied there. In order to learn the appropriate force profile, we record the location and force magnitude for each voxel as it is drilled by the training surgeon (or another expert surgeon). Instead of defining locations with respect to a fixed coordinate system, location is specified in terms of distance from key vulnerable structures, allowing use of the function in any bone model (including patient-specific models) in which the locations of these structures are known.

To determine the appropriate type of function to use, we performed a trial run in which an expert otolaryngologist performed a mastoidectomy on our simulator. Distances from each voxel to six key structures (the sigmoid sinus, facial nerve, dura, incus, chorda tympani nerve, and inner ear) were pre-computed (using a brute-force method, finding, for each voxel and each key structure mesh, the minimum distance between the voxel center and a vertex of the mesh), and then recorded along with applied force magnitude when drilled by the expert.

Distance versus force magnitude plots were constructed for each key structure. An example, for the chorda tympani nerve, is shown in Figure 3A. Examining the upper hull of the plotted raw data for each key structure, there appeared to be a steep, roughly linear increase in maximum forces for small distances from the structure, and then a gradual ascent to a plateau as distance from the structure increased further. For many of the plots, maximum forces then began to decrease again at even further distances, presumably as the drill began to approach other key structures.

Rather than attempt to develop a single, complex, global function for maximum safe force magnitudes, separate functions were constructed for each key structure. Also, since we were primarily concerned with the upper hull of maximum safe forces, sample points were created by calculating, at regular distance intervals, the averages of the five maximum forces at any equal or smaller distance. These sample points were then fitted using linear regression for the initial, steeply sloping region, and logistic regression for the region ascending to a plateau.



A B

Figure 3. A) A plot of applied force magnitudes for voxels as a function of the distance from the voxel to a key structure, in this case the chorda tympani nerve. The raw data points for the 12,573 voxels removed during a mastoidectomy performed on the simulator by an expert surgeon are shown as light dots. Averaged upper bound sample points, shown as crosses, are calculated at regular distance intervals as the average of the five largest force magnitudes at any equal or smaller distance. The initial, steeply sloping region was fitted using linear regression and the region that gradually ascends to a plateau using logistic regression, both shown as solid lines. At somewhat larger distances (off the scale as shown here), the upper hull of the force magnitudes of the raw data starts to decrease again, probably a result of the drill approaching other key structures. B) A plot of instantaneous drill velocity magnitudes for voxels as a function of the distance from the voxel to a key structure, in this case the sigmoid sinus. The raw data points for the 12,573 voxels removed during a mastoidectomy performed on the simulator by an expert surgeon are shown as light dots. Averaged upper bound sample points, shown as crosses, are calculated at regular distance intervals as the average of the five largest velocity magnitudes at any equal or smaller distance. The initial, steeply sloping region was fitted using linear regression and the region that gradually ascends to a plateau using logistic regression, both shown as solid lines. At larger distances, the upper hull of the velocity magnitudes of the raw data starts to decrease again, probably a result of the drill approaching other key structures.

When the trainee then performs a mastoidectomy on the simulator, the force magnitude applied at each drilled voxel can be compared to the safe force function for each of the key structures, evaluated at the voxel's distance from the structure. The trainee's performance can be evaluated quantitatively by summing, for all voxels for which the applied force magnitude exceeds the maximum of the safe force functions at that location, the amount by which the maximum safe force is exceeded. Perhaps more importantly, upon request, the removed voxels can be re-rendered green if removed with a safe force and red if removed at an unsafe force (with darkness shading according to the amount by which the safe force was exceeded), providing the trainee with constructive graphical criticism, clearly showing where he/she should be observing greater caution.

2.2 Drill Velocity

In addition to refraining from applying excessive forces, caution should also be exercised when in proximity to vulnerable structures by maintaining drill velocities at sufficiently small magnitudes. As for the drill forces, we want to develop a function that, given a trainee's drill location (in terms of its distance from key structures), returns the maximum velocity magnitude that can safely be applied there, and to learn this function by recording the location and velocity magnitude for each voxel as it is drilled by an expert surgeon. Velocities are obtained directly from the haptic device's API.

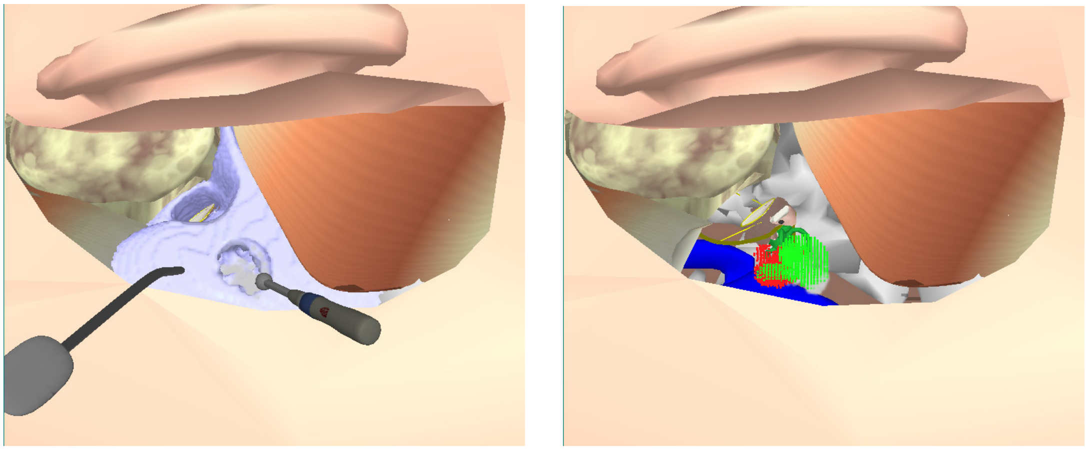
During the trial run described in the previous subsection in which an expert otolaryngologist performed a mastoidectomy on our simulator, we also recorded velocities for each removed voxel. Distance versus velocity magnitude plots were constructed for each key structure. An example, for the sigmoid sinus, is shown in Figure 3B. Examining the upper hull of the plotted raw data for each key structure, the behavior appeared similar to that for force magnitude data, exhibiting a steep, roughly linear increase for small distances, tapering off into a plateau at further distances, and often decreasing again at even further distances. Thus, as before, these regions were fitted with linear and logistic regression, respectively. Quantitative and graphical feedback can then be provided to the trainee based on drill velocities as well as applied forces, showing regions in which he/she should drill more slowly and carefully.

2.3 Visibility

Another type of dangerous behavior in performing a mastoidectomy is drilling bone that is hidden from view, usually by other bone. An otolaryngologist is trained to recognize a variety of subtle but important visual cues that indicate the proximity of vulnerable structures just below the surface of the bone, but these cues cannot be heeded if they cannot be seen.

When the trainee removes a voxel, its visibility is tested using a simple algorithm that traces a line from the voxel to the view point (i.e., the OpenGL camera position). Points along this line are tested for intersection with the bone voxel mesh; if any voxel between the drill and the viewpoint is intersected, the removed voxel must have been hidden. See Figure 2B. As for the other risky behaviors, the trainee's performance can be both scored, according to the percentage of removed voxels that were hidden, and visualized, showing voxels that were visible when removed in green and those that were not in red, as shown in Figure 4. Optionally, selected structures can be prevented from being rendered, and the scene can be rotated, allowing unobstructed visualization of the locations of the removed voxels and their proximities to other structures.

Although ideally no invisible voxels should be removed, the severity of the consequences of removing invisible voxels can vary depending on location. Thus, rather than a binary visible/invisible rating for each voxel, scoring and shading could be graded depending upon distances to key structures, although we have not implemented this.



A B

Figure 4. A) A trainee drills the temporal bone while performing a mastoidectomy on the simulator. In this example, the trainee "saucerized" on the right side, removing only visible bone, while he "undercut" on the left side, removing bone that was hidden by other bone. B) On the final report, it was noted that 33% of voxels were hidden when removed. Here the trainee can visualize where he performed well (visible voxels in green) and where he did not (invisible voxels in red), and realize that undercutting in close proximity to the sigmoid sinus (in blue) was dangerous as he could not see the visual cues indicating the vein's location below the bone surface.

3. Future Work and Conclusions

In this paper we have introduced a methodology for training stylistic surgical techniques that avoid risky behavior. In the context of a mastoidectomy procedure, we evaluated a trainee's ability to maintain an appropriate field of view, as well as to apply forces and velocities consistent with those implicitly defined by an expert surgeon by performing model procedures on the simulator. Although these risky behaviors were identified by an otolaryngologist as among the most important for training specifically in a mastoidectomy simulator, the goals of maintaining visibility in the operating region and of recognizing crucial structures and exercising appropriate caution around them are common in surgery, so many of these ideas may be generalized and applied to other procedures.

The performance of anyone using the simulator is highly dependent upon the quality of the simulator. Thus, as we continue to improve the visual and haptic realism of the simulator, new expert force and velocity profiles will need to be recorded, and will more closely approximate the forces and velocities actually used in the operating room. Other additions to the simulator, such as more realistic stereo rendering and a viewpoint positioning device that mimics the microscope used in the operating room, will also affect the ways in which users can maintain visibility and avoid removing hidden bone. We also hope to employ more sophisticated learning algorithms to better model expert force and velocity profiles, as well as to explore other risky surgical behaviors, perhaps even in the context of other procedures.

4. References

[1] Christopher Sewell, Dan Morris, Nikolas Blevins, Federico Barbagli, and Kenneth Salisbury: An Event-Driven Framework for the Simulation of Complex Surgical Procedures. Proceedings of the International Conference on Medical Imaging Computing and Computer Assisted Intervention (MICCAI), St. Malo, France, September 2004, Springer Lecture Notes in Comptuer Science, Vol. II, pp. 346-354.

[2] David Canfield Smith: Kidsim: Programming Agents Without a Programming Language. Communications of the ACM, July 1994, Vol. 37, No. 7.

[3] Jacob Rosen, Massimiliano Solazzo, Blake Hannaford, and Mika Sinanan: Objective Laparoscopic Skills Assessments of Surgical Residents Using Hidden Markov Models Based on Haptic Information and Tool/Tissue Interactions. Medicine Meets Virtual Reality, Newport Beach, CA, January 2001.

[4] Louai Adhami. An Architecture for Computer Integrated Mini-Invasive Robotic Surgery: Focus on Optimal Planning. PhD Thesis, Ecole des Mines de Paris, 2002.

[5] Dan Morris, Christopher Sewell, Nikolas Blevins, Federico Barbagli, and Kenneth Salisbury: A Collaborative Virtual Environment for the Simulation of Temporal Bone Surgery. Proceedings of the International Conference on Medical Imaging Computing and Computer Assisted Intervention (MICCAI), St. Malo, France, September 2004, Springer Lecture Notes in Comptuer Science, Vol. II, pp. 319-327.

[6] A. Petersik, B. Pflesser, U. Tiede, K.H. Hohne, and R. Leuwer: Haptic Volume Interaction with Anatomic Models at Sub-Voxel Resolution. Proceedings of the IEEE Virtual Reality, Orlando, FL, March 2002, pp. 66-72.