Quantifying Risky Behavior in Surgical Simulation

Christopher Sewell\(^1\)  Dan Morris\(^1\)  Nikolas Blevins\(^2\)  Federico Barbagli\(^3\)  Kenneth Salisbury\(^1\)
Departments of \(^1\)Computer Science and \(^2\)Otolaryngology, Stanford University
\{ceswell, dmmorris, nblevins, barbagli, jks\}@stanford.edu

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 freeform 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; Figure 1A) while avoiding “injurious” actions (such as cutting a nerve; Figure 1B). We have previously proposed an event based framework, including a finite state machine based scripting environment, that allows for the development of such simulations, and have illustrated the feasibility of the methodology in a simulation of mastoidectomy [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 injurious actions, it would be far better for him/her to learn the specific technique developed over many years by expert surgeons that minimizes the probability of injury to the patient. Nevertheless, such criteria are significantly more difficult to specify and to quantify.

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]. 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.

One of the primary components of good technique in a mastoidectomy is exercising appropriate caution when drilling near vulnerable structures, such as the facial nerve or the sigmoid sinus. Caution is exercised both by operating more slowly, and by removing only bone in the field of view (Figure 1C).

In order to assess the risk of a trainee’s performance, a function is needed that associates each location in which bone is removed with a value indicative of the degree of danger involved in drilling there. However, explicitly segmenting the bone and assigning danger values would be very tedious, error-prone, subjective, and applicable only to one specific model.

Instead, we are developing a system that learns this function from simulated procedures performed by expert surgeons. The bone model is discretized into voxels, and each voxel removed is a training example. Its location is specified in terms of its distance from key vulnerable structures, allowing use of the function in any model in which these structures are identified. Its danger value is inferred from the instantaneous rate of bone removal and visibility of the voxel when removed. Areas in which the expert is comfortable operating quickly and removing some bone before exposing it are assumed to be less dangerous than areas in which the expert operates more slowly and refrains from removing any bone prior to exposure.

The risk factor associated with each voxel removed by a trainee is then obtained as the product of the instantaneous rate of bone removal and the danger function evaluated at the voxel’s location, multiplied by a constant $c>1$ if the removed voxel is not visible. In addition to yielding an overall risk score, the simulator can provide the user with a visualization of all removed voxels, colored according to risk value so that he/she can see the regions in which additional caution should be taken.

---