Geometric modelling of the temporal bone for cochlea implant simulation

Catherine Todd  
*University of Wollongong, cath@uow.edu.au*

Fazel Naghdy  
*University of Wollongong, fazel@uow.edu.au*

S. J. O'Leary  
*Royal Victorian Eye and Ear Hospital*


**Publication Details**

Geometric Modelling of the Temporal Bone for Cochlea Implant Simulation

C.A. Todd¹, F. Naghdy¹, S.J. O’Leary²
¹School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, Northfields Ave., NSW 2520, Australia
²Royal Victorian Eye and Ear Hospital, 32 Gisborne St., East Melbourne, VIC 3002, Australia

ABSTRACT

The first stage in the development of a clinically valid surgical simulator for training otologic surgeons in performing cochlea implantation is presented. For this purpose, a geometric model of the temporal bone has been derived from a cadaver specimen using the biomedical image processing software package Analyze (AnalyzeDirect, Inc) and its three-dimensional reconstruction is examined. Simulator construction begins with registration and processing of a Computer Tomography (CT) medical image sequence. Important anatomical structures of the middle and inner ear are identified and segmented from each scan in a semi-automated threshold-based approach. Linear interpolation between image slices produces a three-dimensional volume dataset: the geometrical model. Artefacts are effectively eliminated using a semi-automatic seeded region-growing algorithm and unnecessary bony structures are removed. Once validated by an Ear, Nose and Throat (ENT) specialist, the model may be imported into the Reachin Application Programming Interface (API) (Reachin Technologies AB) for visual and haptic rendering associated with a virtual mastoidectomy. Interaction with the model is realized with haptics interfacing, providing the user with accurate torque and force feedback. Electrode array insertion into the cochlea will be introduced in the final stage of design.

Keywords: haptic, temporal bone, cochlea implant, geometrical model, Analyze, region-growing, otologic.

1. INTRODUCTION

1.1. Insight into Cochlea Implantation

Cochlea implant surgery was initiated in 1978 to aid in the correction of hearing for the profoundly deaf. Since then, over 36,000 adults and children have been fitted with the auditory prosthetic. Recipients range in age, from as young as twelve months to ninety-four years, and degree of success is variable. The cochlea implantation takes place in three stages: access to the middle ear via temporal bone drilling (mastoidectomy), access to the cochlea duct (cochleostomy) and insertion of the multi-electrode array.

For the otologic surgeon specializing in cochlea implantation, methods of training are limited. Medical instruction has traditionally relied on a mentor/trainee approach, yet this has proven to be both time-consuming and costly. Temporal bone drilling laboratories have been established for the training surgeon to practice bone drilling techniques using cadaver specimens. Material of this type is not readily available and is becoming increasingly hard to acquire. In effect, different bone conditions, sizes, shapes, races and ages cannot be represented by the small sample in the laboratory. Facilities of this type are therefore costly and do not represent all possible scenarios that could be encountered in the operating suite. Further, intra-cochlea tissue characteristics change significantly after death, yet post-mortem analysis of the middle and inner ear remains the only practical alternative to live dissection.

There is significant risk associated with an inexperienced medical student performing cochlea implant surgery. In the middle and inner ear regions, the human physiology is quite complex. Damage to muscles, blood vessels or nerves within this area could cause permanent trauma and/or loss of motor sensory functionality: for example, harm to the facial nerve can cause facial palsy. Observation of an experienced ENT specialist does not, however, provide adequate training for the novice physician.
Surgical simulators may provide a valuable supplement to current methods of cochlea implant training. Difficult or unfamiliar procedures may be practiced without endangering the patient. Abnormalities or complications may be replicated in a risk-free environment. Although initial costs associated with establishing the simulator may be incurred, on-going costs of training may be minimized as opposed to those related to experienced instruction and purchase of materials (such as drill bits and cadaver specimens). Objective evaluation of surgeon technique is another potential benefit, offering a new level of ability assessment. Cochlea implant simulation would provide a safe, cost-effective, yet realistic substitute to current medical intervention practices. It also has the potential to provide the manufacturer with a model that can predict, display and monitor over time the behaviour of a prosthesis relative to operational influences such as electrode array location.

Technological advancement in computer hardware and software has led to increased processing speeds, realistic graphics and more recently, development of force-feedback (haptic) devices that compute real-time forces and torques associated with model manipulation. Consequently, virtual representations of human anatomy are becoming life-like, both in look and feel. Real-time interaction and manipulation of a realistic virtual model of the ear has great potential for medical instruction of cochlea implant surgery.

1.2. Haptics

Recent medical simulations have integrated haptic feedback with visual models to provide the user with a more realistic and immersive environment. Haptic rendering forms a closed loop system. The user initiates movement through the haptic device. Upon tool/object interaction, the model deforms accordingly and associated force and torque components are delivered back to the user through the haptic interface.

Sense of touch is often used by the surgeon to infer important information during the operation. Tool/object collision may relay information concerning localized tissue or bone recognition, defects in burr condition, drill responsiveness or off-centred drill bit mounting. Haptic rendering should ideally model these forces and torques as they would occur in real-life surgical interventions, including any imperfections.

Although the benefits of force-feedback are numerous, there are two significant drawbacks. One is cost: initial price of simulator establishment is quite high, especially for systems incorporating a six-degree-of-freedom (6DOF) haptic device. The second is that there is often a trade-off between graphic display quality and haptic algorithm complexity. This is due to real-time constraints imposed on the haptic update loop. Visual update rates can be as low as 20 - 30Hz, whilst acceptable haptic update rates must be at least 500Hz. Complex systems (for example those including force calculation using high order differential equations) are often simplified and reduced to rudimentary approximations to reduce computational overhead.

In this work, a haptic rendered computer simulation for mastoidectomy and cochlea implantation processes will be developed. The first stage of this research will be presented in this paper, following a brief analysis of existing work in this field.

2. EXISTING WORK

2.1. Visual Representation

Three-dimensional modelling of the human ear in a virtual reality environment has attracted global interest. Computer-aided reconstruction of the temporal bone has been implemented using histological sections, computer tomograms (CT) and magnetic resonance images (MRI). Multi-modal datasets have also been used, where high resolution MRI data is combined with CT sequences to provide enhanced detail (such as soft tissue structures).

Image processing and volume generation may be implemented using a manual, semi-automatic or fully automated approach. Anatomical structures in the temporal bone are separated from surrounding parenchyma in a process called segmentation. A fully automated approach resulting in highly accurate feature delineation is difficult and has not yet been achieved. Often, complex structures can only be precisely defined using manual segmentation; however manual techniques are time-consuming. Manual and semi-automatic image processing methods are the most common for this
type of application. Software packages facilitating temporal bone model generation from CT or MRI include SolidWorks\textsuperscript{7-8} (SolidWorks, Inc.), OpenGL\textsuperscript{6} (Silicon Graphics, Inc.) and Analyze\textsuperscript{17} (AnalyzeDirect, Inc.).

2.2. Temporal Bone Dissection

Previous work has focused solely on a mastoidectomy. Three such groups have produced real-time models: the University of Hamburg, Germany; the Ohio State University, Columbus, United States of America and the Center for Advanced Studies, Research and Development, Sardinia, Italy. Tiede et al\textsuperscript{12-14,18-19} achieve superior quality graphics and detailed haptic resolution, with haptic update rates of 6000Hz on a three-degree-of-freedom (3DOF) haptic device. A life-like scenario of the bone drilling operation is provided\textsuperscript{6,20}, where visual feedback is combined with aural and haptic sensations, although haptic algorithms are simplified in a spring-based approach. Agus et al\textsuperscript{21} model burr-bone interactions and effectively simulate fluid movement caused by irrigation. Insight into the forces exerted during bone erosion is provided, but the authors acknowledge further verification is required.

3. CURRENT RESEARCH

Despite emerging interest in the area of surgical simulation, an accurate force-feedback system for training in cochlea implantation has not yet been developed. This is primarily due to real-time constraints and limitations in computer hardware and software capabilities. Development in the area of haptics is only quite recent and will invariably take some time to progress.

In this work, accurate force-feedback associated with a mastoidectomy and cochlea implant insertion will be modelled, using a 6DOF PHANTOM device (SensAble Technologies). Visual and haptic rendering of this process will be implemented using the Reachin API platform.

The focus of this paper is on the three-dimensional reconstruction of the temporal bone, using the Analyze software package. This work forms the basis for simulator construction. Geometric modelling of the temporal bone will be discussed in the following section 4. Geometric Modelling. Current development in Reachin API will also be presented.

4. GEOMETRIC MODELLING

A Siemens Somatom Plus 4 scanner was used to take spiral CT scans of a right temporal bone cadaver, at the Royal Victorian Eye and Ear Hospital (RVEEH) in Melbourne. The cadaveric specimen was acquired from the institute’s temporal bone drilling laboratory and selected for scanning by an experienced ENT surgeon, as it appeared to be in good condition with no structures missing. Upon inspection, skin, muscle, nerves, an artery and vertebrae could be distinguished without visual aid.

An imaging technician positioned the cadaver and scanned it from superior to inferior, anterior to posterior. Cross-sectional resolution was 512 (pixels/centimetre) by 512. Initially, spacing between image planes was 1mm and slices were later reconstructed to provide 0.5mm spacing between each scan. The image sequence was transferred from raw to DICOM data format in a lossless transfer.

4.1. Image Sequence Registration in Analyze

Analyze is a biomedical image processing and visualization application that supports a variety of image modalities (including CT). Image sequences may be viewed, manipulated and measured interactively by the user. A range of functions are available for feature enhancement and segmentation that use either manual or semi-automatic processes. Various interpolation and shading methods may be selected for volume rendering techniques. In this work, modules that were used for image processing include: Import/Export, Image Edit, Morphology, Image Calculator and Volume Renderer. These are all semi-automatic processes that require some degree of user input.

The image sequence generated from scanning was imported into Analyze for processing. Scan files were sorted automatically prior to import, to ensure that the sequence was in chronological order. This is essential. Scan sequences may be stored or transferred out of sequence and will invariably create erroneous data upon volume generation. Once
loaded as a single volume, anisotropic volume elements (voxels) comprising the image sequence were forced to cubic geometry using linear interpolation. Following import into Analyze, individual scans within the volume were viewed using the Image Edit module. With this functionality, anatomical landmarks such as the cochlea first and second turns, were easily identified by visual comparison with published works\textsuperscript{22}.

4.2. Image Processing

Mastoidectomy simulation in this work requires that only bony structures be segmented from surrounding parenchyma. To successfully segment the temporal bone, simple thresholding was applied. This method was chosen due to the high density of bone which often has specific grey levels and exhibits good contrast with softer tissues captured by CT scans\textsuperscript{19}. A semi-automatic threshold-based method was effectively used by Pfleesser\textsuperscript{18} et al to segment thirty objects within CT scans. It should be noted that certain structures still required manual segmentation, including the facial nerve and auditory ossicles.

Simple thresholding was applied in a semi-automatic approach, within the Morphology module, to segment the bony structures comprising the temporal bone. Threshold limits were interactively changed using an interface that facilitated effective and fast user interaction. As these boundary values were varied, a binary slide sequence was updated on the display, showing voxels (in white) whose intensities were within this range.

The temporal bone is comprised of four main constituents: the petrous, mastoid, squamous and tympanic bone sections. These four bony structures were preserved whilst non-bony elements (such as skin and muscle) were eliminated using the threshold approach. An intensity range of 300 to 5904 was selected to satisfy this requirement. Again, published works\textsuperscript{22} were used to cross-reference anatomical landmarks identified in the images, to ensure that no bony information was missing in the processed image. Verification was also carried out using a sub-function of the Morphology module, the Step Editor, as well as inspection by an experienced ENT surgeon. As an output from this stage, a binary volume, formed from processed image scans of the temporal bone, is produced. Pictured below is an original image imported into Analyze (Figure A) and a processed image of the temporal bone (Figure B). Note that the latter is not a binary image, but is formed following application of methods to be discussed below.

![Figure A: Original Image of temporal bone scan slice.](image1.png)  ![Figure B: Processed image of temporal bone scan slice includes vertebrae and cheekbone components.](image2.png)

A binary dataset is produced from the method applied previously. To derive a subset of the original temporal bone volume that consists of voxels belonging only to bony structures within this region, a matrix multiplication is carried out. Image Calculator module provided a quick, easy multiplication of the two arrays: the original grey-scale volume of the
temporal bone and the binary volume. The resultant volume was displayed using Volume Renderer, which provided a three-dimensional perspective with Gradient Shading as the render type.

Volume Renderer functionality allows for interactive display and rotation of the derived volume. Upon rotating the volume about each of its three axes, artefacts or high intensity noise surrounding the temporal bone segments were observable. Initially, artefacts were removed manually using Image Edit. Unwanted parts in each scan were identified, outlined and deleted manually. This proved to be a laborious task that could not easily nor quickly be replicated. A second, semi-automatic method was then chosen for this purpose, using Slice Edit in Volume Renderer. This also proved arduous and a final method for fast elimination of artefacts was instead selected. The latter makes use of the Connect Tool in Volume Renderer. Again this is a semi-automatic process that applies a seeded region grow, with 26-voxel connectivity. The user selects a single seed pixel which is positioned on the mastoid portion of the temporal bone. Once the connection process is initiated, voxels connected to the seed pixel are kept as part of the Region of Interest (ROI). Successful elimination of surrounding artefacts was achieved using this approach (Figure B).

4.3. Final Volume

Interactive rotation and display functionality in Volume Renderer was again used to view the volume produced after elimination of noise surrounding the temporal bone region. An ENT specialist examined the virtual volume and identified inclusion of unnecessary structures: the vertebrae and cheek bone components (Figure B). Such information is not required for mastoidectomy simulation. Another sub-function within Volume Renderer, Object Separator, was effectively used to remove this unwanted data.

To initiate the semi-automatic process, a seed pixel was selected on both sides of the object boundary to identify the regions to be separated. This was done for both the cheek bone and vertebrae structures. Clearly defined, natural edges existed around these components, which made the segmentation process uncomplicated. An ENT surgeon again examined the virtual specimen, using both interactive rotation capabilities offered by Volume Renderer to view a three-dimensional perspective and analysing individual, two-dimensional image slices with Edit Review in the Image Edit module. Extraneous data was effectively eliminated using object separation techniques available in Analyze. The final result is shown below (Figure C, Figure D).

Figure C: Final temporal bone volume produced with Analyze AVW format.

Figure D: Final volume image slice TIFF format.

At the time of geometric modelling, it was undecided that the Reachin API platform would be the software package used for visual and haptic rendering of the virtual cochlea implantation. As such, the dataset was saved in three different formats. By applying a Marching Cubes algorithm available for selection from the Surface Extraction module in...
Analyze, a surface description of the volume was generated (Figure C). The volume file wrapper function in Analyze
was used to create a series of two dimensional files in TIFF format (Figure D). These were similar in appearance to the
final image sequence viewed using Image Edit. The final volume itself was stored in the custom AVW format, yet this
format is not generic and is not recognizable in any program other than Analyze.

Geometric modelling of the temporal bone comprises the first stage in system design. Analyze facilitated fast and
accurate volume generation. Completion of this phase provides a validated model for use in the visual and haptic
rendering stages of the control system, which is currently being implemented in the Reachin API.

5. REACHIN API

Reachin API is a software package that facilitates fast and realistic three-dimensional modelling of virtual reality
applications. It is based on the concept of a single scene-graph that combines haptics and graphics. A scene-graph is
basically a description of a scene using a hierarchical data structure. A scene-graph is composed of objects within the
scene and dependencies between objects. Scene-graph primitives are called nodes and information within the nodes are
termed fields. Field networks are established between fields to enable relationships between objects and events occurring
within the scene to be realized. Object geometry, position and appearance may be defined in the scene, as well as
information describing the scene itself, such as light source type. The scene-graph concept can speed up development
time and also offer optimisations so that rendering performance is enhanced.

Programming in the Reachin API primarily uses C++, but is integrated with scripting languages: Virtual Reality
Modelling Language (VRML) and Python. The latter is important for field event handling implementations. Custom
graphic and haptic algorithms are also available for use (implemented in OpenGL).

There are two rendering engines: a haptic rendering loop - the real-time loop - and a graphics loop. No manual
synchronisation between the two loops is required. The real-time loop is updated at a frequency of 1000Hz while the
graphics loop is updated at 30Hz. Although the two loops are separate, the same data is shared between the graphic and
haptic renderings. Older programming environments, such as Ghost/OpenGL (SensAble Technologies), have required
manual synchronisation and separate datasets for the two loops. These two issues are addressed in the Reachin API,
increasing processing speeds and sparing the user synchronisation issues.

6. VISUAL AND HAPTIC RENDERING IN REACHIN

Visual and haptic rendering of the virtual mastoidectomy and cochlea implantation is currently being implemented in the
Reachin API. The geometric model of the temporal bone is imported into Reachin API for this purpose. Benefits of this
application have been discussed in Section 5. Reachin API.

Three-dimensional texture mapping is currently being applied for visual rendering of the mastoidectomy process. The
two-dimensional TIFF image sequence is converted to Portable Network Graphics (PNG) format for import into Reachin
API. Reachin offers image texture mapping nodes, including Texture3D and Texture2D, which support PNG format.
Conversion to PNG from TIFF was easily achieved using an open-source file conversion utility ‘Tiff2Png’. Batch
mode was not supported however and conversion was performed on an individual, image-by-image basis. Reachin class
ImageTexture3D is currently being integrated into the system for visual rendering of bone drilling.

Initially, simplified 6DOF haptics algorithms will be used to simulate the forces and torques involved in the temporal
bone drilling process. Finally, accurate algorithms that mimic forces and torques associated with a true mastoidectomy
will be introduced into the model. These will be based on existing literature and/or mathematical approximations of the
tool/object interactions. Reachin API supports 6DOF haptics algorithms and offers some custom base classes for 6DOF
implementations; however these remain to be extensively tested on a 6DOF device.
7. STEREOGRAPHIC DISPLAY

Virtual reality applications for surgical simulators of this type must be realistic enough that the user feels completely immersed in his or her environment. This is particularly so if the system is to be developed for training purposes. For this work, a system has been designed and installed to enhance the sense of realism. A framework was built in-house (similar to designs of CSIRO, Canberra and Reachin Technologies AB), that supports the monitor and houses a mirror to view the virtual scene so that the haptic device and virtual stylus have co-location. Stereoscopic equipment, including CrystalEyes shutter glasses and Emitter (Direct2U) were purchased and installed to further enhance the experience. Reachin API supports these utilities, where the system was configured to allow for mirroring and stereo view. Graphics driver configurations were changed to enable stereo vision. In the current system, the user views a stereoscopic image of the virtual environment and moves his or her hand in the same space that they see the stylus movement. The installed system is shown below (Figure E).

Figure E: Stereoscopic Display at University of Wollongong, Australia

7. CONCLUSIONS

A clinically valid surgical simulator for training ENT specialists in performing cochlea implantation is currently under development. The first stage in this process has been completed and involved the geometric modelling of the temporal bone, for this purpose. Image processing and volume generation were performed using Analyze. Temporal bone model derivation is the focus of this paper and comprises the initial phase in simulator construction.

Model development began with the scanning of a temporal bone cadaver specimen, followed by registration and processing of the generated CT scan sequence. A semi-automatic threshold-based approach was used to identify and segment the temporal bone from surrounding parenchyma. A volume dataset was produced via linear interpolation between image slices. Elimination of artefacts was easily achieved, by applying a semi-automatic seeded region-grow to the mastoid section. Unnecessary bony structures including cheek bone and vertebrae were removed using the Analyze connected components method. An ENT specialist provided validation of the model at various stages through model progression, including the final stage of design.
In this paper, we first explored the potential benefits of implementing a surgical simulator of this type for educational purposes. Advantages and complications associated with inclusion of haptic-feedback into the model are considered. A brief analysis of existing work in the area of temporal bone segmentation in image processing is given. Methodology associated with geometric modelling of the temporal bone is discussed, accompanied by results. Reachin API is presented as a suitable platform in which to implement visual and haptic rendering of the model. This work is currently being undertaken. The present system uses stereoscopic viewing to enhance realism. The final stage in simulator design will include insertion of the cochlea implant and will be the first accurate force-feedback model of cochlea implant surgery developed to date.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions from staff at the RVEEH, St Vincent’s Private Hospital (Sydney), the Euroa Centre at the Prince of Wales Hospital (Sydney) and the CSIRO (Canberra). A particular mention is given to Cochlear Ltd who are funding this work. Thanks also to colleagues at the University of Wollongong for their support.

REFERENCES

23. Shaik WV, Tiff2Png Version 0.82; http://www.schaik.com/png/tiff2png.html.