Western Graduate&PostdoctoralStudies

Western University

Scholarship@Western

Electronic Thesis and Dissertation Repository

6-21-2019 10:30 AM

Multi-Atlas Segmentation of the Facial Nerve

Bradley M. Gare
The University of Western Ontario

Supervisor
Ladak, Hanif M.
The University of Western Ontario Co-Supervisor
Agrawal, Sumit K.
The University of Western Ontario

Graduate Program in Electrical and Computer Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Engineering Science
© Bradley M. Gare 2019

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Biomedical Commons

Recommended Citation

Gare, Bradley M., "Multi-Atlas Segmentation of the Facial Nerve" (2019). *Electronic Thesis and Dissertation Repository*. 6290.
https://ir.lib.uwo.ca/etd/6290

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

Abstract

Medical image segmentation is an important step to identify the shape and position of patient anatomy prior to surgical simulation, surgical rehearsal, and surgical planning. It is crucial that the facial nerve (FN) is segmented accurately as damage to this nerve can severely impact facial expression, speech, and taste. Manual segmentation provides accurate results but is time-consuming and labor-intensive; semi-automatic methods of segmentation are more feasible in a clinical setting and can provide accurate results with minimal user involvement.

The objective of this work was to create a novel, open-source, multi-atlas based segmentation algorithm of the entire FN requiring minimal user intervention.

Twenty-eight temporal bones were segmented producing an average Dice metric of 0.76 and an average Hausdorff distance of 0.17 mm which is similar to previously published algorithms. These results indicate that this segmentation approach can accurately segment the FN and greatly reduce time spent with manual segmentation.

Keywords: Facial Nerve, Image Segmentation, Image Registration, Surgical Simulation, Otolaryngology, Atlas-Based Segmentation

Co-Authorship Statement

This master's thesis is an integration of two journal articles. The first article, Chapter 2, is currently submitted and under review in the Journal of *Otology & Neurotology*. The second article, Chapter 3, is currently submitted and under review in the *International Journal of Computer Assisted Radiology and Surgery*.

Chapter 2: Hudson T, Gare BM, Allen DG, Ladak HM, Agrawal SK. Intrinsic Measures and Shape Analysis of the Intratemporal Facial Nerve. *Otology & Neurotology*. Submitted February 25, 2019.

The initial motivation of this study was presented by S. Agrawal and H. Ladak and the primary work of writing, experiment design, and analysis was carried out by T. Hudson. I made two contributions to this work. The first contribution was generating a principle component analysis (PCA) model to augment the results from a PCA T. Hudson had created with the assistance of D. Allen. The second contribution was in evaluating the results of the PCA I created, as well as writing about the PCA and generating figures with it.

Chapter 3: Gare BM, Hudson T, Rohani SA, Allen DG, Agrawal SK, Ladak HM. Multi-Atlas Segmentation of the Facial Nerve from Clinical CT. *International Journal of Computer Assisted Radiology and Surgery*. Submitted April 29, 2019.

The motivation of this study was to investigate the impact of micro-CT atlases and their effect on segmentation algorithms and was proposed by H. Ladak and S. Agrawal. My contribution to this study was the development of the algorithm, analysis of results, and writing of the paper. T. Hudson's contribution was the generation of the ground truth segmentations which were used as a basis to compare the results of my algorithm. D. Allen's contribution was consultation as he worked in parallel to segment different anatomy. A. Rohani's contribution was figure generation and consultation.

Acknowledgments

Firstly, I would like to thank both of my supervisors. Dr. Hanif Ladak has been an amazing supervisor during the extent of my Master's degree and I consider myself extremely fortunate to have had his supervision in these last two years. I was excited to have him as a supervisor before beginning my research due to my experiences with him as a student and having him as a supervisor has only made me respect him more. I was also very fortunate to work under the supervision of Dr. Sumit Agrawal. His ability to combine expertise from a clinical setting with engineering knowledge makes him a perfect supervisor for ensuring that research will be focused towards finding solutions which will have a meaningful impact.

I would also like to thank all of the members of the Auditory Biophysics lab. The former members, Dr. Mai Elfarnawany, John Iyaniwura, Robert Koch, and Carlos Salgado all helped to make the lab a great place to work and were always willing to lend a hand when I was beginning my Master's degree. Current lab members Dr. Alireza Rohani, Dr. Soodeh Nikan, Daniel Allen, and Luke Helpard have also been a pleasure to work with and have always been willing to help me out during my degree. I am grateful to have worked beside all of you and to have learned from you.

Lastly, I would like to thank my family, my girlfriend, and my friends. All of you have been extremely supportive of me and have gone above and beyond to make these last couple years as comfortable for me as possible. I am thankful that all of you were not just here for me, but also looked for ways to do anything you could to make my experience better.

Table of Contents

List of Figures

List of Abbreviations

3D   Three-dimensional

AHD   Average Hausdorff Distance

CT   Computed Tomography

CHL   Conductive Hearing Loss

CI   Cochlear Implants

DSC   Dice Similarity Coefficient

FN   Facial Nerve

HD   Hausdorff Distance

MHL   Mixed Hearing Loss

MI   Mutual Information

micro-CT   Micro Computed Tomography

MSE   Mean Squared Error

NCC   Normalized Cross Correlation

NMI   Normalized Mutual Information

PCA   Principle Component Analysis

ROI   Region of Interest

SNHL   Sensorineural Hearing Loss

Voxel   Volumetric Pixel

Chapter 1

1 Introduction

Hearing loss is a disabling condition affecting many individuals in Canada and world-wide. In 2012, a Canadian survey on disability reported 4% of Canadians as having a hearing disability and 2.3% as being deaf or having no hearing at all (1). The most common treatment for hearing loss is the use of hearing aids, however some individuals are unable to use or do not attain sufficient improvement from hearing aids, for example in cases of profound deafness. In these cases, a cochlear implant (CI) can be used to directly stimulate nerve fibers leading to the brain, creating the sensation of sound much like sound conduction occurs in air in a healthy cochlea. The most common method of implanting a CI is through a surgical procedure called a mastoidectomy, where the mastoid air cells are removed by drilling to allow the insertion of an electrode array through the round window into the cochlea (2,3). In performing a mastoidectomy, one of the difficulties encountered while drilling is to avoid damaging anatomy such as the facial nerve (FN), which is responsible for facial expression and taste. In order to avoid injury during a mastoidectomy, it is crucial that surgically significant anatomy is identified prior to surgery using surgical simulation and medical image segmentation.

Medical image segmentation is the process of identifying the shape and position of anatomical structures in medical images, such as computed tomography (CT) images. A number of methods can be used to segment anatomy (4). Simple methods such as thresholding can be used when structures have consistent or distinct intensity values, however in images with varying intensities or indistinct boundaries, these methods fail to identify the distinctions among various structures. In cases where region boundaries are unclear, methods which consider shape information can provide more plausible segmentations, such as shape models.

However, the use of shape models still leads to difficulty in locally conforming and fully utilizing intensity information (5). To effectively segment anatomy, it is important that both image and shape information are utilized.

This chapter provides relevant background material upon which this thesis work is based.

1.1 Anatomy

1.1.1 The Human Temporal Bone

The temporal bone consists of four parts: the squamous, mastoid, and tympanic parts on the outside of the skull as shown in Figure 1.1, and the petrous part inside the skull. The squamous part is the largest part, located superior to the rest of the temporal bone. It connects posteriorly through the zygomatic process into the zygomatic bone (cheekbone) and also has the mandibular fossa which is a depression which the lower jaw connects to. The mastoid part is located posterior to the squamous part and it can be felt behind the ear. It is known to be a very pneumatized area, having a significant number of air cells and a high degree of variability with respect to the number of the air cells and their size (6). The tympanic part of the temporal bone is located inferior to the squamous part and surrounds the external part of the ear canal. The petrous part located within the skull houses the middle ear containing the tympanic cavity and ossicles and inner ear containing the cochlea and semicircular canals.
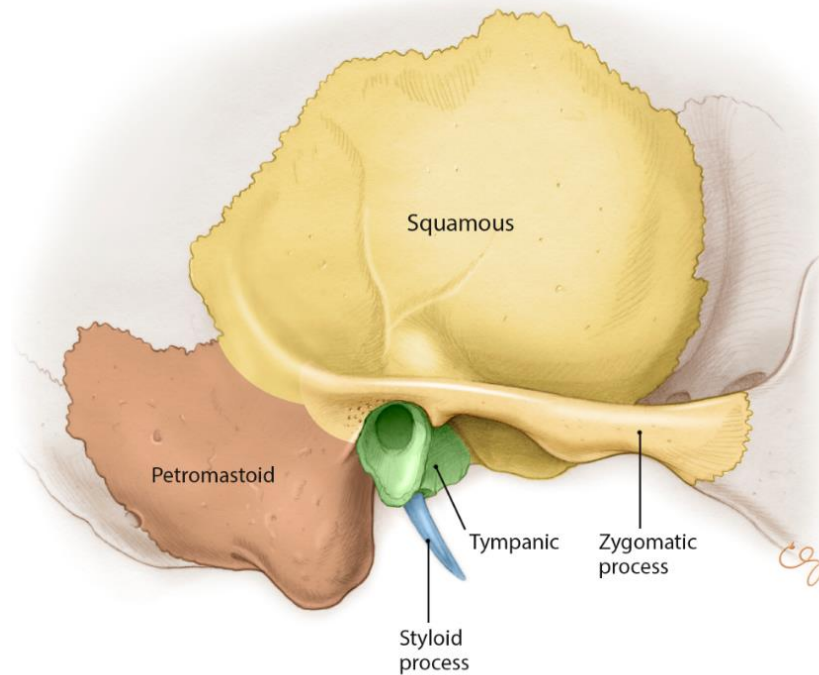
Figure 1.1: Anatomy of the temporal bone showing the squamous part (yellow), mastoid part (brown), and tympanic part (green). Image Courtesy of Stanford Otolaryngology, Head and Neck Surgery (7)

1.1.2 The Facial Nerve

The FN is a thin tubular structure which travels between several structures in the middle and inner ear as seen in Figure 1.2. The FN enters the temporal bone inferiorly at an orifice called the stylomastoid foramen located between the mastoid and squamous parts of the temporal bone. From here, the mastoid segment of the FN moves upward, moving behind the tympanic membrane (eardrum) and the malleus, incus, and stapes until it reaches a bend called the second genu. From here, the tympanic segment of the FN moves further into the ear superior to the cochlea until reaching the labyrinthine segment where it forks and the main nerve travels with the audiovestibular nerve while the petrosal nerve branches off toward the mouth, eyes, and nose to gather other sensory information. While the FN maintains these positional relationships

with other structures in the temporal bone, it still possesses significant variance which presents a challenge for surgeries (8).
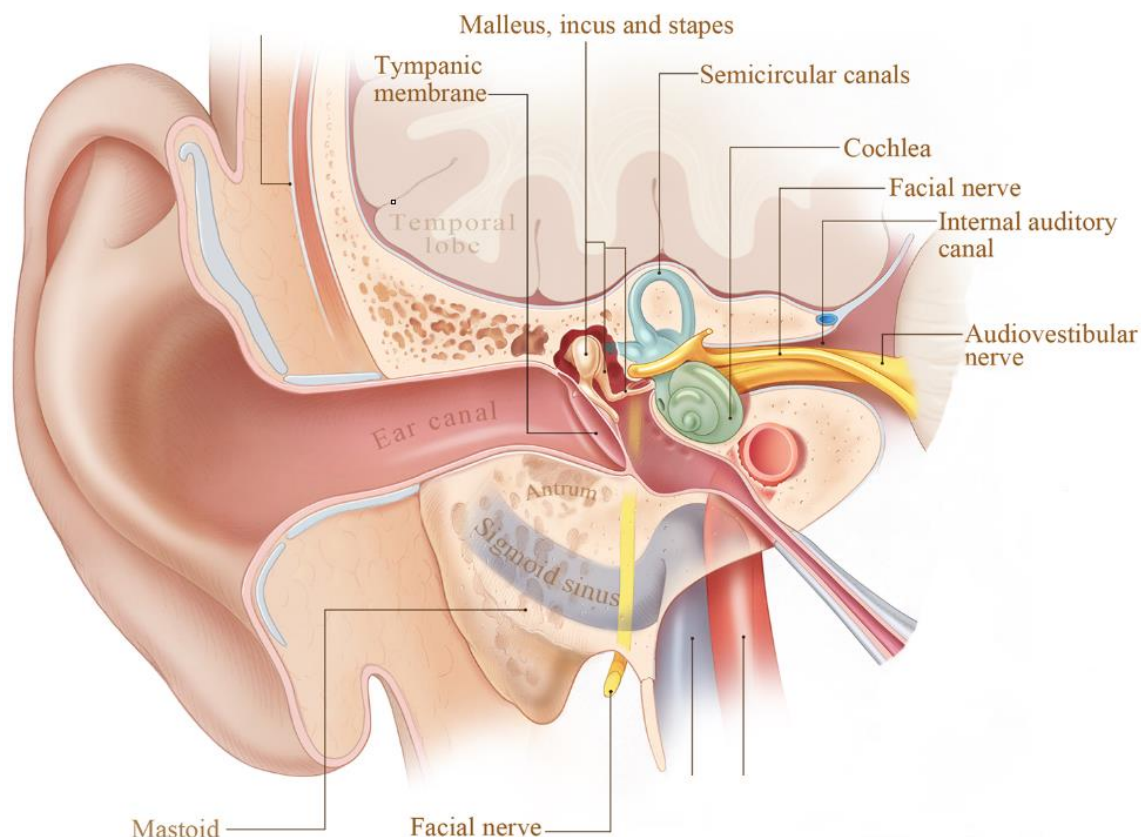
Figure 1.2: Anatomy of the outer, middle, and inner ears with facial nerve shown. Image courtesy of Stanford Otolaryngology, Head and Neck Surgery (7)

1.1.3 Facial Nerve Variation

The FN has significant anatomical variation which is later discussed in Chapter 2. The greatest variation of the FN exists in the angle that the mastoid segment forms with the tympanic segment (Figure 1.3). There is also significant variance perpendicular to the direction of the angle formed by the mastoid and tympanic segments as well variation in the length of the mastoid segment. Other noteworthy variances exist in the length of the tympanic segment. The labyrinthine segment, while having less overall variation, has the most variation with respect

to its size particularly in the length from the geniculate ganglion to the internal auditory meatus. Diameters of each segment also show some variability and branches of the FN such as the chorda tympani and the greater petrosal nerve also demonstrate variation in both the angle and point at which they exit the FN. Finally, the FN is commonly dehiscent (9), meaning that the bone surrounding it can often be discontinuous.
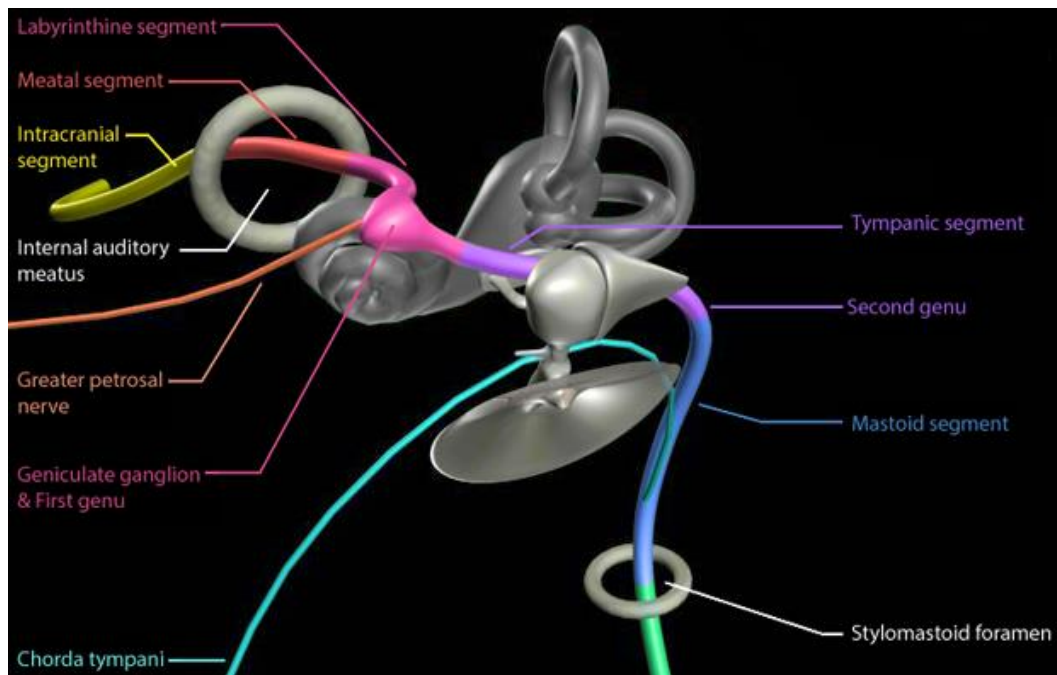
Figure 1.3: Facial segments and surrounding anatomy. Image courtesy of Schulich School of Medicine & Dentistry, Western University (10).

1.2 Hearing Loss and Treatments

There are three main types of hearing loss: conductive hearing loss (CHL), sensorineural hearing loss (SNHL), and mixed hearing loss (MHL) (11). CHL occurs due to difficulties transferring sound waves along the pathway through the ear canal, the tympanic membrane (eardrum), or the ossicles (malleus, incus, and stapes). SNHL is caused by issues in the inner ear or cochlea and can often be caused by damaged hair cells which if functioning properly, would transform vibrations in the cochlea into electric signals and provide the brain with sound

information. MHL is a combination of CHL and SNHL. Different types of hearing loss require different treatments, but in the case of SNHL, devices such as CIs (Figure 1.4) or hearing aids could be used. Hearing aids are typically used in cases of minor hearing loss while CIs are used for profound deafness. CIs are used to directly stimulate the auditory nerve, bypassing the cochlear hair cells and sending sound information directly to the brain.
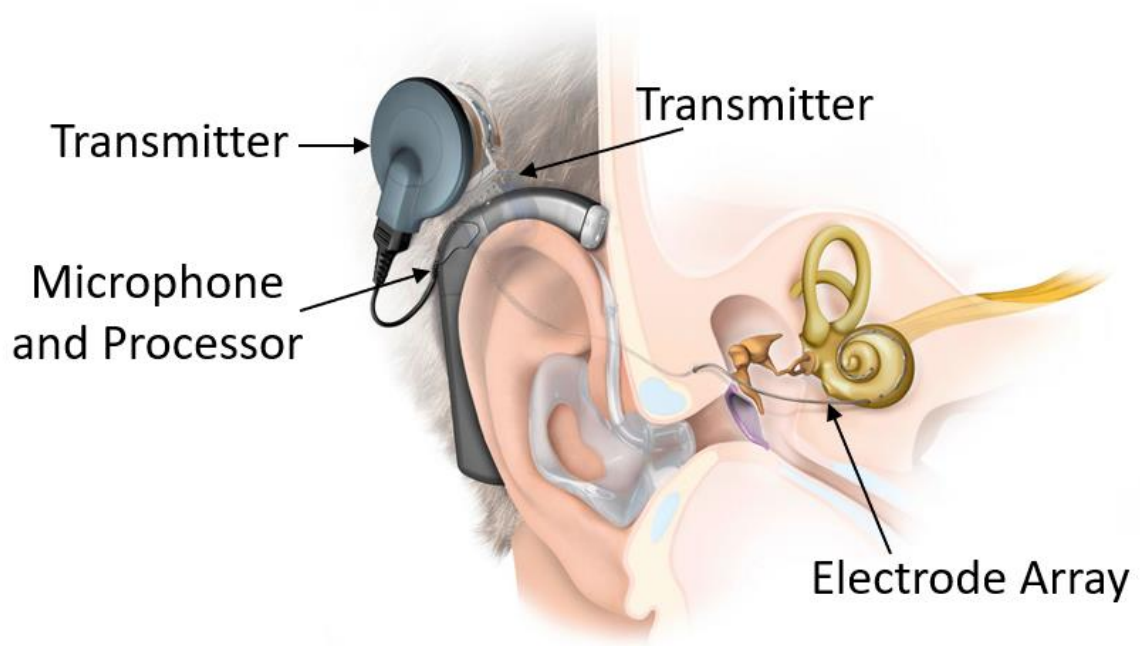
Figure 1.4: Cochlear implant components showing components that are worn externally and those that are internal. Image courtesy of MED-EL GmbH (12)

1.3 Mastoidectomy and Cochlear Implant Insertion

To insert a CI, a surgical procedure called a mastoidectomy must first be performed by drilling away the mastoid air cells from the temporal bone. The mastoid bone can be accessed by pulling the ear forward and making an incision in the skin behind the ear. Once the mastoid bone has been exposed, the mastoid air cells are drilled away until reaching the middle ear.

When the mastoidectomy has been completed, the middle ear will be accessible. The surgeon must then drill through the facial recess, the gap between the FN and the chorda tympani branching off of the FN, which is typically less than 2 mm wide at the round window. This is done in order to gain access to the round window of the cochlea, the location where the CI electrode array is placed. Once the round window has been exposed, the CI's electrode array can be threaded through the round window into the cochlea. Figure 1.5 illustrates the drilling trajectory.
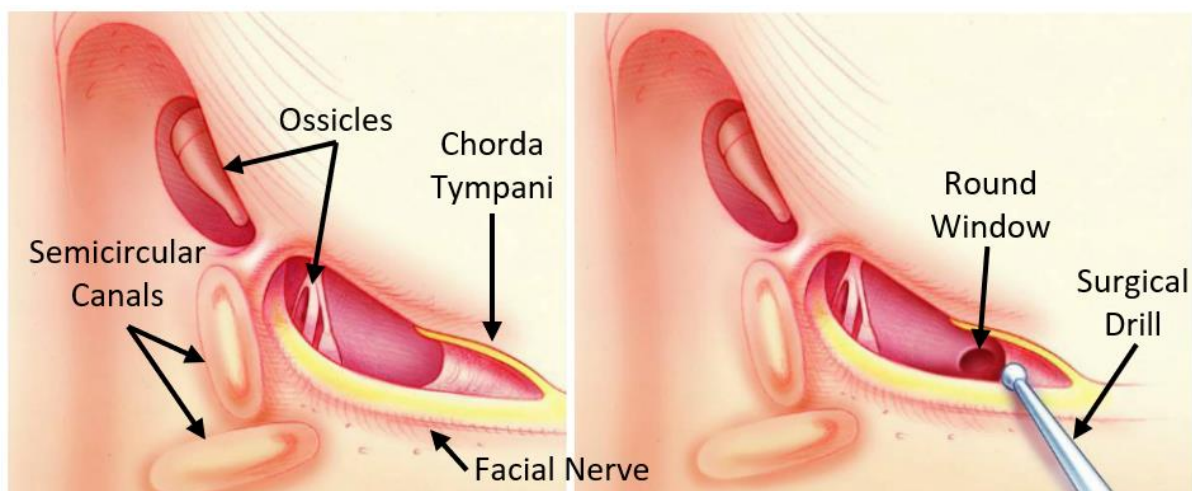
Figure 1.5: Facial recess before (left) and after (right) drilling towards the RW. Image courtesy of Schulich School of Medicine & Dentistry, Western University (10)

In addition to inserting the CI's electrode array, a receiver must be placed into a well drilled into the bone behind the ear. This receiver is connected to an external transmitter which sends a digitized signal of sound signals picked up by the external microphones. The receiver then collects the signals and converts them to electrical pulses which are sent to the electrode array. Finally, the electrode array directly stimulates the nerve fibers in the cochlea and the brain interprets these signals as if they were sound.

1.4 Surgical Simulation

Surgical simulators are tools used in medical education to allow surgical trainees to learn and practice difficult surgical procedures such as mastoidectomy as well as the facial recess approach (13). Surgical simulators offer an alternative to practicing on real patients with associated risk, or on cadavers which are costly and difficult to obtain. Surgical simulators work by uploading specific surgical environments or scenarios into a simulator along with labelled anatomy. Trainees can then practice procedures in a virtual environment with minimal risk. Virtual -reality technologies such as stereoscopic displays and haptic feedback can help to enhance the learning experience and more closely represent real-life procedures. Additionally, with advances in technology, the cost of surgical simulators has significantly decreased (14).

Another benefit of surgical simulators is that they are not simply limited to one example case with general anatomy, which may not represent real-life cases that will be encountered in the operating room. In surgical simulators, multiple different anatomical cases can be incorporated into simulators so that a surgeon can prepare for anatomy which is more representative of a real-life procedure. In addition, patient-specific anatomy can be uploaded allowing the surgeon to practice a procedure on a specific patient's anatomy before performing the procedure on that patient. While surgical simulation of procedures with a patient's specific anatomy is ideal, it requires difficult medical image segmentation and labelling.

1.5 Medical Image Segmentation

Medical Image segmentation is the process of partitioning an image in such a way that individual structures are labelled and are easily identifiable. In this work, we will consider

segmentations to be labelmaps or a 3D matrix representing every volumetric pixel (voxel) in a medical image volume and identifying it as a 1 if it is part of a structure of interest and a 0 if not. This is done to isolate structures of interest which can then be analyzed in terms of shape and position. Medical image segmentation is important for several reasons such as measurements, treatment planning, and in our work, the ability to incorporate a digital structural model into a surgical simulation environment.

1.5.1 Segmentation Methods

There are three main categories of segmentation methods: manual segmentation, semi-automatic segmentation, and automatic segmentation; all methods have their own strengths and weaknesses. Manual segmentation is the most time-consuming method as it requires a user to use a three-dimensional (3D) medical image volume to manually label each structure in every slice of the image. The time required to manually segment images may not be feasible in a real-life surgical workflow. While this method is extremely time-consuming, typically the results produced are of high quality. Additionally, when dealing with difficult image volumes (i.e. images with low contrast, artifacts, and irregular anatomy), a user may be more likely to be able to segment the image properly by utilizing knowledge of shape and orientation as well as the position of structures with respect to others.

In comparison, automatic segmentation can be used to provide image segmentations with minimal user involvement. This is highly desirable, as it requires significantly less user time which can then be used for surgical planning or rehearsal. Difficulties with automatic segmentation can come from the fact that users may not be certain of the segmentation quality of any new segmentation before inspecting it, and programming automatic segmentation tools requires knowledge of information which trained users use to manually segment image volumes as well as the ability to code it. Semi-automatic image segmentation offers a valuable

compromise between manual and automatic segmentation. In semi-automatic segmentation, a user initializes a segmentation by providing information about the desired segmentation and then the rest of the segmentation is automatically produced. Examples of initialization are partially segmenting a volume or placing identifier points along the path of the segmentation. Depending on the semi-automatic method used, the user involvement can vary from less than a minute to the same amount of time as manual segmentation based on the difficulty of the image volumes. A benefit of semi-automatic segmentation over automatic segmentation is that the user can be more assured of segmentation quality since they have helped to make sure that the segmentation has been initialized properly. The result is usually a quality segmentation while not sacrificing copious amounts of time.

1.5.2 Segmentation Evaluation

Medical segmentations are produced for various purposes such as surgical simulation, surgical rehearsal, and surgical planning. Depending on what the segmentations are being used for, different levels of accuracy are acceptable. In order to evaluate the accuracy of the segmentations, various metrics are used to quantitatively evaluate items like volume overlap and closeness of boundaries (15).

One common measure of segmentation accuracy is the Dice Similarity Coefficient (DSC). Equation 1 shows how the DSC is calculated between two segmentations A and B where bars around the segmentations represent the total voxels for segmentations A and B respectively in the denominator, while $A \cap B$ represents overlapping voxels.

$$DSC(A, B) = 2\ \frac{|A \cap B|}{|A| + |B|}$$
<div align="right">Equation 1.1</div>

An ideal value for DSC would be 1.0 as that would imply that all of the voxels in segmentation A overlap with all of the voxels in segmentation B and that both segmentations contain the same number of voxels.

Other measures of segmentation accuracy are the Hausdorff Distance (HD) and average Hausdorff Distance (AHD). The directed HD representing the maximum distance from one set of points to the closest corresponding point in the other set is defined by Equation 1.2 where a represents the segmentation voxels in one image, b represents the closest segmentation voxels in the other image, and $\|d(a,b)\|$ represents the maximum of the closest voxels. HD is the maximum of the two directed HDs and is given by Equation 1.3.

$$h(A,B) = \max_{a \epsilon A} \min_{b \epsilon B} \|d(a,b)\| \qquad\qquad \text{Equation 1.2}$$

$$HD(A,B) = \max(h(A,B), h(B,A)) \qquad\qquad \text{Equation 1.3}$$

HD is an effective way of finding the maximum error between two segmentations, however, it is often more useful to consider the average error as the HD can be susceptible to outliers. To measure the average error, average HD (AHD) is used. AHD (Equation 1.5) is calculated using the directed average HD shown in Equation 1.4.

$$d(A,B) = \frac{1}{N} \sum_{a \epsilon A} \min_{b \epsilon B} \|a - b\| \qquad\qquad \text{Equation 1.4}$$

$$AHD(A,B) = \max(d(A,B), d(B,A)) \qquad\qquad \text{Equation 1.5}$$

The AHD metric is an indicator of the closeness of the boundary of each segmentation to the other.

1.6 Image Registration

Image registration is the process of aligning one image with another image. Registration can be useful for transferring labelmaps, or segmentations, from a labelled image to an unlabeled image and is the approach used in this thesis. The goal of registration is to make both images overlap each other in the most meaningful way possible by not just providing maximum overlap of the two images, but by attempting to maximize the overlap between features in one image with the same features in another image, regardless of how well the entire images

overlap. Typically, registration is an iterative process; a fixed image will be chosen as a target image and a second image will be chosen as the moving image which will be continuously transformed until it best fits the target image. The registration will occur in steps where the moving image is transformed using an optimizer. The fit of the registration is measured with a cost metric, and then if the cost is less than it previously was, the transformed image will be made permanent. This transformed image will be the new moving image and a new transform will be calculated to transform the transformed image to best match the fixed image. A series of transformations and cost calculations will be done until a maximum number of iterations has been reached, or until the improvement in the cost metric falls below a chosen threshold at which it is considered to be a negligible improvement.

1.6.1 Optimizer Functions

Optimizers are numerical tools for solving or achieving the best score for a given cost metric which evaluates the similarity between corresponding points in two images. This is done by moving one image such that features in that image overlap with the same features in the other image. Optimizers are used to determine how an image should be shifted around so that the cost metric measuring the overlap between the two images reaches a locally optimal value for the given cost metric. Some optimizers for determining how the image should shift are gradient descent (16), Newton's method, and Quasi-Newton methods (17). While gradient descent moves along a gradient in order to find the locally minimal value and then updates and repeats, Newton's Method uses the image gradient and a Hessian matrix of second derivatives of the function. Quasi-Newton methods are similar to Newton's method but do not recompute the Hessian matrix at each iteration.

1.6.2 Cost Metrics

To measure the overlap between the two images, or how well they match, quality or cost metrics are implemented (18,19). These metrics help to quantify and evaluate the current overlap by comparing intensities between overlapping regions or by comparing them in a probabilistic measure.

Mean Squared Error (MSE) and Normalized Cross-Correlation (NCC) are examples of simple cost metrics which only consider differences in intensity between two overlapping regions. MSE (Equation 1.6) computes the squared differences in image volumes A and B with respective voxels located at (x, y, z) while NCC (Equation 1.7) is similar but insensitive to shifting and scaling by subtracting the means $\bar{A}$ and $\bar{B}$ and dividing by variances σ_A and σ_B.

$$MSE(A,B) = \frac{1}{N}\sum_x \sum_y \sum_z (A(x,y,z) - B(x,y,z))^2 \qquad \text{Equation 1.6}$$

$$NCC(A,B) = \frac{1}{N}\sum_x \sum_y \sum_z \frac{(A(x,y,z)-\bar{A})(B(x,y,z)-\bar{B})}{\sigma_A \sigma_B} \qquad \text{Equation 1.7}$$

In the case of MSE, a perfect value would be 0, however, due to image noise and differences between images, a value of 0 is not typically achieved in practice. Instead, the optimizer function is used to iteratively minimize MSE. Unlike MSE, NCC is a measure of similarity between images, meaning that a higher value represents a better fit between two images. In the case of NCC, the optimizer iteratively seeks to increase the NCC to get as close as possible to a perfect correlation of 1.0.

A more complicated measure, Mutual Information (MI) is used to assess how much can be known about one image given the information included in the other. While MSE and NCC are faster metrics to implement and execute, they fail to consider how intensities and intensity distributions can change from one image modality to another and therefore are less likely to converge to an optimal solution when registering image volumes produced with

different imaging modalities. MI takes into account intensity values and probabilistic information which makes it much more suitable for registering images with different modalities. MI is calculated using the entropy equation given in Equation 1.8 representing the entropy in a single image where the information gained, $H(A)$, is given by calculating the negative sum of the probability of all voxels, or a values, multiplied by the log of that probability of those voxel values. Equation 1.9 represents entropy between two images.

$$H(A) = -\sum_a p(a) \log_2 p(a) \qquad\qquad \text{Equation 1.8}$$

$$H(A, B) = -\sum_a \sum_b p(a, b) \log_2 p(a, b) \qquad\qquad \text{Equation 1.9}$$

The entropy in a single image should approach 0 if the image has low entropy or the intensity values can be easily predicted. Equation 1.10 shows normalized MI (NMI) which is typically chosen because it avoids degenerate solutions where entropy is minimized by avoiding overlap, while at the same time, normalizes the results such that excessive overlap is not encouraged. An optimal value for NMI is 2 and a worst-case value is 1.

$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)} \qquad\qquad \text{Equation 1.10}$$

1.6.3 Image Transforms

To achieve the best score for each cost metric, optimizers are used to determine how an image should be transformed such that the cost metrics are minimized or maximized. There are two broad categories of transforms which can be used to optimize the cost metric: linear transforms and non-rigid transforms (20). These transforms simply move or modify one image, the moving image, to optimally overlap the fixed or target image. The most common linear transforms are rigid transforms which will shift and rotate an image and affine transforms which will shift, rotate, scale, and shear images to maximize overlap. Non-rigid transforms instead locally warp images such that local shape variability can be more accurately represented.

1.6.3.1 Linear Transforms

Linear transformations are often used to align two images. Figure 1.6 demonstrates a rigid transformation which is the simplest form of linear transform including only translation and rotation. Similarity transforms are more complex allowing for translation, rotation, and scaling. Affine transformations are even more complex than similarity transformations allowing for translation, rotation, scaling, and shear. In the case that you have two identical objects, a linear transformation could be used in order to make them align perfectly. In the case where two images are not the same, yet have some similarities, linear transformations can be useful to provide some overlap. In cases where one has two different images or objects, linear transformations are an effective way to ensure that both images start in the same physical space so that more sophisticated transforms such as non-rigid transforms can take place.
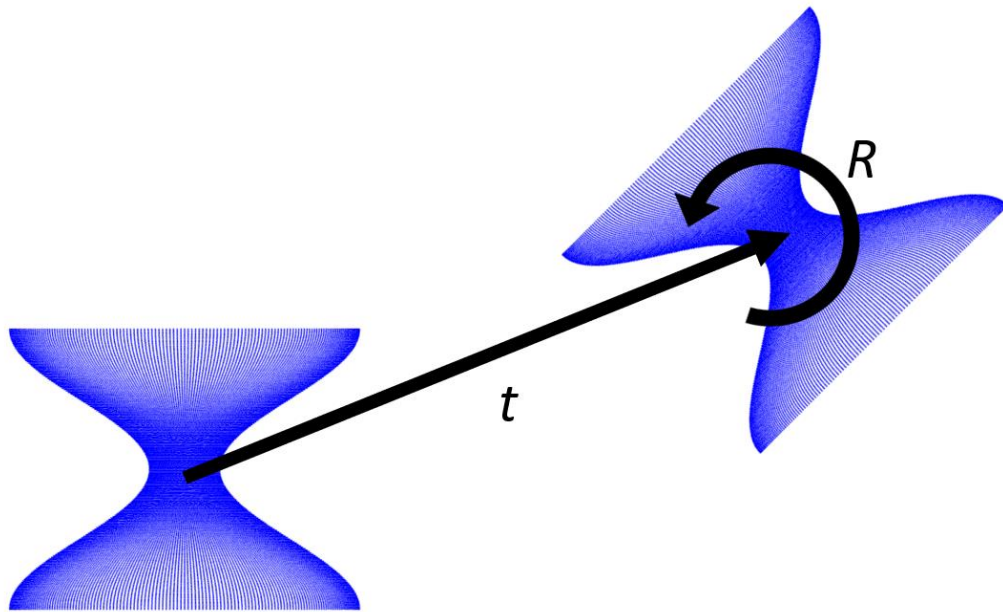
Figure 1.6: Rigid transformation. Translation is shown (t) as well as rotation (R).

1.6.3.2 Non-rigid Transforms

In cases where one has two approximately similar images and they have been aligned well, a non-rigid, or deformable transformation can be used to locally transform an image such that local differences are minimized. There are a significant number of methods which can be used to produce non-rigid transforms. Figure 1.7 demonstrates this concept with a B-Spline transformation whereby a grid made up of many control points is deformed until the two images overlap optimally.

Non-rigid transformations are able to account for local details such as curves and can transform two images or objects with some similarity to become more similar. The drawbacks for non-rigid transforms are that they are far slower than rigid transformations due to the amount of computations required and they also present the risk of unrealistic deformation. In order to avoid undesirable deformation, registration parameters must be chosen carefully.

Moving Image   Target Image   Result
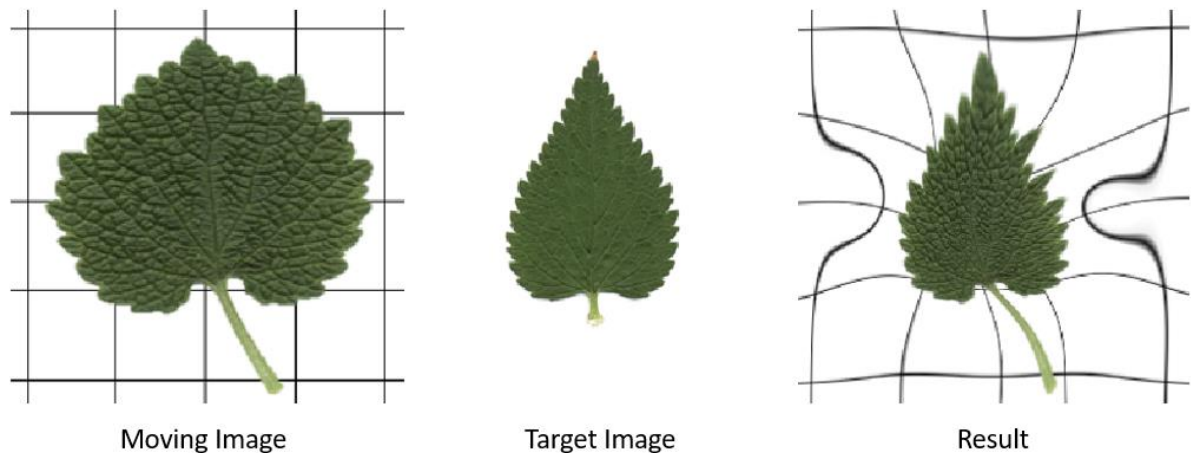
Figure 1.7: Deformable registration.

1.6.4 Iterative Closest Point Registration

The iterative closest point (ICP) transformation (20) is an example of a rigid registration technique. This technique does not consider images or intensities, but instead considers point clouds: a fixed or target point cloud, and a moving or source point cloud. It is discussed here as it is used in Chapter 3. ICP considers every point of the moving point cloud and matches those points to the closest points in the target point cloud. Once this is done, the algorithm measures the error between corresponding points and estimates both the rotation and translation which will best align every source point to the chosen matching point. Once the rotation and translation have been calculated, the moving point cloud is transformed using the transformation and the new error is measured. Then, the process is repeated for the moving cloud from its new position. This process is repeated until either a maximum number of iterations has been reached or until the improvement in error between transformations becomes smaller than a chosen tolerance meaning that the change is deemed insignificant. Figure 1.8 demonstrates an example of the iterative closest point algorithm where on the left, corresponding points are found, and on the right, rotation and translation have taken place resulting in a moved point cloud and new closest points.
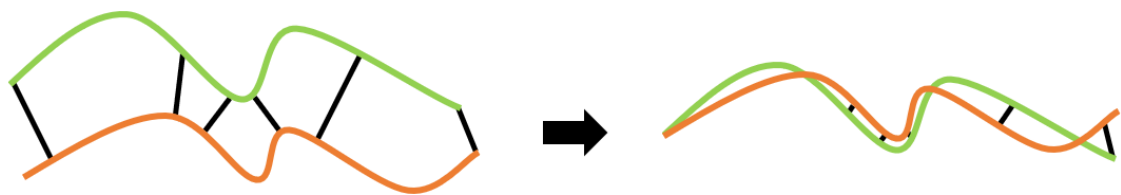
Figure 1.8: Iterative Closest Point algorithm. Unregistered shown left and registered shown right.

1.6.5 Atlas-based Segmentation

Atlas-based segmentation is a powerful segmentation technique. The principle behind this technique is that an atlas, or labeled image volume serves as an example segmentation, and this example segmentation is registered to a new unlabeled image volume.

One of the benefits of using atlas-based segmentation is that the segmentation achieved is based on a manually segmented volume and as such, it takes image information and labels which have been verified to accurately represent realistic anatomy and image appearance. Therefore, during the registration process, the manual segmentation should be transformed in such a way that the segmentation produced for the new image volume should be realistic as well.

Atlas-based segmentation performs well when the labeled and unlabeled regions in the atlas image volume and the image volume to be segmented have similar intensity distributions, and the anatomy represented in the atlas image volume is similar to the anatomy in the image volume to be segmented. In these cases, segmentations in the atlas can be easily transferred to the target image before non-rigid transformations. In cases where the atlas image volume and image volume to be segmented have differing intensity distributions and anatomy, atlas-based segmentations may produce unrealistic segmentations, so a single atlas may not effectively segment all new images (21). This can be mitigated by combining atlas-based approaches with other image segmentation techniques, or by using multi-atlas-based segmentation whereby multiple labeled atlas-volumes are used.

Multiple atlases can be used in two different ways. The first is a computationally intensive method which registers every atlas image volume to the image to be segmented and then utilizes all of the segmentations to produce a final segmentation. The second method is much less computationally demanding and instead of registering every atlas image volume to

the image volume to be segmented, atlas candidates are chosen based on certain criteria and the atlas which is the best fit for these criteria is chosen as the atlas to be used for registration.

1.6.6 Production of Atlas Sets

To select the medical images volume to be used for an atlas or an atlas set, it is important to ensure that as much anatomical variability can be represented as possible. New image volumes to be segmented can be flipped if on the opposite side to the atlas set and then be segmented. Once the segmentation has been completed, the image volume and produced segmentations can simply be mirrored back to fit the original image volume.

1.7 Objectives

The primary objective of this work was (1) to develop a segmentation software tool to accurately segment the FN with minimal user involvement that could be used with a surgical simulation platform. The secondary objectives of this work were (2) to investigate the benefits of using high resolution micro-CT images to create atlases for atlas-based segmentation and (3) to investigate the benefit of providing multiple atlas candidates to overcome previous challenges when using atlas-based segmentation to segment the FN.

1.8 Hypothesis

It was hypothesized that previous challenges associated with atlas-based semi-automatic segmentation of the FN could be overcome by using two unique methods; the first method was to use micro-CT image volumes to create high resolution atlases, and the second method was to use multiple atlas candidates so that an ideal atlas image volume could be selected.

1.9 Novel Contributions

The novel contributions of this work are that the proposed algorithm is the first algorithm to utilize multiple atlases and the first algorithm to use atlases formed with high-resolution micro-CT images. Additionally, the most challenging regions of the intratemporal facial nerve to segment, the stylomastoid foramen and labyrinthine segment, which were not segmented in previous algorithms, are now segmented resulting in a segmentation of the entire intratemporal facial nerve.

References

1. Statistics Canada. Statistics Canada. [Online].; 2016. Available from: https://www150.statcan.gc.ca/n1/daily-quotidien/160229/dq160229c-eng.htm.

2. Kronenberg J, Migirov L. The Role of Mastoidectomy in Cochlear Implant Surgery. Acta Oto-Laryngologica. 2003 February; 123(2).

3. Yawn R, Hunter JB, Sweeney AD, Bennett ML. Cochlear implantation: a biomechanical prosthesis for hearing loss. F1000prime reports. 2015 April; 7.

4. Norouzi A, Rahim MSM, Altameem A, Saba T, Rad AE, Rehman A, et al. Medical Image Segmentation Methods, Algorithms, and Applications. IETE Technical Review. 2014 May; 31(3).

5. van Ginneken B, Frangi AF, Staal JJ, ter Haar Romeny BM, Viergever MA. Active shape model segmentation with optimal features. IEEE Transactions on Medical Imaging. 2002 August; 21(8).

6. Hindi K, Alazzawi S, Raman R, Prepageran N, Rahmat K. Pneumatization of Mastoid Air Cells, Temporal Bone, Ethmoid and Sphenoid Sinuses. Any Correlation? Indian Journal of Otolaryngology and Head & Neck Surgery. 2014 December; 66(4).

7. Jackler R, Gralapp C. Stanford Medicine. Otologic Surgery Atlas. [Online].; 2017 [cited 2019 April 28. Available from: http://med.stanford.edu/sm/ohns-otologic-surgery-atlas/ear-anatomy/.

8. Kalaiarasi R, Kiran AS, Vijayakumar C, Venkataramanan R, Manusrut M, Prabhu R. Anatomical Features of Intratemporal Course of Facial Nerve and its Variations. Cureus. 2018 August; 10(8).

9. Selesnick SH, Lynn-Macrae AG. The incidence of facial nerve dehiscence at surgery for cholesteatoma. Otology & Neurotology. 2001 March; 22(2).

10. Cheng H, Agrawal S, Nikolas Blevins RJ, Gralapp C. Neurotology Learning Module. [Online]. [cited 2019 April 28. Available from: http://www.entonline.ca/neurotology/index.htm.

11. Shearer AE, Hildebrand MS, Smith RJ. Hereditary Hearing Loss and Deafness Overview. In Pagon RA, Bird TD, Dolan CR, Stephens K, Adam MP, editors. GeneReviews. Seattle: University of Washington; 2008.

12. MED-EL. MED-EL Image Gallery. [Online].; 2019 [cited 2019 April 28. Available from: https://www.medel.com/in/image-gallery/.

13. Wiet GJ, Stredney D, Kerwin T, Hittle B, Fernandez SA, Abdel-Rasoul M, et al. Virtual temporal bone dissection system: OSU virtual temporal bone system. The Laryngoscope. 2012 March; 122(S1).

14. George AP, De R. Review of temporal bone dissection teaching: how it was, is and will be. The Journal of Laryngology and Otology. 2010 February; 124(2).

15. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Medical Imaging. 2015 August; 15(1).

16. SpringerLink. Image-Guided Interventions: Technology and Applications Peters T, Cleary K, editors. Boston: Springer Science+Business Media; 2008.

17. Klein S, Staring M, Pluim JPW. Evaluation of Optimization Methods for Nonrigid Medical Image Registration Using Mutual Information and B-Splines. IEEE Transactions on Image Processing. 2007 December; 16(12): p. 2879-2890.

18. Thakur KV, Damodare OH, Sapkal AM. Identification of Suited Quality Metrics for Natural and Medical Images. Signal & Image Processing : An International Journal. 2016 June; 7(3).

19. Pluim JPW, Maintz JBA, Viergever MA. Mutual-information-based registration of medical images: a survey. IEEE Transactions on Medical Imaging. 2003 August; 22(8).

20. Fitzpatrick JM, Sonka M. Handbook of medical imaging. Volume 2, Medical image processing and analysis. 1st ed. Fitzpatrick JM, Sonka M, editors.: SPIE Publications; 2000.

21. Suri JS, Wilson DL, Laxminarayan S, editors. Handbook of Biomedical Image Analysis New York: Kluwer Academic / Plenum Publishers; 2005.

Chapter 2

2 Intrinsic Measures and Shape Analysis of the Intratemporal Facial Nerve

2.1 Introduction

The facial nerve is a fundamental structure in the study of head and neck anatomy. Being the seventh cranial nerve, it serves primarily by innervating the muscles of facial expression, and therefore is necessary for emotional expression, speech and mastication. Additionally, it contributes to taste, lacrimation, salivation, and some cutaneous sensation. The nerve initially emerges from the brainstem at the cerebellopontine angle and travels with the vestibulocochlear nerve in a short segment laterally to the internal auditory canal (IAC). From here, it begins its course through the fallopian canal of the temporal bone. This canal is notable as it is the longest bone-laden nerve canal in the body (1), is subject to great variation in its anatomical course (2), and shares intimate borders with many of the essential structures of the ear. Therefore, there is considerable importance in the knowledge and study of this path for both the anatomist and the otologic surgeon.

The intratemporal path of the facial nerve has been described in a number of previous texts (1,3,4), and a graphic representation can be seen in Figure 2.1. Early quantitative analysis used computed tomography (CT) to measure the length of each segment, ranging from 2.5 to 6mm, 8 to 11mm, and 8.9 to 16mm for the labyrinthine, tympanic, and mastoid, respectively (5). Diameter has been evaluated in several studies, with smaller canals being correlated to increasing incidences of Bell's palsy (6). Fisch *et al.* (7) initially focused on the entrance of the facial nerve from the IAC at the meatal foramen, as it is one of the smallest areas, averaging

only 0.68mm in diameter. Nakashima et al. (8) later described a second bottleneck in the midtympanic segment. More recently, Vianna *et al.* (6) measured the diameter at this tympanic midpoint, along with the labyrinthine and mastoid, as 1.9, 1.7 and 2.6mm respectively. Significant differences between control and Bell's palsy subjects were found in only the tympanic and mastoid measurements. Multiple studies have also evaluated the facial recess as it is implicated surgically in cochlear implantation and other mastoid surgeries. A recent study by Calli *et al.* (9) demonstrated a range of 5.6mm to 18.4mm and a mean of 7.8mm between the takeoff of the chorda tympani and the incus buttress. Other objective studies of the intra-temporal facial nerve have focused on imaging features, concluding that coronal and axial sections from high resolution CT are excellent for visualization of the bony canal (10), and that various magnetic resonance (MR) techniques provide the superior soft tissue delineation needed for diagnosis of many pathologies (e.g. schwannoma, malignancy, neuritis) (11).
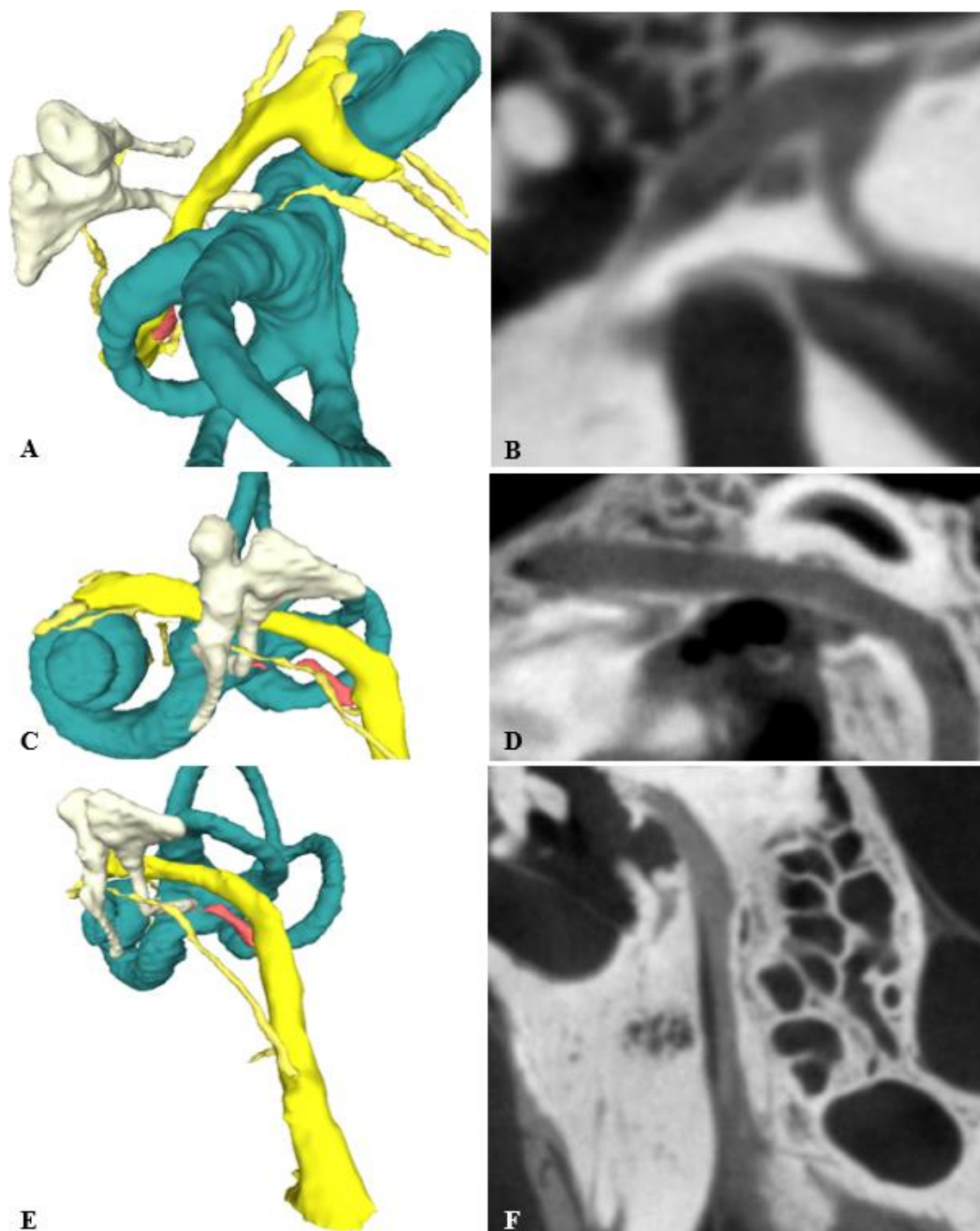
Figure 2.1: 3D and CT representations of the intratemporal portion of the facial nerve. Figures A, C, and E demonstrate spatial orientation of the nerve (yellow) relative to the vestibulocochlear system (blue), ossicles (white), and stapedius muscle (red). Figures A and B demonstrate an axial view of the labyrinthine segment, with the vestibular, cochlear and facial nerves entering from the internal auditory canal. Figures C and D demonstrate a parasagittal view of the tympanic segment, with the nerve deep to the ossicles and near the semicircular canals. Figures E and F demonstrate the mastoid segment, with a clear view of the facial recess between the mastoid segment and the chorda tympani.

The incidence of iatrogenic injury of the facial nerve during mastoid surgery has been estimated to be as high as 1.7% (12). Along the labyrinthine segment, the facial nerve is in very close proximity to the basal turn of the cochlea, ranging from 0.06 to 0.8mm, and both structures can be vulnerable during surgery (13). Just distal to this, the geniculate ganglion has previously been shown to lack a bony covering in 15% of temporal bones (14), therefore requiring caution during middle fossa surgical approaches. Along the tympanic segment, the wall separating the nerve from the middle ear is thin and easily fractured, and in up to 33% of cases dehiscent (15), making it susceptible to pathologic processes of the middle ear, such as acute otitis media and barotrauma. Although the mastoid segment is generally predictable in orientation, and well-preserved surgically, rare anatomical variations in shape exist, such as bifurcation (10).

One promising approach to minimizing the risk of facial nerve injury has been the recent introduction of image-guided minimally invasive (16) and robotic (17) surgical approaches. Another approach has been in enhancing the training environment for training otologic surgeons, by providing a more safe and effective method of developing operative skills and spatial awareness. This is a notable challenge, as previous evidence has demonstrated a reduction in the prevalence of cholesteatoma, and therefore a reduction in total mastoidectomies performed, leaving less opportunity for trainees to develop these skills in residency (18). To address this, there has been an emergence of virtual simulation platforms (19, 20), with a recent study by Locketz *et al.* demonstrating the effectiveness of virtual simulation in improving confidence among junior learners (21).

Both the image-guided surgical techniques and virtual training environments rely on three dimensional volumes obtained from CT or MR scans. The process of segmenting these volumes is quite time-intensive if done manually. Instead, automated segmentation algorithms

are now being applied in otology to evaluate the inner ear (22, 23), facial nerve (24), and combined temporal bone structures (25). These segmentation algorithms have so far used multi-atlas segmentation (25) and statistical shape modelling (23, 24, 26), and deep learning remains a promising method demonstrated in many other organ systems (27, 28). However, developing effective image processing algorithms often requires large datasets, and collection of many high definition temporal bone scans is a noted challenge. Two groups have recently created publicly available data sets, with eight (29) and 51 temporal bone specimens (30), respectively, although there still remains some growth necessary to achieve the statistical power needed for algorithm training and validation.

The present study will aim to expand the objective understanding of the facial nerve anatomy within the temporal bone using high-resolution micro-CT (μCT) and novel statistical shape analysis. Anatomical measurements of the nerve path will be evaluated, and regions of statistical variability in the shape of the facial nerve will be demonstrated. This study will provide statistical models for automated segmentation algorithms, and the models will be made publicly available to all researchers to foster collaborative research.

2.2 Materials and Methods

Ethics approval was obtained from the Department of Anatomy at the Schulich School of Medicine and Dentistry. Three-dimensional μCT images were collected from 40 cadaveric specimens. Specimens were taken from 21 donors, with an even distribution of 20 right and 20 left ears. The temporal bone and its structures were removed from the body and brain for simplicity of the experimental method. The specimens were scanned using the GE Healthcare eXplore Locus μCT scanner (GE Healthcare, Chicago, IL). The scanner was operated with a voltage of 80 kV and current of 0.45 mA. Approximately 900 views were captured at an

incremental angle of 0.4 degrees. Images were reconstructed with an isometric voxel size of 154μm.

The open source software Slicer v4.6.2 (31) was used to analyze the imaging data. All images were aligned using a series of rigid registration steps. To begin, one master image volume was manually rotated into the standard anatomical position to ensure that multiple landmarks, including the semicircular canals, pinna, tragus and ossicles, were correctly oriented. All left-sided temporal bones were mirrored to match the right. Image volumes were aligned to the master volume using a rigid body fiducial registration (as implemented in the Slicer platform), with fiducials centered on the following landmarks: cochlear nerve at the entry point to the cochlea, center of the oval window, and center of the round window.

A series of measurements were made in Slicer to quantify the dimensions of the facial nerve canal, a graphic representation of which can be seen in Figure 2.2. All measurements and segmentations were reached via consensus interpretation by three authors (T.J.H., B.G., and S.K.A.) to minimize interobserver variability. Some measurements were made based on those previously described in the literature, as mentioned above, including diameters at the minimum bottlenecks of the labyrinthine and tympanic segments, and the length of each segment. For the three nerves leaving the mastoid portion - the nerve to the stapedius, chorda tympani and Arnold's nerve (auricular branch of the vagus nerve) – the height from the stylomastoid foramen was measured. Additionally, the angle between the chorda tympani and the facial nerve was recorded, as it is representative of the facial recess. Nerve branches were not traced past the first 2-3cm, as most could not be followed reliably into the surrounding architecture. The three branches from the geniculate ganglion – the greater, lesser and external petrosal nerves – were observed but not reported in the current study, as their position showed little variability and their measurements yielded little clinical significance.
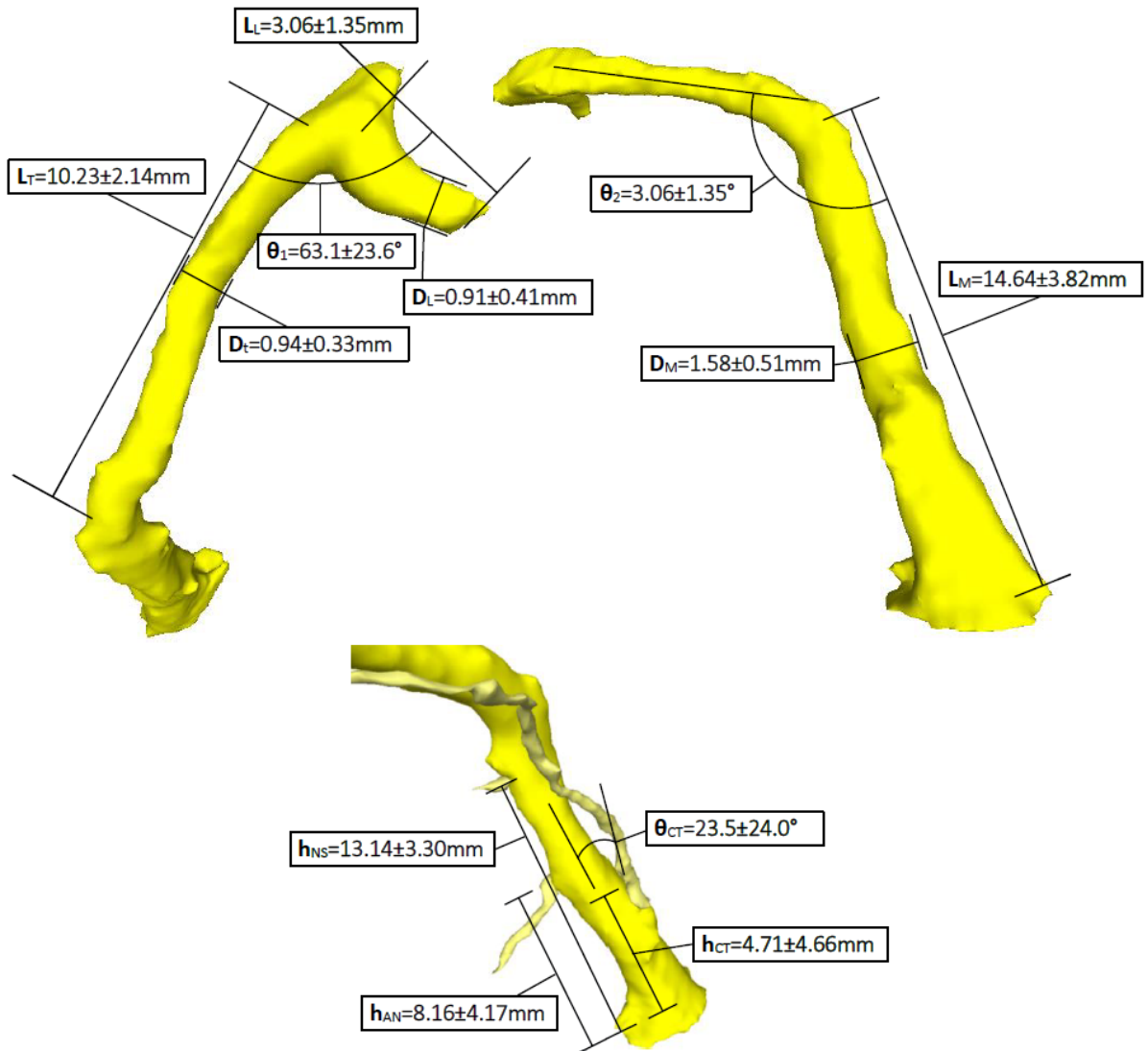
Figure 2.2: Graphical representation of the diameters, lengths and angles of the facial nerve canal. L, D, h and θ represent the length, diameter, height and angle. Subscripts L, T, and M denote the labyrinthine, tympanic and mastoid segments, and subscripts 1 & 2 denote the first and second genu. Nerve to stapedius, Arnold's nerve and Chorda tympani (NS, AN, CT) were defined by their height above the stylomastoid foramen, along with the angle with the angle between the chorda tympani and the facial nerve (representing the facial recess). A 95% confidence interval was reported (two standard deviations).

As part of the measurement analysis, an assessment of facial nerve dehiscence was also performed. Images were first visually inspected along the tympanic segment for apparent absence of bone between the fallopian canal and middle ear cavity. An example can be seen in Figure 2.3, where there is an apparent loss of CT signal in the tympanic segment just lateral to the geniculate ganglion. Images that showed signs of dehiscence were then evaluated further,

where voxel intensities along 2-4 segments travelling from the canal through the region of interest to the cavity were obtained. If at least two of the segments depicted a decrease in signal, with no spike in intensity at the interface, then the region was deemed dehiscent.
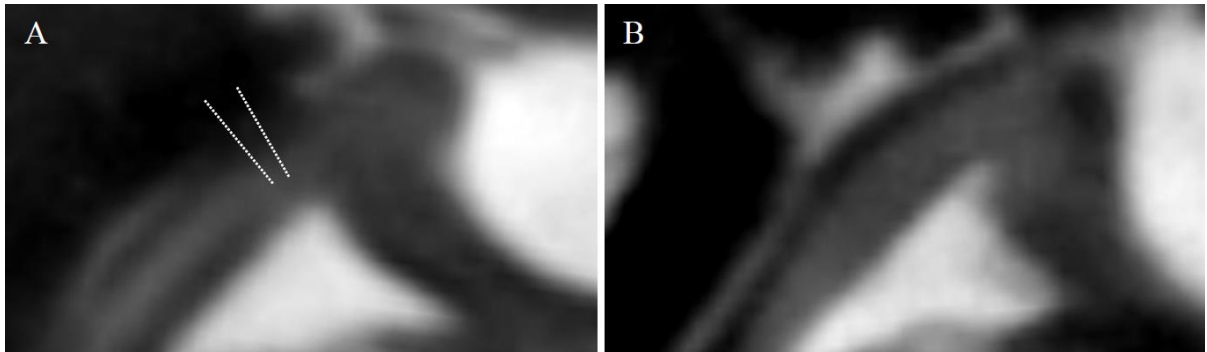
Figure 2.3: Dehiscence between the facial canal and the middle ear cavity just lateral to the geniculate ganglion in one subject (A) on axial view, and a thin but intact bony separation in another (B). Dotted segments drawn in (A) were used to measure for a spike in intensity at the interface, implying presence or absence of bone.

Shape variation of the facial nerve was analyzed using two different methodologies. First, a principle component analysis (PCA) was performed using mathematical calculations well described in the literature (32, 33). To perform a PCA on the facial nerve found in this dataset, 11 landmark fiducials were placed along the canal, as seen in Figure 2.5. Five separate sections of the nerve were analyzed – the labyrinthine, tympanic, and mastoid segments, along with the junctions at the geniculate ganglion and second genu.

The second method created a statistical shape model (SSM) with the open source modelling platform Statismo (34). The nerve canals were initially segmented manually from the μCT data set. Within Statismo, a Gaussian process model was created to determine the mean facial nerve, with the gaussian kernel setting of $\sigma = 50$ and 200 basis functions. This mean nerve was used as the reference shape, and all other facial nerves were fit to it to attain a point-to-point correspondence. With this correspondence, a statistical analysis could be

performed using the platform's mathematic computation, where all of the nerves were projected onto the reference shape to determine the overall shape variability.

2.3 Results

The measurements of the facial nerve canal and its branches can be seen in Figure 2.2. A 95% confidence interval (two standard deviations) was used for all measurements. Image analysis demonstrated that nine of the 40 facial nerves (22.5%) were dehiscent in the tympanic segment. Two came from opposite ears of the same cadaver, and both demonstrated dehiscence in the tympanic segment adjacent to the oval window. The other seven came from individual cadavers, and demonstrated dehiscence just lateral to the geniculate ganglion, as demonstrated in Figure 2.3.

The PCA performed on the 40 sets of landmark fiducials can be seen graphically in Figure 2.4. All four images (2.4A to 2.4D) demonstrate the same mean nerve shape in yellow, and then project six more nerves according to the direction and magnitude provided by the component coefficients from the PCA calculation. Blue through red demonstrates seven shapes produced by $n\sqrt{\lambda_i}$, where n is the index for each shape (-3 in blue, -2 in teal, -1 in green, 0 in yellow, 1 in orange, 2 in dark orange, and 3 in red) and λ_i is the eigenvalue for that component. The value, $\sqrt{\lambda_i}$, was the multiplier chosen as it is a value representing the amount of statistical deviation accounted for by that component, analogous to standard deviations in a Gaussian distribution. Note that these are constant diameter representations and do not capture the true diameters along the nerve.
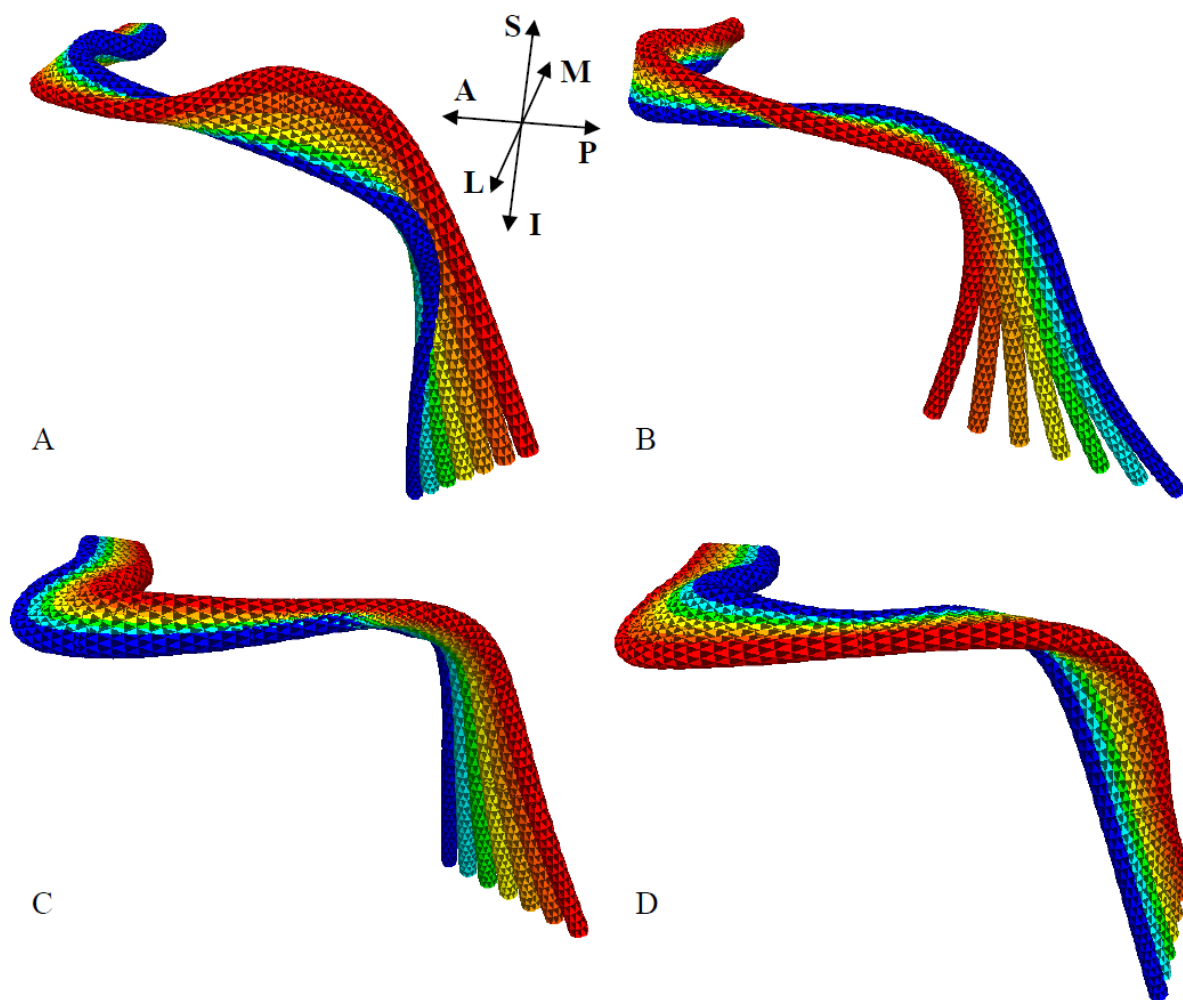
Figure 2.4: Principle components 1-4 (A-Superior/Inferior, B-Anteromedial/Posterolateral, C-Medial/Lateral, D-Anterolateral/Posteromedial). Yellow surface represents the mean shape, with three additional surfaces in each direction demonstrating $\pm 1\sqrt{\lambda_i}$, $\pm 2\sqrt{\lambda_i}$, $\pm 3\sqrt{\lambda_i}$ in variation. Cardinal directions for all figures demonstrated in A (Superior (S), Inferior (I), Anterior (A), Posterior (P), Medial (M), Lateral (L)).

Although PCA does discover numerous modes of variation in complex data, most of the shape variability in this analysis came from the first four components, accounting for 75.3% of the net variation. Here, the first component in the superior/inferior distribution (Figure 2.4A) appeared to capture variability just proximal to the second genu, angling the tympanic segment from inferior to slightly superior in the extrema. The next three components appeared to capture variability primarily in the mastoid segment, where anteromedial/posterolateral

(Figure 2.4B), medial/lateral (Figure 2.4C), and anterolateral/posteromedial (Figure 2.4D) distributions were seen.

Figure 2.5 shows all the component surfaces combined into one view, helping to illustrate the major regions of variability. The labyrinthine segment, geniculate ganglion and proximal tympanic segment remained relatively fixed, with the overall shape preserved and up to 2-3 millimeters of translation. The first major variation was found in the distal tympanic segment, as described above in the first principle component and seen again in Figure 2.5A (violet), where the slope angled from inferior to slightly superior in the extrema. The most variability, however, occurred in the distal mastoid segment and stylomastoid foramen, as seen most clearly in Figure 2.5B. Here, the second and third components (green and blue, respectively) demonstrated a strong tendency for the nerve to vary along their anteromedial/posterolateral and medial/lateral axes. The perpendicular axes demonstrated in components one and four (violet and orange) showed significantly less variation, and it can therefore be inferred that the possible locations of the stylomastoid foramen relative to the temporal bone varies in a cross distribution (seen graphically in Figure 2.5B).
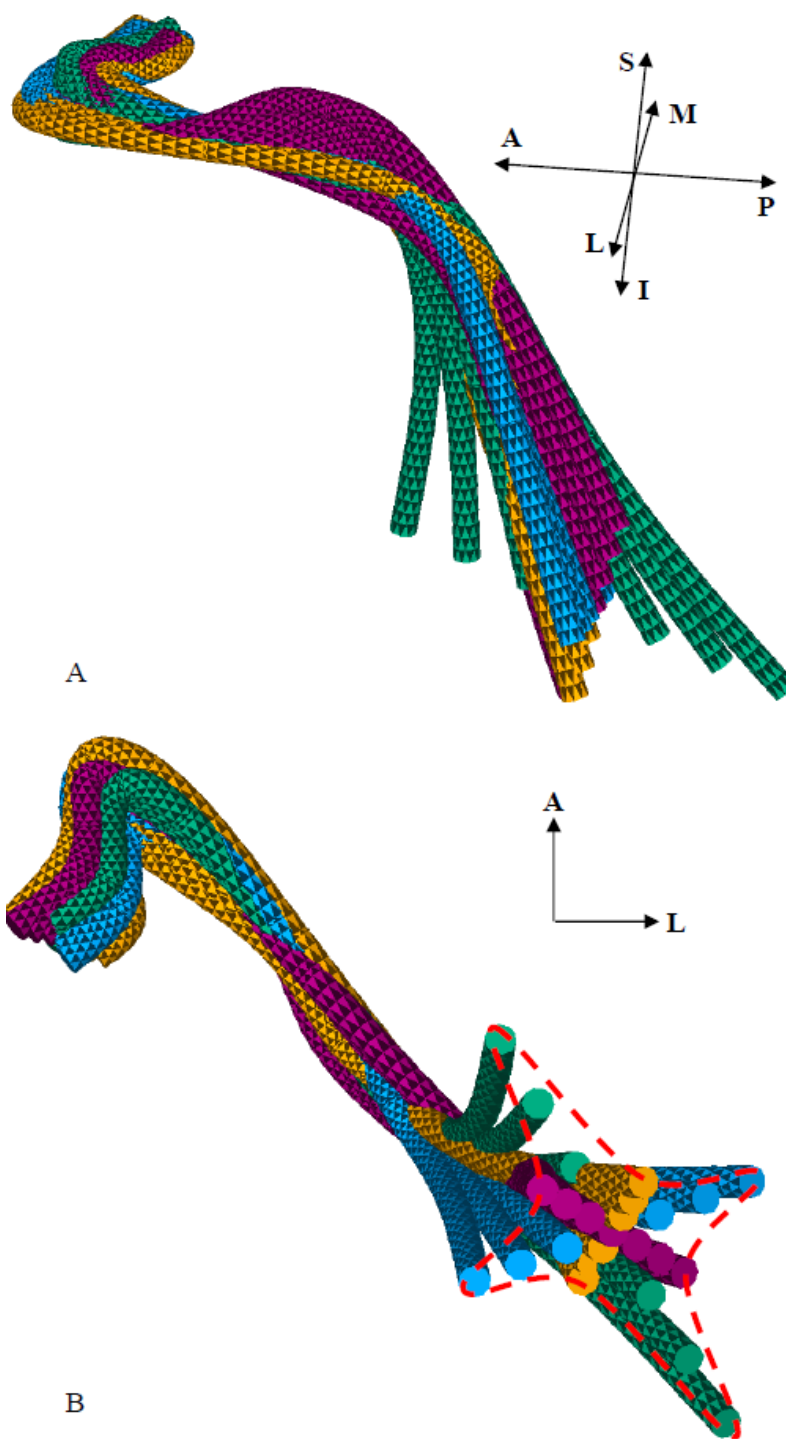
Figure 2.5: All principle component surfaces from above, in diagonal (A) and axial view (from inferior). Component 1 shown in violet (superior/inferior), 2 in green (anteromedial/posterolateral), 3 in blue (medial/lateral), 4 in orange (anterolateral/posteromedial). Dotted red line demonstrates the cross distribution of the stylomastoid foramen.

The SSM created with the Statismo framework demonstrated similar variability patterns as described above in the PCA. Components one through four of the SSM can be seen in Figure 2.6(A)-(D). Component one of the SSM (Figure 2.6A) demonstrated a close match to the first component of the previous PCA (Figure 2.4A), showing a change in superior/inferior angulation of the tympanic segment, along with a moderate amount of anterior/posterior translation of the mastoid segment. Component two of the SSM (Figure 2.6B) demonstrated an isolated variation in the superior/inferior length of the mastoid segment, a change seen partially in each of the first, third and fourth components (Figure 2.4A, 2.4C, 2.4D, respectively) of the PCA. Component three of the SSM (Figure 2.6C) demonstrated a similar medial/lateral displacement of the mastoid as the second component of the PCA (Figure 2.4B), along with some translation of the labyrinthine segment, as seen through all previous components. Component four of the SSM (Figure 2.6D) also demonstrated the anterior/posterior displacement of the mastoid segment as seen previously in the PCA. In summary, the anterior/posterior translation of the labyrinthine segment, the superior/inferior angulation of the tympanic, and the anterior/posterior, medial/lateral and superior/inferior displacements of the mastoid segment were re-demonstrated in the Statismo SSM.
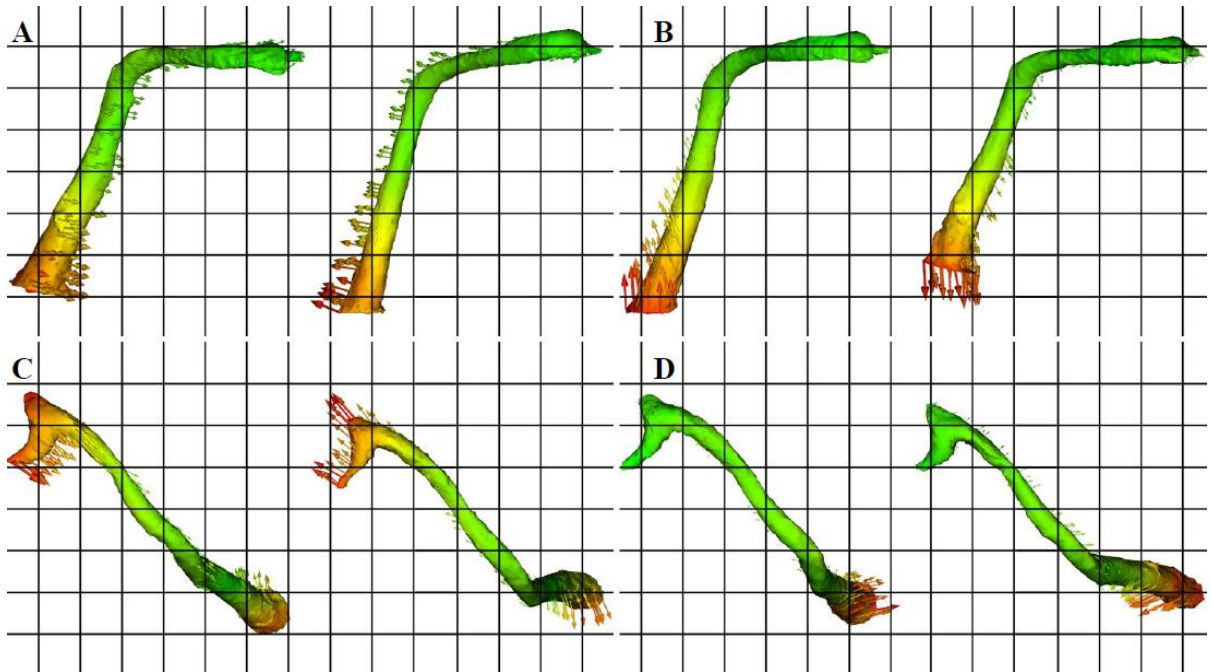
Figure 2.6: Statistical shape model produced by the Statismo framework. This representation demonstrates the variability of components 1-4 (A = superior/inferior, B = anteromedial/posterolateral, C = medial/lateral, D = anterolateral/posteromedial) by showing facial nerves with +/- 2 standard deviations along each component and the distance (arrows) from points to their corresponding points on the mean facial nerve. Larger arrows and color approaching the red portion of the spectrum demonstrate regions of increased variability.

One additional observation of the facial nerve was made from processing the imaging data, where the fallopian canal appeared to split into two at the second genu, and the facial nerve appeared to travel in two separate canals inferiorly through the mastoid bone. This was found in two of the 40 facial nerves (5%), from separate cadavers.

2.4 Discussion

Many of the quantities measured in this analysis were comparable to those found previously in the literature. Regarding the minimum diameters of the labyrinthine and tympanic segments (the bottlenecks), there was close agreement between those found in the current study and in a previous study by Nakashima et al. (8) (0.91 and 0.94mm compared to 0.89 and 0.92mm, respectively). Regarding lengths, reports from an early CT study (5) demonstrated ranges 2-

6mm, 8-11mm, and 9-16mm which were comparable to the 2.11-5.36mm, 7.73-12.91mm, and 11.09-19.63mm measurements found in the current study for labyrinthine, tympanic and mastoid segments, respectively. The ranges reported in the current study, however, were influenced by a small portion of outliers, and the overall mean and standard deviation tends to fit well within the previously reported ranges. Regarding dehiscence, the 22.5% reported in the current study is lower than the 33% quoted in a recent study (15), although a past compilation of surgical study data demonstrated that dehiscence is present in 25% of cases on gross dissection, and can range from 6-33% on histology (35).

The use of μCT technology permitted several additional observations from what was found in the literature. The greater and lesser petrosal nerves were able to be reliably traced anteriorly to the meeting point with the deep petrosal and tympanic plexus branches, respectively. The external petrosal was a very thin branch, had the greatest degree of variation projecting in nearly any direction, and was not identifiable in three of the specimens. This was not surprising, as previous anatomical literature has cited that the external petrosal is an inconstant branch, in some instances joining the greater and lesser petrosal nerves directly from the carotid canal (36). Of the three branches exiting from the mastoid segment, the chorda tympani had the greatest variation in its exit height. Although this height has not been directly reported previously, the standard deviation of 2.33mm is comparable to the deviation of 2.68mm reported for the facial recess width in the literature (9). Additionally, the same study demonstrated a recess angle of 23.58±6.84°, which matches closely the angle of 23.5±24.0° reported in the present study (note that the previous study reported a confidence interval of one standard deviation, while the present study reports two standard deviations). The nerve to the stapedius was the most predictable branch, and subjectively was almost always found branching antero-medially just below the second genu.

The statistical shape analysis added a new perspective to the study of the facial nerve, demonstrating the statistical distribution of nerve paths through the temporal bone. As mentioned, there were two areas with significant deviation: the first being at the distal tympanic segment, with variation on one statistical plane showing a superior to inferior change in angulation; the second being at the distal mastoid segment, with variation primarily along two statistical planes. Otherwise, nerve shape was relatively constant and only varied in position by approximately 2 to 3mm. From a surgical point of view, this distal tympanic site is of particular significance, as a recent case study has demonstrated that the intratympanic facial nerve is commonly injured around the tympanic segment and second genu (37). Iatrogenic facial nerve injury is a devastating complication of otologic surgery, and the reported incidence rate of 1.7% (12) could potentially be minimized through anatomic knowledge, surgical experience, nerve monitoring, and image guidance.

One source of error in the current study was due to the process of manually selecting landmarks. Extensive image segmentation studies in the literature have demonstrated that inter-observer variability can be an issue in image identification. This was mitigated in the present study through consensus interpretation by multiple observers, and through the use of high-resolution images, which clearly delineated borders and areas of interest. Another potential source of error arose from the fiducials chosen for the initial rigid registration. The images were all aligned such that the three fiducials chosen – the cochlear nerve, the oval window, and the round window – were always in approximately the same position. Small variations in the true orientation may have influenced the variability in the models demonstrated; however, substantial portions of the nerve, including the proximal mastoid and tympanic segments, had little shape variability, therefore this likely had a minor effect.

Furthermore, the agreement of these results with those found in previous studies supports that the methods used here produced results within reasonable limits.

Several studies have previously attempted to characterize features of the facial nerve path, including diameters, lengths and presence of dehiscence. This study aimed to expand on this knowledge through a cadaveric imaging analysis using μCT images, providing size measurements and adding new information about angulation, branching and dehiscence. Additionally, this study was the first to use a μCT dataset to create a PCA and SSM of the facial nerve, demonstrating in three dimensions how the facial nerve varies within the temporal bone. Along with providing valuable knowledge of the anatomy, the data processing and statistical models created here will be further used in the development of automatic segmentation algorithms. The images, data and models will be made available to the greater research community via the Auditory Biophysics Laboratory website (abl.uwo.ca).

References

1. May M, Schaitkin BM. The facial nerve New York: Thieme; 2000.

2. Barry J. Anson P, James A. Donaldson M, Raymond L. Warpeha DMP, al e. Surgical Anatomy of the Facial Nerve. Archives of Otolaryngology. 1973; 97(2): p. 201-213.

3. Slattery WH,I, Azizzadeh B. The facial nerve New York: Thieme; 2014.

4. Agur AMR, Grant JCB, Dalley AF,I. Grant's atlas of anatomy Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.

5. Wilbrand H. Multidirectional tomography of the facial canal Stockholm: Acta radiologica: diagnosis; 1975.

6. Vianna M, Adams M, Schachern P, Lazarini PR, Paparella MM, Cureoglu S. Differences in the diameter of facial nerve and facial canal in Bell's Palsy-A 3-dimensional temporal bone study. Otology and Neurotology. 2014; 35(3): p. 514-518.

7. Fisch U, Esslen E. Total Intratemporal Exposure of the Facial Nerve: Pathologic Findings in Bell's Palsy. Archives of Otolaryngology - Head and Neck Surgery. 1972; 95(4): p. 335-341.

8. Nakashima S, Nakashima S, Sando I, Sando I, Takahashi H, Takahashi H, et al. Computer-aided 3-D reconstruction and measurement of the facial canal and facial nerve. I. Cross-sectional area and diameter: Preliminary report. Laryngoscope. 1993; 103(10): p. 1150-1156.

9. Calli C, Pinar E, Oncel S, Tatar B, Tuncbilek MA. Measurements of the facial recess anatomy: implications for sparing the facial nerve and chorda tympani during posterior tympanotomy. Ear, Nose and Throat Journal. 2010; 89(10): p. 490-494.

10. Tüccar E, Tekdemir I, Aslan A, Elhan A, Deda H. Radiological anatomy of the intratemporal course of facial nerve. Clinical Anatomy. 2000; 13(2): p. 83-87.

11. Jäger L, Reiser M. CT and MR imaging of the normal and pathologic conditions of the facial nerve. European Journal of Radiology. 2001; 40(2): p. 133-146.

12. Nilssen ELK, Wormald PJ. Facial nerve palsy in mastoid surgery. The Journal of Laryngology and Otology. 1997; 111(2): p. 113-116.

13. Redleaf MI, Blough RR. Distance from the labyrinthine portion of the facial nerve to the basal turn of the cochlea Temporal bone histopathologic study. Annals of Otology, Rhinology & Laryngology. 1996; 105(4): p. 323-326.

14. Rhoton Jr AL, Pulec JL, Hall GM, Boyd Jr AS. Absence of bone over the geniculate ganglion. Journal of neurosurgery. 1968; 28(1): p. 48-53.

15. Selesnick SH, Lynn-Macrae AG. The incidence of facial nerve dehiscence at surgery for cholesteatoma. Otology and Neurotology. 2001; 22(2): p. 129-132.

16. Labadie RF, Balachandran R, Noble JH, Blachon GS, Mitchell JE, Reda FA, et al. Minimally invasive image-guided cochlear implantation surgery: First report of clinical implementation. The Laryngoscope. 2014; 124(8): p. 1915-1922.

17. Caversaccio M, Gavaghan K, Wimmer W, Williamson T, Ansò J, Mantokoudis G, et al. Robotic cochlear implantation: surgical procedure and first clinical experience. Acta Oto-Laryngologica. 2017; 137(4): p. 447-454.

18. Das-Purkayastha PK, Coulson CJ, Pothier DD, Lai P, Rutka JA. Time trend analysis of mastoidectomy procedures performed in Ontario. Acto Oto-Laryngologica. 2012; 132(1): p. 16-20.

19. Wiet GJ, Stredney D, Kerwin T, Hittle B, Fernandez SA, Abdel-Rasoul M, et al. Virtual temporal bone dissection system: OSU virtual temporal bone system: development and testing. Laryngoscope. 2012; 122(1): p. S1-S12.

20. Chan S, Li P, Locketz G, Salisbury K, Blevins NH. High-fidelity haptic and visual rendering for patient-specific simulation of temporal bone surgery. Computer Assisted Surgery. 2016; 21(1): p. 85-101.

21. Locketz GD, Lui JT, Chan S, Salisbury K, Dort JC, Youngblood P, et al. Anatomy-specific virtual reality simulation in temporal bone dissection: perceived utility and impact on surgeon confidence. Otolaryngology - Head and Neck Surgery. 2017; 156(6): p. 1142-1149.

22. Pietsch M, Aguirre Dávila L, Erfurt P, Avci E, Lenarz T, Kral A. Spiral form of the human cochlea results from spatial constraints. Scientific Reports. 2017; 7(1).

23. Kjer HM, Fagertun J, Wimmer W, Gerber N, Vera S, Barazzetti L, et al. Patient-specific estimation of detailed cochlear shape from clinical CT images. International Journal of Computer Assisted Radiology and Surgery. 2018; 13(3): p. 389-396.

24. Reda FA, Noble JH, Rivas A, McRackan TR, Labadie RF, Dawant BM. Automatic segmentation of the facial nerve and chorda tympani in pediatric CT scans. Medical Physics. 2011; 38(10): p. 5590-5600.

25. Powell KA, Liang T, Hittle B, Stredney D, Kerwin T, Wiet GJ. Atlas-Based Segmentation of Temporal Bone Anatomy. International Journal of Computer Assisted Radiology and Surgery. 2017; 12(11): p. 1937-1944.

26. Schuman TA, Noble JH, Wright CG, Wanna GB, Dawant B, Labadie RF. Anatomic verification of a novel method for precise intrascalar localization of cochlear implant electrodes in adult temporal bones using clinically available computed tomography. The Laryngoscope. 2010; 120(11): p. 2277-2283.

27. Ueda D, Shimazaki A, Miki Y. Technical and clinical overview of deep learning in radiology. Japanese Journal of Radiology. 2019; 37(1): p. 15-33.

28. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, et al. Deep learning in medical imaging and radiation therapy. Medical Physics. 2019; 46(1): p. e1-e36.

29. Sieber D, Erfurt P, John S, Santos GRD, Schurzig D, Sørensen MS, et al. The OpenEar library of 3D models of the human temporal bone based on computed tomography and micro-slicing. Scientific data. 2019; 6.

30. Gerber N, Reyes M, Barazzetti L, Kjer HM, Vera S, Stauber M, et al. A multiscale imaging and modelling dataset of the human inner ear. Scientific Data. 2017; 4.

31. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magnetic Resonance Imaging. 2012; 30(9): p. 1323-1341.

32. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and Intelligent Laboratory Systems. 1987; 2(1): p. 37-52.

33. Abdi H, Williams LJ. Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics. 2010; 2(4): p. 433-459.

34. Lüthi M, Blanc R, Albrecht T, Gass T, Goksel O, Büchler P, et al. Statismo-A framework for PCA based statistical models. Insight Journal. 2012.

35. Baxter A. Dehiscence of the fallopian canal an anatomical study. The Journal of Laryngology and Otology. 1971; 85(6): p. 587-594.

36. Tubbs RS, Menendez J, Loukas M, Shoja MM, Shokouhi G, Salter EG, et al. The petrosal nerves: anatomy, pathology, and surgical considerations. Clinical Anatomy. 2009; 22(5): p. 537-544.

37. Linder T, Mulazimoglu S, El Hadi T, Darrouzet V, Ayache D, Somers T, et al. Iatrogenic facial nerve injuries during chronic otitis media surgery: a multicentre retrospective study. Clinical Otolaryngology. 2017; 42(3): p. 521-527.

Chapter 3

3 Multi-Atlas Segmentation of the Facial Nerve from Clinical CT

3.1 Introduction

Mastoidectomy is a surgical procedure in which a portion of the temporal bone of the skull is removed using a high-speed otologic drill. This procedure is necessary to treat complications from ear infections; benign tumors such as cholesteatoma and acoustic neuromas; malignant tumors; and to insert cochlear implants to treat hearing loss. Although many delicate structures can be injured during this procedure, iatrogenic facial nerve (FN) injury is of most concern to otologic surgeons (1). The FN is responsible for innervating muscles for facial expression and speech, as well as contributing to taste, and the loss of any of these functions can present issues in the social, psychological, and economical aspects of a person's life (2,3).

FN anatomy is conceptually divided into six segments, several of which are at risk of injury during mastoidectomy (Figure 3.1). The FN leaves the brainstem and enters the internal auditory canal. It then forms the labyrinthine segment (LS) and makes a hairpin turn at the geniculate ganglion (GG). The nerve continues through the middle ear as the tympanic segment (TS), makes a turn inferiorly at the second genu (SG), and then travels inferiorly as the mastoid segment (MS) until it exits the skull at the stylomastoid foramen. The anatomy of the FN can be quite variable and difficult to visualize on imaging (4,5). An analysis of 14 cases of iatrogenic FN injury due to mastoidectomy found six cases (43%) of injury around the tympanic segment, four cases (29%) of injury around the second genu and tympanic

segment, and four cases (29%) of injury in the mastoid segment (6). This indicates the anatomic knowledge of the entire course of the FN is important for surgical trainees.
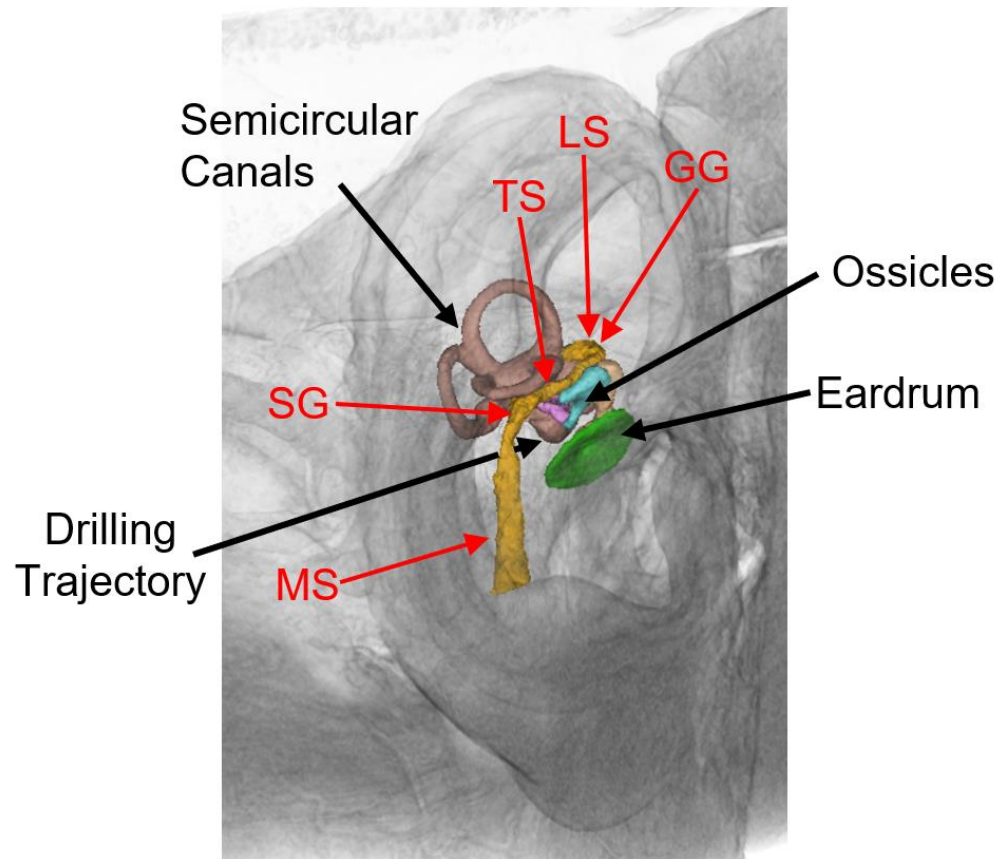
Figure 3.1: Temporal bone anatomy. Red labels show FN anatomy for the mastoid segment (MS), second genu (SG), tympanic segment (TS), geniculate ganglion (GG), and the labyrinthine segment (LS).

Given the risk of injury to the FN and other structures during surgery (7), a number of groups have developed virtual-reality (VR) simulators for practicing mastoidectomy (8,9,10,11). Compared to practicing in a cadaveric temporal bone lab or on patients in the operating room, surgical simulation is always available, minimizes risk, provides anatomic variation, and can offer automated feedback without expert intervention (12). While traditional simulators used a database of anatomies, modern VR simulators can import patient-specific imaging to allow actual surgical rehearsal on a specific patient prior to the operation (13). However, the primary drawback is that these simulators require manual segmentation of the

temporal bone structures by an expert, which is time consuming and reduces the feasibility of using these simulators in actual surgical training programs (9,14).

Several groups have worked on FN segmentation within the temporal bone. Noble and colleagues (15,16,17,18,19) used a navigated optimal medial axis and deformable-model algorithm (NOMAD) framework to localize the tubular structure of the FN. This used a combination of statistical intensity and shape information stored in a model, and atlas-based registration to align the model to their target computed tomography (CT) scan. The CT scans of 15 temporal bones were used to create the model, and a set of 10 parameters were manually and heuristically set to minimize the error between their model and gold standard segmentations. Only the parts of the FN involved in cochlear implant surgery, primarily around the second genu, were included in the automatic segmentation. The model was then tested on 10 clinical CT scans and the maximum error was 0.838 mm (15).

Powell and colleagues (20) used atlas-based segmentation to create gold-standard regions-of-interest (ROI) from 6 cadaveric temporal bone clinical CT scans. Intensity-based segmentation was then used to segment the tympanic and mastoid portions of the FN. The algorithm was validated on 20 temporal bone scans, and the results averaged between the right and left ears were Dice coefficient of 0.71, Hausdorff distance of 0.85, and volume similarity of 0.90.

The objective of this study was to develop a multi-atlas based FN segmentation algorithm which: 1) was based on high-resolution micro-CT atlases, 2) segmented the entire FN including the labyrinthine segment, 3) required minimal user input, and 4) could easily be deployed into an existing open-source toolkit for use in VR simulators.

3.2 Materials & Methods

3.2.1 Image Acquisition of Micro-CT Data

Three-dimensional (3D) micro-CT images were collected from 37 de-identified cadaveric temporal bones. Specimens were taken from 21 donors, with a distribution of 18 right and 19 left ears. Three temporal bones used in the anatomical study were not included because one bone had a rare facial nerve bifurcation and the clinical CT scans of the other two temporal bones could not be obtained. Ethics approval was obtained from the Department of Anatomy and Cell Biology at Western University (Approval Number #19062014). The specimens were scanned using a GE Healthcare eXplore Locus scanner (GE Healthcare, Chicago, IL). The scanner was operated with a voltage of 80 kV and current of 0.45 mA. Approximately 900 views were captured at an incremental angle of 0.4 degrees. Images were reconstructed with an isometric voxel size of 154 μm.

3.2.2 Image Acquisition of Clinical CT Data

All 37 specimens were also scanned at a clinical resolution of 600 μm using a Discovery CT750 HD Clinical Scanner (GE Healthcare, Chicago, IL), equipped with GE's Gemstone CT detector. The Scanner was set to a slice thickness of 625 μm and an x-ray voltage of 120 kV. The acquisition time for each of the 37 specimens was approximately 20 s. The reconstructed voxel size was 235 μm × 235 μm × 625 μm.

### 3.2.3 Formation of Atlas Set

The FNs on all 37 micro-CT were manually segmented using the 3D Slicer software (21). A consensus interpretation of the segmentations was achieved by three authors, including one neurotologist (B.G., T.H., and S.A). For each nerve, a labelmap, a centerline, and a 3D model were created. Principal component analysis and statistical shape modelling was performed on

these segmentations (22), and 9 samples that incorporated the majority of the anatomical variation were chosen for atlas creation. The atlas was created from the high-resolution micro-CT scans of these samples. Using additional atlases didn't improve segmentation metrics, however using less atlases adversely affected the segmentation accuracy.

The low-resolution clinical CT scans of the remaining 28 samples were then used for validation of the algorithm, and these specimens were not used in atlas formation. The manual FN segmentations of the corresponding rigidly aligned micro-CT scans were used as the gold-standard with which to compare the automated segmentations of the clinical CT scans.

3.2.4 Segmentation Algorithm

Figure 3 illustrates the overall flow of our atlas-based segmentation algorithm. To initialize the algorithm, the user selects 4 fiducials on the clinical CT image to be segmented (Step A in Figure 3.2); the image to be segmented will henceforth be referred to as the target image. The selection of these 4 points is the only part of the algorithm that requires user intervention with the remaining steps being fully automatic. The algorithm *approximates* the centerline of the FN from these 4 points (Step B). Each atlas centerline is then rigidly registered to the centerline of the input image (Step C). The atlas that results in the best registration (i.e., closest match) to the new centerline is selected for further processing (Step D). The rigid-body transformation found in (Step C) is then applied to the atlas image volume selected in Step D to approximately align the volume to the target clinical CT image (Step E). Next, the rigidly aligned atlas volume is locally deformed so its FN boundaries better match those of the input image (Step F). Finally, in Step G the optimal rigid registration parameters found in Step C and the local deformation field computed in Step F are applied to the binary labelmap associated with the atlas selected in Step D, and this produces a labeling of FN voxels in the target clinical CT image.
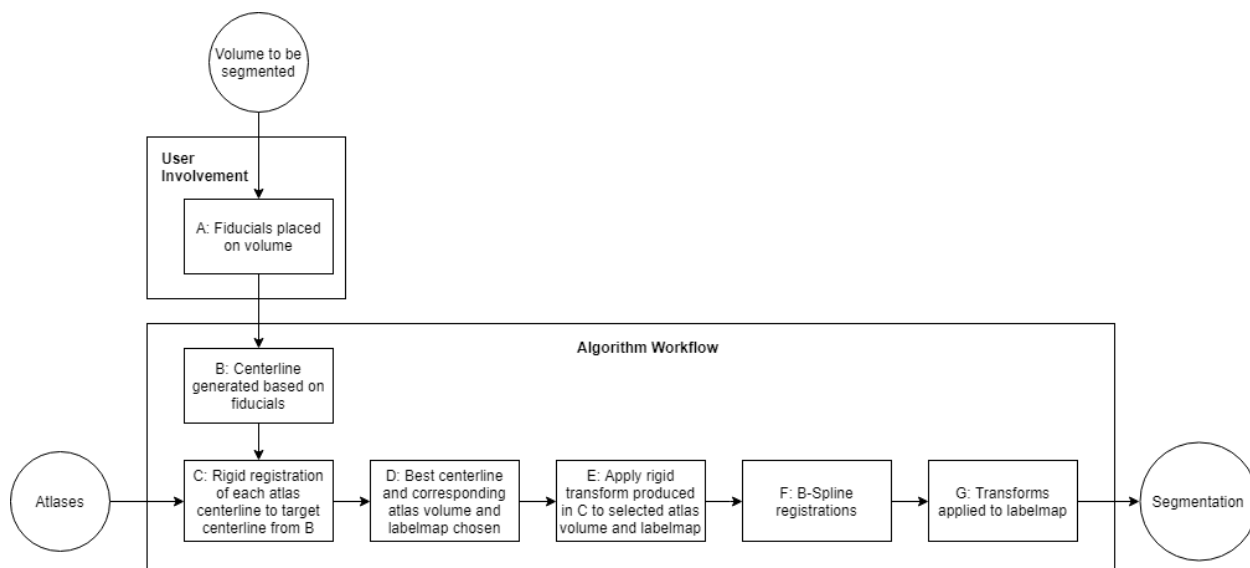
Figure 3.2: Process flowchart. The top-left block 'user involvement' represents the clinician's work. The bottom block 'Algorithm Workflow' represents the algorithm.

A, B: Fiducial Placement and Centerline Generation

The user is asked to place 4 fiducials approximating the beginning of the labyrinthine segment (meatal foramen), first genu (geniculate ganglion), second genu, and stylomastoid foramen. These locations define our region of interest (starting and ending points), as well as the two locations where the FN has significant curvature. These points are easy to identify and the fiducials typically take less than a minute to place. A sample of the four fiducials' placement is given in Figure 3.3.

The fiducials are then connected by straight-line segments to approximate the centerline of the FN. This centerline is automatically linearly sampled to produce 250 uniformly-spaced points along the FN; this is the same number of points used with the atlases.
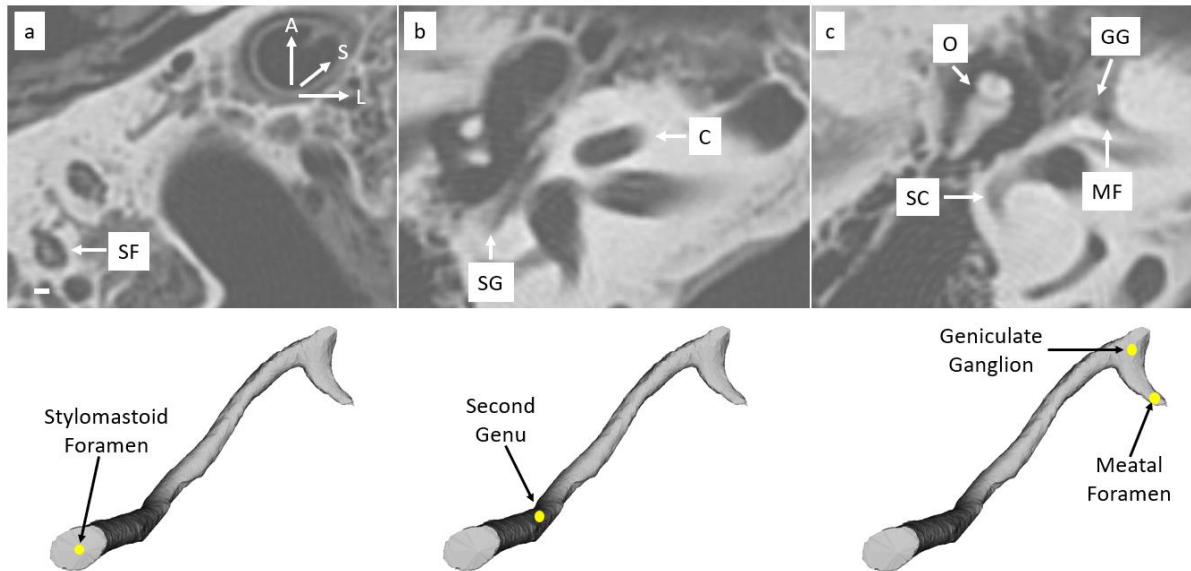
Figure 3.3: Axial view of fiducial locations on a right ear. Fiducials to be placed are shown on the (a) Stylomastoid foramen (SF), (b) Second Genu (SG), (c) Geniculate Ganglion (GG) and Meatal Foramen (MF). Cochlea (C), ossicles (O), and semicircular canals (SC) are shown for reference. Orientation axes are included on image (a) with superior direction going into the image and inferior towards the reader. Bar for scale shown on bottom left of (a) represents 1 mm.

C: Rigid Registration of Centerlines

The iterative closest point (ICP) algorithm is an option to align point clouds, and this was used to automatically and iteratively rotate and translate each of the nine atlas centerlines to align them to the target clinical CT image. Iterations were repeated until the atlas centerline moved no more than 0.001 mm in distance between iterations, or until 500 iterations were reached. Centerlines from each atlas were rigidly registered rather than the entire image volume as this is much less computationally intensive.

D, E: Centerline Selection and Atlas Volume Registration

On completion, ICP computes the root mean squared error (RMSE) of the Euclidean distance between each aligned atlas centerline and the target centerline. The atlas that produced the lowest RMSE was selected for further processing. The rigid transformation estimated for the

selected atlas centerline was then used to transform the associated atlas image volume. This step brings the atlas image volume into overall global alignment with the target image, however local differences in FN shape will still be present.

F: B-Spline Registration

To locally refine the fit of the atlas image volume to the target image, a deformable B-Spline registration was used. B-Spline registrations move points in a grid to deform one image to fit another. Grid sizes of 12×12×12 and 20×20×20 are used for an initial coarse registration and a second fine registration, respectively. Other grid sizes were tested but they either overly restricted deformation or allowed for unrealistic deformation. In order to avoid considering extraneous image information which could both slow down the algorithm and could cause unnatural deformation, a mask was created by automatically dilating the associated atlas labelmap by 4 voxels as this was empirically found to contain all information relevant to the FN in the target image volume.

A second method to avoid unnatural deformation was imposing limits on the maximum displacement of control points to ensure that the displacement was not unrealistically large. These limits were imposed because in many cases, the tympanic segment of the FN can be dehiscent (23) and in dehiscent regions, the boundary of the FN is discontinuous which can result in a B-Spline registration where the atlas image volume locally deforms beyond normal FN anatomy. In the coarse registration, the displacement of control points was limited to a maximum of 0.8 mm, and in the fine registration, the displacement of control points was limited to a maximum of 0.4 mm. These values were found empirically to balance the amount of deformation needed for an accurate segmentation. In both coarse and fine registrations, mutual information was chosen as the cost metric (24) and the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm (25) was used as the optimizer. The

optimization was iterative and terminated when the displacement was less than 0.001 mm, or a predetermined maximum number of iterations was reached. In practice, the algorithm always converged well before 1500 iterations had taken place.

3.2.5 Validation

Evaluation of the segmentation algorithm was conducted by applying the algorithm to the 28 *clinical* CT images that were not selected to form atlases; henceforth, these clinical images will be called the *test images*. The algorithm was applied to each test image, and the resulting automatic segmentation was compared to the gold-standard manual segmentations. In comparison to previous studies, high-resolution micro-CT segmentations were used to create a highly accurate "ground truth" for each of the test samples.

Two metrics were used to compare the ground truth segmentation to the automated segmentation: the average Hausdorff Distance (AHD) and the Dice Similarity Coefficient (DSC) (26). The AHD is an indicator of the closeness between the two FN segmentations, whereas the DSC is an indicator of how well the two segmentations overlap. The metrics were evaluated over the entire segmented FN, as well as under a cropped region of interest (ROI) consisting of the tympanic and mastoid portions of the FN. As this was the first study to include the labyrinthine segment, this ROI was created to allow comparisons against previous studies which did not include this portion.

Mean and standard deviation of the AHD and DSC were calculated. The Shapiro-Wilk test was used to test variables for normality and differences between AHD and DSC for the entire FN compared with the ROI only were analyzed with the Mann-Whitney U test. Both tests were conducted using the Statistical Package for the Social Sciences (SPSS) version 25 (SPSS, Inc, an IBM Company, Chicago, Illinois). A p value of $<.05$ was considered significant.

3.3 Results

Figure 3.4 demonstrates the qualitative difference between segmentation methods. 3D surface models were used to visualize the differences between the ground truth and automated segmentation (Figure 3.4A and 3.4B). Cross-sections of the segmentation results are shown below (Figure 3.4C-F), which demonstrate the differences between the segmentation outlines.
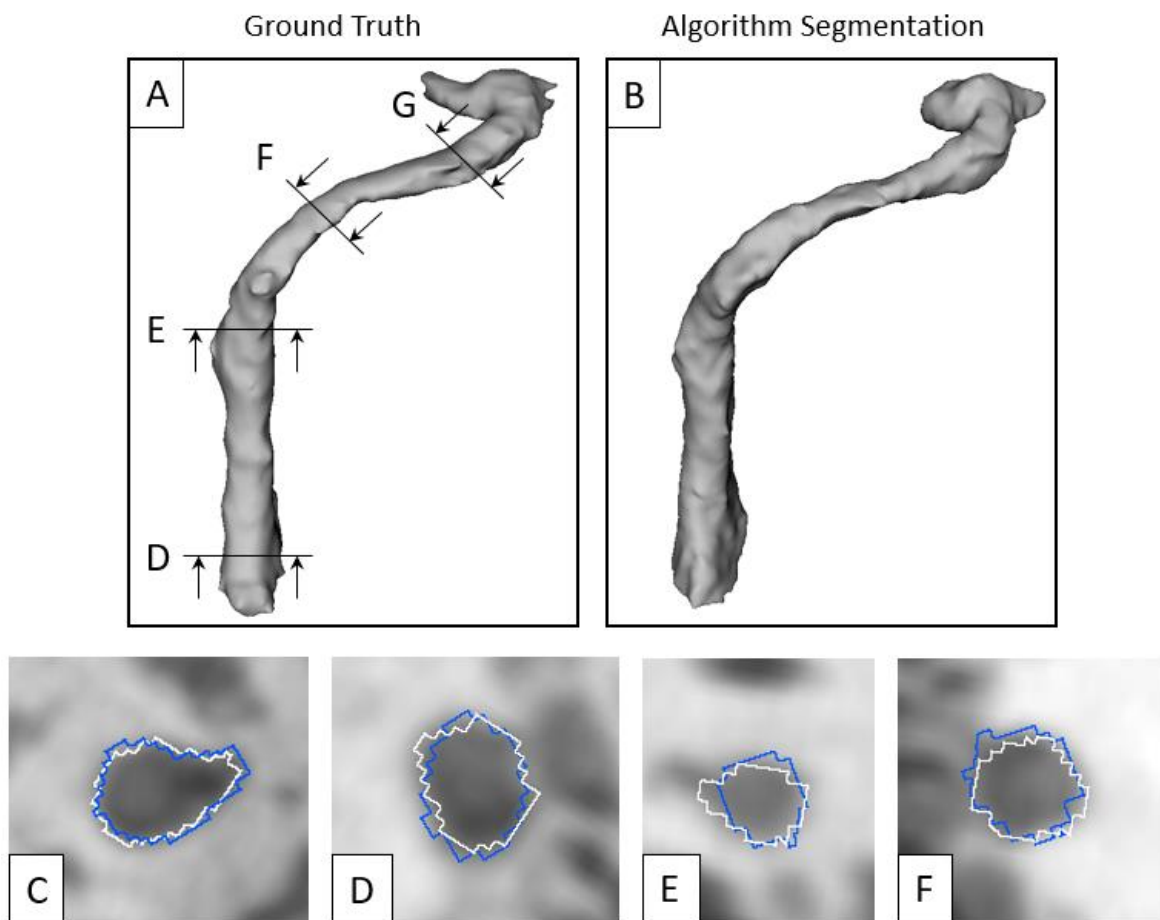
Figure 3.4: Segmentation results for manual and algorithm segmentations. 3D renderings of the segmentations are shown on the top (A, B). Cross-sections of the FNs are selected along the course of the FN as shown in panel (A) and the corresponding cross-sections are illustrated below (C-F) on top of the micro-CT image corresponding to the clinical CT image that was segmented by each algorthim. White represents the manually segmented ground truth while blue represents the algorithm segmentation.

Figure 3.5 shows 3D representations of every ground-truth segmentation which have been shaded to display the Euclidean distance from the closest corresponding points in the ground-truth segmentation and the associated automatic segmentation. In general, the largest distances were seen at the meatal foramen and at the stylomastoid foramen.
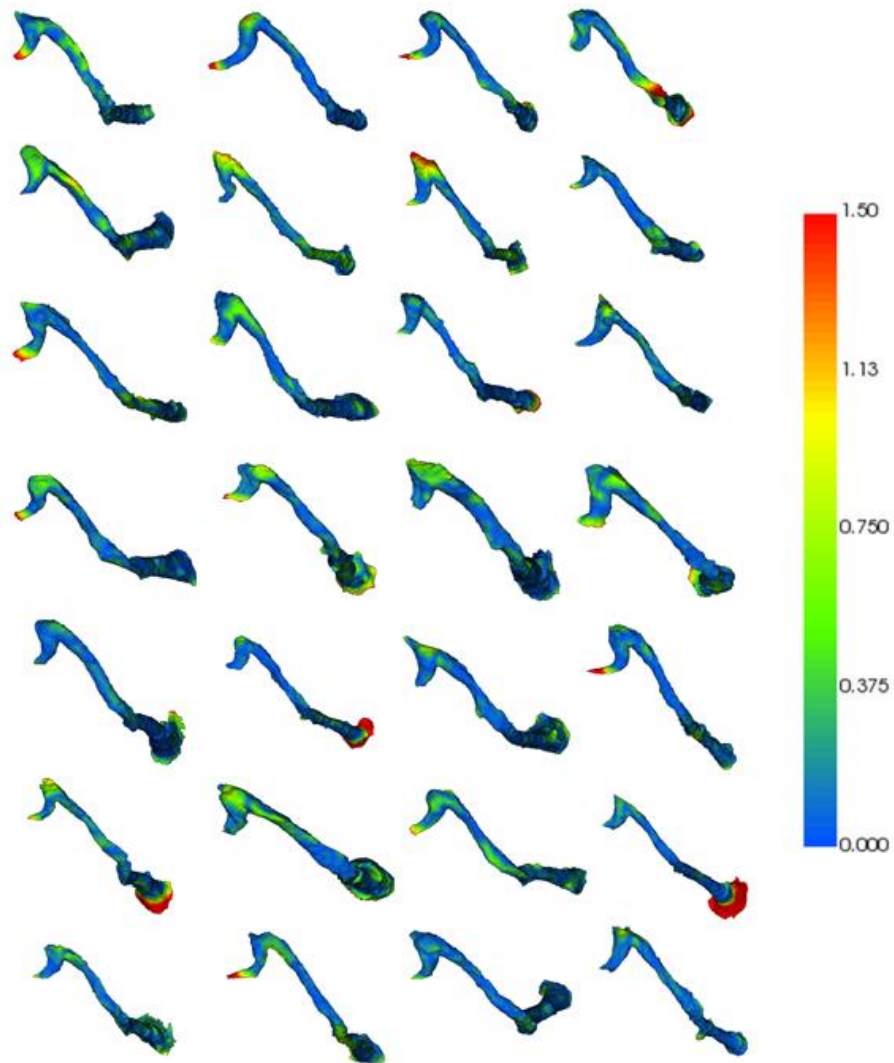
Figure 3.5: Ground-truth segmentations color coded with distance in mm from corresponding points on algorithmic segmentation. Regions are shaded from blue meaning distances close to zero to red meaning distances greater than or equal to 1.5 mm.

Quantitative results are shown in Table 2.1 for the entire FN and over the ROI. When considering the cropped region, a DSC of 0.76 was found which was significantly higher than the DSC of 0.70 found when considering the entire FN, $p < .001$. The AHD also has a significant drop from 0.26 to 0.17 when considering the ROI only, $p < .005$, which is to be expected when the largest distances between corresponding points existed outside of the ROI.

Table 3.1: Metrics comparing the algorithmic segmentations to ground-truth manual segmentations. DSC and AHD are considered for the entire segmentation as well as under a ROI.

Segmentation Type	DSC ± SD	AHD (mm) ± SD
Entire FN	0.70 ± 0.08	0.26 ± 0.10
ROI (tympanic and mastoid segments)	0.76 ± 0.07	0.17 ± 0.05
p-value	.001	.005

3.4 Discussion

The results of this segmentation algorithm are favourable when compared to similar algorithms in the literature. Powell et al (20) examined the tympanic and mastoid portions of the FN, which is similar to the ROI defined in this study. The DSC in this study was slightly higher (0.76 vs. 0.71) and the AHD was slightly lower (0.17 vs. 0.20). These improved results may be due to our use of local B-spline registration to identify contours as opposed to intensity-based registration. In addition, higher-resolution micro-CT atlases were used in this study, and these were chosen from a large sample size to incorporate much of the natural variation in FN anatomy.

Comparing results against Noble et al. (15) is more difficult as they did not report DSC, but reported the distance between structure voxels from automatic to manual segmentations, as well as from manual to automatic segmentations. When averaging these results, their mean distance was 0.136 mm and maximum distance was 0.82 mm. This is similar to the current study which had an AHD of 0.17 mm.

This study is the first to describe an algorithm which segments the entire FN, including the labyrinthine segment. This portion of the nerve is of importance in VR simulation, especially for patient-specific simulation of difficult skull-base cases. The qualitative results of the automated FN segmentations were quite accurate in Figure 3.4 and 2.5. The labelmaps in Figure 3.4C-F did show pixelation despite using high-resolution micro-CT atlases, however this could be refined in a post-processing step as described by Lu et al (27,28).

When examining the DSC and AHD for the entire FN, the values were significantly worse than when limiting the segmentation to the ROI. This is likely due to the significant variation seen at the meatal foramen and at the stylomastoid foramen in Figure 3.5. This variation is not seen when an ROI is used, as there are specific start and stop points to the nerve. Noble et al. (15) specifically describe cropping the automatic segmentations to match the manual segmentations in order to avoid biasing the quantitative analysis.

Despite the worse quantitative results when considering the entire FN, the AHD was still less than a single voxel on average for a clinical CT scan. Therefore, the automated algorithm would be comparable to manual segmentation by an expert, and provide a suitable starting point for surgical training purposes. If the results were to be used to surgical planning or robotic surgery, the results could quickly be reviewed by an expert to ensure there are no errors.

This algorithm was implemented in 3D Slicer to allow for free, easy deployment to surgical trainees using VR simulation for rehearsal. The initial registration of centerlines takes a few seconds, while a B-spline registration can take 2-3 minutes using a modern processor.

3.5 Conclusion

This is the first automated algorithm to segment the entire FN, including the labyrinthine portion. The algorithm is unique in its use of high-resolution micro-CT atlases, and the sub-voxel accuracy when applied to low resolution scans was promising. This algorithm will be deployed as part of an open-source toolkit to allow for easy incorporation in VR simulation and surgical rehearsal.

References

1. Hohman MH, Hadlock TA. Etiology, diagnosis, and management of facial palsy: 2000 patients at a facial nerve center. The Laryngoscope. 2014;: p. E283-E293.

2. Bradbury ET, Simons W, Sanders R. Psychological and social factors in reconstructive surgery for hemi-facial palsy. Journal of Plastic, Reconstructive & Aesthetic Surgery. 2006; 59(3): p. 272-278.

3. VanSwearingen JM, Cohn JF, Turnbull J, Mrzai T, Johnson P. Psychological distress: Linking impairment with disability in facial neuromotor disorders. Otology - Head and Neck Surgery. 1998; 118(6): p. 790-796.

4. Kharat RD, Golhar SV, Patil CY. Study of intratemporal course of facial nerve and its variations — 25 temporal bones dissection. Indian Journal of Otolaryngology and Head & Neck Surgery. 2009; 61(1): p. 39-42.

5. Tüccar E, Tekdemir I, Aslan A, Elhan A, Deda H. Radiological anatomy of the intratemporal course of facial nerve. Clinical Anatomy. 2000; 13(2): p. 83-87.

6. Ryu NG, Kim J. How to Avoid Facial Nerve Injury in Mastoidectomy? Journal of Audiology & Otology. 2016; 20(2): p. 68-72.

7. Babu Nishant M. Intratemporal Facial Nerve Trauma- A Study of 40 Cases. Journal of Evolution of Medical and Dental Sciences. 2018; 7(12): p. 1465-1467.

8. Wiet GJ, Stredney D, Sessanna D, Bryan JA, Welling DB, Schmalbrock P. Virtual temporal bone dissection: An interactive surgical simulator. Otolaryngology - Head and Neck Surgery. 2002; 127(1): p. 79-83.

9. Chan S, Li P, Locketz G, Salisbury K, Blevins NH. High-fidelity haptic and visual rendering for patient-specific simulation of temporal bone surgery. Computer Assisted Surgery. 2016; 21(1): p. 85-101.

10. Sorensen MS, Mosegaard J, Trier P. The visible ear simulator: A public PC application for GPU-accelerated haptic 3D simulation of ear surgery based on the visible ear data. Otology and Neurotology. 2009; 30(4): p. 484-487.

11. Nash R, Sykes R, Majithia A, Arora A, Singh A, Khemani S. Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator. Journal of Laryngology and Otology. 2012; 126(7): p. 663-669.

12. Fried MP, Sadoughi B, Weghorst SJ, Zeltsan M. Construct Validity of the Endoscopic Sinus Surgery Simulator. Archives of Otolaryngology - Head & Neck Surgery. 2007; 133(4): p. 350.

13. Endo K, Sata N, Ishiguro Y, Miki A, Sasanuma H, Sakuma Y, et al. A patient-specific surgical simulator using preoperative imaging data: an interactive simulator using a three-dimensional tactile mouse. Journal of Computational Surgery. 2017; 1(1): p. 1-8.

14. Hassan K, Dort J, Sutherland G, Chan S. Evaluation of Software Tools for Segmentation of Temporal Bone Anatomy. In Volume 220: Medicine Meets Virtual Reality 22; 2016; Los Angeles. p. 130-133.

15. Noble JH, Warren FM, Labadie RF, Dawant BM. Automatic segmentation of the facial nerve and chorda tympani in CT images using spatially dependent feature values. Medical Physics. 2008; 35(12): p. 5375-5384.

16. Noble JH, Warren FM, Labadie RF, Dawant BM. Automatic segmentation of the facial nerve and chorda tympani using image registration and statistical priors. In Proceedings Volume 6914, Medical Imaging 2008: Image Processing; 2008; San Diego.

17. Noble JH, Dawant BM, Warren FM, Labadie RF. Automatic identification and 3D rendering of temporal bone anatomy. Otology and Neurotology. 2009; 30(4): p. 436-442.

18. Reda FA, Noble JH, Rivas A, Labadie RF, Dawant BM. Model-based segmentation of the facial nerve and chorda tympani in pediatric CT scans. In SPIE Medical Imaging; 2011; Orlando.

19. Reda FA, Noble JH, Rivas A, McRackan TR, Labadie RF, Dawant BM. Automatic segmentation of the facial nerve and chorda tympani in pediatric CT scans. Medical Physics. 2011; 38(10): p. 5590-5600.

20. Powell KA, Liang T, Hittle B, Stredney D, Kerwin T, Wiet GJ. Atlas-Based Segmentation of Temporal Bone Anatomy. International Journal of Computer Assisted Radiology and Surgery. 2017; 12(11): p. 1937-1944.

21. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magnetic Resonance Imaging. 2012; 30(9): p. 1323-1341.

22. Hudson T, Gare B, Allen D, Ladak H, Agrawal S. Segmentation of the Facial Nerve and Other Temporal Bone Structures for Patient-specific Simulation in Surgical. In Otology; 2018; Quebec City.

23. Heman-Ackah SE, Gupta S, Lalwani AK. Is facial nerve integrity monitoring of value in chronic ear surgery. The Layrngoscope. 2013; 123(1): p. 2-3.

24. Pluim JPW, Maintz JBA, Viergever MA. Mutual-information-based registration of medical images: a survey. IEEE Transactions on Medical Imaging. 2003; 22(8): p. 986-1004.

25. Nocedal J. Updating quasi-Newton matrices with limited storage. Mathematics of Computation of the American Mathematical Society. 1980; 35(151): p. 773-782.

26. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. BMC Medical Imaging. 2015; 15(1): p. 29.

27. Lu P, Barazzetti L, Chandran V, Gavaghan K, Weber S, Gerber N, et al. Facial nerve image enhancement from CBCT using supervised learning technique. In 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2015; Milano. p. 2964-2967.

28. Lu P, Barazzetti L, Chandran V, Gavaghan K, Weber S, Gerber N, et al. Highly Accurate Facial Nerve Segmentation Refinement From CBCT/CT Imaging Using a Super-

Resolution Classification Approach. IEEE Transactions on Biomedical Engineering. 2018; 65(1): p. 178-188.

29. Arora A, Hall A, Kotecha J, Burgess C, Khemani S, Darzi A, et al. Virtual reality simulation training in temporal bone surgery. Clinical Otolaryngology. 2015; 40(2): p. 153-159.

30. Reda FA, Noble JH, Rivas A, Labadie RF, Dawant BM. Model-based segmentation of the facial nerve and chorda tympani in pediatric CT scans. In Proceedings Volume 7962, Medical Imaging 2011: Image Processing; 2001; Orlando.

31. Oliveira FPM, Tavares JMRS. Medical image registration: A review. Computer Methods in Biomechanics and Biomedical Engineering. 2014; 17(2): p. 73-93.

Chapter 4

4 Conclusions and Future Work

4.1 Conclusions

The purpose of this work was to develop a semi-automatic segmentation algorithm to segment the facial nerve (FN) from clinical computed-tomography (CT) images that minimizes the amount of user involvement required. Previous work was done to model and to analyze the variability of the FN as discussed in Chapter 2. In the process of modeling and analyzing the FN, ground truth segmentations were formed which were used to evaluate the segmentations discussed in Chapter 3.

In Chapter 3, a novel, fiducial-initialized, semi-automatic, multi-atlas based approach to FN segmentation was introduced which utilizes micro-CT atlas information. In previous segmentation approaches, fully atlas-based segmentation approaches were not used due to difficulties in segmenting the FN such as low contrast between the FN and surrounding air cells, as well as dehiscence or discontinuity of bone surrounding the FN. This work differs from previous work by utilizing micro-CT atlas information which offers improved contrast, helping to mitigate the aforementioned issues. Previous results did not include or discuss segmentation of the entire intra-temporal FN, most commonly considering only the mastoid and tympanic segments. The algorithm presented in this work segments the labyrinthine segment which, to the best of our knowledge, has not been segmented as reported previously. Another novelty of the algorithm presented here was the use of multiple atlases. Previous work had used multiple atlases to form an average atlas as well as to create models which could be used for segmentation, however in these processes individual atlas information was lost. The current work more closely represents a classic multi-atlas approach, but instead of producing

a segmentation with every atlas and merging segmentations at the end, the best atlas candidate was chosen for single-atlas segmentation.

The segmentation results in this work were a DSC of 0.76 and an AHD of 0.17 mm for surface voxels within surgically relevant regions and all segmentations successfully followed the course of the FN, making them suitable for surgical simulation. We believe that the use of multiple atlas candidates bypassed previous issues with atlas-based segmentation by choosing an appropriate atlas for segmentation which reasonably represented image volumes to be segmented. Due to the variability of the FN, it is unlikely that a single atlas candidate could be used to provide segmentations which consistently follow the course of the FN. These results were comparable to what has been shown in the existing literature. It is however difficult to fairly compare our results to those reported in the literature due to the use of differing test sets. Previous algorithms were tested on different image volumes with differing anatomy and voxel size and imaging parameters were not always included or consistent with our own. While there is no standardized test set for FN segmentation, we believe our results are competitive since our test set has the largest voxel size of all test sets, meaning that we are producing segmentations with less image information than other algorithms. Considering the limited information used, the results presented here offer a good representation of results that may be seen for segmenting new FNs in a clinical setting. The computation time of this algorithm was approximately five minutes per temporal bone, which is significantly shorter than the time required to manually segment the FN.

4.2 Future Work

To ensure that the present work is robust, future testing would be beneficial. In addition to our test set of image volumes, we have produced segmentations using images acquired from two other scanners. The results of these segmentations were not quantified for accuracy since there were no ground truth segmentations for these image volumes, but visual inspection of these segmentations demonstrated that the FN segmentations accurately followed the course of the FN. Additional testing on image volumes from these two scanners would be beneficial as well as testing on image volumes produced using other scanners in order to ensure that our results could be reproduced elsewhere. Furthermore, fixed cadavers may result in improved contrast when compared to living patients, and it is important that the algorithm is validated outside of fixed cadaveric specimens.

Another direction for future work is investigation of pediatric cases. Children often undergo mastoidectomy to insert CIs, so it is important that the algorithm can handle pediatric cases as well. The primary differences between pediatric and adult FN morphology is that in pediatric temporal bones, the FN has a sharper turn at the second genu, causing the mastoid tip located at the stylomastoid foramen in adult temporal bones to be located inferiomedially to the mastoid tip in pediatric temporal bones. Despite the atlas set being developed to represent the variability of adult FNs, the aforementioned turn at the second genu is the primary feature considered in our atlas selection and it is thereby expected that some of the atlases chosen would be representative of pediatric cases. It is still possible that this atlas set may not perform well in all pediatric cases, and since acquisition of pediatric temporal bones to produce micro-CT atlases for augmenting the current atlas set would be extremely difficult and impractical, it may be suitable to create a separate atlas set specifically for pediatric cases with clinical CT scans instead.

In addition to producing results on a wider test set and considering pediatric cases, future work may adapt our algorithm into a more user-friendly and streamlined process by coding our algorithm in C++ or Python while utilizing the Insight Toolkit. We could then incorporate the algorithm into a 3D Slicer module with a graphical user interface which would allow users to be guided through the process of using the algorithm.

Curriculum Vitae

Name: Bradley Michael Gare, BESc.

Education:

2017 – Present   M.E.Sc. (Electrical & Computer Engineering)

                      Western University

                      London, Ontario, CA

2013 – 2017   BESc. (Computer Engineering)

                      Western University

                      London, Ontario, CA

Honours and Awards:

2017 – 2018   Ontario Graduate Scholarship

2013 – 2017   Continuing Admission Scholarship

2013 – 2017   Dean's Honour List

Related Work Experience:

2018 – 2019   Graduate Teaching Assistant (Introduction to Electrical
                      Engineering – ECE 2238)

2017 – 2018   Graduate Teaching Assistant (Digital Logic Systems – ECE
                      2277)

Publications:

1. Van Osch K, Allen D, **Gare B**, Hudson T, Ladak H, Agrawal S. Morphological analysis of sigmoid sinus anatomy: Clinical applications to neurotological surgery. Journal of Otolaryngology-Head & Neck Surgery. 2019; 48(1), 1-8. doi:10.1186/s40463-019-032
2. Hudson T, **Gare B**, Allen D, Ladak H, Agrawal S. Intrinsic Measures and Shape Analysis of the Intratemporal Facial Nerve. *Journal of Otolaryngology*. Submitted Feb 25, 2019.
3. Rohani A, Allen D, **Gare B**, Zhu N, Ladak H, Agrawal S. High-Resolution Imaging of the Human Incudostapedial Joint Using Synchrotron-Radiation Phase-Contrast Imaging. *Journal of* Microscopy. Submitted Mar 26, 2019.

4. **Gare B**, Hudson T, Rohani A, Allen D, Agrawal S, Ladak H. Multi-Atlas Segmentation of the Facial Nerve from Clinical CT. *International Journal of Computer Assisted Radiology and Surgery*. Submitted April 29, 2019.

Conference Presentations:
* denotes presenting author

1. Allen DG*, Van Osch KA, **Gare B**, Agrawal SK, Ladak HM. Automated digital modeling of the sigmoid sinus from 3D micro-CT images for surgical training. Abs. 41st Midwinter Research Meeting of the Association for Research in Otolaryngology, San Diego, CA, February 2018.
2. Rohani SA, **Gare B***, Iyaniwura J, Zhu N, Agrawal SK, Ladak HM. Optimal parameters for cochlear implant visualization using in-line synchrotron-radiation phase-contrast imaging. Abs. 41st Midwinter Research Meeting of the Association for Research in Otolaryngology, San Diego, CA, February 2018.
3. Van Osch K*, Allen D, **Gare B**, Hudson T, Ladak H, Agrawal S. Morphological analysis of sigmoid sinus anatomy: Clinical applications to neurotological surgery. Canadian Society of Otolaryngology – Head and Neck Surgery Annual Meeting, Quebec, QC, June 2018.
4. Hudson T*, **Gare B**, Allen D, Ladak H, Agrawal S. Segmentation of the facial nerve and other temporal bone structures for patient-specific simulation in surgical otology. Canadian Society of Otolaryngology – Head and Neck Surgery Annual Meeting, Quebec, QC, June 2018.
5. Rohani SA, **Gare B**, Iyaniwura J, Zhu N, Agrawal SK, Ladak HM*. In-line synchrotron-radiation phase-contrast imaging of the implanted cochlea. 15th International Conference on Cochlear Implants and Other Implantable Auditory Technology, Antwerp, Belgium, June 2018.