Segmentation and study of anatomical variability of the cochlea from medical images

Thomas Demarcy

To cite this version:

Thomas Demarcy. Segmentation and study of anatomical variability of the cochlea from medical images. Other. COMUE Université Côte d'Azur (2015 - 2019), 2017. English. NNT : 2017AZUR4052 . tel-01609910

HAL Id: tel-01609910
https://theses.hal.science/tel-01609910v1

Submitted on 4 Oct 2017

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE STIC
SCIENCES ET TECHNOLOGIES DE L'INFORMATION
ET DE LA COMMUNICATION

PHD THESIS

to obtain the title of

PhD of Science

of the University of Nice Sophia Antipolis
Specialty : COMPUTER SCIENCE

Written by
Thomas DEMARCY

Segmentation and Study of Anatomical Variability of the Cochlea from Medical Images

prepared at Inria Sophia Antipolis, ASCLEPIOS Research Team
to be defended on July 4th, 2017

Jury :

Advisors :	Nicholas AYACHE	-	Inria (Asclepios)
	Hervé DELINGETTE	-	Inria (Asclepios)
Reviewers :	Miguel Angel GONZÁLEZ BALLESTER	-	Pompeu Fabra University, Spain
	Rasmus PAULSEN	-	Technical University of Denmark
	Mauricio REYES	-	University of Bern, Switzerland
President :	Maureen CLERC	-	Inria (Athena)
Examinators :	Dan GNANSIA	-	Oticon Medical
Invited :	Nicolas GUEVARA	-	Head and Neck University Institute (IUFC, Nice)
	Charles RAFFAELLI	-	Nice University Hospital (CHU)

Acknowledgments

J'aimerais remercier chaleureusement mes directeurs de thèse Nicholas Ayache et Hervé Delingette pour leur soutien et leurs conseils éclairés au cours de ces trois dernières années. Hervé, merci énormément pour le temps consacré et les excellentes idées qui ont grandement participées aux résultats présentés dans ce manuscrit. Nicholas, merci de m'avoir accepté dans ton équipe, où tu as su réunir toutes ces personnes talentueuses et créer une atmosphère de travail chaleureuse.

À Dan Gnansia, un grand merci pour ton encadrement bienveillant et d'avoir été mon interlocuteur privilégié au sein d'Oticon Medical. J'ai une pensée pour Jonathan Laudanski, qui a initié ce projet avec une grande motivation et qui nous a tragiquement quittés.

Mille mercis aux cliniciens niçois, Nicolas Guevara, Charles Raffaelli et Clair Vandersteen, pour votre vision éclairée de l'otologie et de la médecine computationnelle. Nicolas, merci d'avoir su fédéré par ton enthousiasme chercheurs, cliniciens et industriels autour de ce projet prometteur. Charles, merci d'avoir partagé la passion sans limite de ton travail et de m'avoir convaincu de faire partie de cette aventure. Clair, mon binôme, chien fou et touche à tout de génie, j'espère que tout ceci n'est qu'un début !

Je suis très reconnaissant aux rapporteurs qui ont pris le temps de lire mon manuscrit : Miguel Gonzalez, Rasmus Paulsen et Mauricio Reyes. Merci pour vos commentaires et vos compliments. Merci à Maureen Clerc d'avoir accepté de faire partie de mon jury.

Merci à tous les collègues de l'équipe Asclepios. Jan, merci d'avoir été un co-bureau hors pair, je n'oublierai pas les bons moments passés ensemble que ce soit au travail, à Bratislava ou sur les rivières ! Milky Matt et Chloé, quel dommage qui vous soyez maintenant exilés si loin, j'espère vous revoir vite ! Special thanks to Marc-Mich et Nina, mes compagnons roux des premiers jours. Enfin, et j'espère n'oublier personne, merci à Alan, Anant, Aurélie, Aziz, Bishesh, Federico, Florent, Florian, Hakim, Héloïse, Hervé L, Hugo, Isabelle, Krissy, Loïc, Luigi, Manon, Marco, Marine, Maxime, Mehdi, Mike, Nicoco C, Nicolas C, Nicolas D, Pauline, Pawel, Pietro, Qiao, Rafifou, Roch, Rocio, Shuman, Sofia, Sophie, Thomas, Vikash, Wen, Xavier et Yann sans qui mon quotidien n'aurait pas été le même !

Merci à tous ceux qui ont animé l'Epic Sax House, je pense particulièrement à Zoomizoom, Lionyyy, Nina, Clémencebenardcinéma, Nico, Manon, Zouzemarmar, Charline Leguenec et Schmittyschmitt.

Je remercie également les rostagniens, les kayakistes et mes amis parisiens qui m'ont encouragé à distance, en particulier Paul, Jules, Clément et Clara.

J'ai également une pensée particulière pour Lucile et tous les Auvergnats pour m'avoir supporté et apporté cette bouffée d'air frais dont j'avais tant besoin pendant ces trois ans.

Enfin je remercie toute ma famille pour leur soutien inconditionnel, mes parents que je ne pourrais jamais suffisamment remercier pour l'éducation qu'ils m'ont donnée, et ma jumelle Laura.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

1.1 Context of the Thesis

Computer Tomography (CT) scanning is now used routinely by radiologists and surgeons for diagnosis, surgery planning and postoperative evaluation, to the point that clinicians have access to a large number of data. However, the process of analyzing CT data is complex, time consuming and requires highly trained professionals.

The cochlea, the sensory organ of hearing, is an anatomical structure whose CT images is specifically difficult to analyze. Cochlear internal cavities have complex spiraling shapes and are nearly invisible with clinically available CT scans due to the relative small size of the cochlea with respect to the scanner resolution.

Cochlear implants (CI) are used to treat severe to profound hearing loss by surgically implanting these electronic medical devices into the cochlea. CI, which convert sound energy to electrical stimulation, were the first implanted neural prosthesis and are now the most widespread with approximately 400.000 cochlear-implanted individuals worldwide. The implantation procedure involves drilling through the mastoid part of the temporal bone to open one of the three cochlear chambers and insert an electrode array to directly stimulate the auditory nerve, which induces the sensation of hearing. The surgical procedure performed manually with limited visual and sensitive feedback is challenging. The functional outcome of the implantation often varies among patients and depends on multiple factors, which are not yet clearly identified. In particular, hearing restoration is correlated with the preservation of innervated cochlear structures.

1.2 Focus of the Thesis

Patient-specific three-dimensional reconstruction of cochlea and its substructures could contribute to the improvement of different aspects of cochlear implantation. First of all, it can provide a better anatomical understanding for the clinicians and suggest electrode array design improvements for cochlear implant manufacturers. Quantitative analysis of the anatomical cochlear shape from medical images is essential for diagnosis of shape abnormality such as cochlear hypoplasia, an incomplete development of the cochlea. Using three-dimensional anatomical reconstruction, the cochlear implantation could be optimized to be the least traumatic by selecting the drilling trajectory that avoids sensitive structures such as the facial nerve and that estimates the insertion angle between the electrode and the basal turn of the cochlea.

The insertion can be evaluated postoperatively in order to give feedback for the surgeons and to improve the set-up of the CI by evaluating to exact position of the electrode with respect to cochlear nerve. Acquisition of surgical skills requires a lot of practice and deep anatomical knowledge, thus three-dimensional models could be used for surgery simulation and training.

In the context of this thesis, several clinical and methodological questions arise:

- **What is a proper geometric model to measure the cochlear shape?** The particular spiraling shape of the cochlea leads to a need for new application-specific models in order to estimate anatomical measurements of the cochlea. What is the variability of this shape among a population? Statistics on the quantitative measurements help to establish normal and healthy shapes and detect outliers. Is there intra-patient bilateral symmetry?

- **Can we construct a "good" cochlear shape model?** That is to say a model anatomically correct that can provide meaningful and intuitive clinical interpretation. According to Occam's razor, the law of parsimony, the model should be compact, while providing a good generalization, namely the ability to represent the shape variation across the population. Eventually the shape variability could be embedded in the model in order to generate realistic data or provide prior knowledge for model fitting.

- **What information can we extract from CT images?** Is this information sufficient to extract clinically exploitable measurements such as the maximal depth insertion? To answer these questions one must quantify the uncertainty of the measurements extracted from CT images in relation to more precise acquisition methods such as high-resolution μCT.

- **How can be addressed industrial and clinical challenges of cochlear implantation?** For example, the following questions are of interest: Can we optimize patient-specifically the choice of CI model before the surgery? Can we estimate postoperatively the position of the electrode array with respect to the cochlear internal cavities?

1.3 Structure of the Thesis

The thesis work is presented in a chronological order. Although a given chapter relies on the preceding ones, each chapter is meant to be self-contained.

Chapter 2 provides background knowledge about the anatomy and physiology of hearing, focusing on the cochlea and its imaging. The chapter includes an introduction to cochlear implant and its surgical procedure as well as a short state of the art review on segmentation of the cochlea from medical images.

Chapter 3 aims at defining automatic image processing methods adapted to the spiral shape of the cochlea to study the cochlear shape variability from high-resolution μCT images. This chapter is adapted from [Demarcy 2017].

Chapter 4 aims at developing and evaluating a new parametric cochlear shape model. The model is applied to extract patient-specific clinically relevant metrics such as the maximal insertion depth for different designs of CI electrode array. Thanks to the uncertainty quantification, provided by the model, we can assess the reliability of CT-based segmentation as compared to the ground truth segmentation provided by μCT scans. This chapter is adapted form [Demarcy 2016b] and its clinical application has been presented in [Gnansia 2016].

Chapter 5 describes a joint model of the cochlear shape (and its substructures) model and its appearance within a generative probabilistic Bayesian framework. The proposed segmentation method estimates jointly the shape and appearance parameters and applies an iterative expectation-maximization (EM) strategy that interleaves shape model parameters fitting and image segmentation with mixture of Student's t-distributions. The method was applied to a large database of 987 CT images and allowed the statistical characterization of the cochlear anatomical variability along with the quantification of the bilateral symmetry.

Chapter 6 highlights a clinical application of cochlear substructures segmentation as presented in [Demarcy 2016a], which consists of assessing postoperatively the position of the electrode array with respect to the anatomy, and concludes the thesis.

Chapter 2

Cochlear Anatomy and Imaging

Contents

2.1 Anatomy and Physiology of Hearing

The ear is the organ of hearing and balance. The hearing anatomy is usually described as having three parts:

- **Outer Ear** is the external part of the ear. It consists of the auricle (or pinna) which is the visible part of the external ear and ear canal (or external acoustic meatus), which is around 25 mm deep. It amplifies the sound and focuses it on the eardrum which separates the inner and middle ears.

- **Middle Ear** is the internal part of the ear between the eardrum and the oval window. It contains the eardrum and three ossicles: malleus (or hammer), incus (or anvil) and stapes (or stirrups). The chain of ossicles transfers the compression waves in air at the eardrum into waves in the fluid and membranes of the inner ear. The vibration may be stiffened, when facing with for example very loud sound, by two muscles: the stapedius muscle and the tensor tympani muscle. The footplate of the stapes is connected to the oval window. Through the middle ear area passes also the facial nerve and its branch, the chorda tympani.

- **Inner Ear** is the innermost part of the ear. It consists of a closed and fluid-filled cavity, the bony labyrinth which can be separates in two parts: the cochlea, dedicated to the hearing and the vestibular system, dedicated to the balance (equilibrioception). The membranous labyrinth is inserted within the bony labyrinth and has the same general shape. The vestibular system consists of the three semicircular canals, providing sensory input of rotary movements

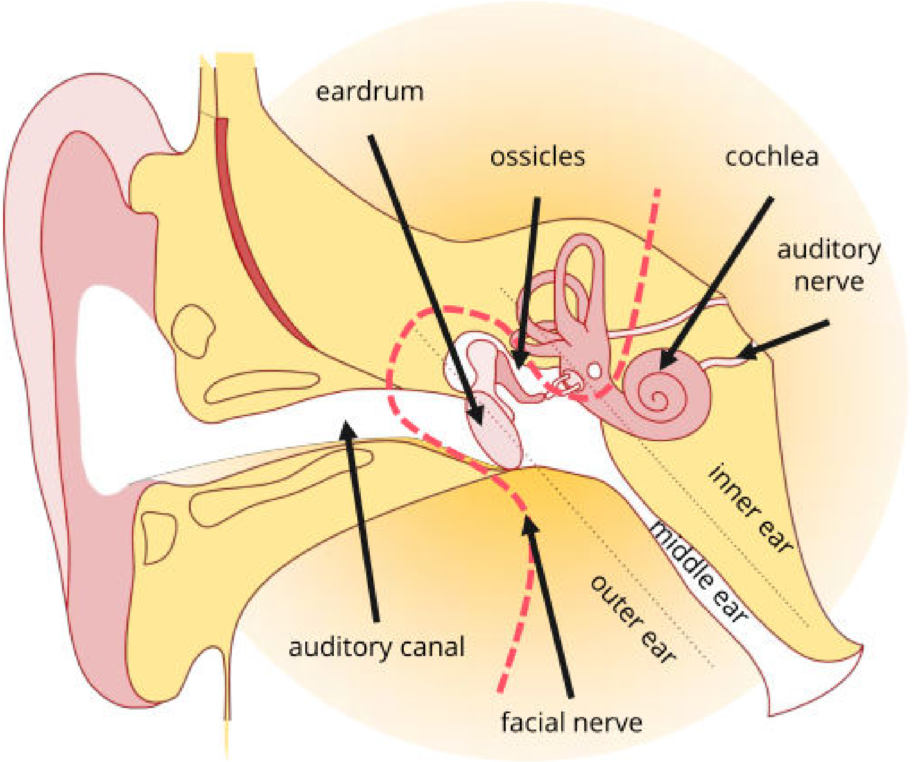


Figure 2.1: Anatomy of hearing adapted from [Miroir 2012]

along the pitch, roll and yaw axes (kinetic balance), and the vestibule, providing sensory input of acceleration and head-tilts in the horizontal plane, in the utricle, and in the vertical plane, in the saccule (static balance). The inner ear has two membrane-covered openings into the middle: the round and oval windows which vibrate with opposite phase since the fluid is nearly incompressible.

The cochlea (Latin word for snail) is the small organ responsible of the sense of hearing. In this small spiral structure, the sound, a mechanical wave is transformed into an electrical pulse (Fig. 2.2). A normal human cochlea completes about two and a half turns. It grows to its definitive size within the first 17 to 19 weeks of gestation, then the otic capsule, also known as the bony labyrinth ossifies. The cochlea, in which waves propagate from the base (near the round and oval window) to the apex (the top), is mainly composed of fluids inside soft tissues and bony walls. The whole cochlea is within the petrous part of the temporal bone. The cochlea is a set of different spiraling substructures called scalae: the scala tympani (or tympanic duct), the scala vestibuli (or vestibular duct) and the scala media (or cochlear duct). The scala tympani and the scala vestibuli are filled with perilymph and communicate at the most apical part of the cochlea, called helicotrema. The scala media is filled with endolymph. Perilymph and endolymph contains different ionic composition and concentration. The cochlear cross-sections look roughly like a cardioid. The cochlea

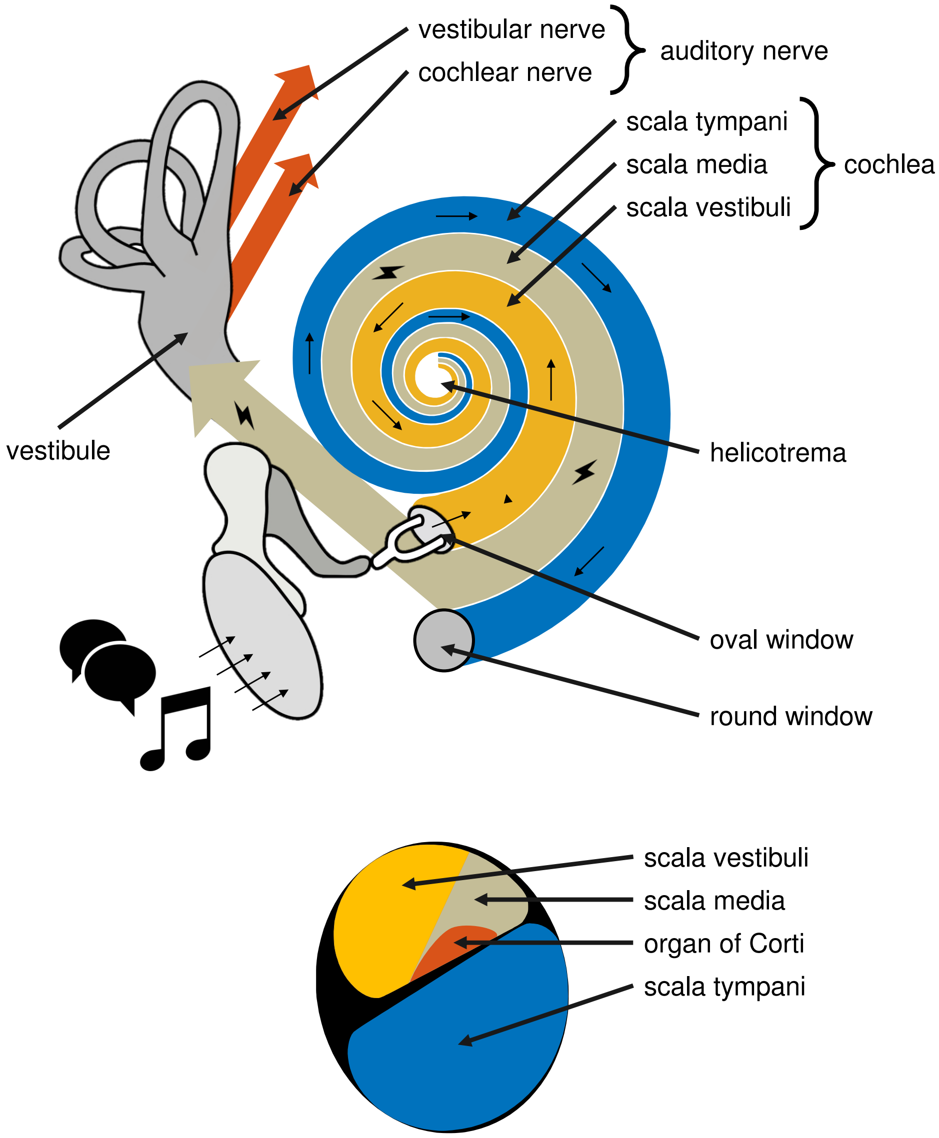


Figure 2.2: Sketch of the physiology of hearing (top) and of the cross-section of the cochlea (bottom). Mechanical waves are represented with arrows and electrical pulses with thunderbolt symbols.

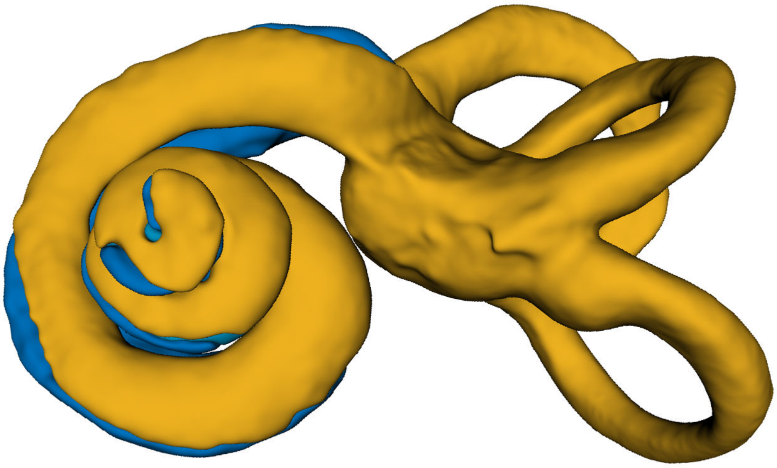


Figure 2.3: Three-dimensional model of the inner ear

is wrapped around the modiolus, a coned-shape porous bone with a theoretical axis of revolution, the modiolar axis. From the modiolus in the radial direction is the osseous spiral lamina and the basilar membrane, which separate the scala tympani and the scala vestibuli. The Reissner's membrane separates the scala media and the scala vestibuli. The organ of Corti (Fig 2.5) is a tiny structure located in the scala media on top of the basilar membrane that contains the hair cells, sensory receptors linked to the auditory nerve. Above the organ of Corti, the tectorial membrane overlies the hair cells and stimulates them in neural processing of sounds.

Frequency Analysis The range of audible sound frequency for human is generally estimated between 20 Hz and 20 kHz [Greenwood 1990]. There is a relationship between the perceived sound frequency and the anatomical location in the cochlea. This relationship is called the tonotopic map. The tonotopy is mostly linked to the mechanical properties of the basilar membrane. Along the length of the cochlear spiral, the width and the stiffness of the basilar membrane gradually vary and it resonates at specific wave-frequency. Basically the resonance frequency reduces from the base of the cochlea to the apex. Neurons can be activated selectively and interpreted by the brain.

2.2 Hearing Loss

Preserving the ear integrity is essential to ensure the perception of a stimulus. Hearing requires good functional outcome of the outer, middle and inner ear in order to best convert acoustic pressure waves into an electrical pulses. Otology is one of the specialties of Otorhinolaryngology and is aimed at treating ear pathologies. There are different types of hearing impairments, such as conductive, neurosensory or mixed deafness. Surgical hearing loss treatments include otosclerosis surgery,

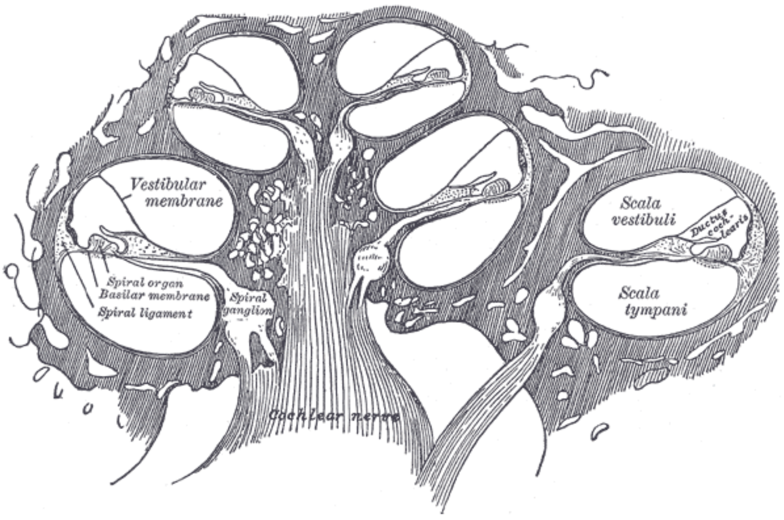


Figure 2.4: Diagrammatic longitudinal section of the cochlea (from [Gray 1918])

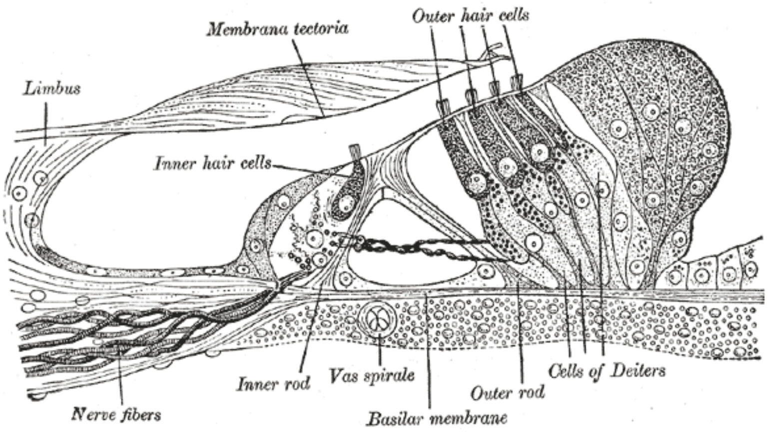


Figure 2.5: Section through the spiral organ of Corti (from [Gray 1918])

tympanoplasty, or cochlear implantation.

Conductive Deafness occurs when the sound waves are not properly conducted in the external ear canal and the middle ear. Damages to the ossicular chain, in the case of otosclerosis notably, can be the cause of conductive hearing loss. Most often, it is reversible and can be treated with hearing aids or surgical procedures.

Sensory Neural Deafness originates from sensory organ, vestibulocochlear nerve or neural part. When the root cause lies in the cochlea, it is caused by loss or degeneration of the hair cells. The amount of preserved hair cells is linked to the severity of the hearing loss (see Table 2.1), in the case of insensitivity of the cochlea direct neural stimulation with an auditory prosthesis is required.

rank	**decibels of hearing loss**
slight	16 to 25 dB HL
mild	26 to 40 dB HL
moderate	41 to 54 dB HL
moderately severe	55 to 70 dB HL
Severe	71 to 90 dB HL
Profound	greater than 91 dB HL

Table 2.1: Severity of hearing loss

Mixed Eventually, hearing loss can be caused by a combination of the two above categories.

2.3 Cochlear Implant Surgery

Cochlear Implant A cochlear implant (CI) is a surgically implanted electronic device that restores a sense of hearing. CI bypass the normal hearing process. Typically a cochlear implant system consists in an external speech processor generally worn behind the ear, which transmits a signal to an array of electrodes inserted in the cochlea, which stimulate the cochlear nerve.

The external processor is equipped with a microphone and several digital signal processors to perform signal processing, which basically consists in noise reduction, automatic gain control and decomposition into series of bandpass-filtered channels.

CI can be considered as one the great medical advances of the 20^{th} century provoking a revolution in the treatment of profound hearing loss [Wilson 2008]. This technology, commonly used in clinical practice since the 1990's, allowed to restore an usable hearing to patients handicapped by severe to profound deafness. Besides, CI transformed the lives of children born deaf allowing them to access the hearing world and to have an intellectual and educational development identical to normally hearing children. More recently, through technological and surgical progress, CI could be proposed to patients with severe hearing loss. The principle of CI is to rehabilitate the sense of sound by electrically stimulating the auditory pathways. CI substitutes the inner ear function by producing significant electrical stimulation around the damaged hair cells and directly stimulating the residual neurons of the auditory nerve, in order to reconnect the afferent signals to the central nervous system. CI consists of an array of about twenty electrodes placed along the cochlea which stimulate the auditory nerve with electrical impulses whose sequences depend on the cochlear tonotopy: the basal portion of the cochlea is normally sensitive to sounds with high pitch whereas the apical part of the cochlea is sensitive to sounds with low pitch. The objective is to transmit a signal to the brain that best reproduces the neuronal encoding of sound in natural hearing, the signal delivered by the implant must have an encoding adapted in frequency and intensity. The implanted patient needs then auditory re-education and speech therapy to learn

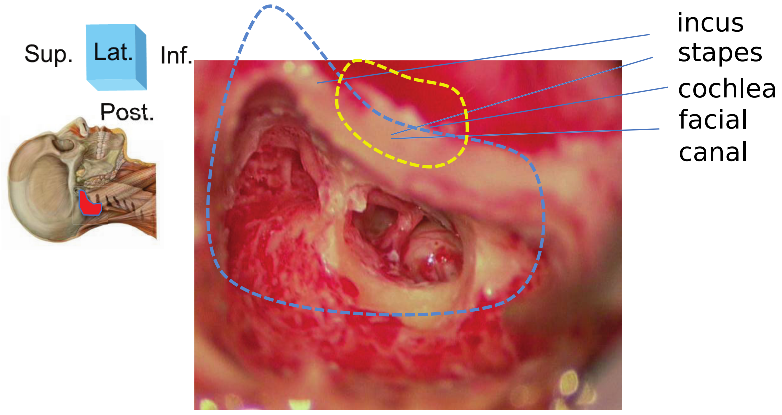


Figure 2.6: Surgical view of the approach for cochlear implantation (right ear). The approach is performed by mastoidectomy (blue dotted lines) and posterior tympanotomy (yellow dotted lines) (from [Nguyen 2011])

how to interpret newly perceived auditory stimuli.

Currently, CI has matured with more than 400,000 patients implanted worldwide. This number is much larger than the sum of all other types of neural prostheses. Furthermore, the restoration of the function provided thus far by CI far exceeds the ones achieved with other neuroprostheses. The major benefit in terms of quality of life compared to its cost allows its coverage by the majority of the health-care systems in the world. Moreover, CI remains a model in the development of other implants such as restoration of sight (retinal implants) or balance (vestibular implants). This remarkable capacity to adapt to the implant is linked to the capacity of the nervous system to create and reorganize neural networks: this is referred to as neuroplasticity. The brain plasticity can be observed throughout the life course, with however a maximum of effectiveness during the childhood. Thus creating the human-implant interface at the earliest performs the best functional outcome. The age of implantation is anti-correlated to the implant performance. Children implanted before the age of two generally perform better than those implanted over age five. Although some studies have confirmed that neural plasticity persists with postlingually deaf adult patients implanted with CI, the results are nevertheless inconsistent and dependent on multiple factors, which are not yet clearly identified.

Currently, the overall challenge with cochlear implantation is twofold. First of all, there is a limitation of speech recognition performance, which is progressing very slowly since the major evolution of the last century, which are the evolution from mono-electrode to multi-electrode and the evolution from simultaneous to sequential stimulation. Average speech recognition rates stagnate around 60% for monosyllabic word understanding in quiet environment. Understanding in noisy environments is even more limited. Secondly, there is a major interindividual variability, some patients recognize 100% of the words in quiet environment while others nearly 0%.

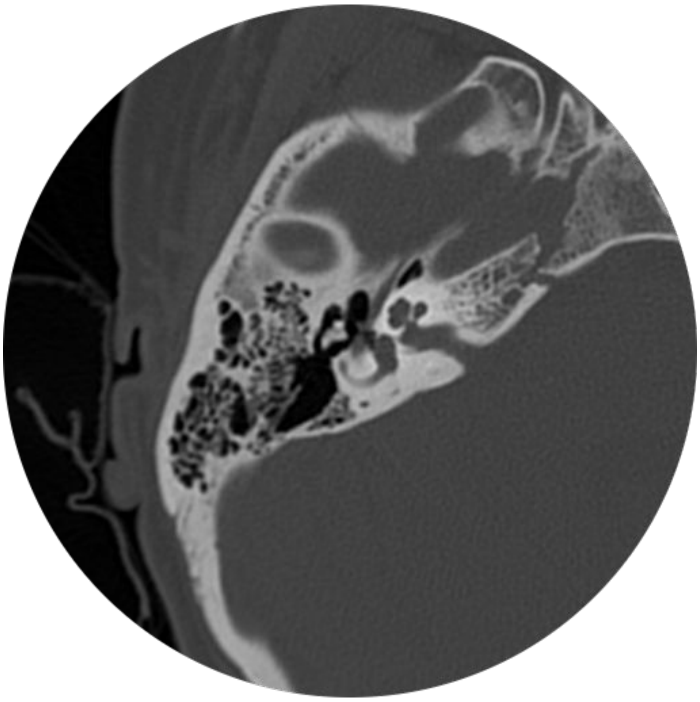


Figure 2.7: Temporal bone CT image

Surgical Procedure During the CI surgery, the insertion can lead to lesions of the modiolus, osseous spiral lamina or basilar membrane, resulting in degeneration of the hair and spiral ganglion cells. Because the viability of these cells is correlated with neuronal survival and speech expression, it is crucial to avoid destroying them during surgery and to minimize injury to the previously cited anatomic structures with so-called atraumatic or minimally invasive surgery techniques [Weber 2017]. An improved knowledge of the cochlea shape variability is not only essential for diagnosis of shape abnormality, atraumatic surgery planning and post-operative insertion assessment but it also provides a better anatomical understanding for the clinicians and can suggest electrode array design improvements for cochlear implant manufacturers.

Cochlear implantation is performed through a mastoidectomy and a posterior tympanotomy. The insertion of the electrode array is performed manually by the ear, nose, and throat (ENT) surgeon with limited visual and sensitive feedback.

Electrode Array Design Several parameters of the design of the electrode array can vary depending of the products proposed by cochlear implant manufacturers: the length, the number of electrodes, the electrode spacing, the width, the rigidity, the curvature, etc. The link between those parameters and the clinical outcomes such as the insertion trauma and the quality of hearing restoration is not completely understood and controlled.

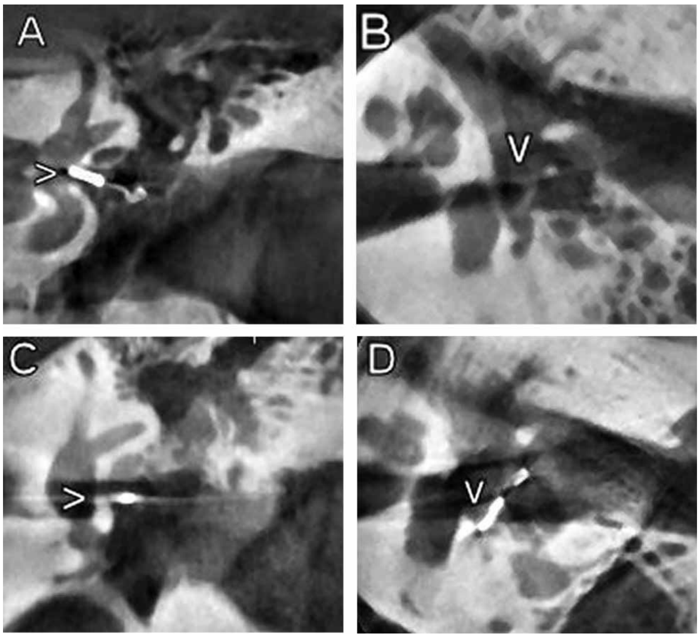


Figure 2.8: Temporal bone CBCT image (from [Peltonen 2009]) acquired to investigate postoperatively an ossicular prosthesis (arrow).

2.4 Inner Ear Imaging

Clinical CT and Cone Beam CT (CBCT) are commonly used to provide anatomical information about the cochlea and its surrounding structures. However, the relative small size of the cochlea with respect to the typical voxel size, restrains the observation of intracochlear details. The cochlea measures about 8.5x7x5 mm^3 while the typical CT image spacing is larger than 0.2 mm. The complete spiral can be difficult to recognize especially at the apex. The basilar membrane that separates the scala tympani and the scala vestibuli is partially visible at the first cochlear basal turn. CBCT (Fig. 2.8) has several advantages over CT (Fig. 2.7): the acquisition is faster with a lower dose exposure for the patient and a superior spatial resolution so that few cochlear structures, such as the osseous spiral lamina, seem identified more easily [Zou 2015]. However the reconstruction algorithm used for CBCT does not produce image as anatomically correct as classical CT. Cochlear implant electrodes produce metal artifacts in post-implantation CT [Reda 2014b], which makes difficult to characterize intracochlear anatomy (Fig. 2.9).

Micro-computed Tomography (μCT) is designed to assess the anatomy with high accuracy by acquiring images with voxel size of few micrometers (usually between 5 and 50 μm). This imaging modality is typically used to scan small animals or biomedical samples. Due to the limited size of the object that can be scanned (with a diameter smaller than 100 mm), it is not possible to acquire μCT scan of *in vivo* temporal bones. High resolution imaging can be performed on cadaveric specimens,

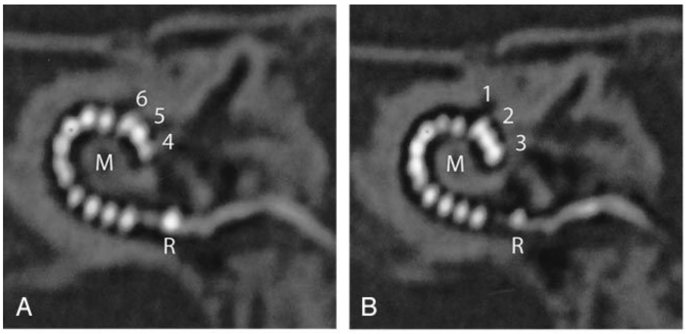


Figure 2.9: CT scan of an implanted cochlea (from [Verbist 2005])

after cropping the temporal bone around the bony labyrinth. It allows visualization of the fine cochlear structures, especially the thin bony structures that are well contrasted such as the osseous spiral lamina [Lane 2004]. With μCT modality, it becomes possible to assess the position of the scala tympani and the scala vestibuli. However, without chemical sample preparation and with a resolution above 10 μm the Reissner's membrane, that separates the scala media and the scala vestibuli is still not observable. Chemical sample preparations include dehydration (or drying), freezing, decalcification, fixation and staining (Fig. 2.10), and can produce artifacts such as tissue distortion.

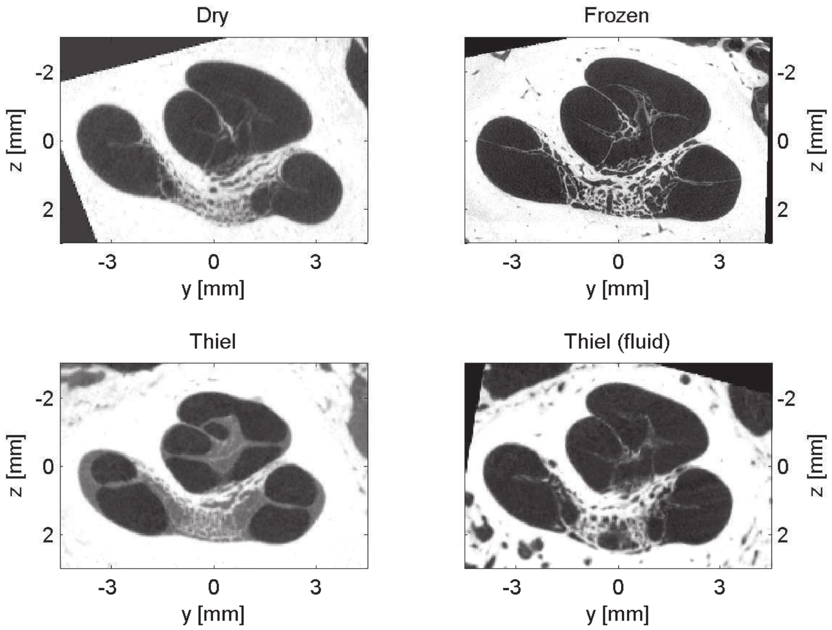


Figure 2.10: μCT scans of the cochlea with sample preparation (from [Kjer 2015b]). Thiel-solution was used for fixation with cochlea specimen without fluid (left) and fluid-filled specimen (right).

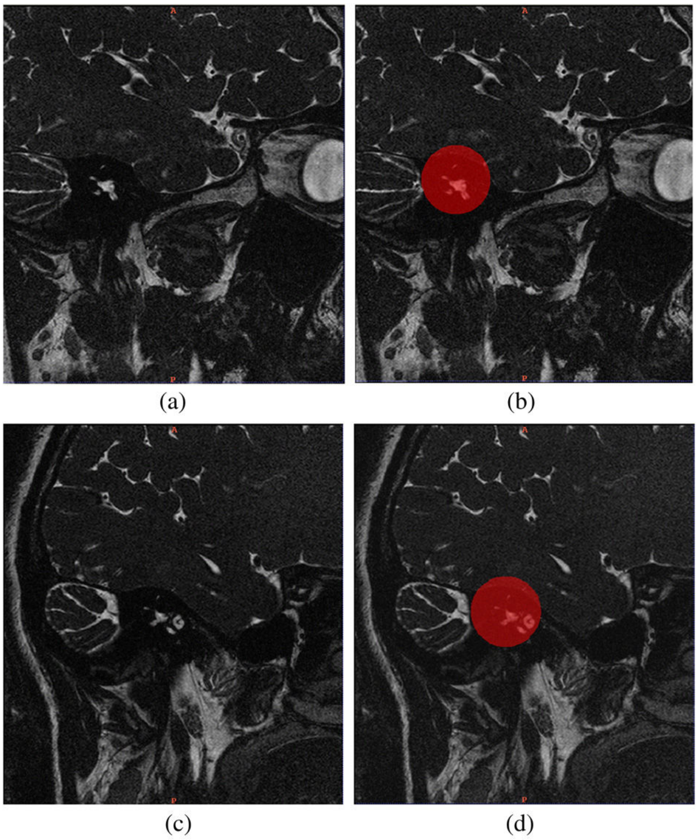


Figure 2.11: Temporal bone MRI sagittal slices (from [Shi 2011]). The cochlear region is indicated by semi-transparent red color.

MRI can be used as a complementary imaging modality (Fig. 2.11). Fluids within the membranous labyrinth can be imaged with good quality [Lane 2005, Shi 2011] and fusion of CT and MRI imaging can be used to identified nerves, such as the facial nerve and the chorda tympani [Bartling 2005].

2.5 Cochlear Segmentation From CT Images

The segmentation of intracochlear structures represents a great challenge. CT images of the temporal bone have low resolution with respect to the small size of the anatomical structures and the topology of the cochlea is complex. The main chambers of the cochlea, the scala tympani and the scala vestibuli, form a nested double helix where the spatial division between each turn (basal, middle and apical) is weakly visible and where the stiff structural elements separating the scala tympani and the scala vestibuli, such as the basilar membrane and the spiral lamina ossea are weakly visible or invisible. Finally the cochlea is filled with fluids which can be similarly found in the vestibular system and other neighboring structures, with similar appearance in CT images.

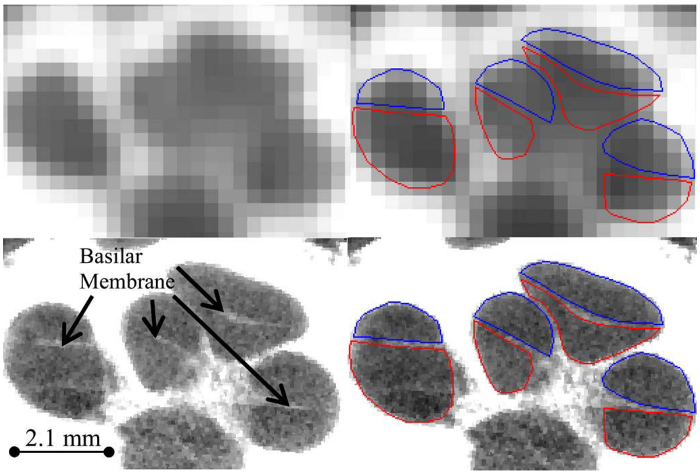


Figure 2.12: Example of cochlear segmentation (from [Noble 2011]). Segmentation of the scala tympani (red) and the scala vestibuli (blue) are shown on CT (top) and μCT (bottom) images.

This challenging task requires additional information to achieve successful segmentation. Anatomical atlases and shape models can be combined with intensity information to provide sufficient prior knowledge.

In order to performed segmentation of the cochlea, two main approaches have been previously developed can be summarized as follows.

First set of methods consists in an optimization procedure to perform parametric model fitting [Baker 2005]. A parametric model-image registration procedure is used to align the model surface with the image gradient.

Second set of methods consists in a non rigid registration procedure between an input CT image and an atlas embedding statistical shape prior [Noble 2010, Noble 2011, Reda 2011, Noble 2012, Reda 2012, Noble 2013, Reda 2013, Kjer 2014a, Reda 2014a, Reda 2014b, Kjer 2015d]. Anatomical modeling is performed from high-resolution datasets, usually ex-vivo μCT scans. The shape prior is provided by a statistical shape model (SSM) of the intra-cochlear anatomy, which guides the registration between a CT scan and an enhanced CT-atlas (Fig. 2.12).

CHAPTER 3

Analysis of Human Cochlea Shape Variability from μCT Images

Contents

This chapter is adapted from [Demarcy 2017].
Thomas Demarcy, Clair Vandersteen, Nicolas Guevara, Charles Raffaelli, Dan Gnansia, Nicholas Ayache and Hervé Delingette. *Automated Analysis of Human Cochlea Shape Variability From Segmented μCT images.* Computerized Medical Imaging and Graphics, volume 59, pages 1-12, 2017.

Abstract

The aim of this study is to define an automated and reproducible framework for cochlear anatomical analysis from high-resolution segmented images and to provide a comprehensive and objective shape variability study suitable for cochlear implant design and surgery planning. For the scala tympani, the scala vestibuli and the

whole cochlea, the variability of the arc lengths and the radial and longitudinal components of the lateral, central and modiolar paths are studied. The robustness of the automated cochlear coordinate system estimation is validated with synthetic and real data. Cochlear cross-sections are statistically analyzed using area, height and width measurements. The cross-section tilt angle is objectively measured and this data documents a significant feature for occurrence of surgical trauma.

3.1 Introduction

The human cochlea is a spiral-shaped structure located in the inner ear. Essential organ of audition, its largest diameter is about 8.5 mm at the basal turn and its height about 7 mm [Dimopoulos 1990]. Cochlear structures include three scalae (or ducts) with complex morphology and showing interindividual variability in size and shape.

Cochlear implant surgery is an effective treatment for severe to profound sensorineural hearing loss. The classical implantation procedure involves drilling a path through the mastoid in order to reach the tympanic cavity allowing the surgeon to open the scala tympani (one of the three cochlear ducts) and insert an electrode array. The functional outcomes for these subjects are linked to the potential insertion traumas [Aschendorff 2007, Finley 2009]. The insertion can lead to lesions of the modiolus, osseous spiral lamina or basilar membrane, resulting in degeneration of the ciliated [Adunka 2004] and spiral ganglion cells [Leake 1999]. Because the viability of these cells is correlated with neuronal survival [Nadol 1997] and speech expression [Xu 2012, Fayad 2006], it is crucial to avoid destroying them during surgery and to minimize injury to the previously cited anatomic structures with so-called atraumatic surgery techniques [James 2005].

An improved knowledge of the cochlea shape variability is not only essential for diagnosis of shape abnormality, atraumatic surgery planning and post-operative insertion assessment but it also provides a better anatomical understanding for the clinicians and can suggest electrode array design improvements for cochlear implant manufacturers.

Various methods have been used for the analysis of human cochlear anatomical structures. Methods using several type of plastic casts have been widely employed to evaluate the dimensions of cochlear anatomy [Zrunek 1980, Zrunek 1981, Dimopoulos 1990, Hatsushika 1990, Wysocki 1999, Erixon 2009, Rask-Andersen 2011], but they do not allow accurate preservation of the geometric relationships of fragile cochlear structures and artifact-free measurements. Microdissections enable measurements of basic external dimensions of the cochlea like the width of the cochlear base [Wright 2005], but as plastic casts, it cannot provide detailed three-dimensional reconstruction.

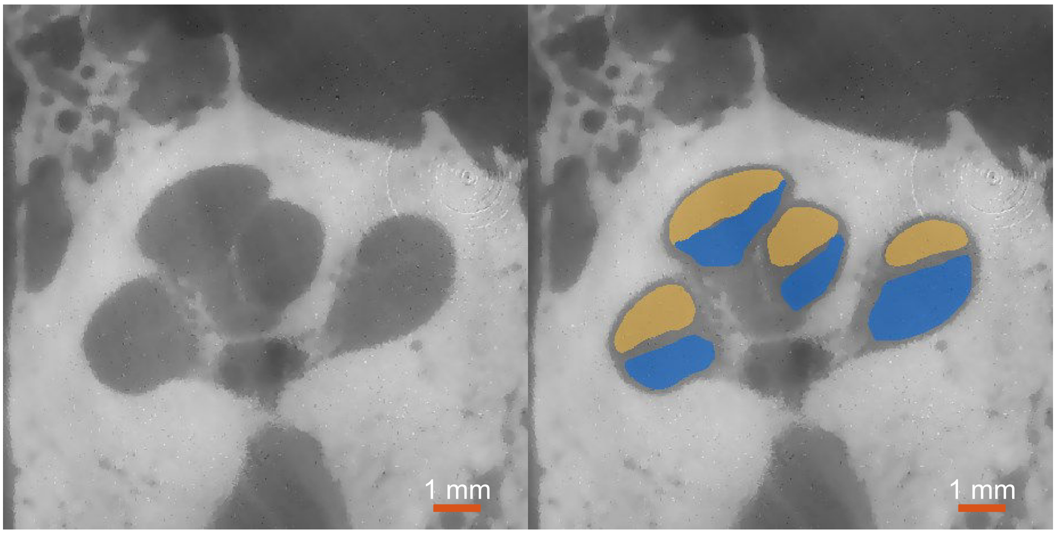
Histological sections offer the best image resolutions for the examination of fine cochlear structure [Hardy 1938, Gulya 1996, Biedron 2010, Makary 2010]. However just as plastic casts and microdissection it is a destructive method. Indeed, the method needs several chemical steps such as dehydration, decalcification, fixation and staining, which may induce some tissue distortion and shrinkage [Buytaert 2011, Rau 2013].

Medical imaging techniques provide nondestructive methods. CT and MRI are common preoperative procedures. For the purpose of cochlear anatomy analysis, CT provides more information and even though MRI could be a complementary imaging modality [Thorne 1999, Gibson 2012], most of the studies published to date have focused on CT [Skinner 2002, Escudé 2006, Baker 2008, van der Marel 2014]. However, conventional CT scan gives rather poor image resolution with respect to the cochlear dimensions.

As important cochlear anatomical structures such as the basilar membrane are not visible in CT imaging, analyses of the cochlea are commonly performed using μCT images [Verbist 2009, Poznyakovskiy 2011, Gunz 2012, Shin 2013, Avci 2014, Ceresa 2014]. In order to improve contrast, temporal bone samples can be fixed and stained before imaging [Poznyakovskiy 2011, Kjer 2015b]. Cochlear fluids are commonly removed [Postnov 2006, Poznyakovskiy 2011, Avci 2014, Ceresa 2014] but this operation requires to drill through the round and oval windows and to replace the perilymph with air by suction. The influence of this method on the geometry of the membranous structures is unknown [Rau 2013] and μCT scanning with fluid removal cannot be considered as nondestructive.

Previous analyses include measurements of the length of the cochlea and the number of cochlear turns [Hardy 1938, Kawano 1996, Skinner 2002, Escudé 2006, Stakhovskaya 2007, Erixon 2009, Gunz 2012, Shin 2013, van der Marel 2014], measurements of the heights, widths and sectional areas of the scala tympani [Zrunek 1980, Zrunek 1981, Hatsushika 1990, Gulya 1996, Wysocki 1999, Thorne 1999, Biedron 2010, Avci 2014], the scala vestibuli [Zrunek 1981, Gulya 1996, Wysocki 1999, Thorne 1999, Biedron 2010] and the cochlea [Erixon 2009, Shin 2013], measurements of the radial component of the cochlear centerlines (also called cochlear curvature) [Cohen 1996, Baker 2008, van der Marel 2014], and studies of the longitudinal component of the centerlines of the scala tympani [Avci 2014] and the cochlea itself [Verbist 2009].

In this chapter, the variability of the human cochlea was studied from nine μCT scans with a nondestructive preparation technique. The anatomical study provides a comprehensive set of measurements on the basis of new methods, which avoid inter-expert variability of manual measurements and can cope with the presence of noise and outliers. The central path analysis includes not only measurements of the scala tympani but also of the scala vestibuli and the whole cochlea, including lateral and modiolar paths. The centerlines were extracted with an image processing pipeline and expressed in an non-ambiguous cylindrical "cochlear" coordinate system, estimated with an original robust method. From the centerlines, cross-sections of the tympanic and vestibular ducts were acquired in a new non-ambiguous moving frame. The heights, widths, areas and their respective variability across the nine



(a) (b)

Figure 3.1: (a) μCT image acquired on patient #4. An anisotropic diffusion filter was applied to reduce image noise. (b) Segmentations of the scala tympani (blue) and the scala vestibuli (yellow). The round window is not present in this slice.

samples, were estimated along the centerline. The cross-section tilt angle was for the first time quantified.

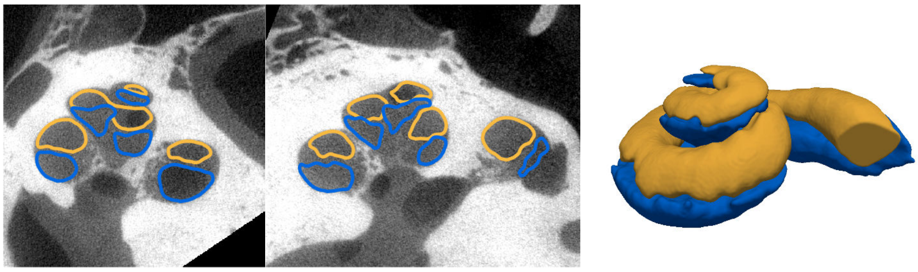
3.2 Materials and Methods

3.2.1 Nondestructive Preparation of Human Cochleae

Nine healthy human temporal bones (5 right and 4 left sides) were obtained from cadavers harvested within 24 hours after death. An otologic surgeon harvested the entire temporal bone keeping intact the mastoid, the tympanic cavity and cochlear fluids in order to prevent any damage of the membranous structures. Using an *in vivo* high-resolution μCT scanner (GE eXplore speCZT system), μCT images with isotropic voxel size of 24.79 μm were acquired in the laboratory Transporters, Imaging and Radiotherapy in Oncology (TIRO, UMR-E4320) in Nice under the supervision of Pr Thierry Pourcher. Unlike previous studies no destructive preparation techniques, such as cochlear fluids removal, were performed. The contrast level was, however, sufficient to perform manual segmentation (Fig 3.1).

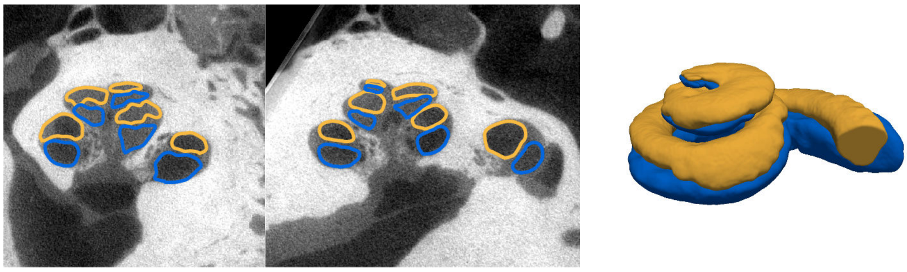
3.2.2 Interactive Segmentation

Automatic segmentation of the cochlear structures has been widely studied over the last decade [Xianfen 2005, Schuman 2010, Noble 2011, Reda 2013, Reda 2014a, Reda 2014b, Franz 2014]. Image segmentation is not the main focus of this study but a prerequisite for the shape analysis of delineated structures.



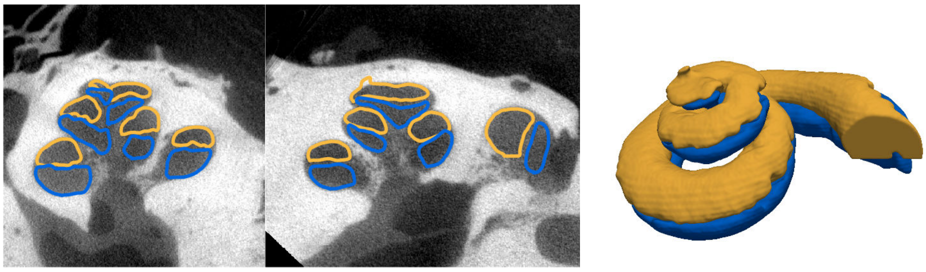
(a) (b) (c)

Figure 3.2: Patient 1
Manually segmented μCT slices in the yz- (a) and xz-planes (b) and volume (c) of the ST (blue) and SV(orange).



(a) (b) (c)

Figure 3.3: Patient 2



(a) (b) (c)

Figure 3.4: Patient 3

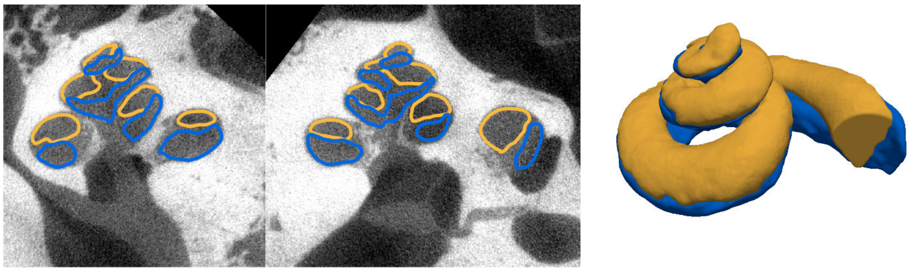

(a) (b) (c)

Figure 3.5: Patient 4

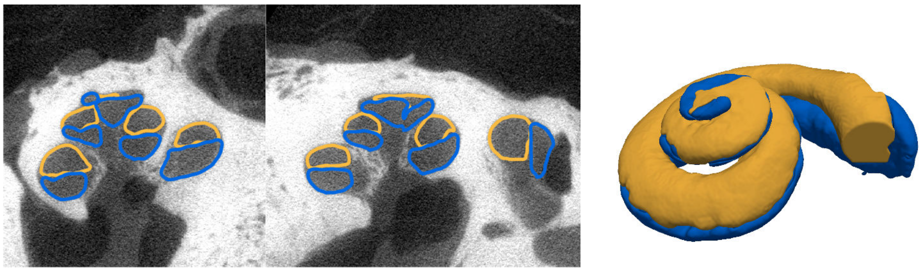

(a) (b) (c)

Figure 3.6: Patient 5.

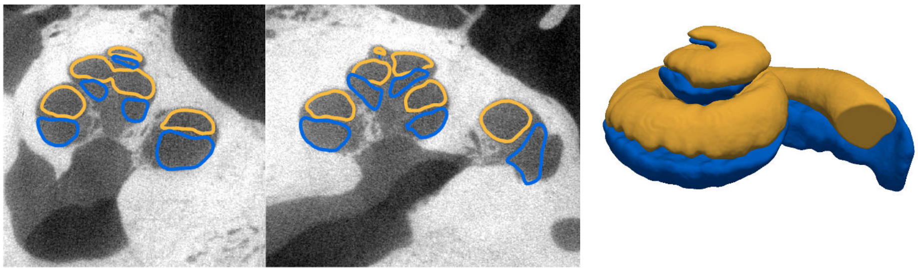
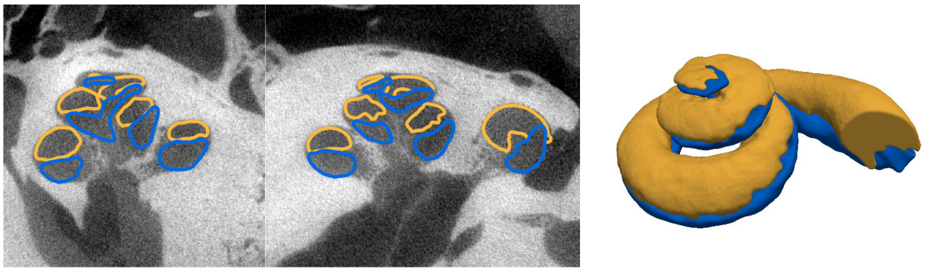
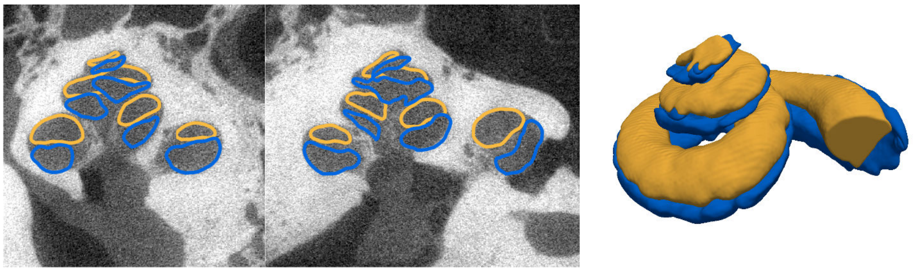

(a) (b) (c)

Figure 3.7: Patient 6



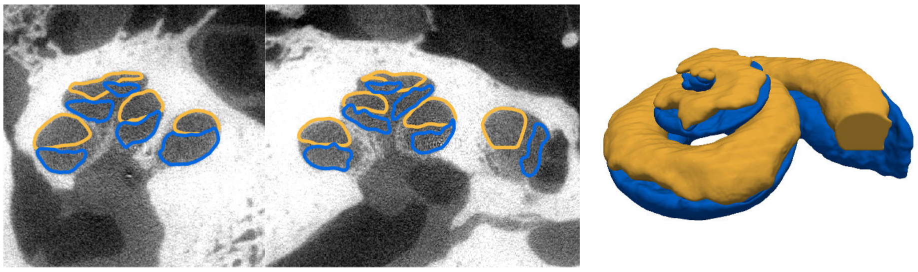
(a) (b) (c)

Figure 3.8: Patient 7



(a) (b) (c)

Figure 3.9: Patient 8



(a) (b) (c)

Figure 3.10: Patient 9

The μCT images were cropped around the inner ear region. In order to improve the quality of the images before segmentation an anisotropic diffusion filter [Perona 1990] was applied (Fig. 3.1a), which denoises the image while preserving edge contrast. Histogram equalization was then performed to enhance the contrast of the images. Each image was segmented using the interactive semi automated tool GeoS [Criminisi 2008] by a head and neck imaging expert (Fig. 3.1b). The image resolution is high enough to identify the basilar membrane but insufficient to identify the Reissner's membrane and distinguish the scala media (or cochlear duct) from the scala vestibuli. The first segmented area corresponds to the round window, the second to the scala tympani and the third to the scala vestibuli, the scala media and the semi-circular canals taken together. Here we focused on the anatomy of the cochlea, discarding the vestibule. Henceforth, unless otherwise specified, the label "scala vestibuli" refers to both vestibular and cochlear ducts (as frequently in other studies [Gulya 1996, Wysocki 1999, Yoo 2000a, Postnov 2006, Meshik 2010, Noble 2011, Braun 2012]). Fig. 3.2-3.10 present the nine segmented μCT images.

3.2.3 Automated Centerline Extraction

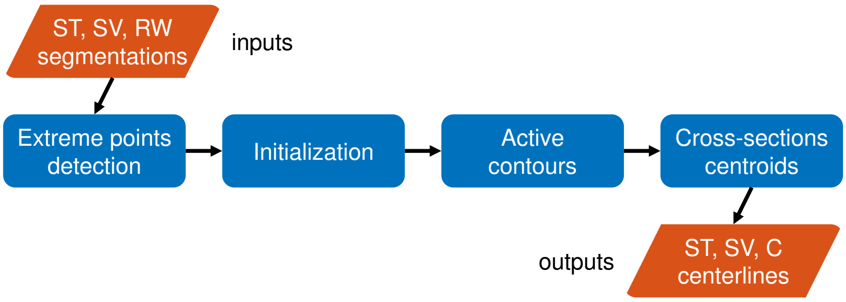


Figure 3.11: Flowchart of the automated centerline extraction. (Abbreviations: ST, scala tympani; SV, scala vestibuli; RW, round window; C, cochlea).

The challenge of cochlear centerline extraction has already been dealt with [Baker 2004, Verbist 2009, Poznyakovskiy 2011, Gunz 2012]. A flowchart of the method applied in this study and the results are respectively given on Fig. 3.11 and Fig. 3.12. We performed a standard active contours approach [Kass 1988] driven by the signed distance function which determines the distance of a given point from the shape boundary of the segmented anatomical structure. The method is followed by an adjustment using cross-section centroids to ensure that the centerline is properly centered. Right cochleae are similar to right-handed helices and symmetrically for left cochleae. Therefore to compare all centerlines, left cochleae are flipped in an arbitrary direction.

The cochlea is topologically equivalent to a cylinder, which means that the cochlear shape can be obtained by sweeping a planar surface along an axis. The planar surface has a varying shape along the axis, it is referred as the cross-sections.

Its axis is not straight and is set as the centerline (or central path), we defined as the centerline as the space curve inside the cochlea which maximizes its distance to the shape boundary.

With this definition, several image processing methods can be used for cochlear centerlines extraction. One of the most intuitive methods is skeletonization, a topological skeleton consists of a thin version of shape represented by one-dimensional branches. It is obtained by computing the locus of the centers of maximally inscribed spheres. This method was previously applied for the cochlear centerlines extraction problem [Verbist 2009] but is very sensitive to small variations of the shape, the resulting skeletons are highly noisy and correcting techniques, such as wave propagation, must be used. [Gunz 2012] claimed that skeleton are so inaccurate for cochlea that it is more appropriate to create manual skeletons. [Baker 2004] proposed an alternative technique based on principal flow, which is able to deal with non-circular cross-sections but also generates noisy outliers. [Poznyakovskiy 2011] developed a segmentation method that extracts the centerline as well as the cross-section contours, the algorithm iteratively computes the centerline from the mass center of the cross-section and using Kalman filter and segments the cross-section contours from the centerline normal planes using active contours until convergence. This method works well on μCT scans with cochlear fluids removal, providing high contrast between emptied structures filled with air and bony structures.

The problem of cochlear centerlines extraction is a specific problem of tubular structure extraction, which has already been considerably studied by the computer vision and medical imaging communities [Lesage 2009], in particular for vascular network segmentation or virtual endoscopy [Deschamps 2001].

Extreme Points Detection of the Scala Tympani The starting and target positions of the scala tympani were extracted automatically from the segmentations. The starting point was set as the voxel labeled “scala tympani” closest to the round window centroid. The detection of the target point is not straightforward. The most apical point is aimed. For this purpose, a roughly approximated modiolar axis (z-axis in the cochlear coordinate system by [Verbist 2010], see Fig. 3.13 in subsection 3.2.4) is needed. Since the cochlear diameter is about 8.5 mm at the basal turn and the cochlear height is about 7 mm, the axis of least inertia, defined as the smallest component using principal component analysis (PCA) of a voxel-based shape, is roughly aligned with the modiolar axis . So the voxel labeled “scala tympani”, whose projection on this axis is furthest from the starting point, corresponds to the most apical point. This point was chosen as the target point.

Active Contours A signed distance function which determines the distance of a given point from the shape boundary of the segmented anatomical structure can be defined from segmented surface structures. In this study, inner points have negative values whereas outer points have positive values.

Active contours can be used to solve the problem of tubular structure centerline extraction. The basic idea is to find a global minimum of an energy E,

$$E(\mathbf{x}) = \int_0^1 \left(P(\mathbf{x}(s)) + w_1 \left\| \frac{\partial \mathbf{x}}{\partial s} \right\|^2 + w_2 \left\| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right\|^2 \right) \mathrm{d}s \tag{3.1}$$

where $\mathbf{x}(s)$ is the three-dimensional centerline parameterized by $s \in [0,1]$ the normalized curvilinear coordinate of the curve. In our case $\mathbf{x}(0)$ and $\mathbf{x}(1)$ are set to what we called respectively the starting and target points (extreme points). The first term P is chosen as the signed distance function previously introduced. The two other terms are regularization terms called internal energy. The tension and the bending of the centerline can be controlled respectively by w_1 and w_2.

The energy is similar to the one introduced in the seminal work of [Kass 1988] coined as the snake method. The minimization procedure is based on a finite difference approximations (see appendix A) of the Euler-Lagrange differential equation,

$$\frac{\partial \mathbf{x}}{\partial t} = -\nabla P(\mathbf{x}) + w_1 \frac{\partial^2 \mathbf{x}}{\partial s^2} - w_2 \frac{\partial^4 \mathbf{x}}{\partial s^4}. \tag{3.2}$$

The centerline $\mathbf{x}$ evolves with respect to an artificial time t, and is initialized with contour $\mathbf{x}_0$, which must be close enough to avoid falling in an undesired local minimum. The iterations end when the change in the value of the energy E is smaller than a threshold value ε, i.e. $|E(\mathbf{x}(t+\Delta t)) - E(\mathbf{x}(t))| < \varepsilon$, the number of iterations is also limited. It can be noted that snake methods have been already used for segmenting cochlear images by [Yoo 2001].

Initialization The initialization is based on the [Dijkstra 1959]'s algorithm after defining a graph of voxels in 6-adjacency. Each voxel was considered as a node and the weight between neighboring voxels was set as the mean signed distance of the two adjacent voxels. An offset is added to keep positive weights. Even if the results suffer from approximation errors (it uses the L_1 norm) and tends to follow the modiolar wall (opposite the lateral wall) of the scala tympani, it leads to a good initialization of the helico-spiral centerline.

Cross-section Centroids The centerline obtained with the snake algorithm allows us to define tangent vectors along the centerline. To ensure that the centerline is properly centered, a slightly corrected centerline is computed from the locus of the cross-sections centroids computed by slicing the segmentations in planes whose normals are aligned with the centerline tangent vectors.

Initialization of the Scala Vestibuli The centerline of the scala vestibuli was also extracted with the same approach but with a different initialization step, only the initialization step differs from the pipeline used for the scala tympani. An initialized vestibular centerline $\mathbf{x}_0^{SV}$ can be estimated directly from the tympanic centerline $\mathbf{x}^{ST}$.

$$\mathbf{x}_0^{SV}(s) = \mathbf{x}^{ST}(s) + f(s)\hat{\mathbf{z}} \tag{3.3}$$

where $\hat{\mathbf{z}}$ is the unit vector of the modiolar axis (z-axis in the basal cochlear coordinate system, see subsection 3.2.4), and f is an affine function of the normalized parameter s, expressed in millimeters.

$$f(s) = 1 - 0.8s \tag{3.4}$$

based on the measurement of the heights of the tympanic and vestibular ducts found in [Wysocki 1999]. As for the scala tympani, even if it suffers from approximation errors, this initialization method is good enough to play the role of $\mathbf{x}_0$ in the minimization procedure.

Centerlines Extraction of the Cochlea The centerline of the union of the tympanic and vestibular ducts (called here the whole cochlea) x^C was finally computed from the scalae tympani and vestibuli as the mean centerline weighted by the cross-sectional areas (respectively A^{ST} and A^{SV}) computed again by slicing the segmentations in planes whose normal vectors are the centerlines tangents.

$$\mathbf{x}^C(s) = \frac{A^{ST}(s)\mathbf{x}^{ST}(s) + A^{SV}(s)\mathbf{x}^{SV}(s)}{A^{ST}(s) + A^{SV}(s)}. \tag{3.5}$$

3.2.4 Robust Modiolar Axis Estimation

The cochlear coordinate system (x, y, z) by [Verbist 2010] provide a standard and unambiguous framework for anatomic studies of the cochlea. It can be defined independently from a modality-centered anatomical coordinate system. The modiolus is a cone-shape bony structure around which the first two turns of the cochlea are wrapped. The z-axis is chosen as the modiolar-axis. As [Verbist 2010] point out, there exist two main alternatives for the origin: the helicotrema (at the apex) with the z-axis oriented from the apex to the base of the modiolus, or the base with the reversed orientation. To remove any ambiguity, we call the former option “apical cochlear coordinate system” (ACCS) [Verbist 2010] and the latter “basal cochlear coordinate system” (BCCS) (see Fig. 3.13). In both coordinate systems, the xz-plane passes through the center of the round window, x values being positive at the level of the round window. The y-axis is set in order to have $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ being an orthonormal basis. In this study we chose the BCCS. A flowchart of the methods used for anatomical analysis of the cochlea is given on Fig. 3.14.

One of the major difficulties is to define the modiolar axis. Most of previous studies define this axis by manually adjusted multiplanar reconstruction [Baker 2008, van der Marel 2014] or simply by setting manually two points [Poznyakovskiy 2008, Verbist 2009, Kjer 2015a]. [Escudé 2006] proposed maximization of dark pixel area using minimum intensity projection. [Yoo 2000b] determined the modiolar axis using three nonlinear least square minimization based algorithms.

Misalignment of the modiolar axis may greatly impact the measurements and can be a cause of non reproducibility of the results.

Helico-spiral Axis Inference The challenge is to estimate the unknown modiolar axis $\hat{\mathbf{z}}$ from the cochlear centerline $\mathbf{x}$ expressed in another coordinate system. We propose a method based on the intrinsic geometrical properties of the centerlines, more specifically based on the locus of the centers of curvature of the centerline.

The cochlear centerline is most frequently described as a logarithmic spiral [Cohen 1996, Xu 2000, Yoo 2000a, Yoo 2000b, Baker 2008]. In a cylindrical coordinate system (where $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{z}})$ are the local radial, circumferential and longitudinal unit vectors) an ideal logarithmic helico-spiral can be written as:

$$\mathbf{x} = ae^{-b\theta}\hat{\mathbf{r}} + c\theta\hat{\mathbf{z}} \tag{3.6}$$

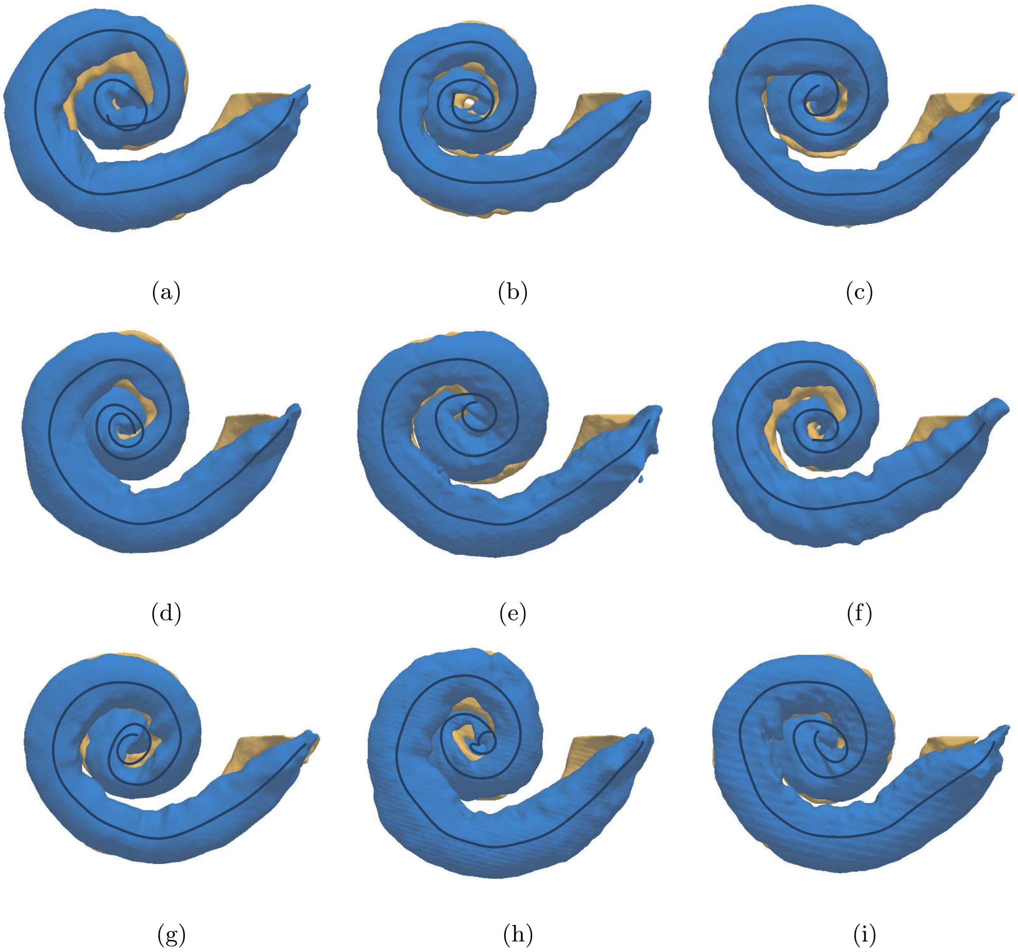
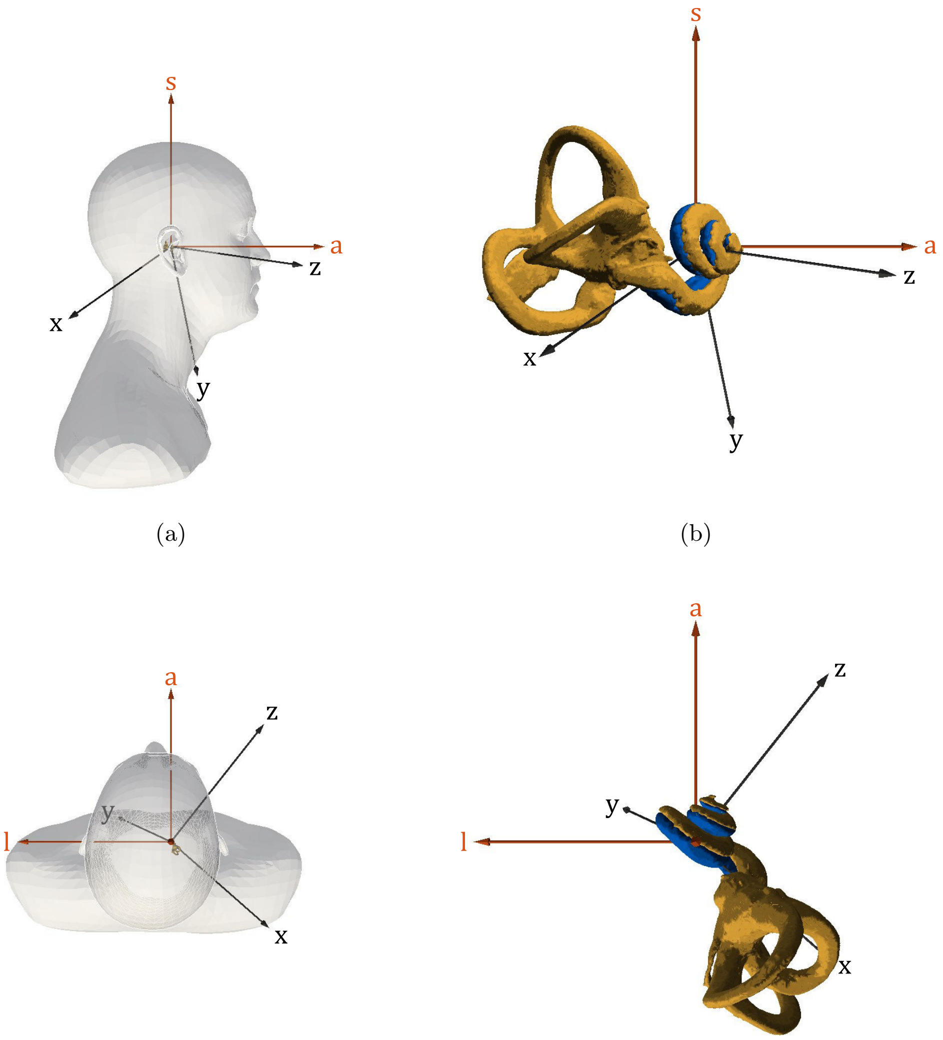

(a) (b) (c)
(d) (e) (f)
(g) (h) (i)

Figure 3.12: Nine ST centerlines overlaid over their corresponding three-dimensional segmented cochleae.



(a) (b)

(c) (d)

Figure 3.13: Comparison of modality-centered anatomical coordinate system (l, a, s) and BCCS (x, y, z). (a) Left sagittal view from posterior to anterior (a-axis) and from inferior to superior (s-axis) and (b) its corresponding close-up view. (c) Superior transverse view from left (l-axis) to right and from posterior to anterior (a-axis) and (d) its corresponding close-up view. The segmented scala tympani (blue) and the bony labyrinth (yellow) of the patient #1 (right cochlea) are represented. The z-axis oriented from the base to the apex of the modiolus, the x-axis passes through the center of the round window. There is no simple relationship (such as coplanarity) between axes and planes from the two coordinate systems. The angle between a and z-axes is know to be around 45° [Xu 2000], which is not exactly the case here.

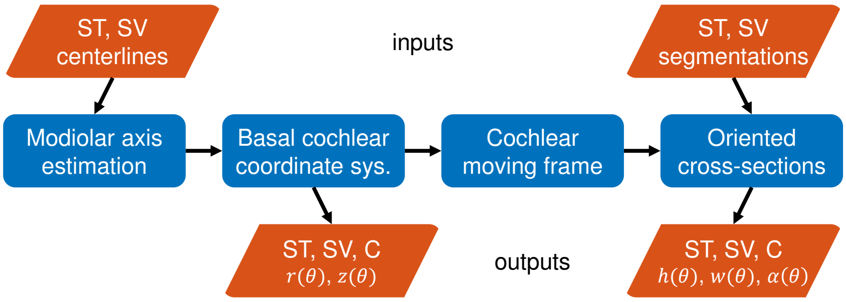


Figure 3.14: Flowchart of the shape parameterization. The centerline is derived in its radial $r(\theta)$ and longitudinal $z(\theta)$ components in a cylindrical coordinate system called BCCS. The cross-sections are expressed in a cochlear moving frame allowing for measurements of their heights $h(\theta)$, widths $w(\theta)$ and tilt angles $\alpha(\theta)$. (Abbreviations: ST, scala tympani; SV, scala vestibuli; C, cochlea).

Table 3.1: Logarithmic helico-spiral constants for cochlear model. The values for the constants are given by [Clark 2011].

a (mm)	**b** (rad^{-1})	**c** (mm·rad^{-1})	$\boldsymbol{\theta_0}$ (rad)	$\boldsymbol{\theta_1}$ (rad)
3.762	0.07546	0.1751	0	15.71

where a, b and c are positive constants. This cochlear model, completely parameterized by $\theta \in [\theta_0, \theta_1]$, is close to the one given by [Clark 2011], discarding the piecewise definition of the radial function used to model the basal turn.

The radial and longitudinal components of $\mathbf{x}$ are:

$$r(\theta) = ae^{-b\theta} \tag{3.7}$$

$$z(\theta) = c\theta. \tag{3.8}$$

Let us consider an infinitesimal part of the curve in the cylindrical coordinate system

$$\mathrm{d}s^2 = \mathrm{d}r^2(\theta) + r^2(\theta)\mathrm{d}\theta^2 + \mathrm{d}z^2(\theta) \tag{3.9}$$

$$\mathrm{d}s^2 = b^2r^2(\theta)\mathrm{d}\theta^2 + r^2(\theta)\mathrm{d}\theta^2 + c^2\mathrm{d}\theta^2 \tag{3.10}$$

$$\mathrm{d}s = \sqrt{(b^2+1)r^2(\theta) + c^2}\mathrm{d}\theta. \tag{3.11}$$

We assume that $c^2 \ll r(\theta)^2 \ \forall\theta \in [\theta_0, \theta_1]$ since, in practice, using the values in table 3.1,

$$2.2\times 10^{-3}\approx \frac{c^2}{r(\theta_0)^2}\leq \frac{c^2}{r(\theta)^2}\leq \frac{c^2}{r(\theta_1)^2}\approx 2.3\times 10^{-2} \tag{3.12}$$

.

We define differential arc length $d\tilde{s}$ as follows,

$$\mathrm{d}\tilde{s}=\sqrt{b^2+1}r(\theta)\mathrm{d}\theta \tag{3.13}$$

Let us define the function γ and ε,

$$\begin{aligned}\gamma=\frac{\mathrm{d}\tilde{s}}{\mathrm{d}s}&=\frac{\sqrt{b^2+1}r(\theta)}{\sqrt{(b^2+1)r^2(\theta)+c^2}}\\&=1-\frac{c^2}{2(b^2+1)r^2(\theta)}+o\left(\frac{c^2}{r^2(\theta)}\right)\approx 1\end{aligned} \tag{3.14}$$

$$\begin{aligned}\varepsilon=\left|\frac{\mathrm{d}^2\tilde{s}}{\mathrm{d}s^2}\right|&=\left|\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}\tilde{s}}{\mathrm{d}s}\right)\right|\\&=\frac{\sqrt{b^2+1}br(\theta)c^2}{\left((b^2+1)^2r^2(\theta)+c^2\right)^2}\propto\frac{c^2}{r^3(\theta)}\ll 1.\end{aligned} \tag{3.15}$$

The tangent unit vector $\hat{\mathbf{t}}$ is defined as:

$$\hat{\mathbf{t}}=\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}=\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tilde{s}}\frac{\mathrm{d}\tilde{s}}{\mathrm{d}s}=\frac{-br(\theta)\hat{\mathbf{r}}+r(\theta)\hat{\boldsymbol{\theta}}+c\hat{\mathbf{z}}}{\sqrt{b^2+1}r(\theta)}\gamma \tag{3.16}$$

$$\begin{aligned}\kappa\hat{\mathbf{n}}=\frac{\mathrm{d}\hat{\mathbf{t}}}{\mathrm{d}s}&=\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}s^2}=\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tilde{s}}\frac{\mathrm{d}\tilde{s}}{\mathrm{d}s}\right)\\&=\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}\tilde{s}^2}\left(\frac{\mathrm{d}\tilde{s}}{\mathrm{d}s}\right)^2+\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tilde{s}}\frac{\mathrm{d}^2\tilde{s}}{\mathrm{d}\tilde{s}^2}=\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}\tilde{s}^2}\gamma^2-\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\tilde{s}}\varepsilon\end{aligned} \tag{3.17}$$

where κ is the curvature and $\hat{\mathbf{n}}$ is the normal unit vector. Using the approximations of γ and ε,

$$\kappa\hat{\mathbf{n}}\approx\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}\tilde{s}^2}=\frac{-r(\theta)\hat{\mathbf{r}}-br(\theta)\hat{\boldsymbol{\theta}}+bc\hat{\mathbf{z}}}{(b^2+1)r^2(\theta)} \tag{3.18}$$

$$\begin{aligned}\kappa^2=\|\kappa\hat{\mathbf{n}}\|^2&\approx\frac{(b^2+1)r^2(\theta)+b^2c^2}{\left((b^2+1)^2r^2(\theta)\right)^2}\\&\approx\frac{1}{(b^2+1)r^2(\theta)}+\frac{b^2c^2}{\left((b^2+1)^2r^2(\theta)\right)^2}\end{aligned} \tag{3.19}$$

where the second term in the right-hand side equation ($\propto c^2/r^4(\theta)$) can be neglected, the curvature is then,

$$\kappa\approx\frac{1}{\sqrt{b^2+1}r(\theta)}. \tag{3.20}$$

We can now approximate the following terms,

$$\frac{\hat{\mathbf{t}}}{\kappa} \approx -br(\theta)\hat{\mathbf{r}} + r(\theta)\hat{\boldsymbol{\theta}} + c\hat{\mathbf{z}} \tag{3.21}$$

$$\frac{\hat{\mathbf{n}}}{\kappa} \approx -r(\theta)\hat{\mathbf{r}} - br(\theta)\hat{\boldsymbol{\theta}} + bc\hat{\mathbf{z}}. \tag{3.22}$$

The evolute $\mathbf{e}$ of curve $\mathbf{x}$ is the locus of all its centers of curvature, the evolute of an logarithmic helico-spiral has the parametric equation:

$$\mathbf{e} = \mathbf{x} + \frac{\hat{\mathbf{n}}}{\kappa} \approx -br(\theta)\hat{\boldsymbol{\theta}} + c \cdot (\theta + b)\hat{\mathbf{z}}. \tag{3.23}$$

We propose to build a corrected evolute $\mathbf{e_c}$ which is a sole function of the modiolar axis enabling to remove the circumferential dependency,

$$\begin{aligned} \mathbf{e_c} &= \mathbf{x} + \frac{b}{1+b^2}\frac{\hat{\mathbf{t}}}{\kappa} + \frac{1}{1+b^2}\frac{\hat{\mathbf{n}}}{\kappa} \\ &\approx c \cdot \left(\theta + \frac{2b}{1+b^2}\right)\hat{\mathbf{z}}. \end{aligned} \tag{3.24}$$

The corrected evolute estimates the modiolar axis, aligned with $\hat{\mathbf{z}}$, using only the intrinsic geometrical properties of the cochlear centerline, the curve $\mathbf{x}$ i.e. its position, its local tangent and normal unit vector $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$, its local curvature κ, and a constant parameter b. The result is independent from the modality-centered coordinate system (defined in the axial, coronal and sagittal planes).

The parameter b is a widely studied cochlear measurement, here it was set to the optimum value found by [Cohen 1996] among 30 patients ($b = 0.075$ rad^{-1}).

Robust Principal Component Analysis (PCA) The modiolar axis is thus computed from the cochlear centerline as the locus of the corrected evolute. In practice, discretization errors (see A for more details) and model errors (the centerline may not be a perfect logarithmic spiral) cause the discrete corrected evolute points to lie close to the modiolar axis but not exactly on it. Thus it is necessary to estimate the line which best approximates the discrete corrected evolute in a robust manner in order to cope with the presence of noise and outliers. This is equivalent to performing robust PCA as we are looking for the first principle component of the cloud of points. Robust PCA is performed using an extension of the expectation maximization (EM) algorithm called ECME [Liu 1995] fitting Student's t-distributions instead of Gaussian distributions. The Student's t-distribution can be interpreted as an infinite weighted sum of normal distributions with the same mean and variance determined by a gamma distribution. In the Maximization step the mean and the variance are updated by computing a weighted sum of the data.

In the cylindrical coordinate system (r, θ, z) (Fig. 3.15) constructed from the BCCS (x, y, z), the radial component $r(\theta)$ and the longitudinal component $z(\theta)$ of the centerlines were extracted for the tympanic and vestibular ducts and the cochlea.

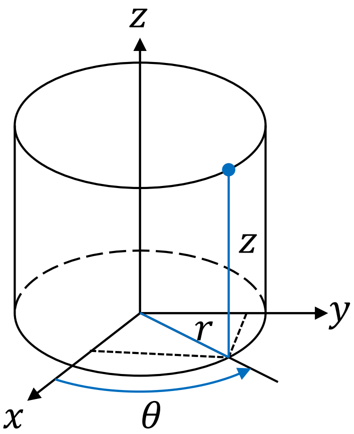


Figure 3.15: Cylindrical coordinates.

3.2.5 Cochlear Moving Frame

In order to evaluate the internal dimensions of the cochlea, cross-section measurements were performed in radial planes (a.k.a mid-modiolar cross-sections) [Zrunek 1980, Zrunek 1981, Erixon 2009], in parallel planes such as histological sections [Biedron 2010] and in planes normal to the centerline [Poznyakovskiy 2011, Avci 2014]. This last method does not induce measurement errors (e.g. the section of a circular cylinder performed in plane not normal to the axis produces an ellipse).

Planes normal to the centerline can be defined in the Frenet-Serret moving frame (where $(\hat{\mathbf{t}}, \hat{\mathbf{n}}, \hat{\mathbf{b}})$ are the basis vectors). This coordinate system is traditionally used for centerline-based cochlear models [Viergever 1978], however, its moving frame presents the significant drawback that the binormal unit vector may be ill-defined at inflection points (see Fig. 3.17a). Instead, to overcome this drawback, let us define the vectors:

$$\hat{\mathbf{u}} = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{t}}}{\|\hat{\mathbf{z}} \times \hat{\mathbf{t}}\|} \text{ (a)} \qquad\qquad \hat{\mathbf{v}} = \hat{\mathbf{t}} \times \hat{\mathbf{u}} \text{ (b)} \tag{3.25}$$

where $\hat{\mathbf{z}}$ is the unit vector of modiolar axis, $\hat{\mathbf{t}}$ is the unit vector tangent to the centerline and the symbol × denotes the cross product. $(\hat{\mathbf{t}}, \hat{\mathbf{u}}, \hat{\mathbf{v}})$ defines a moving frame which is always well defined because $\hat{\mathbf{t}}$ (and $\hat{\mathbf{z}}$) is never degenerated (see Fig. 3.17b). The cross-sectional measurements can then be performed locally in the uv-planes.

Oriented Cross-sections Along the cochlear centerline the cross-sections of the scalae tympani and vestibuli were acquired in the normal uv-planes. The cross-sections may have problems with accuracy at the two extremities of the centerline. Near the round window the cross-sections of the label “scala vestibuli” can leak into the vestibule as well as the cross-sections may not be perfectly defined at the apical turn since both scalae are supposed to merge at the apex. Potential errors are detected with the change in the cross-sectional areas. An abrupt increase of

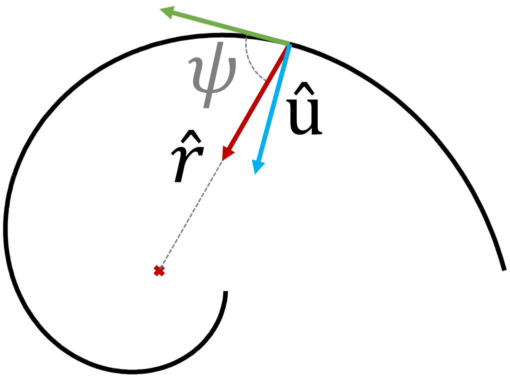


Figure 3.16: $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ can be linked to the radial $r(\theta)$ and longitudinal component $z(\theta)$ of the central paths. $\hat{\mathbf{u}}$ is defined in the xy-plane and $\angle\hat{\mathbf{u}}\hat{\mathbf{r}}$ the angle between $\hat{\mathbf{u}}$ and $\hat{\mathbf{r}}$, the radial unit vector, is the complementary to the logarithmic spiral pitch $\psi = \cot^{-1}(b)$ where b is constant for a logarithmic spiral (a.k.a. equiangular spiral) (as in Eq. 3.6). $\angle\hat{\mathbf{v}}\hat{\mathbf{z}}$ the angle between $\hat{\mathbf{v}}$ and $\hat{\mathbf{z}}$, is the longitudinal growth rate, $\angle\hat{\mathbf{v}}\hat{\mathbf{z}}(\theta) = \arctan(\mathrm{d}z/\mathrm{d}\theta)$.

one of the areas is indeed related to a leakage. The improper cross-sections were constrained to vary smoothly, by imposing incorrect cross-sections to be inside the morphological dilatation of their correct neighbor.

Cross-sections are fitted independently for each normal plane with ellipses using classical principal component analysis. The major axes define the widths of the ducts w_{ST} and w_{SV} and the minor axes define the heights of the ducts h_{ST} and h_{SV}. The ellipses also provide an orientation. Let us call α_{ST} (respectively α_{SV}) the angle between $\hat{\mathbf{u}}$ (i.e. the xy-plane) and the major axis of the scala tympani (respectively the scala vestibuli). The cochlear tilt angle α is defined as the mean of α_{ST} and α_{SV} (see Fig. 3.31b). Positive values of the tilt angle mean that locally the scala tympani is more lateral than the scala vestibuli and conversely for negative values (to be consistent with the profile tilt angles in [Clark 2011]).

x^{ST}_{lat} (respectively x^{ST}_{mod}) that represents the set of points defining the paths along the tympanic lateral (respectively modiolar) walls can be computed from the tympanic centerline x^{ST} using the cross-section measurements.

$$x^{ST}_{lat}(s) = x^{ST}(s) - \frac{w_{ST}}{2}(\cos(\alpha_{ST})\hat{\mathbf{u}} - \sin(\alpha_{ST})\hat{\mathbf{v}}) \tag{3.26a}$$

$$x^{ST}_{mod}(s) = x^{ST}(s) + \frac{w_{ST}}{2}(\cos(\alpha_{ST})\hat{\mathbf{u}} - \sin(\alpha_{ST})\hat{\mathbf{v}}). \tag{3.26b}$$

These values can be computed similarly on the scala vestibuli.

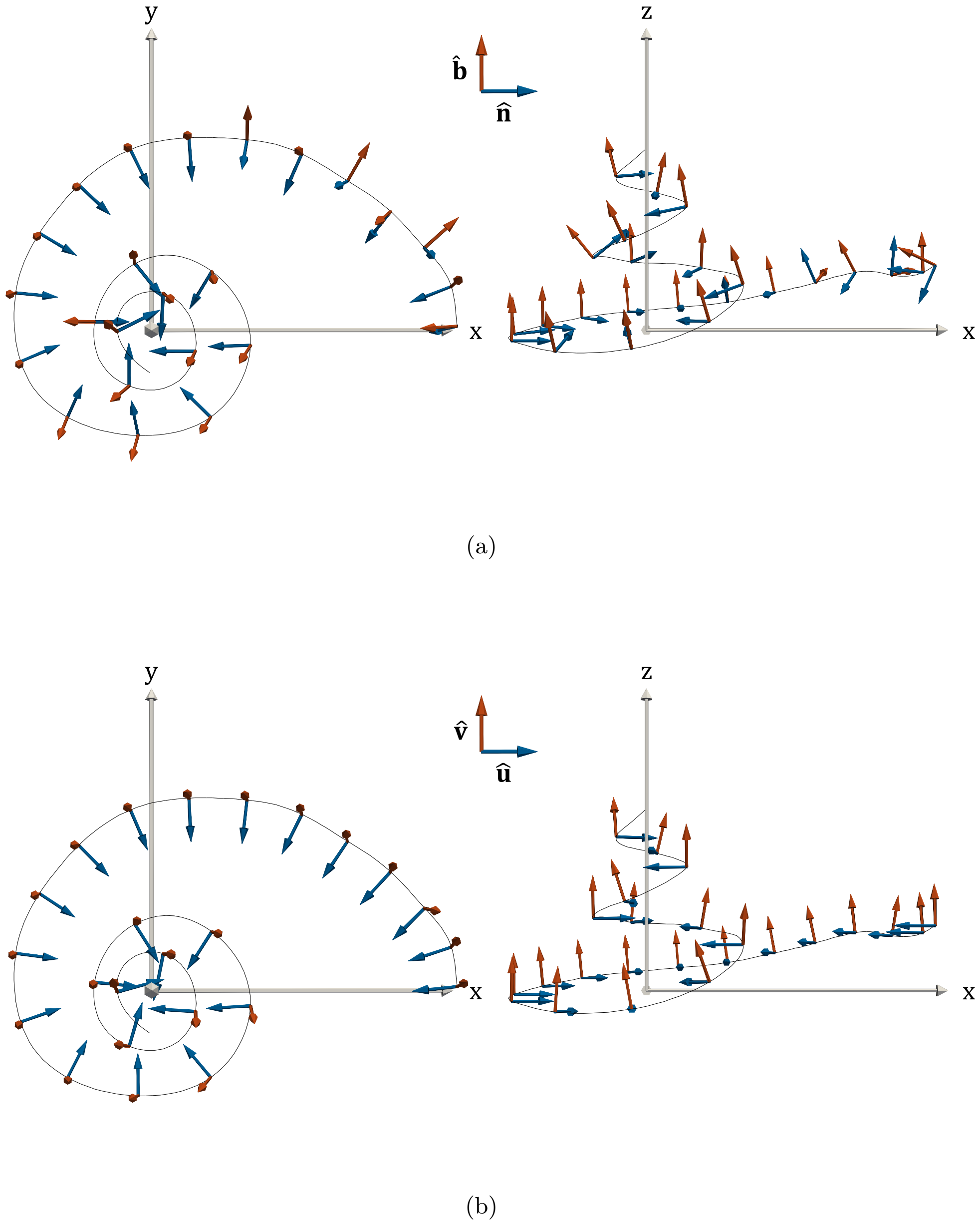


Figure 3.17: (a) Frenet-Serret frame. The start of the first basal turn is known to be not perfectly logarithmic [Cohen 1996] and can present inflection points. Here the normal unit vector $\hat{\mathbf{n}}$ (blue) and the binormal unit vector $\hat{\mathbf{b}}$ (red) vary singularly, as we can see notably within first five represented moving frames. (b) New cochlear moving frame. The vectors $\hat{\mathbf{u}}$ (blue) and $\hat{\mathbf{v}}$ (red) vary smoothly along the centerline.

3.3 Results

3.3.1 Centerlines Measurements

Table 3.2 shows the variability of basic measurements determined from the individual cochlear centerlines. The number of cochlear turn is statistically equal to 2.57 ± 0.28 (mean and ± 1 standard deviation), the length of the cochlear centerline is equal to 24.4 mm ± 1.0. No significant correlation was found between these two measurements ($R^2 = 0.39$). Most of the variability occurs at the apical turn. Fig. 3.18 illustrates the arc length along the scala tympani following the central path (i.e. centerline), the lateral or modiolar walls. A projected representation of the automatically extracted lateral, central and modiolar paths of a scala tympani is shown in Fig. 3.19.

Fig. 3.20 presents the individual radial component of the tympanic centerlines and Fig. 3.21, the mean radial component of the two scalae and their union. In the first part of the basal turn the scala tympani is sightly more lateral than the scala vestibuli. The radial component after 180° are roughly the same following a logarithmic form. The standard deviation of the radial component of the cochlear centerline ($r(\theta)$) averaged over the patients and the angular coordinate θ (between 0 and 900°) is equal to 0.11 mm.

Fig. 3.22 demonstrates the individual longitudinal component of the tympanic centerlines and Fig. 3.23, the mean longitudinal component of the two scalae and their union. The standard deviation of the cochlear longitudinal component of the cochlear centerline ($z(\theta)$) averaged over the patients and the angular coordinate θ (between 0 and 900°) is equal to 0.46 mm, which is more than four times the value obtained for the radial component.

3.3.2 Modiolar Axis Estimation

Automated vs. Manual Estimation In order to compare the current method, four cochlear anatomy experts estimated the modiolar axis by manually setting two points, several times for each temporal bones.

Let us consider the modiolar axis as a line $L = \{\mathbf{p} + s\hat{\mathbf{z}} | s \in \mathbb{R}\}$, where $\hat{\mathbf{z}}$ is a unit vector. We evaluate respectively the angle of rotation and the distance between two axes L_i and L_j using,

$$d_r(L_i, L_j) = \left|\sin^{-1}(\|\hat{\mathbf{z}}_i \times \hat{\mathbf{z}}_j\|)\right| \in [0, \pi] \tag{3.27a}$$

$$d_t(L_i, L_j) = \left|\frac{\hat{\mathbf{z}}_i \times \hat{\mathbf{z}}_j}{\|\hat{\mathbf{z}}_i \times \hat{\mathbf{z}}_j\|} \cdot (\mathbf{p}_j - \mathbf{p}_i)\right| \geq 0 \tag{3.27b}$$

The inter- and intra-expert standard deviation equal respectively to 10.0° and 8.7° for the angle of rotation and 0.14 mm and 0.12 mm for the distance. The mean angle and distance between the automatically estimated and manually defined modiolar axis are respectively 10.2° and 0.28 mm. Fig. 3.25 shows an example of automatic and manual axis estimation.

patient	cochlear turns	cochlear length (mm)	ST length (mm)	SV length (mm)
1	2.32	23.6	25.3	22.9
2	2.57	23.5	25.3	22.9
3	2.74	26.1	26.7	27.0
4	3.08	25.5	27.8	25.6
5	2.14	24.5	25.8	24.5
6	2.32	23.0	24.3	22.6
7	2.74	24.6	26.0	24.6
8	2.54	23.9	26.5	23.2
9	2.67	24.6	26.8	24.2

Table 3.2: Measurements determined from the individual cochlear centerlines

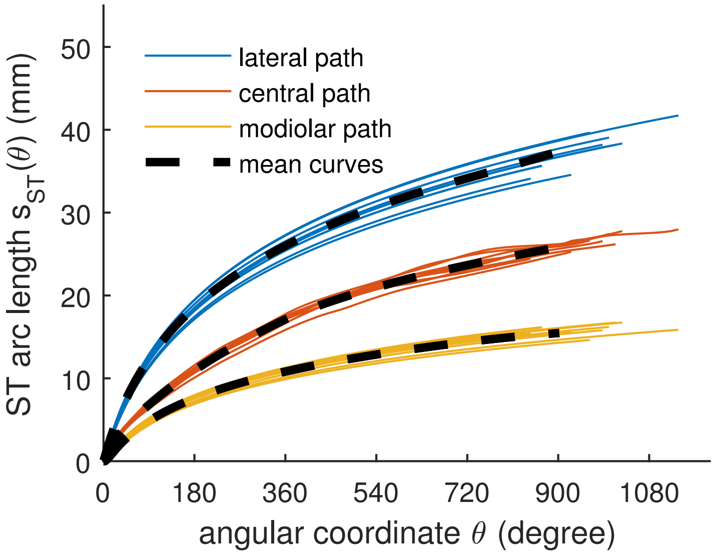


Figure 3.18: Individual and mean arc length of the lateral, central and modiolar paths of the scala tympani as a function of the angular coordinate.

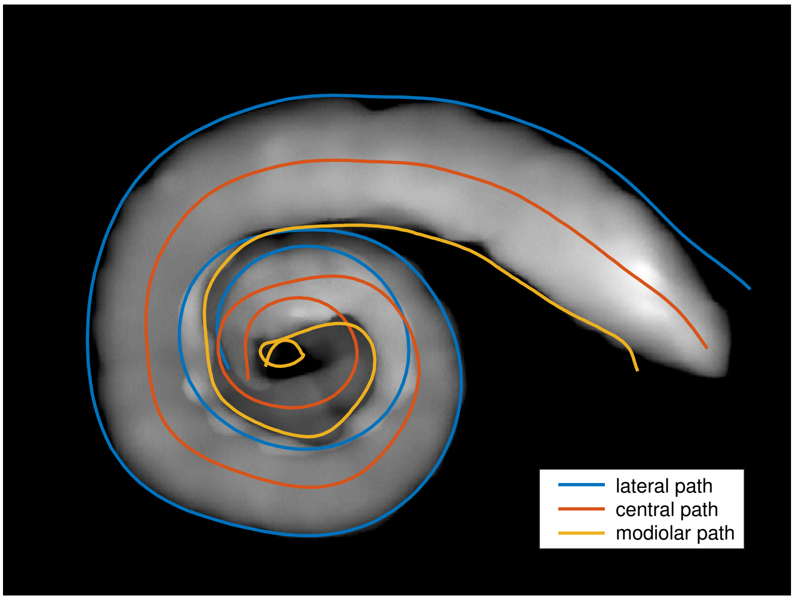


Figure 3.19: Lateral, central and modiolar paths of a scala tympani (patient #2) projected into the xy-plane. The mean intensity projection of the segmentation is in the background.

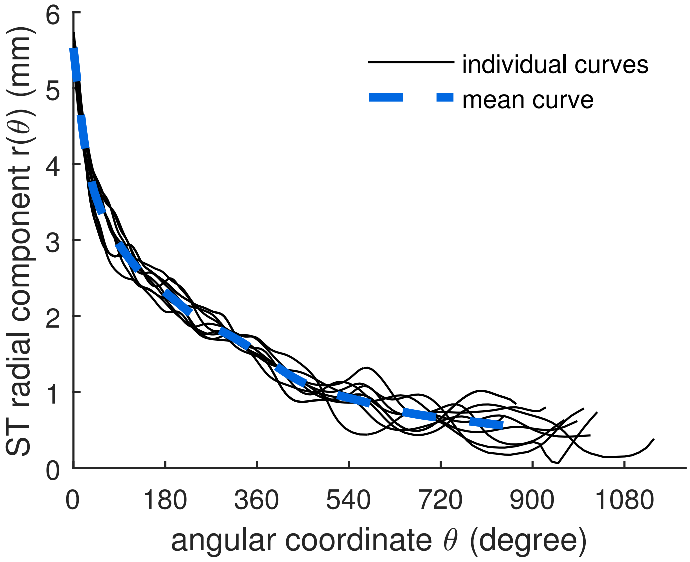


Figure 3.20: Individual and mean radial component of the tympanic centerline (ST).

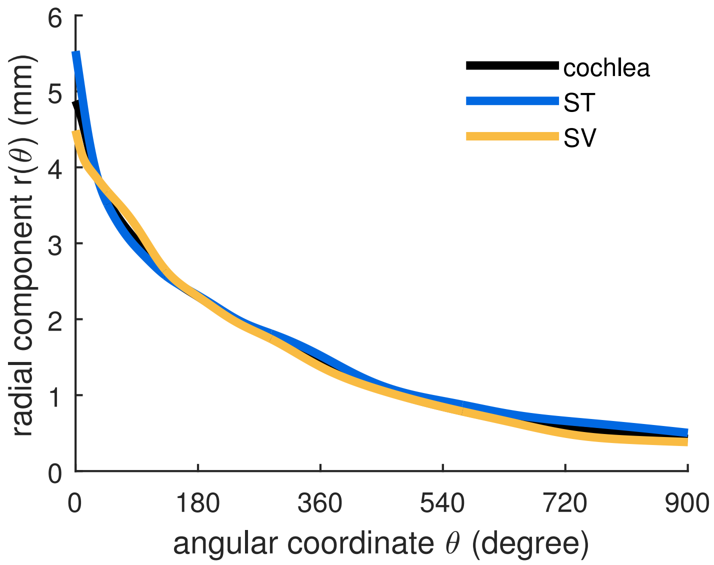


Figure 3.21: Mean radial component of the tympanic (ST) and vestibular (SV) central paths as a function of the angular coordinate.

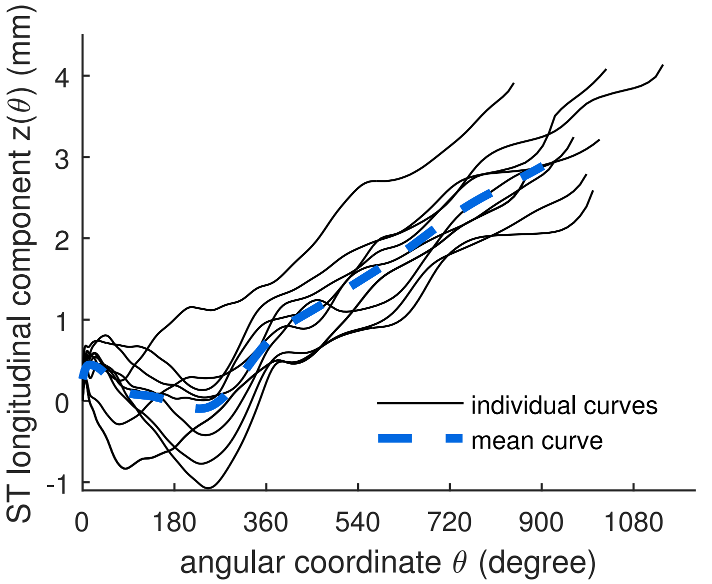


Figure 3.22: Individual and mean longitudinal component of the tympanic centerline (ST). The origin corresponds to the centers of the round windows.

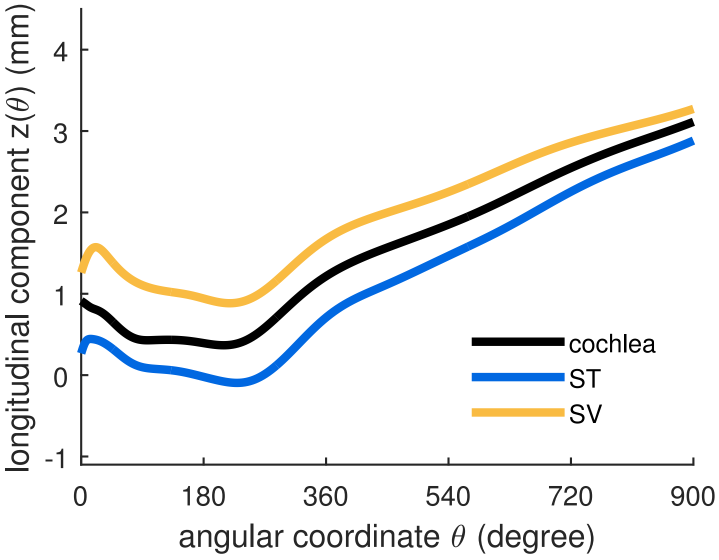


Figure 3.23: Mean longitudinal component of the tympanic (ST) and vestibular (SV) central paths as a function of the angular coordinate.

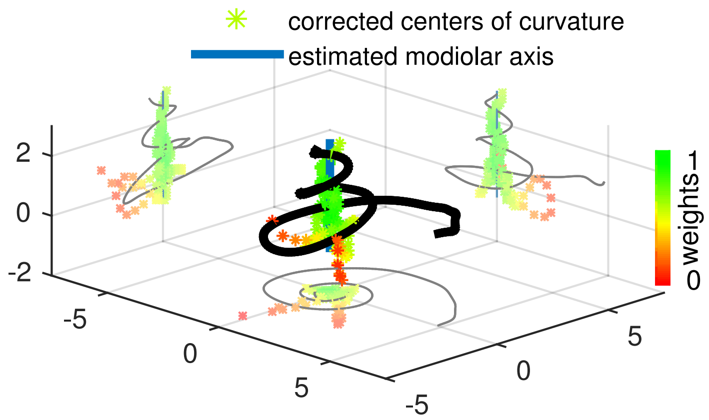


Figure 3.24: Modiolar axis estimation from a real cochlear centerline (patient #3). A robust PCA of a multivariate Student's t-distribution is performed on the corrected centers of curvature. The weights are shown with the color map, the green points are weighted more in the Student's t-distribution whereas the red points belong to the tail of the distribution and are almost not taken into account.

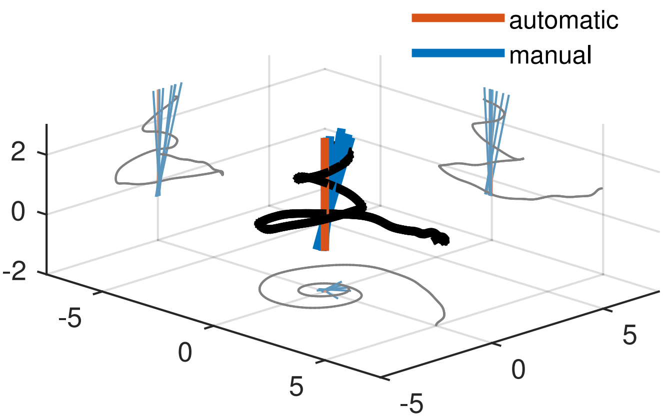


Figure 3.25: Automatic and manual modiolar axis estimation from a real cochlear centerline (patient #1).

parameters	$\boldsymbol{\mu_{A1}}$	$\boldsymbol{TH_{A1}}$	$\boldsymbol{\mu_{A2}}$ (mm)	$\boldsymbol{TH_{A2}}$ (mm)	$\boldsymbol{C_{init}}$ (mm)
values	10^{-5}	10^{-5}	10^{-3}	10^{-3}	2.8

Table 3.3: Parameters used to implement [Yoo 2000b]'s algorithm A

Least Square Minimization Method vs. Evolute-based [Yoo 2000b] designed two algorithms for the estimation of the modiolar axis from centerlines. Algorithm A utilized the longitudinal variation, while algorithm B considers the radial variation. They claimed that the former one is more accurate and suitable for three-dimensional data, which led us to implement it. Algorithm A is based on nonlinear optimization using steepest-descent algorithm and requires five parameters (μ_{A1}, TH_{A1}, μ_{A2}, TH_{A1}, C_{init}) whose values were not all specified in the article. We performed ourselves parameters tuning in order to get good results with the synthesized curves designed by the authors (summarized in Table 3.3).

The method described in subsection 3.2.4 is compared with the method introduced by [Yoo 2000b] using their own parametric model and angular sampling in order to perform a accurate comparison. Their centerline model is based on [Fowler 1992]'s model with a logarithmic spiral and an exponential term for the radial and the longitudinal component of the centerline. Fig. 3.26 shows the model for which the ideal axis is known.

[Yoo 2000b]'s method repeatedly updates the rotation parameters and the parameter C using the steepest-decent algorithm. The final rotation error depends on the initialization. The parameter C is an amplitude parameter that is used to model the longitudinal component of the helico-spiral. C depends on the (uniform) sampling (150 vertices for a total length of 35 mm). A criterion of convergence is that for each sampled point $i = 1, \ldots, n$, C needs to be greater than difference between the adjacent longitudinal component $(\mathbf{x}_i - \mathbf{x}_{i-1}) \cdot \hat{\mathbf{z}}$. Fig. 3.27 shows the rotation error with an optimized value of C_{init} (2.8 mm) but with different initial rotation around the x-axis (R_x), [Yoo 2000b] have tested the convexity of their objective function within a small orientation change (-10° to 10°), but in this study, μCT images of the temporal bones are acquired in random orientations.

We evaluated the robustness of both methods by adding Gaussian noise to the original centerline. For every noise amplitude we generated 100 noisy centerlines and evaluated the mean rotation error. Since [Yoo 2000b]'s algorithm needs to be initialized, we used the most favorable conditions, i.e. the values in Table 3.3 with the initial rotation around the x and y-axis (R_x and R_y) equal to 45° and 0° and different C_{init} values. Our method is run only once since it does not depend on parameter C_{init} and it uses the Frenet-Serret formulas which are independent from the initial coordinate system. The mean errors are plotted in Fig. 3.28.

We also estimated the distance (d_t) between the estimated modiolar axis and the ideal one, it is equal 0.005 μm, which would represent 0.2 pixel on a μCT image, while [Yoo 2000b] claimed 0 mm error.

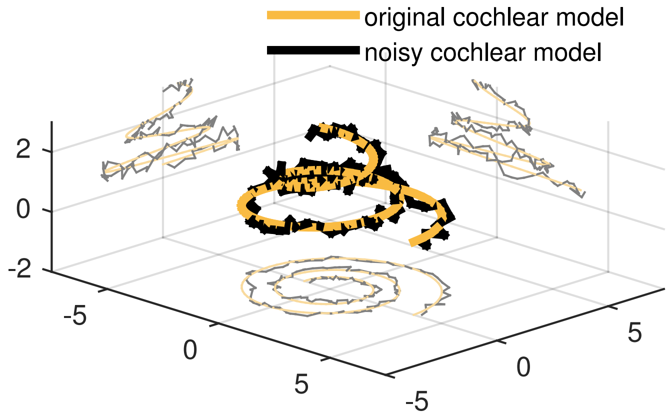


Figure 3.26: Synthetic cochlear model as defined by [Yoo 2000b] with and without random Gaussian noise at three different scale (0.23 mm, 0.94 mm, 3.76 mm) with a total amplitude of 0.18 mm.

Application to Real Data Fig. 3.24 presents the method applied on real patient data, while Fig. 3.29 and Fig. 3.30 express the influence of the estimated modiolar

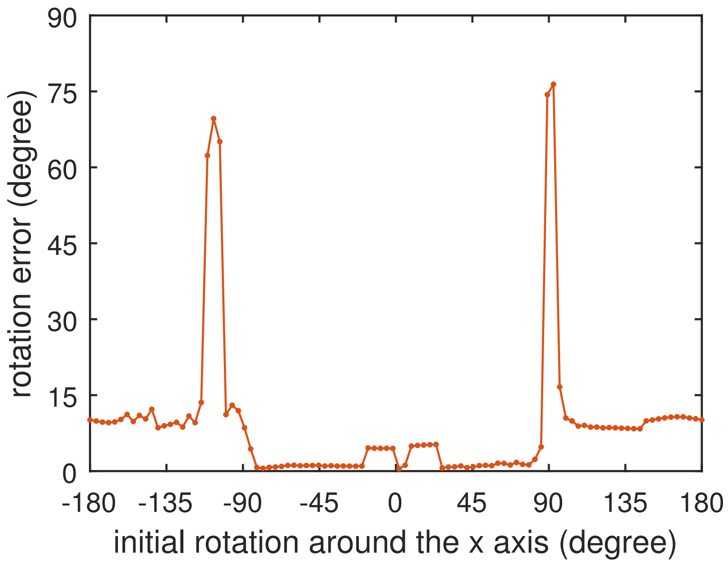


Figure 3.27: Sensitivity of the algorithm A developed by [Yoo 2000b] with respect to the initial rotation R_x.

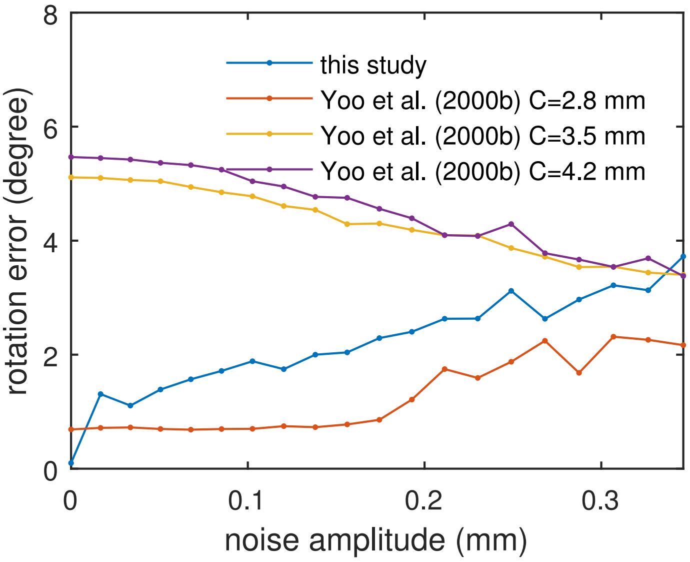


Figure 3.28: Noise sensitivity of the algorithm developed in this study and algorithm A developed by [Yoo 2000b] with different C_{init} values.

axis on the radial and longitudinal components of the centerlines. With a rotation of the modiolar axis of 11.5°, the projected components can be significantly different. [Avci 2014] defined categories of the longitudinal component of the tympanic centerline. Among other categories, the “rollercoaster” category was defined as a decreasing longitudinal component from the round window, changing to an increasing curve between 5 and 10 mm. The “sloping” category follows an increasing curve without any significant decreasing trends, local peaks located between 10 and 15 mm being also a feature of this category. Following this taxonomy, the estimated projection in Fig. 3.30 would have been classified as a “rollercoaster” whereas the perturbed one appears more to correspond to the “sloping” category with a peak around 10 mm.

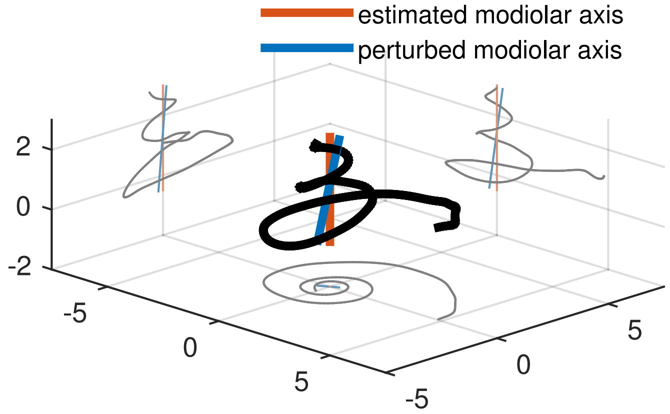


Figure 3.29: Influence of the axis estimation on the centerline projected onto the modiolar axis (called longitudinal component). The angle between the original and the modified axis is equal to 11.5°.

3.3.3 Cross-sections Measurements

Fig. 3.31 gives a sample of 24 cross-sections (represented in the uv-planes) computed from the three-dimensional segmentations.

The mean cross-section area (Fig. 3.33), height (Fig. 3.34) and width (Fig. 3.35) of the tympanic and vestibular ducts are presented. After 360°the width of the scala tympani and vestibuli tend to be equal.

Fig. 3.32 presents the individual tilt angle α, the mean angle between $\hat{\mathbf{u}}$ and the major axes of the tympanic and vestibular scalae. Examples of the influence of the tilt angle can be appreciated in Fig. 3.31a, b and c.

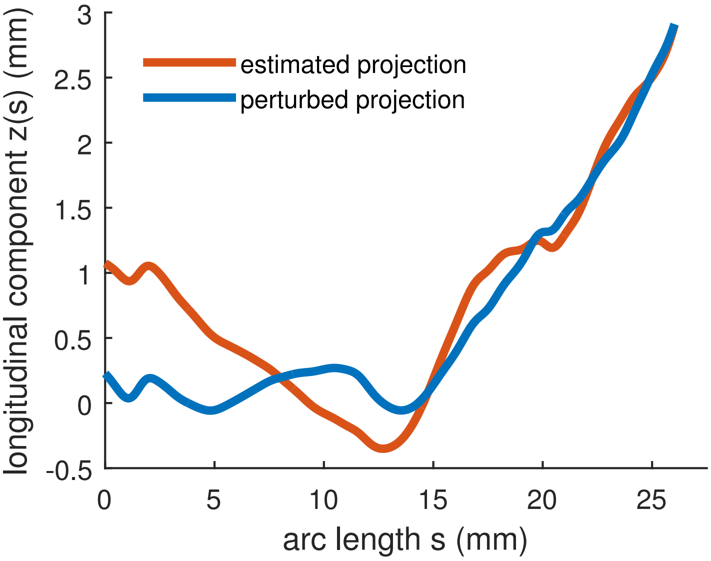


Figure 3.30: Longitudinal component of the cochlear centerline as a function of the arc length. Both longitudinal components are extracted from the same cochlear centerline.

3.4 Discussion

3.4.1 Parameterization Framework

The distance along the cochlear central path s and its angular parameterization θ are both used in the literature. The distance can be measured from the different centerlines (scala tympani, scala vestibuli or the whole cochlea), which produces different parameterization (see the different final length in table 3.2). Depending on the extraction method chosen, the cochlear central path may be shifted closer to the lateral (as in [Wysocki 1999, Avci 2014]) or modiolar wall which would entail again highly different parameterization (see Fig. 3.18 for the scala tympani). The angular coordinate on the other hand is independent from the centerline. Therefore, as concluded by [Verbist 2010] to allow accurate comparisons the angular parameterization should be preferred.

A significant variability is observed for the longitudinal components. The origin chosen for the coordinate system aligns all round windows but not the apexes, whereas the apical coordinate system introduced by [Verbist 2010] would have aligned all apexes but not the round windows.

The cochlear moving frame defined in this chapter allows one to easily measure the profile tilt angle. To our knowledge this chapter provides a first quantification of the tilt angle along the full centerline. [Clark 2011] introduced this parameter but without any estimation from real data. Their proposed model captures some of the variability (Fig. 3.32) but our systematic measurements provide a new insight

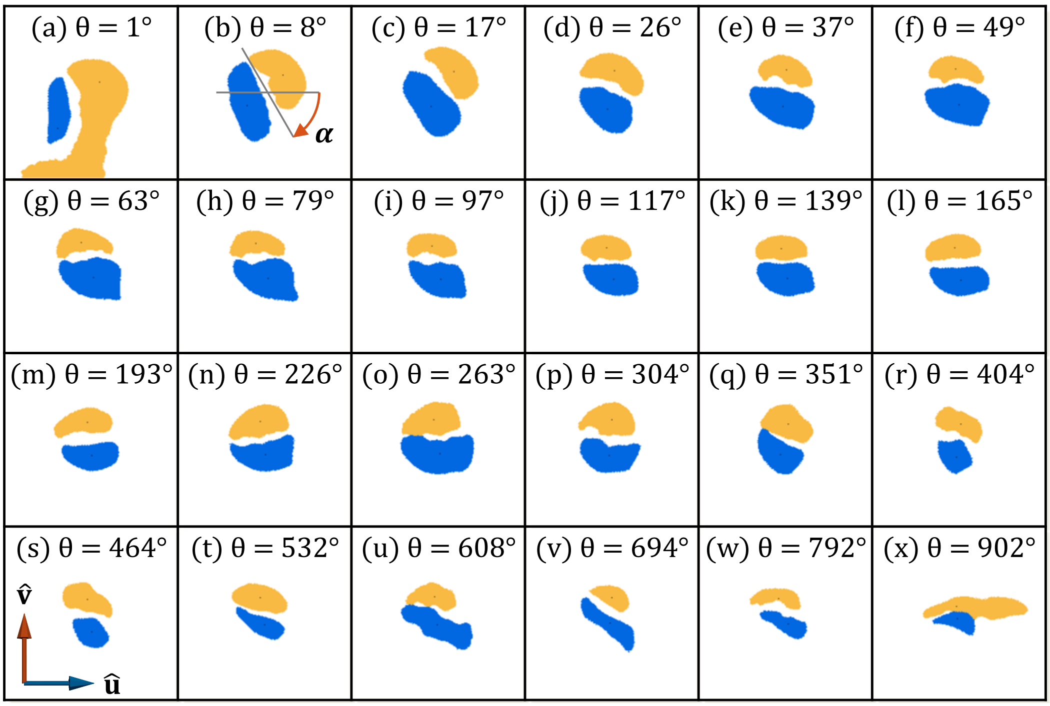


Figure 3.31: Cross-sections (patient #3), with the scala tympani (blue) and the scala vestibuli (yellow) from the base to the apex. Subfigure (a) illustrates a leakage of the label "scala vestibuli" to the vestibule. Subfigure (x) illustrates the difficulty to have well defined centerlines at the apex, where both scalae merge.

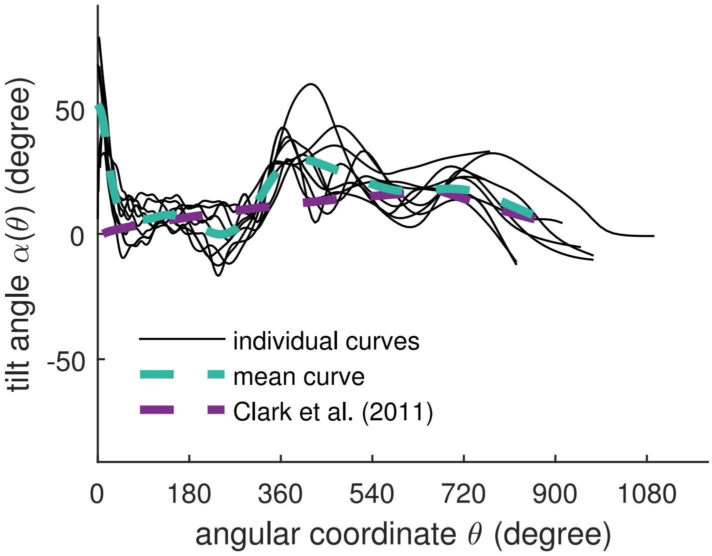


Figure 3.32: Individual and mean cochlear tilt angle α, the mean angle between $\hat{\mathbf{u}}$ and the major axes of the tympanic and vestibular scalae.

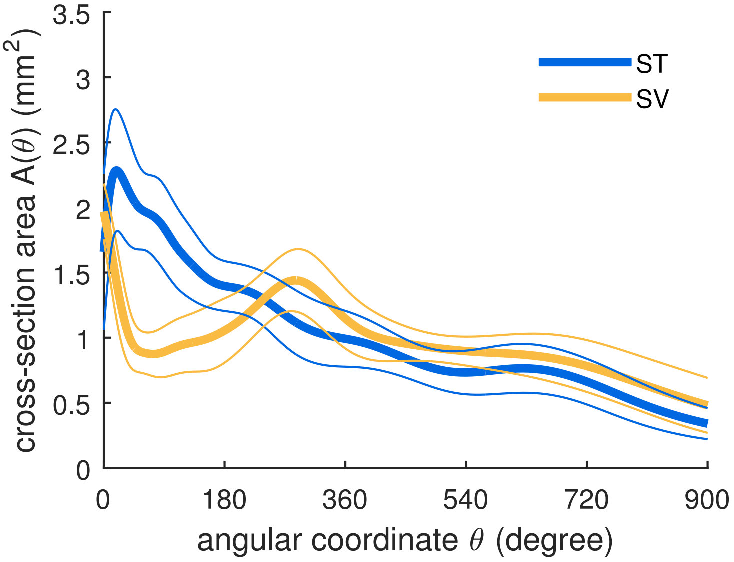


Figure 3.33: Cross-section area of the tympanic and vestibular scalae. The thinner lines represent ± 1 standard deviation.

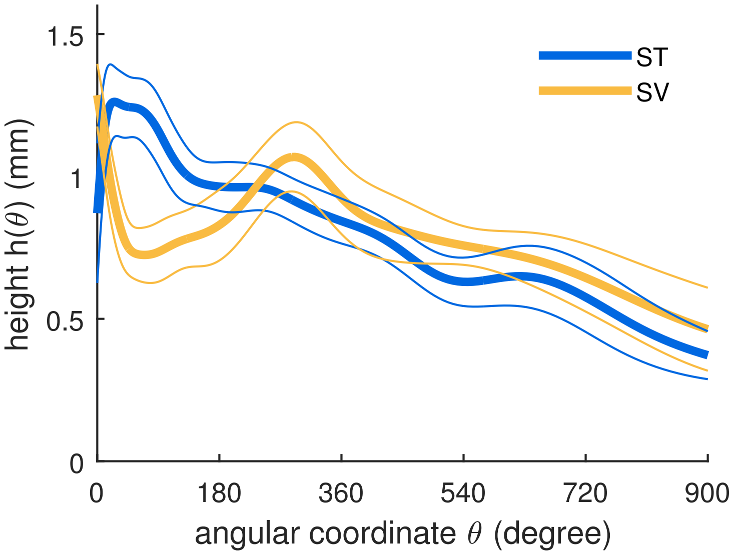


Figure 3.34: Heights of the tympanic and vestibular scalae.

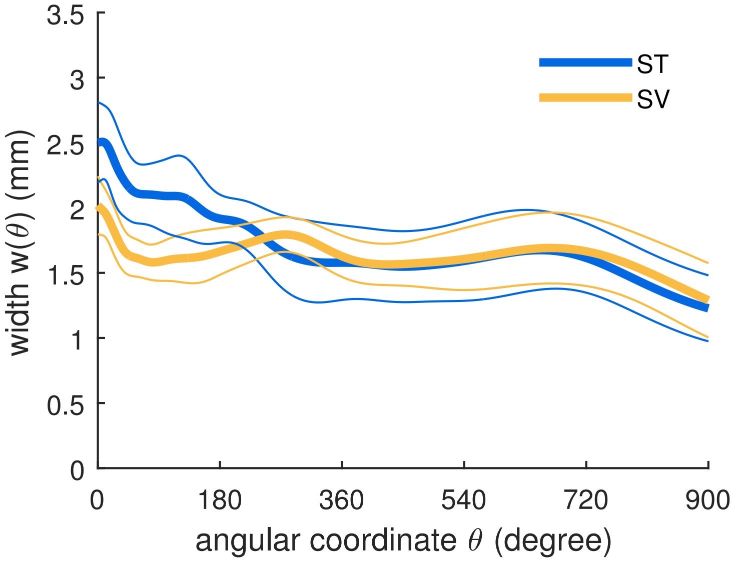


Figure 3.35: Width of the tympanic and vestibular scalae.

in the cochlear structure. [Gibson 2012] measured the rotation of the osseous spiral lamina in the hook region. The measurements are assessed at four locations (1, 3, 5 and 7 mm distal to the round window) along the basal turn of the cochlea. Slices of MRI images are extracted using an oblique sagittal plane. The acquisition of this plane called Pöschl projection is radiologist-dependent and the rotations measured are difficult to compare with the tilt angle.

The tilt angle measurements shows a relatively repetitive pattern. Close to the round window, the tilt angle is on average greater than 45°. Between 60 and 360°, the tilt angle falls around 0°, making the radial component of the tympanic and vestibular centerlines coincident. This decrease of the tilt angle along the basal turn matches with the prominent rotation noted in [Gibson 2012]'s study which has been performed between approximately 10 and 100°. After 360° the tilt angle turns positive before becoming flat again at the apex.

Additionally, the cochlear moving frame allows to deduct unambiguously the lateral and the modiolar path from the central one.

3.4.2 Variability Study

When the number of cochlear turns is not specifically detailed [Kawano 1996, Braun 2012, Kjer 2015a], many authors transcribed the mean, the standard deviation, the range and the number of the measurements [Erixon 2009, Shin 2013, Avci 2014]. For illustration purposes, Fig. 3.36 shows sets of data values which fulfill the measures listed above (the standard deviation was arbitrarily set to 0.14 for the data in [Erixon 2009]). A normal distribution was fit to the concatenated data. Comparing the results of this current study with previous findings shows that, on one hand the mean number of cochlear turn (2.58) falls extremely close to the mode of the normal distribution (2.59), on the other hand the standard deviation (0.28) is broader than what is found in the literature (0.11 [Kawano 1996], 0.09 [Shin 2013], 0.17 [Avci 2014]). This result could be explained by the presence of two extreme cases. Patient #5 (2.14) is close the lower end of the range observed by [Erixon 2009] (2.15) and according to [Jackler 2009]'s classification is not malformed. Cochleae with 3 turns such as the second outlier (3.08) were previously reported by [Tian 2006].

Irrespective to the parameterization, our cross-sections measurements are very similar to the ones from [Wysocki 1999] and [Thorne 1999], our height measurements being also comparable with diameter measurements obtained from inscribed circles [Biedron 2010, Avci 2014]. Tympanic cross-section area appears greater to the vestibular one in the basal turn and conversely in the middle and apical turns (as shown in Fig .3.33) match with [Gulya 1996, Wysocki 1999]'s observations.

The radial component of the different cochlear ducts shows low variability. In accordance with the results from previous studies [Cohen 1996, Baker 2008], a logarithmic spiral seems to describe well the projected cochlear shape.

The longitudinal component of the cochlea, as described by [Verbist 2009], shows

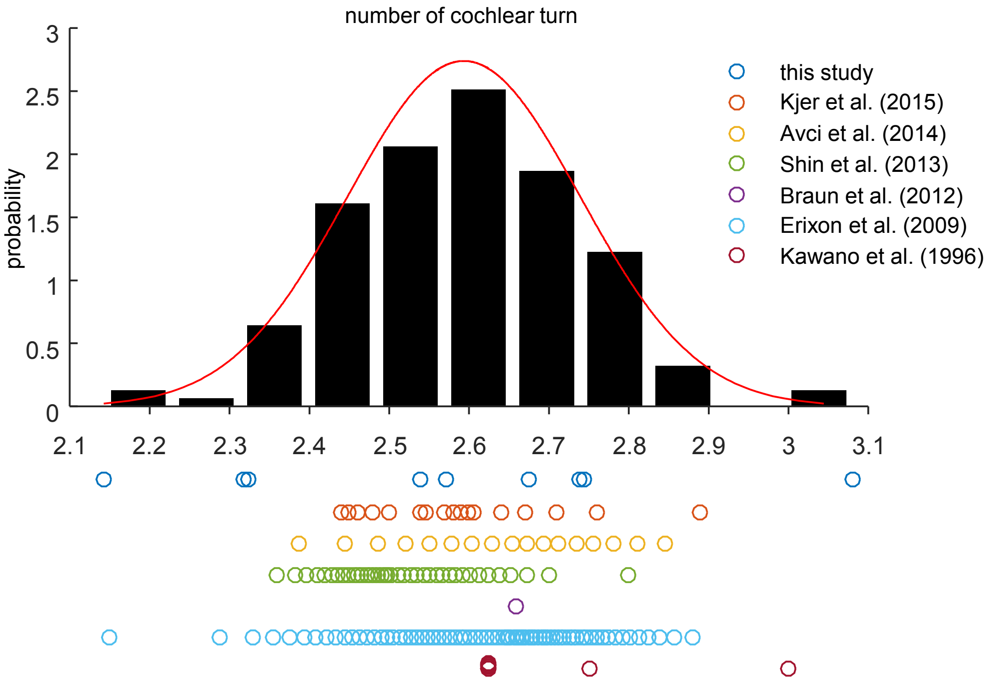


Figure 3.36: Comparison of measurements of the number of cochlear turns. A normal distribution was fitted to the concatenated measurements.

a pattern similar to the one in Fig. 3.23, which is a decreasing slope, changing to an increasing curve. The "proximal short rising" described by the authors has not been observed in our mean cochlear centerline data but can still be noticed in the mean tympanic and vestibular centerlines. The longitudinal component of the scala tympani was studied by [Avci 2014]. Unlike [Avci 2014] all the scala tympani showed the same profile, which would have been classified as a "rollercoaster" (see subsection 3.3.2).

Finally, despite the challenge in shape analysis at the apex, the results of this study have the advantage of being meaningful even after the second turn (720°).

3.4.3 Modiolar Axis Estimation

The comparison between manual and subsection 3.3.2 shows that automatically estimated axis is on average sightly further from the axis defined by the experts than the inter- and intra-expert standard deviation. More specifically, the rotation error is similar to the inter-expert variability while the translation is greater. The variability of the manually defined axes comes from the estimation of the apical axial point rather than the basal axial point. The experts tend to choose a point directly on the centerline rather than at the hypothetical center of the helico-spiral (see Fig. 3.25). Theoretically the radius of a logarithmic spiral converges to zero but never reaches it. We can reasonably assume these results can be explained by a bias that occurs with the manual estimation.

Moreover we have shown an example of the influence of a rotation error of 11.5° (Fig. 3.29), which is of the same order of magnitude as the mean angle error between the automatically and manually estimated axis (10.2°). In some cases the influence is large enough to change the profile of the longitudinal component. For these reasons we can hypothesize that the fact that only one class of [Avci 2014]'s classification is found in this study can be explained either by a lack of variability of our database or by the influence of our more reproducible modiolar axis estimation method compared with the manual one.

The comparison between the two estimation methods of the modiolar axis shows that the method of [Yoo 2000b] provides a good estimate only when the parameter C_{init} is correctly estimated and when the initialization angle is within a close range of the true value. However, the parameter C_{init} cannot be estimated easily since it depends on the curve discretization and the initial coordinate system. Our method on the other hand is an unbiased estimator of the axis in the absence of noise and does not depend on the choice of an additional extrinsic parameter or the reference frame. Its performance in the presence of noise is fairly similar (up to 2° difference) to the best estimate of [Yoo 2000b] and is robust to outliers. Since our approach relies on computing the center of curvature (Eq.) in the supplementary materials), of a noisy centerline, it could be improved by considering scale sensitive discrete curvature computation.

3.4.4 Implication for Cochlear Implantation

The preparation and the segmentation of the human cochlea allow us to assess with confidence meaningful information for cochlear implantation and even after the first two basal cochlear turns.

The centerlines of the scala tympani are useful to identify the potential location of insertion trauma. Most longitudinal components show an important "bump" at 184.5° ± 80.9 (mean and ± 1 standard deviation), which matches well with the empirical position (around 180°), observed by surgeons, of cochlear trauma and basilar perforation sites [Eshraghi 2003, Wardrop 2005, Nguyen 2012].

In order to make the centerline projections significant, one must pay attention on how to define correctly the cochlear coordinate system.

Fig. 3.32 shows a distinctive tilt angle, on average greater than 45° close to the round window (as in Fig. 3.31b). Such a feature may make the electrode array deviate in the scala vestibuli shortly after an insertion through a cochleostomy. The tilt angle may have implication for occurrence of trauma and cochleostomy site selection [Briggs 2009].

Straight electrodes tend to follow the lateral wall of the scala tympani whereas perimodiolar electrodes are designed to follow the modiolar wall. The measure of the lateral, central and modiolar arc lengths (Fig. 3.18) could provide, as in [Escudé 2006], information about ideal electrode array length according to the type of electrode and the desired insertion depth.

The height of the scala tympani was measured as the minor axis of the cross-sections, it is also comparable with the diameter of the inscribed circle (subsection 3.4.2). The results are independent from the cross-section plane orientation. It provides an upper limit for the dimension of an electrode array. We showed (Fig. 3.34) that it is of primary importance to have an electrode diameter smaller than 0.5 mm in order to be atraumatic after 360°.

3.5 Conclusion

We designed an automated method for the cochlear centerline extraction given a segmented cochlea image. A robust method for modiolar axis estimation was developed, validated on a synthetic cochlear model, compared with manual estimation and another algorithm from the literature and finally applied on segmented images of nine temporal bones. Objective geometrical measurement were performed on the tympanic and vestibular duct segmentations assessing, among others, the radial and longitudinal components of the lateral, central and modiolar paths. Cross-section measurements were estimated within a new non-ambiguous moving frame. Close to the round window, the cross-section tilt angle shows an important feature for insertion trauma comprehension. This study was meant to be reproducible and suitable for a larger database in order to improve significantly the knowledge of the cochlea shape variability. Finally, this study could provide the key measurements to validate a parametric shape model of the cochlea and its inner structures. This model could be used to provide prior anatomical information required for analyzing clinical CT images. Indeed due to spatial resolution limitation, conventional CT images alone cannot provide enough fine information about anatomical structures relevant for cochlear implantation [Shin 2013]. Specifically the basilar membrane that delimits the scala tympani is nearly invisible with clinically available imaging techniques. Based on the presented results, further studies will analyzed the human cochlear shape variability from large CT image databases.

Chapter 4

A New Parametric Cochlear Shape Model

Contents

This chapter is adapted from [Demarcy 2016b].
Thomas Demarcy, Clair Vandersteen, Nicolas Guevara, Charles Raffaelli, Dan Gnansia, Nicholas Ayache and Hervé Delingette. *Uncertainty Quantification of Cochlear Implant Insertion from CT images.* Clinical Image-Based Procedures. Translational Research in Medical Imaging: 5th International Workshop. Held in Conjunction with MICCAI 2016, Athens, Greece, pages 27-35, 2016.

Abstract

Cochlear implants (CI) are used to treat severe hearing loss by surgically inserting an electrode array into the cochlea. Since current electrodes are designed with various insertion depth, ENT surgeons must choose the implant that will maximize the insertion depth without causing any trauma based on preoperative CT images. In this chapter, we propose a novel framework for estimating the insertion depth and its uncertainty from segmented CT images based on a new parametric shape model. Our method relies on the posterior probability estimation of the model parameters using stochastic sampling and a careful evaluation of the model complexity compared to CT and μCT images. The results indicate that preoperative CT images can be used by ENT surgeons to safely select patient-specific cochlear implants.

4.1 Introduction

A cochlear implant (CI) is a surgically implanted device used to treat severe to profound sensorineural hearing loss. The implantation procedure involves drilling through the mastoid to open one of the three cochlear ducts, the scala tympani (ST), and insert an electrode array to directly stimulate the auditory nerve, which induces the sensation of hearing. The post-operative hearing restoration is correlated with the preservation of innervated cochlear structure, such as the modiolus and the osseous spiral lamina, and the viability of hair cells [Nadol 1997].

Therefore for a successful CI insertion, it is crucial that the CI is fully inserted in the ST without traumatizing the neighboring structures. This is a difficult task as deeply inserted electrodes are more likely to stimulate wide cochlear regions but also to damage sensitive internal structures. Current electrode designs include arrays with different lengths, diameters, flexibilities and shapes (straight and preformed). Based on the cochlear morphology selecting the patient-appropriate electrode is a difficult decision for the surgeon [van der Marel 2014].

For routine CI surgery, a conventional CT is usually acquired for insertion planning and abnormality diagnosis. However, the anatomical information that can be extracted is limited. Thus, important structures, such as the basilar membrane that separates the ST from other intracochlear cavities, are not visible. On the other hand, high resolution μCT images leads to high quality observation of the cochlear cavities but can only be acquired on cadaveric temporal bones.

Several authors have devised reconstruction methods of the cochlea from CT images by incorporating shape information extracted from μCT images. In particular, [Noble 2011] and [Kjer 2015b] created statistical shape models of the cochlea based on high-resolution segmented μCT images. Those shape models are created from a small number of μCT images (typically 10) and therefore may not represent well the generality of cochlear shapes that can bias the CT anatomical reconstruction. [Baker 2005] used a parametric model based on 9 parameters to describe the cochlear as a spiral shell surface. This model was fit to CT images by assuming that the surface model matches high gradient voxels.

In this chapter, we aim at estimating to which extent a surgeon can choose a proper CI design for a specific patient based on CT imaging. More specifically, we consider 3 types of implant designs based on their positioning behavior (see Fig. 4.10) and evaluate for each design the uncertainty in their maximal insertion depth. If this uncertainty is too large then there is a risk of damaging the ST during the insertion by making a wrong choice. For this uncertainty quantification, we take specific care of the bias-variance tradeoff induced by the choice of the geometric model. Indeed, considering an oversimplified model of the cochlea will typically lead to an underestimation of the uncertainty whereas an overparameterized model would conversely lead to an overestimation of uncertainty.

Therefore, we introduce in this chapter a new parametric model of the cochlea and estimate the posterior distribution of its parameters using Markov Chain Monte Carlo (MCMC) method with non informative priors. We devised likelihood functions that relate this parametric shape with the segmentation of 9 pairs of CT and μCT images. The risk of overparameterization is evaluated by measuring the entropy of those posterior probabilities leading to possible correlation between parameters. This generic approach leads to a principled estimation of the probability of CI insertion depths for each of the 9 CT and μCT cases.

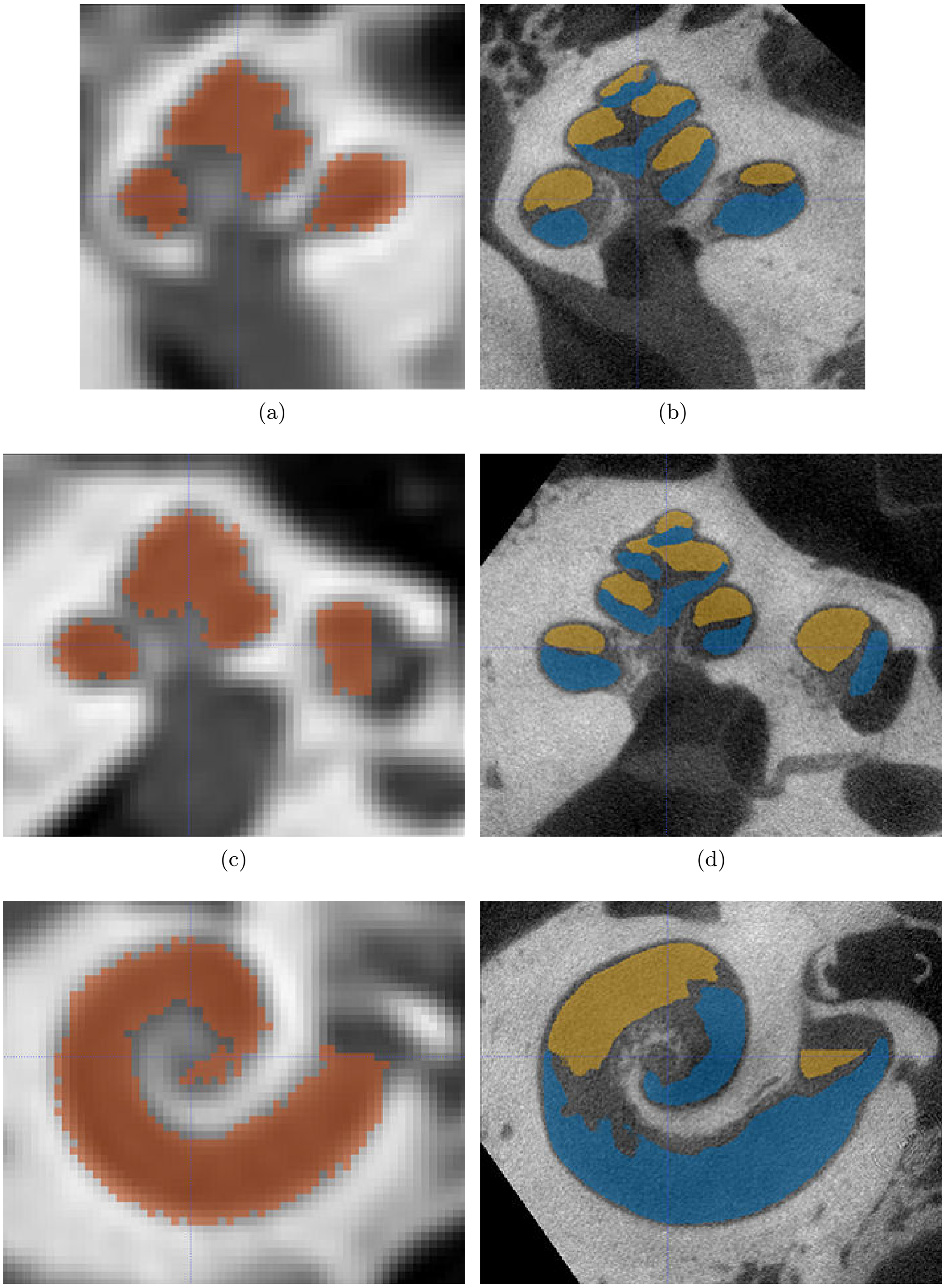
4.2 Methods

4.2.1 Data

Healthy temporal bones from 9 different cadavers were scanned using CT and μCT scanners. CT scans were acquired at the Imaging Center of the Nice University Hospital using a GE LightSpeed VCT CT system under the supervision of Pr Charles Raffaelli, while μCT scans were acquired in the laboratory Transporters, Imaging and Radiotherapy in Oncology (TIRO, UMR-E4320) in Nice using the GE eXplore speCZT120 scanner under the supervision of Pr Thierry Pourcher. Unlike CT images, which have a voxel size of 0.1875x0.1875 x0.25 mm^3 (here resampled to 0.2x0.2x0.2 mm^3) the resolution of μCT images (0.025 mm per voxel) is high enough to identify the basilar membrane that separates the ST from the scala vestibuli (SV) and the scala media. The scala media represents a negligible part of the cochlear anatomy, for simplicity purposes, both SV and scala media will be referred as the SV. Since intracochlear anatomy are not visible in CT images, only the cochlea was manually segmented by an head and neck imaging expert, while the ST and the SV were segmented in μCT images (see Fig. 4.1). All images were rigidly registered using a pyramidal block-matching algorithm [Ourselin 2000] and aligned in a cochlear coordinate system [Verbist 2010].

4.2.2 Parametric Cochlear Shape Model

Since we have a very limited number of high resolution images of the cochlea, we cannot use statistical shape models to represent the generality of those shapes. Instead, we propose a novel parametric model $\mathcal{M}$ of the 3 spiraling surfaces: the whole cochlea, the scala tympani and scala vestibuli (see Fig. 4.2). The cochlea corresponds to the surface enclosing the 2 scalae and we introduce a compact parameterization $\mathcal{T} = \{\tau_i\}$ based on 22 parameters for describing the 3 surfaces. This model extends in several ways the ones previously proposed in the literature [Baker 2005] as to properly capture the complex longitudinal profile of the centerline and the specific shapes of the cross-sections detailed in clinical studies [Wysocki 1999]. More precisely, in this novel model, the cochlea and two scalae can be seen as generalized cylinders, i.e cross-sections swept along a spiral curve. This centerline is parameterized in a cylindrical coordinate system by its radial $r(\theta)$ and longitudinal $z(\theta)$ functions of the angular coordinate θ within a given interval $[0, \theta_f]$. The cross-sections of the ST and SV are modeled by a closed planar curve on which a varying affinity transformation is applied along the centerline, parameterized by an angle of rotation $\alpha(\theta)$ and



(a) (b)

(c) (d)

(e) (f)

Figure 4.1: Slices of CT (a,c,e) and μCT (b,d,f) with segmented cochlea (red), ST (blue) and SV (yellow).

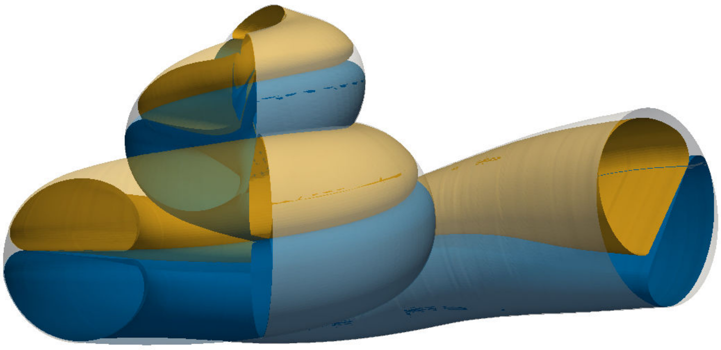


Figure 4.2: Parametric model with the ST (blue), the SV (yellow) and the whole cochlea (translucent white).

two scaling parameters $w(\theta)$ and $h(\theta)$. In particular, the three modeled anatomical structures shared the same centerline, the tympanic and vestibular cross-sections are modeled with two half pseudo-cardioids within the same oriented plane while the cochlear cross-section corresponds the minimal circumscribed ellipse of the union of the tympanic and vestibular cross-sections (see Fig. 4.10). The center of the ellipse is on the centerline. Eventually the shapes are fully described by 7 one-dimensional functions of θ: $r(\theta)$, $z(\theta)$, $\alpha(\theta)$, $w_{ST}(\theta)$, $w_{SV}(\theta)$, $h_{ST}(\theta)$, $h_{SV}(\theta)$, combinations of simple functions (i.e polynomial, logarithmic, ...) of θ.

4.2.2.1 Previous models

Historically, cochlear models first described the cochlear centerline prior to emergence of solid shape models.

Early Models The cochlea had long been identified as shell-like spiral structure [Leidy 1883], hence it etymology derived from the Latin word for snail-shell.

Parametric modeling of seashell was emphasized before parametric shape modeling of cochlea [Thompson 1917]. As described in the review of shell parametric models by [Stone 1996], [Moseley 1838] was the first to model properly shell shape using polar coordinates.

[Ketten 1990] identified two cochlear canal spiral types in cetaceans : Archimedian spiral (4.1) and logarithmic spiral (4.2) while, at the same time, [Voie 1990]

parameters	a	b	c	d	θ_0	θ_1	θ_2	θ_f
values	3.76	0.0755	3.82	0.269	0.0873	0.180	1.74	15.9

Table 4.1: Parameters of Cohen's model (4.4) with $\theta \in [\theta_1, \theta_f]$ (in radians). a and c are in millimeters, all others are dimensionless.

noted the resemblance of the guinea pig scala tympani with an Euler spiral (4.3) .

$$r(\theta) = a\theta \tag{4.1}$$

$$r(\theta) = ae^{-b\theta} \tag{4.2}$$

$$x(l) = \int_0^l \cos s^2 \mathrm{d}s \tag{4.3a}$$

$$y(l) = \int_0^l \sin s^2 \mathrm{d}s \tag{4.3b}$$

Cohen's Model [Cohen 1996] introduced a two-dimensional parametric model describing the radial component $r(\theta)$ of the electrode array placed in the human scala tympani. The radial component consists in a piecewise logarithmic function. The piecewise defined functions allow to take into account the shape of the basal turn.

$$r(\theta) = \begin{cases} c(1 - d\log(\theta - \theta_0)) & \text{if } \theta < \theta_2 \quad (4.4a) \\ ae^{-b\theta} & \text{else.} \quad (4.4b) \end{cases}$$

Ketten's Model [Ketten 1998] introduced a three-dimensional parametric model of the human cochlea based on their previous work [Ketten 1990] adopting the Archimedian spiral (Eq 4.1) as radial component $r(\theta)$ of the cochlear central path. A major motivation was to measure the cochlear length easily so the vertical component was defined such as the Pythagorean theorem relates, on one hand the arc length of the Archimedian spiral in the xy-plane s_{pol} (defined in Eq 4.5) and the height h of the cochlea and, and on the other hand, the cochlear arc length s_{cyl} (Eq 4.6) .

$$s_{pol}(\theta) = \int_0^\theta \sqrt{r(t)^2 + \left(\frac{\mathrm{d}r(t)}{\mathrm{d}t}\right)^2} \mathrm{d}t \tag{4.5}$$

$$s_{cyl}(\theta_{max}) = \sqrt{s_{pol}(\theta_{max})^2 + h^2}. \tag{4.6}$$

The vertical component defined as above would be suitable with Ketten's definition (Eq 4.5 and Eq 4.6):

$$z(\theta) = \frac{s_{pol}(\theta)}{s_{pol}(\theta_{max})} h. \tag{4.7}$$

Despite its reasonably complex formulation and its lack of anatomical relevance, Ketten's model offers the advantage that it only needs three parameters to describe the three-dimensional cochlear central path. Besides, the parameters are physically meaningful: θ_{max} depicts the number of cochlear turns, a is a scale factor representing the cochlear diameter at the basal turn and h is the cochlear height.

Helico-spiral Models Several other three-dimensional models were built from two-dimensional models adding the longitudinal component $z(\theta)$. Yoo et al. modeled the helico-spiral cochlear shape adapting, first the seashell model of [Fowler 1992] with a logarithmic spiral (Eq 4.2) and an exponential longitudinal component (Eq 4.8) [Yoo 2000b], and second Cohen's model (Eq 4.4) by adding a linear longitudinal component (Eq 4.9) [Yoo 2000a].

$$z(\theta) = ce^{d\theta} \tag{4.8}$$

$$z(\theta) = e(\theta - \theta_1) \tag{4.9}$$

Solid Shape Models Based on the centerline modeling, solid shape models were introduced by modeling the cross-sectional shapes. [Manoussaki 2000] modeled the cross-section of the combined scala vestibuli and scala media as a constant rectangle, [Baker 2005] and [Zhang 2006] modeled the cochlear cross-section as an ellipse (Fig. 4.3) and [Clark 2011] modeled the cross-section of the scala tympani as semi-circular ends connected by straight segments. The latter can be considered as the most realistic three-dimensional parametric shape model.

Statistical Cochlear Models Other approaches try to learn the intra-cochlear anatomy using Statistical Shape Model (SSM), Active Shape Model (ASM) or Atlas-based procedure [Noble 2011, Noble 2012, Noble 2013, Noble 2013, Kjer 2014a, Kjer 2015d, Romera Romero 2016]. Statistical Shape Model are usually built from high-resolution μCT. The basic idea is to describe the shape as a set of points. After aligning the training shape using rigid registration and defining a mapping from the landmarks of one shape to another, the dimensionality of the data is reduced using most frequently Principle Component Analysis (PCA). For a detailed explanation of the SSM methodology please refer to [Paulsen 2004, Heimann 2009].

4.2.2.2 New Parametric Model Equations

Centerline The centerline is a curve defined in a cylindrical coordinate system (where $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{z}})$ are the local radial, circumferential and longitudinal unit vectors) as,

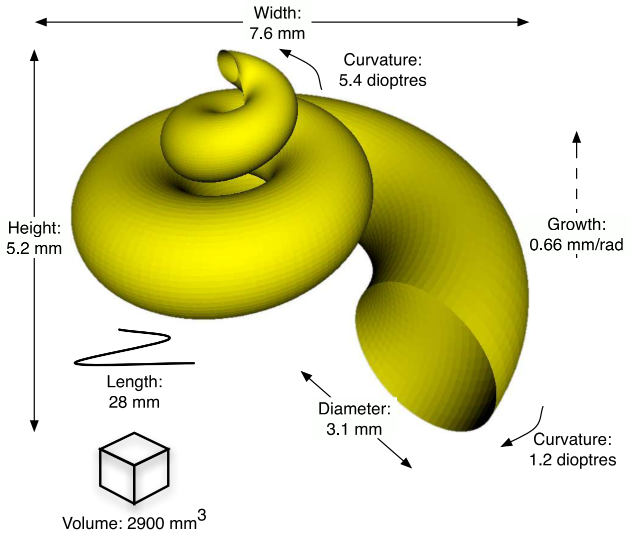


Figure 4.3: Example of cochlear solid shape model from [Baker 2008]

$$\mathbf{x}_c(\theta) = r(\theta)\hat{\mathbf{r}}(\theta) + z(\theta)\hat{\mathbf{z}}. \tag{4.10}$$

The radial component is defined piecewise with a polynomial function and a logarithmic function,

$$r(\theta) = \begin{cases} p_2\theta^2 + p_1\theta + p_0 & \text{if } \theta < \theta_2 \quad (4.11a) \\ ae^{-b\theta} & \text{else.} \quad (4.11b) \end{cases}$$

The longitudinal component is defined as a sum of two terms,

$$z(\theta) = z_0(\theta) * \frac{1}{w}\Pi\left(\frac{\theta}{w}\right) + z_1(\theta) \tag{4.12}$$

where the symbol $*$ denotes the convolution and Π is the rectangle function that is null outside $[-1/2, 1/2]$ and unity inside, the constant w parameterized the width of the rectangle function and thus the span of the smoothing on $z_0(\theta)$.

$$z_0(\theta) = \begin{cases} q_1\theta + q_0 & \text{if } \theta < \theta_1 \quad (4.13a) \\ r_1\theta + r_0 & \text{else} \quad (4.13b) \end{cases}$$

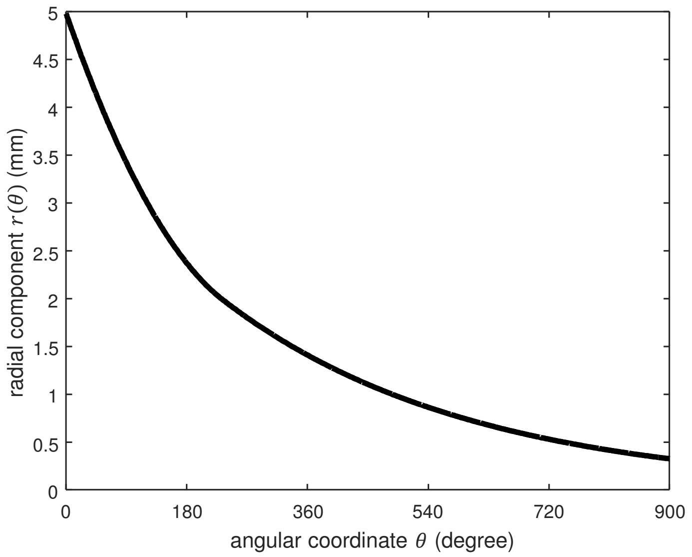


Figure 4.4: Radial component of the centerline.

scale factor ($a < a$)

spiral pitch, $\psi = \cot^{-1} b$ ($\psi < \psi$)

angular threshold ($\theta_2 < \theta_2$)

radius at origin ($p_0 < p_0$)

Figure 4.5: Illustration of the influence of the parameters a, ψ, θ_2, p_0.

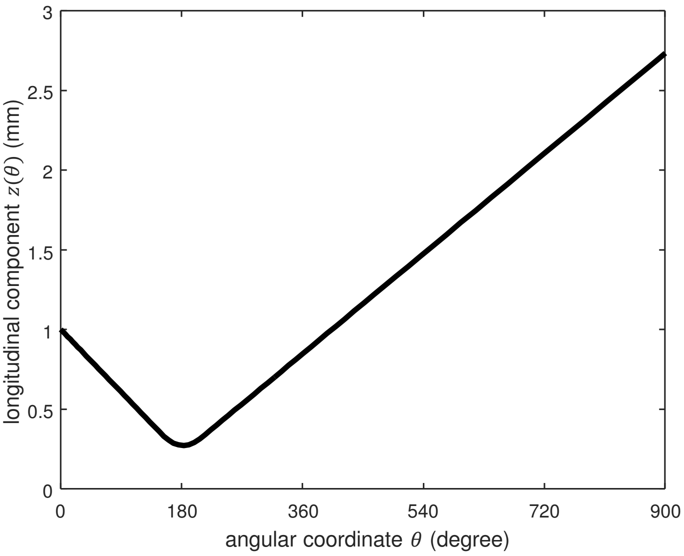


Figure 4.6: First term z_0 of longitudinal component of the centerline.

$$z_0(\theta) * \frac{1}{w}\Pi\left(\frac{\theta}{w}\right) = \begin{cases} q_1\theta + q_0 & \text{if } \theta < \theta_1 - \frac{w}{2} \quad (4.14a) \\ \frac{q_1}{2}\left(\theta_1^2 - \left(\theta - \frac{w}{2}\right)^2\right) + q_0\left(\theta_1 - \left(\theta - \frac{w}{2}\right)\right) & \\ \quad +\frac{r_1}{2}\left(\left(\theta + \frac{w}{2}\right)^2 - \theta_1^2\right) + r_0\left(\left(\theta + \frac{w}{2}\right) - \theta_1\right) & \\ & \text{if } \theta_1 - \frac{w}{2} \leq \theta < \theta_1 + \frac{w}{2} \quad (4.14b) \\ r_1\theta + r_0 & \text{else} \quad (4.14c) \end{cases}$$

$$z_1(\theta) = ke^{-\frac{(\theta-\mu)^2}{2\sigma^2}} cos(\omega\theta + \varphi) \tag{4.15}$$

Cross-section The cross-section shapes are defined for the scala tympani, the scala vestibuli and the cochlea as a whole. Respectively, the subscripts ST, SV and Co are used for identification purposes.

$$u_0(\phi) = \begin{cases} \frac{2}{9}\left(2\cos\phi - \cos 2\phi + \frac{1}{2}\right) + \frac{5}{9}\sqrt{1-2\tau} & \text{if } 0 \leq \phi \leq \phi_0 \quad (4.16a) \\ \frac{5}{9}(\tau\cos(f(\phi))) & \text{if } \phi_0 \leq \phi \leq \pi \quad (4.16b) \\ g(\phi) & \text{if } \pi \leq \phi \leq 2\pi \quad (4.16c) \end{cases}$$

negative slope ($q_1 < q_1$)

intercept ($q_0 < q_0$)

angular threshold ($\theta_1 < \theta_1$)

positive slope ($r_1 < r_1$)

Figure 4.7: Illustration of the influence of the parameters q_1, q_0, θ_1, r_1.

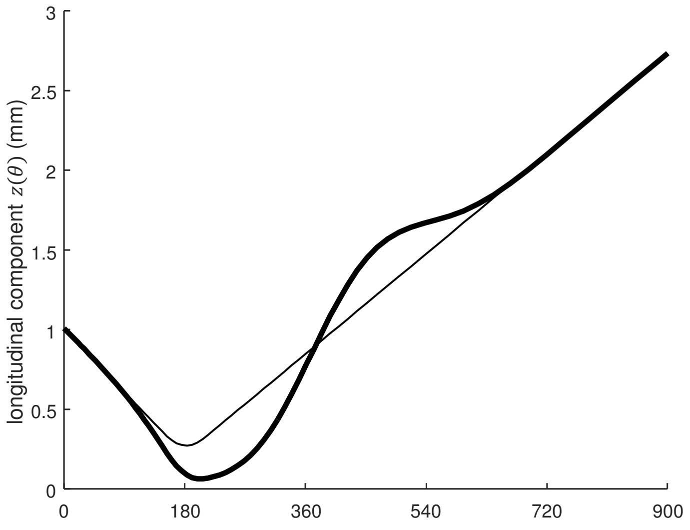


Figure 4.8: Second term z_1 of longitudinal component of the centerline.

$$
v_0(\phi) = \begin{cases} \dfrac{2}{3\sqrt{3}}(2\sin\phi - \sin 2\phi) & \text{if } 0 \leq \phi \leq \phi_0 \quad (4.17\text{a}) \\ \tau \sin(f(\phi)) + \tau & \text{if } \phi_0 \leq \phi \leq \pi \quad (4.17\text{b}) \\ 0 & \text{if } \pi \leq \phi \leq 2\pi \quad (4.17\text{c}) \end{cases}
$$

Eq. 4.16a and 4.17a define a cardioid segment, Eq. 4.16b and 4.17b define an ellipse segment and Eq. 4.16c and 4.17c define a line segment. $\tau \in [0, 1/2]$. The affine functions $f(\phi)$ and $g(\phi)$ and the angular threshold ϕ_0 are defined as,

$$
f(\phi) = \frac{\dfrac{\pi}{2} + \arcsin\left(\dfrac{\tau}{1-\tau}\right)}{\pi - \phi_0}(x - \phi_0) + \pi - \arcsin\left(\frac{\tau}{1-\tau}\right) \tag{4.18}
$$

$$
g(\phi) = \frac{5}{9\pi}\left(\frac{3}{5} + \sqrt{1 - 2\tau}\right)(x - \pi) \tag{4.19}
$$

$$
\phi_0 = \arccos\left(1 - \frac{\sqrt{A^2 + B^2}}{2}\right) \tag{4.20}
$$

amplitude ($k < k$)

center of the peak ($\mu < \mu$)

angular frequency ($\omega < \omega$)

phase ($\phi < \phi$)

Figure 4.9: Illustration of the influence of the parameters k, μ, ω, ϕ.

with,

$$A = \frac{-5\sqrt{1-2\tau}}{2(1-\tau)} - \frac{3}{2} \tag{4.21}$$

$$B = \frac{3\sqrt{3}\tau}{2(1-\tau)}. \tag{4.22}$$

Introducing the following parametric equations parameterized by $\phi \in [0, 2\pi]$,

$$\begin{bmatrix} \tilde{u}_{ST}(\theta,\phi) \\ \tilde{v}_{ST}(\theta,\phi) \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} w_{ST}(\theta) & 0 \\ 0 & -h_{ST}(\theta) \end{bmatrix} \begin{bmatrix} u_0(\phi) \\ v_0(\phi) \end{bmatrix} \tag{4.23}$$

$$\begin{bmatrix} \tilde{u}_{SV}(\theta,\phi) \\ \tilde{v}_{SV}(\theta,\phi) \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} w_{SV}(\theta) & 0 \\ 0 & h_{SV}(\theta) \end{bmatrix} \begin{bmatrix} u_0(\phi) \\ v_0(\phi) \end{bmatrix}. \tag{4.24}$$

Let us define the Löwner-John ellipse [Henk 2012, John 2014], parameterized by $\phi \in [0, 2\pi]$,

$$\mathbf{e}_{LJ}(\theta,\phi) = \begin{bmatrix} t_1(\theta) \\ t_2(\theta) \end{bmatrix} + \begin{bmatrix} T_{1,1}(\theta) & T_{1,2}(\theta) \\ T_{2,1}(\theta) & T_{2,2}(\theta) \end{bmatrix} \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \tag{4.25}$$

as a minimal circumscribed ellipse of:

$$\bigcup_{\phi\in[0,2\pi]} \left\{ \begin{bmatrix} \tilde{u}_{ST}(\theta,\phi) \\ \tilde{v}_{ST}(\theta,\phi) \end{bmatrix}, \begin{bmatrix} \tilde{u}_{SV}(\theta,\phi) \\ \tilde{v}_{SV}(\theta,\phi) \end{bmatrix} \right\}. \tag{4.26}$$

$\forall i \in \{ST, SV\}$,

$$\begin{bmatrix} u_i(\theta,\phi) \\ v_i(\theta,\phi) \end{bmatrix} = \begin{bmatrix} \cos\alpha(\theta) & \sin\alpha(\theta) \\ -\sin\alpha(\theta) & \cos\alpha(\theta) \end{bmatrix} \left(\begin{bmatrix} \tilde{u}_i(\theta,\phi) \\ \tilde{v}_i(\theta,\phi) \end{bmatrix} - \begin{bmatrix} t_1(\theta) \\ t_2(\theta) \end{bmatrix} \right) \tag{4.27}$$

$$\begin{bmatrix} u_{Co}(\theta,\phi) \\ v_{Co}(\theta,\phi) \end{bmatrix} = \begin{bmatrix} \cos\alpha(\theta) & \sin\alpha(\theta) \\ -\sin\alpha(\theta) & \cos\alpha(\theta) \end{bmatrix} \begin{bmatrix} T_{1,1}(\theta) & T_{1,2}(\theta) \\ T_{2,1}(\theta) & T_{2,2}(\theta) \end{bmatrix} \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix} \tag{4.28}$$

Cross-section Parameters Along the Centerline From now on, unless otherwise specified, the variables using the subscripts i are defined for $i \in \{ST, SV, Co\}$. The cross-section parametric equations $u_i(\theta,\phi)$ and $v_i(\theta,\phi)$ represent functions of the parameters $w_{ST}(\theta)$, $h_{ST}(\theta)$, $w_{SV}(\theta)$, $h_{SV}(\theta)$, $\alpha(\theta)$, β and τ.

$$w_{ST}(\theta) = \overline{w}_{ST}(\theta) + w_{ST1}\theta + w_{ST0} \tag{4.29}$$

$$\alpha(\theta) = \overline{\alpha}(\theta) \tag{4.30}$$

were $\overline{w}_{ST}(\theta)$ and $\overline{\alpha}(\theta)$ are constant polynomial functions (Table 4.2). $h_{ST}(\theta)$, $w_{SV}(\theta)$, $h_{SV}(\theta)$ are as defined in the equivalent manner as for $w_{ST}(\theta)$. β and τ are constants set to 0.25 and 0.04.

Full Model The full parametric cochlear shapes $\mathbf{x}_i(\theta, \phi)$ can be written as generalized cylinders [Voie 1990],

$$\mathbf{x}_i(\theta, \phi) = \mathbf{x}_c(\theta) + u_i(\theta, \phi)\hat{\mathbf{u}}(\theta) + v_i(\theta, \phi)\hat{\mathbf{v}}(\theta) \tag{4.31}$$

parameterized by $(\theta, \phi) \in [0, \theta_{max}] \times [0, 2\pi]$, and where $(\hat{\mathbf{t}}, \hat{\mathbf{u}}, \hat{\mathbf{v}})$ define a moving frame, with $\hat{\mathbf{t}}(\theta)$ the tangent unit vector of $\mathbf{x}_c(\theta)$, such as,

$$\hat{\mathbf{t}}(\theta) = \frac{\mathrm{d}\mathbf{x}_c(\theta)}{\mathrm{d}s(\theta)} \tag{4.32}$$

$$\hat{\mathbf{u}}(\theta) = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{t}}(\theta)}{\|\hat{\mathbf{z}} \times \hat{\mathbf{t}}(\theta)\|} \tag{4.33}$$

$$\hat{\mathbf{v}}(\theta) = \hat{\mathbf{t}}(\theta) \times \hat{\mathbf{u}}(\theta) \tag{4.34}$$

where the symbol × denotes the cross product and the infinitesimal part of the curve $\mathrm{d}s(\theta)$,

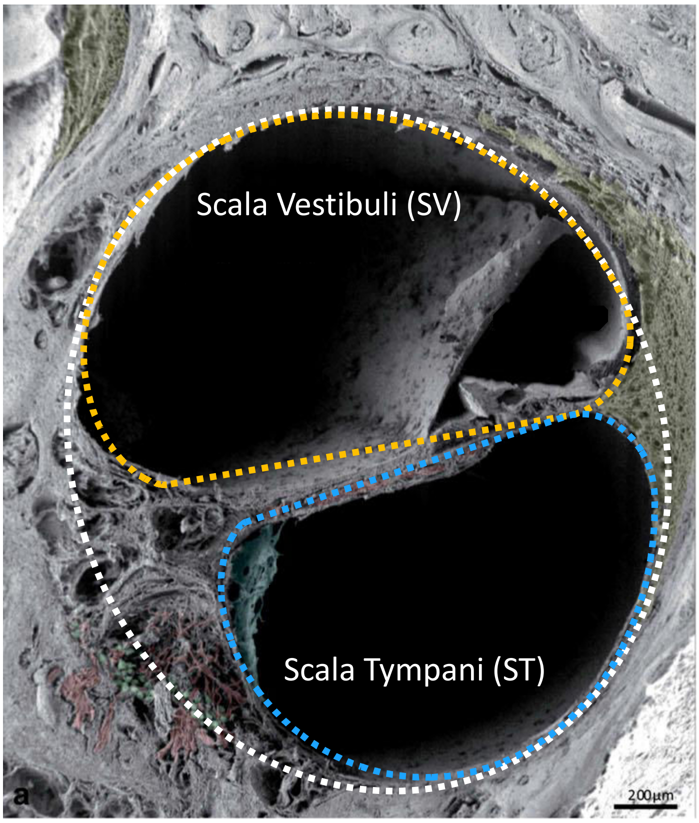


Figure 4.10: Parametric cross-sections (ST in blue, SV in orange, cochlea in white) fitted to a microscopic images from [Rask-Andersen 2012], here $\theta \approx 3\pi/4$.

p	$\overline{\alpha}$	$\overline{w}_{ST}$	$\overline{h}_{ST}$	$\overline{w}_{SV}$	$\overline{h}_{SV}$
0	0.966	2.59	0.952	2.09	1.27
1	-1.15	-0.815	1.15	-0.937	-1.03
2	-0.784	0.545	-1.36	0.593	0.632
3	2.17	-0.196	0.670	-0.165	-0.166
4	-1.60	3.73×10^{-2}	-0.180	2.39×10^{-2}	2.29×10^{-2}
5	0.620	-4.00×10^{-3}	2.95×10^{-2}	-1.94×10^{-3}	-1.80×10^{-3}
6	-0.150	2.52×10^{-4}	-3.08×10^{-3}	8.85×10^{-5}	8.16×10^{-5}
7	2.43×10^{-2}	-9.25×10^{-6}	2.12×10^{-4}	-2.13×10^{-6}	-1.97×10^{-6}
8	-2.76×10^{-3}	1.84×10^{-7}	-9.47×10^{-6}	2.09×10^{-8}	1.97×10^{-8}
9	2.23×10^{-4}	-1.52×10^{-9}	2.67×10^{-7}		
10	-1.31×10^{-5}		-4.29×10^{-9}		
11	5.45×10^{-7}		3.01×10^{-11}		
12	-1.59×10^{-8}				
13	3.08×10^{-10}				
14	-3.56×10^{-12}				
15	1.86×10^{-14}				

Table 4.2: Polynomial coefficients of degree p for cross-section functions $\overline{\alpha}(\theta)$, $\overline{w}_{ST}(\theta)$, $\overline{h}_{ST}(\theta)$, $\overline{w}_{SV}(\theta)$ and $\overline{h}_{SV}(\theta)$

$$\mathrm{d}s(\theta) = \sqrt{\mathrm{d}r^2(\theta) + r^2(\theta)\mathrm{d}\theta^2 + \mathrm{d}z(\theta)} \tag{4.35}$$

$$\mathrm{d}s(\theta) = \sqrt{\left(\frac{\mathrm{d}r(\theta)}{\mathrm{d}\theta}\right)^2 + r^2(\theta) + \left(\frac{\mathrm{d}z(\theta)}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta \tag{4.36}$$

4.2.2.3 Surface Self-intersection Avoidance

Local Self-intersection The generalized cylinders $\mathbf{x}_i(\theta, \phi)$ have local self-intersections if the radius of their cross-sections is greater than the radius of curvature of the axis $\mathbf{x}_c(\theta)$ [Zerroug 1996, Gansca 2002]. More specifically, they have local self-intersections if and only if the following condition is not satisfied:

$$(\mathbf{x}_i(\theta, \phi) - \mathbf{x}_c(\theta)) \cdot \hat{\mathbf{n}}(\theta) \leq \frac{1}{\kappa(\theta)} \tag{4.37}$$

where $\hat{\mathbf{n}}(\theta)$ is the normal unit vector of $\mathbf{x}_c(\theta)$ and $\kappa(\theta)$ its local curvature, such as,

$$\kappa(\theta)\hat{\mathbf{n}}(\theta) = \frac{\mathrm{d}\hat{\mathbf{t}}}{\mathrm{d}s} \tag{4.38}$$

Global Self-intersection Several kinds of global self-intersections are possible but the way the shapes were constructed restricts the set of possible anomalies.

The full parametric shapes $\mathbf{x}_i(\theta,\phi)$ can also be written in a cylindrical coordinate system (in a non-unique manner) as,

$$\mathbf{x}_i(\theta,\phi) = \tilde{r}_i(\theta,\phi)\hat{\mathbf{r}}(\tilde{\theta}_i(\theta,\phi)) + \tilde{z}_i(\theta,\phi)\hat{\mathbf{z}} \tag{4.39}$$

with $\tilde{r}_i(\theta,\phi)$, $\tilde{\theta}_i(\theta,\phi)$ and $\tilde{z}_i(\theta,\phi)$ are defined in $\mathbb{R}$ and calculated as follows,

$$\tilde{\theta}_{0,i}(\theta,\phi) = \operatorname{atan2}(\mathbf{x}_i(\theta,\phi)\cdot\hat{\mathbf{y}}, \mathbf{x}_i(\theta,\phi)\cdot\hat{\mathbf{x}}) \tag{4.40}$$

where the function $\operatorname{atan2}(y,x)$ is the four-quadrant inverse tangent and returns values in $[-\pi,\pi]$ based on the values of y and x.

$$n_i^*(\theta,\phi) = \arg\min_{n\in\mathbb{Z}}\{|\tilde{\theta}_0(\theta,\phi) - \theta + n\pi|\} \tag{4.41}$$

$$\tilde{\theta}_i(\theta,\phi) = \tilde{\theta}_{0,i}(\theta,\phi) + n_i^*(\theta,\phi)\pi \tag{4.42}$$

$$\tilde{r}_i(\theta,\phi) = (-1)^{n_i^*(\theta,\phi)}\sqrt{(\mathbf{x}_i(\theta,\phi)\cdot\hat{\mathbf{x}})^2 + (\mathbf{x}_i(\theta,\phi)\cdot\hat{\mathbf{y}})^2} \tag{4.43}$$

$$\tilde{z}_i(\theta,\phi) = \mathbf{x}_i(\theta,\phi)\cdot\hat{\mathbf{z}} \tag{4.44}$$

There could be global self-intersection between $\{\mathbf{x}_i(\theta,\phi)|\tilde{\theta}_i(\theta,\phi) = x\}$ and $\{\mathbf{x}_i(\theta,\phi)|\tilde{\theta}_i(\theta,\phi) = x + 2\pi\}$, what might be called inter-turn self-intersection.

We introduce,

$$\mathbf{l}(\theta) = \mathbf{x}_c(\theta + 2\pi) - \mathbf{x}_c(\theta) \tag{4.45}$$

$$l(\theta) = \|\mathbf{l}(\theta)\| \qquad\qquad \hat{\mathbf{l}}(\theta) = \frac{\mathbf{l}(\theta)}{l(\theta)} \tag{4.46}$$

To improve the readability $\tilde{\theta}_i(\theta,\phi)$ is now written $\tilde{\theta}_i$. The following inequalities are sufficient conditions to avoid inter-turn self-intersections,

$$\begin{cases}
(\mathbf{x}_{Co}(\theta,\phi) - \mathbf{x}_c(\tilde{\theta}_{Co}))\cdot\hat{\mathbf{l}}(\tilde{\theta}_{Co}) \leq \lambda(\tilde{\theta}_{Co})l(\tilde{\theta}_{Co}) & \text{(4.47a)}\\
(\mathbf{x}_{Co}(\theta,\phi) - \mathbf{x}_c(\tilde{\theta}_{Co}))\cdot\hat{\mathbf{l}}(\tilde{\theta}_{Co} - 2\pi) \geq (\lambda(\tilde{\theta}_{Co} - 2\pi) - 1)l(\tilde{\theta}_{Co} - 2\pi) & \text{(4.47b)}\\
(\mathbf{x}_{SV}(\theta,\phi) - \mathbf{x}_c(\tilde{\theta}_{SV}))\cdot\hat{\mathbf{l}}(\tilde{\theta}_{SV}) \leq \lambda(\tilde{\theta}_{SV})l(\tilde{\theta}_{ST}) & \text{(4.47c)}\\
(\mathbf{x}_{ST}(\theta,\phi) - \mathbf{x}_c(\tilde{\theta}_{ST}))\cdot\hat{\mathbf{l}}(\tilde{\theta}_{ST} - 2\pi) \geq (\lambda(\tilde{\theta}_{ST} - 2\pi) - 1)l(\tilde{\theta}_{ST} - 2\pi) & \text{(4.47d)}
\end{cases}$$

where λ could be any function $\mathbb{R} \to [0,1]$.

In practice,

$$r_{\max}(x) = \max_{\theta,\phi|\tilde{\theta}_{Co}(\theta,\phi)=x}(\mathbf{x}_{Co}(\theta,\phi) - \mathbf{x}_c(x))\cdot\hat{\mathbf{l}}(x) \tag{4.48}$$

$$r_{\min}(x) = \min_{\theta,\phi|\tilde{\theta}_{Co}(\theta,\phi)=x} (\mathbf{x}_{Co}(\theta,\phi) - \mathbf{x}_c(x)) \cdot \hat{\mathbf{l}}(x - 2\pi) \tag{4.49}$$

$$\lambda(x) = \frac{r_{\max}(x)^2 + l(x)^2 - r_{\min}(x + 2\pi)^2}{2l(x)^2} \tag{4.50}$$

Self-intersection correction The conditions to avoid local (Eq. 4.37) and global (Eq. 4.47) self-intersections can be written in a general way as,

$$\mathbf{x} \cdot \hat{\mathbf{u}} \leq f \tag{4.51}$$

where $\mathbf{x}$ parametrizes the shape, $\hat{\mathbf{u}}$ represents any mobile unit vector and f is a scalar function. In order to avoid the self-intersection the following substitution, which satisfied Eq. 4.37 and 4.47, is proposed:

$$\mathbf{x} \leftarrow \frac{\min\{\mathbf{x} \cdot \hat{\mathbf{u}}, f\}}{\mathbf{x} \cdot \hat{\mathbf{u}}} \mathbf{x} \tag{4.52}$$

4.2.3 Parameters Posterior Probability

Given a binary manual segmentation $\mathbf{S}$ of the cochlea from CT imaging, we want to estimate the posterior probability $p(\mathcal{T}|\mathbf{S}) \propto p(\mathbf{S}|\mathcal{T})\ p(\mathcal{T})$ proportional to the product of the likelihood $p(\mathbf{S}|\mathcal{T})$ and the prior $p(\mathcal{T})$.

Likelihood The likelihood measures the discrepancy between the known segmentation $\mathbf{S}$ and the parametric model $\mathcal{M}(\mathcal{T})$. The shape model can be rasterized, we obtain a binary filled image $\mathcal{R}(\mathcal{T})$ which can be compared to the manual segmentation. Note that the rigid transformation is known after the alignment in cochlear coordinate system [Verbist 2010]. The log-likelihood was chosen to be proportional to the negative square Dice index $s_2(\mathcal{R}(\mathcal{T}), \mathbf{S})$ between the rasterized parametric model and the manually segmented cochlea, $p(\mathbf{S}|\mathcal{T}) \propto \exp(-s_2^2(\mathcal{R}(\mathcal{T}), \mathbf{S})/\sigma^2)$. The square Dice allows to further penalize the shape with low Dice index (e.g. less than 0.7) and σ was set to 0.1 after multiple tests as to provide sufficiently spread posterior distribution.

Prior The prior is chosen to be as uninformative as possible while authorizing an efficient stochastic sampling. We chose an uniform prior for all 22 parameters within a carefully chosen range of values. From 5 manually segmented cochlear shapes from 5 μCT images we have extracted the 7 one-dimensional functions of θ modeling the centerline and the cross-sections using a Dijkstra algorithm combined with an active contour estimation. θ was discretized and subsampled 1000 times. The 22 parameters were least-square fit on the subsampled centerline and cochlear points. This has provided us with an histogram of each parameter value from the 5 combined datasets, and eventually the parameter range for the prior was set to the average value plus or minus 3 standard deviations.

Posterior Estimation We use the Metropolis-Hastings Markov Chain Monte-Carlo method for estimating the posterior distribution of the 22 parameters. We choose Gaussian proposal distributions with standard deviations equal to 0.3% of the whole parameter range used in the prior distribution. Since the parameter range is finite, we use a bounce-back projection whenever the random walk leads a parameter to leave this range.

Posterior From μCT Images In μCT images, the scala tympani and vestibuli can be segmented separately as $\mathbf{S}_{ST}$ and $\mathbf{S}_{SV}$ thus requiring a different likelihood function. The 2 scalae generated by the model $\mathcal{M}(\mathcal{T})$ are separately rasterized as $\mathcal{R}_{ST}(\mathcal{T})$ and $\mathcal{R}_{SV}(\mathcal{T})$ and compared to the 2 manual segmentations using a single multi-structure Dice index $s_3(\mathcal{R}_{ST}(\mathcal{T}), \mathcal{R}_{SV}(\mathcal{T}), \mathbf{S}_{ST}, \mathbf{S}_{SV})$. This index is computed as the weighted average of the 2 Dice indices associated with the 2 scalae. The likelihood function is then $p(\mathbf{S}_{ST}, \mathbf{S}_{SV}|\mathcal{T}) \propto \exp(-s_3^2/\sigma^2)$.

4.2.4 Controlling Model Complexity

We want to limit the extent of overestimation of uncertainty induced by our rich parametric model. Therefore, we look at the observability of each parameter through its marginalized posterior distribution $p(\tau_i|\mathbf{S}) = \int \int_{\tau_j \neq \tau_i} p(\mathcal{T}|\mathbf{S})\ d\tau_j$. In an ideal scenario, all model parameters should be observable thus indicating that we have not overparameterized the cochlear shape. Therefore we consider the *information gain* $\mathcal{IG}(\tau_i) = -\int_{\tau_i} p(\tau_i)\ \log p(\tau_i)\ d\tau_i + \int_{\tau_i} p(\tau_i|\mathbf{S})\ \log p(\tau_i|\mathbf{S})\ d\tau_i$ computed as difference of entropy between the prior (uniform) distribution and the marginal posterior distribution. The entropy is estimated by binning the distributions using 256 bins covering the range defined by the uniform prior. A low information gain indicates either that the parameter has no observed influence on the shape or that it is correlated with another set of parameters such that many combinations of them lead to the same shape. To test if we are in the former situation, we simply check if the parameter i decreases significantly the likelihood around the maximum *a posteriori* (MAP) by plotting the probability $p(\tau_i|\mathbf{S}, \mathcal{T}_{-i}^{\mathrm{MAP}})$.

4.2.5 Clinical Metrics

We consider three types of electrodes having the same constant diameter of 0.5 mm. Straight electrodes follow the lateral (outer) wall of the ST, whereas perimodiolar ones follow the modiolar (inner) wall of the ST and mid-scala electrodes are located in the geometric center of the cross-section (see Fig. 4.10).

For a given parameter $\mathcal{T}$ and a certain type of electrode, it is relatively simple to compute its trajectory in the ST, by considering each cross-section of the parametric shape model and positioning the center of the CI relative to the inner and outer wall. Furthermore, the maximum insertion depth of a CI $l^{Max}(\mathcal{T})$ can be computed by the arc length of the curve defined by the locus of the electrode positions and by testing if the inscribed circle of the ST boundaries is larger than the electrode. We propose

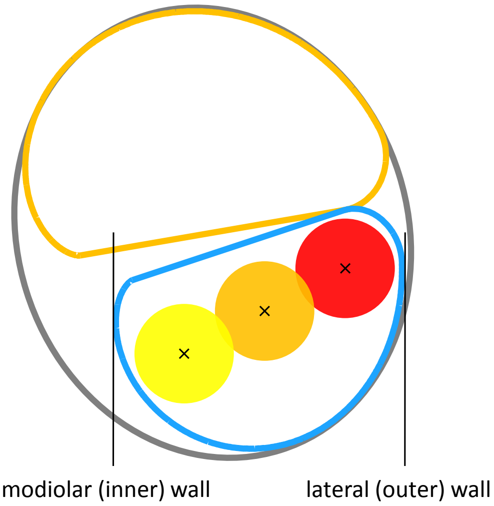


Figure 4.11: Lateral wall (red), mid-scala (orange) and perimodiolar (yellow) positions of a 0.5 mm diameter electrode.

to estimate the posterior probability $p(l^{\text{Max}}|\mathbf{S})$ for each CI type by marginalizing over the set of cochlea parameters : $p(l^{\text{Max}}|\mathbf{S}) = \int_{\mathcal{T}} p(\mathcal{T}|\mathbf{S})\ l^{\text{Max}}(\mathcal{T})\ d\mathcal{T}$. Similarly, we can compute the prior probability of insertion depth which is governed by the prior of the set of parameters : $p(l^{\text{Max}}(\mathcal{T})) = \int_{\tau_i} p(\mathcal{T}) l^{\text{Max}}(\mathcal{T})\ d\tau_i$.

4.3 Results

4.3.1 Model Complexity Evaluation

For each image, 20,000 iterations of the MCMC estimation were performed using a 3.6 GHz Intel Xeon processor machine. The computational time per iteration is less than 4 s for the CT images and less than 20 s for the μCT images. The MCMC mean acceptance rate is 0.38.

The Dice index between the samples corresponding to the maximum *a posteriori* probability (MAP) and the manual segmentations are summarized in Table. 4.3. Note that s_3 indices are lower on μCT because it considers more substructures (ST and SV) than s_2 indices on CT (cochlea only). A careful inspection of the two structures in Fig 4.15 suggests that our parametric model has enough degree of freedom to account the complexity of the cochlear shape. The model even appears to regularize the incomplete manual segmentation without overfitting the noise. The mean surface error between the segmented μCT images and the maximum a posteriori models estimated from segmented CT images is less than 0.3 mm. This error depends on the complexity of the model, the rigid registration and the segmentations (independently performed for each modality) but still comparable with the score of 0.2 mm obtained with statistical shape models for cochlear substructures

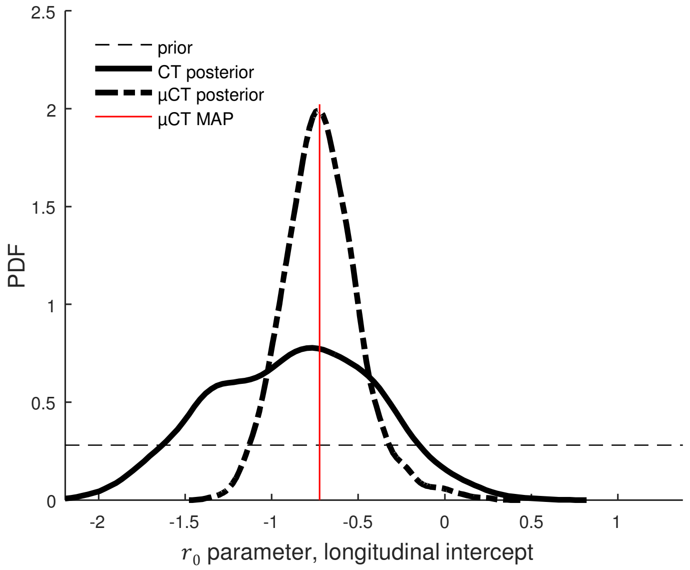


Figure 4.12: Estimated posterior for the parameter r_0 with patient #1. $\mathcal{IG}_{CT} = 0.91, \mathcal{IG}_{\mu CT} = 2.03$.

segmentations in CT [Noble 2011].

On μCT scans, 78% of the cross-sections parameters have an *information gain* greater than 0.1, while the mean information gain over the 22 parameters is $\overline{\mathcal{IG}} = 0.41$. Furthermore, we checked that on μCT scans, for all parameters, any local variation leads to a significant decrease of likelihood $p(\tau_i|\mathbf{S}, \mathcal{T}_{-i}^{\mathrm{MAP}})$ and thus showing an influence on the observed shape (as in Fig. 4.14). This implies that some parameters might be correlated and that shapes may be described by different parameters combinations. Thus we may slightly overestimate the uncertainty (and minimize bias) which is preferable than underestimating it through an oversimplified model. Setting some of those parameters to a constant may be a too strong assumption given that only 9 patient data are considered and therefore we decided to keep the current set of 22 parameters.

On CT scans, 28% of the cross-sections parameters have an *information gain* greater than 0.1 and $\overline{\mathcal{IG}} = 0.23$. The information gain is smaller for CT images than μCT images, which is expected as far less details are visible. In particular, the two scalae are not distinguishable making their model parameters unidentifiable.

4.3.2 CT Uncertainty Evaluation

We evaluate the posterior probability of the maximal insertion depth $p(l^{Max}|\mathbf{S})$ for each patient, modality and electrode design. Their cumulative distribution function (CDF) can be clinically interpreted, as it expresses the probability that the maximal

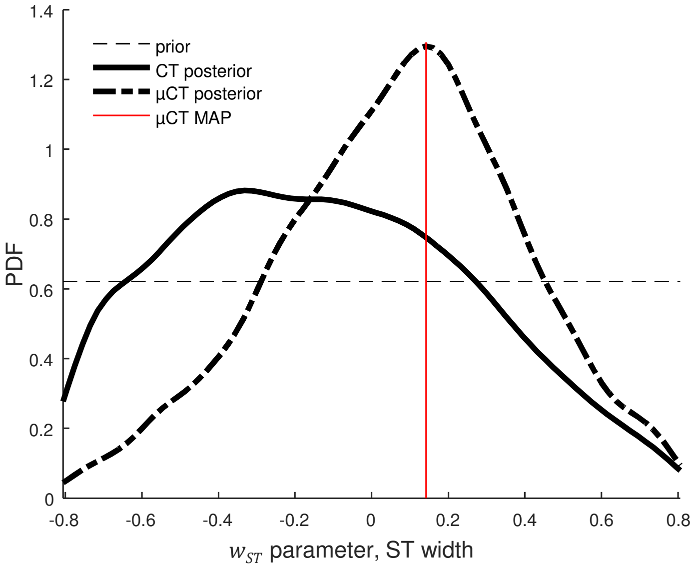


Figure 4.13: Estimated posterior for the parameter w_{ST0} with patient #1. $\mathcal{IG}_{CT} = 0.16, \mathcal{IG}_{\mu CT} = 0.38$.

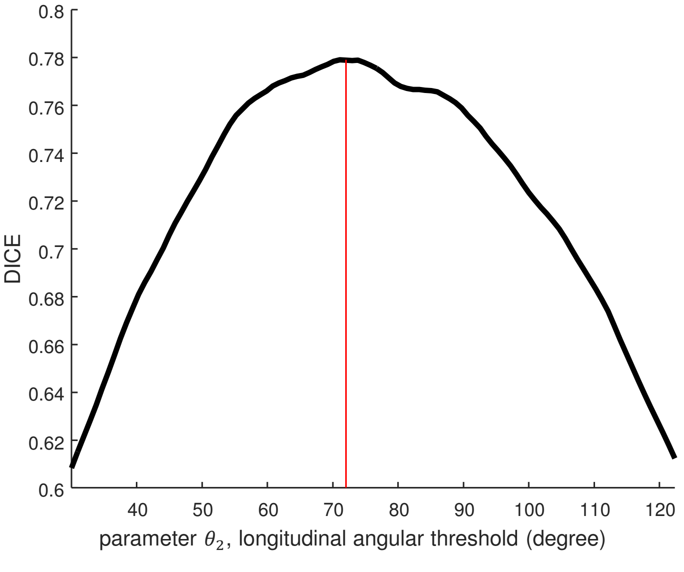


Figure 4.14: Example of $p(\tau_i|\mathbf{S}, \mathcal{T}_{-i}^{\mathrm{MAP}})$ with $i = \theta_2$.

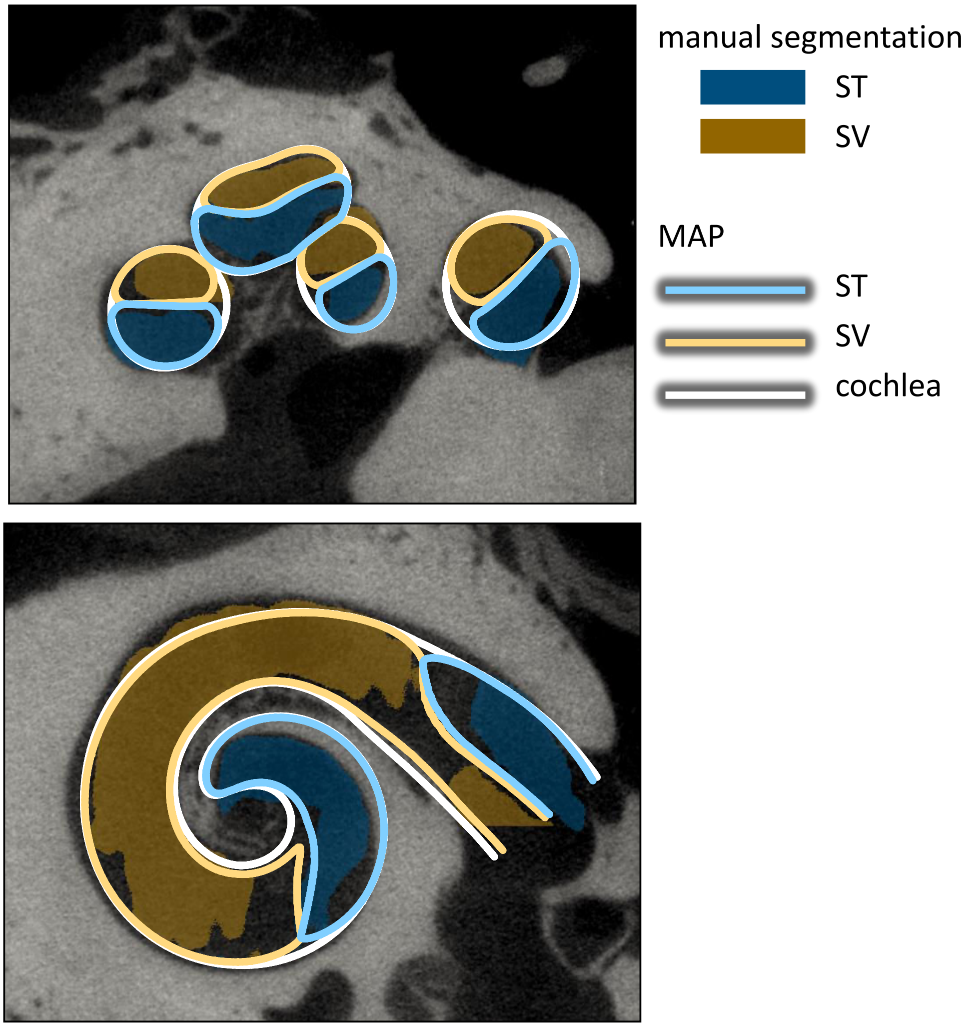


Figure 4.15: Shape models of the cochlea (light line) of the MAP of patient 1 with the segmented ST (blue) and SV (orange) on μCT images.

patient	with μCT	with CT
1	0.78	0.82
2	0.78	0.86
3	0.76	0.75
4	0.79	0.82
5	0.76	0.80
6	0.78	0.79
7	0.75	0.76
8	0.79	0.79
9	0.79	0.79
mean	**0.77**	**0.80**

Table 4.3: Dice indices between the MAP and manual segmentation.

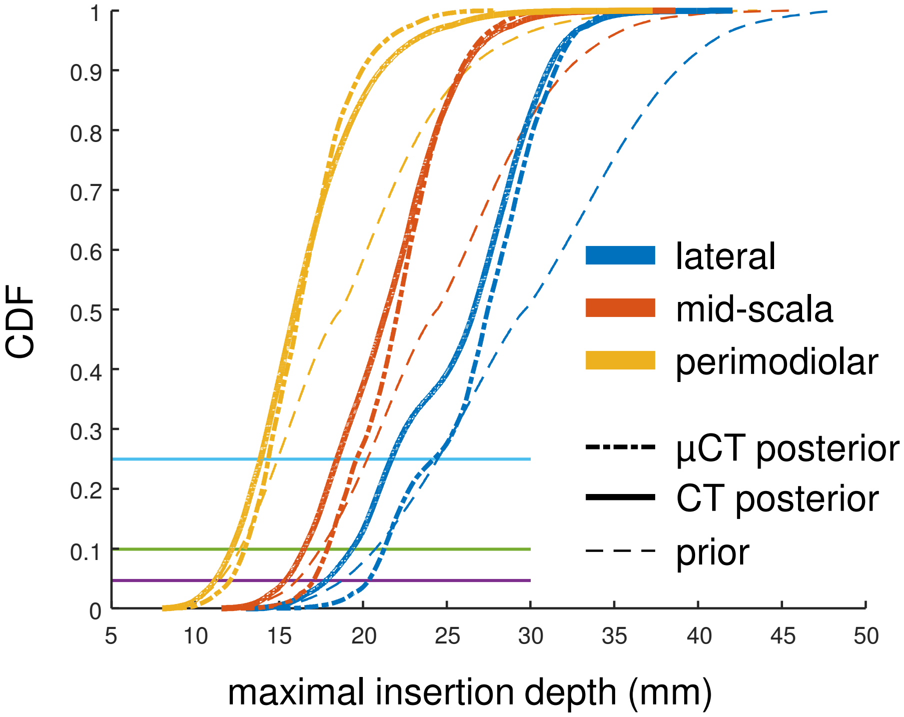


Figure 4.16: CDF of the maximal insertion depth estimation for Patient #1.

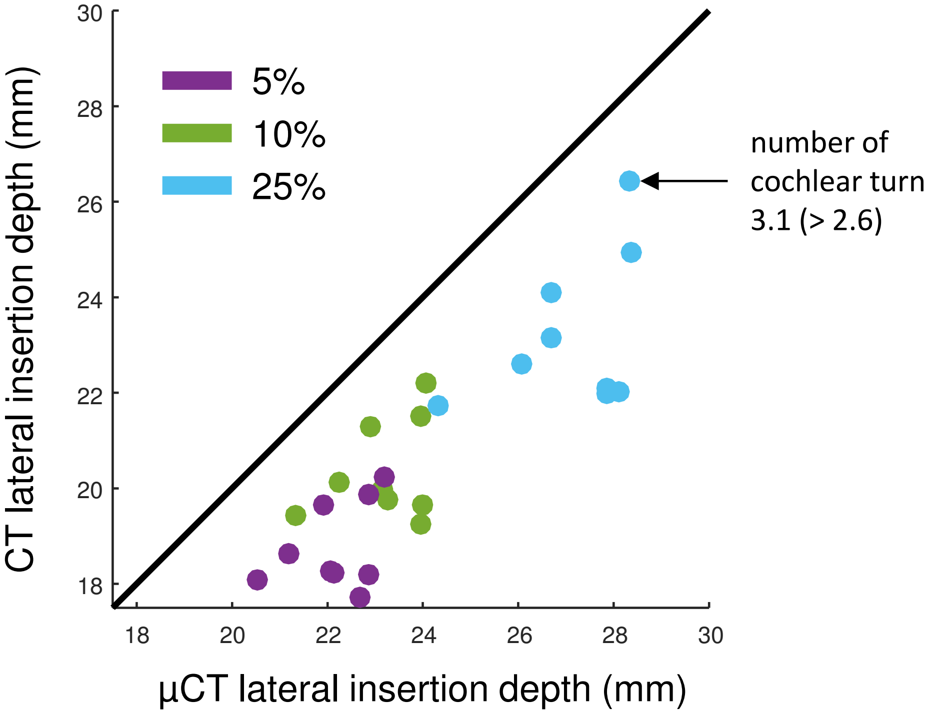


Figure 4.17: Maximal insertion depth estimation discrepancy between CT and μCT for electrodes following lateral wall at different quantiles (5%, 10% and 25%). Note that the lateral position is the least favorable result in terms of discrepancy between modalities (see Table 4.5).

μCT posterior	**CT posterior**	**prior**
3.42	4.14	5.54

Table 4.4: Standard deviation of probability distribution function of the maximum insertion depth averaged over all patients and electrode designs (mm).

lateral	**midscala**	**perimodiolar**
2.34	1.32	0.92

Table 4.5: Discrepancy between CT and μCT averaged over all patients (mm).

insertion depth of a cochlea is less or equal than a given value. Therefore if an electrode has a length l, it also indicates the probability to traumatize the cochlea (if fully inserted). Hence maximal insertion depth corresponding to a CDF of 5%, can be understood as a 95% chance that the electrode actually fits in the ST. The CDF accounts for the uncertainty in the whole shape, including cochlear length or diameter. A cochlea with a longer or larger ST would naturally result in a CDF shifted to the right.

The mean standard deviation of the distributions across the patients and electrode designs (see Table 4.4) shows that uncertainty with CT images is greater than μCT images but still more informative than the prior. To evaluate the bias of maximal insertion depth estimated from CT images we measure the mean discrepancy between the estimation from μCT and CT images. Fig. 4.17 shows the estimation differences between modalities for the worse case, namely straight electrodes. We must stress that all maximal insertion depths are underestimated with CT images. The ST is usually larger than the SV at the first basal turn [Wysocki 1999] and this information is not explicitly embedded in the prior. Since only little cross-section information can be inferred from CT images, we could hypothesize that the diameters of the ST are more likely to be underestimated with CT images, leading to underestimate insertion depth.

4.4 Conclusion

In this study, we have proposed a novel parametric model for detailed cochlea shape reconstruction. We evaluated its complexity in order to optimize the uncertainty quantification of intracochlear shapes from CT images. Based on anatomical considerations, our results introduce a measurements of the risk of trauma given a cochlear design and an insertion depth. Most of the CI have a linear electrode depth between 10 and 30 mm, corresponding to the range within which our results are the most revealing. For this data set, the maximal insertion depth spans a 4 mm range. One cochlea (Patient 4) presents a deeper maximal insertion depth than others, we observed that it had a high number of cochlear turns (3.08 compared to an average of 2.6) which was confirmed by a radiologist on μCT. This exemplifies the importance of providing a patient-specific estimation of the maximal insertion depth.

Our experiments show that under the best possible conditions (careful image segmentation, stochastic sampling of a detailed cochlear model), classical preoperative CT images could be used by ENT surgeons to safely select a patient-specific CI. Indeed, the discrepancy is limited (maximum of 2.34 mm for the lateral position) and always lead to an underestimation of the maximal insertion depth from CT images which is more safe for the patient. In future work, more data will be considered to improve the correlation between CT and μCT predictions and to estimate more thoroughly the bias between both modalities in order to apply a correction.

Chapter 5

Joint Shape and Intensity Model-Based Segmentation of CT Images of the Cochlea

---

Contents

Abstract

Cochlear implants (CI) are used to treat severe hearing loss by surgically inserting an electrode array into the cochlea. Cochlear internal cavities have complex spiraling shapes and are nearly invisible with clinically available CT scans due to the relative small size of the cochlea with respect to the scanner resolution. In this chapter we propose a joint model of the cochlear shape (and its substructures) model and its appearance within a generative probabilistic Bayesian framework. The proposed

segmentation method performed on CT images is compared to high-resolution manually segmented μCT images and applied to a large database of 987 CT images. This allows the statistical characterization of the cochlear anatomical variability along with the quantification of the bilateral symmetry.

5.1 Introduction

Manual segmentations are expensive and time consuming to produce, thus a fully-automated segmentation method is preferable. The segmentation of intracochlear structures represents a great challenge. CT images of the temporal bone have low resolution with respect to the small size of the anatomical structures and the topology of the cochlea is complex. The cochlea measures about 8.5x7x5 mm^3 while the typical CT image spacing is larger than 0.2 mm. The main chambers of the cochlea, the scala tympani and the scala vestibuli, form a nested double helix where the spatial division between each turn (basal, middle and apical) is weakly visible and where the stiff structural elements separating the scala tympani and the scala vestibuli, such as the basilar membrane and the spiral lamina ossea are weakly visible or invisible. Finally the cochlea is filled with fluids which can be similarly found in the vestibular system and other neighboring structures, with similar appearance in CT images.

This challenging task requires additional information to achieve successful segmentation. Anatomical atlases and shape models can be combined with intensity information to provide sufficient prior knowledge. While those methods are commonly available for major organs such as the brain or the heart, few cochlear shape models have been developed.

5.1.1 Detailed Cochlear Shape Model Fitting from CT Images

Cochlear Model Fitting In order to fit cochlea models to CT images, two main approaches can be summarized as follows.

A first set of methods consists in an optimization procedure to perform parametric model fitting [Baker 2005]. A parametric model-image registration procedure is used to align the model surface with the salient image features, for instance captured by high intensity gradient voxels.

A second set of methods consists in a non rigid registration procedure between an input CT image and an atlas embedding statistical shape prior [Noble 2010, Noble 2011, Reda 2011, Noble 2012, Reda 2012, Noble 2013, Reda 2013, Kjer 2014a, Reda 2014a, Reda 2014b, Kjer 2015d]. Anatomical models are extracted from high-resolution datasets, usually ex-vivo μCT scans. The shape prior is provided by a statistical shape model (SSM) of the intra-cochlear anatomy, which guides the registration between a CT scan and an enhanced CT-atlas.

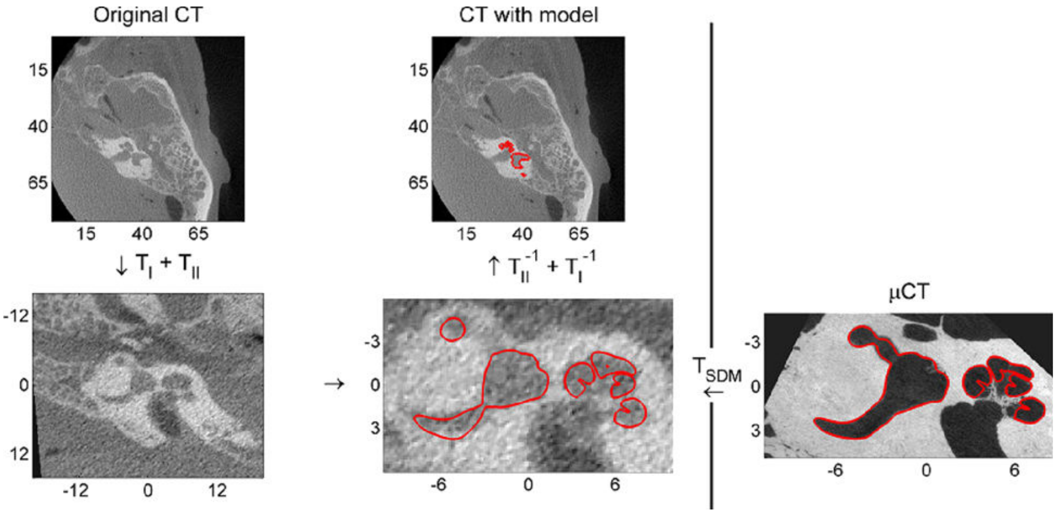


Figure 5.1: Example of labyrinth segmentation with SSM (from [Kjer 2015d]).

5.1.2 Joint Shape and Intensity Model for Segmentation

Anatomical structures defined by weakly visible boundaries in medical images are often segmented by defining an appearance model of the structures of interest and by encoding the spatial prior information [Pohl 2006a, Heimann 2007]. To do so, a strong shape model needs to be **realistic** and to provides good **generalization**, i.e. the ability to represent realistically shapes that are not present in the training set. Most of the cochlear segmentation algorithms incorporating shape prior use Statistical Shape Models (SSM) [Noble 2010, Noble 2011, Reda 2011, Noble 2012, Reda 2012, Noble 2013, Reda 2013, Kjer 2014a, Reda 2014a, Reda 2014b, Kjer 2015d] as introduced by [Cootes 1995]. However when the number of training observations is limited (in our case $N = 9$), the shape variability might be too constrained and the model may not generalize well enough. Alternatively to SSM, an analytic parametric shape model can be defined “manually” but may also be considered too simple and not realistic enough since it creates an idealized shape model with few parameters. As highlighted in section 4.3.1 there is an inherent problem in shape fitting, since the number of degrees of freedom of SMM or parametric models may be arbitrarily increased, but possibly leading to the issue of overparameterization. In this case, the parameters are difficult to estimate uniquely because they are correlated and are not easily interpretable.

To overcome this limitation due to a low number of training shapes, we propose **to move away from a simple model fitting** approach where the resulting shape is an instance of a compact parametric model. Instead, we propose to combine this parametric model with an appearance model such that the shape model constrains the output segmentation without restricting it to a low dimensional space.

The goal is to unify the cochlear shape (and its substructures) model fitting with a local appearance model within a Bayesian framework. Using a generative proba-

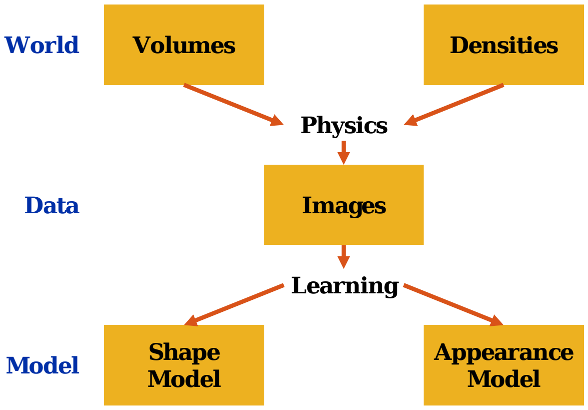


Figure 5.2: Modeling the contingency of CT images through shape and appearance model (adapted from [Eslami 2013a]).

bilistic model, which describes the label of each voxel having shape and appearance parameters as hidden variables, the problem is formulated as an incomplete data problem marginalizing over all possible labels for which the maximum *a posteriori* is estimated. The proposed method estimates jointly the shape and appearance parameters and applies an iterative expectation-maximization (EM) strategy that interleaves shape model parameters fitting and image segmentation with a mixture of Student's t-distributions.

Our approach relies on a generative probabilistic model as it is a natural way to describe the image content. As such, the estimated parameters and variables have a clear interpretation and the underlying hypothesis are well understood. As our approach alternates between the optimization of shape and intensity parameters, it is related to prior work on joint segmentation and registration approaches.

Related Previous Work on Segmentation and Registration Object segmentation can be performed using generative probabilistic models and exhibit good generalization [Eslami 2013b]. In medical images analysis, most of the studies have focused on generative models for brain imaging segmentation [Wyatt 2003, Pohl 2005, Ashburner 2005, Riklin-Raviv 2009, Patenaude 2011, Wu 2013, Puonti 2016] In neuroimaging studies, several challenges must be addressed such as the presence of a bias field or the presence of brain lesions, thus intensity of magnetic resonance images (MRI) needs often to be corrected and brain tumors need to be specifically handled differently than healthy tissues. Several articles have demonstrated that by using an approach that exploits and solves different problems of neuroimaging in a joint way could improve the results as opposed to performing them sequentially. We can cite for example, joint image registration, tissue classification and bias correction [Ashburner 2005], joint segmentation and registration [Pohl 2006a, Riklin-Raviv 2009], joint registration and estimations of tumor displacement [Gooya 2011] or joint label fusion and multi-atlas registration [Wu 2013].

Joint segmentation and registration methods benefit from the fact that given a registered label image, the segmentation task becomes an easier problem to solve and conversely, given a pair of segmented images the registration task is simplified. The idea behind the combined method presented in [Pohl 2006a, Riklin-Raviv 2009] is to register an atlas to an image and simultaneously estimate the contour of the structures of interest. This requires first to model within a Bayesian framework the relationship between the segmentation, the image data and the registration parameters.

Contribution The aforementioned methods the shape constraints were modeled from a multi-atlas learned from annotated training images and deformed onto the target image using an image registration method. In contrast to these methods, in our approach the shape constraints are described by the parameters of a parametric shape model. Different deformations were previously used: rigid transformations [Wyatt 2003], hierarchical (global and structure-dependent) affine deformations [Pohl 2005], B-spline [Riklin-Raviv 2009] or diffeomorphic Demons [Wu 2013]. It has been shown that due to its complex spiral shape, natural shape variability of the cochlea such as deformations under local torque forces are very difficult to obtained with classical deformation methods such as B-spline [Kjer 2015a]. We tackle this challenge by using a parametric shape model defined in a cylindrical coordinate system that is specific to cochlear structures.

Since the proposed segmentation method is not an instance of the shape but a combination of shape and intensity priors, a new cochlear parameterization in a very low-dimensional space is introduced. This compact shape model addresses the problem of shape and pose parameters correlation by eliminating most shape parameters strongly correlated to the pose and by estimating simultaneously the pose and the shape. An underparameterization model is sufficient because the local shape variability can be taken into account by the intensity model.

Furthermore, instead of using a Gaussian mixture for the intensity probability distribution [Pohl 2006a, Patenaude 2011, Puonti 2016], we used a mixture of Student's t-distributions in order to make the intensity model estimation more robust to image artifacts (heavy-tailed distributions).

5.2 Material and Methods

5.2.1 Overview

A graphical abstract of the method developed for intracochlear structures segmentation is shown in Fig. 5.3.

The rigid registration method used to orient the cochlea in its standard coordinate system will be first presented. The hypothesis of the generative probabilistic model and then the inference method will be introduced in a generic way for multi-class segmentation. The specific approach for cochlear segmentation will be described in the following order: shape modeling, intensity modeling and final hard segmentation. Finally, the implementation details will be further presented.

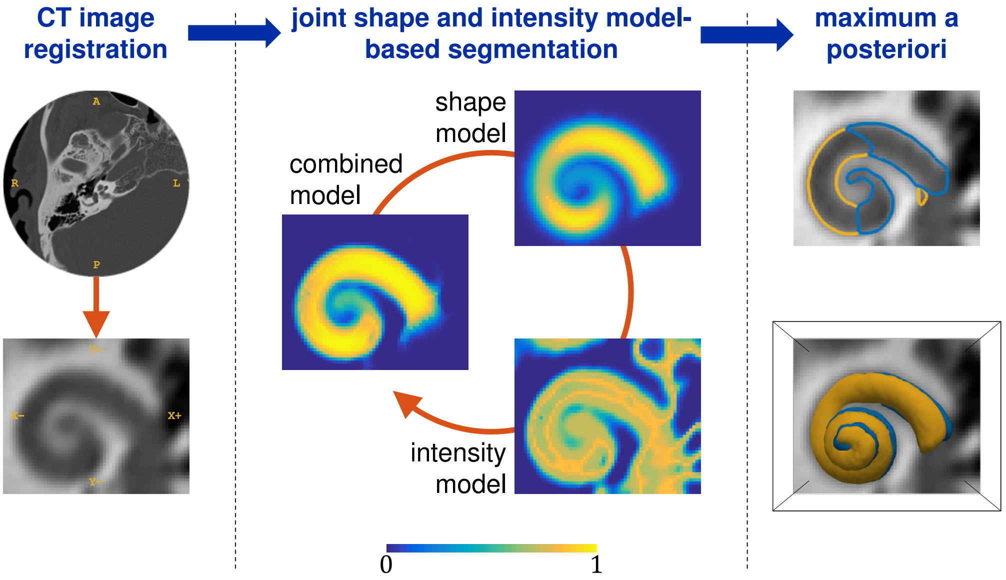


Figure 5.3: Overview of the method. Rigid registration is first performed in order to orient the cochlea in the cochlear coordinate system centered on a small region of interest. Intra-cochlear structures probabilities are estimated by iteratively optimizing the shape and intensity model parameters so that it minimizes the discrepancy between the posterior probability of the label knowing the shape and intensity (combined model) and the probability knowing the shape only (shape model) on one hand, and knowing the intensity only (intensity model) on the other hand. Eventually hard segmentation is performed with the maximum a posteriori (MAP) of intracochlear structures probabilities. Resulting isocontours and isosurfaces are represented.

5.2.2 Data

Two datasets were available. **Dataset #1** consists in 9 cadaveric temporal bones (4 left and 5 right sides) with low-resolution CT and high-resolution μCT acquisitions performed in inconsistent radiological planes. The images, noted $I_i^{\#1,\mathrm{CT}}$ and $I_i^{\#1,\mu\mathrm{CT}}$, for $i = 1, \ldots, 9$, are segmented. Indeed, a head and neck imaging expert produced manual segmentations of the scala tympani and the scala vestibuli using interactive segmentation software ITK-Snap [Yushkevich 2006] and GeoS [Criminisi 2008].

Dataset #2 consists in 987 clinical CT images (452 left and 535 right sides) of temporal bones acquired in standard imaging plane, noted $I_i^{\#2}$ for $i = 1, \ldots, 987$.

All CT scans were acquired at the Imaging Center of the Nice University Hospital using a GE LightSpeed VCT CT system under the supervision of Pr Charles Raffaelli. All μCT scans were acquired in the laboratory Transporters, Imaging and Radiotherapy in Oncology (TIRO, UMR-E4320) in Nice using the GE eXplore speCZT120 scanner under the supervision of Pr Thierry Pourcher.

5.2.3 Rigid Registration

We seek to initialize the segmentation by aligning roughly all images in the same orientation.

5.2.3.1 Cochlear Coordinate System

The Stenvers' plane [Lane 2015] is an oblique radiological reconstruction plane that is in the plane of the superior semicircular canal (also known as the short axis of the temporal bone). The Stenvers' plane is also roughly aligned with the cochlea coordinate system, defined by 16 cochlear implantation experts [Verbist 2010], which includes an origin in addition to an orientation.

The cochlear coordinate system (CCS) appears as a consensual choice between the radiologists and otologists, and we seek to aligned all CT and μCT images in the coordinate system where the z-axis coincides with the modiolar axis, the xz-plane passes through the center of the round window (Fig. 3.13).

5.2.3.2 Temporal Bone Image Registration

Image registration is a well known challenge in medical image analysis and requires to define several main components: a transformation model (rigid or non-rigid), a cost function (similarity metric and regularization), and an optimization method. For a comprehensive reviews for theory and application the reader could refer to [Maintz 1998, Sotiras 2013]. Registration methods depend on the modality and they have been applied to temporal bone CT and μCT images [Whiting 2001, Kjer 2015b, Kjer 2015a, Kjer 2015c], to pre- and intra-operative CT images [Reda 2012] and to fuse complementary information of CT and MRI [Bartling 2005, Neri 2005].

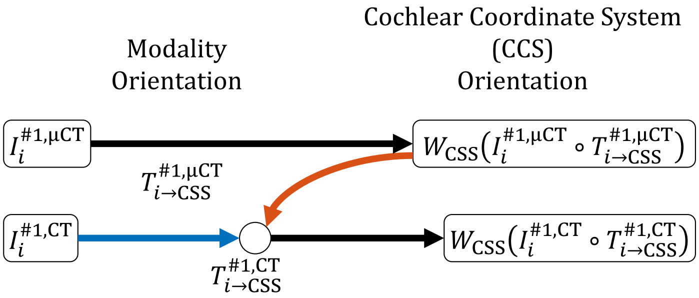


Figure 5.4: Summary of the registration procedures applied to dataset #1. Rounded rectangles represent images, black arrows represent rigid transformation T, connectors ○ represent image registration where the red arrow indicates the fixed image and the blue arrow the moving image.

The goal here is to estimate and analyze the cochlear shape variability. To preserve the global anatomy, only rigid registration is considered. Unlike the cochlea, the temporal bone and its pneumatization show a very large interpatient variability. Pneumatization consists in air cells in the temporal bone that have different patterns and sizes based on environmental and genetic factors. It is initiated after birth and is completed during childhood. To avoid the potential nuisance caused by the temporal bone high variability, we apply our image processing pipeline on a small region of interest (ROI) around the cochlea.

5.2.3.3 Dataset #1 Registration

From the manually segmented μCT images, we write $T_{i\to \text{CSS}}^{\#1,\mu\text{CT}}$ the rigid transformation retrieved from the **robust modiolar axis estimation** method (subsection 3.2.4), such that $I_i^{\#1,\mu\text{CT}} \circ T_{i\to \text{CSS}}^{\#1,\mu\text{CT}}$ corresponds to the μCT images in the cochlear coordinate system. All applied rigid transforms are resampled using trilinear interpolation.

We also define $W_{\text{CSS}}(\cdot)$, as the function that resamples and crops an image so that the spacing is $[0.2, 0.2, 0.2]$ mm, the physical extent is 12x10x10 mm^3 and the offset is $[-4.8, -4, -4]$ mm, corresponding to a bounding box around the cochlea known to measure roughly 8.5x7x5 mm^3 [Dimopoulos 1990].

Multimodal image registration can be performed between CT, $I_i^{\#1,\text{CT}}$, and μCT images from the same temporal bone. We set $T_{i\to \text{CSS}}^{\#1,\text{CT}}$ as,

$$T_{i\to \text{CSS}}^{\#1,\text{CT}} = \arg\min_{T\in SE(3)} \sum \left(W_{\text{CSS}}\left(I_i^{\#1,\mu\text{CT}} \circ T_{i\to \text{CSS}}^{\#1,\text{CT}}\right) - I_i^{\#1,\mu\text{CT}} \circ T\right)^2 , \tag{5.1}$$

where $SE(3)$ is the space of rigid transformations. The downsampled μCT image is very similar to the CT image since they are acquired on the same patient, have the same spacing and the intensities correspond to the Hounsfield unit (HU). For this reason, the sum of squared differences (SDD) is used as the dissimilarity measure. The registration were performed using the block-matching algorithm called Baloo

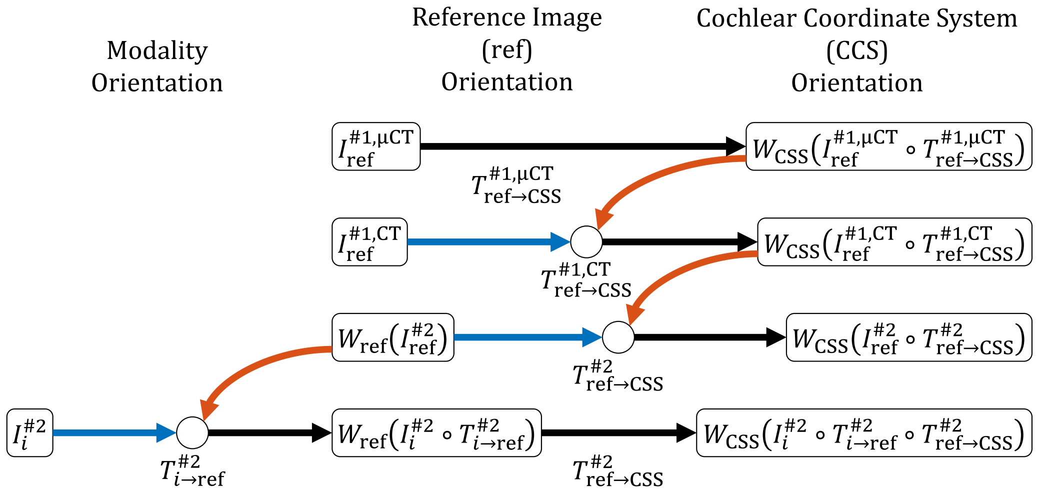


Figure 5.5: Summary of the registration procedures applied to dataset #2. Rounded rectangles represent images, black arrows represent rigid transformation T, connectors ○ represent image registration where the red arrow indicates the fixed image and the blue arrow the moving image.

[Ourselin 2000] from the software MedInria [Toussaint 2007]. Fig. 5.4 summarized the registration procedures applied to the dataset #1.

5.2.3.4 Dataset #2 Registration

No manual segmentation was available to register dataset #2. The transformations from the dataset #1 (defined with the robust modiolar axis estimation) can therefore be used. A reference couple of CT and μCT images, $I_{\text{ref}}^{\#1,\text{CT}}$ and $I_{\text{ref}}^{\#1,\mu\text{CT}}$ was chosen as the medoid with respect to the intensity sum squared difference, i.e the image whose average dissimilarity with all images in the dataset #1 is minimal,

$$\begin{aligned}\text{ref}^{\#1} = \arg\min_{i=1,\dots,9} \sum_{j=1}^{9}\sum_{n=1}^{N} \Big(& W_{\text{CSS}}\left(I_i^{\#1,\text{CT}} \circ T_{i\to\text{CSS}}^{\#1,\text{CT}}\right)(n) \\ & -W_{\text{CSS}}\left(I_j^{\#1,\text{CT}} \circ T_{j\to\text{CSS}}^{\#1,\text{CT}}\right)(n)\Big)^2 ,\end{aligned} \tag{5.2}$$

where $I(n)$ represents the intensity of voxel n in image I. Similarly we choose a reference image from the dataset #2, $I_{\text{ref}}^{\#2}$, in order estimate only once the transformation (noted $T_{\text{ref}\to\text{CSS}}^{\#2}$) between the standard imaging plane in which all images $I_i^{\#2}$ from dataset #2 have been acquired and the cochlear coordinate system. The reference image was chosen, once again, as the medoid image among the dataset #2,

$$\text{ref}^{\#2} = \arg\min_{i=1,\dots,987} \sum_{j=1}^{987}\sum_{n=1}^{N} \left(I_i^{\#2}(n) - I_j^{\#2}(n)\right)^2 . \tag{5.3}$$

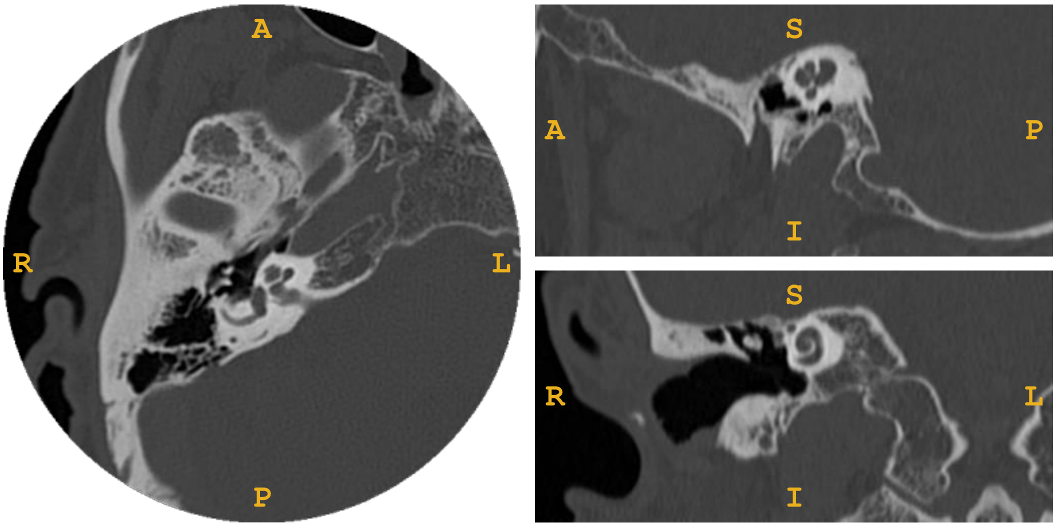


Figure 5.6: Image $I^{\#2}_{31}$ from the dataset #2 shown in the axial, sagittal and coronal planes

Since, the field of view of $I^{\#2}_{\text{ref}}$ is much larger than the resampled and cropped version of $I^{\#1,\text{CT}}_{i} \circ T^{\#1,\text{CT}}_{i\to\text{CSS}}$ (typically 512x512x200 compared to 60x50x50 voxels), we initialized the transformation $T^{\#2}_{\text{ref}\to\text{CSS}}$ by manually selecting the center of the cochlea. We define another ROI around the cochlea with the function $W_{\text{ref}}(\cdot)$ which creates an image with a spacing of $[0.1875, 0.1875, 0.25]$ mm (corresponding to the typical spacing of a temporal bone CT image), a physical extent of 30x30x30 mm^3 and an offset of $[-15, -15, -15]$ mm. The rigid transformation was estimated with a registration procedure minimizing the following criteria,

$$T^{\#2}_{\text{ref}\to\text{CSS}} = \arg\min_{T\in SE(3)} NC\left(W_{\text{CSS}}\left(I^{\#1,\text{CT}}_{\text{ref}} \circ T^{\#1,\text{CT}}_{\text{ref}\to\text{CSS}}\right), W_{\text{ref}}\left(I^{\#2}_{\text{ref}}\right) \circ T\right) , \tag{5.4}$$

where the similarity criterion NC is the normalized correlation criterion (as suggested by [Jenkinson 2001] for monomodal image registration),

$$NC(X, Y) = \frac{\sum(XY)}{\sqrt{\sum X^2}\sqrt{\sum Y^2}} . \tag{5.5}$$

For every CT images $I^{\#2}_{i}$ for $i = 1, \ldots, 987$ we performed a rigid registration using $I^{\#2}_{\text{ref}}$ as a reference image. This has the advantage of reducing the registration parameter space to explore since all images are relatively “close”. The transformation is then estimated as,

$$T^{\#2}_{i\to\text{ref}} = \arg\min_{T\in SE(3)} NC\left(W_{\text{ref}}\left(I^{\#2}_{\text{ref}}\right), I^{\#2}_{i} \circ T\right) . \tag{5.6}$$

The registration were performed using the FMRIB’s Linear Image Registration Tool from the FSL library [Jenkinson 2001, Jenkinson 2002].

Eventually, images $I^{\#2}_{i} \circ T^{\#2}_{i\to\text{ref}} \circ T^{\#2}_{\text{ref}\to\text{CSS}}$ are expressed in the cochlear coordinate system. Fig. 5.5 summarizes the registration procedures applied to dataset #2.

5.2.3.5 Laterality

The presented pipeline is valid for right ears. An equivalent pipeline was applied to the left ears with a different reference image $I_{\text{ref}}^{\#2}$ (chosen as the medoid among the left ears) except that the images were first flipped along the x-axis.

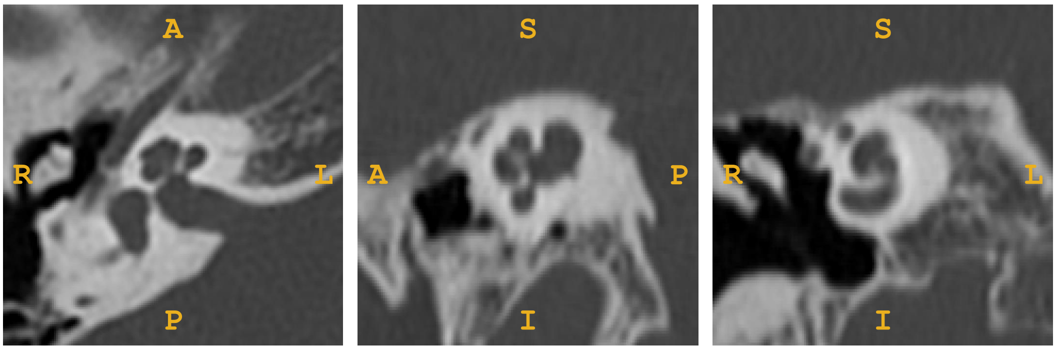


Figure 5.7: $W_{\text{ref}}\left(I_{31}^{\#2} \circ T_{31\to\text{ref}}^{\#2}\right)$, image $I_{31}^{\#2}$ from the dataset #2 registered to the reference image $I_{\text{ref}}^{\#2}$ cropped around the cochlea in a 30x30x30 mm^3 box, shown in the axial, sagittal and coronal planes

5.2.4 Generic Multi-class Approach for Joint Shape and Intensity Model-based Segmentation

5.2.4.1 Model Hypothesis

Shape Model Fitting Classical multi-class parametric shape model fitting consists in the estimation of θ_S the parameters of the shape model S in order to fit an observed image I. A graphical representation of shape fitting is shown in Fig. 5.9. Let Z_n be the tissue class of the voxels n, such as $Z_n \in \{z_i\}_{i=1,\ldots,L}$, with L different classes (for example, z_1 for the foreground and z_2 for the background). We write

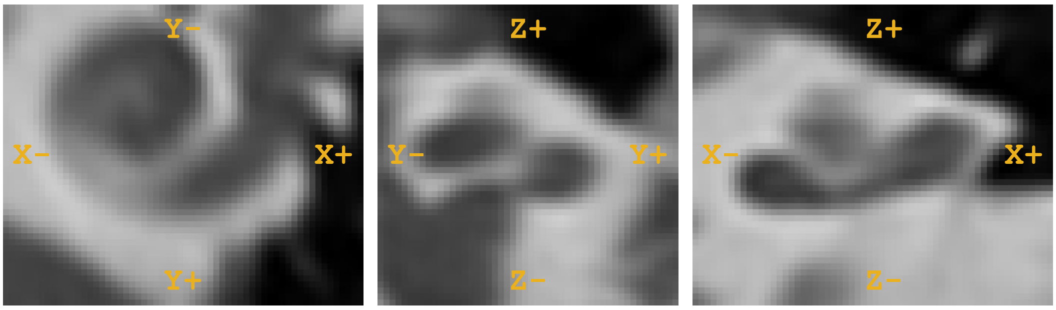


Figure 5.8: $W_{\text{CSS}}\left(I_31^{\#2} \circ T_{31\to\text{ref}}^{\#2} \circ T_{\text{ref}\to\text{CSS}}^{\#2}\right)$, image $I_{31}^{\#2}$ from the dataset #2 aligned in the cochlear coordinate system cropped around the cochlea in a 12x10x10 mm^3 box

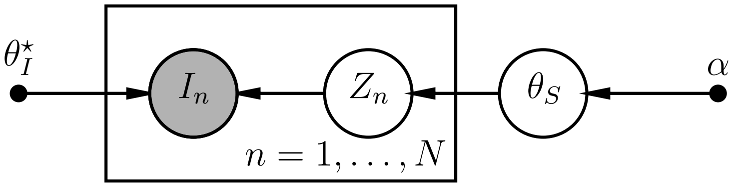


Figure 5.9: Graphical model for the for shape fitting. In this model, I_n the intensity of the voxel for $n = 1, \ldots, N$ is linked to the latent label Z_n and the intensity model parameter $\theta_I^\star$. The label Z_n is linked to shape model parameter θ_S. The observed (known) variables are shaded.

$Z = \{Z_n\}_{n=1,\ldots,N}$ as the set of voxel categorical variables. It is assumed that the N voxels of the image I are conditionally independent given Z_n.

A classical approach is to estimate the maximum of the likelihood $\mathcal{L}(\theta_S|I)$ of fitting a shape model S parameterized by θ_S given the image I (ML), or the maximum *a posteriori* (MAP)

$$\begin{aligned}
\hat{\theta}_S^{\text{ML}} &= \underset{\theta_S}{\operatorname{argmax}}(\mathcal{L}(\theta_S|I)) \\
&= \underset{\theta_S}{\operatorname{argmax}}(p(I|\theta_S)) \\
\hat{\theta}_S^{\text{MAP}} &= \underset{\theta_S}{\operatorname{argmax}}(p(\theta_S|I)) \\
&= \underset{\theta_S}{\operatorname{argmax}}(p(I|\theta_S)p(\theta_S))\,.
\end{aligned} \tag{5.7}$$

Given $\hat{\theta}_S$ the final segmentation can be obtained by assigning the class with the maximum probability.

Joint Shape and Intensity Modeling In order to perform joint shape and intensity model-based segmentation, the intensity model parameter $\theta_I^\star$ is no longer considered as a model parameter but a random variable θ_I. We introduce the model whose graphical representation is shown in Fig. 5.10.

Following the conditional dependence between the variables, the joint probability of the observed and hidden variables $p(\theta_S, \theta_I, Z_n, I_n)$ factorizes as,

$$p(\theta_S, \theta_I, Z_n, I_n) = p(I_n|Z_n, \theta_I)p(\theta_I|\beta)p(Z_n|\theta_S)p(\theta_S|\alpha)\,. \tag{5.8}$$

Let us paraphrase the following probabilities:

$$p(Z_n = z_i|\theta_S) \tag{5.9}$$

is the probability that the voxel n belongs to the label z_i knowing the shape parameters θ_S. Note that we extend the definition of the shape to include its pose.

$$p(I_n|Z_n = z_i, \theta_I) \tag{5.10}$$

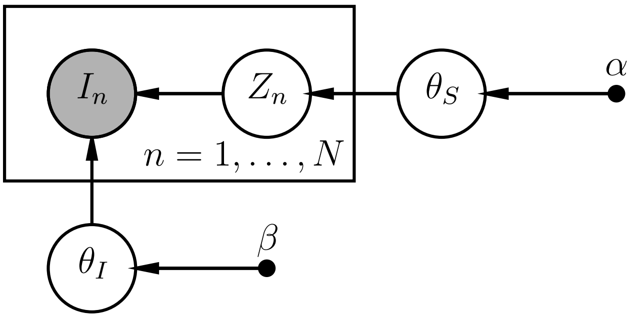


Figure 5.10: Graphical model for the generic multi-class approach for joint shape and intensity model-based segmentation. In this model, I_n the intensity of the voxel $n = 1, \ldots, N$ is linked to the latent label Z_n and the intensity model parameter θ_I. The label Z_n is linked to shape model parameter θ_S. The observed (known) variables are shaded.

is the probability of observing intensity I_n knowing the label $Z_n = z_i$ and the parameters that characterize the intensity of this label,

We seek to estimate the posterior probability that the voxel n belongs to the label z_i,

$$p(Z_n = z_i | \theta_S, \theta_I, I_n)\,, \tag{5.11}$$

which can also be rewritten following Bayes rules,

$$p(Z_n = z_i | \theta_S, \theta_I, I_n) = \frac{p(I_n | Z_n = z_i, \theta_I) p(Z_n = z_i | \theta_S)}{\sum_{i=1}^{L} p(I_n | Z_n = z_i, \theta_I) p(Z_n = z_i | \theta_S)}\,. \tag{5.12}$$

5.2.4.2 Model Inference

We seek to estimate the MAP, that is to find the optimal parameters $\hat{\theta}_S$ and $\hat{\theta}_I$ which maximize the following log-posterior probability,

$$\begin{aligned}\log p(\theta_S, \theta_I | I) = & \sum_{n=1}^{N} \log \left(\sum_{i=1}^{L} p(I_n | Z_n = z_i, \theta_I) p(Z_n = z_i | \theta_S) \right) \\ & + N \left(\log p(\theta_S | \alpha) + \log p(\theta_I | \beta) \right) + \text{cst}\,,\end{aligned} \tag{5.13}$$

where the first term corresponds to the log-likelihood, $\log \mathcal{L}$, and the second term corresponds to the log-prior and cst is a normalization constant corresponding to the log-evidence term. Since the log-likelihood is intractable because of the logarithm of a sum, we use the **expectation maximization** (EM) algorithm which takes an alternative strategy. Instead of considering the log-likelihood which is difficult because it involves the latent variable Z, we consider the complete log-likelihood

$\log \mathcal{L}_c$ as if Z were observable,

$$\begin{aligned}\log \mathcal{L}_c(\theta_S, \theta_I) &= \log p(I, Z|\theta_S, \theta_I)\\ &= \sum_{n=1}^{N} \log p(I_n, Z_n|\theta_S, \theta_I)\\ &= \sum_{n=1}^{N} \log p(I_n|Z_n, \theta_I) + \log p(Z_n|\theta_S)\,.\end{aligned} \tag{5.14}$$

Let $\theta = \{\theta_S, \theta_I\}$ be the whole set of shape and intensity parameters. Since Z is hidden, the complete log-likelihood cannot be evaluated directly, therefore we consider its conditional expectation (under the posterior distribution of Z), noted Q

$$\begin{aligned}Q(\theta, \theta^{[t]}) &= \mathrm{E}_Z[\log \mathcal{L}_c(\theta)|I, \theta^{[t]}]\\ &= \sum_{n=1}^{N}\sum_{i=1}^{L} p(Z_n = z_i|\theta^{[t]}, I_n)\\ &\qquad (\log p(I_n|Z_n = z_i, \theta_I) + \log p(Z_n = z_i|\theta_S))\\ &\leq \sum_{n=1}^{N}\sum_{i=1}^{L} (\log p(I_n|Z_n = z_i, \theta_I) + \log p(Z_n = z_i|\theta_S))\\ &\leq \sum_{n=1}^{N} \log\left(\sum_{i=1}^{L} p(I_n|Z_n = z_i, \theta_I) p(Z_n = z_i|\theta_S)\right) = \log \mathcal{L}\,.\end{aligned} \tag{5.15}$$

Using the Jensen's inequality, the auxiliary function $Q(\theta, \theta^{[t]})$ is a lower bound of the log-likelihood $\log \mathcal{L}$.

Following Bayes rules, the posterior of the hidden variable $p(Z_n = z_i|\theta^{[t]}, I_n)$ in equation 5.15 can be written,

$$\begin{aligned}p(Z_n = z_i|\theta^{[t]}, I_n) &= \frac{p(I_n|Z_n = z_i, \theta^{[t]}) p(Z_n = z_i|\theta^{[t]})}{\sum_{i=1}^{L} p(I_n|Z_n = z_i, \theta^{[t]}) p(Z_n = z_i|\theta^{[t]})}\\ &= \frac{p(I_n|Z_n = z_i, \theta_I^{[t]}) p(Z_n = z_i|\theta_S^{[t]})}{\sum_{i=1}^{L} p(I_n|Z_n = z_i, \theta_I^{[t]}) p(Z_n = z_i|\theta_S^{[t]})}\,.\end{aligned} \tag{5.16}$$

In the EM algorithm we alternate between expectation steps (E-steps) and maximization steps (M-steps) (Fig. 5.11). In the M-step, we update the estimation of

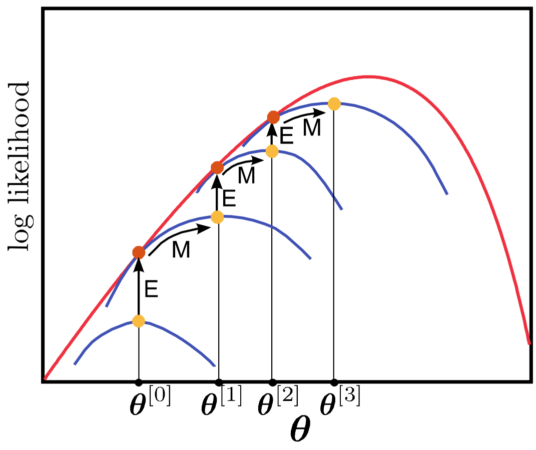


Figure 5.11: EM algorithm (adapted from [Prince 2012]). The red and blue curves represent respectively the log-likelihood, $\log \mathcal{L}$, and the lower bound, $Q(\theta, \theta^{[t]})$.

θ_S and θ_I by maximizing the Q function,

$$
\begin{aligned}
\hat{\theta}_S^{[t]} &= \operatorname*{argmax}_{\theta_S} Q(\theta, \theta^{[t-1]}) \\
&= \operatorname*{argmax}_{\theta_S} \sum_{n=1}^{N} \sum_{i=1}^{L} p(Z_n = z_i | \theta^{[t-1]}, I_n) \log p(Z_n = z_i | \theta_S) \\
&= \operatorname*{argmax}_{\theta_S} \sum_{n=1}^{N} \sum_{i=1}^{L} -p(Z_n = z_i | \theta^{[t-1]}, I_n) \log \frac{p(Z_n = z_i | \theta^{[t-1]}, I_n)}{p(Z_n = z_i | \theta_S)} \\
&= \operatorname*{argmin}_{\theta_S} D_{\mathrm{KL}}(p(Z | \theta^{[t-1]}, I) \| p(Z | \theta_S)) \\
\hat{\theta}_I^{[t]} &= \operatorname*{argmax}_{\theta_I} Q(\theta, \theta^{[t-1]}) \\
&= \operatorname*{argmax}_{\theta_I} \sum_{n=1}^{N} \sum_{i=1}^{L} p(Z_n = z_i | \theta^{[t-1]}, I_n) \log p(I_n | Z_n = z_i, \theta_I) \\
&= \operatorname*{argmax}_{\theta_I} \sum_{n=1}^{N} \sum_{i=1}^{L} -p(Z_n = z_i | \theta^{[t-1]}, I_n) \log \frac{p(Z_n = z_i | \theta^{[t-1]}, I_n)}{p(I_n | Z_n = z_i, \theta_I)} \\
&= \operatorname*{argmin}_{\theta_I} D_{\mathrm{KL}}(p(Z | \theta^{[t-1]}, I) \| p(I | Z, \theta_I)) \,,
\end{aligned}
\tag{5.17}
$$

where the maximization procedure for each parameter is equivalent to the minimization of the Kullback-Leibler divergence between $p(Z|\theta^{[t-1]}, I)$ and $p(Z|\theta_S)$ for the shape model parameters and $p(I|Z, \theta_I)$ for the intensity model parameters. The log-prior is now added to perform **maximum *a posteriori*** (MAP) estimation,

$$
\begin{aligned}
\hat{\theta}_S^{[t]} &= \operatorname*{argmax}_{\theta_S} \Bigg(N \log p(\theta_S|\alpha) \\
&\qquad + \sum_{n=1}^{N} \sum_{i=1}^{L} p(Z_n = z_i|\theta^{[t-1]}, I_n) \log p(Z_n = z_i|\theta_S) \Bigg) \\
&= \operatorname*{argmin}_{\theta_S} D_{\mathrm{KL}}(p(Z|\theta^{[t-1]}, I)\|p(Z|\theta_S)) - N \log p(\theta_S|\alpha) \\
\hat{\theta}_I^{[t]} &= \operatorname*{argmax}_{\theta_I} \Bigg(N \log p(\theta_I|\beta) \\
&\qquad + \sum_{n=1}^{N} \sum_{i=1}^{L} p(Z_n = z_i|\theta^{[t-1]}, I_n) \log p(I_n|Z_n = z_i, \theta_I) \Bigg) \\
&= \operatorname*{argmin}_{\theta_I} D_{\mathrm{KL}}(p(Z|\theta^{[t-1]}, I)\|p(I|Z, \theta_I)) - N \log p(\theta_I|\beta) \,.
\end{aligned}
\tag{5.18}
$$

In the E-step we evaluate the new value of $p(Z_n = z_i|\theta^{[t]}, I_n)$ to update the Q function.

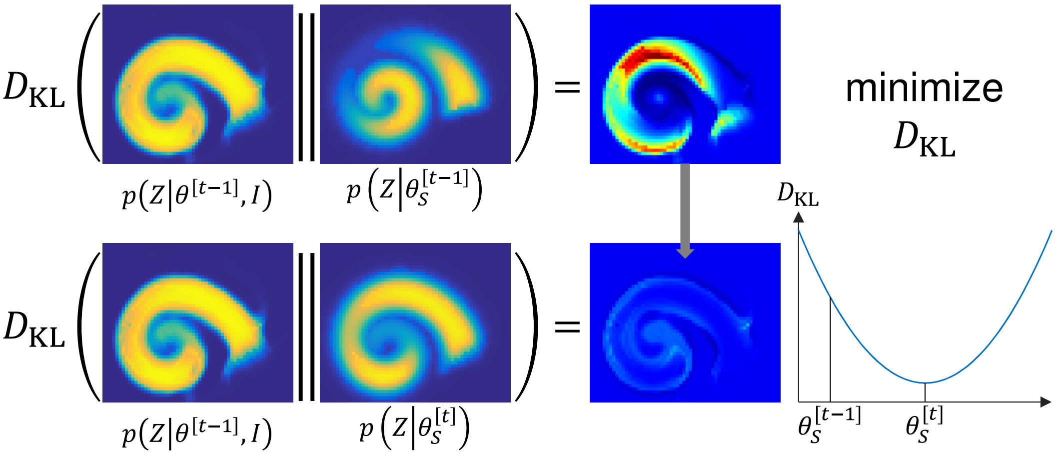


Figure 5.12: Graphical illustration of the M-step used to update $\theta_S^{[t]}$

Summary In a nutshell, the algorithm proceeds with 3 probability maps for each class:

- $p(Z|\theta_S)$, probability defined by the shape model
- $p(I|Z, \theta_I)$, probability defined by the intensity model
- $p(Z|\theta, I)$, joint posterior probability due to the shape and intensity model

The algorithm adjusts iteratively the shape and intensity model parameters in the following manner:

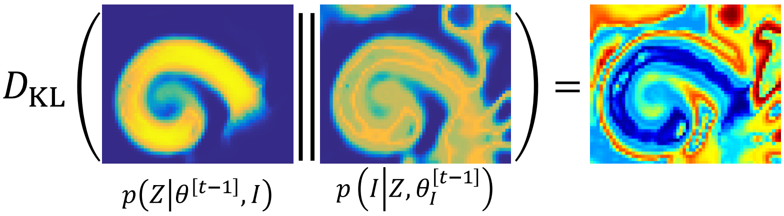


Figure 5.13: Kullback-Leibler divergence between the joint and the intensity model probability maps, $D_{\mathrm{KL}}(p(Z|\theta^{[t-1]}, I) \| p(I|Z, \theta_I))$.

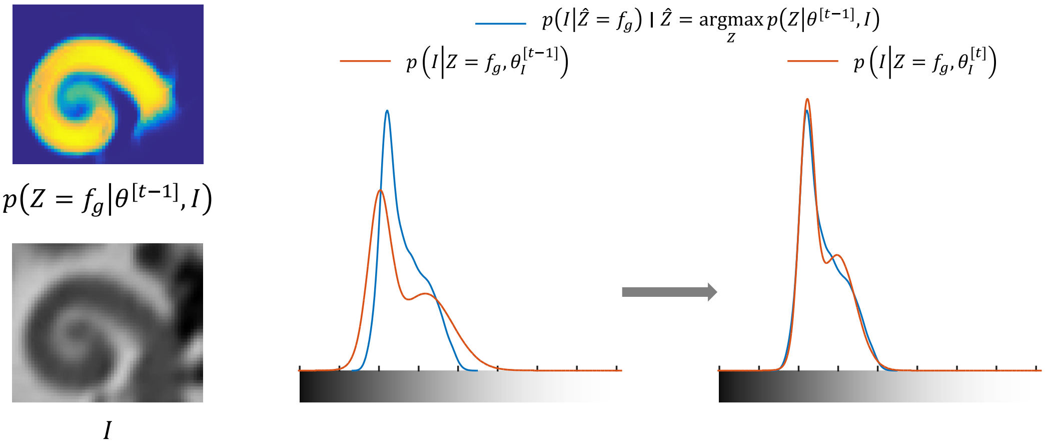


Figure 5.14: Graphical illustration of the M-step used to update $\theta_I^{[t]}$ by minimizing $D_{\mathrm{KL}}(p(Z|\theta^{[t-1]}, I) \| p(I|Z, \theta_I))$ with respect to θ_I. The illustration shows the probability distribution for the label cochlea (or foreground f_g).

1. M-step, update $\theta_S^{[t]}$ (Fig. 5.12) by minimizing the divergence between the shape model and the joint posterior probability maps

2. E-step, update the joint posterior probability with the new shape model parameters

3. M-step, update $\theta_I^{[t]}$ (Fig. 5.14) by minimizing the divergence between the intensity model and the joint posterior probability maps

4. E-step, update the joint posterior probability with the new intensity model parameters

5.2.5 Specific Approach for Cochlear Segmentation

5.2.5.1 Specific Shape Model Probability Map

We seek to define $p(Z|\theta_S)$. The shape model is specified to include or yield metrics that are clinically relevant.

More precisely, the shape model parameters θ_S characterize three surfaces: the cochlea, the scala tympani (ST) and the scala vestibuli (SV), the two scalae being enclosed by the cochlea. Each surface is fully parametric, oriented and represented in the physical space as already detailed in section 4.2.2.2.

The probability that voxel n belongs to the label z_i knowing the shape parameters θ_S, is evaluated using the logit function of the signed distance $d(Z_n = z_i, \theta_S)$ between the center of the voxel n and the surface shape model $S(\theta_S)$ of the cochlea, which is similar to the LogOdds representation [Pohl 2006b].

$$p(Z_n = z_i|\theta_S) = \frac{1}{1 + \exp\left(c \cdot d(Z_n = z_i, \theta_S)\right)}, \tag{5.19}$$

where c is a positive constant. The signed distance function $d(Z_n = z_i, \theta_S)$ returns positive values inside the shape and negative values outside.

We consider 4 labels: the cochlea (or foreground f_g), the scala tympani (ST), the scala vestibuli (SV) and the background (b_g). Since the cochlea encloses the two scalae, we assume that the cochlea is the union of the scala tympani and the scala vestibuli.

$$f_g = ST \cup SV\,, \tag{5.20}$$

furthermore the background is the absolute complement of the cochlea

$$b_g = f_g^{\complement}\,, \tag{5.21}$$

$$d(Z_n = b_g, \theta_S) = -d(Z_n = f_g, \theta_S)\,, \tag{5.22}$$

$$p(Z_n = b_g|\theta_S) = 1 - p(Z_n = f_g|\theta_S)\,. \tag{5.23}$$

An example of $p(Z = f_g|\theta_S)$ is shown for a slice of the CT in Fig. 5.12 illustrating "shape model".

5.2.5.2 Specific Cochlear Shape Model

The surfaces $S(Z = z_i, \theta_S)$ are hierarchically described in Fig. 5.15. The parameters θ_S include **pose** (with three **position** parameters and three **orientation** parameters) and **shape** parameterization.

The cochlear structures are defined as generalized cylinders, i.e. as **cross-sections** swept along a **centerline**.

The centerline is parameterized in a cylindrical coordinate system by its **radial** and **longitudinal** coordinates within a given interval which defines the **length** of the centerline.

The cross-sections are modeled by a closed **planar shape** on which a varying **affinity** is applied along the centerline. The scala tympani and the scala vestibuli are modeled with two half pseudo-cardioids while the cochlear cross-section corresponds to the minimal circumscribed ellipse of the union of the tympanic and vestibular cross-sections.

The affinity of cross-section is parameterized by a **rotation**, a **width** and a **height scaling**.

Eventually each of the derived component of the model is represented by a vector or an one-dimensional function of the angular coordinate. One-dimensional functions are themselves parameterized by combinations of simple functions (i.e polynomial, logarithmic, etc.).

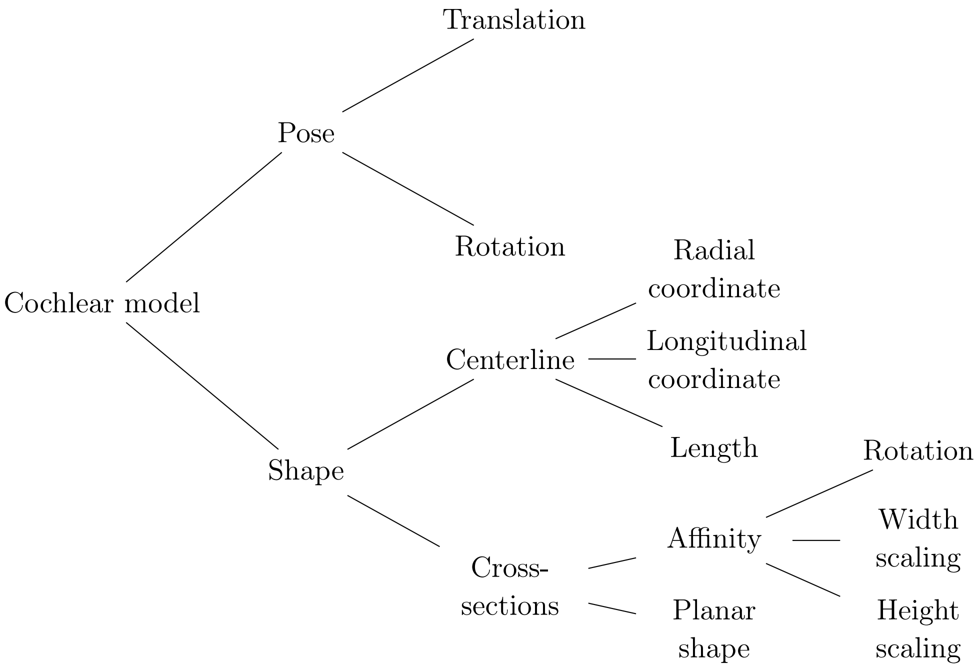


Figure 5.15: Derivation of the model parameters describing the surface shape model S

The shape parameters probability distribution were modeled using either an uniform distribution (uninformative prior) or a Dirac δ distribution (fixed parameter). Since the segmentation is not an instance of the shape model and part of the variability is taken into account by the intensity model parameters θ_I, we can use a very compact shape model. Without considering the fixed parameters θ_S includes:

- 3 translation parameters, t_x, t_y and t_z (in the direction x, y and z)
- 3 rotation parameters, r_x, r_y and r_z (about the x-, y- and z-axis)
- 1 scale parameter of the radial component of the centerline, a
- 2 parameters of longitudinal component of the centerline, w_1 and w_2
- 1 parameter of the length (or more precisely the number of cochlear turn) $\theta_{\max}$

A small number of degrees of freedom was chosen in order to limit as much as possible the correlation between the shape parameters.

Radial Component The radial component is defined piecewise with a polynomial function and a logarithmic function of the angular coordinate θ_c in the cylindrical coordinate system,

$$r(\theta_c) = \begin{cases} p_2\theta_c^2 + p_1\theta_c + p_0 & \text{if } \theta_c < \theta_2 \\ ae^{-b\theta_c} & \text{else} \end{cases} \tag{5.24}$$

By constraining $r(\theta_c)$ to be continuously differentiable, 4 free parameters remain, a, b, θ_2 and p_0 since,

$$p_2 = \frac{C_1\theta_2 - C_2 + p_0}{\theta_2^2} \quad p_1 = \frac{-C_1\theta_2 + 2C_2 + 2p_0}{\theta_2}\,, \tag{5.25}$$

where

$$C_2 = ae^{-b\theta_2} \quad C_1 = -C_2 b\,. \tag{5.26}$$

The parameter b is a widely studied cochlear measurements [Cohen 1996] that can be fixed ($b = 0.15$ rad^{-1}). The parameters a, θ_2 and p_0 are correlated with the rotation parameter r_z since

$$r(\theta_c + r_z|a, b, \theta_2, p_0) \Leftrightarrow r(\theta_c|a', b, \theta_2', p_0')\,, \tag{5.27}$$

where

$$a' = ae^{-br_z},\ \theta_2' = \theta_2 - r_z,\ p_0' = p_2 r_z^2 + p_1 r_z + p_0\,. \tag{5.28}$$

To avoid this correlation we fixed the base of the spiral by setting θ_2 and p_0 (respectively to $5\pi/6$ and 5 mm).

Longitudinal Component Concerning the longitudinal component of the centerline we used the following simplified version,

$$z(\theta_c) = \alpha e^{-\beta\theta_c}\cos(\theta_c + \phi) + q_1\theta_c + q_0\,, \tag{5.29}$$

which is the sum of an exponentially damped sinusoidal and a linear function. The parameter q_0 is linearly correlated with the translation parameter t_z, the slope parameter p_1 varies sightly and α, β and ϕ have complex correlation with the rotation parameters.

To overcome this complex correlation and reduce the number of degrees of freedom we performed a principal component analysis of the parameters $\{\alpha, \beta, \phi, q_1, q_0\}$ obtained by minimizing the sum of squared differences between the fitted function and discretized longitudinal component extracted manually from 20 CT images. The centerlines were extracted by manually setting 17 points, the projection on the z-axis was performed with the robust modiolar axis estimation method (subsection 3.2.4). Two principal components weighted by parameters w_1 and w_2, were kept because in practice only the initial amplitude α and the phase ϕ of the exponentially damped sinusoidal part influence significantly the centerline. Note that α

and ϕ are respectively roughly aligned with w_2 and w_1,

$$\begin{bmatrix}\alpha\\ \beta\\ \phi\\ q_1\\ q_0\end{bmatrix} = \begin{bmatrix} 1.100 & -0.089 & 0.473\\ 0.310 & -0.211 & 0.209\\ 0.571 & 0.558 & 0.105\\ 0.224 & -0.028 & 0.014\\ -0.035 & 0.227 & -0.114\end{bmatrix} \cdot \begin{bmatrix}1\\ w_1\\ w_2\end{bmatrix} \tag{5.30}$$

Cross-sectional Parameters All cross-sectional parameters are fixed because they deformed the shape only locally and this variability is well captured by the intensity model. The cross-sectional parameters are comprehensively described in Table 4.2.

5.2.5.3 Specific Intensity Model

Mixture of Student's t-distributions The probability of observing intensity I_n knowing the label Z_n is parameterized with mixtures of Student's t-distributions.

$$p(I_n|Z_n = z_i, \theta_I) = \sum_{k=1}^{K_i} \phi_{i,k}\, t(I_n|\mu_{i,k}, \sigma_{i,k}, \nu_{i,k})\,, \tag{5.31}$$

where K_i corresponds to the number of components of the mixture for the label z_i, the weights $\phi_{i,k}$ are positive and sum to one and the Student's t-distribution is

$$t(I_n|\mu, \sigma, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}} \left(1 + \frac{(I_n - \mu)^2}{\sigma^2\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}, \tag{5.32}$$

where $\Gamma(\cdot)$ is the gamma function.

We assume that the scala tympani, the scala vestibuli and the cochlea share the same probability of observing the intensity I_n,

$$p(I_n|Z_n = ST, \theta_I) = p(I_n|Z_n = SV, \theta_I) = p(I_n|Z_n = f_g, \theta_I)\,. \tag{5.33}$$

An example of $p(I|Z = f_g, \theta_I)$ is shown for a slice of the CT in Fig. 5.3 illustrating quotesintensity model.

Intensity Distributions in Cochlear CT images The dimension of the region of interest around the cochlea is 12x10x10 mm^3.

In this region, the **background** (b_g) consists mainly in **fluids** (perilymph and endolymph in the vestibular labyrinth), **extremely dense bone** (bony labyrinth), **air** and **less dense bone** (pneumatized temporal bone).

The **foreground** (the cochlea, f_g) contains mainly **fluids** (perilymph and endolymph), **bony structures** (cochlear walls) and negligible stiff structures (basilar membrane, organ of Corti, Reissner's membrane, etc.).

The set of images is CT scans, the Hounsfield unit (HU) is applicable here (Table 5.1).

Substance	Anatomy	HU
fluid	perilymph, endolymph	0
dense bone	bony labyrinth	2000
air	pneumatization	-1000
less dense bone	temporal bone, bony walls	700

Table 5.1: Hounsfield unit (HU) of the main structures of the cochlear region

Based on the anatomy, the number of modes K_{f_g} and K_{b_g} of the mixtures of Student's t-distributions of respectively the foreground and background are set to 2 (expected at 0 HU and 700 HU) and 4 (expected at 0 HU, 2000 HU, -1000 HU and 700 HU).

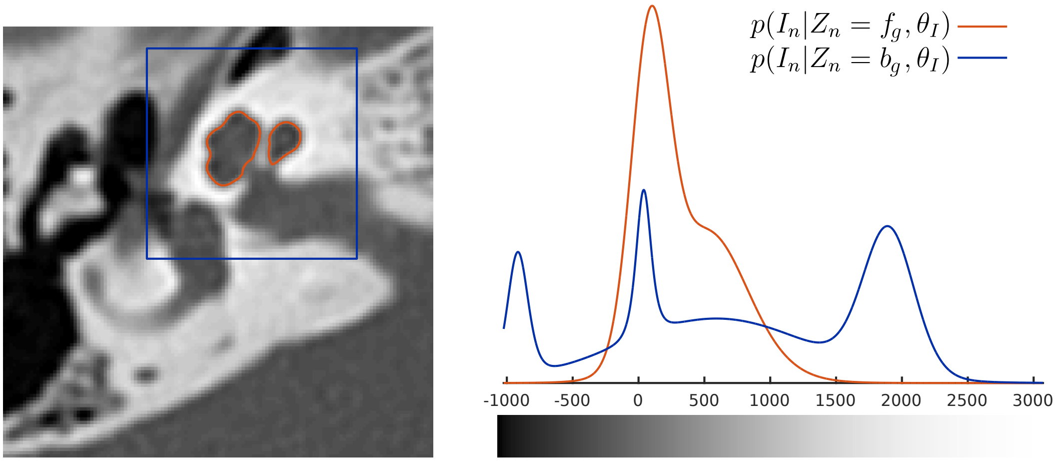


Figure 5.16: Example of intensity probability distributions of the foreground (f_g, in red) and the background (b_g, in blue) as functions of the Hounsfield unit.

5.2.5.4 Hard Segmentation

Given the considered labels and their relations (Eq. 5.20 and 5.21), two sets of label can be considered $\{f_g, b_g\}$ or $\{ST, SV, b_g\}$.

The hard segmentation $\hat{Z}_n$ of the voxel n is then obtained by MAP which is equivalent to local majority voting,

$$\hat{Z}_n = \underset{z_i}{\operatorname{argmax}}(p(Z_n = z_i|\hat{\theta}, I_n)) , \tag{5.34}$$

where $\hat{\theta}$ is computed using the EM algorithm. An example of $\hat{Z}_n$ is shown on the right side of Fig. 5.3 as isocontours and isosurfaces.

5.2.6 Implementation

5.2.6.1 Signed Distance Evaluation

We call signed distance a function that has positive values when a point p is inside the label ($p \in z_i$) and negative values outside the label ($p \in z_i^\complement$). In order to evaluate $d(Z = z_i, \theta_S)$, we use Algorithm 1 which computes the signed distance function with a coarse-to-fine approach. First parametric surfaces were discretized into triangular meshes so that all edges length are of the same order of magnitude (0.30 ± 0.15 mm), which leads to meshes M with about 2500 points and 5000 triangles.

The distance is first approximated by rasterizing each mesh triangle in the image space (with the [Bresenham 1965]'s line algorithm). The holes of the resulting binary image BW are then filled (using morphological reconstruction [Soille 1999]), a hole being a set of background connected components that are not adjacent to the 6 sides of the three-dimensional image. The Euclidean distance between inside voxels and the closest outside voxel is approximated inside the binary image and conversely outside (but with negative values) (with the fast algorithm described in [Maurer 2003]).

The values of the voxels close to the mesh (within a given margin $|d| < T = 0.6$ mm) are reevaluated by computing the exact signed distance between the center of the voxels and the mesh (as described in Algorithm 2). This algorithm requires the computation of the signed distance between a point and an oriented triangle, Algorithm 4 in appendix B was used.

Eventually, the evaluation of the exact signed distance $d(Z = z_i, \theta_S)$ is time-consuming and the running time is $O(Nn_p)$, where N is the number of voxels and n_p is the number of points in the triangle surface mesh.

The constant c in Eq. 5.19 is set to 4 mm^{-1}, such that the margin of 0.6 mm covers all voxels that have a probability to belong to a given label between 0.08 and 0.92. We assume that voxels far away from the surface shape (where $p \to 1$ inside and $p \to 0$ outside) do not influence the optimization procedure so that a first approximation is sufficient.

5.2.6.2 Optimization Algorithm

During the M-step of the EM algorithm, the equivalent minimization of the Kullback-Leibler divergences (Eq. 5.17) were performed using the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970] considering unconstrained functions (`fminunc` in MATLAB). The shape parameters θ_S were rescaled so that minimum and the maximum values of the uniform priors corresponds to -5 and 5. The gradient was estimated with central finite differences with a step size of 0.1 for both shape and intensity parameters.

Algorithm 1: Signed distance map

```
Input: parametric surfaces $S(Z = z_i, \theta_S)$ for all labels $\{z_i\}_{i=1,\dots,L}$ and image $I$
Output: signed distance maps $d(Z = z_i, \theta_S)$ for all labels $\{z_i\}_{i=1,\dots,L}$
for $i = 1, \dots, L-1$ do
    $M \leftarrow$ discretize $S(Z = z_i, \theta_S)$ in a triangle mesh ($n_p$ points, $\{p_i\}_{i=1,\dots,n_p}$ and $n_t$ triangles, $\{v_{1,i}, v_{2,i}, v_{3,i}\}_{i=1,\dots,n_t}$);
    $BW \leftarrow$ rasterize all triangles [Bresenham 1965];
    $BW \leftarrow$ fill in topological holes [Soille 1999];
    $d(Z = z_i, \theta_S) \leftarrow$ points (center of $I_n$) to binary image $BW$ signed Euclidean distance [Maurer 2003] ;
    for $n = 1, \dots, N$ do
        if $|d(Z_n = z_i, \theta_S)| < T$ then
            $d(Z_n = z_i, \theta_S) \leftarrow$ PointToMesh$(M, p)$, where $p$ is the center of $I_n$;
        end
    end
end
for $n = 1, \dots, N$ do
    $d(Z_n = z_L, \theta_S) = -\max_{i=1,\dots,L-1}(d(Z_n = z_i, \theta_S))$;
end
```

Algorithm 2: `PointToMesh`(p, M)
Point to triangle mesh signed Euclidean distance

```
Input: point, $p$
       oriented mesh $M = \{n_p$ points, $\{p_i\}_{i=1,\dots,n_p}$
       and $n_t$ triangles, $\{v_{1,i}, v_{2,i}, v_{3,i}\}_{i=1,\dots,n_t}\}$
Output: signed Euclidean distance $d$
$\{d_i\}_{i=1,\dots,n_p} \leftarrow$ Euclidean distances between point $p$ and each point of the triangular mesh $\{p_i\}_{i=1,\dots,n_p}$ ;
$i^* \leftarrow \mathrm{argmin}_i\, d_i$ ;
for $i = 1, \dots, n_t$ do
    if $p_i^* \in \{v_{1,i}, v_{2,i}, v_{3,i}\}$ then
        if $|$PointToTriangle$(p, v_{1,i}, v_{2,i}, v_{3,i}))| < |d|$ then
            $d \leftarrow$ PointToTriangle$(p, v_{1,i}, v_{2,i}, v_{3,i})$ ;
        end
    end
end
```

5.3 Results

Two datasets were segmented.

Algorithm 3: EM algorithm

```
$\hat{\theta}^{[t-1]} \leftarrow \theta_0$ ;
$\hat{\theta}^{[t]} \leftarrow \theta_0$ ;
$\Delta Q \leftarrow \infty$ ;
while $\Delta Q > \varepsilon$ do
    for $i = 1, \ldots, L-1$ do
        $P_i \leftarrow p(Z = z_i | \theta^{[t-1]}, I)$ ;
    end
    $P_L \leftarrow 1 - \sum_{i=1}^{L-1} P_i$ ;
    $Q_S(x) \leftarrow N \log p(x|\alpha) + \sum_{n=1}^{N} \sum_{i=1}^{L} P_{i,n} \log p(Z_n = z_i | x)$ ;
    $Q_{I,f_g}(x) \leftarrow N \log p(x|\beta_{f_g}) + \sum_{n=1}^{N} P_{i|z_i=f_g,n} \log p(I_n | Z_n = f_g, x)$ ;
    $Q_{I,b_g}(x) \leftarrow N \log p(x|\beta_{b_g}) + \sum_{n=1}^{N} P_{i|z_i=b_g,n} \log p(I_n | Z_n = b_g, x)$ ;
    $\hat{\theta}^{[t]} \leftarrow \texttt{BFGS}(Q_S(x), \hat{\theta}^{[t]})$ ;
    $\hat{\theta}^{[t]} \leftarrow \texttt{BFGS}(Q_{I,f_g}(x), \hat{\theta}^{[t]})$ ;
    $\hat{\theta}^{[t]} \leftarrow \texttt{BFGS}(Q_{I,b_g}(x), \hat{\theta}^{[t]})$ ;
    $\Delta Q \leftarrow Q_S(\hat{\theta}^{[t]}) + Q_{I,f_g}(\hat{\theta}^{[t]}) + Q_{I,b_g}(\hat{\theta}^{[t]}) - Q$ ;
    $Q \leftarrow Q + \Delta Q$ ;
    $\hat{\theta}^{[t-1]} \leftarrow \hat{\theta}^{[t]}$ ;
end
```

5.3.1 Dataset #1 Segmentation

A summary of the available segmentations is given in Table 5.2. When comparing different modalities, the quantitative results will actually reflect both the segmentation and the registration quality. The segmentation are obtained from the probability maps using hard segmentation as described in section 5.2.5.4.

	f_g ($ST \cup SV$)		ST		SV	
	CT	μCT	CT	μCT	CT	μCT
manual segmentation Z_0	✓	✓		✓		✓
model fitting S	✓		✓		✓	
automatic segmentation $\hat{Z}$	✓		✓		✓	

Table 5.2: Available segmentations for dataset #1

We compared quantitatively the fitted parametric shape model S and the automatic segmentation $\hat{Z}$ with manual segmentation Z_0.

The Dice score (Eq 5.38) was computed between the rasterization of the shape model S and the automatic segmentation $\hat{Z}$ on one hand and the manual segmentation Z_0 on the other hand. The Hausdorff distance was computed symmetrically between the surface shape model of the cochlea and the isosurface of the automatic

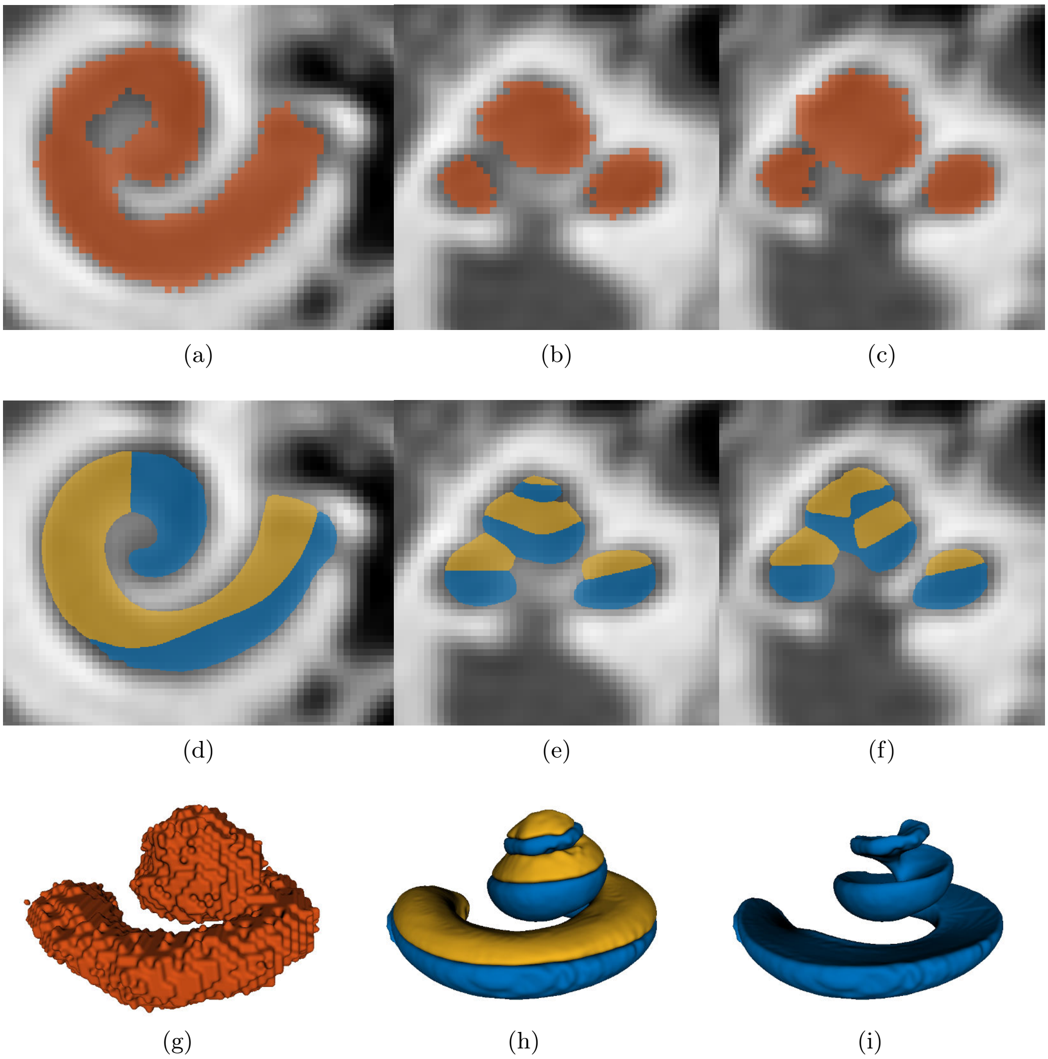


Figure 5.17: Comparison of the manual segmentation Z_0 (a), (b), (c) and (g) and the automatic segmentation $\hat{Z}$ (d), (e), (f), (h) and (i) on CT images

segmentation on one side and the isosurface of the manual segmentation on the other side. The parametric surfaces were discretized in triangle meshes so that all edge lengths are of the same order of magnitude (0.30 ± 0.15 mm) and the meshes have the same number of vertices, in order to guarantee the homogeneity of the measurement. Table 5.3 presents the mean results obtained on dataset #1. The compared labels are described as *method*/*modality*/*label*, where the *method* can be manual segmentation Z_0, model fitting S or automatic segmentation $\hat{Z}$, the *modality* may be CT or μCT and the *label* my be the scala tympani (ST), the scala vestibuli (SV), the cochlea (or foreground f_g).

To compare the cochlea f_g and the union of the scalae $ST \cup SV$, only the Dice is evaluated since the Hausdorff distance expresses a surface error and the isosurface of the union of different labels is not equivalent to the union of the isosurfaces of different labels.

Detailed results on the nine cochleae are presented in the appendix C.

		Dice	**Hausdorff distance (mm)**				
Compared labels		**score**	50%	68%	95%	100%	
$Z_0/\text{CT}/f_g$	$S/\text{CT}/f_g$	0.79	0.13	0.21	0.48	1.42	Table C.1
	$\hat{Z}/\text{CT}/f_g$	**0.84**	**0.10**	**0.16**	**0.40**	**1.31**	
$Z_0/\mu\text{CT}/ST \cup SV$	$S/\text{CT}/f_g$	0.70	-	-	-	-	Table C.2
	$\hat{Z}/\text{CT}/f_g$	**0.78**	-	-	-	-	
	$Z_0/\text{CT}/f_g$	0.76	-	-	-	-	
$Z_0/\mu\text{CT}/ST$	$S/\text{CT}/ST$	0.72	0.11	0.17	0.40	0.88	Table C.3
	$\hat{Z}/\text{CT}/ST$	**0.77**	**0.08**	**0.14**	**0.37**	**0.87**	
$Z_0/\mu\text{CT}/SV$	$S/\text{CT}/SV$	0.70	0.10	0.16	0.37	**0.96**	Table C.4
	$\hat{Z}/\text{CT}/SV$	**0.73**	**0.08**	**0.14**	**0.34**	0.97	

Table 5.3: Mean model fitting and segmentation results obtained on dataset #1.

Comparison with the State of the Art Previous works already performed cochlear anatomy segmentation and several metrics have been used to evaluate the segmentation quality. Since the goal is to performed segmentation of conventional CT and despite the great interest of the methods, μCT segmentation methods [Kjer 2014b, Ruiz Pujadas 2016a, Ruiz Pujadas 2016b] are not involved in the comparison.

Above all, it is important to stress that the Dice-score obtained by comparing the registered manual segmentation performed on CT ($Z_0/\text{CT}/f_g$) and the manual segmentation performed by the same expert on μCT ($Z_0/\mu\text{CT}/ST \cup SV$) is on average equal to 0.76 which is not as good as the results obtained with automatic segmentation ($\hat{Z}/\text{CT}/f_g$, 0.78), as detailed in Table C.2. This highlights the limitations (large inter-rater variability) of manual CT images segmentation (Fig. 5.17).

Cochlear shape fitting was studied by Gavin Baker, his model-image registration framework presented in [Baker 2005] and detailed in [Baker 2008] was validated by evaluating the precision, the recall and the Jaccard index between fitted shape and synthetic noisy images. As a reminder, when comparing the results of a segmentation Z and a reference mask A, we write

$$\text{precision} = \frac{A \cap Z}{Z} \tag{5.35}$$

$$\text{recall} = \frac{A \cap Z}{A} \tag{5.36}$$

$$\text{Jaccard index} = \frac{A \cap Z}{A \cup Z} \tag{5.37}$$

$$\text{Dice score} = 2\frac{A \cap Z}{A + Z} = \text{F}_1 \text{ score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}\,. \tag{5.38}$$

In [Baker 2008], a cochlear shape model was fitted to four CT images of the temporal bone. After being manually initialized the precision was evaluated by comparing with manual segmentations. The method performed with a precision of 0.72 ± 0.09, while our method presented in this study provides a precision of 0.80 ± 0.06 considering model fitting only and 0.81 ± 0.05 with joint shape and intensity model-based segmentation.

The problem and validation described in [Noble 2011] are similar to the problem dealt with in this study. Indeed, in [Noble 2011] manually segmented ST and SV from high-resolution images were compared to five automatically segmented conventional CT in order to obtain quantitative measurements. This is equivalent to compare the label $Z_0/\mu\text{CT}/ST$ and the label $\hat{Z}/\text{CT}/ST$, as in Table C.3 (and similarly for SV in Table C.4). The authors used a four modes active shape model trained with a leave-one-out approach and claimed a Dice score of 0.77 for the scala tympani and 0.72 for the scala vestibuli, which is close to the score of 0.77 and 0.73 obtained here. Their segmentation method applied to five cadaveric cochleae achieved an average mean and maximum surface errors of 0.21 mm and 0.8 mm, while our method achieved 0.12 mm and 0.92 mm. [Noble 2012, Noble 2013] applied also the method for spiral ganglion segmentation with similar outcomes.

Almost the same validation (i.e comparing detailed μCT based segmentation and CT) was performed by [Kjer 2015d], with the exception that the label considered was a unique shape of what we can call the labyrinth, i.e. the vestibule, semicircular canals, and cochlea. The fairest comparison would be to compare with segmentation errors between the label $Z_0/\mu\text{CT}/ST \cup SV$ and the label $\hat{Z}/\text{CT}/f_g$, as in Table C.2. For the same reason that prevented us to compute the Hausdorff distance, the authors refrained from computing the symmetric distances and computed the mean (± 1 std) and maximum surface errors from the automatic CT segmentation to the manual μCT segmentation: 1.15 mm and 0.22 ± 0.16 mm for one case and 0.89 mm and 0.23 ± 0.18 mm for the other, while the average measurements calculated here are 0.88 mm and 0.12 ± 0.13 mm.

Hans Martin Kjer's PhD thesis [Kjer 2015b], shows promising results that it would be interesting to compare to. In the last chapter, the testing dataset has no

Study	**Comparison**	**Metric**		Proposed method (N = 9)
Baker 2008 (N = 4)	$S/\mathrm{CT}/f_g$ vs $Z_0/\mathrm{CT}/f_g$	Precision	0.72 ± 0.09	**0.80 ± 0.06**
Noble 2011 (N = 5)	$\hat{Z}/\mathrm{CT}/ST$ vs $Z_0/\mu\mathrm{CT}/ST$	Dice score	**0.77**	**0.77**
		Mean surface error	0.18	**0.12**
		Hausdorff distance	**0.8**	0.87
	$\hat{Z}/\mathrm{CT}/SV$ vs $Z_0/\mu\mathrm{CT}/SV$	Dice score	0.72	**0.73**
		Mean surface error	0.22	**0.12**
		Hausdorff distance	**0.8**	0.97
Kjer 2015 (N = 2)	$\hat{Z}/\mathrm{CT}/f_g$ vs $Z_0/\mu\mathrm{CT}/ST \cup SV$	Mean (± 1 std) surface error CT to μCT	0.22 ± 0.17	**0.12 ± 0.13**

Table 5.4: Summary of the comparison with previous studies. N indicates the number of segmented images

ground truth segmentation making quantitative evaluation impossible. The training dataset is composed of manually segmented μCT and a fitted statistical deformation model is evaluated on downsampled μCT images to CT resolutions. By avoiding the multi-modal registration, the study achieved labyrinth segmentation with a Dice score of 0.88 ± 0.02, a mean surface error of 0.11 ± 0.01 and a Hausdorff distance of 0.58 ± 0.11.

Other studies [Reda 2013, Reda 2014a, Reda 2014b] performed ST and SV segmentation by using the active shape model-based method developed in [Noble 2011].

5.3.2 Dataset #2 Segmentation

Since no manual segmentation was available for all datasets, a qualitative visual check was performed by a head and neck imaging expert using 15 slices in 3 orientation planes per cochlea (see Fig. 5.19 as an example of such panel).

The quality of the automatic segmentation was classified into 4 categories : failure (the segmentation does not match the intensity), major imperfection (over 100 voxels are misclassified), minor imperfection (less 100 voxels are misclassified) and success. Major and minor imperfections were themselves classified into 2 categories: undersegmentation and oversegmentation.

95% of the segmentations were considered satisfactory as reported in Table 5.5. Failed segmentations are due to pathological cochleae (Fig. 5.25) or bad initialization procedure (Fig. 5.26).

		#	%
Success		937	95%
minor imperfection	undersegmentation	5	2%
	oversegmentation	15	
major imperfection	undersegmentation	6	2.5%
	oversegmentation	19	
Failure		5	0.5%

Table 5.5: Qualitative segmentation evaluation on dataset #2

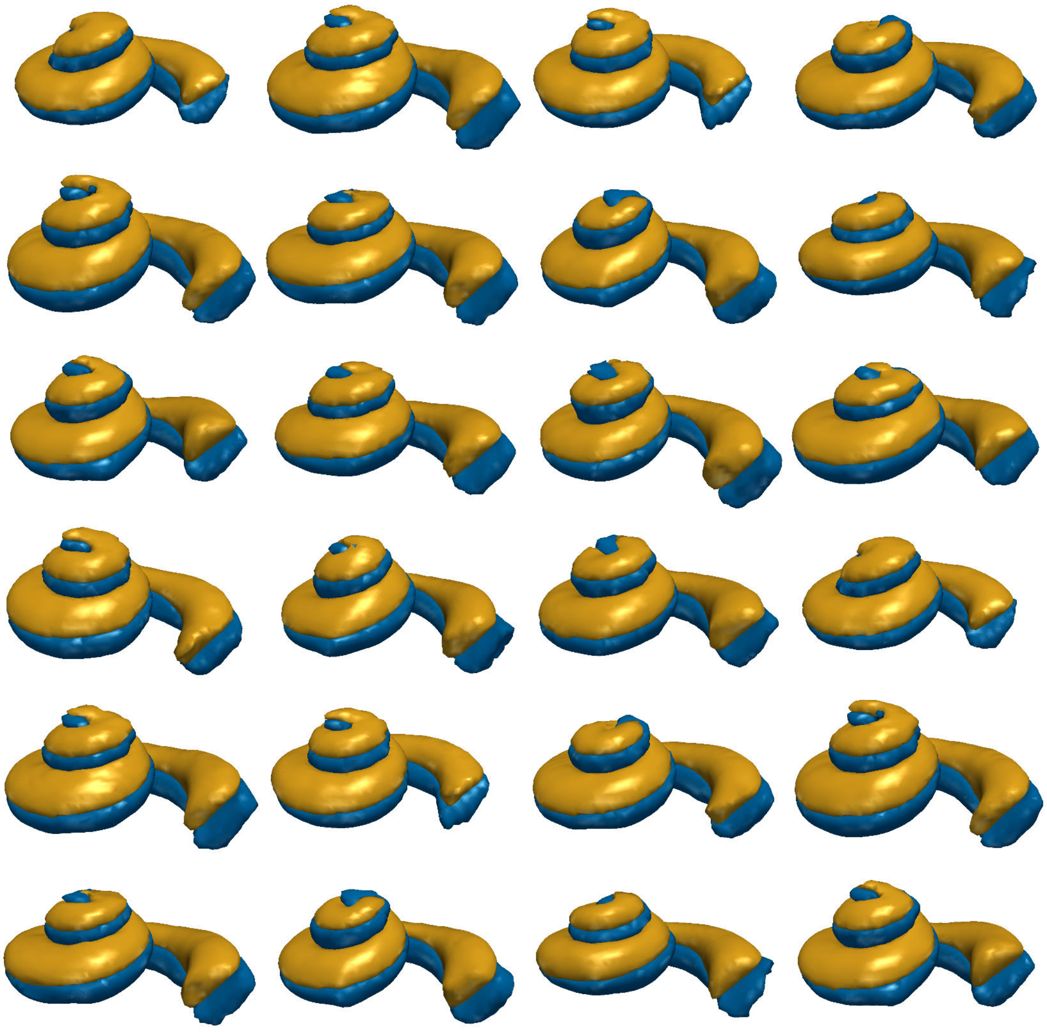


Figure 5.18: Examples of three-dimensional shapes of the ST (blue) and SV (orange) obtained with automatic segmentation from CT images.

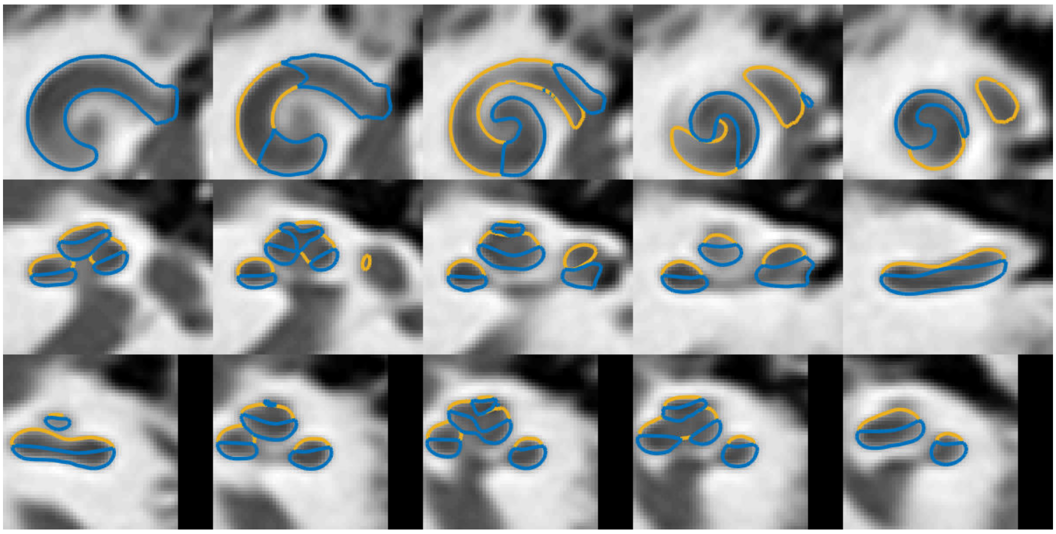


Figure 5.19: Example of automatic segmentation considered as a “success”

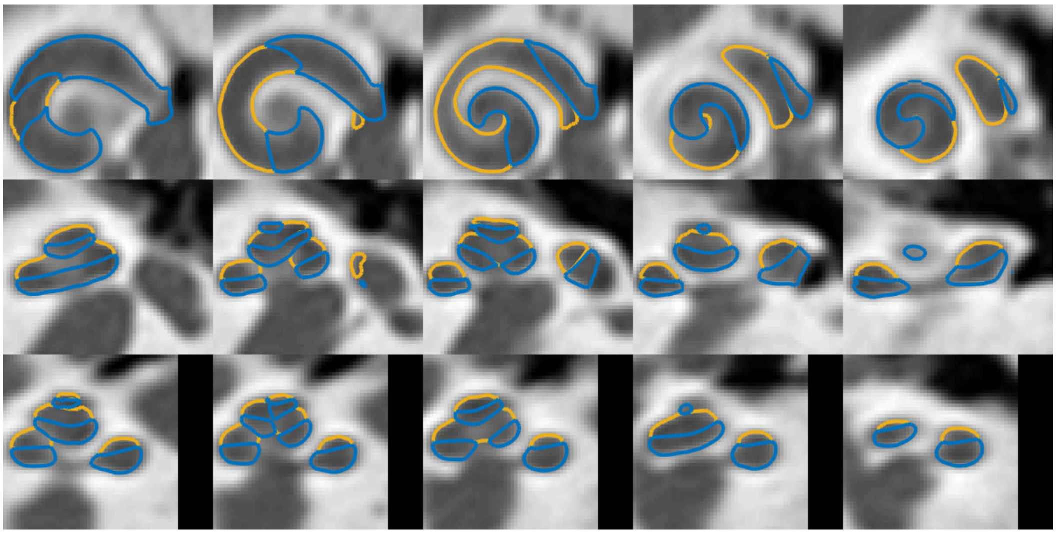


Figure 5.20: Example of automatic segmentation considered as a “success”

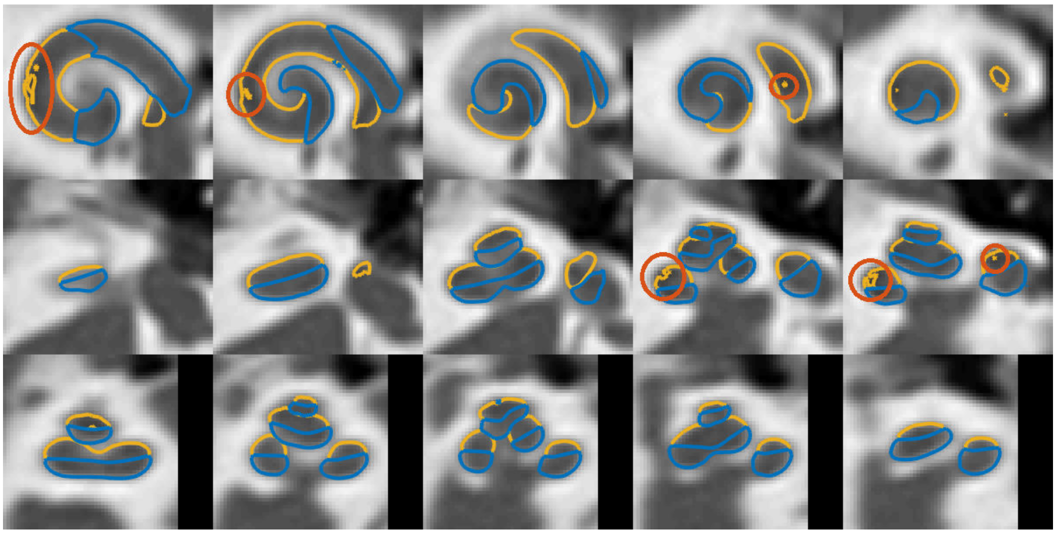


Figure 5.21: Example of automatic segmentation considered as a “minor imperfection” with undersegmentation. Segmentation errors are marked in red.

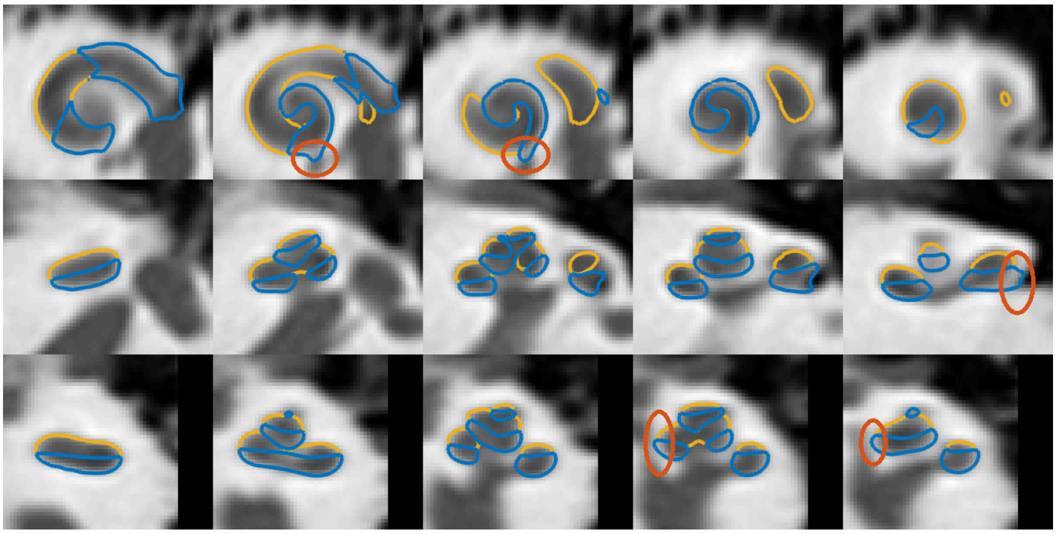


Figure 5.22: Example of automatic segmentation considered as a “minor imperfection” with oversegmentation. Segmentation errors are marked in red.

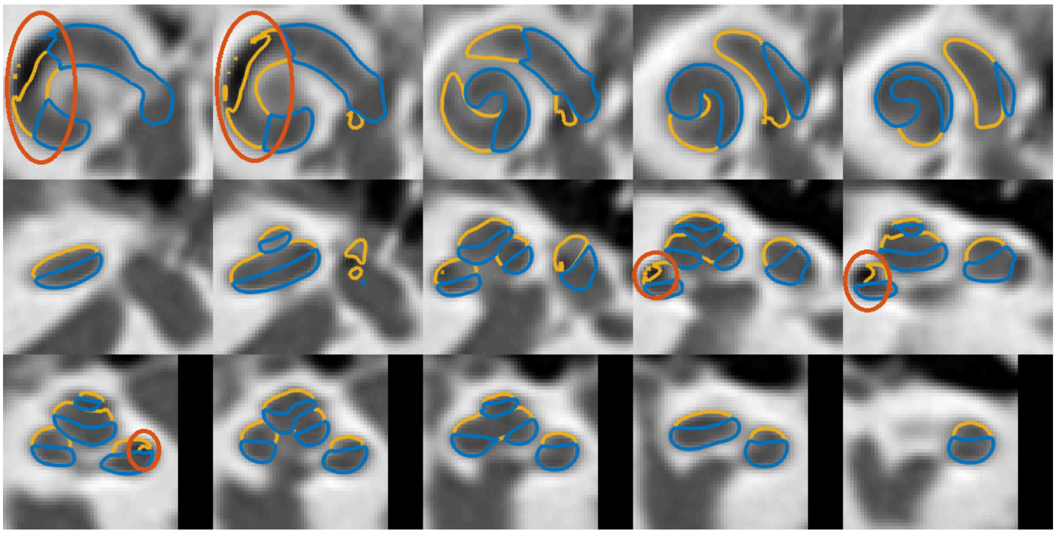


Figure 5.23: Example of automatic segmentation considered as a “major imperfection” with undersegmentation. Segmentation errors are marked in red.

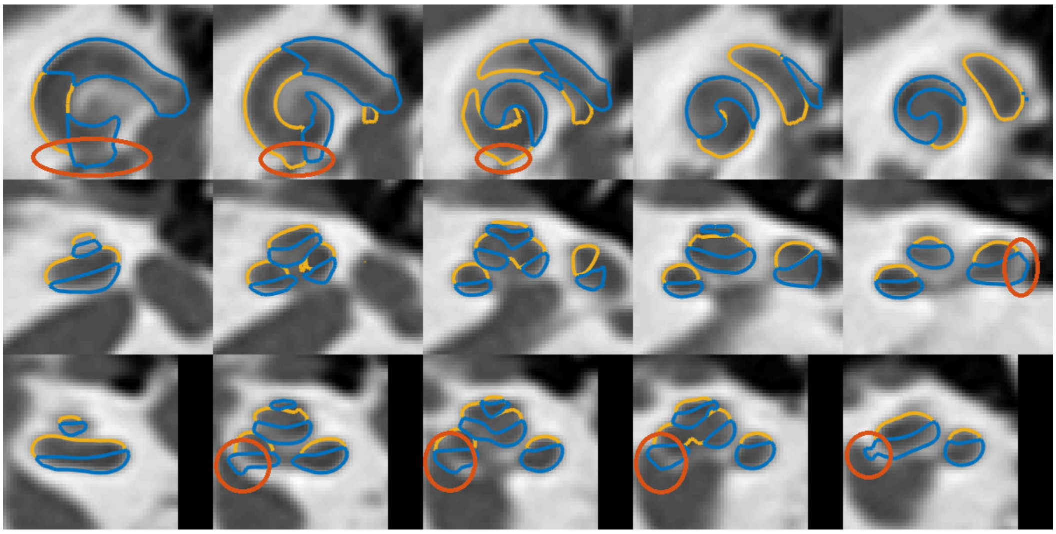


Figure 5.24: Example of automatic segmentation considered as a “major imperfection” with oversegmentation. Segmentation errors are marked in red.

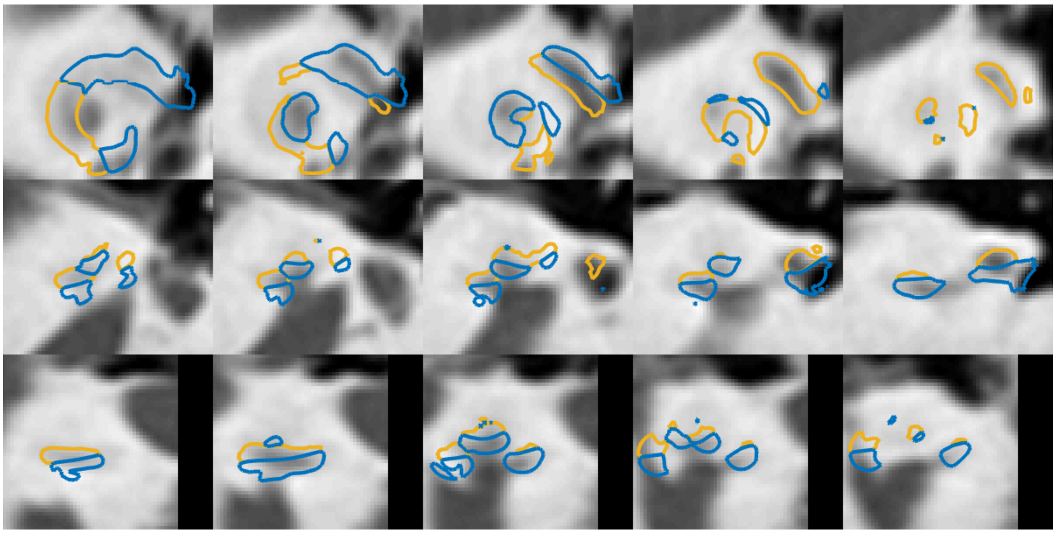


Figure 5.25: Example of automatic segmentation considered as a “failure”

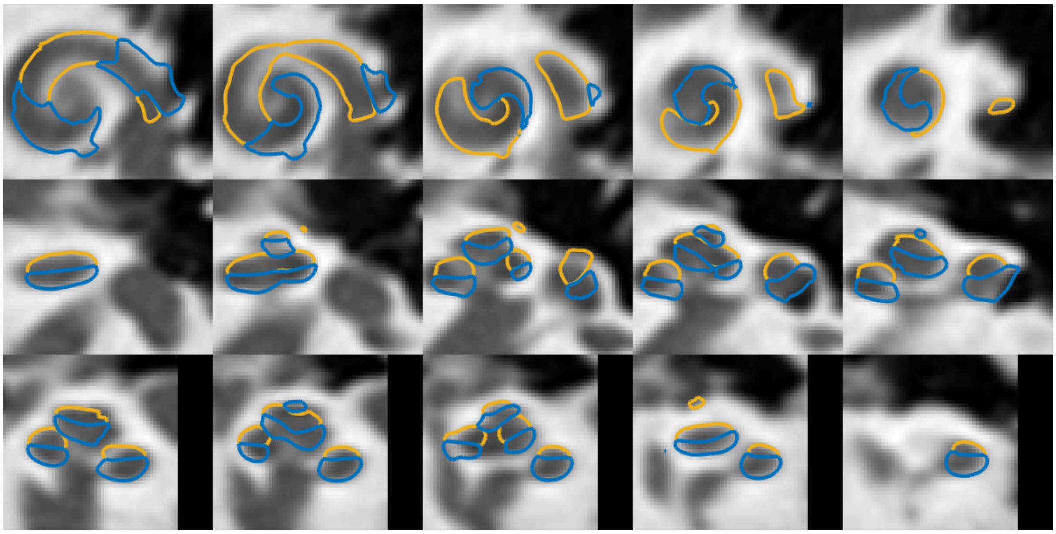


Figure 5.26: Example of automatic segmentation considered as a “failure”

	w_1	w_2	a	$\theta_{\max}$
w_1	0.12	-0.03	0.01	0.01
w_2	-0.03	0.52	-0.01	-0.03
a	0.01	-0.01	0.27	-0.02
$\theta_{\max}$	0.01	-0.03	-0.02	0.09

Table 5.6: Covariance matrix of the 4 shape parameters

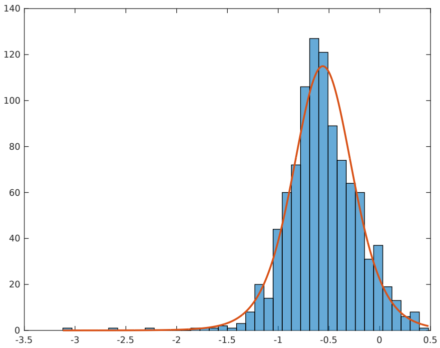


Figure 5.27: Shape parameter w_1

5.3.3 Anatomical Variability on Dataset #2

We analyze the variability of the 4 shape parameters estimated with the large dataset. As a reminder, the parameters are initially scaled so that the minimum and the maximum values of the uniform priors corresponds to -5 and 5. The absolute value of the Pearson correlation coefficient between each shape parameters is lower than 0.16. The covariance matrix (Table 5.6) is diagonal dominant confirming that parameters are mostly independent.

Shape parameters of the longitudinal component, w_1 (Fig. 5.27), the radial component, a (Fig. 5.30a) and the length $\theta_{\max}$ (Fig. 5.30b) of the centerline follow roughly normal distributions (slightly skewed) while the parameter w_2 is bimodal follows the distribution of the mixture of two normal distributions. Parameter w_2 modeled the longitudinal component of the centerline. Fig. 5.29 shows the typical longitudinal profile of the centerline for each of the two modes by evaluating the cochlear with the mean value of each mode. Most of the variation occurs at the beginning of the basal turn. The first mode ($N = 306$) generates a cochlea with almost a **straight** longitudinal component of the centerline, while the second mode ($N = 681$) generates a cochlea that has been classified by [Avci 2014] as a **rollercoaster**. The classification of [Avci 2014] was created from 16 cadaveric μCT images.

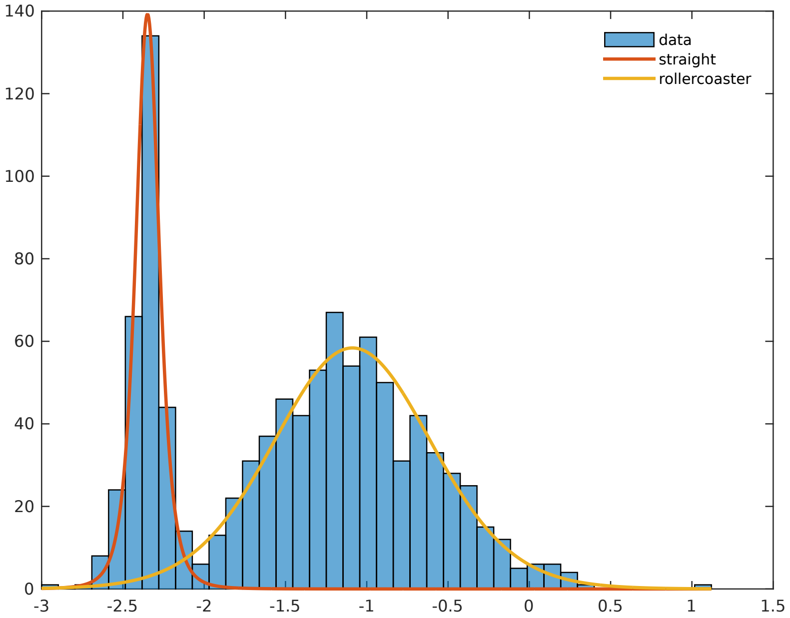


Figure 5.28: Shape parameter w_2

5.3.4 Bilateral Symmetry

The bilateral symmetry was reported by [Reda 2014a], where the hypothesis of symmetry was supported from data acquired on 10 patients. In order to characterize it, we evaluate the shape variability between pairs of left and right cochleae compared with interpatient cochlear shape variability.

The anatomical variability study has confirmed the existence of two categories of cochlear shape, one called "straight" and the other one called "rollercoaster". After classifying each right and left cochleae, we statistically examine the database. Based on the value of w_2, 69% of all the cochleae are categorized as "rollercoaster" (i.e $w_2 > -2.1$), 64% of the patients have the left and right ears within the same category (50% with two "rollercoaster" and 14% with two "straight" cochleae) and 36% of the patients have one cochlea in one category and the other one in the other category.

Note that the fraction of the bilateral cochlea belonging to the same category (64%) is slightly higher than the probability that two independent cochleae belong to the same category (57%), $p(\text{rollercoaster})^2 + p(\text{straight})^2$.

We also characterize the symmetry by evaluating the distance between cochleae in the shape parameters space by using the Mahalanobis distance (Table 5.7) d_M,

$$\bar{d}_M = \frac{\sum_i \sum_j \sqrt{(\theta_S(i) - \theta_S(j))^T \Sigma^{-1} (\theta_S(i) - \theta_S(j))}}{\sum_i \sum_j 1}, \tag{5.39}$$

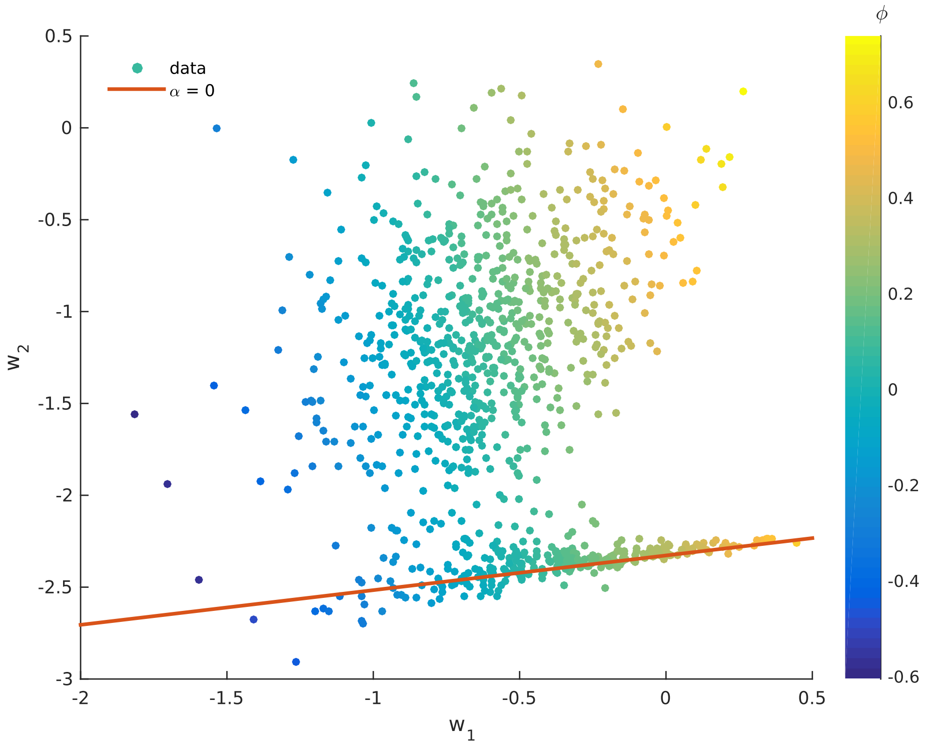
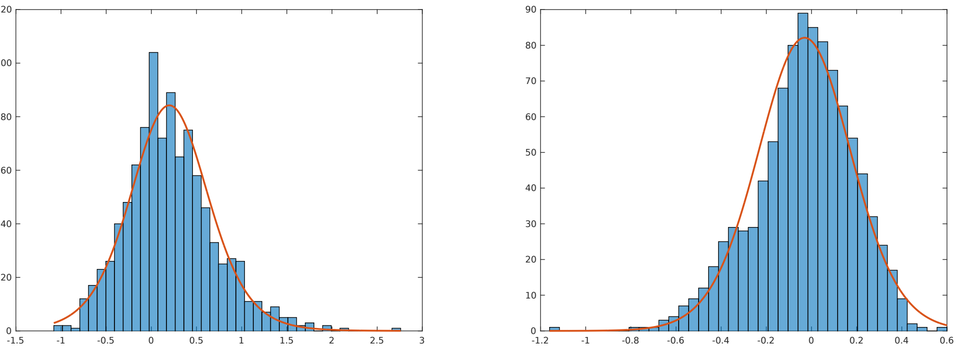


Figure 5.29: Shape parameters of the longitudinal component of the centerline. The right line represents $\alpha(w_1, w_2) = 0$, The color map represents $\phi(w_1, w_2)$. α, correlated with w_2, and ϕ, correlated with w_1, are respectively the initial amplitude and the phase of the exponentially damped sinusoidal part.



(a) (b)

Figure 5.30: Shape parameter a (a) and $\theta_{\max}$ (b)

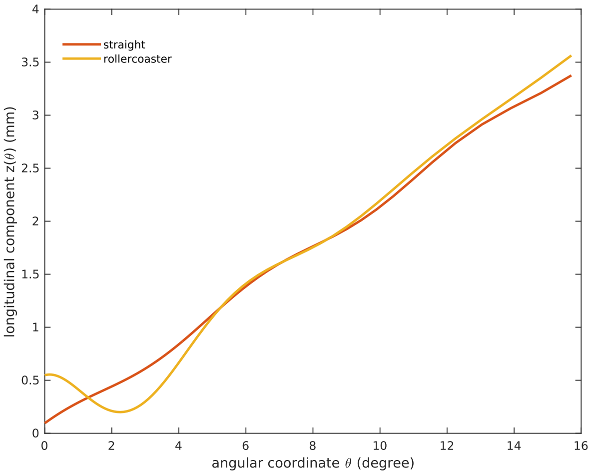


Figure 5.31: Longitudinal components of the centerline generated with the shape parameter w_2 at the mean value of each mode.

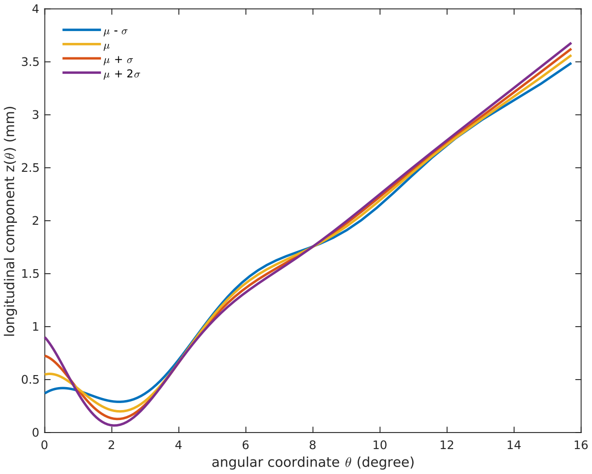


Figure 5.32: Longitudinal component of the centerline generated with values of the shape parameter w_2 taken around the mean of the second model, called “roller-coaster”. Similarly, variations of the parameter w_1 around its mean mainly affect the position of the bump (minimum).

	Mahalanobis distance	Dice score
intrapatient bilateral	0.61	0.91
interpatient	0.74	0.87

Table 5.7: Quantitative evaluation of the shape variability between pairs of left and right cochleae compared with interpatient cochlear shape variability

where $\theta_S(i)$ are the shape parameters (without the pose) of the i^{th} cochlea and Σ is the covariance matrix of $[\theta_S(i)]_{i=1,\dots,987}$ (Table 5.6). For the intrapatient bilateral Mahalanobis i and j were such as they indicate the left and right cochleae of the patient, while for the interpatient Mahalanobis distance i and j are indices of different patients.

Finally to take advantage of the output of the automatic segmentation, we performed fine registration between all segmented CT by removing the rigid transformation, which has been estimated within the cochlear shape parameters. The Dice score between the segmented cochleae was compared (Table 5.7).

5.4 Discussion

We presented a new method for detailed cochlea segmentation from CT images with a joint shape and intensity model. The method was applied to two different datasets, a small one with high-resolution manually segmented μCT on cadaveric patients and a large one based on clinical CT patient images without any ground truth for which a quantitative evaluation was not available.

Dataset #1 For the first dataset we have provided the most common quantitative scores and compared then as fairly as possible with previous methods. It seems that the parametric shape model alone is realistic enough to overcome the limitations of the prior work, pointed out by [Baker 2005]. The segmentation method combining shape and intensity model performed better compared to previous methods with most of the metrics, although the small number of cases ($N \leq 9$) considered make this statement difficult to assert. Despite having a mean surface errors at least 50% larger than our proposed method, SSM-based method performed slightly better with respect to the Hausdorff distance. We can interpret this as a result of the strict constraint of SSM in the allowable shapes. Thus it offers robust performance with limited deformation but poor generalization. We can also note that our model cannot determine the border between the cochlea and the vestibule. Indeed, there is no clear anatomical and visible boundary since the vestibule communicates with the scala vestibuli, and is filled with perilymph and thus having the same intensity than the cochlea. This fuzzy anatomical boundary is an important source of error. To overcome this limitation, our proposed parametric model could be extended by adding the vestibule and the semicircular canals using a 3-torus model similar to what was done by [Bradshaw 2010, Zhang 2014, Kjer 2015b].

Dataset #2 To the best of our knowledge, this is the first time that tridiemensional reconstruction of the cochlea was reported on such a large dataset. The shape variability study performed on 987 cochleae indicates that 2 categories of cochlear shapes can be identified, instead of the 3 categories from a previous study [Avci 2014]. Given the strong bimodal distribution of the data, we could assume that the “intermediate” category is not a proper category and can be corrected by robust modiolar axis estimation. Thanks to the large dataset, a statistical study can be performed and indicates that 69% of the cochleae present a typical “rollercoaster” bump just before 180 degrees.

Several studies [Reda 2014a, Reda 2014b] conclude to a high bilateral symmetry. Our study mitigates this statement, indeed if the interpatient cochlear shape variability is greater than the intrapatient bilateral variability, we have shown that the symmetry is not obvious as 36% of the patient have two cochleae that can be categorized differently with respect to their longitudinal profiles.

Chapter 6

Conclusion

6.1 Current Clinical and Industrial Applications

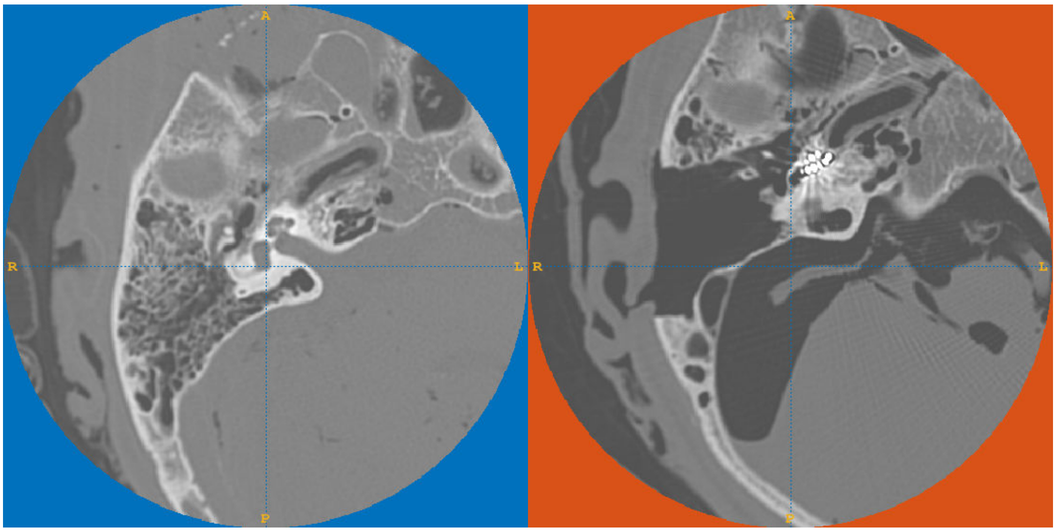
Our knowledge of the cochlear shape and its anatomical variation in the population is currently limited. By studying the anatomical variability, we obtained statistics on parameters that are clinically relevant (number of cochlear turn, scala tympani diameter, maximal insertion depth, etc.). It can provide a better anatomical understanding for the clinicians and suggest electrode array design improvements for cochlear implant manufacturers.

CT imaging is used routinely as a non-invasive method for diagnosis, cochlear implant surgery planning and post-operative evaluation. Automatic cochlear cavities segmentation methods such as the one proposed in this thesis, may help the **diagnosis** by identifying pathological cochlear shapes or **assist preoperatively the surgeon** to choose the optimal electrode design and the optimal drilling trajectory.

Finally, very preliminary study has been performed to **estimate the position of the electrode array** in the post-operative scan with respect to the scala tympani identified in the preoperative scan [Demarcy 2016a]. To illustrate this clinical application, Pre- and postoperative CT images (Fig. 6.1) have been rigidly registered. Postoperative CT images present metal artifacts due to the electrode array (Fig. 6.1b), which can be easily segmented with image thresholding (Fig. 6.2b). Knowing the design of electrode array (Fig. 6.2a), it is possible to deform the model of electrode array so that it shares the same centerline as the skeletonization of the segmentation (Fig. 6.2c). The fusion of the cochlear shape model and the inserted electrode model allows to estimate the position of the electrode relative to the intracochlear cavities and the basilar membrane (Fig. 6.3). An evaluation of a critical factor of the surgery can be assessed: knowing the position and the insertion depth, the electrode array can be optimally activated, by deactivating electrodes located closed to basilar membrane trauma.

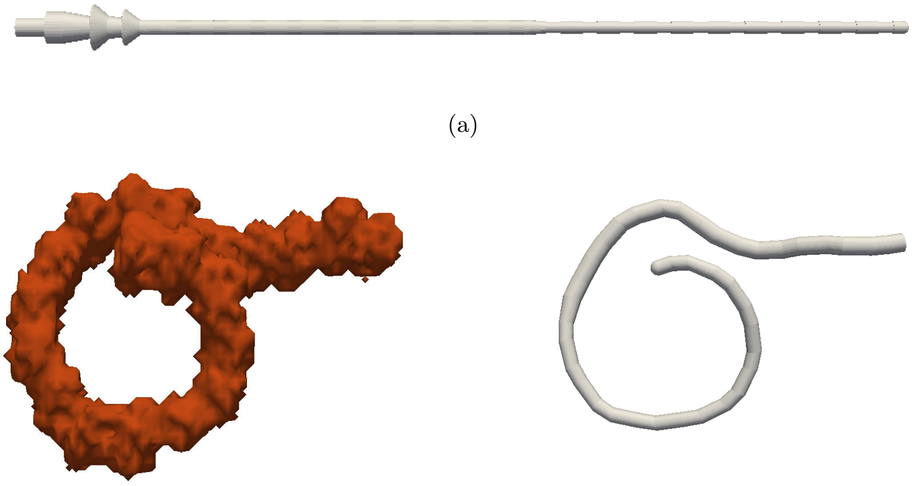
6.2 Contributions

The segmentation of the cochlea from medical images was challenging and several difficulties needed to be overcome. This thesis introduced original and significant contributions, specifically:



(a) (b)

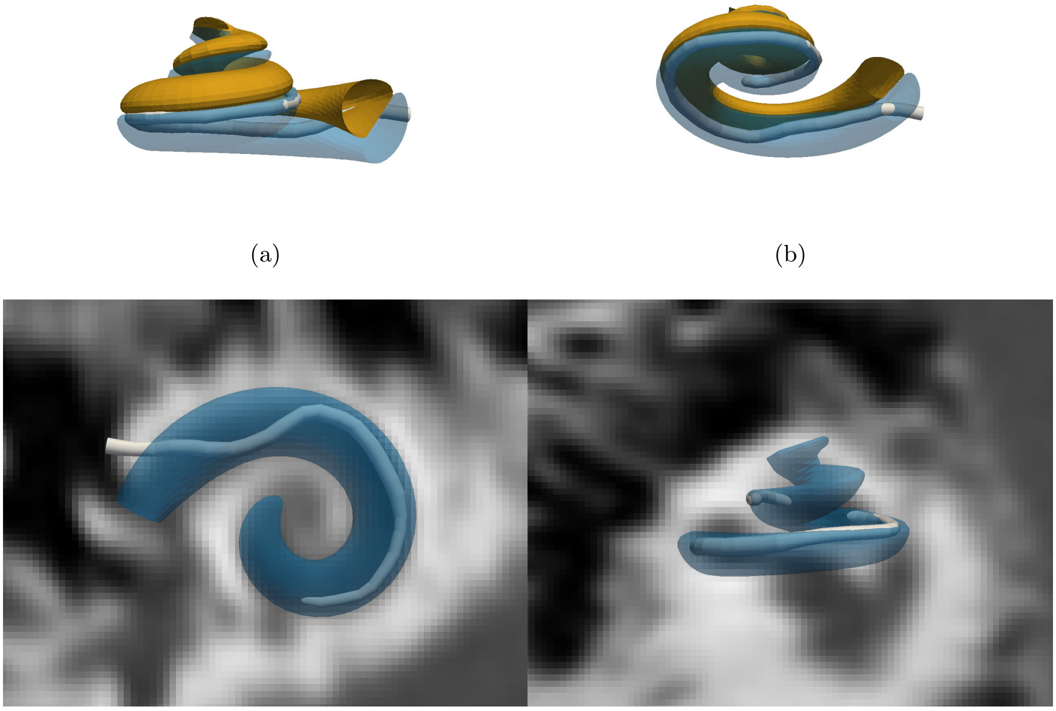
Figure 6.1: (a) Pre- (b) Postoperative CT images



(a)

(b) (c)

Figure 6.2: (a) Electrode array model before insertion. (b) Electrode segmentation from postoperative CT with metal artifacts. (c) Electrode model fitted to postoperative CT.



(a) (b)

(c) (d)

Figure 6.3: Post-implantation electrode position obtained by rigid registration between pre- and postoperative CT image with cochlear and electrode model fitting.

Figure 6.4: Microdissected cochlea performed by [Vandersteen 2015]

6.2.1 Cochlear Shape Description

An automated and reproducible framework for cochlear shape analysis was introduced. Particularly, a robust modiolar axis estimation was proposed from cochlear centerlines. It was applied on segmented μCT images and validated on a synthetic cochlear model, quantitatively compared to manual axis estimation and another algorithm from the literature. A new non-ambiguous moving frame was described and allowed rigorous cross-section measurements and cross-section modeling [Demarcy 2017].

We proposed a new cochlear shape model modeling the scala tympani and the scala vestibuli, fully parametric, realistic, with self-intersection avoidance [Demarcy 2016b]. Compared to previous work, the model, with 22 parameters, represents not only shapes with straight longitudinal component of the centerline but also typical “rollercoaster” shapes [Avci 2014]. The sections are no longer modeled by constant ellipses but half pseudo-cardioids with varying width and height. We proposed as well a compact cochlear shape model with only 4 parameters.

6.2.2 Cochlear Shape Variability Analysis

Thanks to the framework proposed for cochlear shape variability analysis, the cross-section tilt angle was measured from μCT images quantitatively along the cochlear centerline for the first time [Demarcy 2017]. The tilt angle may have implication for occurrence of trauma and cochleostomy site selection [Briggs 2009].

The shape variability was studied with a large database ($N = 987$) and two categories of cochlear shape were identified and their proportions were measured on CT images. 69% of cochleae present a “dip” in the longitudinal component within the first basal turn which can have implications with respect to the surgical gesture and the design of the electrodes.

The bilateral symmetry in cochlear anatomy was quantified and the statement of high symmetry was mitigated. Indeed, if the interpatient cochlear shape variability is greater than the intrapatient bilateral variability, we have shown that the symmetry is not observed very broadly as 36% of the patient have two cochleae that can be categorized differently with respect to their longitudinal profiles.

6.2.3 Cochlear Model-based Segmentation From CT Images

Using the new parametric model, CT-based measurements uncertainty was quantified and compared to the ground truth high-resolution segmentation. In addition to the shape parameters, such as the number of cochlear turns or the cross-section parameters, the uncertainty of the maximal insertion depth given a patient and an electrode array design was quantified. One of the findings of this study was that despite their limited resolution classical CT images could be used by ENT surgeons to improve the surgery by safely selecting a patient-specific CI for instance [Demarcy 2016b].

We introduced a new cochlear segmentation method within a generative probabilistic Bayesian framework for CT images. We unify cochlear shape model fitting with a local appearance model using a model which describes the label of each voxel having shape and appearance parameters as hidden variables. The method applies an iterative expectation-maximization (EM) strategy that interleaves shape model parameters fitting and image segmentation with a mixture of Student's t-distributions. The results show that using the joint intensity and shape model the scala tympani and the scala vestibuli can be segmented from CT images with a mean surface error of 0.12 mm, which is at least 33% lower than previous studies.

6.2.4 Clinical and Industrial Applications

Automatic cochlear segmentation methods can be used to extract parameters that are clinically relevant and provide diagnosis of shape abnormality, for instance pathological cochleae with 1.5 turns instead of the expected 2.5 are diagnosed with cochlear hypoplasia such as the Mondini malformation [Sennaroglu 2002]. We introduced a measurements of the risk of trauma given different electrode array designs (lengths, diameters, flexibilities and shapes) and gave a proof of concept of the estimation of the electrode array position with respect to the cochlea cavities from clinically available CT images [Demarcy 2016a].

6.3 Perspectives

In the continuation of this work of thesis the following perspectives are envisaged.

Methodological Perspectives It is important to predict and analyze the uncertainty of the shape parameters and the estimated clinical measurements: variational Bayesian methods, stochastic sampling or estimation based on the Hessian matrix could be tested.

Modeling and Validation Our proposed parametric model only describes the cochlea and could be extended by adding the vestibule and the semicircular canals, using a 3-torus model similar to what was done by [Bradshaw 2010, Zhang 2014, Kjer 2015b]. Furthermore, the method should be tested on a wider database including pediatric cases as it represents an important cohort of implanted patients. A more thorough evaluation must be performed by more independent raters and could be validated on a larger dataset of segmented high-resolution images (CT and μCT images) of the inner ear to be released by the Hear-EU project [Gerber 2017].

Scientific Applications The contributions of this thesis open the doors for more scientific collaborations. A realistic and patient-specific cochlear could be used to improve electrophysiological simulation studied by Kai Dang from the Inria team *Athena* [Dang 2015] or mechanical insertion simulation studied by Inria team *Defrost* and the Inserm research structure *Minimally invasive and robotic surgical rehabilitation of hearing* (UMR 1159) [Goury 2016]. A study of the correlation between the anatomical position of the implant, the insertion quality (complete, incomplete, traumatic) electrophysiological measurements (interaction in particular) and the functional results of implantation could be carried out by Dr Clair Vandersteen and Dr Julien Lahmar under the supervision of Pr Nicolas Guevara from the Nice University Hospital [Vandersteen 2015]. The segmentation of the cochlear substructure could be validated by comparing the relative position of the electrode array estimated from CT images and observed from microdissection (Fig. 6.4). This requires to apply the segmentation method, to estimate the electrode position CT images acquired from cadaveric temporal bones and to perform a careful microdissection. Under the supervision of Pr Charles Raffaelli, the temporal bone database could be increased to include more pathological and pediatric data. Using our cochlear segmentation method, a comprehensive study of the density of the bony labyrinth could be conducted in order to improved the identification of grade 2 otospongiosis, which is characterized by a calcification of the cochlea.

Clinical and Industrial Perspectives On the clinical and industrial side, a software application could be provided to ENT surgeons enabling them to select the most suitable cochlear implant model for a given patient before surgery. Surgical skills require practice and deep anatomical knowledge. By developing a cochlear insertion simulator, it could be possible to virtually augment the surgical experience before the implantation, with patient specific or randomly generated cochlear models. Indeed due to the complexity of the temporal anatomy and the lack of preoperative landmarks to planned the ideal axis of electrode insertion, it has been shown that experienced clinician determined better insertion approaches [Vandersteen 2015, Torres 2015].

Given the broad potential for clinical and industrial applications, the models and methods developed in this thesis would profit from future research.

List of Publications

The presented work lead to several published and submitted publications.

Journal Papers

- [Demarcy 2017]
  Thomas Demarcy, Clair Vandersteen, Nicolas Guevara, Charles Raffaelli, Dan Gnansia, Nicholas Ayache and Hervé Delingette. *Automated Analysis of Human Cochlea Shape Variability From Segmented μCT images.* Computerized Medical Imaging and Graphics, volume 59, pages 1-12, 2017.

- To be submitted to IEEE Transactions on Medical Imaging
  Thomas Demarcy, Clair Vandersteen, Charles Raffaelli, Nicolas Guevara, Dan Gnansia, Nicholas Ayache and Hervé Delingette. *Joint Shape and Intensity Model-Based Segmentation of CT Images of the Cochlea.*

- [Vandersteen 2015]
  Clair Vandersteen, **Thomas Demarcy**, Coralie Roger, Éric Fontas, Charles Raffaelli, Nicholas Ayache, Hervé Delingette and Nicolas Guevera. *Impact of the Surgical Experience on Cochleostomy Location : A Comparative Temporal Bone Study Between Endaural and Posterior Tympanotomy Approaches for Cochlear Implantation.* European Archives of Oto-Rhino-Laryngology, volume 273, issue 9, pages 2355–2361, 2015.

Peer-Reviewed Archived Conference and Workshop Papers

- [Demarcy 2016b] (oral and proceedings)
  Thomas Demarcy, Clair Vandersteen, Nicolas Guevara, Charles Raffaelli, Dan Gnansia, Nicholas Ayache and Hervé Delingette. *Uncertainty Quantification of Cochlear Implant Insertion from CT images.* Clinical Image-Based Procedures. Translational Research in Medical Imaging: 5th International Workshop. Held in Conjunction with MICCAI 2016, Athens, Greece, pages 27-35, 2016.

- [Gnansia 2016] (oral and proceedings)
  Dan Gnansia, **Thomas Demarcy**, Clair Vandersteen, Charles Raffaelli, Nicolas Guevara, Hervé Delingette and Nicholas Ayache. *Optimal Electrode Diameter in Relation to Volume of the Cochlea.* European Annals of Otorhinolaryngology, Head and Neck Diseases, volume 133, pages S66-S67, 2016.

Conference Abstracts

- [Demarcy 2016a] (oral)
  Thomas Demarcy, Clair Vandersteen, Dan Gnansia, Charles Raffaelli, Nicholas Ayache, Hervé Delingette, Nicolas Guevera. *Estimation of Postoperative Cochlear Implant Electrode-Array Position From Clinical Computed Tomography.* Annales Française d'Oto-Rhino-Laryngologie et de Pathologie Cervico-faciale, 2016.

- [Vandersteen 2014] (best poster award)
  Clair Vandersteen, **Thomas Demarcy**, Hervé Delingette, Charles Raffaelli, Jonathan Laudanski, Thierry Pourcher, Jacques Darcourt, Philippe Franker, Dan Gnansia, Nicholas Ayache and Nicolas Guevara. *Teaching Tool for Advanced Visualization of Temporal Bone Structures by Fusion of μCT and CT Scan Images.* In 8th International Symposium of Objective Measures in Auditory Implants, 2014.

Appendix A

Equations for the discretization scheme of the centerline

In practice the continuous curve $\mathbf{x}$ is approximated by a piecewise linear curve denoted $\{\mathbf{x}_i\}_{i=1}^{n}$.

In subsection 3.2.3 and more specifically in Eq. 4.3, $\mathbf{x}$ has also a temporal dimension t. $\{\mathbf{x}_i\}_{i=1}^{n}$ are rediscretized at each time step using linear interpolation to obtain all line segments $\mathbf{x}_i\mathbf{x}_{i+1}$ of equal length Δs. Derivatives were approximated using finite difference methods,

$$\frac{\partial \mathbf{x}(t)}{\partial t} = \frac{-\mathbf{x}(t) + \mathbf{x}(t+\Delta t)}{\Delta t} + O(\Delta t) \tag{A.1}$$

$$\frac{\partial^2 \mathbf{x}_i}{\partial s^2} = \frac{\mathbf{x}_{i-1} - 2\mathbf{x}_i + \mathbf{x}_{i+1}}{(\Delta s)^2} + O((\Delta s)^2) \tag{A.2}$$

$$\frac{\partial^4 \mathbf{x}_i}{\partial s^4} = \frac{\mathbf{x}_{i-2} - 4\mathbf{x}_{i-1} + 6\mathbf{x}_i - 4\mathbf{x}_{i+1} + \mathbf{x}_{i+2}}{(\Delta s)^4} + O((\Delta s)^2) \tag{A.3}$$

$\nabla P(\mathbf{x})$ was approximated using trilinear interpolation onto the gradient maps in the three dimensions computed from the distance map P using the Prewitt operator.

In subsection 3.2.4, line segments $\mathbf{x}_i\mathbf{x}_{i+1}$ are not constant. The left and right discrete tangent unit vectors, respectively $\mathbf{t}_{i-}$ and $\mathbf{t}_{i+}$ are estimated:

$$\mathbf{t}_{i-} = \frac{\mathbf{x}_i - \mathbf{x}_{i-1}}{\|\mathbf{x}_i - \mathbf{x}_{i-1}\|} \tag{A.4}$$

$$\mathbf{t}_{i+} = \frac{\mathbf{x}_{i+1} - \mathbf{x}_i}{\|\mathbf{x}_{i+1} - \mathbf{x}_i\|}. \tag{A.5}$$

The discrete tangent and normal unit vectors, $\mathbf{t}_i$ and $\mathbf{n}_i$ can be calculated as follows:

$$\mathbf{t}_i = \frac{\mathbf{t}_{i-} + \mathbf{t}_{i+}}{\|\mathbf{t}_{i-} + \mathbf{t}_{i+}\|} \tag{A.6}$$

$$\mathbf{n}_i = \frac{\mathbf{t}_{i+} - \mathbf{t}_{i-}}{\|\mathbf{t}_{i+} - \mathbf{t}_{i-}\|}. \tag{A.7}$$

The discrete curvature κ_i can be set as,

$$\kappa_i = \frac{2\left\|\mathbf{t}_{i+} - \mathbf{t}_{i-}\right\|}{\left\|\mathbf{x}_{i+1} - \mathbf{x}_i\right\| + \left\|\mathbf{x}_i - \mathbf{x}_{i-1}\right\|} \tag{A.8}$$

using a discretization of Serret-Frenet formulas or,

$$\kappa_i = \frac{2\left\|(\mathbf{x}_{i+1} - \mathbf{x}_i) \times (\mathbf{x}_i - \mathbf{x}_{i-1})\right\|}{\left\|\mathbf{x}_{i+1} - \mathbf{x}_i\right\| \left\|\mathbf{x}_i - \mathbf{x}_{i-1}\right\| \left\|\mathbf{x}_{i+1} - \mathbf{x}_{i-1}\right\|} \tag{A.9}$$

as the inverse of the circumradius of the triangle $\mathbf{x}_{i-1}\,\mathbf{x}_i\,\mathbf{x}_{i+1}$.

Appendix B

Additional algorithm: Point to triangle signed Euclidean distance

Algorithm 4 can be interpreted as follow: Lines 1 to 5, compute the normal n to the plane defined by $\{v_1, v_2, v_3\}$, the signed distance between d between the point p and the plane and its projection p_0. Lines 6 to 8, express the projected point p_0 in the barycentric coordinate system of the triangle such as,

$$p_0 = \lambda_1 \, v_1 + \lambda_2 \, v_2 + \lambda_3 \, v_3 \tag{B.1}$$

The coefficients λ are used to determine if the projected point is inside or outside the triangle. If it is outside, we check if the closest point in the triangle belongs to a vertex (lines 12 and 16) or an edge (line 18).

We use the convention that surface normal are oriented toward the outside and that the signed distance is positive inside, so line 22 returns the opposite value.

Algorithm 4: `PointToTriangle`(p, v_1, v_2, v_3)
Point to triangle signed Euclidean distance

```
Input: point, $p$
oriented triangle, $\{v_1, v_2, v_3\}$
Output: signed Euclidean distance $d$
$n \leftarrow (v_2 - v_1) \times (v_3 - v_1)$;
$\lambda \leftarrow \|n\|$;
$n \leftarrow n/\lambda$;
$d \leftarrow (p - v_1) \cdot n$;
$p_0 \leftarrow p - d\, n$;
$\lambda_1 \leftarrow \det[v_2 - p_0, v_3 - p_0, n]/\lambda$;
$\lambda_2 \leftarrow \det[v_3 - p_0, v_1 - p_0, n]/\lambda$;
$\lambda_3 \leftarrow 1 - \lambda_1 - \lambda_2$;
sort $\{(\lambda_i, v_i)\}$ such as $\lambda_1 \geq \lambda_2 \geq \lambda_3$;
if $\lambda_3 < 0$ then
    if $\lambda_2 < 0$ then
        $d \leftarrow \|p - v_1\| \operatorname{sgn}(d)$ ;
    else
        $\lambda_2 \leftarrow ((p - v_1) \cdot (v2 - v1))/((v2 - v1) \cdot (v2 - v1))$;
        if $\lambda_2 < 0$ then
            $d \leftarrow \|p - v_1\| \operatorname{sgn}(d)$ ;
        else
            $d \leftarrow \|p - ((1 - \lambda_2)v_1 + \lambda_2\, v_2)\| \operatorname{sgn}(d)$ ;
        end
    end
end
$d \leftarrow -d$;
```

Appendix C

Additional tables for model fitting and segmentation results obtained on dataset #1

	Dice score		Hausdorff distance (mm)							
			50%		68%		95%		100%	
#	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$
1	0.79	0.85	0.14	0.10	0.21	0.15	0.43	0.35	1.17	0.93
2	0.80	0.87	0.13	0.08	0.20	0.12	0.44	0.31	1.07	1.02
3	0.75	0.78	0.14	0.12	0.21	0.18	0.50	0.49	1.30	1.12
4	0.79	0.82	0.12	0.10	0.19	0.15	0.49	0.47	1.19	1.42
5	0.78	0.83	0.14	0.11	0.21	0.18	0.55	0.44	1.17	1.15
6	0.78	0.84	0.18	0.13	0.26	0.19	0.54	0.38	1.13	1.17
7	0.79	0.85	0.14	0.10	0.21	0.15	0.53	0.36	1.13	1.05
8	0.80	0.86	0.12	0.08	0.20	0.14	0.52	0.44	1.33	1.02
9	0.82	0.85	0.11	0.10	0.17	0.15	0.38	0.35	3.30	2.91
mean	0.79	0.84	0.13	0.10	0.21	0.16	0.48	0.40	1.42	1.31

Table C.1: Model fitting and segmentation results obtained on dataset #1 between $Z_0/\mathrm{CT}/f_g$, the manual segmentation of the cochlea on CT and $S/\mathrm{CT}/f_g$ and $\hat{Z}/\mathrm{CT}/f_g$, respectively the fitted shape and the automatic segmentation of the cochlea on CT.

	Dice score		
#	S	$\hat{Z}$	Z_0
1	0.71	0.78	0.74
2	0.69	0.76	0.73
3	0.67	0.76	0.73
4	0.71	0.80	0.78
5	0.71	0.80	0.77
6	0.69	0.78	0.76
7	0.66	0.75	0.75
8	0.72	0.80	0.79
9	0.73	0.82	0.81
mean	0.70	0.78	0.76

Table C.2: Model fitting and segmentation results obtained on dataset #1 between $Z_0/\mu\text{CT}/ST \cup SV$, the union of the manual segmentation of the ST and SV on μCT and $S/\text{CT}/f_g$, $\hat{Z}/\text{CT}/f_g$ and $Z_0/\text{CT}/f_g$ respectively the fitted shape, the automatic segmentation and the manual segmentation of the cochlea on CT.

	Dice score		**Hausdorff distance (mm)** 50%		68%		95%		100%	
#	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$
1	0.72	0.76	0.11	0.08	0.18	0.15	0.44	0.39	0.80	0.76
2	0.67	0.73	0.16	0.10	0.22	0.17	0.45	0.40	0.87	0.78
3	0.70	0.74	0.12	0.09	0.18	0.15	0.39	0.39	0.89	1.03
4	0.73	0.78	0.10	0.07	0.16	0.12	0.38	0.35	1.00	0.96
5	0.73	0.77	0.10	0.08	0.16	0.13	0.40	0.38	1.27	1.26
6	0.73	0.77	0.10	0.08	0.16	0.14	0.43	0.42	0.72	0.74
7	0.71	0.76	0.13	0.08	0.18	0.14	0.33	0.31	0.72	0.69
8	0.75	0.80	0.09	0.06	0.15	0.11	0.39	0.36	0.91	0.90
9	0.73	0.79	0.10	0.07	0.17	0.12	0.37	0.32	0.75	0.69
mean	0.72	0.77	0.11	0.08	0.17	0.14	0.40	0.37	0.88	0.87

Table C.3: Model fitting and segmentation results obtained on dataset #1 between $Z_0/\mu\text{CT}/ST$, the manual segmentation of the ST on μCT and $S/\text{CT}/ST$ and $\hat{Z}/\text{CT}/ST$, respectively the fitted shape and the automatic segmentation of the cochlea on CT.

	Dice score		**Hausdorff distance (mm)**							
			50%		68%		95%		100%	
#	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$	S	$\hat{Z}$
1	0.67	0.73	0.12	0.10	0.19	0.15	0.40	0.34	0.77	0.82
2	0.67	0.68	0.11	0.09	0.18	0.16	0.38	0.39	0.60	0.73
3	0.70	0.74	0.10	0.08	0.15	0.14	0.38	0.34	1.19	1.19
4	0.71	0.73	0.10	0.09	0.16	0.15	0.33	0.31	0.77	0.51
5	0.70	0.73	0.10	0.08	0.16	0.15	0.36	0.36	0.71	0.69
6	0.66	0.75	0.12	0.08	0.19	0.14	0.37	0.32	0.64	1.39
7	0.74	0.77	0.08	0.06	0.12	0.12	0.32	0.28	0.75	0.68
8	0.72	0.73	0.08	0.08	0.14	0.14	0.40	0.37	2.45	2.08
9	0.71	0.75	0.10	0.09	0.18	0.15	0.38	0.33	0.72	0.62
mean	0.70	0.73	0.10	0.08	0.16	0.14	0.37	0.34	0.96	0.97

Table C.4: Model fitting and segmentation results obtained on dataset #1 between $Z_0/\mu\text{CT}/SV$, the manual segmentation of the ST on μCT and $S/\text{CT}/SV$ and $\hat{Z}/\text{CT}/SV$, respectively the fitted shape and the automatic segmentation of the cochlea on CT.

Bibliography

[Adunka 2004] Oliver Adunka, Marc H Unkelbach, Martin Mack, Markus Hambek, Wolfgang Gstoettner and Jan Kiefer. *Cochlear implantation via the round window membrane minimizes trauma to cochlear structures: a histologically controlled insertion study.* Acta Oto-Laryngologica, vol. 124, no. 7, pages 807–12, sep 2004. (Cited on page 18.)

[Aschendorff 2007] A Aschendorff, J Kromeier, T Klenzner and Roland Laszig. *Quality control after insertion of the nucleus contour and contour advance electrode in adults.* Ear and Hearing, 2007. (Cited on page 18.)

[Ashburner 2005] John Ashburner and Karl J. Friston. *Unified segmentation.* NeuroImage, vol. 26, no. 3, pages 839–851, 2005. (Cited on page 82.)

[Avci 2014] Ersin Avci, Tim Nauwelaers, Thomas Lenarz, Volkmar Hamacher and Andrej Kral. *Variations in microanatomy of the human cochlea.* The Journal of Comparative Neurology, vol. 00, pages 1–17, mar 2014. (Cited on pages 19, 33, 44, 45, 48, 49, 50, 113, 117 and 122.)

[Baker 2004] Gavin Baker and Nick Barnes. *Principal flow for tubular objects with non-circular cross-sections.* Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., pages 750–753 Vol.3, 2004. (Cited on pages 24 and 25.)

[Baker 2005] Gavin Baker and Nick Barnes. *Model-image registration of parametric shape models: fitting a shell to the cochlea.* Insight Journal, 2005. (Cited on pages 16, 54, 55, 59, 80, 105 and 117.)

[Baker 2008] Gavin Baker. *Tracking, modelling and registration of anatomical objects: the human cochlea.* PhD thesis, The University of Melbourne, 2008. (Cited on pages vi, 19, 27, 28, 49, 60, 105 and 106.)

[Bartling 2005] Soenke Heinrich Bartling, Kersten Peldschus, Thomas Rodt, Florian Kral, Herbert Matthies, Ron Kikinis and Hartmut Becker. *Registration and fusion of CT and MRI of the temporal bone.* Journal of computer assisted tomography, vol. 29, no. 3, pages 305–10, 2005. (Cited on pages 15 and 85.)

[Biedron 2010] Slavomir Biedron, Andreas Prescher, Justus Ilgner and Martin Westhofen. *The internal dimensions of the cochlear scalae with special reference to cochlear electrode insertion trauma.* Otology & Neurotology, vol. 31, no. 5, pages 731–7, jul 2010. (Cited on pages 18, 19, 33 and 49.)

[Bradshaw 2010] Andrew P Bradshaw, Ian S Curthoys, Michael J Todd, John S Magnussen, David S Taubman, Swee T Aw and G Michael Halmagyi. *A mathematical model of human semicircular canal geometry: a new basis for* *interpreting vestibular physiology.* Journal of the Association for Research in Otolaryngology, vol. 11, no. 2, pages 145–59, jun 2010. (Cited on pages 117 and 123.)

[Braun 2012] Katharina Braun, Frank Böhnke and Thomas Stark. *Three-dimensional representation of the human cochlea using micro-computed tomography data: presenting an anatomical model for further numerical calculations.* Acta oto-laryngologica, vol. 132, no. 6, pages 603–13, jun 2012. (Cited on pages 24 and 48.)

[Bresenham 1965] J E Bresenham. *Algorithm for computer control of a digital plotter.* IBM Systems Journal, vol. 4, no. 1, pages 25–30, 1965. (Cited on pages 101 and 102.)

[Briggs 2009] Robert J.S Briggs, Michael Tykocinski, Katrina Stidham and Joseph B Roberson. *Cochleostomy site: Implications for electrode placement and hearing preservation.* Acta Oto-Laryngologica, jul 2009. (Cited on pages 51 and 122.)

[Broyden 1970] Charles George Broyden. *The convergence of a class of double-rank minimization algorithms.* IMA Journal of Applied Mathematics, vol. 6, no. 1, pages 76–90, 1970. (Cited on page 101.)

[Buytaert 2011] Jan A N Buytaert, Wasil H M Salih, Manual Dierick, Patric Jacobs and Joris J J Dirckx. *Realistic 3D computer model of the gerbil middle ear, featuring accurate morphology of bone and soft tissue structures.* JARO - Journal of the Association for Research in Otolaryngology, vol. 12, no. 6, pages 681–696, 2011. (Cited on page 19.)

[Ceresa 2014] Mario Ceresa, Nerea Mangago Lopez, Hector Dejea Velardo, Noemi Carranza Herrezuelo, Pavel Mistrik, Hans Martin Kjer, Sergio Vera, Rasmus Reinhold Paulsen and Miguel Angel González Ballester. *Patient-Specific Simulation of Implant Placement and Function for Cochlear Implantation Surgery Planning.* Medical Image Computing and Computer-Assisted Intervention, pages 49–56, 2014. (Cited on page 19.)

[Clark 2011] James R Clark, Frank M Warren and Jake J Abbott. *A Scalable Model for Human Scala-Tympani Phantoms.* Journal of Medical Devices, vol. 5, no. 1, page 014501, 2011. (Cited on pages 30, 34, 45 and 59.)

[Cohen 1996] Lawrence T Cohen, Jin Xu, Shi Ang Xu and Graeme M Clark. *Improved and simplified methods for specifying positions of the electrode bands of a cochlear implant array.* The American Journal of Otology, vol. 17, no. 6, pages 859–865, 1996. (Cited on pages 19, 28, 32, 35, 49, 58 and 98.)

[Cootes 1995] Tim F Cootes, Christopher J Taylor, David H Cooper and Jim Graham. *Active Shape Models-Their Training and Application.* Computer Vision and Image Understanding, vol. 61, no. 1, pages 38–59, jan 1995. (Cited on page 81.)

[Criminisi 2008] Antonio Criminisi, Toby Sharp and Andrew Blake. *GeoS: Geodesic Image Segmentation.* ECCV, pages 99–112, 2008. (Cited on pages 24 and 85.)

[Dang 2015] Kai Dang, Maureen Clerc, Clair Vandersteen, Nicolas Guevara and Dan Gnansia. *In situ validation of a parametric model of electrical field distribution in an implanted cochlea.* International IEEE/EMBS Conference on Neural Engineering, NER, vol. 2015-July, pages 667–670, 2015. (Cited on page 123.)

[Demarcy 2016a] Thomas Demarcy, Clair Vandersteen, Dan Gnansia, Charles Raffaelli, Nicholas Ayache, Hervé Delingette and Nicolas Guevara. *Estimation of postoperative cochlear implant electrode-array position from clinical computed tomography.* Annales françaises d'Oto-rhino-laryngologie et de Pathologie Cervico-faciale, oct 2016. (Cited on pages 3, 119, 123 and 126.)

[Demarcy 2016b] Thomas Demarcy, Clair Vandersteen, Charles Raffaelli, Dan Gnansia, Nicolas Guevara, Nicholas Ayache and Hervé Delingette. Uncertainty Quantification of Cochlear Implant Insertion from CT Images, pages 27–35. Springer International Publishing, Cham, 2016. (Cited on pages 3, 53, 122 and 125.)

[Demarcy 2017] Thomas Demarcy, Clair Vandersteen, Nicolas Guevara, Charles Raffaelli, Dan Gnansia, Nicholas Ayache and Hervé Delingette. *Automated analysis of human cochlea shape variability from segmented μ CT images.* Computerized Medical Imaging and Graphics, vol. 59, pages 1–12, 2017. (Cited on pages 2, 17, 122 and 125.)

[Deschamps 2001] Thomas Deschamps and Laurent D Cohen. *Fast extraction of minimal paths in 3D images and applications to virtual endoscopy.* Medical Image Analysis, vol. 5, no. 4, pages 281–299, dec 2001. (Cited on page 25.)

[Dijkstra 1959] E. W. Dijkstra. *A note on two problems in connexion with graphs.* Numerische Mathematik, vol. 1, pages 269–271, 1959. (Cited on page 26.)

[Dimopoulos 1990] Panos Dimopoulos and Catharina Muren. *Anatomic variations of the cochlea and relations to other temporal bone structures.* Acta Radiologica, vol. 31, no. 5, pages 439–44, sep 1990. (Cited on pages 18 and 86.)

[Erixon 2009] Elsa Erixon, Herman Högstorp, Karin Wadin and Helge Rask-Andersen. *Variational anatomy of the human cochlea: implications for cochlear implantation.* Otology & Neurotology, vol. 30, no. 1, pages 14–22, jan 2009. (Cited on pages 18, 19, 33, 48 and 49.)

[Escudé 2006] Bernard Escudé, Chris James, Olivier Deguine, Nadine Cochard, Elias Eter and Bernard Fraysse. *The size of the cochlea and predictions* *of insertion depth angles for cochlear implant electrodes.* Audiology & Neurotology, vol. 11 Suppl 1, no. suppl 1, pages 27–33, jan 2006. (Cited on pages 19, 27 and 51.)

[Eshraghi 2003] Adrien A Eshraghi, Nathaniel W Yang and Thomas J Balkany. *Comparative study of cochlear damage with three perimodiolar electrode designs.* The Laryngoscope, vol. 113, no. 3, pages 415–9, mar 2003. (Cited on page 51.)

[Eslami 2013a] Abouzar Eslami, Athanasios Karamalis, Amin Katouzian and Nassir Navab. *Segmentation by retrieval with guided random walks: application to left ventricle segmentation in MRI.* Medical Image Analysis, vol. 17, no. 2, pages 236–53, feb 2013. (Cited on page 82.)

[Eslami 2013b] S M Ali Eslami. *Generative Probabilistic Models for Object Segmentation.* PhD thesis, School of Informatics University of Edinburgh, 2013. (Cited on page 82.)

[Fayad 2006] Jose N Fayad and Fred H Linthicum. *Multichannel cochlear implants: relation of histopathology to performance.* The Laryngoscope, vol. 116, no. 8, pages 1310–20, aug 2006. (Cited on page 18.)

[Finley 2009] Charles C Finley, Timothy A Holden, Laura K Holden, Bruce R Whiting, Richard A Chole, J Gail Neely, Timothy E Hullar and Margaret W Skinner. *NIH Public Access.* October, vol. 29, no. 7, pages 920–928, 2009. (Cited on page 18.)

[Fletcher 1970] Roger Fletcher. *A new approach to variable metric algorithms.* The computer journal, vol. 13, no. 3, pages 317–322, 1970. (Cited on page 101.)

[Fowler 1992] Deborah R. Fowler, Hans Meinhardt and Przemyslaw Prusinkiewicz. *Modeling seashells.* ACM SIGGRAPH Computer Graphics, vol. 26, no. 2, pages 379–387, jul 1992. (Cited on pages 41 and 59.)

[Franz 2014] Daniela Franz, Mathias Hofer, Matthias Pfeifle, Markus Pirlich, Marc Stamminger and Thomas Wittenberg. *Wizard-Based Segmentation for Cochlear Implant Planning.* In Thomas Martin Deserno, Heinz Handels, Hans-Peter Meinzer and Thomas Tolxdorff, editors, Bildverarbeitung für die Medizin 2014 SE - 49, Informatik aktuell, pages 258–263. Springer Berlin Heidelberg, 2014. (Cited on page 20.)

[Gansca 2002] I Gansca, W F Bronsvoort, G Coman and L Tambulea. *Self-intersection avoidance and integral properties of generalized cylinders.* Computer Aided Geometric Design, vol. 19, no. 9, pages 695–707, dec 2002. (Cited on page 67.)

[Gerber 2017] Nicolas Gerber, Mauricio Reyes, Livia Barazzetti, Hans Martin Kjer, Sergio Vera, Martin Stauber, Pavel Mistrik, Mario Ceresa, Nerea Mangado, Wilhelm Wimmer, Thomas Stark, Rasmus Paulsen, Stefan Weber, Marco Caversaccio and Miguel Angel González Ballester. *Multiscale Imaging Dataset of the Human Inner Ear*. To be published, 2017. (Cited on page 123.)

[Gibson 2012] Daren Gibson, Michael Brian Gluth, Andy Whyte and Marcus David Atlas. *Rotation of the osseous spiral lamina from the hook region along the basal turn of the cochlea: results of a magnetic resonance image anatomical study using high-resolution DRIVE sequences.* Surgical and radiologic anatomy : SRA, vol. 34, no. 8, pages 781–5, oct 2012. (Cited on pages 19 and 48.)

[Gnansia 2016] Dan Gnansia, Thomas Demarcy, Clair Vandersteen, Charles Raffaelli, Nicolas Guevara, Hervé Delingette and Nicholas Ayache. *Optimal electrode diameter in relation to volume of the cochlea.* European Annals of Otorhinolaryngology, Head and Neck Diseases, vol. 133, pages S66–S67, 2016. (Cited on pages 3 and 125.)

[Goldfarb 1970] Donald Goldfarb. *A family of variable-metric methods derived by variational means.* Mathematics of computation, vol. 24, no. 109, pages 23–26, 1970. (Cited on page 101.)

[Gooya 2011] A Gooya, G Biros and Christos Davatzikos. *Deformable Registration of Glioma Images Using EM Algorithm and Diffusion Reaction Modeling.* Medical Imaging, IEEE . . . , vol. 30, no. 2, pages 375–390, 2011. (Cited on page 82.)

[Goury 2016] Olivier Goury, Yann Nguyen, Renato Torres, Jeremie Dequidt and Christian Duriez. Numerical Simulation of Cochlear-Implant Surgery: Towards Patient-Specific Planning, pages 500–507. Springer International Publishing, Cham, 2016. (Cited on page 124.)

[Gray 1918] Henry Gray. Anatomy of the human body. Lea & Febiger, 1918. (Cited on page 9.)

[Greenwood 1990] Donald D Greenwood. *A cochlear frequency-position function for several species—29 years later.* The Journal of the Acoustical Society of America, vol. 87, no. 6, pages 2592–2605, 1990. (Cited on page 8.)

[Gulya 1996] A J Gulya and R L Steenerson. *The scala vestibuli for cochlear implantation. An anatomic study.* Archives of Otolaryngology - Head & Neck Surgery, vol. 122, no. 2, pages 130–132, feb 1996. (Cited on pages 18, 19, 24 and 49.)

[Gunz 2012] Philipp Gunz, Marissa Ramsier, Melanie Kuhrig, Jean-Jacques Hublin and Fred Spoor. *The mammalian bony labyrinth reconsidered, introducing* *a comprehensive geometric morphometric approach.* Journal of Anatomy, vol. 220, no. 6, pages 529–43, jun 2012. (Cited on pages 19, 24 and 25.)

[Hardy 1938] Mary Hardy. *The length of the organ of Corti in man.* American Journal of Anatomy, vol. 62, no. 2, pages 291–311, jan 1938. (Cited on pages 18 and 19.)

[Hatsushika 1990] S-I Hatsushika, R K Shepherd, Y C Tong, Graeme M Clark and S Funasaka. *Dimensions of the Scala Tympani in the Human and Cat with Reference to Cochlear Implants.* Annals of Otology, Rhinology & Laryngology, vol. 99, no. 11, pages 871–876, nov 1990. (Cited on pages 18 and 19.)

[Heimann 2007] Tobias Heimann, Sascha Münzing, Hans-Peter Meinzer and Ivo Wolf. *A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation.* Inf Process Med Imaging, vol. 20, pages 1–12, 2007. (Cited on page 81.)

[Heimann 2009] Tobias Heimann and Hans-Peter Meinzer. *Statistical shape models for 3D medical image segmentation: a review.* Medical Image Analysis, vol. 13, no. 4, pages 543–63, aug 2009. (Cited on page 59.)

[Henk 2012] Martin Henk. *Löwner-John Ellipsoids.* Optimization Stories, Documenta Mathematica, vol. I, no. 2012, pages 95–106, 2012. (Cited on page 65.)

[Jackler 2009] Robert K Jackler, William M Luxfor and William F House. *Congenital malformations of the inner ear: A classification based on embryogenesis.* The Laryngoscope, vol. 97, no. S40, pages 2–14, oct 2009. (Cited on page 49.)

[James 2005] Chris James, Klaus Albegger, Rolf Battmer, Sandro Burdo, Naima Deggouj, Olivier Deguine, Norbert Dillier, Michel Gersdorff, Roland Laszig, Thomas Lenarz, Manuel Manrique Rodriguez, Michel Mondain, F Erwin Offeciers, Ángel Ramos Macías, Richard Ramsden, Olivier Sterkers, Ernst Von Wallenberg, Benno Weber and Bernard Fraysse. *Preservation of residual hearing with cochlear implantation: How and why.* Acta Oto-Laryngologica, vol. 125, no. 5, pages 481–491, may 2005. (Cited on page 18.)

[Jenkinson 2001] Mark Jenkinson and Stephen M Smith. *A global optimisation method for robust affine registration of brain images.* Medical Image Analysis, vol. 5, no. 2, pages 143–156, 2001. (Cited on page 88.)

[Jenkinson 2002] Mark Jenkinson, Peter Bannister, Michael Brady and Stephen M Smith. *Improved optimization for the robust and accurate linear registration and motion correction of brain images.* NeuroImage, vol. 17, no. 2, pages 825–841, 2002. (Cited on page 88.)

[John 2014] Fritz John. Extremum Problems with Inequalities as Subsidiary Conditions, pages 197–215. Springer Basel, Basel, 2014. (Cited on page 65.)

[Kass 1988] Michael Kass, Andrew Witkin and Demetri Terzopoulos. *Snakes: Active contour models.* International Journal of Computer Vision, vol. 1, no. 4, pages 321–331, 1988. (Cited on pages 24 and 26.)

[Kawano 1996] A Kawano, H L Seldon and Graeme M Clark. *Computer-Aided Three-Dimensional Reconstruction in Human Cochlear Maps: Measurement of the Lengths of Organ of Corti, Outer Wall, Inner Wall, and Rosenthal's Canal.* Annals of Otology, Rhinology & Laryngology, vol. 105, no. 9, pages 701–709, sep 1996. (Cited on pages 19, 48 and 49.)

[Ketten 1990] Darlene R Ketten and Douglas Wartzok. *Three-Dimensional Reconstructions of the Dolphin Ear.* In JeanetteA. Thomas and RonaldA. Kastelein, editors, Sensory Abilities of Cetaceans SE - 6, volume 196 of *NATO ASI Series*, pages 81–105. Springer US, 1990. (Cited on pages 57 and 58.)

[Ketten 1998] Darlene R Ketten, Margaret W Skinner, Ge Wang, Michael W Vannier, George A Gates and J Gail Neely. *In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays.* Annals of Otology, Rhinology & Laryngology, vol. 175, no. 12, page 1998, nov 1998. (Cited on page 58.)

[Kjer 2014a] Hans Martin Kjer, Jens Fagertun, Sergio Vera, Miguel Angel González Ballester and Rasmus Reinhold Paulsen. *Shape modelling of the inner ear from micro-CT data.* Shape Symposium, 2014. (Cited on pages 16, 59, 80 and 81.)

[Kjer 2014b] Hans Martin Kjer, Sergio Vera, Frederic Pérez, Miguel Angel González Ballester and Rasmus Reinhold Paulsen. *Semi-automatic anatomical measurements on microCT 3D surface models.* In International Conference on Cochlear Implants and Other Implantable Auditory Technologies, Munich, Germany, page 711, 2014. (Cited on page 105.)

[Kjer 2015a] Hans Martin Kjer, Jens Fagertun, Sergio Vera, Debora Gil, Miguel Angel González Ballester and Rasmus Reinhold Paulsen. *Free-form image registration of human cochlear μCT data using skeleton similarity as anatomical prior.* Pattern Recognition Letters, page 0000, jul 2015. (Cited on pages 27, 48, 83 and 85.)

[Kjer 2015b] Hans Martin Kjer and Rasmus Reinhold Paulsen. *Modelling of the Human Inner Ear Anatomy and Variability for Cochlear Implant Applications.* PhD thesis, Technical University of Denmark (DTU), 2015. (Cited on pages 14, 19, 54, 85, 106, 117 and 123.)

[Kjer 2015c] Hans Martin Kjer, Sergio Vera, Jens Fagertun, Debora Gil, Miguel Angel González Ballester and Rasmus Reinhold Paulsen. *Image registration of cochlear muCT data using heat distribution similarity.* In Lecture Notes in

Computer Science (Scandinavian Conference on Image Analysis - SCIA), volume 9127, pages 234–245. Springer, 2015. (Cited on page 85.)

[Kjer 2015d] Hans Martin Kjer, Sergio Vera, Jens Fagertun, Frederic Pérez, Javier Herrero Jover, Miguel Angel González Ballester, Rasmus Reinhold Paulsen and Miguel Angel Gonzalez Ballester. *Predicting detailed inner ear anatomy from pre-operational CT for cochlear implant surgery.* International Journal of Computer Assisted Radiology and Surgery (Computer Assisted Radiology and Surgery, Barcelona, Spain), vol. 10, pages S98—-99, 2015. (Cited on pages 16, 59, 80, 81 and 106.)

[Lane 2004] John I Lane, Robert J Witte, Colin L W Driscoll, John J Camp and Richard A Robb. *Imaging microscopy of the middle and inner ear: Part I: CT microscopy.* Clinical Anatomy, vol. 17, no. 8, pages 607–612, 2004. (Cited on page 14.)

[Lane 2005] John I Lane, Robert J Witte, Odell W Henson, Colin L W Driscoll, John J Camp and Richard A Robb. *Imaging microscopy of the middle and inner ear: Part II: MR microscopy.* Clinical Anatomy, vol. 18, no. 6, pages 409–415, 2005. (Cited on page 15.)

[Lane 2015] John I Lane. MultiPlanar Reformation in CT of the Temporal Bone, pages 367–380. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. (Cited on page 85.)

[Leake 1999] P A Leake, G T Hradek and R L Snyder. *Chronic electrical stimulation by a cochlear implant promotes survival of spiral ganglion neurons after neonatal deafness.* The Journal of Comparative Neurology, vol. 412, no. 4, pages 543–62, oct 1999. (Cited on page 18.)

[Leidy 1883] Joseph Leidy. *A Study of the Human Temporal Bone - II.* Science, vol. 1, no. 17, pages 475–477, jun 1883. (Cited on page 57.)

[Lesage 2009] David Lesage, Elsa D Angelini, Isabelle Bloch and Gareth Funka-Lea. *A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes.* Medical Image Analysis, vol. 13, no. 6, pages 819–45, dec 2009. (Cited on page 25.)

[Liu 1995] Chuanhai Liu and Donald B Rubin. *ML estimation of the t distribution using EM and its extensions, ECM and ECME.* Statistica Sinica, vol. 5, pages 19–39, 1995. (Cited on page 32.)

[Maintz 1998] J.B.Antoine Maintz and Max A. Viergever. *A survey of medical image registration.* Medical Image Analysis, vol. 2, no. 1, pages 1–36, mar 1998. (Cited on page 85.)

[Makary 2010] Chadi Makary, Jennifer Shin, Paul Caruso, Hugh D Curtin and Saumil N Merchant. *A histological study of scala communis with radiological* *implications.* Audiology & Neurotology, vol. 15, no. 6, pages 383–93, jan 2010. (Cited on page 18.)

[Manoussaki 2000] Daphne Manoussaki and Richard S Chadwick. *Effects of Geometry on Fluid Loading in a Coiled Cochlea.* SIAM Journal on Applied Mathematics, vol. 61, no. 2, pages 369–386, jan 2000. (Cited on page 59.)

[Maurer 2003] C R Maurer and V Raghavan. *A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 2, pages 265–270, 2003. (Cited on pages 101 and 102.)

[Meshik 2010] Xenia Meshik, Timothy A Holden, Richard A Chole and Timothy E Hullar. *Optimal cochlear implant insertion vectors.* Otology & neurotology, vol. 31, no. 1, pages 58–63, 2010. (Cited on page 24.)

[Miroir 2012] Mathieu Miroir, Yann Nguyen, Guillaume Kazmitcheff, Evelyne Ferrary, Olivier Sterkers and Alexis Bozorg Grayeli. *Friction force measurement during cochlear implant insertion: application to a force-controlled insertion tool design.* Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology, vol. 33, no. 6, pages 1092–100, aug 2012. (Cited on page 6.)

[Moseley 1838] H Moseley. *On the Geometrical Forms of Turbinated and Discoid Shells.* Philosophical Transactions of the Royal Society of London, vol. 128, pages 351–370, jan 1838. (Cited on page 57.)

[Nadol 1997] Joseph B Nadol. *Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation.* Otolaryngology - Head and Neck Surgery, vol. 117, no. 3, pages 220–228, sep 1997. (Cited on pages 18 and 54.)

[Neri 2005] Emanuele Neri, Stefano Berrettini, Luca Salvatori, Francesca Forli, Stefano Sellari Franceschini and Carlo Bartolozzi. *3-D CT and MRI coregistration in the assessment of cochlear implantation.* Medical science monitor : international medical journal of experimental and clinical research, vol. 11, no. 10, pages MT63–7, oct 2005. (Cited on page 85.)

[Nguyen 2011] Yann Nguyen. *Accés mini invasif à la cochlée, application à l'implantation cochléaire.* PhD thesis, Université Pierre et Marie Curie, 2011. (Cited on page 11.)

[Nguyen 2012] Yann Nguyen, Mathieu Miroir, Guillaume Kazmitcheff, Jasmine Sutter, Morad Bensidhoum, Evelyne Ferrary, Olivier Sterkers and Alexis Bozorg Grayeli. *Cochlear implant insertion forces in microdissected human cochlea to evaluate a prototype array.* Audiology & Neurotology, vol. 17, no. 5, pages 290–8, jan 2012. (Cited on page 51.)

[Noble 2010] Jack H Noble, Robert B Rutherford, Robert Frederick Labadie, Omid Majdani and Benoit M Dawant. *Modeling and segmentation of intra-cochlear anatomy in conventional CT*. In Benoit M. Dawant and David R. Haynor, editors, SPIE Medical Imaging, pages 762302–762302–9. International Society for Optics and Photonics, mar 2010. (Cited on pages 16, 80 and 81.)

[Noble 2011] Jack H Noble, Robert Frederick Labadie, Omid Majdani and Benoit M Dawant. *Automatic segmentation of intracochlear anatomy in conventional CT*. IEEE Transactions on Biomedical Engineering, vol. 58, no. 9, pages 2625–32, sep 2011. (Cited on pages 16, 20, 24, 54, 59, 72, 80, 81, 106 and 107.)

[Noble 2012] Jack H Noble, René H Gifford, Robert Frederick Labadie and Benoit M Dawant. *Statistical shape model segmentation and frequency mapping of cochlear implant stimulation targets in CT*. Medical Image Computing and Computer-Assisted Intervention, vol. 15, no. Pt 2, pages 421–8, jan 2012. (Cited on pages 16, 59, 80, 81 and 106.)

[Noble 2013] Jack H Noble, Robert Frederick Labadie, René H Gifford and Benoit M Dawant. *Image-Guidance enables new methods for customizing cochlear implant stimulation strategies*. IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 21, no. 5, pages 820–829, 2013. (Cited on pages 16, 59, 80, 81 and 106.)

[Ourselin 2000] Sébastien Ourselin, A Roche, S Prima and Nicholas Ayache. *Block Matching : A General Framework to Improve Robustness of Rigid Registration of Medical Images*. Medical Image Computing and Computer-Assisted Intervention, pages 557–566, 2000. (Cited on pages 55 and 87.)

[Patenaude 2011] Brian Patenaude, Stephen M Smith, David N Kennedy and Mark Jenkinson. *A Bayesian model of shape and appearance for subcortical brain segmentation*. NeuroImage, vol. 56, no. 3, pages 907–922, 2011. (Cited on pages 82 and 83.)

[Paulsen 2004] Rasmus Reinhold Paulsen. *Statistical Shape Analysis of the Human Ear Canal with Application to In-the-Ear Hearing Aid Design*. PhD thesis, Technical University of Denmark, 2004. (Cited on page 59.)

[Peltonen 2009] L I Peltonen, Antti A Aarnisalo, Y Käser, M K Kortesniemi, S Robinson, A Suomalainen and J Jero. *Cone-beam computed tomography: A new method for imaging of the temporal bone*. Acta Radiologica, vol. 50, no. 5, pages 543–548, jun 2009. (Cited on page 13.)

[Perona 1990] P Perona and J Malik. *Scale-space and edge detection using anisotropic diffusion*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pages 629–639, jul 1990. (Cited on page 24.)

[Pohl 2005] Kilian M Pohl, John Fisher, James J Levitt, Martha E Shenton, Ron Kikinis, W Eric L Grimson and William M Wells. *A unifying approach to registration, segmentation, and intensity correction.* Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 3749 LNCS, pages 310–318, 2005. (Cited on pages 82 and 83.)

[Pohl 2006a] Kilian M Pohl, John Fisher, W Eric L Grimson, Ron Kikinis and William M Wells. *A Bayesian model for joint segmentation and registration.* NeuroImage, vol. 31, no. 1, pages 228–239, 2006. (Cited on pages 81, 82 and 83.)

[Pohl 2006b] Kilian M Pohl, John Fisher, Martha E Shenton, Robert W Mccarley, W Eric L Grimson, Ron Kikinis, William M Wells and W L Eric. *Logarithm Odds Maps for Shape Representation.* Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, vol. 9, no. Pt 2, pages 955–963, 2006. (Cited on page 96.)

[Postnov 2006] A Postnov, Andrzej Zarowski, N De Clerck, F Vanpoucke, F Erwin Offeciers, D Van Dyck and S Peeters. *High resolution micro-CT scanning as an innovative tool for evaluation of the surgical positioning of cochlear implant electrodes.* Acta Oto-Laryngologica, vol. 126, no. 5, pages 467–74, may 2006. (Cited on pages 19 and 24.)

[Poznyakovskiy 2008] Anton A Poznyakovskiy, Thomas Zahnert, Yannis Kalaidzidis, Rolf Schmidt, Björn Fischer, Johannes Baumgart and Yury M Yarin. *The creation of geometric three-dimensional models of the inner ear based on micro computer tomography data.* Hearing Research, vol. 243, no. 1-2, pages 95–104, sep 2008. (Cited on page 27.)

[Poznyakovskiy 2011] Anton A Poznyakovskiy, Thomas Zahnert, Yannis Kalaidzidis, Nikoloz Lazurashvili, Rolf Schmidt, Hans-Jürgen Hardtke, Björn Fischer and Yury M Yarin. *A segmentation method to obtain a complete geometry model of the hearing organ.* Hearing Research, vol. 282, no. 1-2, pages 25–34, dec 2011. (Cited on pages 19, 24, 25 and 33.)

[Prince 2012] Simon J D Prince. Computer vision: models, learning, and inference. Computer Vision: Models, Learning, and Inference. Cambridge University Press, 2012. (Cited on page 93.)

[Puonti 2016] Oula Puonti, Juan Eugenio Iglesias and Koen Van Leemput. *Fast and sequence-adaptive whole-brain segmentation using parametric Bayesian modeling.* NeuroImage, vol. 143, pages 235–249, 2016. (Cited on pages 82 and 83.)

[Rask-Andersen 2011] Helge Rask-Andersen, Elsa Erixon, Anders Kinnefors, Hubert Löwenheim, Anneliese Schrott-Fischer and Wei Liu. *Anatomy of the human cochlea – implications for cochlear implantation*. Cochlear Implants International, vol. 12, no. s1, pages S13–S8, may 2011. (Cited on page 18.)

[Rask-Andersen 2012] Helge Rask-Andersen, Wei Liu, Elsa Erixon, Anders Kinnefors, Kristian Pfaller, Annelies Schrott-Fischer and Rudolf Glueckert. *Human cochlea: anatomical characteristics and their relevance for cochlear implantation*. The Anatomical Record, vol. 295, no. 11, pages 1791–811, nov 2012. (Cited on page 66.)

[Rau 2013] Thomas S Rau, Waldemar Würfel, Thomas Lenarz and Omid Majdani. *Three-dimensional histological specimen preparation for accurate imaging and spatial reconstruction of the middle and inner ear*. International Journal of Computer Assisted Radiology and Surgery, vol. 8, no. 4, pages 481–509, apr 2013. (Cited on page 19.)

[Reda 2011] Fitsum A Reda, Jack H Noble, Alejandro Rivas, Theodore R McRackan, Robert Frederick Labadie and Benoit M Dawant. *Automatic segmentation of the facial nerve and chorda tympani in pediatric CT scans*. Medical Physics, vol. 38, no. 10, pages 5590–600, oct 2011. (Cited on pages 16, 80 and 81.)

[Reda 2012] Fitsum A Reda, Jack H Noble, Robert Frederick Labadie and Benoit M Dawant. *Automatic pre- to intra-operative CT registration for image-guided cochlear implant surgery*. IEEE Transactions on Biomedical Engineering, vol. 59, no. 11, pages 3070–7, nov 2012. (Cited on pages 16, 80, 81 and 85.)

[Reda 2013] Fitsum A Reda, Benoit M Dawant, Theodore R McRackan, Robert Frederick Labadie and Jack H Noble. *Automatic segmentation of intra-cochlear anatomy in post-implantation CT*. Proc. SPIE, vol. 8671, pages 86710I–86710I–9, mar 2013. (Cited on pages 16, 20, 80, 81 and 107.)

[Reda 2014a] Fitsum A Reda, Theodore R McRackan, Robert Frederick Labadie, Benoit M Dawant and Jack H Noble. *Automatic segmentation of intra-cochlear anatomy in post-implantation CT of unilateral cochlear implant recipients*. Medical Image Analysis, vol. 18, no. 3, pages 605–615, feb 2014. (Cited on pages 16, 20, 80, 81, 107, 114 and 118.)

[Reda 2014b] Fitsum A Reda, Jack H Noble, Robert Frederick Labadie and Benoit M Dawant. *An artifact-robust, shape library-based algorithm for automatic segmentation of inner ear anatomy in post-cochlear-implantation CT*. In Sebastien Ourselin and Martin A. Styner, editors, SPIE Medical Imaging, volume 9034, page 90342V. International Society for Optics and Photonics, mar 2014. (Cited on pages 13, 16, 20, 80, 81, 107 and 118.)

[Riklin-Raviv 2009] Tammy Riklin-Raviv, Koen Van Leemput, William M Wells and Polina Golland. *Joint Segmentation of Image Ensembles via Latent Atlases Tammy.* Medical Image Computing and Computer-Assisted Intervention, 2009. (Cited on pages 82 and 83.)

[Romera Romero 2016] Jordi Romera Romero, Hans Martin Kjer, Gemma Piella, Mario Ceresa and Miguel Angel González Ballester. *Multi-region statistical shape model for cochlear implantation.* In SPIE Medical Imaging, San Diego, USA, volume 9784, pages 97840T1—-8. SPIE, 2016. (Cited on page 59.)

[Ruiz Pujadas 2016a] Esmeralda Ruiz Pujadas, Hans Martin Kjer, Gemma Piella, Mario Ceresa and Miguel Angel González Ballester. *Random walks with shape prior for cochlea segmentation in ex vivo CT.* International Journal of Computer Assisted Radiology and Surgery, vol. 11, no. 9, pages 1647–1659, 2016. (Cited on page 105.)

[Ruiz Pujadas 2016b] Esmeralda Ruiz Pujadas, Hans Martin Kjer, Sergio Vera, Mario Ceresa and Miguel Angel González Ballester. *Cochlea segmentation using iterated random walks with shape prior*, 2016. (Cited on page 105.)

[Schuman 2010] Theodore A Schuman, Jack H Noble, Charles G Wright, George B Wanna, Benoit M Dawant and Robert Frederick Labadie. *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, vol. 120, no. 11, pages 2277–83, nov 2010. (Cited on page 20.)

[Sennaroglu 2002] Levent Sennaroglu and Isil Saatci. *A New Classification for Cochleovestibular Malformations.* The Laryngoscope, vol. 112, no. 12, pages 2230–2241, 2002. (Cited on page 123.)

[Shanno 1970] David F Shanno. *Conditioning of quasi-Newton methods for function minimization.* Mathematics of computation, vol. 24, no. 111, pages 647–656, 1970. (Cited on page 101.)

[Shi 2011] Lin Shi, Defeng Wang, Winnie C W Chu, Geoffrey R Burwell, Tien-Tsin Wong, Pheng Ann Heng and Jack C Y Cheng. *Automatic MRI segmentation and morphoanatomy analysis of the vestibular system in adolescent idiopathic scoliosis.* NeuroImage, vol. 54 Suppl 1, pages S180–8, jan 2011. (Cited on page 15.)

[Shin 2013] Kang-Jae Shin, Ju-Young Lee, Jeong-Nam Kim, Ja-Young Yoo, Chuog Shin, Wu-Chul Song and Ki-Seok Koh. *Quantitative analysis of the cochlea using three-dimensional reconstruction based on microcomputed tomographic images.* The Anatomical Record, vol. 296, no. 7, pages 1083–8, jul 2013. (Cited on pages 19, 48, 49 and 52.)

[Skinner 2002] Margaret W Skinner, Darlene R Ketten, Laura K Holden, Gary W Harding, Peter G Smith, George A Gates, J Gail Neely, G Robert Kletzker, Barry Brunsden and Barbara Blocker. *CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients.* Journal of the Association for Research in Otolaryngology : JARO, vol. 3, no. 3, pages 332–50, sep 2002. (Cited on page 19.)

[Soille 1999] Pierre Soille. Morphological image analysis: principles and applications. Springer Science & Business Media, 1999. (Cited on pages 101 and 102.)

[Sotiras 2013] Aristeidis Sotiras, Christos Davatzikos and Nikos Paragios. *Deformable Medical Image Registration: A Survey.* IEEE transactions on medical imaging, vol. 32, no. 7, pages 1153–1190, jul 2013. (Cited on page 85.)

[Stakhovskaya 2007] Olga Stakhovskaya, Divya Sridhar, Ben H Bonham and Patricia A Leake. *Frequency map for the human cochlear spiral ganglion: Implications for cochlear implants.* JARO - Journal of the Association for Research in Otolaryngology, vol. 8, no. 2, pages 220–233, 2007. (Cited on page 19.)

[Stone 1996] JR Stone. *The evolution of ideas: a phylogeny of shell models.* American Naturalist, 1996. (Cited on page 57.)

[Thompson 1917] D'Arcy Wentworth Thompson. *On growth and form*, 1917. (Cited on page 57.)

[Thorne 1999] M Thorne, A N Salt, J E DeMott, M M Henson, Odell W Henson and S L Gewalt. *Cochlear fluid space dimensions for six species derived from reconstructions of three-dimensional magnetic resonance images.* The Laryngoscope, vol. 109, no. 10, pages 1661–8, oct 1999. (Cited on pages 19 and 49.)

[Tian 2006] Qing Tian, Fred H Linthicum and Jose N Fayad. *Human cochleae with three turns: an unreported malformation.* The Laryngoscope, vol. 116, no. 5, pages 800–3, may 2006. (Cited on page 49.)

[Torres 2015] Renato Torres, Guillaume Kazmitcheff, Daniele Bernardeschi, Daniele De Seta, Jean Loup Bensimon, Evelyne Ferrary, Olivier Sterkers and Yann Nguyen. *Variability of the mental representation of the cochlear anatomy during cochlear implantation.* European Archives of Oto-Rhino-Laryngology, sep 2015. (Cited on page 124.)

[Toussaint 2007] Nicolas Toussaint, Jean-Christophe Souplet and Pierre Fillard. *MedINRIA: Medical Image Navigation and Research Tool by INRIA.* In Proc. of MICCAI'07 Workshop on Interaction in medical image analysis and visualization, Brisbane, Australia, Australia, 2007. (Cited on page 87.)

[van der Marel 2014] Kim S van der Marel, Jeroen Johannes Briaire, Ron Wolterbeek, Jorien Snel-Bongers, Berit M Verbist and Johan H M Frijns. *Diversity in cochlear morphology and its influence on cochlear implant electrode position.* Ear and Hearing, vol. 35, no. 1, pages e9–20, jan 2014. (Cited on pages 19, 27 and 54.)

[Vandersteen 2014] Clair Vandersteen, Thomas Demarcy, Hervé Delingette, Charles Raffaelli, Jonathan Laudanski, Thierry Pourcher, Jacques Darcourt, Philippe Franken, Dan Gnansia, Nicholas Ayache and Nicolas Guevara. *Teaching tool for advanced visualization of temporal bone structures by fusion of* μ*CT and CT scan images.* In 8th International Symposium on Objective Measures in Auditory Implants, 2014. (Cited on page 126.)

[Vandersteen 2015] Clair Vandersteen, Thomas Demarcy, Coralie Roger, Eric Fontas, Charles Raffaelli, Nicholas Ayache, Hervé Delingette and Nicolas Guevara. *Impact of the surgical experience on cochleostomy location : a comparative temporal bone study between endaural and posterior tympanotomy approaches for cochlear implantation.* European Archives of Oto-Rhino-Laryngology, vol. 273, no. 9, pages 2355–2361, 2015. (Cited on pages 121, 124 and 125.)

[Verbist 2005] Berit M Verbist, Johan H M Frijns, Jakob Geleijns and Mark A van Buchem. *Multisection CT as a Valuable Tool in the Postoperative Assessment of Cochlear Implant Patients.* American Journal of Neuroradiology, vol. 26, no. 2, pages 424 LP – 429, feb 2005. (Cited on page 14.)

[Verbist 2009] Berit M Verbist, Luca Ferrarini, Jeroen Johannes Briaire, Andrzej Zarowski, Faiza Admiraal-behloul, Hans Olofsen, Johan H C Reiber and Johan H M Frijns. *Anatomic considerations of cochlear morphology and its implications for insertion trauma in cochlear implant surgery.* Otology & Neurotology, vol. 30, no. 4, pages 471–7, jun 2009. (Cited on pages 19, 24, 25, 27 and 49.)

[Verbist 2010] Berit M Verbist, Margaret W Skinner, Lawrence T Cohen, Patricia A Leake, Chris James, Colette Boëx, Timothy A Holden, Charles C Finley, Peter S Roland, J Thomas Roland, Matt Haller, Jim F Patrick, Claude N Jolly, Mike A Faltys, Jeroen Johannes Briaire and Johan H M Frijns. *Consensus panel on a cochlear coordinate system applicable in histologic, physiologic, and radiologic studies of the human cochlea.* Otology & Neurotology, vol. 31, no. 5, pages 722–30, jul 2010. (Cited on pages 25, 27, 45, 55, 69 and 85.)

[Viergever 1978] Max A. Viergever. *Basilar membrane motion in a spiral-shaped cochlea.* The Journal of the Acoustical Society of America, vol. 64, no. 4, page 1048, oct 1978. (Cited on page 33.)

[Voie 1990] Arne H Voie and Francis A Spelman. *Analysis Of The Guinea Pig Cochlea Using A General Cylindrical Coordinate System.* In IEEE Engineering in Medicine & Biology Society, pages 206–207. IEEE, 1990. (Cited on pages 57 and 66.)

[Wardrop 2005] Peter Wardrop, David Whinney, Stephen J Rebscher, J Thomas Roland, William Luxford and Patricia A Leake. *A temporal bone study of insertion trauma and intracochlear position of cochlear implant electrodes. I: Comparison of Nucleus banded and Nucleus Contour*^{TM} *electrodes.* Hearing Research, vol. 203, no. 1-2, pages 54–67, may 2005. (Cited on page 51.)

[Weber 2017] Stefan Weber, Kate Gavaghan, Wilhelm Wimmer, Tom Williamson, Nicolas Gerber, Juan Anso, Brett Bell, Arne Feldmann, Christoph Rathgeb, Marco Matulic, Manuel Stebinger, Daniel Schneider, Georgios Mantokoudis, Olivier Scheidegger, Franca Wagner, Martin Kompis and Marco Caversaccio. *Instrument flight to the inner ear.* Science Robotics, vol. 2, no. 4, mar 2017. (Cited on page 12.)

[Whiting 2001] B R Whiting, K T Bae and Margaret W Skinner. *Cochlear implants: three-dimensional localization by means of coregistration of CT and conventional radiographs.* Radiology, vol. 221, no. 2, pages 543–9, nov 2001. (Cited on page 85.)

[Wilson 2008] Blake S Wilson and Michael F Dorman. *Cochlear implants: a remarkable past and a brilliant future.* Hearing Research, vol. 242, no. 1-2, pages 3–21, aug 2008. (Cited on page 10.)

[Wright 2005] Charles G Wright and Peter S Roland. *Temporal bone microdissection for anatomic study of cochlear implant electrodes.* Cochlear Implants International, vol. 6, no. 4, pages 159–168, 2005. (Cited on page 18.)

[Wu 2013] Guorong Wu, Qian Wang, Daoqiang Zhang, Feiping Nie, Heng Huang and Dinggang Shen. *A generative probability model of joint label fusion for multi-atlas based brain segmentation.* Medical Image Analysis, vol. 18, no. 6, pages 881–890, nov 2013. (Cited on pages 82 and 83.)

[Wyatt 2003] Paul P Wyatt and J Alison Noble. *MAP MRF joint segmentation and registration of medical images.* Medical Image Analysis, vol. 7, no. 4, pages 539–552, 2003. (Cited on pages 82 and 83.)

[Wysocki 1999] Jarosław Wysocki. *Dimensions of the human vestibular and tympanic scalae.* Hearing Research, vol. 135, no. 1-2, pages 39–46, 1999. (Cited on pages 18, 19, 24, 27, 45, 49, 55 and 77.)

[Xianfen 2005] Diao Xianfen, Chen Siping, Liang Changhong and Wang Yuanmei. *3D semi-automatic segmentation of the cochlea and inner ear.* Engineering in Medicine and Biology Society, vol. 6, pages 6285–6288, jan 2005. (Cited on page 20.)

[Xu 2000] Jin Xu, Shi Ang Xu, Lawrence T Cohen and Graeme M Clark. *Cochlear view: postoperative radiography for cochlear implantation.* The American Journal of Otology, vol. 21, no. 1, pages 49–56, jan 2000. (Cited on pages 28 and 29.)

[Xu 2012] Helen X Xu, Grace H Kim, Eugene P Snissarenko, Sebahattin Cureoglu and Michael M Paparella. *Multi-channel cochlear implant histopathology: are fewer spiral ganglion cells really related to better clinical performance?* Acta Oto-Laryngologica, vol. 132, no. 5, pages 482–90, may 2012. (Cited on page 18.)

[Yoo 2000a] Sun K Yoo, Ge Wang, Jay T Rubinstein, Margaret W Skinner and Michael W Vannier. *Three-dimensional modeling and visualization of the cochlea on the Internet.* IEEE Transactions on Information Technology in Biomedicine, vol. 4, no. 2, pages 144–51, jun 2000. (Cited on pages 24, 28 and 59.)

[Yoo 2000b] Sun K Yoo, Ge Wang, Jay T Rubinstein and Michael W Vannier. *Three-dimensional geometric modeling of the cochlea using helico-spiral approximation.* IEEE Transactions on Biomedical Engineering, vol. 47, no. 10, pages 1392–402, oct 2000. (Cited on pages ix, 27, 28, 41, 42, 43, 50 and 59.)

[Yoo 2001] Sun K Yoo, Ge Wang, Jay T Rubinstein and Michael W Vannier. *Semi-automatic segmentation of the cochlea using real-time volume rendering and regional adaptive snake modeling.* Journal of Digital Imaging, vol. 14, no. 4, pages 173–181, dec 2001. (Cited on page 26.)

[Yushkevich 2006] Paul A Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C Gee and Guido Gerig. *User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability.* NeuroImage, vol. 31, no. 3, pages 1116–28, jul 2006. (Cited on page 85.)

[Zerroug 1996] Mourad Zerroug and Ramakant Nevatia. *Three-dimensional descriptions based on the analysis of the invariant and quasi-invariant properties of some curved-axis generalized cylinders.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 3, pages 237–253, 1996. (Cited on page 67.)

[Zhang 2006] Jian Zhang, Kai Xu, Nabil Simaan and Spiros Manolidis. *A pilot study of robot-assisted cochlear implant surgery using steerable electrode arrays.* Medical Image Computing and Computer-Assisted Intervention Intervention, vol. 9, no. Pt 1, pages 33–40, 2006. (Cited on page 59.)

[Zhang 2014] Minqi Zhang, Fang Li, Xingce Wang, Zhongke Wu, Shi-Qing Xin, Lok-Ming Lui, Lin Shi, Defeng Wang and Ying He. *Automatic registration* *of vestibular systems with exact landmark correspondence.* Graphical Models, apr 2014. (Cited on pages 117 and 123.)

[Zou 2015] Jing Zou, Jaakko Lähelmä, Juha Koivisto, Anandhan Dhanasingh, Claude N Jolly, Antti A Aarnisalo, Jan Wolff and Ilmari Pyykkö. *Imaging cochlear implantation with round window insertion in human temporal bones and cochlear morphological variation using high-resolution cone beam CT.* Acta oto-laryngologica, vol. 135, no. 5, 2015. (Cited on page 13.)

[Zrunek 1980] M Zrunek, M Lischka, I Hochmair-Desoyer and K Burian. *Dimensions of the scala tympani in relation to the diameters of multichannel electrodes.* Archives Of Oto-Rhino-Laryngology, vol. 229, pages 159–165, 1980. (Cited on pages 18, 19 and 33.)

[Zrunek 1981] M Zrunek and M Lischka. *Dimensions of the scala vestibuli and sectional areas of both scales.* Archives Of Oto-Rhino-Laryngology, vol. 233, no. 4151, pages 99–104, 1981. (Cited on pages 18, 19 and 33.)

Segmentation and Study of Anatomical Variability of the Cochlea from Medical Images

Abstract: Cochlear implants (CI) are used to treat hearing loss by surgically inserting an electrode array into the organ of hearing, the cochlea.

Pre- and post-operative CT images are used routinely for surgery planning and evaluation of cochlear implantation. However, due to the small size and the complex topology of the cochlea, the anatomical information that can be extracted from the images is limited.

The first focus of this work aims at defining automatic image processing methods adapted to the spiral shape of the cochlea to study the cochlear shape variability from high-resolution μCT images.

The second focus aims at developing and evaluating a new parametric cochlear shape model. The model is applied to extract patient-specific clinically relevant metrics such as the maximal insertion depth of CI electrode arrays. Thanks to the uncertainty quantification, provided by the model, we can assess the reliability of CT-based segmentation as compared to the ground truth segmentation provided by μCT scans.

Finally, the last focus concerns a joint model of the cochlear shape (and its substructures) model and its appearance within a generative probabilistic Bayesian framework. The proposed segmentation method was applied to a large database of 987 CT images and allowed the statistical characterization of the cochlear anatomical variability along with the quantification of the bilateral symmetry.

This work paves the way to novel clinical applications such as improved diagnosis by identifying pathological cochlear shapes; preoperative optimal electrode design and insertion axis planning; postoperative electrode position estimation and implantation evaluation; and cochlear implantation simulation.

Keywords: cochlea, segmentation, shape model, shape variability