Segmentation Algorithms for Ear Image Data towards Biomechanical Studies

Ana Ferreira[a], Fernanda Gentil[b], João Manuel R. S. Tavares[a]

[a] *Instituto de Engenharia Mecânica e Gestão Industrial, Faculdade de Engenharia, Universidade do Porto, PORTO, PORTUGAL*

Emails: aiferreira@inegi, tavares@fe.up.pt

[b] *Escola Superior de Tecnologia da Saúde do Porto, Clínica ORL – Dr. Eurico Almeida, IDMEC-Polo FEUP, PORTO, PORTUGAL*

Email: fernanda.fgnanda@gmail.com

Corresponding author:

Professor João Manuel R. S. Tavares

Departamento de Engenharia Mecânica

Faculdade de Engenharia da Universidade do Porto

Rua Dr. Roberto Frias

4200-465 Porto

Portugal

Phone: +351 225 081 487, Fax: +351 225 081 445

Email: tavares@fe.up.pt, url: www.fe.up.pt/~tavares

Segmentation Algorithms for Ear Image Data towards Biomechanical Studies

ABSTRACT

In the recent years, the segmentation, i.e. the identification, of ear structures in Video-otoscopy (VO), Computerized Tomography (CT) and Magnetic Resonance (MR) image data has gained significantly importance in the medical imaging area, particularly those in CT and MR imaging. Segmentation is the fundamental step of any automated technique for supporting the medical diagnosis and, in particular, in biomechanics studies, to build realistic geometric models of ear structures. In this paper, a review of the algorithms used in ear segmentation is presented. The review includes an introduction to the usually biomechanical modeling approaches and also to the common imaging modalities. Afterwards, several segmentation algorithms for ear image data are described, their specificities and difficulties as well their advantages and disadvantages are identified and analyzed using experimental examples. Finally, the conclusions are presented as well as a discussion about possible trends for future research concerning the ear segmentation.

Keywords: Biomedical Engineering, Medical Imaging, Analogue Circuits, Multibody, Finite Element Modeling, Thresholding, Clustering, Deformable Models, Atlas, Review.

1. Introduction

The human ear is the most complex organ of the human sensory system (Moller 2006, Seeley et al. 2004). Its hearing receptors convert sound waves into nerve impulses, and its equilibrium receptors are associated with the movements of the head. The vestibulocochlear nerve is responsible for transmitting impulses from these receptors to the brain.

Anatomically, the auditory system consists of the outer ear, middle ear and inner ear, the auditory pathways and the auditory cortex. The outer ear captures the sound waves, which travel through the external auditory canal until they reach the eardrum. This causes the membrane and the attached chain of auditory ossicles to vibrate. The vibrations are passed to the ossicles, which transmit them to the cochlea. The cochlea contains tubes filled with fluid, inside one of these tubes tiny hair cells pick up the vibrations and convert them into nerve impulses. These impulses are delivered to the brain via the hearing nerve, which interprets the impulses as sound. Figure 1 depicts the anatomy of the human ear and illustrates how the sound waves travel through its main components.

Heredity, toxins, drugs and infections are some factors that can have consequences either in the function or in the shape of the auditory system (Costa 2008). Otosclerosis, characterized by the abnormal ossification of the stapes – smallest bone in the human body – is one of the major causes of deafness in adults. People with this condition have large problems in communication, since adolescence, being worse in the adulthood (Niparko 1994). Tinnitus, a common phenomenon defined as an unwanted auditory perception of internal origin, usually localized and rarely heard by others (Meyerhoff and Cooper 1991), affects 17% of the general population and about 33% of the elderly (Jastreboff and Hazell 1993) and can cause very discomfort interfering with people life quality, leading to anxiety and depression that can result in suicide (Lewis et al. 1994). So, it is common believed there is a pressing need for further studies in this area.

Computational models can be used to simulate the anatomic structure of the ear in order to, not only help radiological diagnosis, surgical planning and teaching, but also to better understand the relationship between its structures and function, simulate pathologies of the ear, comparing them with normal ear and improve the design of prosthesis. In general, there are two groups of models: (1) lumped parameter models (e.g., the analogue circuit models, mechanical models or multibody models) and (2) distributed parameter models (including Finite Element Models) (Volandri et al. 2012). There are numerous approaches in the literature describing how to obtain different models of the (animal and human) ear.

The segmentation of the image data to be studied is a prerequisite step for modeling and analysis the ear represented. A number of approaches have been presented for that purpose, most of them employing manual tracing of the contours on each slice (Jun et al. 2005, Sim and Puria 2008). However, since this procedure is very time consuming and subjective, attention has been focused on the development of both semi-automatic and automatic algorithms, but, particularly due to the complex shape and reduced dimension of the structures involved, several difficulties still persist.

In the following sections, biomechanical modeling approaches are introduced, and afterwards, the image segmentation algorithms are classified into four groups: Thresholding, Clustering, Deformable models and Atlas based. The definition of each group, an overview of how each algorithm is implemented and a discussion of its advantages and disadvantages are exposed using illustrative experimental cases. It should be noted that we do not intend to propose a new algorithm for human ear segmentation, but to present, evaluate and discuss solutions that can be suitable for the building of geometric models of ear structures from medical images, mainly for biomechanical studies. Furthermore, the main guidelines that an effective algorithm should adopt for a successful segmentation of the human ear will be pointed out.

The outline of the paper is as follows: The next section introduces the biomechanical modeling approaches applied to the ear. Section 3 provides the necessary background about medical image segmentation, including an overview about the current ear imaging modalities. Then, a review of segmentation algorithms that have been used in ear image data is presented, and examples of their experimental results are illustrated. In Section 4, the advantages and disadvantages of each segmentation algorithm group are pointed out as well as their main guidelines. Finally, in Section 5, the conclusions are presented and the possible trends in this are towards the effective segmentation of the human ear structures for realistic biomechanical simulations are identified.

2. Biomechanical Modeling on Ear Structures

For many ears, several modeling approaches have been successfully developed to refine the understanding and simulation of the hearing process. In the literature, models for the ear can be classified into two broad groups (Volandri et al. 2012). The first group consists of lumped parameter models, including electro-acoustical analogue circuit models (Goode et al. 1994,

Hudde et al. 1997, Kringlebotn 1988, Parent and Allen 2010, Peake et al. 1992, Puria and Allen 1998, Rosowski and Merchant 1995, Shera and Zweig 1991, Zwislocki 1962), electronic and signal processing based models (Chitore et al. 1983), mechanical (Stieger et al. 2007, Yao et al. 2010) or multi-body (Eiber and Schiehlen 1995, Volandri et al. 2012, Wegel and Lane 1924, Wright 2005) models. The second group is composed of distributed parameter models, including analytical asymptotic models (Rabbitt and Holmes 1986), but mainly by Finite Element Models for the outer (Fay et al. 2005, Funnell et al. 1987, Funnell and Laszlo 1978, Funnell and Laszlo 1982, Gan et al. 2009, Gan et al. 2004, Gan et al. 2007, Gan et al. 2006, Gan and Wang 2007, Lee et al. 2010a, Lesser and Williams 1988, Prendergast et al. 1999a, Prendergast et al. 1999b, Williams and Lesser 1990, Zhang and Gan 2011), middle (Beer et al. 1999, Bornitz et al. 2010, Bornitz et al. 1999, Chou et al. 2011, Ferris and Prendergast 2000, Gan et al. 2004, Gan et al. 2007, Gan et al. 2006, Gan and Wang 2007, Gentil et al. 2005, Gentil et al. 2011, Gentil et al. 2012, Koike et al. 2002, Ladak and Funnell 1996, Lee et al. 2010a, Lee et al. 2006, Lesser et al. 1991, Prendergast et al. 1999a, Sun et al. 2002, Wada et al. 1992, Williams et al. 1995, Zhang and Gan 2011, Zhao et al. 2009) and inner (Gan et al. 2007, Gan and Wang 2007, Zhang and Gan 2011) ear.

2.1. Lumped parameter models

Lumped parameter models use the analogy between acoustics and electrical engineering. Physical components with acoustic properties are then represented as behaving similarly to the standard electronic components: mass components are modeled as inductors, stiffness-walled cavities containing air are modeled as capacitors and damping components are approximated as resistors. The firsts to present a model of the ear function using a transformer analogy with analogue circuit models were (Wegel and Lane 1924). Later, Zwislocki et al. (Zwislocki 1962) presented a key work in the field by modeling the entire ossicular chain of a cat. In their approach, numerical values derived from impedance measurements on normal and pathological ears were used. The results showed that changes in analog parameters corresponding to identified anatomical changes produce the same effect on its impedance characteristics as measured at the eardrum and that the input impedance of the analogue agrees with the experimental error with the acoustic impedance at the eardrum. On the other hand, the biomechanical models based on multi-body systems collect rigid and/or flexible bodies that are constrained by kinematic joints and contacts, acting upon by a set of internal and/or external forces (Figure 2). Essentially, the dynamic behavior of a multi-body

system is defined by solving equations of motion, usually derived from the Newton-Euler equations or Lagrange's equations (Wright 2005).

2.2. Distributed parameter models

Distributed parameter models includes mainly the Finite Element Method (FEM), which is a mathematical method of discretization (subdivision) from a continuum medium from a smaller sub-domain (elements), while maintaining the same properties as the original medium. The behavior of these elements can be described by differential equations and resolved by mathematical models using computer analysis. Actually, FEM is one of the most powerful tools to simulate mechanical problems, allowing an analysis with a high level of complexity, from geometric models (Belytschko and Moran 2000). In this method, a continuous system is divided into a finite number of parts, called elements. In each element, the solution is obtained from nodes, ensuring their boundary conditions, turning a problem with an infinite number of freedom degrees in the continuum, into another finite problem (Belytschko and Moran 2000). From the conceptual viewpoint, the process of building mathematical FEM models firstly needs to know the geometry of the ear structures. Then, the material properties containing information on the internal constitution of the different structures also must be described. The influences of the surrounding environment must be described by the boundary conditions as well as the changes of the physical quantities, which interact at the model boundaries in different form. In order to import geometrical data into the mathematical model, the volumes of the different ear structures should be firstly imaged. Then, to extract selected regions of interest from the image data, an effective segmentation algorithm is required.

The first FEM model of the ear was built in the cat and dates from 1978 (Funnell and Laszlo 1978). This model was further refined in collaboration with other authors (Funnell et al. 1987, Funnell and Laszlo 1982, Ladak and Funnell 1996). Other FEM models have been developed from the geometry of the human ear, considering the tympanic membrane, the ossicles and the cochlear impedance; then the inclusion of some ligaments and tendons (Beer et al. 1999, Koike et al. 2002, Wada et al. 1992, Williams and Lesser 1990). Since then, other FEM models have been developed to simulate the static and dynamic behavior of the model (Ferris and Prendergast 2000, Prendergast et al. 1999a). Many of these studies compare their results with experimental data. However, all these FEM models represent the behavior of the ear, taking the capsular ligaments as a continuous medium between the ossicles, not

presenting any analysis about the activation of the muscles of the ear. In the work of (Gentil et al. 2005, Gentil et al. 2011, Gentil et al. 2012), a formulation of contact was used in the simulation of the capsular ligaments, considering the ligaments with hyperelastic behavior (Figure 3). They also used a constitutive model to simulate the active and passive function of the middle ear muscles.

As far as our knowledge, the geometric models used in aforementioned biomechanical studies were built from a complete set of histological section images, i.e. image slices, that were manually segmented (Gan et al. 2004, Gan et al. 2006, Gentil et al. 2011, Gentil et al. 2012, Liu et al. 2009, Sun et al. 2002), using published data (Koike et al. 2002) or anatomic models available online as, for example, at http://audilab.bmed.mcgill.ca/~daren/3Dear/index.html, (Volandri et al. 2011). Hence, new solutions to build geometrical models for the imaged ear structures, in particular, fully automated, are demanded. Besides, all the biomechanical modeling approaches presented in the literature have a measurement error, caused either by the method of scanning or by the algorithms used to construct the geometry of the structures. Therefore, there is also a pressing need to minimizing this error, in order to make the biomechanical models more realistic and more customized to the patient under study.

3. Building of Geometric Models from Medical Images

There are two modes of building 3D geometric models from medical images. In one mode, the structures can be directly segmented from the 3D volumetric data from different medical imaging modalities, by using segmentation techniques such as region growing or deformable models (described in detail in Section 3.2.). Methods that used 3D segmentation for ear data images are scarce available on literature. The work of Xianfen et al. (Xianfen et al. 2005) is the only known example of how the cochlea and the semicircular canals can be well segmented with a Level set algorithm applied on 3D spiral CT images. In the other mode, the segmentation is done in each slice of the volumetric data, and the resulting 2D contours are used to create the 3D models using interpolation algorithms.

We hereby concentrate our review on the second mode of building geometric models. Hence, in the next sections, we introduce the common medical modalities used in ear biomechanical simulations. Then, algorithms that have been applied to segment the outer, middle and inner ear in images from these modalities are addressed.

3.1. Imaging modalities

Medical image segmentation consists of extracting some anatomical structures from various medical imaging modalities. Video-otoscopy (VO), Computerized Tomography (CT) and Magnetic Resonance (MR) are often used imaging techniques for the study of the ear. Figures 4 and 5 illustrate a slice example of an ear CT and MR images, respectively. As far as our knowledge, there are few works made in the segmentation of ear structures concerning VO (Xie et al. 2005, Comunello et al. 2009) and MR (Shi et al. 2010, Melhem et al. 1998, Tabrizi 2003, Folowosele et al. 2004) images, being the CT the most used image modality for this purpose.

In clinical routine, the VO is the commonly used image acquisition process for consultants examine pathological alterations, especially in the ear canal and eardrum. The acquired examination data is stored using a digital video file format. The main disadvantage of VO images is that they present irregular illumination, which leaves some image regions brighter or darker than the average color of a given structure. It also turns out to obtain a good structural targeting into a difficult task. These characteristics together with the low contrast of the boundaries of anatomical structures make structures problematic to be segmented automatically.

In CT, the images are reconstructed from a large number of X-rays to obtain structural information about the human body. X-rays are based on its property that all matters and tissues differ in their ability to absorb X-rays (Prince and Links 2006). It is primarily used for the imaging of bony structures, appearing white on the CT data. It is also used for searching for geometrical data of the cochlea at certain important regions (Spoor and Zonneveld 1998) and in surgical assessment for cochlear implants candidacy (Todd et al. 2009). Numerous artifacts can occur on CT images, such as partial volume effect, streak, motion, beam-hardening, ring and bloom artifacts (Popilock et al. 2008).

MR is the most widely used technique in the field of radio imaging (Macovski 1983, Prince and Links 2006). It is based on the achievement of a variable image contrast by using different pulse sequence and by changing parameters corresponding to longitudinal and transversal relaxation times. Signal intensities on those two times weighted images relate to specific tissue characteristics (Hendee and Morgan 1984). The contrast on MR images is a factor dependent on pulse sequence parameters. Partial volume effect, intensity inhomogeneity, motion, wrap around, Gibbs ringing are some artifacts that can occur on MR

images. MR is often used to create soft tissues models of the cochlea. The major disadvantage of this technique lies in the difficulty or even inability to display bony structures. Nevertheless, its main application has concentrated on the reconstruction of the fluid chambers of the inner ear (Counter et al. 2000, Thorne et al. 1999). MR also allows enhanced 3D visualization, especially when the inner ear has to be evaluated. With MR the different portions of the facial and vestibulo-cochlear nerve can be depicted to very high details (Rodt et al. 2002).

Both CT and MR images suffer from partial volume effects and motion artifacts. CT has inferior soft tissue contrast when compared to MR. Also, in the case of MR image quality is not so good as CT. MR is relatively safe and unlike CT modality, can be used as often as necessary.

The ossicles (malleus, incus and stapes), the tympanic membrane and the external ear canal are the types of ear structures that are better represented in CT images. Although in CT images the cochlea, the semicircular canals and the vestibule are also visible, is in MR images that those structures are better represented (see Figures 4 and 5). In MR images the facial and vestibular nerves are also well represented (see Figure 5).

3.2. Segmentation Algorithms for Ear Structures

Segmentation is one of the most important techniques for image analysis (Sonka et al. 2008). Its purpose is to partition an image into non-overlapping, component regions that are homogeneous with respect to some characteristic, such as intensity or texture (Gonzalez and Woods 1992, Haralick and Shapiro 1985). Numerous approaches regarding image segmentation techniques are available in the literature (Gonzalez and Woods 1992, Pham et al. 2000, Sharma and Aggarwal 2010, Sonka et al. 2008, Withey and Koles 2007). Segmentation of the human ear includes the outlining labyrinth (cochlea, semicircular canals and vestibule), the ossicles, facial and vestibular nerves, external ear canal and tympanic membrane. A manual segmentation, including some anatomical structures on axial and coronal CT slices are depicted in Figures 4b) and c), respectively.

In order to identify relevant publications on the article's subject, during January to March 2012, a literature review was performed. The following databases were explored to identify thesis, articles, conference papers and reviews: ISI Web of Knowledge, Scopus, Google-Scholar and PubMed. The search was limited to manuscripts to which the authors had full access to and published in English. For each database, the search was accomplished

considering the following keywords: computational vision, segmentation approaches, and ear structures in CT, MRI or VO images. A first selection was accomplished considering the titles and the abstracts of the publications. Then, the duplicated titles were excluded, and the full texts were analyzed; only the publications that included a segmentation approach, in a whole or parts, were included. Further searches were conducted in the World Wide Web using the search engine Google to identify books, standards and publications from regulatory authorities. The remaining articles presented in this section are by way of example.

The publications addressed in this study are indicated in Table 1.

In this paper, we divide the segmentation algorithms into four groups: (1) Thresholding, (2) Clustering, (3) Deformable models and (4) Atlas based. Applications of each type to the ear are illustrated to further state their characteristics.

3.2.1 Thresholding

Thresholding is one of the most used segmentation algorithms in digital images. They basically create a portioning of the image based on quantifiable features like image intensity or gradient magnitude. Thresholding algorithms can roughly be categorized into two groups, namely: Global thresholding and Local thresholding, according to histogram or local properties of the image, respectively. Otsu method is one of the most well-known segmentation algorithms that uses global thresholding (Otsu 1979). Local thresholding algorithms can further be divided into edge based, region based ones and hybrids. Edge-based algorithms use edge detectors to find edges in the image. Laplacian (Davis 1975), Sobel (Davis 1975) and Canny (Canny 1986) operators are some examples of edge detectors. Laplacian finds edges by looking for zero crossings after filtering the image with a specified filter; Sobel finds edges using the Sobel approximation to the derivative, returning edges at those points where the gradient of the image is maximum; Canny operator finds edges by looking for local maxima of the gradient of the image. The gradient is calculated using the derivative of a Gaussian filter. The algorithm uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. As such, this algorithm is less likely than the others to be fooled by noise, and more likely to detect true weak edges.

Region based algorithms examine pixels in an image and build disjoint regions by merging neighborhood pixels with homogeneous properties based on a predefined similarity criterion. Region growing is the simplest region based algorithm, and it starts by selecting a pixel or a

group of pixels called seed points, which belong to the structure of interest. Then, the neighboring pixels of each seed point are inspected and those with properties similar to the original seeds are added to the region that the seeds belong to, and thus, the region is growing as shape is also changing. The procedure stops when no more pixels can be added. For a region growing algorithm to be automatic and therefore no initial seed dependence needed, statistical information and a prior knowledge can be integrated into the algorithms (Dehmeshki J. et al. 2003, Pohle R. and KD. 2001). Even so, due to the intrinsic dependence on intensity of the region growing algorithms, they tend to have difficulties to control the leakage or eliminate the influence of partial volume effect.

Finally, hybrid algorithms fuse region information with a boundary detector to complete the segmentation. A typical hybrid algorithm is the watershed, which combine the image intensity with the gradient information. It is based on the assumption that the gradient magnitude of the image is a topographic surface. The gradient local minimum of each region is like a valley from which the water will rise up. The position where each two valleys are converging gives rise to a boundary called watershed line. Each local minimum is then surrounded by the watershed line, which represents a segmentation region.

Global thresholding is one of the most used algorithms to identify the inner ear, localized in the temporal bone (Lee et al. 2010b, Melhem et al. 1998, Rodt et al. 2002). This method was applied in MR (Melhem et al. 1998) and CT images (Lee et al. 2010b, Rodt et al. 2002). Local thresholding, using region growing or watershed is also used to segment ear structures. In (Seemann et al. 1999), authors define an individual threshold based interval density value for each anatomical structure in order to perform an interactive volume-growing segmentation, especially of the temporal bone on spiral-CT images. In the same image modality, a Connected threshold region growing algorithm was applied in (Todd et al. 2009) to extract the external ear canal and the cochlea. Connected threshold region growing requires the user to specify the index of a seed point and the lower and upper threshold limits. Pixels are included into the region of interest if their intensity values are within the threshold range specified. In order to iterate through pixels within the image and establish the region of interest (ROI), the connected threshold applies a flood iterator for visiting neighboring pixels. To perform the segmentation of the semicircular canals, but in the case of micro-CT, a watershed algorithm was used, designed for boundary determination in situations where objects appear to overlap or are blurred together (Bradshaw et al. 2010). The strategy used to reconstructing a complete semicircular canal is then by combined with an automated tracking system using Active contours (see Section 2.2.3).

3.2.2 Clustering

Clustering is a discovering process that organizes image structures into clusters such that the structures within a given cluster have a high degree of similarity, whereas structures belonging to different clusters have a high degree of dissimilarity (Kaufman and Rousseeuw 2005). It is considered an unsupervised learning technique since it does not require a training data to be efficient. In order to compensate for the lack of training data, clustering methods train themselves using available data (Pham et al. 2000). They although require an initial segmentation (or equivalent, initial parameters). One of the most commonly used algorithms for clustering is the fuzzy c-mean (Bezdek et al. 1993), which generalizes the k-means algorithm (Bezdek et al. 1993) allowing for soft segmentation based on fuzzy set theory (Zadeh 1965). The vestibular system presented in the inner ear is segmented in (Shi et al. 2010) by combining a clustering algorithm with a deformable model (see the next section). Automatic MR segmentation of the vestibular system involves the following steps: region of interest (ROI) extraction, resampling to make the image isotropic, edge-preserving filtering, k-means clustering and fine-tune using a deformable model. The k-means was applied as a pre-segmentation step to categorize the voxels into background and foreground based on their signal intensities. The foreground cluster contained several connected components, among which the largest was chosen as a coarsely defined vestibular region.

3.2.3. Deformable models

Deformable models were introduced by Kass et al. (Michael Kass et al. 1988) as a deformable contour in 2D and generalized to 3D by Terzopoulos and Metaxas (Terzopoulos and Metaxas 1991). Later, deformable models with the capacity of topological transformation were developed. Deformable models can be classified into Parametric and Geometric deformable models depending on the representation way of the contour.

Typical parametric deformable models are the Active parametric contours (Michael Kass et al. 1988), also called snakes, which were the first deformable models used for medical image analysis. The main concept associated to the snakes is its energy. Similar to a physics process, the energy of the contour is composed by two terms: internal energy, which depends on the elasticity and rigidity of the model, and external energy associated to image characteristics. The final contour is obtained by an energy minimizing formulation, corresponding to an equilibrium situation between the internal and external forces associated

to the image. However, in non-interactive applications, initial contours of the snake model should be placed near the region of interest to guarantee a good performance. On the other hand, the shape of the region of interest has to be well known from the beginning, since deformable models are parametric and incapable of topological transformations without additional algorithms (McInerney and Terzopoulos 1996). Gradient information of the input image can be incorporated into the snake model, originating a method called Gradient Vector Field (GVF). The GVF is distinguished from nearly all previous snake formulations because its external forces cannot be written as the negative gradient of a potential function. Therefore, it cannot be formulated using the standard energy minimization framework; instead, it is specified directly from a force balance condition. In (Xie et al. 2005) a Generalized Gradient Vector Field snake (GGVF) algorithm was applied to VO images with the aim to delineate the tympanic membrane boundaries and to detect color abnormalities in the tympanic membrane. This geometric GGVF snake is useful to delineate boundaries with small gaps and tympanic membrane boundaries present this feature. The GGVF snake presents advantages over the traditional snake, which demonstrate its efficiency to segment the tympanic membrane boundaries, such as its insensitivity to initialization and its ability to move into boundary concavities. Furthermore, GGVF snake does not need prior knowledge about whether to shrink or expand toward the boundary. The GGVF snake also has a large capture range, which means that, barring interference from other objects, it can be initialized far away from the boundary. This increase capture range is achieved through a diffusion process that does not blur the edges themselves, as such multi-resolution methods are not needed (Xu and Prince 1998). Bradshaw et al. (Bradshaw et al. 2010) used a 2D B-spline snake to reconstruct cross-sectional slices of the semicircular canal taken from CT imaging. Tabrizi (Tabrizi 2003) used two different active contour approaches, i.e., parametric active contours and discrete dynamic contours and compared them in the segmentation of middle ear images from MR images. These two algorithms showed successfully similar boundary identification results. However, the original active contour has, intrinsically, some limitations. The small capture range and the convergence of the algorithm are mostly dependent of the initial position. Besides, it also has difficulties in progressing into boundary concavities. To overcome some of these difficulties, Yoo et al. (Yoo et al. 2001) used a coarse-to-fine strategy to segment the cochlea in human spiral-CT images. In coarse segmentation, intensity range and volume-of-interest of the cochlea were defined. In fine segmentation, the output of course segmentation was refined, and the cochlea was identified from mixed surrounding structures by an elaborated snake algorithm. To compensate the

inconsistencies between adjacent contours, Poznyakovskiy et al. (Poznyakovskiy et al. 2008) added to the standard snake approach a new energy linking the contours on consecutive slices of the cross-sections on pig cochlea micro-CT images to segment the cochlea. Noble et al. (Noble et al. 2011) proposed an algorithm based on a deformable model of the cochlea and its components to automatically segment intracochlear structures. To build such model, they first manually segment intracochlear structures in a series of scans from micro-CT images, which were used to build the active shape model (Vasconcelos and Tavares 2008). The procedure they used for segmentation was: first, the built model was placed in the input image to initialize the segmentation; then, better solutions are found while deforming the shapes only in ways that are described by the pre-computed modes of the variation and finally, after iterative shape adjustments, the shape converges, and the segmentation was complete.

Geometric deformable models are characterized by Level set algorithms and include the following models: Mumford-Shah (Mumford and Shah 1989), Chan & Vese (Chan and Vese 2001) and Malladi et al. (Malladi et al. 1995). They are based on a curve evolution that is related to the geometric characteristics of the region of interest. These models adjust to the topology of the target, and they can easily adapt its shape. The main idea of the Level set is to minimize a function solving the corresponding Partial Differential Equation (PDE). The algorithm involves a contour implicitly by manipulating the higher dimensional function. Typical geometric deformable models include Level set algorithm. Level set is involved in the image segmentation problem by asking the user to draw a contour outside or inside the object, and then the contour will shrink or extend. The procedure will be ended when the contour meets the boundary of the object to be segmented. The drawback of the Level set algorithms is the definition of a proper speed function as it plays the main role in controlling the direction of contour shrink or extending as well as in finding the endpoint of the procedure. Medical segmentation methods of this class can be divided into two subclasses; 2D Level set methods and 3D Level set methods. The approach described by Xianfen et al. (Xianfen et al. 2005) is a 3D Level set method that requires a low level of intervention for the segmentation of cochlea and semicircular canals in spiral-CT images. The user locates a sphere contour in the cochlea region and uses it as the initial contour to run the Level set. Then, the 3D narrow band Level set algorithm was used to finish fine segmentation. In the final step of the 3D narrow band Level set, the segmented results are rendered with the Marching Cubes Algorithm (Lorensen and Cline 1987). The Mumford-Shah algorithm (Mumford and Shah 1989) was used by Comunello et al. (Comunello et al. 2009) to segment the tympanic membrane of VO data images. This algorithm presents effective in image

segmentation of the tympanic membrane, because it has high robustness in the presence of noise and in the choice of place to start the segmentation (Tsai et al. 2001). In addition, this algorithm guarantees that no segment leakage between structures occurs, and it also allows knowledge the quantitative information about tympanic membrane perforations, so it is indicated for clinic diagnosis.

3.2.4. Atlas based

Atlas based segmentation has become a standard paradigm for exploiting prior knowledge in medical image segmentation (Duay et al. 2005). The main idea of this approach is to generate an atlas by compiling some prior information about a structure and use this atlas to aid segmentation of similar structures. This information can be the contour of an object in a 2D image. After generating the atlas, it is placed near to the desired contour and registered to the input images by some local transformation. The registered atlas gives the segmentation result. Various registration techniques can be used in the registration process (Hill et al. 2001, Zitova and Flusser 2003). Generating the atlas based on a single sample is inadvisable, because the selected sample may not be a typical one and besides it does not may contain any information of variability, which cannot determine whether a deformed shape is an acceptable shape or not. One method that helps model anatomical variability is the use of Probabilistic atlas (Thompson and Toga 1997), which represents the spatial distribution of probability that a pixel belongs to a particular object (Hyunjin et al. 2003). The disadvantage of these is that it requires a lot of data to be collected. An atlas based registration process was used in the work of Noble et al. (Noble et al. 2009, Noble et al. 2010) to automatically identify the labyrinth, ossicles and external auditory canals in CT database images. For the segmentation of the facial nerve and chorda tympani, topological similarity between images cannot be assumed to the highly variable pneumatized bone. Therefore, facial nerve and chorda tympani were identified using a novel method that combines an atlas based approach with a minimum cost path finding algorithm. Christensen et al. (Christensen et al. 2003), used a combination of an atlas based approach with a minimum cost path finding algorithm for automatically segment the cochlea, the vestibule, the semicircular canals and the internal auditory canal from CT data.

3.3. Experimental Results: examples and discussion

In this section, some of the algorithms introduced in the previous section are applied on CT images in order to illustrate their use, and discuss their main advantages and disadvantages.

3.3.1. Examples

The segmentation results of Otsu method, Canny edge detector, region growing and watershed algorithms are illustrated in Figures 6 to 9. The segmentation result from Otsu method is not satisfactory (Figure 6). Although it can be observed the successfully identification of some structures, such as the ossicles and the semicircular canal, other structures, like the cochlea and the external auditory canal, are far away from being clearly identified. Otsu method is limited by the considerable amount of noise presented in the input image, by the small size of the structures and by the large variance of the background intensities. Therefore, Otsu method is useful as an initial step for further segmentation. Figure 7 shows the result of applying Canny edge detector using standard parameters in the Matlab Image Processing Toolbox (The MathWorks, Inc., USA). Usually, an anisotropic diffusion filter, such the one proposed in (Perona and Malik 1990), is applied before the Canny edge segmentation for enhancing and smoothing the original image. This filter blurs areas of low contrast and enhances the edges, as a high pass filter. Thus, this filter is used to reduce the noise of the input image. The result shows that boundaries obtained are discontinuous, incomplete or wrongly connected. Changing parameters or applying different filters do not reveal any solution once all the boundaries of the objects were detected whereas all other edges were removed or most of the edges were obtained whereas an amount of noise was increased. These results are due to the noise presented in the image and partial volume effect. An example of a region growing algorithm application is shown in Figure 8. For each region that needed to be segmented, a seed point is manually defined. The seed is then iteratively grown by comparing all unallocated adjacent pixels to the same region. A measure of similarity based on the difference between the intensity value of the pixel and the mean value of the region is used. The pixel with the smallest difference is allocated to the region. This process stopped when the intensity difference between the region mean and the new pixel became higher than a certain predefined threshold value. Results show that the areas corresponding to the ossicles and the semicircular canal are successfully segmented. However, the segmentation of the cochlea and the external auditory canal is not so satisfactory, due to the influence of the intensity dependence of these algorithms. The

boundary of the vestibule leaks in the upward direction and the facial nerve is almost erased. Figure 9 shows the results of applying a watershed algorithm. It is observed that a complete segmentation of the image was accomplished; however, due to the presence of several pixels with local minimums of gradient magnitude, the resultant over segmentation is appreciable.

Figure 10 presents the result by using a fuzzy c-means algorithm. An anisotropic diffusion filter was applied (Perona and Malik 1990) before the fuzzy c-means segmentation for enhancing and smoothing the original image. Four clusters were defined with initial mean intensities. The clustering process stopped when the maximum number of iterations was reached. Results show that the boundaries of the external auditory canal, the ossicles and the semicircular canal are successfully segmented. However, the boundaries of the vestibule, cochlea and facial nerve are not so effectively segmented, due to noise and outliers of the image.

Figures 11 and 12 illustrate the segmentation result of applying a snake algorithm and a Level set algorithm, respectively, both proposed in (Lankton and Tannenbaum 2008). It is observed that boundaries are regular and smooth due to the items defined in the speed function of the deformable models. However, the external auditory canal and the cochlea were not totally segmented, either in the snake algorithm as in the Level set algorithm. In our experiments, we defined the stopping criterion for the algorithms as being the same for all structures, which means that, in some cases, it prevents the contours to keep moving (Figures 11 - right and 12 - right). In other cases, the moving contours may leak or shrink to disappear after long time evolution. Nevertheless, with these parameters, the algorithms provide a good segmentation of the ossicles, the facial nerve and the semicircular canal.

3.3.2. Discussion

The segmentation of ear structures is still an open area for more research. Unfortunately, an objective comparison of their performance is not conceivable, or at least, not fair, due to the lack of an accepted manual segmentation common dataset. Nevertheless, some approaches are well-established in the literature in order to measure the results from automatic and manual processes, such as the relative intersection between the areas (Korfiatis et al. 2007), the Hausdorf distance (Ma et al. 2011) or the receiver operating characteristic (ROC) analysis (Gruszauskas et al. 2009, Gruszauskas et al. 2008).

There are several structures of the ear that have been successfully segmented using thresholding algorithms. Thresholding algorithms are fast, computationally efficient and

inexpensive. However, due to noise sensitivity and intensity non-uniformity of the original images, threshold based segmentation can cause segmented regions with inner holes or even wrongly connected regions. In the most common medical images, segmentation results using this algorithm alone are not satisfactory. Therefore, thresholding algorithms are usually used as a pre-processing step for posterior segmentation algorithms.

Clustering algorithms are simple, general and computational efficient, due to the lack of spatial modeling. However, they are very sensitive to noise, to intensity inhomogeneity and in the case of fuzzy c-means they could also be sensitive to the number of clusters, the initial partition and the stopping criterion. Pixels that belong to the same anatomical structure with inhomogeneous features may be grouped into different clusters. Many algorithms were introduced to make fuzzy c-means robust against noise and inhomogeneity but most of them still are not flawless (Acton and Mukherjee 2000, Catte et al. 1992, Zhang and Chen 2004). As such, good results could be achieved in the case of structures with large shape variations in medical images.

As we can verify from the state-of-the-art of section 2.2, the methods that were used by most authors in the ear segmentation were the deformable models. Deformable models have the ability to directly generate closed parametric curves from images and to be smoothness to noise and artifacts. Moreover, deformable models can be implemented on the continuum space and achieve sub pixel accuracy (Xu et al. 2000), a highly desirable property for medical imaging applications. Geometric deformable models have advantages over Parametric models due to their parameterization independence, intrinsic behavior and easy implementation (Xiao et al. 2003). However, a long demanded advantage of Geometric deformable models is the ability to handle topology changes, crucial in applications where the object to be segmented has a known topology that must be preserved (Suri et al. 2007). Comparing with the other two segmentation algorithm groups, deformable models are more flexible and can be used for more complex segmentation. However, usually, thresholding, clustering and deformable models based algorithms require manual interaction. Besides, the selection of appropriate parameters on the deformable models procedures constitutes a challenge (Sharma and Aggarwal 2010), since it is critical to the final segmentation results.

Atlas based algorithms are general applicable, robust and have a high computationally complexity. They are simple to implement since only a registration framework and a number of pre-segmented datasets are required. Expert knowledge is required to build these datasets. They also have the ability to incorporate prior information and employ global image information. The major disadvantage of these procedures is that every single atlas can only be

applied to a small number of specific images whose shapes are similar to the ones used to building the atlas. Therefore, when there is not enough contrast between tissues, the atlas based methods are the best choice (Balafar et al. 2010). Because of this lack of contrast and topological variation of the images, atlas based methods alone do not lead to results that are sufficient accurate (Noble et al. 2008). Both deformable models and atlas based algorithms are sensitive to the initial definition of the contours. Comparing with thresholding, clustering and atlas based algorithms, only deformable models based algorithms are able to handle structures with complex topology. Deformable models are promising because they incorporate prior knowledge about the location, size and shape of the anatomical structures of interest. However, parameters must be selected properly to get satisfied results.

From the observation of the experimental results, some of them presented in the previous section, we can point out that for the segmentation of the cochlea, using a fuzzy c-means or a region growing algorithm may be the best solution; for the segmentation of the semicircular canal, Otsu, region growing, watershed or deformable models will work perfectly; in the case of the vestibule, Otsu, watershed and deformable models proved to be the most indicated algorithms; for the facial nerve, just the deformable models are advised because only the Level set algorithm worked successfully for this structure; in the case of the external ear canal, if the parameters of the speed function were well defined, deformable models may be the most recommended algorithms, otherwise, a region growing or a fuzzy c-means algorithm will be effective; for the segmentation of the ossicles, all algorithms presented are recommended.

Fully automated, less user dependent and more efficient segmentation algorithms should be developed using prior shape information on the structures to be segmented. A possible way to design such algorithms could be by using improved image atlas, effective registration techniques and combining multiple segmentation approaches.

4. Conclusions and Future Trends

Image segmentation algorithms are essential for the construction of realistic biomechanical models of the human ear, which could be helpful for the simulation, understanding, diagnosis and treatment of ear disorders. In this paper, we reviewed some of the works regarding the ear biomechanical modeling focusing our analysis in two main groups: lumped parameter models and distributed parameter models. As far as our concern, there is no model built on

the basis of a totally automatic, subject-specific approach. In order to aim that, algorithms for the ear segmentation in VO, CT and MR images were reviewed.

Segmentation algorithms were classified into four categories: thresholding, clustering, deformable models and atlas based. A critical description and analysis of the state of the art in this field were provided. Some experiments applying these algorithms were illustrated in axial and coronal CT ear images. The experiments confirmed that the segmentation of ear structures is still an open area for more research since various drawbacks and weaknesses of the current methods must still be addressed.

During this review, we have identified the following trends and perspectives for future developments concerning the analysis of ear images: The acquisition of micro-CT images of specific parts of the ear may provide a modality for imaging the small structures involved, such as the ossicles, the semicircular canals and the facial nerve, with very high spatial resolution. The acquisition of CT and MR images with a lower spacing between slices or higher slice thickness, i.e. with higher Z-axis resolutions, may lead to more data about the structures involved and also to enhance the contrast between such structures and the image background. In order to segment the ear structures usually influenced by the partial volume effect and intensity inhomogeneity, such as the external auditory canal and the cochlea, image cues should be combined with the expected relative position of the structures; for example, using a 3D model atlas and applying the algorithms based on deformable models. So as to fulfill the segmentation of structures whose boundaries are barely defined or incomplete through the image appearance, which is often the case of the facial nerve and the ossicles, a good solution can be the combination of restrictions on shape variations with a prior shape model.

To conclude, the future direction of the research concerning the analysis of ear images, both in terms of medical diagnosis and biomechanical simulation, will be towards the use of imaging acquisition processes with superior special image resolution and contrast, the developing of more accurate, efficient, automated and faster computational algorithms based on previous knowledge about the structures involved. Also, the registration, i.e. the fusion, of CT and MR data could improve the possibility of determining adjacent structures, such as nerve structures, soft-tissues masses and tumors.

Acknowledgments

This work was partially done in the scope of the projects “Methodologies to Analyze Organs from Complex Medical Images – Applications to Female Pelvic Cavity” and “Bio-computational study of tinnitus”, with references PTDC/EEA-CRO/103320/2008 and PTDC/SAU-BEB/104992/2008, respectively, financially supported by Fundação para a Ciência e a Tecnologia, in Portugal.

References

Acton ST, Mukherjee DP. 2000. Scale space classification using area morphology. IEEE Trans Image Process. 9(4):623-635.

Balafar M, Ramli A, Saripan M, Mashohor S. 2010. Review of brain MRI image segmentation methods. Artif Intell Rev. 33(3):261-274.

Beer HJ, Bornitz M, Hardtke HJ, Schmidt R, Hofmann G, Vogel U, Zahnert T, Huttenbrink KB. 1999. Modelling of components of the human middle ear and simulation of their dynamic behaviour. Audiol Neurootol. 4(3-4):156-162.

Belytschko T, Moran L. 2000. Nonlinear Finite Element for Continua and Structures. Wiley.

Bezdek JC, Hall LO, Clarke LP. 1993. Review of MR image segmentation techniques using pattern recognition. Med Phys. 20(4):1033-1048.

Bornitz M, Hardtke H-J, Zahnert T. 2010. Evaluation of implantable actuators by means of a middle ear simulation model. Hearing Res. 263(1–2):145-151.

Bornitz M, Zahnert T, Hardtke H, Huttenbrink K. 1999. Identification of parameters for the middle ear model. Audiol Neurootol. 4(3-4):163-169.

Bradshaw AP, Curthoys IS, Todd MJ, Magnussen JS, Taubman DS, Aw ST, Halmagyi GM. 2010. A Mathematical Model of Human Semicircular Canal Geometry: A New Basis for Interpreting Vestibular Physiology. JARO: Journal of the Association for Research in Otolaryngology. 11:145-159.

Canny J. 1986. A Computational Approach to Edge Detection. IEEE Trans Pattern Anal Mach Intell. 8(6):679-698.

Catte F, Lions PL, Morel JM, Coll T. 1992. Image selective smoothing and edge detection by nonlinear diffusion. SIAM J Numer Anal. 29(1):182-193.

Chan TF, Vese LA. 2001. Active contours without edges. IEEE Trans Image Process. 10(2):266-277

Chitore D, Saxena S, Mukhopadhyay P. 1983. Electronic model of the middle ear. Med Biol Eng Comput. 21(2):176-178.

Chou CF, Yu JF, Chen CK. 2011. The natural vibration characteristics of human ossicles. Chang Gung Med J. 34(2):160-165.

Christensen GE, He J, Dill JA, Rubinstein JT, Vannier MW, Wang G. 2003. Automatic Measurement of the Labyrinth Using Image Registration and a Deformable Inner Ear Atlas. Academic Radiol J. 10:988-999.

Comunello E, Wangenheim A, Junior VH, Dornelles C, Costa SS. 2009. A computational method for the semi-automated quantitative analysis of tympanic membrane perforations and tympanosclerosis. Comput Biol Med. 39(10):889-895.

Costa MFG. 2008. Biomechanical study of the middle ear (In Portuguese) [PhD dissertation]. [Porto (Portugal)]: Universidade do Porto.

Counter SA, Bjelke B, Borg E, Klason T, Chen Z, Duan ML. 2000. Magnetic resonance imaging of the membranous labyrinth during in vivo gadolinium (Gd-DTPA-BMA) uptake in the normal and lesioned cochlea. Neuroreport. 11(18):3979-3983.

Davis LS. 1975. A survey of edge detection techniques. Comput Graph Image Process. 4(3):248-270.

Dehmeshki J., Ye X., J. C. 2003. Shape Based Region growing Using Derivatives of 3D Medical Images: Application to Semiautomated Detection of Pulmonary Nodules. Proceedings of the Proceedings of the 2003 International Conference on Image Processing. Barcelona. Spain.

Duay V, Houhou N, Thiran JP. 2005. Atlas-based segmentation of medical images locally constrained by Level sets. Paper presented at: ICIP 2005. Proceedings of the IEEE International Conference on Image Processing. Genoa. Italy.

Eiber A, Schiehlen W. 1995. Reconstruction of Hearing by Mechatronic Devices. Paper presented at: ICRAM 1995. Proceedings of International Conference on Recent Advances in Mechatronics. Istambul. Turkey.

Fay J, Puria S, Decraemer WF, Steele C. 2005. Three approaches for estimating the elastic modulus of the tympanic membrane. J Biomech. 38(9):1807-1815.

Ferris P, Prendergast PJ. 2000. Middle-ear dynamics before and after ossicular replacement. J Biomech. 33(5):581-590.

Folowosele FO, Camp JJ, Brey RH. 2004. 3D imaging and modeling of the middle and inner ear. Proceedings of SPIE: Medical imaging, Visualization, Image-Guided Procedures and Displays. 5367:508-515.

Funnell WR, Decraemer WF, Khanna SM. 1987. On the damped frequency response of a finite-element model of the cat eardrum. J Acoust Soc Am. 81(6):1851-1859.

Funnell WR, Laszlo CA. 1978. Modeling of the cat eardrum as a thin shell using the finite-element method. J Acoust Soc Am. 63(5):1461-1467.

Funnell WR, Laszlo CA. 1982. A critical review of experimental observations on ear-drum structure and function. J Otorhinolaryngol Relat Spec. 44(4):181-205.

Gan RZ, Cheng T, Dai C, Yang F, Wood MW. 2009. Finite element modeling of sound transmission with perforations of tympanic membrane. J Acoust Soc Am. 126(1):243-253.

Gan RZ, Feng B, Sun Q. 2004. Three-dimensional finite element modeling of human ear for sound transmission. Ann Biomed Eng. 32(6):847-859.

Gan RZ, Reeves BP, Wang X. 2007. Modeling of sound transmission from ear canal to cochlea. Ann Biomed Eng. 35(12):2180-2195.

Gan RZ, Sun Q, Feng B, Wood MW. 2006. Acoustic-structural coupled finite element analysis for sound transmission in human ear-pressure distributions. Med Eng Phys. 28(5):395-404.

Gan RZ, Wang X. 2007. Multifield coupled finite element analysis for sound transmission in otitis media with effusion. J Acoust Soc Am. 122(6):3527-3538.

Gentil F, Jorge RMN, Ferreira AJM, Parente M, Moreira M, Almeida E. Biomechanical Study of Middle Ear. 2005. Proceedings of the VIII International Conference on Computational Plasticity. Barcelona. Spain.

Gentil F, Parente M, Martins P, Garbe C, Jorge RN, Ferreira A, Tavares JM. 2011. The influence of the mechanical behaviour of the middle ear ligaments: a finite element analysis. Proc Inst Mech Eng H. 225(1):68-76.

Gentil F, Parente M, Martins P, Garbe C, Paco J, Ferreira AJ, Tavares JM, Jorge RN. 2012. The influence of muscles activation on the dynamical behaviour of the tympano-ossicular system of the middle ear. Comput Methods Biomech Biomed Engin. 19:19.

Gonzalez RC, Woods EE. 1992. Digital Image Processing. 2nd ed. Prentice Hall.

Goode RL, Killion M, Nakamura K, Nishihara S. 1994. New knowledge about the function of the human middle ear: development of an improved analog model. Am J Otol. 15(2):145-154.

Gruszauskas NP, Drukker K, Giger ML, Chang RF, Sennett CA, Moon WK, Pesce LL. 2009. Breast US computer-aided diagnosis system: robustness across urban populations in South Korea and the United States. Radiology. 253(3):661-671.

Gruszauskas NP, Drukker K, Giger ML, Sennett CA, Pesce LL. 2008. Performance of breast ultrasound computer-aided diagnosis: dependence on image selection. Acad Radiol. 15(10):1234-1245.

Haralick RM, Shapiro LG. 1985. Image segmentation techniques. Comput Vis Graph Image Process. 29(1):100-132.

Hendee WR, Morgan CJ. 1984. Magnetic resonance imaging. Part I-physical principles. West J Med. 141(4):491-500.

Hill DLG, Batchelor PG, Holden M, Hawkes DJ. 2001. Medical image registration. Phys Med Biol. 46(3):R1-R45.

Hudde H, Weistenh, fer C. 1997. A Three-Dimensional Circuit Model of the Middle Ear. Acta Acust United Ac. 83(3):535-549.

Hyunjin P, Bland PH, Meyer CR. 2003. Construction of an abdominal probabilistic atlas and its application in segmentation. IEEE Transactions on Medical Imaging. 22(4):483-492.

Jun BC, Song SW, Cho JE, Park CS, Lee DH, Chang KH, Yeo SW. 2005. Three-dimensional reconstruction based on images from spiral high-resolution computed tomography of the temporal bone: anatomy and clinical application. J Laryngol Otol. 119(9):693-698.

Kaufman L, Rousseeuw P. 2005. Finding Groups in Data: An Introduction to Cluster Analysis. 99th ed. Wiley-Interscience.

Koike T, Wada H, Kobayashi T. 2002. Modeling of the human middle ear using the finite-element method. J Acoust Soc Am. 111(3):1306-1317.

Korfiatis P, Skiadopoulos S, Sakellaropoulos P, Kalogeropoulou C, Costaridou L. 2007. Automated 3D segmentation of lung fields in thin slice CT exploiting wavelet preprocessing. Proceedings of the 12th International Conference on Computer Analysis of Images and Patterns. Vienna. Austria.

Kringlebotn M. 1988. Network model for the human middle ear. Scan Audiol. 17(2):75-85.

Ladak HM, Funnell WR. 1996. Finite-element modeling of the normal and surgically repaired cat middle ear. J Acoust Soc Am. 100(2):933-944.

Lankton S, Tannenbaum A. 2008. Localizing Region-Based Active Contours. IEEE Transactions on Image Processing. 17(11):2029-2039.

Lee C-F, Chen P-R, Lee W-J, Chou Y-F, Chen J-H, Liu T-C. 2010a. Computer aided modeling of human mastoid cavity biomechanics using finite element analysis. J Adv Signal Process. 2010:1-9.

Lee CF, Chen PR, Lee WJ, Chen JH, Liu TC. 2006. Three-dimensional reconstruction and modeling of middle ear biomechanics by high-resolution computed tomography and finite element analysis. Laryngoscope. 116(5):711-716.

Lee DH, Chan S, Salisbury C, Kim N, Salisbury K, Puria S, Blevins NH. 2010b. Reconstruction and exploration of virtual middle-ear models derived from micro-CT datasets. Hearing Res. 263(1-2):198-203.

Lesser TH, Williams KR. 1988. The tympanic membrane in cross section: a finite element analysis. J Laryngol Otol. 102(3):209-214.

Lesser TH, Williams KR, Blayney AW. 1991. Mechanics and materials in middle ear reconstruction. Clin Otolaryngol Allied Sci. 16(1):29-32.

Liu Y, Li S, Sun X. 2009. Numerical analysis of ossicular chain lesion of human ear. Acta Mech Sinica. 25:241-247.

Lorensen WE, Cline HE. 1987. Marching cubes: A high resolution 3D surface construction algorithm. Comput Graph. 21(4):163-169.

Ma Z, Jorge RN, Mascarenhas T, Tavares JM. 2011. Novel approach to segment the inner and outer boundaries of the bladder wall in T2-weighted magnetic resonance images. Ann Biomed Eng. 39(8):2287-2297.

Macovski A. 1983. Medical imaging systems. 1st ed. Prentice-Hall.

Malladi R, Sethian JA, Vemuri BC. 1995. Shape modeling with front propagation: a Level set approach. IEEE Transactions on Pattern Anal Mach Intell. 17(2):158-175.

McInerney T, Terzopoulos D. 1996. Deformable models in medical image analysis: a survey. Med Image Anal. 1(2):91-108.

Melhem ER, Shakir H, Bakthavachalam S, MacDonald CB, Gira J, Caruthers SD, Jara H. 1998. Inner ear volumetric measurements using high-resolution 3D T2-weighted fast spin-echo MR imaging: Initial experience in healthy subjects. American Journal of Neuroradiology. 19(10):1819-1822.

Michael Kass, Abdrew P. Witkin, Terzopoulos D. 1988. snakes: Active contour models. Int J Comput Vision. 1(4):321-331.

Moller AR. 2006. Hearing: Anatomy, Physiology, and Disorders of the Auditory System. 2nd ed. San Diego: Academic Press.

Mumford D, Shah J. 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Commun Pure Appl Math. 42(5):577-685.

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

Noble JH, Labadie RF, Majdani O, Dawant BM. 2011. Automatic segmentation of intracochlear anatomy in conventional CT. IEEE Trans Biomed Eng. 58(9):2625-2632.

Noble JH, Rutherford RB, Labadie RF, Majdani O, Dawant BM. 2010. Modeling and segmentation of intra-cochlear anatomy in conventional CT. Paper presented at: Medical Imaging 2010: Image Processing. Proceedings of SPIE Progress in Biomedical Optics and Imaging. San Diego. California. USA.

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

Otsu N. 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man and Cybernetics. 9(1):62-66.

Parent P, Allen JB. 2010. Time-domain "wave" model of the human tympanic membrane. Hearing Res. 263(1–2):152-167.

Peake WT, Rosowski JJ, Lynch TJ. 1992. Middle-ear transmission - acoustic versus ossicular coupling in cat and human. Hearing Res. 57(2):245-268.

Pham DL, Xu C, Prince JL. 2000. Current methods in medical image segmentation. Annu Rev Biomed Eng. 2(1):315-337.

Pohle R., KD. T. 2001. Segmentation of Medical Images Using Adaptative Region growing . Paper presented at: Proceedings of the SPIE Medical Imaging. San Diego. California. USA.

Popilock R, Sandrasagaren K, Harris L, Kaser KA. 2008. CT artifact recognition for the nuclear technologist. J Nucl Med Tech. 36(2):79-81.

Poznyakovskiy AA, Zahnert T, Kalaidzidis Y, Schmidt R, Fischer B, Baumgart J, Yarin YM. 2008. The creation of geometric three-dimensional models of the inner ear based on micro computer tomography data. Hearing Res. 243(1-2):95-104.

Prendergast PJ, Ferris P, Rice HJ, Blayney AW. 1999a. Vibro-acoustic modelling of the outer and middle ear using the finite-element method. Audiol Neurootol. 4(3-4):185-191.

Prendergast PJ, Kelly DJ, Rafferty M, Blayney AW. 1999b. The effect of ventilation tubes on stresses and vibration motion in the tympanic membrane: a finite element analysis. Clin Otolaryngol Allied Sci. 24(6):542-548.

Prince JL, Links JM. 2006. Medical Imaging Signals and Systems. 1st ed. Pearson Prentice Hall.

Puria S, Allen JB. 1998. Measurements and model of the cat middle ear: Evidence of tympanic membrane acoustic delay. J Acoust Soc Am. 104(6):3463-3481.

Rabbitt RD, Holmes MH. 1986. A fibrous dynamic continuum model of the tympanic membrane. J Acoust Soc Am. 80(6):1716-1728.

Rodt T, Ratiu P, Becker H, Bartling S, Kacher DF, Anderson M, Jolesz FA, Kikinis R. 2002. 3D visualisation of the middle ear and adjacent structures using reconstructed multi-slice CT datasets, correlating 3D images and virtual endoscopy to the 2D cross-sectional images. Neuroradiology. 44(9):783-790.

Rosowski JJ, Merchant SN. 1995. Mechanical and acoustic analysis of middle ear reconstruction. Am J Otol. 16(4):486-497.

Seeley, Stephens, Tate. 2004. The Special Senses. Anatomy and Physiology. 6th ed. The MacGraw-Hill Companies.

Seemann MD, Seemann O, Bonel H, Suckfull M, Englmeier KH, Naumann A, Allen CM, Reiser MF. 1999. Evaluation of the middle and inner ear structures: comparison of hybrid rendering, virtual endoscopy and axial 2D source images. Eur Radiol. 9(9):1851-1858.

Sharma N, Aggarwal LM. 2010. Automated medical image segmentation techniques. J Med Phys. 35(1):3-14.

Shera CA, Zweig G. 1991. Phenomenological characterization of eardrum transduction. J Acoust Soc Am. 90(1):253-262.

Shi L, Wang D, Chu WCW, Burwell GR, Wong T-T, Heng PA, Cheng JCY. 2010. Automatic MRI segmentation and morphoanatomy analysis of the vestibular system in adolescent idiopathic scoliosis. NeuroImage. 54(1): 180-188.

Sim JH, Puria S. 2008. Soft tissue morphometry of the malleus-incus complex from micro-CT imaging. J Assoc Res Otolaryngol. 9(1): 5-21.

Sonka M, Hlavac V, Boyle R. 2008. Image processing, analysis, and machine vision. 2nd ed. Toronto: CL-Engineering.

Spoor F, Zonneveld F. 1998. Comparative review of the human bony labyrinth. American J Phys Anthropol. 107(S27):211-251.

Stieger C, Bernhard H, Waeckerlin D, Kornpis M, Burger J, Haeusler R. 2007. Human temporal bones versus mechanical model to evaluate three middle ear transducers. J Rehabil Res Dev. 44(3):407-415.

Sun Q, Chang KH, Dormer KJ, Dyer RK, Jr., Gan RZ. 2002. An advanced computer-aided geometric modeling and fabrication method for human middle ear. Med Eng Phys. 24(9):595-606.

Suri JS, Farag AA, Wang Y, Guo Q, Zhu Y. 2007. Medical Image Segmentation Based On Deformable Models And Its Applications. New York: Springer. p. 209-260.

Tabrizi JH. 2003. Using Active Contours for Segmentation of Middle-Ear Images [Master thesis]. [Montréal, Québec]: McGill University.

Terzopoulos D, Metaxas D. 1991. Dynamic 3D models with local and global deformations: deformable superquadrics. IEEE Transactions on Pattern Anal Mach Intell. 13(7):703-714.

Thompson PM, Toga AW. 1997. Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. Med Image Anal. 1(4):271-294.

Thorne M, Salt AN, DeMott JE, Henson MM, Henson OW, Jr., Gewalt SL. 1999. Cochlear fluid space dimensions for six species derived from reconstructions of three-dimensional magnetic resonance images. Laryngoscope. 109(10):1661-1668.

Todd C, Kirillov M, Tarabichi M, Naghdy F, Naghdy G. 2009. An analysis of medical image processing methods for segmentation of the inner ear. Paper presented at: ISADIS 2009. Proceedings of the IADIS Multiconference, Computer Graphycs, Visualization, Computer Vision and Image Processing. Algarve. Portugal.

Tsai A, Yezzi A, Jr., Willsky AS. 2001. Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification. IEEE Trans Image Process. 10(8):1169-1186.

Vasconcelos M, Tavares J. 2008. Methods to Automatically Built Point Distribuition Models for Objects like Hand Palms and Faces Represented in Images. Comp Mod Eng Sci. Tech Science Press. 36(3): 213-241.

Volandri G, Di Puccio F, Forte P, Carmignani C. 2011. Biomechanics of the tympanic membrane. J Biomech. 44(7):1219-1236.

Volandri G, Puccio FD, Forte P, Manetti S. 2012. Model-oriented Review and Multi-Body Simulation of the Ossicular Chain of the Human Middle Ear. Med Eng Phys (in press).

Wada H, Metoki T, Kobayashi T. 1992. Analysis of dynamic behavior of human middle ear using a finite-element method. J Acoust Soc Am. 92(6):3157-3168.

Wegel RL, Lane CE. 1924. The auditory masking of one pure tone by another and its probable relation to the dynamics of the inner ear. Phys Rev. 23(2):266-285.

Williams KR, Blayney AW, Lesser THJ. 1995. A 3-D finite element analysis of the natural frequencies of vibration of a stapes prosthesis replacement reconstruction of the middle ear. Clin Otolaryngol Allied Sci. 20(1):36-44.

Williams KR, Lesser TH. 1990. A finite element analysis of the natural frequencies of vibration of the human tympanic membrane. Part I. Br J Audiol. 24(5):319-327.

Withey DJ, Koles ZJ. Medical Image Segmentation: Methods and Software. Paper presented at: Joint Meeting of the 6th International Symposium on NFSI-ICFBI 2007. Proceedings of the Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging. Hangzhou. China.

Wright T. 2005. The Linear and Nonlinear Biomechanics of the Middle Ear [Doctoral thesis]. [Sweden]: KTH Royal Institute of Technology.

Xianfen D, Siping C, Changhong L, Yuanmei W. 2005. 3D semi-automatic segmentation of the cochlea and inner ear. IEEE Eng Med Biol Soc. 6:6285-6288.

Xiao H, Chenyang X, Prince JL. 2003. A topology preserving Level set method for geometric deformable models. IEEE Transactions on Pattern Anal Mach Intell. 25(6):755-768.

Xie X, Mirmehdi M, Maw R, A. H. 2005. Detecting Abnormalities in Tympanic Membrane Images. Medical Image Understanding an Analysis. Proceedings of the 9th Medical Image Undestanding and Analysis (MIUA). Bristol. UK

Xu C, Pham DL, Prince JL. 2000. Medical Image Segmentation Using Deformable Models. Handbook of Medical Imaging. 2:129-174

Xu C, Prince JL. 1998. Generalized gradient vector flow external forces for active contours. Signal Process. 71(2):131-139.

Yao W-J, Zhou H-C, Hu B-L, Huang X-S, Li X-Q. 2010. Research on Ossicular Chain Mechanics Model. Mathematical Problems in Engineering. 1-14.

Yoo KS, Wang G, Rubinstein JT, Vannier MW. 2001. Semiautomatic segmentation of the cochlea using real-time volume rendering and regional adaptive snake modeling. J Digit Imaging. 14(4):173-181.

Zadeh LA. 1965. Fuzzy sets. Information and Control. 8(3):338-353.

Zhang D-Q, Chen S-C. 2004. A novel kernelized Fuzzy C-means algorithm with application in medical image segmentation. Artif Intell Med. 32(1):37-50.

Zhang X, Gan RZ. 2011. A comprehensive model of human ear for analysis of implantable hearing devices. IEEE Trans Biomed Eng. 58(10):3024-3027.

Zhao F, Koike T, Wang J, Sienz H, Meredith R. 2009. Finite element analysis of the middle ear transfer functions and related pathologies. Med Eng Phys. 31(8):907-916.

Zitova B, Flusser J. 2003. Image registration methods: a survey. Image Vision Comput. 21(11):977-1000.

Zwislocki J. 1962. Analysis of the Middle-Ear Function. Part I: Input Impedance. J Acoust Soc Am. 34(9B):1514-1523.

FIGURE CAPTIONS

Figure 1: Illustration of the outer, middle and inner ear. The outer ear includes the auditory canal. The middle ear includes the tympanic membrane and three tiny bones for hearing. The bones are called the hammer (malleus), anvil (incus) and stirrup (stapes) to reflect their shapes. The inner ear (labyrinth) contains the semicircular canals, vestibule for balance, and the cochlea for hearing.

Figure 2: Geometric scheme of a multi-body model (adapted from (Volandri et al. 2012)).

Figure 3: A Finite Element Model built for the middle ear ossicles, eardrum, ligaments and muscles.

Figure 4: a) Original axial slice from a CT image, b) ROI image selected manually from the axial CT image showing the malleus, incus, vestibule, semicircular canal; c) ROI image selected manually from a coronal CT image showing the external auditory canal, malleus, facial nerve and cochlea.

Figure 5: Original slice from a MR image, b) ROI image selected manually from the MR image showing the cochlea, semicircular canal, vestibular nerve and facial nerve.

Figure 6: Segmentation results using the Otsu algorithm: on the left, in an axial CT image; on the right, in a coronal CT image.

Figure 7: Segmentation results using the Canny edge detector: on the left, in an axial CT image; on the right, in a coronal CT image.

Figure 8: Segmentation results using a region growing algorithm: on the left, in an axial CT image; on the right, in a coronal CT image.

Figure 9: Segmentation results using a watershed algorithm: on the left, in an axial CT image; on the right, in a coronal CT image.

Figure 10: Segmentation results using a fuzzy c-means algorithm: on the left, in an axial CT image; on the right, in a coronal CT image.

Figure 11: Segmentation results using a snake algorithm: on the left, in an axial CT image; on the right, in a coronal CT image.

Figure 12: Segmentation results using a Level set algorithm: on the left, in an axial CT image; on the right, in a coronal CT image.

TABLE CAPTIONS

Table 1: Segmentation methods and imaging techniques that have been used in ear anatomical structures studies.

FIGURES

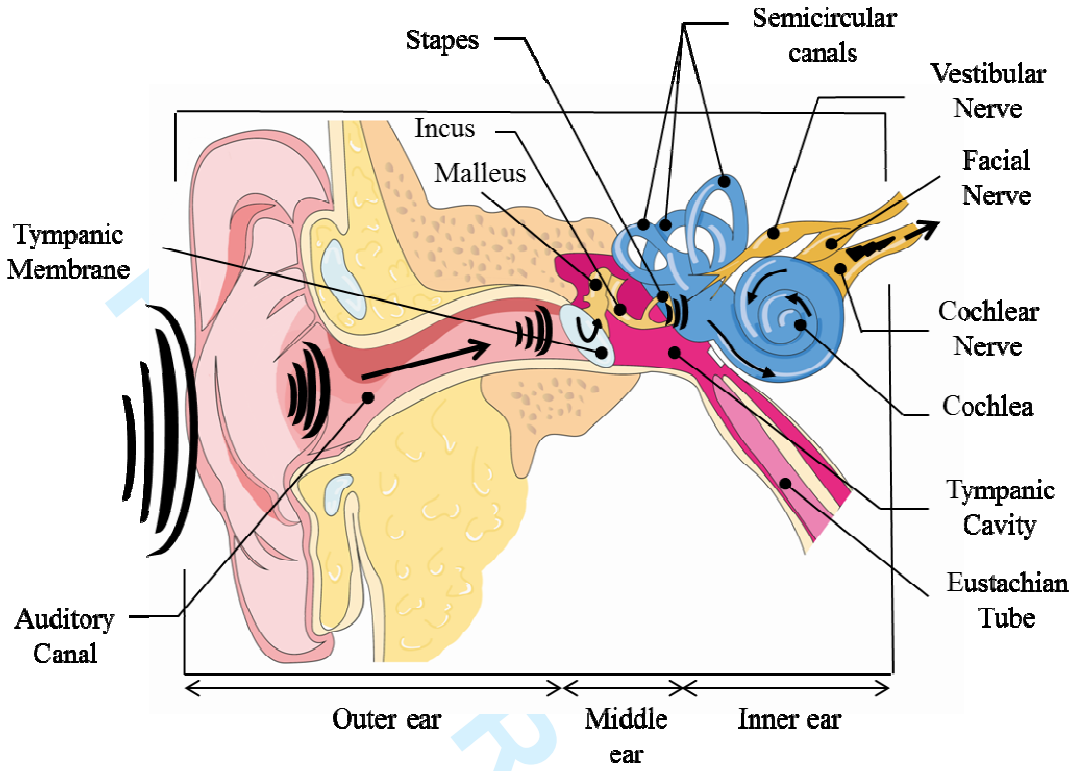

Figure 1

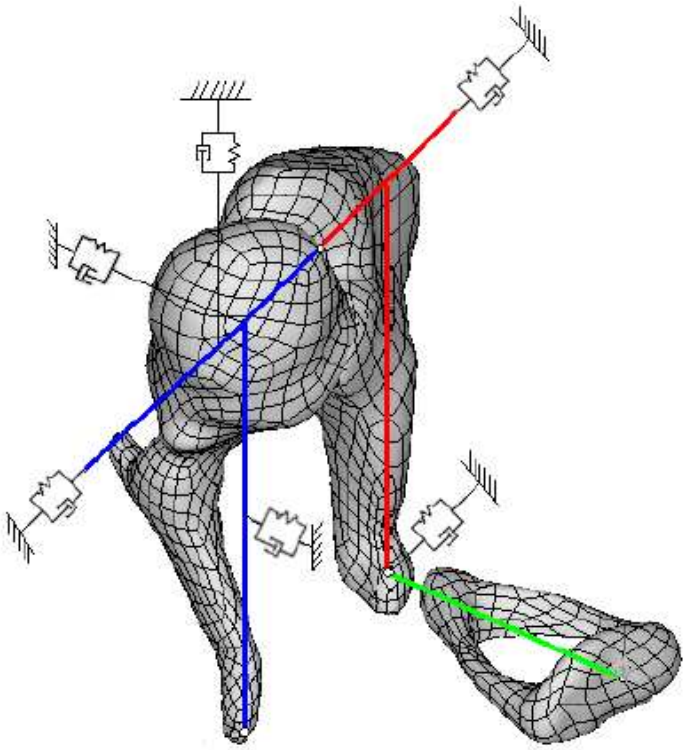

Figure 2

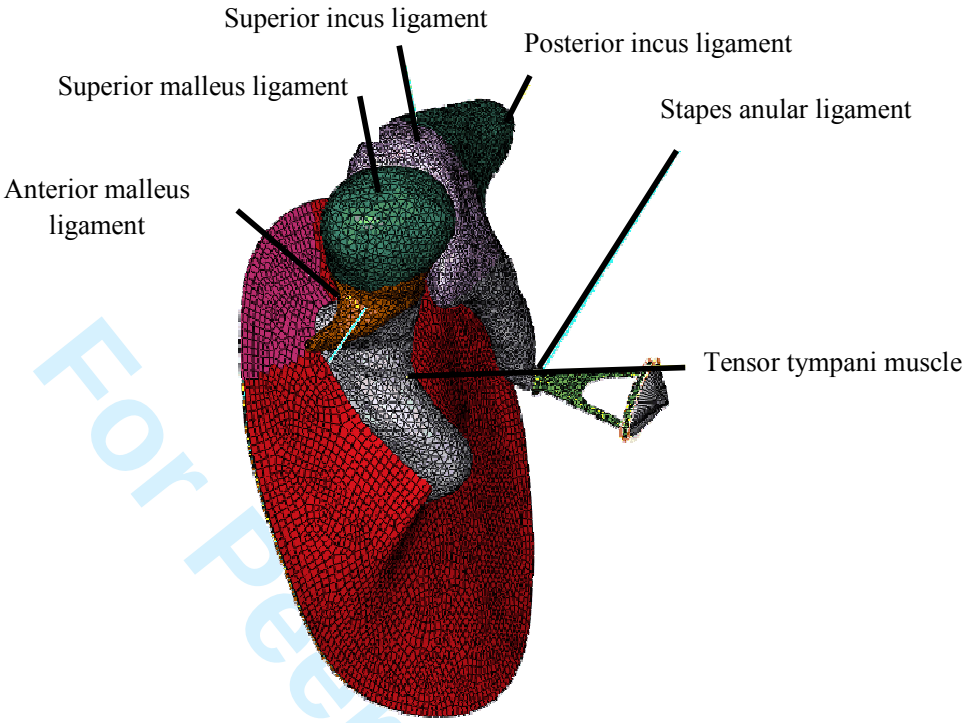

Figure 3

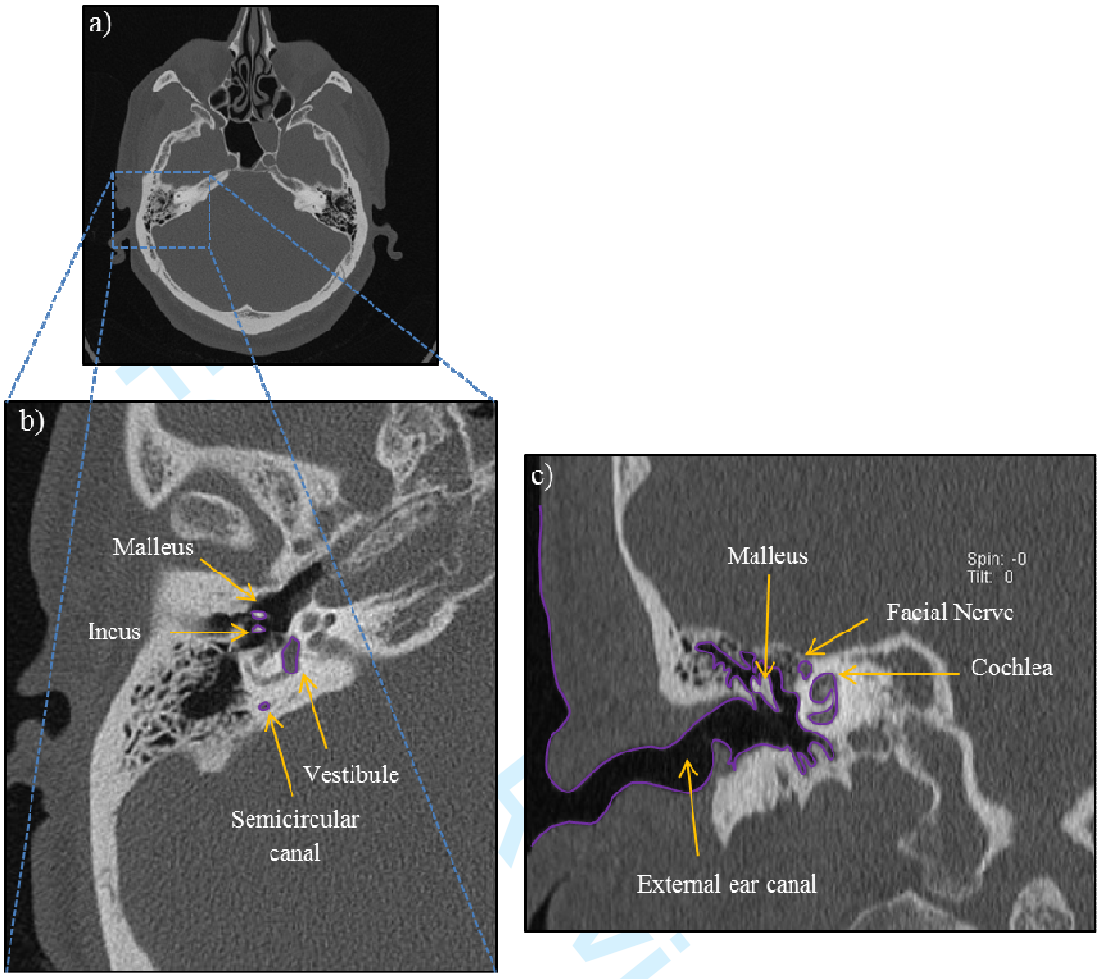

Figure 4

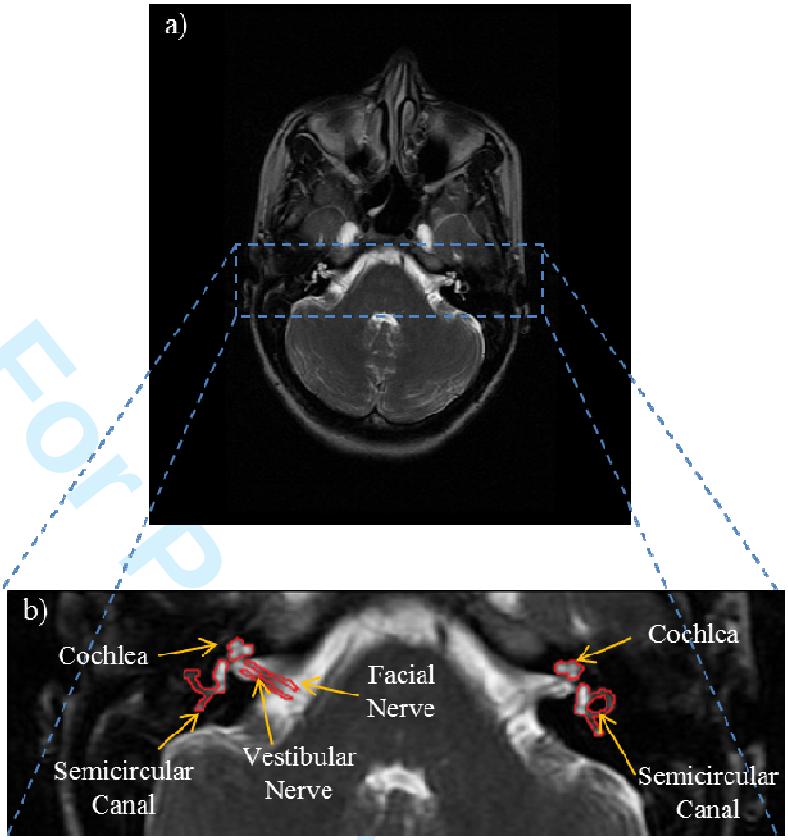

Figure 5

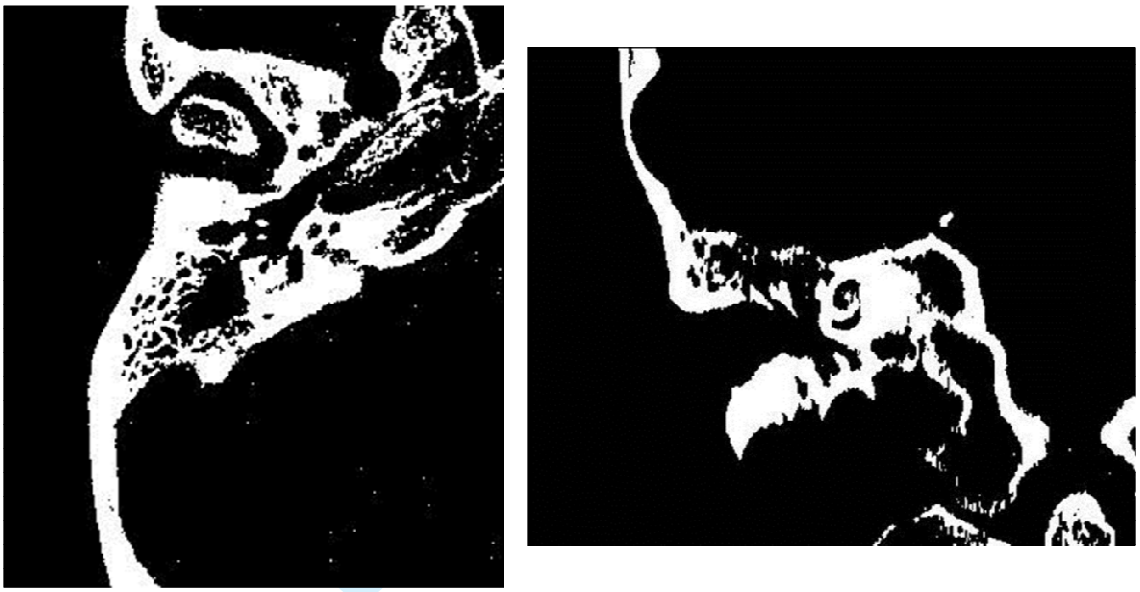

Figure 6

Figure 7

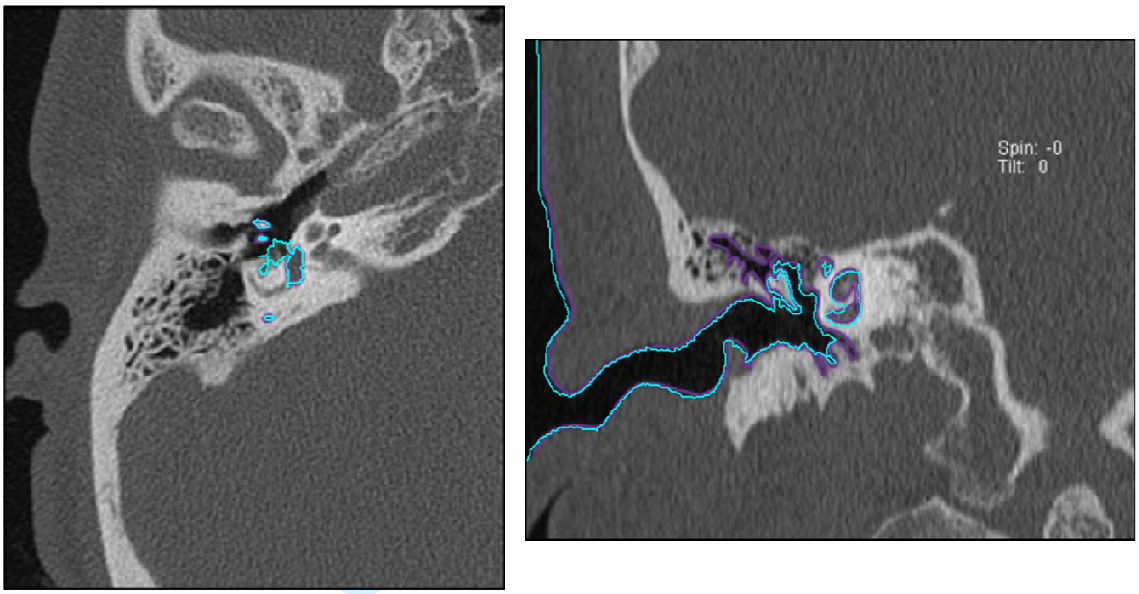

Figure 8

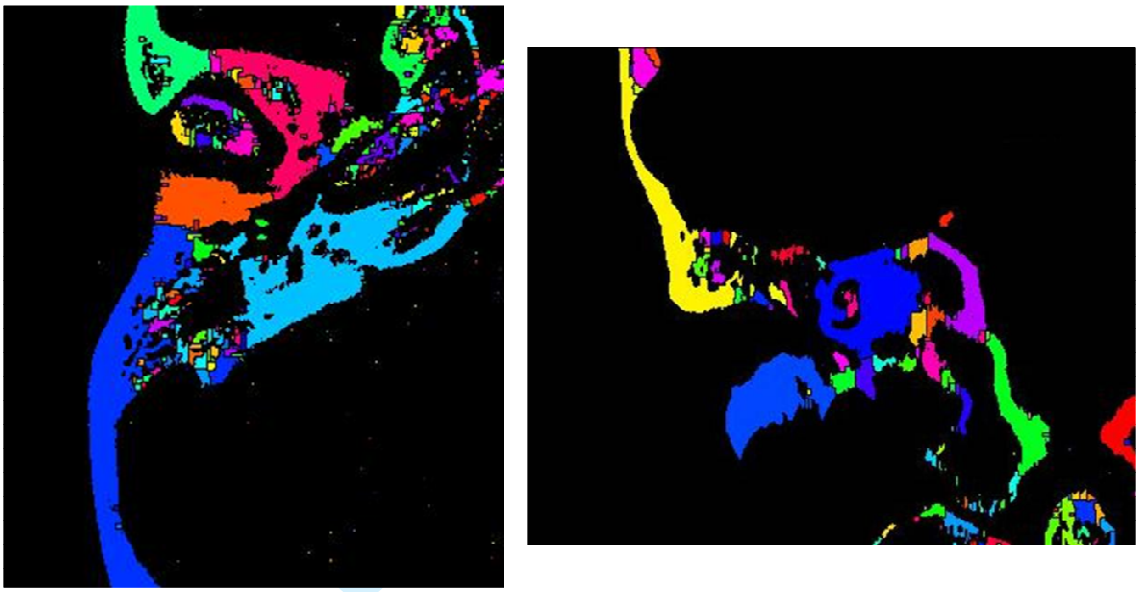

Figure 9

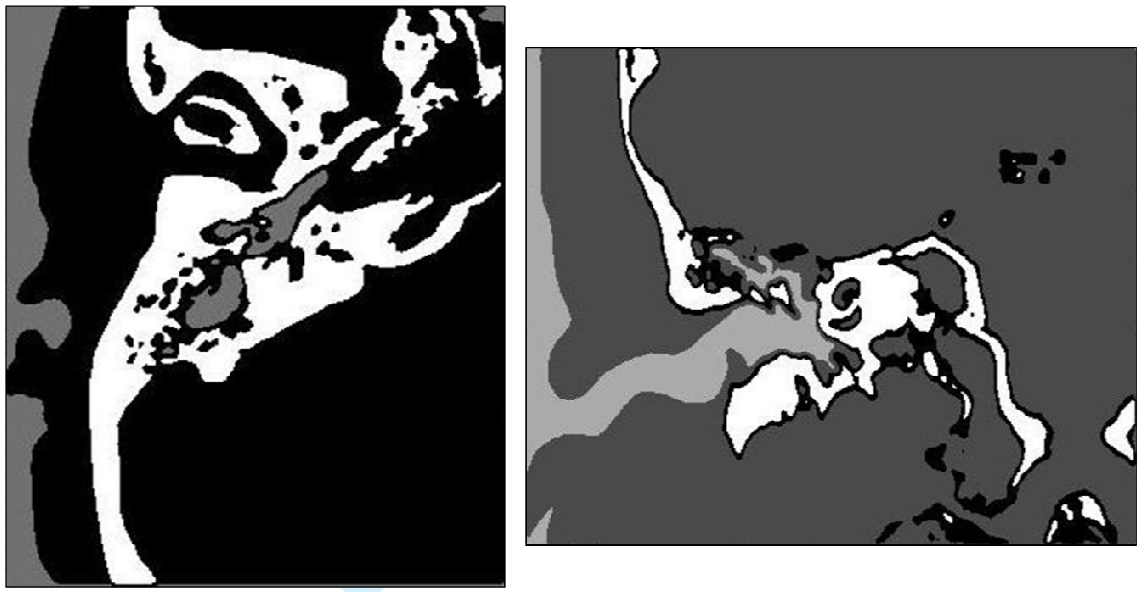

Figure 10

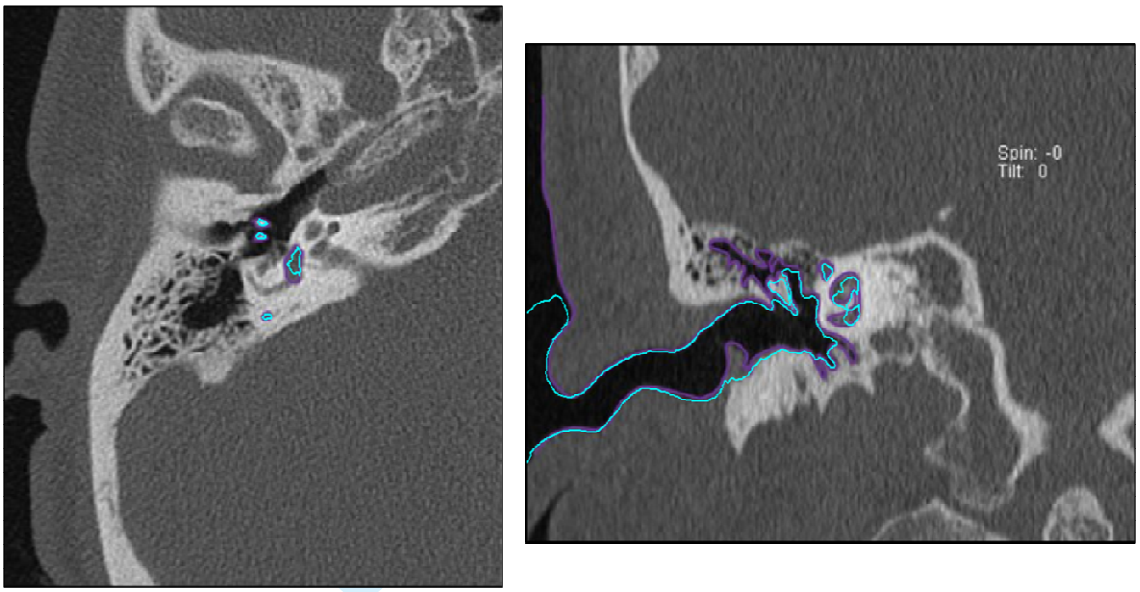

Figure 11

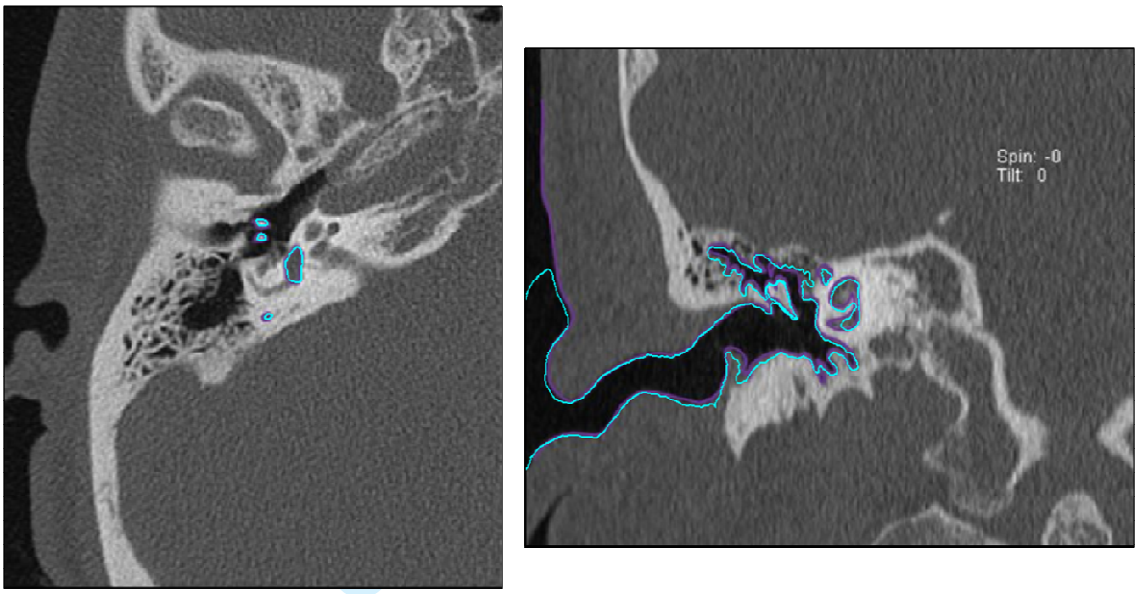

Figure 12

TABLES

Table 1

Author(s)	Title	Segmentation Method	Imaging Modality	Anatomical Structure(s)
Sim and Puria 2008	Soft tissue morphometry of the malleus-incus complex from micro-CT imaging	**Manual**	Micro-CT	**Middle Ear** Ossicles
Jun et al. 2005	Three-dimensional reconstruction based on images from spiral high-resolution computed tomography of the temporal bone: anatomy and clinical application		Spiral-CT	**Inner Ear**
Melhem et al. 1998	Inner ear volumetric measurements using high-resolution 3D T2-weighted fast spin-echo MR imaging: Initial experience in healthy subjects	**Thresholding** Global Thresholding	MR	**Inner Ear**
Lee et al. 2010	Reconstruction and exploration of virtual middle-ear models derived from micro-CT datasets		CT	**Middle Ear** Ossicles
Rodt et al. 2002	3D visualisation of the middle ear and adjacent structures using reconstructed multi-slice CT datasets, correlating 3D images and virtual endoscopy to the 2D cross-sectional images			
Seemann et al. 1999	Evaluation of the middle and inner ear structures: comparison of hybrid rendering, virtual endoscopy and axial 2D source images	**Thresholding** Region Growing	Spiral-CT	**Middle Ear** Ossicles
Todd et al. 2009	An analysis of medical image processing methods for segmentation of the inner ear			**External Ear** Auditory Canal **Inner Ear** Cochlea
Bradshaw et al. 2010	A Mathematical Model of Human Semicircular Canal Geometry: A New Basis for Interpreting Vestibular Physiology	**Thresholding** Watershed	CT	**Inner Ear** Semicircular Canals

Shi et al. 2010	Automatic MRI segmentation and morphoanatomy analysis of the vestibular system in adolescent idiopathic scoliosis	**Clustering**	MR	**Inner Ear** Vestibular System
Bradshaw et al. 2010	A Mathematical Model of Human Semicircular Canal Geometry: A New Basis for Interpreting Vestibular Physiology	**Deformable Models** Snake	CT	**Inner Ear** Semicircular Canals
Xie et al. 2005	Detecting Abnormalities in Tympanic Membrane Images. Medical Image Understanding an Analysis		Video-Otoscopy	**Outer Ear** Tympanic Membrane
Tabrizi 2003	Using Active Contours for Segmentation of Middle-Ear Images		MR	**Middle Ear**
Yoo et al.2001	Semiautomatic segmentation of the cochlea using real-time volume rendering and regional adaptive snake modeling		Spiral-CT	**Inner Ear** Cochlea
Poznyakovskiy et al. 2008	The creation of geometric three-dimensional models of the inner ear based on micro computer tomography data		Micro-CT	
Noble et al. 2011	Automatic segmentation of intracochlear anatomy in conventional CT			
Xianfen et al. 2005	3D semi-automatic segmentation of the cochlea and inner ear	**Deformable Models** Level Set	Spiral-CT	**Inner Ear** Cochlea Semicircular Canals
Comunello et al. 2009	A computational method for the semi-automated quantitative analysis of tympanic membrane perforations and tympanosclerosis		Video-Otoscopy	**Outer Ear** Tympanic Membrane
Tabrizi 2003	Using Active Contours for Segmentation of Middle-Ear Images		MR	**Middle Ear**
Noble et al. 2009 Noble et al. 2010	Automatic identification and 3D rendering of temporal bone anatomy	**Atlas**	CT	**Inner Ear** **Middle ear** Ossicles **Outer ear** Auditory canal

Chistensen et al. 2003	Automatic Measurement of the Labyrinth Using Image Registration and a Deformable Inner Ear Atlas			**Inner Ear** Cochlea Vestibule Semicircular canal **Outer Ear** Auditory canal