Computational algorithms for the segmentation of the human ear

E.M. Barroso, Zhen Ma & João Manuel R. S. Tavares

Faculdade de Engenharia da Universidade do Porto / Instituto de Engenharia Mecânica e Gestão Industrial, Porto, Portugal

Fernanda Gentil

Escola Superior de Tecnologia da Saúde do Porto / Clínica ORL – Dr. Eurico Almeida / IDMEC-Polo FEUP, Porto, Portugal

ABSTRACT: The main goal of this project is to identify an efficient segmentation algorithm for each anatomic structure of the ear. Therefore, in this paper, it is presented and analyzed computational algorithms that have been used to segment structures in images, especially of the human ear in Computed Tomography (CT) images.

1 INTRODUCTION

The organs of hearing and balance constitute the human auditory system, which can be divided into three main parts: external ear, middle ear and inner ear. The ear is by far the most complex organ of the human sensory system (Moller, 2006; Seeley, 2004).

A number of approaches have been presented for reconstructing and visualizing 3D models of the ear with diverse goals, such as with education purposes (Jun, 2005), to build customized biomechanical models (Deacraemer, 2003; Sim, 2008) and in the assessment of pre-operative procedures (Hussong, 2009; Rau, 2009). In order to attain these goals, different types of medical images have been used, acquired by several imaging techniques; for instance, Computerized Tomography (CT-standard, Micro-CT, Spiral-CT) (Christensen, 2003; Poznyakovskiy, 2008; Xianfen, 2005), Magnetic Resonance (MR-standard, Micro-MR) (Shi, 2010; Liu, 2007; Lane, 2005) and Histological processing (Liu, 2007). Using these imaging modalities, anatomical features of the ear have been studied in cats, guinea pigs, chinchillas and humans (Sim, 2008; Liu, 2007).

Solutions of image processing and analysis are essential to attain realistic geometric models for the anatomical structures of the ear. Particularly, the segmentation of the ear structures in images is crucial to build patient-customized biomechanical models to be successfully used in computational simulations. From this simulation, the understanding of the connections between the ear structures and their functions becomes easier as well as the optimization of prosthetic implants.

The study and optimization of cochlear implant systems can be an important application area of the realistic and accurate modeling of the ear. In fact, the position of the implanted electrodes has been identified as one of the most important variables in speech recognition, and the geometric modeling of the ear can facilitate the optimization of the electrode positions, which can be an important step towards efficient traumatic cochlear implant surgeries. Up to now, only manual insertion tools or insertion aids exist, providing the possibility to insert the electrode using a fixed insertion technique that is not adjustable to the patient (Hussong, 2010; Rau, 2010). Thus, based on accurate computational simulations the planning of surgical procedures can be enhanced (Tuck-Lee, 2008).

The biomechanical modeling of the ear also presents a key role in diagnosis and treatment of middle and inner ear diseases, because these two processes are hampered by the small size of the structures and by their hidden locations in the temporal bone (Seemann, 1999). In addition, through the computational modeling of the inner ear, anatomical abnormalities of the bony labyrinth can be easier identified. Therefore, it is possible to create templates that standardize the abnormal configurations (Melhem, 1998).

Image segmentation is a common task in Computational Vision and an important factor for the success of any efficient and accurate image analysis solution. Extracting the structures' contours in medical images, for example, by finding the image edges, can help doctors in detecting more efficient anomalies in visual inspections. However, the segmentation of structures in medical images is normally performed manually, requiring, for example, that medical technicians sketch the desired contours using pointing devices, such as a mouse or a trackball, which is very time-consuming and prone to errors. To overcome the disadvantages of manual segmentation, modern mathematical and physical techniques have been incorporated into the development of computational segmentation algorithms. These incorporations have greatly enhanced the accuracy of the segmentation results (Ma, 2010).

Segmentation is usually regarded as a task of image analysis and, based on the technique adopted, the current segmentation algorithms can be divided into three classes: thresholding, clustering and deformable models (Ma, 2010).

The objective of this work is to review image segmentation algorithms that have been used to segment the structures of the human ear. Hence, the identified algorithms will be analyzed, and their advantages and disadvantages will be pointed out, and some of their results will be presented and discussed.

The paper is organized as follows. In section 2, a review on the segmentation algorithms is made. Afterwards, the characteristics of the algorithms are illustrated through experiments on ear images. In section 4, the advantages and disadvantages of each type are summarized. In the last section, the conclusions are presented.

2 SEGMENTATION ALGORITHMS

In this section, three classes of segmentation algorithms are reviewed. Hence, the common features of each class are identified and discussed, and their advantages and disadvantages are summarized. Additionally, the applications of algorithms of each class on ear images are illustrated to further depict their main characteristics.

2.1 *Algorithms based on Thresholding*

Thresholding is a common region segmentation method. In this technique, threshold values are defined, and the original image is divided into groups of pixels that have values within the ranges defined by the thresholds and groups of pixels with values beyond the ranges (Bankman, 2000). Thresholding is a simple, yet often an effective means to segment images in which the represented structures have distinct intensity levels, or other quantifiable feature. The algorithms are usually performed interactively, based on the visual assessment of the resulting segmentations (Pham, 2000).

There are several thresholding techniques, some of them are based on the image histogram, and others are based on local properties or local gradient.

The global thresholding is the most intuitive approach, and is termed “global” as just one threshold value is selected for the entire image, which is defined based on the image histogram. Instead, when the threshold value depends on local properties, it is classified as “local”. Local thresholding algorithms can be further classified as edge-based, region-based or hybrid. The fundamental goal of edge detection algorithms is to identify the borders of the structures in an image that can be then used to extract features like corners, lines or curves. Canny, Sobel and Laplacian operators are examples of edge detectors (Ma, 2010).

The region based algorithms are another type of thresholding-based algorithms and their idea comes from the observation that quantifiable features inside a structure tend to be homogeneous (Ma, 2010). Therefore, these algorithms aim to search for the image pixels with similar feature values. Common examples of this type of algorithms are the region growing and the split and merge algorithms (Bankman, 2000).

Finally, hybrid algorithms combine different image cues to complete the segmentation and a typical example is the watershed algorithm (Ma, 2010).

Thresholding-based algorithms have been used to segment anatomic structures of the middle and inner ear (ossicles, cochlea, bone labyrinth), in Micro-CT, Magnetic Resonance and Spiral-CT images (Lee, 2010; Rodt, 2002; Xianfen, 2005; Melhem, 1998).

2.2 *Algorithms based on Clustering*

Pattern recognition techniques can be used to perform the segmentation of structures in images, and clustering techniques have been very commonly applied in medical image segmentation (Ma, 2010). Clustering techniques have the same goal as classifier methods that seek to partition the feature space derived from the original image (Pham, 2000). These techniques can be divided into three main classes: supervised, unsupervised and semi-supervised (Sutton, 2000; Ma, 2010).

The supervised techniques are frequently used, and they include k-nearest neighbor (KNN) and maximum likelihood (ML) algorithms, supervised artificial neural networks (ANN), support vector machines (SVM), active shape models (ASM) and active appearance models (AAM) (Ma, 2010). A training data set is needed to perform supervised classification in order to extract the structure information, and the key issue of supervised clustering is the guidance provided by the labeled data (Zhu, 2010).

Unsupervised classification techniques extract the features of the structure to be segmented from the classified points, and examples of such techniques include fuzzy C-means (FCM) and iterative self-organizing data analysis technique (ISODATA) algorithms and unsupervised neural networks (Ma, 2010). As already referred, the unsupervised methods explore the intrinsic structure data to segment the input image into regions with different statistics. However, these methods often fail to achieve the desired results, especially if the wanted segmentations include regions with very dissimilar characteristics. On the other hand, supervised image segmentation methods first build the classifier from a labeled training set. Although these methods are likely to perform better, labeling the training set is usually very time consuming.

Semi-supervised image segmentation is the last type of clustering technique and the methods overcome the problems of the traditional clustering techniques by inferring the segmentation from partially labeled images. For example, in Figueiredo (2007) a simple fully deterministic generalized expectation-maximization algorithm (GEM) is described, which is a semi-supervised mixture-based clustering algorithm.

To conclude, the semi-supervised clustering is the recent type of clustering and takes advantage of the user's labels to attain the segmentation, while minimizing the labeling process; and this algorithm has been applied to problems of symptoms classification in medical image databases with promising results (Figueiredo, 2007).

Clustering technique was used by Shi (2010) in magnetic resonance images to segment the vestibular system. However, to overcome the limitations of the clustering techniques in the segmentation of the ear structures, the clustering algorithm was combined with a deformable model.

2.3 *Algorithms based on Deformable models*

Algorithms based on deformable models are more flexible when compared to the types previously described, and have been successfully used in complex segmentations (Ma, 2010). According to the way that is used for tracking the moving contours, deformable models can be further classified into parametric or geometric models.

Parametric deformable models represent curves and surfaces explicitly in their parametric forms during the model deformation. This representation allows direct interaction with the model and can lead to a compact representation for fast real-time implementation. However, the adaptation of the model to new topologies, such in splitting or merging, during the model deformation, can hamper the use of parametric models. Parametric models include active contours (or snakes), active contours with statistical techniques integrated and generalized gradient vector flow snakes (GGVF).

On the other hand, geometric deformable models can handle topological changes more naturally. These models, based on the theory of curve evolution and the level set method, represent curves and surfaces implicitly as a level set of a higher-dimensional scalar function (McInerney, 1996). The geometric models include level set algorithms, Malladi's algorithm (geodesic active contour – GAC) and Chan-Vese model.

Deformable models have been widely used in several Computational Vision applications, including the segmentation of the external and inner ear. Xie (2005) and Comunello (2009) applied the generalized gradient vector flow snake and active contours segmentation method of Mumford-Shah, respectively, to segment the tympanic membrane of the external ear. Xianfen (2005), Poznyakovskiy (2008) and Bradshaw (2010) used active contours and level set algorithms to segment the cochlea and the semi-circular canals of the inner ear.

3 RESULTS

In this section, it is described the pre-processing technique that was used to remove the noise in the original images and to decrease the computational cost, as well as the results of the tested segmentation algorithms.

3.1 *Pre-processing*

To enhance and smooth the original CT images of the temporal bone, an anisotropic diffusion Gaussian filter was used. The anisotropic diffusion filter blurs areas of low contrast and enhances the areas of high contrast (edges), as it works as a high pass filter. Thus, a band-pass filter, which initially performs a Gaussian filtering and then applies an anisotropic diffusion filter, was used to remove the noise presented in the input image, Figure 1.

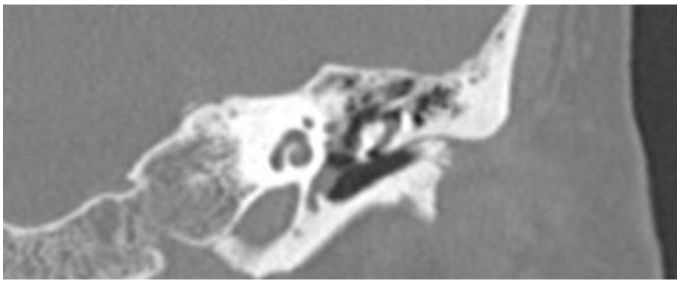

Figure 1. Computerized tomography image of a temporal bone

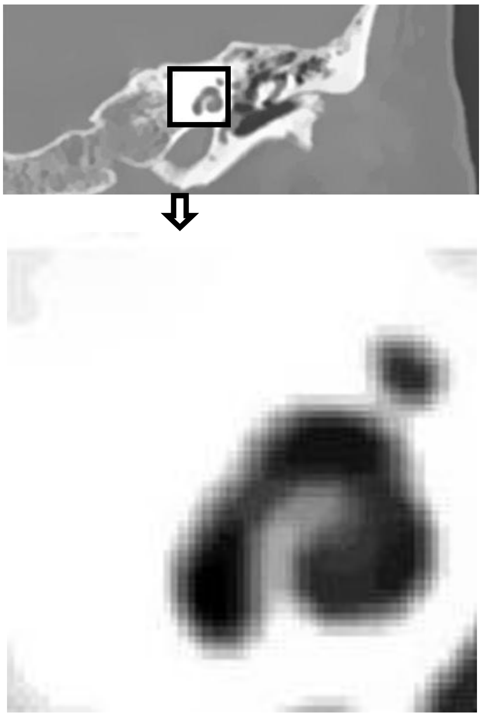

Figure 2. Selection of the ROI from the filtered image

To reduce the required computational time and also the computational cost, the region of the inner ear, i.e. region of interest (ROI), was first selected from the filtered image by searching for the pixels with highest intensity values, Figure 2. As illustrated in Figure 2, the membranous labyrinth of the inner ear is fully enclosed by the resultant ROI. Then, from the membranous labyrinth, it is possible to create a region that represents the boundary of the temporal bone. By this way, the image size is reduced and the structures of the interest are fully inside the ROI.

3.2 *Segmentation*

Figure 1 presents a CT image of the temporal bone, from which the appreciable influences of partial volume effects (PVE) can be easily seen. The segmentation results of Otsu method, Canny edge detector, a region growing and a watershed algorithms are illustrated in Figures 3(a)-(d).

The boundaries obtained by the Canny edge detector are continuous, but this algorithm usually presents edges discontinuity due to the noises and PVE. Besides, the spatial relationships of the edge points are not reflected; as such, most of the detected boundaries are incomplete or wrongly connected. For the region growing algorithm, the boundaries of the cochlea and semicircular canal are well segmented. The watershed algorithm gives a complete segmentation of the image. However, over segmentation can be seen in the area between the cochlea and the temporal bone, because there are a lot of pixels with local maximum of gradient magnitude. Finally, the area of the segmented objects by the Otsu method is excessively large.

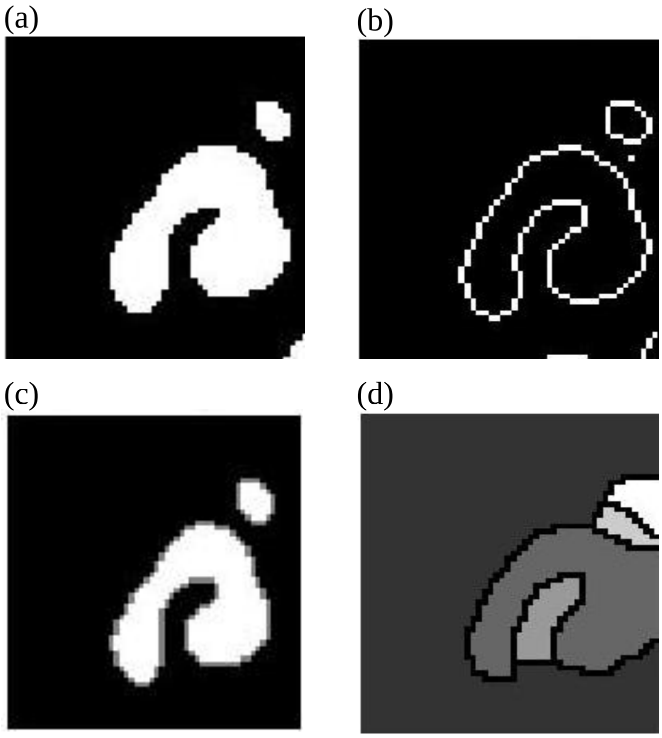

Figure 3. Results of Otsu method (a); Canny edge detection algorithm (b); region growing algorithm (c) and watershed algorithm (d) when is used the region of interest obtained from figure 1, which is illustrated in figure 2.

Figure 4 illustrates the segmentation results of the snake algorithm, Chan-Vese's model and algorithm proposed by Li & Xu.

Parametric deformable models are widely used in structure segmentation and 3D reconstruction. However, the computational complexity such as parameterization of the contours, handling of topological changes, and re-distribution of the contour points, considerably restricts their applications. The snake algorithm, Figure 4(b), has poor convergence to boundaries with larger curvatures. Additionally, the algorithm performance has a high dependence on the initial contour.

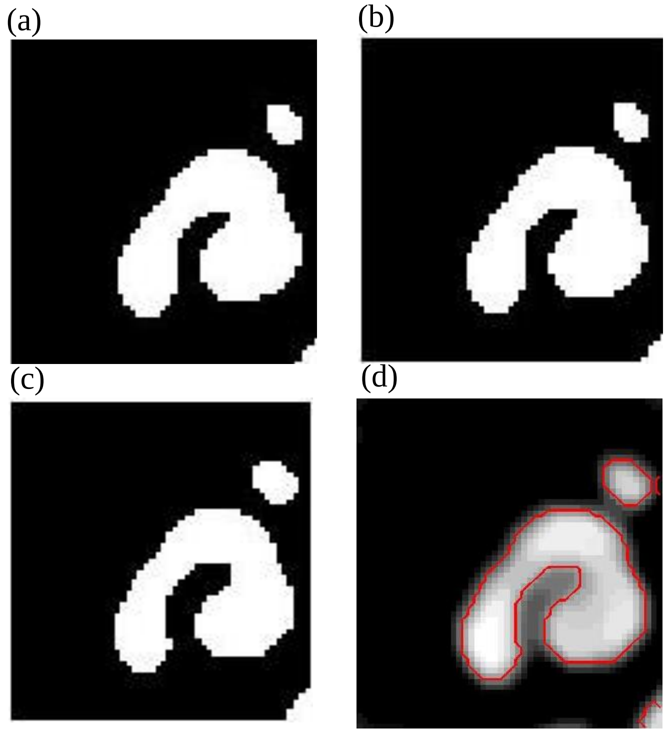

Figure 4. Illustration of the mask, which is obtained by the thresholding algorithm, (a) used to perform the snake algorithm (b), the results of the Chan Vese's model (c) and of the level set of Li and Xu (d).

Figures 4(c) and 4(d) present the segmentation results of the Chan-Vese model and the algorithm proposed in the level set of Li and Xu. When these two results are compared with the results of the snake algorithm, Figure 4(b), one can see that the boundaries of the last two (deformable models) are more regular and less influenced by noise. The regulating effects of internal forces make the boundary shape more reasonable and less influenced by noise.

4 DISCUSSION

Thresholding is a simple, yet frequently effective means to segment images in which different structures have distinct intensities, or other quantifiable features; it is normally performed interactively, based on the operator's visual assessment of the resulting segmentation (Pham, 2000). Moreover, thresholding is often used as an initial step in a sequence of image processing operations. Its main limitations are that its performance is sensitive to the influence of noise and intensity inhomogeneity, and it typically does not take into account the spatial characteristics of the structure to be segmented. In addition, only two classes are generated by using the simplest form of thresholding. Also, artifacts may corrupt the histogram of the image and make the segmentation more difficult (Bankman, 2000).

Usually, the level set methods have slower speed of convergence than the parametric deformable models, due to their computational complexity, and parametric deformable models are sensitive to the initial conditions. On the other hand, geometric deformable models can automatically handle topology changes and allow multiple simultaneous boundary identifications. Specifically, algorithms based on geometric deformable models aim to eliminate noise influence, prevent leakage, enhance accuracy and efficiency, and make the algorithms more automatic and less dependent on the initial conditions (Tsai, 2001; Wang, 2007).

In conclusion, deformable models are promising to segment medical images, because these models can easily incorporate statistical information and other techniques, and the segmented boundaries have regular geometric properties. This way, they can provide contours with regular geometric properties.

5 CONCLUSIONS

In this paper, current segmentation algorithms were classified into three types and their respective characteristics were summarized. Applications of some of the present algorithms to segment ear structures were illustrated. The experimental examples were also used to further state the distinct characteristics of each type of algorithms.

Deformable models seem to be promising for the segmentation of the inner ear because they can easily incorporate statistical information in order to improve their performance/efficiency/effectiveness.

ACKNOWLEDGEMENTS

This work was partially done in the scope of the projects "Methodologies to Analyze Organs from Complex Medical Images – Applications to Female Pelvic Cavity", "Aberrant Crypt Foci and Human Colorectal Polyps: mathematical modeling and endoscopic image processing", "Cardiovascular Imaging Modeling and Simulation - SIMCARD" and "Bio-computational study of tinnitus", with references PTDC/EEA-CRO/103320/2008, UTAustin/MAT/0009/2008, UTAustin/CA/0047/2008 and PTDC/SAU-BEB/104992/2008, respectively, financially supported by Fundação para a Ciência e a Tecnologia (FCT), in Portugal.

The second author would like to thank FCT for his PhD grant with reference SFRH/BD/43768/2008.

6 REFERENCES

Bankman, Isaac N.(2000). *Handbook of Medical Imaging Processing and Analysis*. San Diego: Academic Press, Reprint.

Bradshaw, Adrew P., Ian S. Curthoys, Michael J. Todd, John S. Magnussen, David S. Taubman, Swee T. Aw, and G. Michael Halmagyi. (2010). A Mathematical Model of Human Semicircular Canal Geometry: A New Basis for Interpreting Vestibular Physiology. *Journal of the Association for Research in Otolaryngology* 11: 145-59.

Christensen, Gray E., Jianchun He, John A. Dill, Jay T. Rubinstein, Michael W. Vannier, and Ge Wang. (2003). Automatic Measurement of the Labyrinth Using Image Registration and a Deformable Inner Ear Atlas. *Academic Radiology Journal* 10: 988-99.

Comunello, Eros, Aldo von Wangenheim, Vilson Heck Junior, Cristina Dornelles, and Sady Selamen Costa. (2009). A Computational Method for the Semi-Automated Quantitative Analysis of Tympanic Membrane Perforations and Tympanosclerosis. *Computers in Biology and Medicine* 39: 889-95.

Decraemer, W. F., J. J. J. Dirckx, and W. R. J. Funnell. (2003). Three-Dimensional Modelling of the Middle-Ear Ossicular Chain Using a Commercial High-Resolution X-Ray Ct Scanner. *Journal of the Association for Research in Otolaryngology* 4: 250-63.

Figueiredo, Mário A. T. 2007. Semi-Supervised Clustering: Application to Image Segmentation. In *Advances in Data Analysis* edited by Reinhold Decker and Hans -J. Lenz, 39-50. Berlin: Springer Berlin Heidelberg.

Hussong, Andreas, Thomas S. Rau, Tobias Ortmaier, Bodo Heimann, Thomas Lenarz, and Omid Majadani. (2009). An Automated Insertion Tool for Cochlear Implants: Another Step Towards Atraumetic Cochlear Implant Surgery. *Int Journal CARS* 5: 163-71.

Jun, Beom-Cho, Sun-Wha Song, Ju-Eun Cho, Chan-Soon Park, Dong-Hee Lee, Ki-Hong Chang, and Sang-Won Yeo. (2005). Three-Dimensional Reconstruction Based on Images from Spiral High-Resolution Computed Tomography of the Temporal Bone: Anatomy and Clinical Application. *The Journal of Laryngology & Otology* 119: 693-98.

Lane, John I., Robert J. Witte, Odell W. Henson, Colin L. W. Driscoll, John Camp, and Richard A. Robb. (2005). Imaging Microscopy of the Middle and Inner Ear Part Ii: Mr Microscopy. *Clinical Anatomy Wiley-Liss* 18: 409-15.

Lee, Dong H., Sonny Chan, Curt Salisbury, Namkeum Kim, Kenneth Salisbury, Sunil Puria, and Nikolas H. Blevins. (2010). Reconstruction and Exploration of Virtual Middle-Ear Models Derived from Micro-Ct Datasets. *Hearing Research* 263: 198-203.

Liu, Bo, Xiu L. Gao, Hong X. Yin, Shu Q. Luo, and Jing Lu. (2007). A Detailed 3d Model of the Guinea Pig Cochlea. *Brain Struct Funct* 212: 212-30.

Ma, Zhen, João Manuel R. S. Tavares, Renato Natal Jorge, and T. Mascarenhas. (2010). A Review of Algorithms for Medical Image Segmentation and Their Applications to the Female Pelvic Cavity. *Computer Methods in Biomechanics and Biomedical Engineering* 13: 235-46.

McInerney, Tim, and Demetri Terzopoulos. (1996). Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis* 2: 91-108.

Melhem, Elias R., Huzeifa Shakir, Sivi Bakthavachalam, C. Bruce MacDonald, John Gira, Shelton D. Caruthers, and Hernan Jara. (1998). Inner Ear Volumetric Measurements

Using High-Resolution 3d T2-Weighted Fast Spin-Echo Mr Imaging: Initial Experience in Healthy Subjects. *American Journal Of Neuroradiol* 19: 1819-22.

Moller, A. R.2006. *Hearing: Anatomy, Physiology, and Disorders of the Auditory System, Second Edition*. San Diego: Academic Press, Reprint.

Pham, Dzung L., Chenyang Xu, and Jerry L. Prince. (2000). Current Methods in Medical Image Segmentation. *Annu. Rev. Biomed. Eng.* 2: 315-37.

Poznyakovskiy, Anton A., Thomas Zahnert, Yannis Kalaidzidis, Rolf Schmidt, Bjorn Fischer, Johannes Baumgart, and Yury M. Yarin. (2008). The Creation of Geometric Three-Dimensional Models of the Inner Ear Based on Micro Computer Tomography Data. *Hearing Research* 243: 95-104.

Rau, Thomas S., Andreas Hussong, Martin Leinung, Thomas Lenarz, and Omid Majdani. (2010). Automated Insertion of Performed Cochlear Implant Electrodes: Evaluation of Curling Behaviour and Insertion Forces on an Artificial Cochlear Model. *Int Journal CARS* 5: 173-81.

Rodt, T., P. Ratiu, H. Becker, S. Bartling, D. F. Kacher, M. Anderson, F. A. Jolesz, and R. Kikinis. (2002). 3d Visualisation of the Middle Ear and Adjacent Structures Using Reconstructed Multi-Slice Ct Datasets, Correlating 3d Images and Virtual Endoscopy Images. *Neuroradiology* 44: 783-90.

Seeley, Stephens, and Tate. (2004). *The Special Senses*, Anatomy and Physiology, Sixth Edition: The MacGraw-Hill Companies, Reprint.

Seemann, M. D., O. Seemann, H. Bonél, M. Suckfull, K.-H. Englmeier, A. Naumann, C. M. Allen, and M. F. Reiser. (1999). Evaluation of the Middle and Inner Ear Structures: Comparison of Hybrid Rendering, Virtual Endoscopy and Axial 2d Source Images. *European Radiology* 9: 1851-58.

Shi, Lin, Defeng Wang, Winnie C. W. Chu, Geoffrey R. Burwell, Tien-Tsin Wong, Pheng Ann Heng, and Jack C. Y. Cheng. (2010). Automatic Mri Segmentation and Morphoanatomy Analysis of the Vestibular System in Adolescent Idiopathic Scoliosis. *NeuroImage*: 9.

Sim, Jae Hoon, and Sunil Puria. (2008). Soft Tissue Morphometry of the Malleus-Incus Complex from Micro-Ct Imaging. *Journal of the Association for Research in Otolaryngology* 9: 5-21.

Sutton, Melanie A., James C. Bezdek, and Tobias C. Cahoon. 2000. Image Segmentation by Fuzzy Clustering: Methods and Issues. In *Handbook of Medical Imaging Processing and Analysis*. San Diego: Academic Press Series in Biomedical Engineering.

Tsai, Anthony Yezzi Andy, and Alan S. Willsky. (2001). Curve Evolution Implementation of the Mumford-Shah Functional for Image Segmentation, Denoising, Interpolation, and Magnification. *IEEE transactions on Image Processing* 10: 1169-86.

Tuck-Lee, James P., Peter M. Pinsky, Charles R. Steele, and Sunil Puria. (2008). Finite Element Modeling of Acoustical-Mechanical Coupling in the Cat Middle Ear. *Journal Acoustical Society of America* 124: 348-62.

Wang, Yonggang, Yun Zhu, and Qiang Guo. (2007). Medical Image Segmentation Based on Deformable Models and Its Applications. In *Deformable Models Theory and Biomaterial Applications*, 209-60: Springer New York.

Xianfen, Diao, Chen Siping, Liang Changhong, and Wang Yuanmei. 2005. 3d Semi-Automatic Segmentation of the Cochlea and Inner Ear. In *Engineering in Medicine and Biology 27th Annual Conference*, edited by IEEE. Shangai, China.

Xie, Xianghua, Majid Mirmehdi, Richard Maw, and Amanda Hall. (2005). Detecting Abnormalities in Tympanic Membrane Images. *Medical Image Understanding and Analysis*: 19–22.

Zhu, Yanping. (2010). An Efficient Supervised Clustering Algorithm Based on Neural Networks. In *3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, edited by IEEE, 265-68.