Eurographics Workshop on 3D Object Retrieval (2018)
A. Telea, T. Theoharis, and R. C. Veltkamp (Editors)

Automatic extraction of complex 3D structures Application to the inner ear segmentation from Cone Beam CT digital volumes

Florian Beguet[1,2,3], Jean-Luc Mari[1], Thierry Cresson[3], Matthieu Schmittbuhl[3,4], Jacques A. de Guise[2,3]

[1]Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France
[2]Ecole de Technologie Supérieure, Montréal, Canada
[3]Laboratoire de recherche en Imagerie et Orthopédie (LIO), CRCHUM, Montréal QC, Canada
[4]Faculté de médecine dentaire, Université de Montréal, Canada

Abstract

We present an automatic approach for the retrieval of a complex structure within a 3D digital volume, using a generic deformable surface model. We apply this approach to the inner ear reconstruction of Cone Beam CT(CBCT) 3D data. The proposed method is based on a single prior shape initialization followed by two steps. A geometric rigid adjustment allows a close fit to inner ear boundaries. Finally, a Laplacian mesh deformation method is used to iteratively refine the mesh. Preliminary results are promising in terms of several similarity metrics.

CCS Concepts

•***Computing methodologies*** → *Complex structure segmentation;* •***Hardware*** → *Cone Beam CT; CGAL;*

1. Introduction

1.1. Context

Automatic segmentation of complex topological structures from their volume representations is an important step towards a better understanding of their morphology and deformations. This research field finds many applications in medical imaging and geology [LTRRB08].

In this paper, we will focus on the extraction of three-dimensional models of the human inner ear from Cone Beam CT volumes. The inner ear is a labyrinthine space responsible for balance and hearing, and is composed of a vestibule, a cochlea and semicircular canals (Figure 1). This structure has a typical length and width of 20 mm and 13 mm, respectively. Anatomical knowledge of the inner ear is crucial for diverse research fields *e.g* cochlea implantation [RTP*16] and congenital malformation. Automatic retrieval of the inner ear boundaries through 3D modeling is a precondition for shape analysis and diagnosis. Three-dimensional models of this anatomical structure could also lead to a better understanding of hearing mechanics.

Multi-slice computed tomography (MSCT) is widely used for temporal bone imaging. However, Cone beam computerized tomography (CBCT) presents the advantage of having a shorter acquisition time and a higher spatial resolution while exposing patients to less ionizing radiation than conventional CT-scans [RTP*16]. As a result, this imaging modality is being increasingly used for this purpose. Nonetheless, the image contrast is lower due to the beam hardening phenomenon and scatter radiations causing artifacts and increased noise in CBCT images [Yan16], thus preventing soft tissue visualization.

1.2. Problem statement

One of the main challenges to inner ear retrieval in CBCT volumes is the impossibility to automatically suppress the surrounding bone structures visible in Figure 2.A. These structures have similar intensity distributions in CBCT images, which causes difficulties to detect their boundaries. Also, the shape of the inner ear is topologically complex : the cochlea is a snail-shaped object of $2\frac{1}{2}/2\frac{1}{3}$ turns and the semi-circular canals correspond to 3D objects characterized by surface of genus 3 (Figure 1). Conservation of these complex features during the segmentation process of the inner ear remains a difficult challenge.

1.3. Related work

Manual segmentation is often used as a way to obtain accurate 3D models of the inner ear in spite of the fact that it remains a laborious and time consuming process [VA95, BBS12]. To alleviate drawbacks of manual contouring, some research teams have proposed semi-automated method based on region growing [YWR*00], active contours [YWRV01] or level-sets [XSCY05] on CT-scan images. Simple semi-automatic algorithms implemented in some commercial softwares are also used in the literature [WNB*06, WZYL15].

By defining priors such as foreground and background or by placing seeds inside the desired region, these techniques solved the major leaking problems but the lack of shape constraints still makes them prone to leakage into neighboring. These techniques require the supervision of a user to achieve satisfactory precision. Therefore, the resulting models are often incomplete, and while that might be sufficient for ordinary shape analysis and length measurements, it remains inadequate for deeper analyses.

In order to further constrain the segmentation result and prevent leakage problems, *Noble et al.* [Nob11] proposed to impose global shape control based on an active shape model algorithm. The shape model was based on five models built upon manually segmented cochleae on micro-CT images. Their algorithm has the advantage of provide a complete segmentation of the bony labyrinth with a

DOI: 10.2312/3dor.20181048

differentiation between the scalaes (chambers inside cochlea), even though this differentiation is often not visible in CT-scan images.

1.4. Contribution

In previous works, authors mentioned the difficulties associated with the use of low-resolution imaging approaches to accurately detect inner ear structures. User supervision is usually required to obtain an acceptable 3D model. Our contribution consists of an accurate 3D segmentation of the inner ear from CBCT images *without supervision*. This method retains specific topological attributes of the inner ear using an iterative Laplacian optimization algorithm to apply a deformation on an *a priori* mesh. It is inspired from a liver segmentation approach, developed by Chartrand *et al.* [CCC*16]. Our preliminary results include a preliminary validation of this approach.

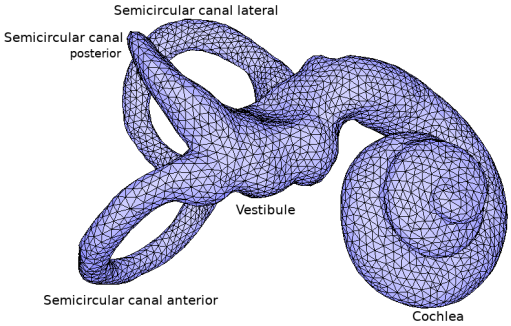

Figure 1: *A priori model of left inner ear.*

2. Description of the segmentation approach

2.1. Overview

In this approach, an *a priori* 3D model is used as input in a shape fitting algorithm. As shown in figure 2, we took a three-pronged approach to the segmentation process: (1) automatic initialization, (2) geometric adjustment, and (3) accurate segmentation. The first step is designed to initialize the prior shape inside the dataset, based on *a priori* knowledge. Then a rigid geometric adjustment is carried out for optimal mesh positioning closed to the boundaries of the structure of interest. Finally, an iterative deformation using a Laplacian mesh optimization method is performed to obtain an accurate three-dimensional geometric mesh of the patient's inner ear.

2.2. Generic 3D model of the inner ear

The *a priori* 3D model used in this approach was obtained from magnetic resonance images (MRI) of a human cadaver ear from Drs. O.W. Henson, Jr. and Miriam Henson (University of North Carolina at Chapel Hill.) Segmented images are available on their website `http://cbaweb2.med.unc.edu/henson_mrm/` and the mesh reconstruction has been performed by W. Robert J. Funnell (McGill University) and is available in VRML format on this website `http://audilab.bmed.mcgill.ca/~daren/3Dear/index.html`. The model shown in Figure 1 has been modified to have a homogeneous point distribution. This reduces the calculation time of the deformation method and prevent redundancies during the vertex matching step. To this end, an isotropic surface remeshing process [AVDI02] has been applied.

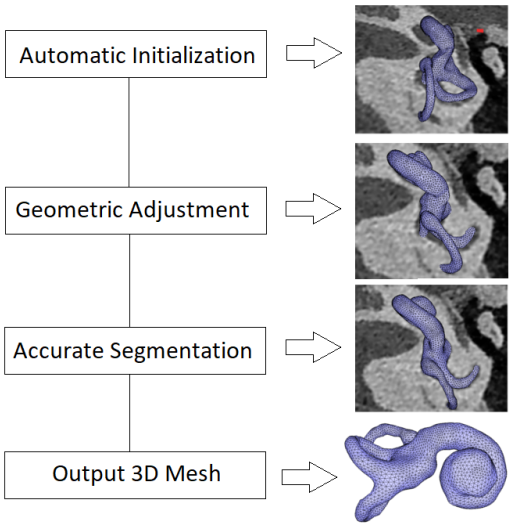

Figure 2: *Global scheme of the methodology.*

2.3. Automatic initialization of the 3D model

The purpose of this step is to automatically retrieve tympanic cavity on CBCT images to initialize the position of the *a priori* mesh. This anatomical point does not have a high morphological variability and can easily be found automatically. First, the volume is cut in two parts - the front and the back of the head - to avoid confusion with the sinus cavities. The back of the head is also split in two to find separately left and right features (Figure 3.A).
The following step is to find the deepest voxel located in front of the opening of the external auditory canal. For each slice of both volumes of interest, a binary image is computed using a threshold to differentiate the air and the soft tissues from hard tissues. To remove noise, we successively apply morphological closing and opening operations. Then, we fill all holes inside the temporal bone (Figure 3.B/C) and keep the deepest voxel of value 0 to the right (or to the left respectively). The most distant voxel position in all slices is used to initialize the model at a predefined distance and orientation (Figure 3.D). A hole filling operation prevents air cells in the temporal bone and the auditory tube to interfere with the detection process.

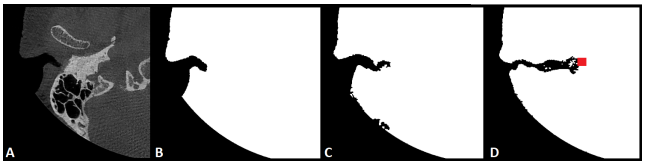

Figure 3: *Image A shows a regular greyscale CBCT slice of the region of interest. Images B,C and D show different slices of the binarized region of interest. The deepest point inside the cavity is marked with red.*

2.4. Geometric adjustment

The proximity of the input mesh S with the target location allows us to search for a rigid transformation matrix that would adjust the position and orientation of the 3D model.
For each vertex $v_i \in S$, an intensity $\Omega_i(v_i)$ is obtained with linear interpolation. On CBCT data, the inner ear has a specific range

of intensity. We rescale the interpolated intensities between 0 and 1, after applying a Gaussian transfer function based on the known mean intensity μ_o and standard deviation σ_o of voxel representing the inner ear. These values were derived from an intensity profile sampled in a slice of the CBCT volume. Voxels that represent the inner ear boundary are attributed a value of 1 in our rescaled intensity distribution, $\Omega_i^{\sigma_o}(v_i)$. To find the boundary, we compute the *sum of absolute differences* (SAD) with a windowed step function $H(x)$.

$$C_i(v_i) = \sum_{j=-\Phi}^{\Phi} |\Omega_i^{\sigma_o}(v_i + j)H(j)| \quad (1)$$

where Φ is the half-length of the similarity window.
In order to proceed to the model registration, we need to find the rigid transformation that will minimize the sum of all SADs. First, a template search is applied in the neighborhood of the initialized model to find the best translation. At each iteration, the value of the sum is compared to the previous smallest value. Finally, a gradient descent is applied to find the rotation that minimizes the sum.

2.5. Accurate deformable fitting of the 3D model

The following step is designed to perform surface deformation in order to achieve an accurate segmentation out of the 3D model boundary. This step uses an iterative Laplacian mesh optimization method [NISA06]. The use of a Laplacian-based method is adequate in our case, because the outer surface of the inner ear is fairly smooth. For each iteration, every vertex of the mesh is paired to a feature point in the volumetric image with a confidence value. Then vertex positions are computed with this target and they are optimized, preserving the local smoothness of the shape.

2.5.1. Target retrieval

The feature-matching step allows the identification of potential target points corresponding to the inner ear boundary inside the volumetric image. The detection of such features is performed along the normals of the 3D surface mesh. For each vertex v_i of the input mesh, intensity profiles $P_i(s)$ are sampled following the normal of v_i . The parameter $s \in [\frac{L}{2r}, \frac{L}{2r}]$ depends on the current step r and on the total initial length L. It is gradually decreased at each iteration to refine the deformation.
To find the boundary target point, we compute the *sum of absolute differences* as seen in equation 1 for each point of the profile. The target point t_i is obtained from the minimum SAD value location along the normal profile. Then, a confidence weight λ_i is set to $\lambda_i = 2\phi - \min(C_i(s))$. 2ϕ is the maximum value that can take $C_i(s)$.

2.5.2. Weighted deformation

Following the matching step, every vertex v_i is assigned to a target t_i with a confidence weight λ_i. The mesh is then deformed by solving an optimization problem that satisfies the target points locations and preserves the smoothness and topology of the input shape. New positions V' are computed by minimizing an energy function (equation 2) [CCC*16].

$$E_{\mathcal{L}}(V') = \alpha \sum_{i=1}^{n} w_i^2 (t_i - v_i')^2 + \|\mathcal{L}V'\|^2 \quad (2)$$

where α is the attraction strength parameter that weighs the volume energy - a squared sum between vertices and targets - weighted by w_i. The internal energy is given by the product of the current position of the vertices and the Laplacian matrix. Minimizing the quadratic energy functional results in an overdetermined linear system of equations $Ax = b$ where A is a $2n \times n$ matrix (equation 3) [CCC*16].

$$\begin{bmatrix} \mathcal{L} \\ W_p \end{bmatrix} V' = \begin{bmatrix} \Delta_p \\ W_p T \end{bmatrix} \quad (3)$$

where W_p are the weights attributed to each point w_i, and Δ_p is a set of length vectors of the local curvature. A high vector length value indicates a strong feature, while a null value is characteristic of a plane surface. These points have first been rescaled by a Gaussian function of mean μ and standard deviation σ based on the absolute vertex displacement distribution and multiplied by the global attraction strength parameter α. Matrix T contains coordinates of target t_i. In addition, the weight of targets located far from the actual surface of the mesh are progressively reduced to prevent vertices being attracted by false boundaries in the complex neighborhood. It is also important to note that in this method, the number of iterations is fixed in order to define the length L of profiles and the influence of α on w_i.

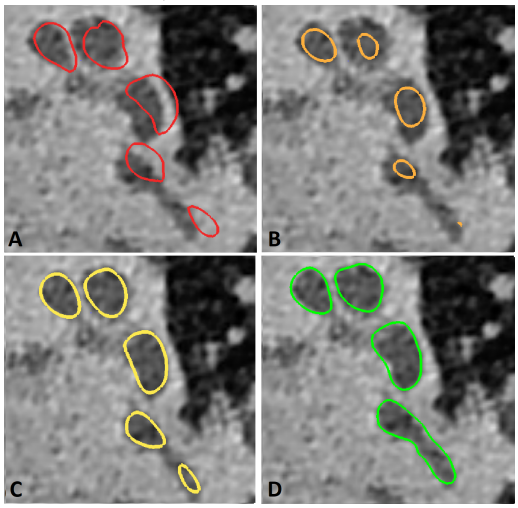

Figure 4: *Cross sections of left the inner ear model registered with equation 1 (A), at the first iteration (B), at the fifth iteration (C) and the final iteration (D).*

3. Results and validation

Our preliminary results include an evaluation of the proposed algorithm on one 8-bit CBCT acquisition obtained from a Newtom 5G unit (Newtom, Verona, Italy) in high resolution scan mode with a field of view of 15×5 cm. The reconstructed volume is displayed on a 3D grid of size $1020 \times 1020 \times 345$ with a voxel resolution of 0.150 mm in all three directions. The initial length L of the deformation algorithm is 8 with a profile spacing of 0.15 mm and the number of iterations is fixed to 15.
To estimate the accuracy of the segmentation method, we compare the results of our automatic extraction with those obtained from a manual approach carried out with 3D Slicer (https://www.slicer.org/). Quantitative comparisons are performed by

using the following metrics: Dice Similarity Coefficient (DSC), Jaccard Similarity Coefficient (JSC), Average Symetric Distance (ASD), Root-mean-square Deviation (RMSD) and Haussdorf distance (Hauss) described in [TH15]. These measures are used as means to quantify the overlaps percentage of segmentations, the mean distance of errors and the robustness to outliers.

Results in Table 1 show also a comparison of the proposed method with *Noble et al.* algorithm [Nob11]. To be more relevant, this comparison should be improved with more dataset in future work. We argue that our method yields a better global accuracy with a higher similarity index score and a smaller average symmetric distance. In addition, the proposed method does not require several meshes to create the prior shape. An unique approximate model could be considered sufficient. However, our method might be more sensitive to outliers since we obtain a higher Hausdorff distance. Figure 5 shows the distance mapped onto the 3D surface. The yellow part under the cochlea is weakly delimited in CBCT imaging. Furthermore, a strong *a priori* on the curvature limits deformations in this area in order to avoid serious potential errors. In addition, Figure 4.D demonstrates that the surface is able to fit accurately with the inner ear boundary, in spite of a rough initialization.

	DSC	JSC	ASD	RMSD	Hauss
Our method	90.6%	83.3%	0.15mm	0.13mm	1.07mm
Noble et al.	75%	-	0.21mm	-	0.80mm

Table 1: *Measures of performance applied to a left inner ear segmentation*

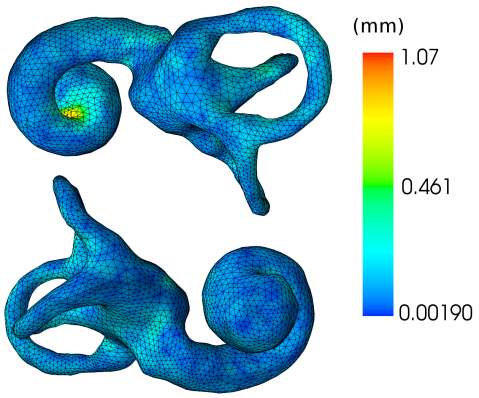

Figure 5: *Symmetric distance mapped over the surface of inner ear.*

4. Conclusion and future work

We have presented preliminary results of an automatic extraction of the inner ear within CBCT volume. Our first objective was to automatically initialize an *a priori* model and adjust its position and orientation inside the volume. Then, our second objective was to perform a deformation based on Laplacian optimization to allow iterative refinement of our input mesh based on the boundary information obtained from the CBCT volume. A preliminary validation shows the accuracy of our method compared with the *Noble et al.* [Nob11] automatic segmentation.

As future work, we intend to create a validation framework on a more extensive database in order to study the robustness of our approach. The application of our algorithm on the structures of the ossicular chain in the middle ear is also considered. In its current state, our method does not manage topological changes such as the absence of semicircular canals. A prior feature search is considered to adapt our *a priori* model to fit the constraints imposed by these particular cases.

Acknowledgements

This work was supported in part by the Natural Sciences and Research Council of Canada and by Useful Progress Canada Inc. under the Collaborative Research and Development program (NSERC CRDPJ 488452-15). The authors would like thank Jérôme Harrison for his contribution to the validation of the results and Nils Olofsson for his reading of this paper.

References

[AVDI02] ALLIEZ P., VERDIÈRE É. C. D., DEVILLERS O., ISENBURG M.: Isotropic Surface Remeshing. *RR-4594, INRIA* (2002). 2

[BBS12] BRAUN K., BÖHNKE F., STARK T.: Three-dimensional representation of the human cochlea using micro-computed tomography data: Presenting an anatomical model for further numerical calculations. *Acta Oto-Laryngologica 132*, 6 (2012), 603–613. 1

[CCC*16] CHARTRAND G., CRESSON T., CHAV R., GOTRA A., TANG A., GUISE J.: Liver Segmentation on CT and MR Using Laplacian Mesh Optimization. 1–12. 2, 3

[LTRRB08] LE TRONG E., ROZENBAUM O., ROUET J.-L., BRUAND A.: A simple methodology to segment X-ray tomographic images of a multiphasic building stone. *Image Analysis and Stereology 27* (2008), 175–182. 1

[NISA06] NEALEN A., IGARASHI T., SORKINE O., ALEXA M.: Laplacian mesh optimization. *Siggraph* (2006), 381. 3

[Nob11] NOBLE JH, LABADIE RF, MAJDANI O D. B.: Automatic Segmentation of Intra-Cochlear Anatomy in Conventional CT. *IEEE transactions on bio-medical engineering. 58*, 9 (2011), 2625–2632. 1, 4

[RTP*16] RAZAFINDRANALY V., TRUY E., PIALAT J. B., MARTINON A., BOURHIS M., BOUBLAY N., FAURE F., LTAIEF-BOUDRIGUA A.: Cone Beam CT versus multislice CT: Radiologic diagnostic agreement in the postoperative assessment of cochlear implantation. *Otology and Neurotology 37*, 9 (2016), 1246–1254. 1

[TH15] TAHA A. A., HANBURY A.: Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging 15*, 1 (2015). 4

[VA95] VOIE A.H. S. F.: Three-dimensional reconstruction of the cochlea from two-dimensional images of optical sections. *Comput. Med. Imag. Graph 19* (1995), 377–384. 1

[WNB*06] WANG H., NORTHROP C., BURGESS B., LIBERMAN M. C., MERCHANT S. N.: Three-dimensional virtual model of the human temporal bone: a stand-alone, downloadable teaching tool. *Otology & neurotology 27*, 4 (2006), 452–457. 1

[WZYL15] WEI X. F., ZHANG X. Y., YUAN W. U., LI Y. U. N. S.: Accuracy of computer-aided geometric three-dimensional reconstruction of the human petrous bone based on serial unstained celloidin sections. *Experimental and Therapeutic Medicine 9*, 4 (2015), 1113–1118. 1

[XSCY05] XIANFEN D., SIPING C., CHANGHONG L., YUANMEI W.: 3D semi-automatic segmentation of the cochlea and inner ear. *IEEE Engineering in Medicine and Biology Society. 6* (2005), 6285–8. 1

[Yan16] YANG C. C.: Characterization of scattered X-Ray photons in dental cone-beam computed tomography. *PLoS ONE 11*, 3 (2016), 1–14. 1

[YWR*00] YOO S. K., WANG G., RUBINSTEIN J. T., SKINNER M. W., VANNIER M. W.: Three-dimensional modeling and visualization of the cochlea on the internet. *IEEE Transactions on Information Technology in Biomedicine 4*, 2 (June 2000), 144–151. 1

[YWRV01] YOO S. K., WANG G., RUBINSTEIN J. T., VANNIER M. W.: Semiautomatic segmentation of the cochlea using real-time volume rendering and regional adaptive snake modeling., 2001. 1