



Published in final edited form as:

Otol Neurotol. 2009 June ; 30(4): 436–442. doi:10.1097/MAO.0b013e31819e61ed.

Automatic Identification and 3-D Rendering of Temporal Bone Anatomy

Jack H. Noble, B.E.^a, Benoit M. Dawant, Ph.D.^a, Frank M. Warren, M.D.^b, and Robert F. Labadie, M.D., Ph.D.^c

^aDepartment of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235, USA

^bDivision of Otolaryngology, University of Utah, Salt Lake City, UT 84132

^cDepartment of Otolaryngology, Vanderbilt University Medical Center, Nashville, TN 37232, USA

Abstract

Hypothesis—Using automated methods, vital anatomy of the middle ear can be identified in CT scans and used to create 3-D renderings.

Background—While difficult to master, clinicians compile 2-D data from CT scans to envision 3-D anatomy. Computer programs exist which can render 3-D surfaces but are limited in that ear structures, e.g. the facial nerve, can only be visualized after time-intensive manual identification for each scan. Herein, we present results from novel computer algorithms which automatically identify temporal bone anatomy (external auditory canal, ossicles, labyrinth, facial nerve, and chorda tympani).

Methods—An atlas of the labyrinth, ossicles, and auditory canal was created by manually identifying the structures in a “normal” temporal bone CT scan. Using well accepted techniques, these structures were automatically identified in (n=14) unknown CT images by deforming the atlas to match the unknown volumes. Another automatic localization algorithm was implemented to identify the position of the facial nerve and chorda tympani. Results were compared to manual identification by measuring false positive and false negative error.

Results—The labyrinth, ossicles, and auditory canal were identified with mean errors below 0.5 mm. The mean errors in facial nerve and chorda tympani identification were below 0.3 mm.

Conclusions—Automated identification of temporal bone anatomy is achievable. The presented combination of techniques was successful in accurately identifying temporal bone anatomy. These results were obtained in less than 10 minutes per patient scan using standard computing equipment.

INTRODUCTION

Expert knowledge of temporal bone anatomy requires years of training due to the highly complex shape and spatial arrangement of anatomical structures. Surgeons become very

adept at using 2-D CT scans to mentally visualize normal 3-D anatomy. However, novel surgical approaches and pathological variations impede this ability on a case specific basis. While some surgeons become adept at this even more difficult task (mental visualization of atypical anatomy), current technology does not allow sharing or verification of the mentally visualized anatomy. Furthermore, this approach does not allow for precise quantitative measurement of spatial relationships.

One solution to this problem is to create case-specific anatomical models by identifying temporal bone anatomy in CT images. These models can then be presented to the clinician in the form of 3-D visualizations and would allow precise measurement of tolerances between critical structures and pathology, thus providing extremely useful information to the clinician for the purpose of surgical planning. Towards this end, previous work has focused on extracting anatomical surfaces using a combination of thresholding and manual identification techniques. Thresholding consists of identifying structures based on their range of CT intensity values. Thresholding is non-specific, often including neighboring anatomical areas which have the same intensity as the anatomy of interest. Manual identification consists of using trained observers, e.g. radiologists or radiology technicians, to trace anatomy of interest on each slice of a CT scan. This process is time intensive, often requiring hours to delineate the anatomy of a single patient.

Using a combination of thresholding and manual identification, Jun et al [1] extracted the facial nerve, ossicles, labyrinth, and internal auditory canal in twenty spiral high-resolution CT volumes. The authors utilized this information to report on spatial relationships within the temporal bone. Wang *et al.* [2] created an extensive library of surfaces by combining serial, computer-assisted manual delineations of 2-D photographic images of histologically-sectioned, frozen temporal bone. Wang's intent was to create a tool for studying anatomical relationships. The main limitation of these and other works [3,4,5] is that they require manual delineation of anatomy.

In this work, we have developed a combination of techniques to automatically identify the vital anatomy of the temporal bone in unknown CT scans. We focused our study on structures of the temporal bone which are relevant to otologic surgery, specifically, the facial nerve, chorda tympani, external auditory canal, ossicles, and labyrinth. The methods we have developed require less than 10 minutes to fully process a single scan on a standard computer.

MATERIALS AND METHODS

The process used to automatically identify the labyrinth, ossicles, and external auditory canal relies on atlas-based registration, a common technique in the field of medical imaging. The principle of atlas-based registration is that an image of a known subject can be transformed automatically such that the anatomical structures of the known subject are made to overlap with the corresponding structures in the image of an unknown subject. Given a perfect registration, transformed labels from the known atlas exactly identify the location of the structures in the unknown image. Figure 1 shows an example of atlas-based registration as is typically used in neurosurgical applications. The underlying assumption of this method

is that the images of different subjects are topologically similar such that a one-to-one mapping between all corresponding anatomical structures can be established via a smooth transformation. For patients with normal anatomy, this assumption is valid in the anatomical regions surrounding the labyrinth, ossicles, and external auditory canal. Using atlas-based registration methods described previously [6,7,8] and an atlas constructed with a CT of a “normal” subject, we created a registration approach to allow labeling of these structures on temporal bone CT’s.

For the anatomical region surrounding the facial nerve and chorda tympani, topological similarity between images cannot be assumed due to the highly variable pneumatized bone. Therefore, the facial nerve and chorda tympani are identified using another approach, the navigated optimal medial axis and deformable-model algorithm (NOMAD) [8]. NOMAD is a general framework for localizing tubular structures. Statistical a-priori intensity and shape information about the structure is stored in a model. Atlas-based registration is used to roughly align this model information to an unknown CT. Using the model information, the optimal axis of the structure is identified. The full structure is then identified by expanding this centerline using deformable-model (ballooning) techniques.

To validate our process, we quantified automated identification error as follows: (1) The temporal bone structures were manually identified in all CT scans by a student rater then verified and corrected by an experienced surgeon. (2) Binary volumes were generated from the manual delineations, with a value of 1 indicating an internal voxel and 0 being an external voxel. (3) Surface voxels were identified in both the automatic and manually generated volumes. (4) For each voxel on the automatic surface, the distance to the closest manual surface voxel was computed. We call this the false positive error distance (FP). Similarly, for each voxel on the manual surface, the distance to the closest automatic surface voxel was computed, which we call the false negative error distance (FN) (See Figure 2). We compute both FP and FN errors because, as shown in Figure 2, the FP and FN errors are not necessarily the same for a given point. In fact, to properly characterize identification errors, computing both distances is necessary.

RESULTS

Our software takes as input clinically-applicable CT images. The software runtime on a typical Windows based machine (we use a 2.4 GHz Intel Xeon processor machine) is approximately 6 minutes for a typical temporal bone CT scan ($768 \times 768 \times 300$ voxels). Voxel dimensions of the scans used in this study were approximately $0.35 \times 0.35 \times 0.4$ mm³. The software outputs the identified structures in the form of surface files that can be loaded into generic CT viewing software.

Identification error was measured individually for the facial nerve, chorda tympani, ossicles, external auditory canal, and labyrinth. Table 1 shows error results for the surfaces automatically localized using our methods. As can be seen from this table, excellent results are achieved for the facial nerve and chorda tympani, with no error exceeding 1.15 mm. Overall mean errors for the labyrinth, ossicles, and external auditory canal is below two

voxels, with low median errors for each volume. Over the vast majority of these surfaces, there is excellent agreement between manually and automatically generated surfaces.

Relatively high maximum errors were seen at the labyrinth and auditory canal in three of fourteen ears. For the labyrinth, this error occurs at the most superior portion of the superior semicircular canal. The automatic identification errs laterally in this region and causes a sharp increase in the FN error. Error in the auditory canal occurs at the anterior wall of the tympanic membrane. In general, the membrane is poorly contrasted in CT. Where the error was observed, the membrane was not visible in the image, making accurate manual or automatic identification very difficult. The maximum errors at these points were on the order of 6 voxels.

As briefly mentioned in the previous section, atlas-based registration techniques alone are not appropriate for identification of the facial nerve and chorda tympani. To verify this conjecture, we applied the same atlas-based methods that were used for the other structures of the temporal bone to identify the facial nerve and chorda. The results of this approach are shown in Table 2. Mean and median errors with this approach are reasonable, suggesting that in most cases, atlas-based techniques perform adequately. However, maximum errors exceed 4 mm. As the facial recess, defined by the facial nerve and chorda tympani, has dimensions on this order, errors this large are deemed unacceptable for neurotologic applications. This can be visually appreciated in Figure 3.

We hypothesize that these errors occur for a variety of reasons. First and foremost, the pneumatized bone surrounding these structures introduces substantial topological variation between images. Second, these structures often have poor contrast with surrounding tissues. Third, these structures are very thin, as small as 1–2 voxels in diameter. The registration process will thus often place more emphasis on larger surrounding structures. It is clear, however, that the NOMAD algorithm successfully accounts for these difficulties.

Figure 4 shows the qualitative results of our localization algorithms. A 3D visualization of the structures and a CT cross section of the structures are shown for selected ears. All surfaces appear qualitatively accurate. While difficult to appreciate in 2-D presentation, these surface renderings are rotatable allowing appreciation of 3-D relationships from any perspective angle.

DISCUSSION

To safely and precisely perform neurotologic surgery, surgeons require intimate understanding of the 3-D spatial relationships of the temporal bone. For most common procedures, this understanding becomes second-nature as surgeons easily translate 2-D CT scans into mental 3-D models. However, in cases of altered pathology and/or uncommon surgical approaches this task proves difficult. In such situations, computer graphics can be used to enhance the surgeon's knowledge by providing virtual 3-D models that can be viewed from multiple angles. To date, such applications have been limited by the amount of time and effort necessary to manually delineate the anatomy of interest. Manual localization – manually identifying a structure in each and every CT slice – is time consuming, often

requiring hours to perform. Thresholding-based, automated algorithms do not provide the submillimetric resolution necessary for neurotologic surgery. For clinical applications, the ideal system would be accurate and automated requiring little to no input from the end-user.

We present herein initial results of a method that, for normal anatomy, achieves this ideal – accurate, automated anatomical identification and 3-D presentation of temporal bone anatomy from clinical CT scans. Using atlas-based registration (transforming a known specimen to match an unknown specimen), we have shown that the labyrinth, ossicles, and external auditory canal can be repeatedly identified with a mean false positive error of ~0.41 mm and a mean false negative error of ~0.45 mm. However, due to inter-subject variations in the temporal bone, atlas-based registration methods alone are not appropriate for identification of the facial nerve and chorda tympani. The NOMAD algorithm was designed to handle these variations and results in facial nerve and chorda identification with mean errors of only ~0.25 mm. Thus, using these methods together in a seamless software package, error for identification of the anatomy was typically within 2 voxels.

Although this initial work reports on limited temporal bone anatomy, we anticipate that identification of other anatomical structures (e.g. the internal auditory canals, carotid artery, jugular bulb, endolymphatic sac - ongoing work in our laboratory) will have similarly low error. Also, the presented localization techniques were tested only on patients exhibiting normal anatomy. Experimentation has yet to be performed on abnormal temporal bones, such as those exhibiting congenital malformations. One can speculate that the algorithms would perform less reliably in such situations due to the potential for greater registration error because of the topological inconsistencies between the atlas (known) and unknown scans.

The error results of this small study suggest that our methods are accurate enough to be used in variety of applications. However, no identification method is perfect, and it is important that the clinician is made aware of error in automatic structure identification. In order to explicitly let the clinician know where error exists, the software we have developed first presents identified structures as shown in the sagittal, coronal, and axial views of Figure 5. In these images, for example, the outlining of the cochlea and facial nerve are quite good, but the outlining of the I-S joint in the coronal scan is slightly inaccurate. Such qualitative assessment of the results must be performed by the user before the software creates the 3-D visualizations, which appear in the lower right panel of Figure 5, or allows measurement of spatial relationships. It is our intent that this process will alert the end-user to imperfections that could occur on a case-specific basis.

There are several applications in the field of neurotology in which we foresee a potential role for such software. (1) From an educational perspective, it would enhance teaching, allowing 3-D relationships to be appreciated from multiple perspective angles. (2) From a clinical standpoint, prior to surgery, a patient specific 3-D model of temporal bone anatomy could be used as a final anatomical check-point (e.g. “How wide is the facial recess on this patient?” “How far lateral does the facial nerve come with respect to the lateral aspect of the horizontal semicircular canal?”). Furthermore, with the addition of other anatomical structures (ongoing work in our lab), we foresee being able to analyze different surgical

approaches before cutting skin (e.g. “Would an infracochlear approach to the petrous apex be feasible?”).

A clinical application of interest to the authors, which motivated the development of this automated localization technique, is a novel surgical approach—percutaneous cochlear implantation. In this previously described technique [10, 11, 12], a single drill hole is made from the lateral surface of the skull surface to the cochlea avoiding vital anatomy, most notably the facial nerve. The drilling trajectory is planned pre-operatively on CT scans and then implemented with image-guided surgical technology. During early work in this project, it was discovered that trajectory planning was difficult due to poor visualization of the spatial arrangement of the facial nerve and chorda tympani in 2-D CT slices. However, as all of the vital anatomy can be automatically localized prior to planning, the drilling trajectory can be visualized and verified in a 3D environment. Furthermore, ongoing work in our laboratory has allowed automated localization of intracochlear structures and planning of the surgical trajectory which maximizes scala tympani insertion [13].

Acknowledgments

This work was supported in part by NIH grants R01DC008408 and R01EB006193.

References

1. Jun BC, Song SW, Cho JE, et al. Three-dimensional reconstruction based on images from spiral high-resolution computed tomography of the temporal bone: anatomy and clinical application. *Laryngology & Otology*. 2005; 119:693–8.
2. Wang H, Merchant SN, Sorensen MS. A Downloadable Three-Dimensional Virtual Model of the Visible Ear. *ORL*. 2007; 69:63–7. [PubMed: 17124433]
3. Nakashima S, Sando I, Tkahashi H, et al. Computer-aided 3-D reconstruction and measurement of the facial canal and facial nerve. I. Cross-sectional area and diameter: preliminary report. *Laryngoscope*. 1993; 103:1150–6. [PubMed: 8412453]
4. Takagi A, Sando I, Takahashi H. Computer-aided three-dimensional reconstruction and measurement of semicircular canals and their cristae in man. *Acta Otolaryngol*. 1989; 107:362–5. [PubMed: 2756825]
5. Green JD Jr, Marion MS, Erickson BJ, et al. Three-dimensional Reconstruction of the temporal bone. *Laryngoscope*. 1990; 100:1–4. [PubMed: 2293695]
6. Dawant BM, Hartmann SL, Thirion JP, Maes F, Vandermeulen A, Demaerel P. Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations. I. Methodology and validation on normal subjects. *IEEE Trans Med Imag*. 1999; 18–10:906–16.
7. Rohde GK, Aldroubi A, Dawant BM. The adaptive bases algorithm for intensity-based nonrigid image registration. *IEEE Trans Med Imag*. 2003; 22:1470–9.
8. Noble JH, Warren FH, Labadie RF, Dawant BM. Automatic segmentation of the facial nerve and chorda tympani using image registration and statistical priors. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2008; 6914:69140P–10.
9. Sethian, J. *Level Set Methods and Fast Marching Methods*. 2. Cambridge, MA: Cambridge Univ. Press; 1999.
10. Labadie RF, Choudhury P, Cetinkaya E, Balachandran R, Haynes DS, Fenlon M, Jusczyk S, Fitzpatrick JM. Minimally-Invasive, Image-Guided, Facial-Recess Approach to the Middle Ear: Demonstration of the Concept of Percutaneous Cochlear Access In-Vitro. *Otol Neurotol*. 2005; 26:557–562. [PubMed: 16015146]

11. Warren, FM.; Labadie, RF.; Balachandran, R.; Fitzpatrick, JM. Percutaneous Cochlear Access Using Bone-Mounted, Customized Drill Guides: Demonstration of Concept In-Vitro. presented at Mtg. Am. Otologic Soc; Chicago, IL. May 2006;
12. Labadie RF, Noble JH, Dawant BM, Balachandran R, Majdani O, Fitzpatrick JM. Clinical Validation of Percutaneous Cochlear Implant Surgery: Initial Report. *Laryngoscope*. 2008:118. pages pending.
13. Noble JH, Warren FM, Labadie RF, Dawant BM, Fitzpatrick JM. Determination of drill paths for percutaneous cochlear access accounting for target positioning error. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2007; 6509:25.

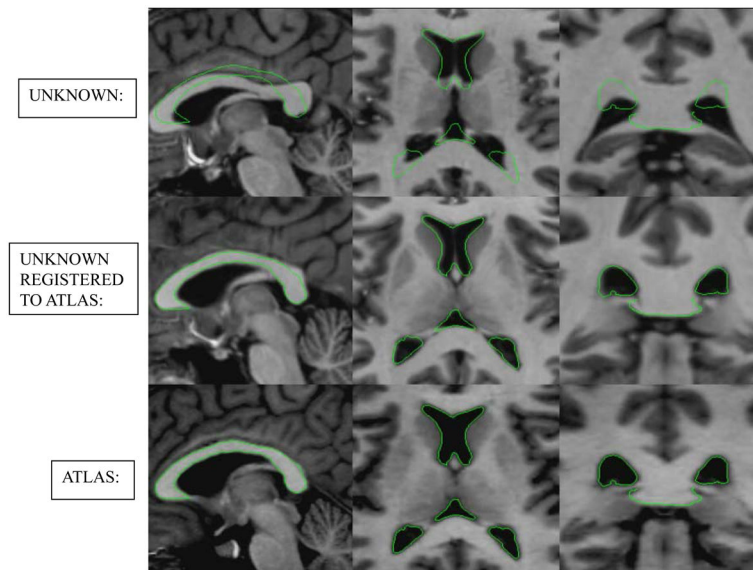


Figure 1. Example of inter-patient image registration. A registration atlas and target temporal bone images are shown, as well as the target image registered to the atlas
 Example of atlas based registration. Shown are delineations of the ventricles in the (left-to-right) sagittal, axial, and coronal views of the MR in the bottom row. These delineations are overlaid on (top row) the MR of another subject and (middle row) the same subject from the top row after registration to the image in the bottom row.

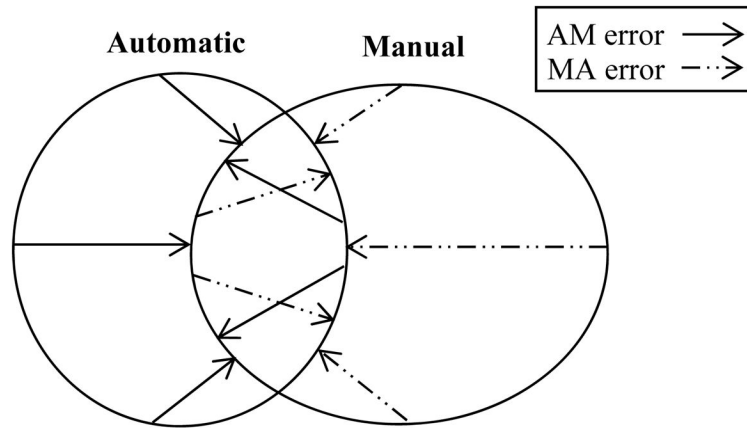


Figure 2. Methods for calculating error between automatically and manually generated structure surfaces. The types of error are visually illustrated in a 2D example
Methods for calculating error between automatically and manually generated structure surfaces. The AM error along the automatic structure surface is calculated as the closest distance to the manual surface. Similarly, MA error is calculated for the manual surface as the closest distance to the automatically generated surface. Zero error is seen where the contours overlap.

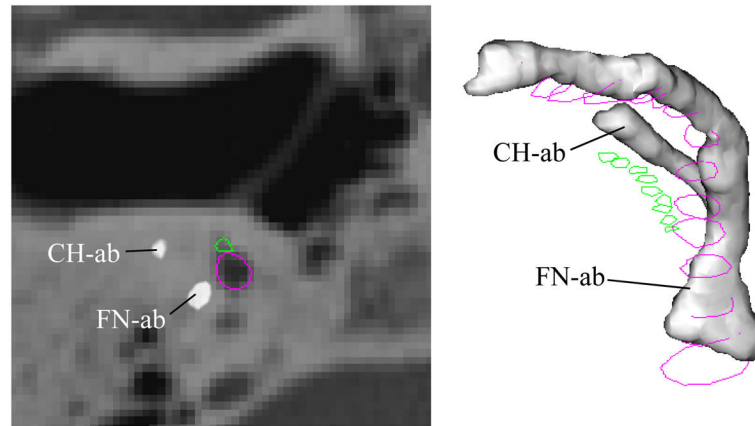


Figure 3. This figure illustrates disagreement between manual identification of the facial nerve and chorda and the localization results using atlas-based methods
Facial nerve and chorda segmentation results using atlas-based methods (FN-ab, CH-ab) for Volume 1, right ear. The true solutions for the facial nerve and chorda are delineated in purple and green. A significant error can be seen between the automatic and true solutions.

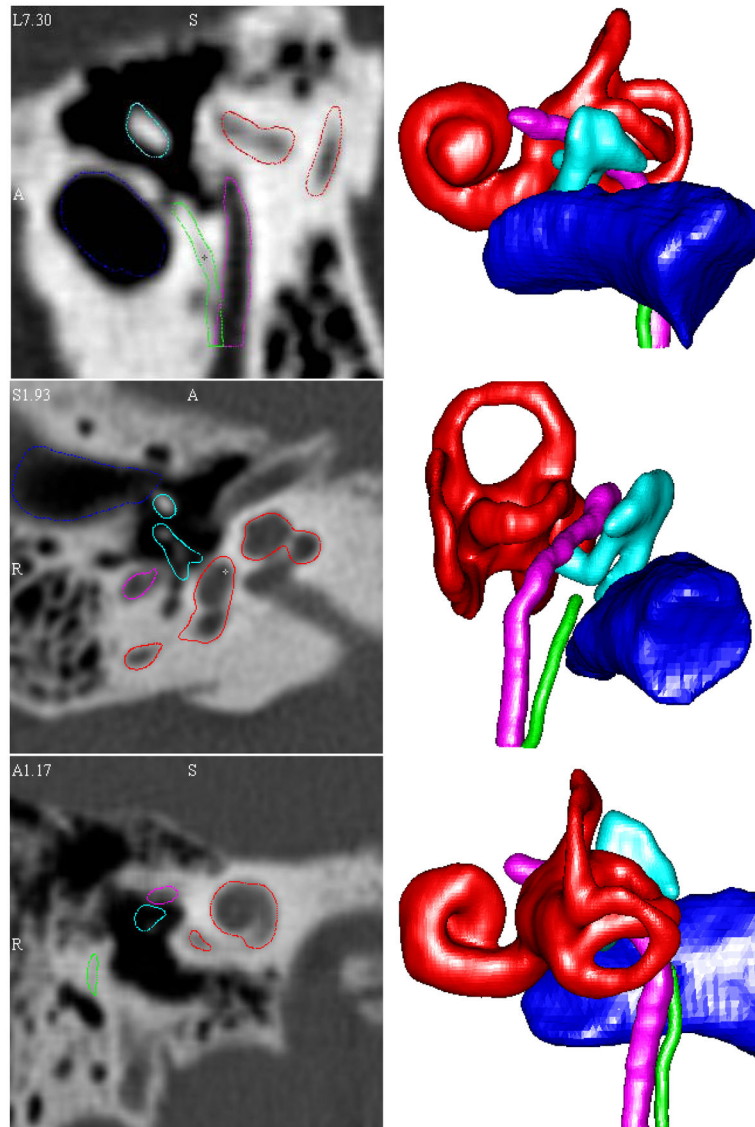


Figure 4. 3D visualizations and CT cross section of automatic localization results
3D renderings and CT cross section automatic segmentation results for three volumes. All segmentation results appear qualitatively accurate.

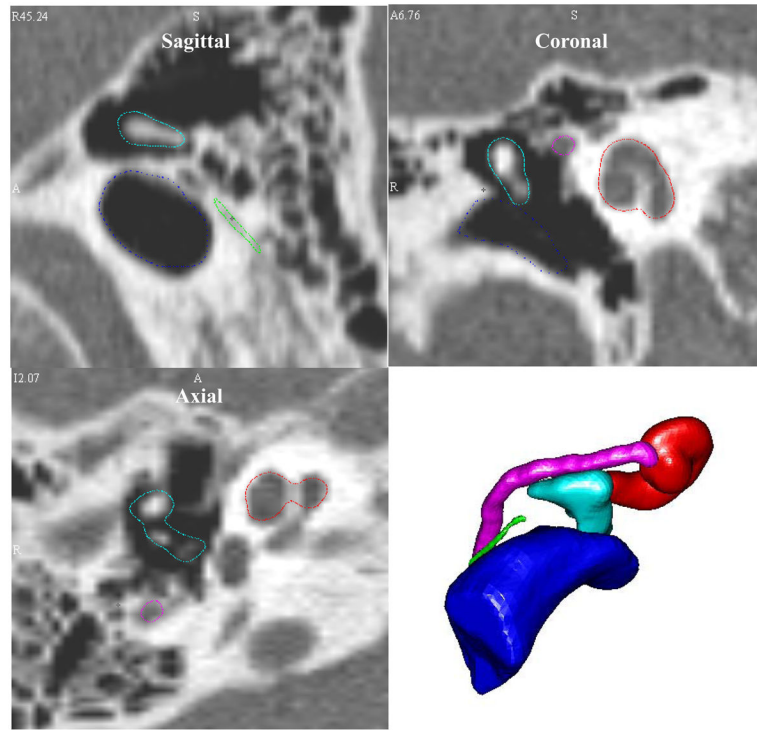


Figure 5. Localized structure contours in the CT planes and 3D visualizations presented to the end-user.

Table 1

This table shows error calculations for the facial nerve, chorda tympani, external auditory canal, labyrinth, and ossicles, which were localized using the methods presented in this work. All measurements are in millimeters. (To be placed on O&N's web resource)

	Atlas-Based Registration			NOMAD	
	Auditory Canal	Labyrinth	Ossicles	Facial Nerve	Chorda
Max FN	2.551	2.081	1.774	0.685	1.154
Mean FN	0.513	0.361	0.442	0.155	0.283
Median FN	0.461	0.332	0.384	0.000	0.340
Max FP	1.667	2.615	1.541	0.859	1.032
Mean FP	0.425	0.477	0.324	0.201	0.321
Median FP	0.400	0.444	0.332	0.298	0.342

Table 2

This table shows error calculations for the facial nerve and chorda segmented using solely registration-based techniques. All measurements are in millimeters.

Atlas-Based Registration		
	Facial Nerve	Chorda
Max FN	2.863	4.185
Mean FN	0.509	1.013
Median FN	0.354	0.711
Max FP	3.876	3.493
Mean FP	0.545	0.984
Median FP	0.354	0.728

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript