Influence of Image Quality on Semi-Automatic 3D Reconstructions of the Lateral Skull Base for Cochlear Implantation

F. Heckel¹, 2, S. Zidowitz¹, T. Neumuth², M. Tittmann³, M. Pirlich³, M. Hofer³

¹Fraunhofer Institute for Medical Image Computing MEVIS, Bremen, Germany
²Innovation Center Computer Assisted Surgery, Leipzig, Germany
³Leipzig University Hospital, Department of ENT-Surgery, Leipzig, Germany

Kontakt: frank.heckel@mevis.fraunhofer.de

Abstract:
Planning a cochlear implantation requires a 3D reconstruction of the risk structures usually from CT images. However, if the data does not fulfill specific image quality requirements, 3D reconstruction methods could fail or give inaccurate results. This study analyzes the correlation of image quality with the accuracy and reproducibility of semi-automatic 3D reconstructions of structures of the lateral skull base. We analyzed CT data from 12 patients where the risk structures of the lateral skull base were segmented by multiple experts using a dedicated semi-automatic tool. The quality of each image was rated by an experienced ENT surgeon with respect to various criteria. We found a correlation of the segmentation quality for some of the structures with the signal-to-noise ratio as well as the resolution of the images. Imaging artifacts influenced the reproducibility but not the accuracy of the reconstruction.

Keywords: image quality assessment, semi-automatic segmentation, 3D reconstruction, cochlear implantation planning, computed tomography

1 Introduction

3D medical images acquired by computed tomography (CT) for example are ubiquitous in today’s surgery. For surgical interventions and particularly if there is high risk of iatrogenic damage, e.g., due to a narrow surgical site or poor anatomical landmarks, the importance of expressive imaging increases. Interventions such as a cochlear implantation (CI) can be seen as a prototype for this. In CI a thin electrode is inserted into the cochlea in order to stimulate the hearing nerve of hearing impaired patients [1]. Common risk structures are the facial nerve, the brain, and important blood vessels for instance.

Planning a CI requires, beside others, the transition of the CT image into the surgical site as seen by the surgeon during the intervention. However, this crucial cognitive task requires a lot of experience. A 3D reconstruction of the risk structures from CT, as shown in Figure 1 could support surgeons in this process [2].

Figure 1: Cochlear implant (CI) planning example (image courtesy of D. Franz).
Reconstructing the structures by manual segmentation methods is time-consuming and lacks reproducibility, so it is therefore not transferable into clinical routine. Thus, (semi-)automatic methods are preferably used. It is generally known, however, that such methods have specific requirements on image quality \cite{3}. If these requirements are not fulfilled, 3D reconstruction methods in particular and computer-assisted systems in general could fail or give inaccurate results \cite{4,5}. Even for physicians, low-quality images might lead to inaccurate diagnosis \cite{6,7}.

This study addresses the influence of image quality on the quality of semi-automatic 3D reconstructions of structures of the lateral skull base for CI planning. We analyzed the correlation of different quality characteristics with the accuracy and reproducibility of semi-automatic 3D reconstructions by the wizard-based segmentation tool proposed by Franz et al. \cite{8}. As the name implies, this tool provides a wizard that guides the user through the segmentation process of each relevant structure. The segmentation algorithm combines anatomical knowledge about the shape and the size of each structure with a region growing process and morphological operations. The user can guide the segmentation process by interactively adapting parameters and seed points for example.

## 2 Materials and Methods

CT data from 12 patients was analyzed in a retrospective study. The images were acquired at Leipzig University Hospital and had a resolution ranging from $0.19 \times 0.19 \times 0.19$ mm to $0.393 \times 0.393 \times 0.393$ mm. On image showed a slice thickness of 0.5 mm. Each image was rated by an experienced ENT surgeon in a visual grading analysis using an absolute grading on an adjectival scale. The annotations have been done using a tool that has been developed using MeVisLab\cite{1} (see Fig. 2). Ratings and comments as well as meta-information for each image was stored in an XML file for further analysis. The following information was collected: The overall suitability of the image for CI planning; the coverage of the overall anatomical region (lateral skull base) and the target region (cochlea and the risk structures); general image quality characteristics (signal-to-noise ratio, resolution); and CT-specific imaging artifacts (motion, partial volume, beam hardening, metal, ring).

In order to evaluate the influence of the above criteria on the quality of 3D reconstructions for CI planning, we used preliminary data from a currently running study, which evaluates the reproducibility of the segmentation algorithm by Franz et al. for the following structures: acoustic canal, ossicles, tympanic cavity, facial nerve, chorda

\footnote{\url{http://www.mevislab.de}}
Figure 3: Boxplots of the quality of all segmentation results for each structure depending on the specific image quality criteria (AC = acoustic canal, OS = ossicles, TC = tympanic cavity, NF = facial nerve, NC = chorda tympani, RW = round window, CO = cochlea, vestibule, inner ear canal, SC = semicircular canal).
tympani, round window, cochlea / vestibule / inner ear canal, semicircular canal. 15 physicians participated in the study. Each CT image was processed by two to four physicians in a randomized order. The segmentation quality was measured in terms of the Jaccard index (volume overlap) with respect to one reference segmentation that has been generated by an experienced ENT surgeon in cooperation with an experienced neuroradiologist. We have analyzed the correlation between the segmentation quality of each individual structure and each patient with the image quality ratings. We consider the average Jaccard index of all segmentation results of a specific structure in a specific CT image to be a measure for the accuracy of the segmentation algorithm. In order to assess the reproducibility, we computed the difference between the maximum and the minimum Jaccard index of all segmentation results of a specific structure in a specific CT image. Spearman’s rank correlation coefficient $\rho$ has been used to measure the correlation, as the ratings are given on an ordinal scale.

3 Results

The coverage was rated as perfect for each image, meaning all relevant anatomical structure were visible. The images did not show CT-specific artifacts such as motion, metal, beam hardening, or rings. Partial volume artifacts were not present or not disturbing. The signal-to-noise ratio (SNR) and the resolution were not always sufficient, though, and for one image the SNR was even rated as bad.

Figure 3 shows boxplots of the Jaccard index of all segmentation results (i.e., all CT images and all participants), for each structure and depending on the most relevant quality criteria. In addition, the overall suitability as rated by the ENT surgeon is shown. For most structures, the diagrams show a trend toward better segmentation results with an increasing image quality. Only the segmentation of the acoustic canal does not seem to benefit from better image quality.

The correlation of the average Jaccard index and its maximum difference with the quality criteria is shown in Figures 4 and 5. Only the reconstruction accuracy of the ossicles, the cochlea, and the semicircular canal seem to be influenced by image quality. Again, our results show a negative correlation of image quality with the
reconstruction accuracy of the acoustic canal. These results are also consistent with the overall suitability rating. Partial volume artifacts do not seem to be correlated with segmentation accuracy in our data. Concerning the reproducibility, we found a slight correlation with the resolution for the facial nerve, the chorda tympani, and the cochlea. Partial volume artifacts seem to influence the reconstruction reproducibility of the ossicles, the tympanic cavity, the cochlea, and the semicircular canal.

4 Discussion

The reconstruction of structures from CT images is an ill-defined problem, as there is often no clear border between a structure and the surrounding tissues. This can lead to significant variations of segmentation results by different users or algorithms, also depending on the amount of user interaction. Reasons for this include partial volume artifacts, low contrast, and noise, but also differing interpretations for example [9].

Even though modern CT scanners typically provide very good images, their quality characteristics are not always sufficient. Both the resolution and the SNR are not always optimal. Our results suggest that this suboptimal quality indeed influences the quality of 3D reconstructions that are generated from those images via semi-automatic segmentation methods during cochlear implantation planning. This affects both the accuracy and the reproducibility of the reconstructions. The segmentation results of the acoustic canal showed the best accuracy and reproducibility compared to the other structures. It can be concluded that the acoustic canal can be reliably reconstructed even under suboptimal image quality conditions. The negative correlation between image and segmentation quality therefore rather shows a limitation of the study, whose main limitation is the small number of images from only one site that have only been rated by one expert.

As absolute grading studies are known to have a low reproducibility, more readers with different levels of experience have to be included in future studies. Alternatively, relative grading might be used [10]. In addition, only the Jaccard index was available as an indicator for the accuracy and reproducibility of the 3D reconstructions. As the distance to risk structures is an important criterion during the surgical intervention, surface-based measures such as the maximum and average surface distance would be suitable complementary
measures for segmentation quality. The site where the images were acquired is a university hospital with modern equipment and high standards. Thus, images with poor quality are rare. Further studies with more data from divergent radiology practices are necessary to verify the results. A focus in next evaluations could also be a more user friendly GUI for the annotation tool dedicated to physicians. Finally, mathematical measures for image quality and model observers should be considered as well [11,10]. Such methods would also allow an automatic estimation of whether a specific image is suitable for implantation planning before the surgeon starts processing the data. This way, insufficient images might be detected as early as possible in the clinical workflow, e.g., immediately after image acquisition.

5 Acknowledgment

This work was funded by the German Ministry of Education and Research (BMBF) as part of the project OR.NET (reference number: 16KT1232).

6 References