



# HFUS Imaging of the Cochlea: A Feasibility Study for Anatomical Identification by Registration with MicroCT

Lucas Lavenir, Nabil Zemiti, Mohamed Akkari, Gérard Subsol, Frederic Venail, Philippe Poignet

## ► To cite this version:

Lucas Lavenir, Nabil Zemiti, Mohamed Akkari, Gérard Subsol, Frederic Venail, et al.. HFUS Imaging of the Cochlea: A Feasibility Study for Anatomical Identification by Registration with MicroCT. *Annals of Biomedical Engineering*, 2021, 49 (5), pp.1308-1317. 10.1007/s10439-020-02671-1 . lirmm-02960692

**HAL Id: lirmm-02960692**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-02960692>**

Submitted on 7 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# HFUS Imaging of the Cochlea: A Feasibility Study for Anatomical Identification by Registration with MicroCT

Lucas Lavenir<sup>1</sup>, Nabil Zemiti<sup>1</sup>, Mohamed Akkari<sup>2</sup>, Gérard Subsol<sup>1</sup>, Frédéric Venail<sup>2,3</sup>  
and Philippe Poignet<sup>1</sup>

<sup>1</sup>LIRMM, University of Montpellier, CNRS, Montpellier, France

<sup>2</sup>Department of ENT and Head and Neck Surgery, University Hospital Gui de Chauliac,  
University of Montpellier, France

<sup>3</sup>Institute for Neurosciences of Montpellier, INSERM U105, France

**Abstract**—Cochlear implantation consists in electrically stimulating the auditory nerve by inserting an electrode array inside the cochlea, a bony structure of the inner ear. In the absence of any visual feedback, the insertion results in many cases of damages of the internal structures. This paper presents a feasibility study on intraoperative imaging and identification of cochlear structures with high-frequency ultrasound (HFUS). 6 *ex-vivo* guinea pig cochleae were subjected to both US and microcomputed tomography ( $\mu$ CT) we respectively referred as intraoperative and preoperative modalities. For each sample, registration based on simulating US from the scanner was performed to allow a precise matching between the visible structures. According to two otologists, the procedure led to a Target Registration Error (TRE) of  $0.32 \text{ mm} \pm 0.05$ . Thanks to referring to a better preoperative anatomical representation, we were able to intraoperatively identify the modiolus, both scalae vestibuli and tympani and deduce the location of the basilar membrane, all of which is of great interest for cochlear implantation. Our main objective is to extend this procedure to the human case and thus provide a new tool for inner ear surgery.

**Keywords**—Cochlea, high-frequency ultrasound, microcomputed tomography, US/CT registration, computer-assisted surgery.

## I. INTRODUCTION

Hearing loss is a major public health problem as 466 million people in the world suffer from hearing disorders. By 2050, the World Health Organization announced that this number could rise to 900 million, i.e. 10% of the world population. Among the main causes of hearing disabilities, the most common is sensorineural hearing loss. The latter occurs in the inner ear and directly affects the internal structures of the cochlea. For instance, hair cell degeneration is caused by presbycusis and ototoxic agents. Whereas both the Reissner's and basilar membranes, which respectively separate the scalae tympani, media and vestibuli, are affected by Ménière's disease and noise exposure. Nevertheless, since the cochlea cannot regenerate itself, cochlear implantation (CI) is the standard hearing restoration care strategy. This surgical procedure consists of inserting an electrode array into the scala tympani through the round window to electrically stimulate the cochlear nerve by bypassing the impaired hair cells. CI has resulted in significant benefits for implanted patients thanks to improvements in the technique in recent decades. The indication, which was restricted to profound deafness cases, has been extended to individuals with residual low-frequency hearing. Yet hearing preservation (HP) is quite challenging since electrode array insertion may cause irreversible anatomical damage<sup>7,17</sup>. Even with the most modern surgical techniques, CI resulted in HP declining over time<sup>23</sup>. Presently, verification of the preservation of the

internal structures of the cochlea is performed postoperatively thanks to cone beam computed tomography (CBCT).

#### A. Imaging the cochlea

Therefore CI would benefit from computer-assisted surgery (CAS) by providing real-time visual feedback of the electrode array insertion. Yet current radiological resources like computed tomography (CT) and magnetic resonance imaging (MRI) are not suitable for inner ear surgery due to their limited resolution. Actually they do not allow visualization of internal cochlea structures such as the basilar membrane (BM), whose thickness ranges from 0.2 to 5  $\mu\text{m}$  and width from 126  $\mu\text{m}$  (base) to 418  $\mu\text{m}$  (apex)<sup>13</sup>. Authors of other studies have proposed to use optical coherence tomography (OCT)<sup>11,19</sup>, which provides high resolution images of the cochlea. Nonetheless, prior thinning or decalcification is needed as the temporal bone is too optically dense for OCT<sup>11</sup>, which is inconsistent with *in-vivo* procedures. More recently, the latest CBCT devices such as the NewTom 5G® provide better resolution for temporal bone imaging than conventional CT<sup>15</sup>, i.e. up to 0.15 mm resolution, thus justifying their use for postoperative verification in CI. Unfortunately, these devices are unsuitable for intraoperative imaging due to their bulk and radiation exposure.

High-frequency ultrasound (HFUS, >20 MHz) seems to be a promising alternative since it can be used in a cost-effective way without any risk, while providing high spatial resolution. In 2009, Brown *et al.*<sup>5,24</sup> were the first to publish 2D-HFUS imaging (50 MHz) of the basilar membrane of an *ex-vivo* human cochlea. However, attenuation is a major concern when it comes to using HFUS to image a bony structure. As the cochlea is deeply anchored in the temporal bone, imaging the inside with an ultrasound probe was restricted to soft tissue accesses. The oval window being covered by one of the ossicles, the round window remained the only viable imaging spot, which considerably limited the field of view. Recently, Landry *et al.* managed to accurately visualize the inside of an *ex-vivo* human cochlea using HFUS (50 MHz), but at the expense of the integrity of the inner ear since they performed decalcification<sup>12</sup>. Actually, the frequency must be lowered when increasing the penetration depth, which decreases the resolution and in turn lowers the accuracy of the information obtained on the inside of the cochlea.

#### B. Registering 3D-US with CT

A solution to maximize information would be to fuse intraoperative US with a better anatomical rendering preoperative modality, like CBCT. This compounded modality could provide an intraoperative view of the internal structures and the electrode array with a high resolution representation of the cochlea to help identify the anatomical structures. However, both modalities need first to be registered and since CBCT provides volume, 3D-ultrasound acquisition is more suited.

US/CT registration is a well-known image analysis problem but as reported in the literature<sup>14,25-29</sup>, there is no superior technique and each algorithm depends on the application. Hence, many studies have focused on surface-based methods when imaging bony structures since algorithms cannot rely on standard features like edges or corners due to acoustic shadowing and distortion in US imaging. The reference for registering two surfaces as point clouds or meshes is actually the well-known iterative closest point (ICP) algorithm. But this method is highly susceptible to local minima and convergence can be critical. In 2007, Moghari *et al.* improved both robustness and accuracy thanks to the unscented Kalman filter (UKF) to register US and CT images<sup>16</sup>. Validation was performed on a scaphoid phantom bone with a TRE of 0.27 mm for the UKF and 2.37 mm for the ICP. The main drawback is the need for segmentation in ultrasound data, which can considerably extend the operative time<sup>28</sup> or induce operator-dependent bias in accuracy if done manually. So in 2010, Rasoulia *et al.* focused on automating the segmentation part via dynamic

programming<sup>21</sup>. Yet regarding the CT-data resolution, this method exhibited low accuracy with a TRE of 2.47 mm on spine phantoms. More recent studies also proposed to use phase information, Radon transform, new bone descriptor or peak detection for extracting the bone surface<sup>4,10,18,26</sup> but without any significant improvement in the registration error.

Performances of surface-based approaches directly depend on the segmentation which often remains complex and leads to partial detection. On the contrary, intensity-based methods are segmentation-free and more accurate since they are based on a similarity measure. Besides both registrations can be combined to roughly align the two modalities and then enhance the accuracy<sup>26</sup>. Common similarity measures like cross-correlation or mutual information nevertheless tend to fail with multi-modal analysis like US/CT registration. Ultrasound imaging actually relies on acoustic physics whereas computed tomography is based on X-ray absorption measurements and hence the same areas of interest can lead to remarkable differences. This phenomenon is even more significant with bony structures since CT images are not hampered by attenuation, contrary to US. Thus, Wein *et al.* proposed an interesting alternative by directly simulating ultrasound imaging from CT and even developed their own similarity measure<sup>25</sup>. They managed to reproduce the transmission and reflection of the sound waves along tissue interfaces according to CT Hounsfield units. Yet their approach led to a TRE of 8.1 mm, which is significant regarding the data resolution. More recently, they adapted their reasoning to the US/MRI registration case and obtained a fiducial registration error (FRE) of 2.51 mm on brain images<sup>8</sup>. In the studies<sup>3,27</sup>, Brendel *et al.* and Winter *et al.* developed a similar approach for spinal surgery by taking acoustic attenuation inherent to bony structures into account. They proposed an estimation of the surface that would be visible with US imaging based on casting of virtual ultrasound waves in the preoperative CT-volumes. Thanks to this segmentation-free procedure, they achieved a submillimetric precision to the estimated optimum. The most critical point of this method is the assessment of the orientation of the transducer for simulating virtual ultrasound waves. Yet, as Brown *et al.* also suggested, the cochlea could be imaged *in-vivo* by inserting a probe through the ear drum, leading to accurate knowledge of the transducer position.

### C. Aim of the study

This paper aims at assessing the feasibility of intraoperatively imaging internal tissues of the cochlea via ultrasound. A second objective is also to assess the relevance of registration with a better rendering modality for more precise identification of the structures of interest. This could be the first step in the development of a real-time guiding system for intracochlear navigation. Compared to the approach of Winter *et al.*<sup>27</sup>, we aimed at reproducing a monomodal problem by registering two binarized volumes from US and ultrasound-simulated images. To the best of our knowledge, this is the first study where an intensity-based method is used for registering binary volumes.

## II. MATERIALS AND METHODS

Since we are conducting a preliminary cochlear imaging study, we decided to carry out our experiments under ideal conditions, which we describe hereafter. This will allow us to study the feasibility of such an approach and, in case of positive results, further studies will focus on how to more closely replicate clinical conditions, as also discussed in section IV.

### A. Specimen preparation

We used 6 ex-vivo guinea pig cochleae, which closely resemble the human model<sup>9</sup> and are routinely used in cochlear implant studies. They were removed from adult female Hartley albino guinea pigs (350 – 500g) with no middle ear effusion or disease verified during cochlea harvesting.

Our experiments were conducted with the approval of the local ethics Committee for the Protection of Persons (CPP Sud Méditerranée IV, N°14461). A lethal intraperitoneal injection of sodium-pentobarbital (300 mg/kg) was performed under general anesthesia (sevoflurane gas anesthesia). The left and right temporal bones were removed under a microscope. The bullae were opened and the cochleae were fixed with a paraformaldehyde 4% injection through the round window membrane with a leakage hole drilled at the apex of the cochlea. Animal care and handling followed the animal welfare guidelines of the “Institut National de la Santé et de la Recherche Médicale” (INSERM), under the approval of the French “Ministère de l’Agriculture et de la Forêt”.

## B. Data acquisition

### *Preoperative data*

We used  $\mu$ CT as the preoperative modality to avoid being limited by the imaging procedure. Actually, it is more suited for cochlear analysis<sup>2,22</sup> with better rendering<sup>6</sup> and a higher signal-to-noise ratio than conventional CT. Microtomographic data acquisition was performed using the EasyTom 150® (X-ray source, 70 kV, 290  $\mu$ A, isotropic resolution 21.17  $\mu$ m). Multiplanar reconstruction was applied to the native images using the manufacturer’s X-Act® software, thus producing an isotropic volume of the cochlea.

### *Intraoperative data*

In addition, guinea pig cochleae are less dense than those of humans, so high-frequency is no longer a limiting factor for imaging the internal structures with ultrasound, contrary to the previous findings<sup>5,12,24</sup>. We thus also used HFUS for intraoperative imaging, which offers a higher spatial resolution than conventional US. The ultrasound data were acquired with the Vevo 3100 Imaging Platform® (FUJIFILM Visual Sonics Inc., Toronto, Canada). We used the MX550D 40 MHz center frequency probe in B-mode, which allowed a 15 mm penetration depth, with 1 mm offset and 40  $\mu$ m axial resolution. To acquire 3D-data, the transducer was mounted on a 3D-motor stage (VS-20005, FUJIFILM Visual Sonics Inc.), which allowed to perform a translational scanning (Fig. 1). For each acquisition, the sample was placed in a container filled with water to allow acoustic coupling and high contrast. Before the insertion, it was fixed on a 3-D printed holder which was meant to lift the cochlea from the bottom of the container where ultrasound may reflect. Thanks to a mechanical positioning system, the operator then manually adjusted the height of the transducer attached to the motor stage to center the focus on the target object. The path to be travelled consisted in a series of steps, in a direction perpendicular to the imaging plane, which distance was also adjusted by the operator to cover the whole cochlea. At each step, the transducer acquired one two-dimensional slice with a fixed thickness of 76  $\mu$ m. These slices were then combined in a volume using the manufacturer’s Vevolab® software.

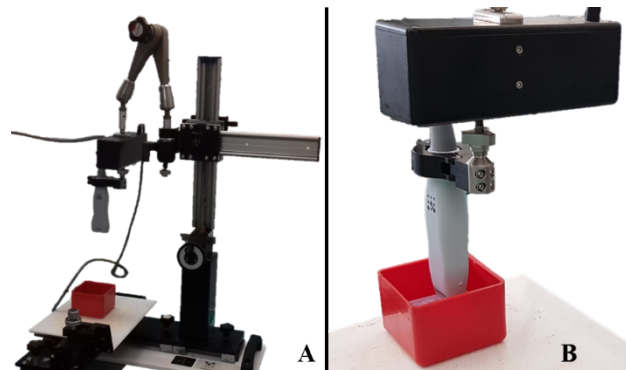


Fig. 1. Setup for intraoperative acquisitions: A. Mechanical positioning system holding the motor stage (black box) on which the MX550D was mounted. B. Scanning of a sample which was placed in a container filled with water for acoustic coupling.

### *C. Data preprocessing*

All the data were preprocessed using the CImg Library, an open-source C++ toolkit for image processing. Registration and merging of the final volumes were performed using FIJI® software. Fig. 2 summarizes the different processing steps of the procedure we describe hereafter.

#### *Adapting the data*

First, both preoperative and intraoperative datasets were adapted to enhance the different processing steps of the procedure. As the native  $\mu$ CT volumes required more than 1 GB of RAM, we decided to downsample the data to reduce the computational effort for any further processing step. A factor of 2 led to a new isotropic resolution of 42.34  $\mu\text{m}$ , which appeared as a good trade-off between the sampling rate and the need for memory. In addition, native US volumes were cropped to remove unnecessary metadata of the Vevo 3100 Imaging Platform®.

#### *Preregistration*

Since our registration algorithm is based on an intensity-based method, preregistration of both modalities is also required. Indeed, these approaches only provide local accuracy and hence they require initial alignment. For each sample, we asked an otologist to select four paired points on the cochlea in both modalities. Then, via least squares minimization, we determined the rotation and translation parameters to compute the associated transformation and preregister the datasets. We denoted US and  $\mu$ CT volumes as the template and moving images for registration, respectively.

#### *Simulating the US*

As mentioned above, due to substantial differences in the underlying US and  $\mu$ CT imaging physics, the information in the preoperative and intraoperative datasets was not comparable. Consequently, standard multi-modal similarity measures tend to fail for registration. As a solution, we propose to reproduce a mono-modal problem by tailoring our data to better fit conventional criteria.

First, we applied global thresholding to both modalities in order to extract the bony part of the samples. Although it is considered to be the simplest method for segmentation in image processing, it is commonly used in conventional CT imaging to distinguish tissues according to the Hounsfield scale. However, the  $\mu$ CT scanner was not calibrated and we had to visually define the threshold for the preoperative volumes. Despite our discussion about the difficulties encountered in trying to segment bone surfaces in US imaging in section I.B, thresholding was also suitable for scanned samples immersed in water baths as our experiments were conducted under ideal conditions. The threshold value was again defined visually.

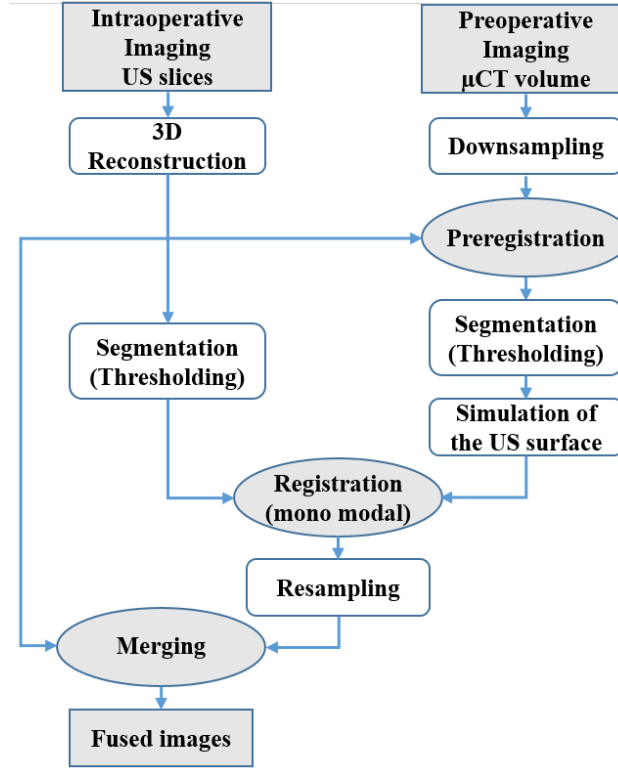


Fig. 2. Pipeline of the overall procedure for multi-modal registration of the cochlea

This basic segmentation led to binary images for both modalities with white voxels for the bony structure and black ones for the background. Yet, the two stacks were still not comparable. The segmented intraoperative image contained information relative to a surface whereas the one provided by the segmented preoperative image referred to a volume. Thus, the next step was to cast virtual ultrasound waves to simulate the surface of the binary  $\mu$ CT image that would be visible with ultrasound imaging. Since the US transducer scan was performed along the  $z$  direction, the simulation process merely consisted in generating white voxels after hitting those on the surface along the  $y$ -axis. However, because of the attenuation phenomenon, the surface information must not be restricted to a single voxel. Thus, in order to gain consistency between the two processed images, we decided to replicate the attenuation depth of the binary US images when hitting a surface voxel in the binary  $\mu$ CT images. The aforesaid depth was specific to each sample and its computation consisted in two steps. First, for each slice of a stack, we computed the average depth along the  $y$ -axis in the binary US image. Then, the mean of the averaged depth of the full stack is computed.

#### D. Registration and merging

We applied an automatic rigid registration algorithm based on maximization of the standard correlation coefficient between “US-simulated”  $\mu$ CT-images and segmented US-images, where the latter was again the template. As both images were binary, this criterion maximized the overlapping of black and white voxels. However, we could reasonably assume that several local minima existed as no assumptions were made on the shape of the cost function. Therefore, because of the low quantity of white voxels compared to black ones, a weak variation in the overlaying of the entire volumes could lead to significant variations in the final registration of the white voxels. Consequently, in order to minimize the quantity of black voxels, we reduced the initial region of

interest in the US-data and applied the same window to the  $\mu$ CT-volumes just after the preregistration.

Finally, the resulting transformation generated by the registration process was applied to the  $\mu$ CT native image, leading to aligned  $\mu$ CT and US-images. Then we resampled the aligned  $\mu$ CT to perfectly match the US native window and assigned a color to each modality (red for US, green for  $\mu$ CT) in order to merge them into a single image (Fig. 3).

All six datasets, including the segmented US and “US-simulated”  $\mu$ CT-images, were successfully registered using our method. For each sample, the registration took a few seconds to perform on CPU on a computer with an Intel® Xeon™ i7-8750H 2.20 GHz processor.

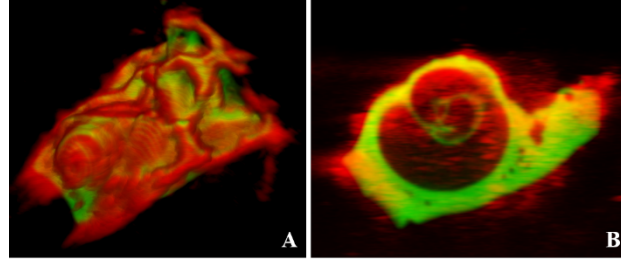


Fig. 3. Resulting alignment for cochlea n°6 after registration and merging of both modalities (US in red and  $\mu$ CT in green): A. 3D rendering. B. Example of merged slices of the same sample.

#### E. Registration error evaluation

A relatively simple matching of the US native and registered native  $\mu$ CT-volumes allowed a qualitative registration evaluation. Each cochlea was screened all along the z-axis by an otologist experienced in cochlear imaging interpretation.

Once we determined that the obtained registration was qualitatively satisfactory and enabled clear identification of the structures while navigating within the cochlea, we then sought to quantitatively evaluate the registration.

Nevertheless, in the absence of any ground truth with regard to the US/ $\mu$ CT registration of our dataset, we could only assess our method by manually selecting remarkable paired-points on the registered volumes. Two otologists were thus asked to place six landmarks in the native US slices and the corresponding ones in the registered native  $\mu$ CT. These fiducials were selected on the basis of their ability to be easily identified in both modalities. The US and  $\mu$ CT-coordinates of each point were predicted to be identical in case of a successful alignment. We thus averaged the sum of distances between the fiducials placed on the  $\mu$ CT and US-images to evaluate the registration error. After removal of the outlier distances, we reported the TRE ( $\pm$  standard error) for each observer and for the overall procedure. All calculations discussed here are presented in the next section.

Intra- and inter-rater reliabilities were not calculated since the high variability of the otic capsule due to the removal of the temporal bone did not allow us to place repeatable landmarks.

### III. RESULTS

#### A. Registration

A typical example of the processing steps after preregistration is presented in Fig. 4 for cochlea n°1.



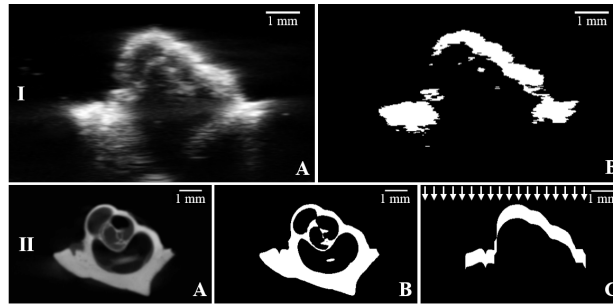


Fig. 4. Processing steps after preregistration for cochlea n°1: I.A. Native US image. I.B. Binary US image after segmentation (thresholding). II.A. Native  $\mu$ CT image. II.B. Binary  $\mu$ CT image after segmentation (thresholding). II.C. Binary  $\mu$ CT image after segmentation and "US simulation". White arrows show the orthogonal orientation of the "US-simulated" waves.

Here we respectively focused on two corresponding slices of the native US and  $\mu$ CT-volumes for a better understanding of the situation. Firstly, both native US (Fig. 4.I.A) and  $\mu$ CT (Fig. 4.II.A) slices were thresholded (Fig. 4.I.B and 4.II.B) to extract the bony part. Then acoustic attenuation simulation was performed on the binary  $\mu$ CT-slice (Fig. 4.II.C), whose depth was averaged in the segmented US volume to reproduce the same quantity of white voxels.

Once segmented, the resulting volumes were registered according to the standard correlation coefficient. The obtained transformation was applied to the native  $\mu$ CT-dataset, which was then resampled to perfectly match the US native dimensions. Thus, we were able to navigate inside the cochlea and use the combination of both modalities to identify the anatomical structures of interest (Fig. 5).

The TRE for each sample and each otologist are given in Table. I. The highest registration error was obtained on cochlea n°4, with  $0.32 \pm 0.08$  mm for otologist n°1 and  $0.50 \pm 0.19$  mm for otologist n°2. The lowest registration errors were obtained on cochlea n°3, with  $0.17 \pm 0.02$  mm for otologist n°1, and on cochlea n°2, with  $0.23 \pm 0.05$  mm for otologist n°2. On average, accuracy was higher for otologist n°1 than for otologist n°2, with respective TRE of  $0.25 \pm 0.04$  mm and  $0.38 \pm 0.10$  mm. Overall, the algorithm achieved a TRE of  $0.32 \pm 0.05$  mm. Calculations and statistical analyses were performed using Matlab®.

Cochlea	Target Registration Error (in mm)		
	Otologist n°1	Otologist n°2	Overall
1	$0.25 \pm 0.08$	$0.27 \pm 0.17$	$0.26 \pm 0.12$
2	$0.31 \pm 0.13$	$0.23 \pm 0.05$	$0.27 \pm 0.10$
3	$0.17 \pm 0.02$	$0.44 \pm 0.30$	$0.32 \pm 0.27$
4	$0.32 \pm 0.08$	$0.50 \pm 0.19$	$0.41 \pm 0.17$
5	$0.18 \pm 0.04$	$0.48 \pm 0.24$	$0.34 \pm 0.23$
6	$0.26 \pm 0.11$	$0.36 \pm 0.30$	$0.31 \pm 0.22$
Mean	$0.25 \pm 0.04$	$0.38 \pm 0.10$	$0.32 \pm 0.05$

Table. I. Calculation of the mean registration error and standard deviation for cochlear fiducials.

### B. Visualization of internal cochlea structures

For all the samples, the use of HFUS enabled to image beyond the bony surface of the cochlea that was exposed to the transducer. Despite their high spatial resolution, HFUS is hampered by both distortion and attenuation after having crossed the surface, which only allowed to visualize large and dense structures. In all cases, the modiolus was the only one to be clearly identifiable contrary to the different conducts, which were impossible to distinguish without additional information. On the other hand,  $\mu$ CT provided a 3D volume of the entire samples. In addition to the modiolus, the coils of the cochlea were easily identifiable, in which we could locate the scalae tympani (ST) and vestibuli (SV). Since the BM was not visible in this modality, we identified both conducts in each coil by referring to the part which was closer to the apex for the SV and to the base of the cochlea for the ST. Despite the poor anatomical rendering of ultrasound, registration of

both US and  $\mu$ CT volumes enabled precise identification of the scalae tympani (ST) and vestibuli (SV) at the apex of the cochlea. Moreover, by merging the US with the registered  $\mu$ CT-volumes, it was also possible to localize the BM. Although the width of this membrane falls under the resolution of both modalities,  $\mu$ CT allowed us to deduce its location by considering another structure of interest called the osseous spiral lamina (OSL). It consists of a tent-shaped extension of the modiolus, clearly visible in the preoperative modality (Fig. 4. III), which actually supports the aforesaid membrane. Consequently, by referring to the tip of the OSL, it was possible to infer the location of the BM in the  $\mu$ CT images and thus thanks to registration, in the US images.

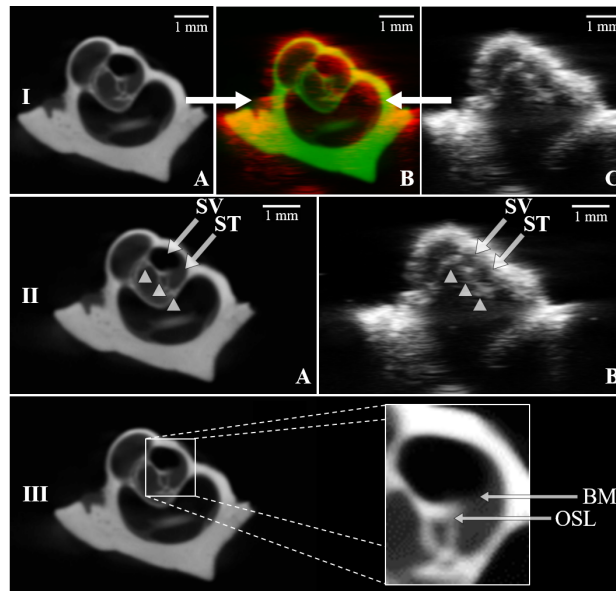


Fig. 5. Cochlea n°1: I. Merging of the two processed modalities: A.  $\mu$ CT slice after registration. B. Merged US (red) and  $\mu$ CT (green) slices. C. US native slice. II. Visualization and identification of internal structures: modiolus (white arrowheads), scalae tympani (ST) and vestibuli (SV) in the  $\mu$ CT image after registration (A) and its corresponding US image (B). III. Visualization of the osseous spiral lamina (OSL) which allows us to deduce the location of the basilar membrane (BM).

#### IV. DISCUSSION

The proposed approach enabled intraoperative imaging of internal structures of six *ex-vivo* guinea pig cochleae, such as the modiolus (Fig. 5). Thanks to the registration with a preoperative  $\mu$ CT volume, we managed to accurately identify the scalae tympani and vestibuli in the US images. According to the orientation of the OSL in the preoperative data, we could also deduce the location of the BM in the corresponding intraoperative slices, which is of great interest for CI. Actually, during the implant insertion, it may be possible to intraoperatively visualize the electrode array since it is supposed to be denser than the modiolus. With the location of the electrode with regard to the BM, the surgeon could avoid anatomical damage and increase HP rate.

Nevertheless, applications to surgical procedures suggest the use of a miniaturized transducer to be inserted through the eardrum. Although few works reported inner ear imaging<sup>5,24</sup>, miniaturization is no longer a limiting factor<sup>1</sup>. However, a trade-off between resolution and compactness still remains.

Overall, the estimated TRE of the algorithm was  $0.32 \pm 0.05$  mm according to both otologists. Compared to previously reported researches, there is a significant improvement in registration. In their more recent study on femur, tibia and fibula, Wein et al<sup>26</sup> actually reported a registration error of 3.7 mm. Even in their comparative study of the most modern techniques, Pandey et al<sup>18</sup> obtained a TRE of 2.44 mm with the most accurate method. Yet, it should be interpreted with regard to the

dimensions of the studied organ and the resolution of the imaging devices. One must keep in mind that the size of the cochlea is only of a few millimeters. Procedures could be hampered by such a registration error since the dimensions of certain structures of interest are under 0.32 mm. This is the case for CI as it requires precise location of the basilar membrane, whose width ranges from 0.1 mm to 0.4 mm. Besides, it can reasonably be assumed that the method we used for evaluating the algorithm induced accuracy bias. Although we asked two ENT surgeons familiar with the cochlea anatomy to place fiducials on both modalities, the TRE of each observer clearly highlighted that the evaluations were operator-dependent (Table. I). Otologist n°1 was more experienced in US-imaging of the inner ear than otologist n°2. His TRE ( $0.25 \pm 0.04$  mm) was significantly lower, compared to the latter ( $0.38 \pm 0.10$  mm). Hence, the estimated accuracy directly depends on the level of expertise of the ENT surgeon evaluating our algorithm. Therefore further studies will focus on finding an objective reference registration to enable correct assessment of the registration.

Thus, in order to combine our pipeline with inner ear procedures, improvements regarding the accuracy will be required. Generally speaking, the accuracy of any registration algorithm mainly depends on two building blocks: the metric and the optimizer. In our approach, the first one was a classical standard correlation criterion. Since we were dealing with a set of binary images, this metric was meant to maximize the overlap of both white and black voxels. The major drawback is that the information of the surface to be registered is limited to white voxels, which appeared to be outnumbered by black voxels. Thus, as mentioned before, a significant misalignment between the two surfaces may still result in a high value for the metric, which is not necessarily associated to a perfect alignment. One improvement could be to define a tailored cost function which only takes into account the white voxels like the one proposed by Winter *et al.*<sup>27</sup>. Compared to other approaches, the strength of Winter's method relies in the absence of any complex processing of the intraoperative data, which, as stated in section I, is an error prone task. Not to mention that it can be implemented in a fully automatic way. Here, accuracy was directly related to the quality of the US segmentation, which threshold was defined visually and was so operator dependent. Yet, the effectiveness of this metric relies on the quality of the surface extracted from the preoperative images. In order to gain consistency between US and US-simulated images, one more improvement could be to consider attenuation with an exponential law. Thus it would no longer lead to a one-voxel-depth surface but to a more realistic structure with gray values. Finally, beyond these processing considerations, the choice of the optimizer has also a significant impact on the outcome of the algorithm. In our case, we are looking at a 6-dimensional space: 3 translations and 3 rotations for a rigid mapping. Since common algorithms are often susceptible to local minima, the best solution may not be easy to find. Consequently, a last proposal could be to use global optimizers that have proven to be really effective for complex objective function, like the CMA-ES<sup>27</sup> or the BOBYQA<sup>20</sup> derivative free algorithms.

Nevertheless, one may not consider our study to be representative for human surgery since it was carried out on animal specimens. Yet guinea pig cochleae are similar to those of humans and with some improvements this procedure could very likely be extended to clinical cases. One main limitation concerns US attenuation in the human case. Contrary to guinea pig cochleae, the otic capsule is denser in humans, which overrules the use of HFUS. Hence, we also conducted an experiment to investigate the use of a 5-14 MHz US probe (Ultrasonix® 4DL14-5/38 3D/4D-linear probe) on an *ex-vivo* human cochlea which allowed us to probe deeper into bony structures, but at the expense of the image resolution (Fig. 6). Registration with a higher resolution preoperative modality such as CBCT is necessary to obtain more detailed information, thus confirming the relevance of our approach.

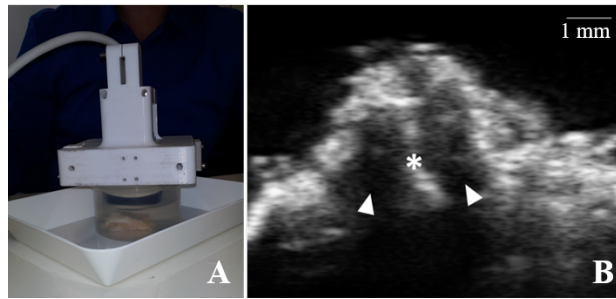


Fig. 6. Evaluation of feasibility of US imaging of ex-vivo human cochlea performed at our lab. with an Ultrasonix 14MHz US Probe. A. View of our experimental set-up. B. 2D B mode view of the cochlea obtained with our setup showing the basal turn (white arrowheads) and the modiolus (star).

We conclude that echography combined with registered computed tomography makes possible intraoperative imaging of the cochlea with accurate identification and location of structures of interest like the scalae vestibuli and tympani and the basilar membrane. This offers extremely promising perspectives for inner ear surgery procedures. Yet, to get closer to clinical conditions, further studies will aim at designing a dedicated miniaturized US-probe and extending this approach on cadaver temporal bones.

### CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

### ACKNOWLEDGMENTS

This work was supported by the French “Fondation pour l’Audition” and by the French National Agency for Research (Agence Nationale pour la Recherche, ANR) within the Investissements d’Avenir Program (Labex CAMI, ANR-11-LABX0004, the Equipex ROBOTEX, ANR-10-EQPX-44-01). We thank Dr Renaud Lebrun (Montpellier Rio Imaging platform and LabEx CeMEB - Institut des Sciences de l’Evolution - Université de Montpellier - CNRS UMR 5554- Campus Triolet Place Eugène Bataillon 34095 Montpellier Cedex 5) for his contribution in the  $\mu$ CT acquisitions and Dr Pierre Sicard (IPAM platform - CHU Arnaud De Villeneuve - UMR INSERM 1046 - CNRS 9214 - 371 Avenue du Doyen Gaston Giraud 34295 Montpellier Cedex) for his contribution in the HFUS acquisitions.

### REFERENCES

- <sup>1</sup>Bezanson, A., R. Adamson, and J. A. Brown. Fabrication and performance of a miniaturized 64-element high-frequency endoscopic phased array. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 61:33–43, 2014.
- <sup>2</sup>Braun, K., F. Böhnke, and T. Stark. Three-dimensional representation of the human cochlea using micro-computed tomography data: Presenting an anatomical model for further numerical calculations. *Acta Otolaryngol. (Stockh.)* 132:603–613, 2012.
- <sup>3</sup>Brendel, B., S. W. A. Rick, M. Stockheim, and H. Ermert. Registration of 3D CT and Ultrasound Datasets of the Spine using Bone Structures. *Comput. Aided Surg.* 7:146–155, 2002.
- <sup>4</sup>Brounstein, A., I. Hacihaliloglu, P. Guy, A. Hodgson, and R. Abugharbieh. Towards Real-Time 3D US to CT Bone Image Registration Using Phase and Curvature Feature Based GMM Matching. *Med Image Comput Comput Assist Interv.* 6891:235-242, 2011.

- <sup>5</sup>Brown, J. A., Z. Torbatian, R. B. Adamson, R. Van Wijhe, R. J. Pennings, G. R. Lockwood, and M. L. Bance. High-Frequency Ex vivo Ultrasound Imaging of the Auditory System. *Ultrasound Med. Biol.* 35:1899–1907, 2009.
- <sup>6</sup>Buytaert, J., S. Johnson, M. Dierick, W. Salih, P. Santi. MicroCT versus sTSLIM 3D Imaging of the Mouse Cochlea. *J. Histochem. Cytochem.* 61:382-95, 2013.
- <sup>7</sup>De Seta, D. Quality of insertion in cochlear implants: a clinical and temporal bone study, 2016.
- <sup>8</sup>Fuerst, B., W. Wein, M. Müller, and N. Navab. Automatic ultrasound–MRI registration for neurosurgery using the 2D and 3D LC2 Metric. *Med. Image Anal.* 18:1312–1319, 2014.
- <sup>9</sup>Goksu, N., N. Karademir, R. Haziroglu, I. Bayramoglu, Y. Kemaloglu, and N. Akyeldiz. Anatomy of the Guinea Pig Temporal Bone. *Ann. Otol. Rhinol. Laryngol.* 101:699–704, 1992.
- <sup>10</sup>Hacihaliloglu, I., D. R. Wilson, M. Gilbert, M. A. Hunt, and P. Abolmaesumi. Non-iterative partial view 3D ultrasound to CT registration in ultrasound-guided computer-assisted orthopedic surgery. *Int. J. Comput. Assist. Radiol. Surg.* 8:157–168, 2013.
- <sup>11</sup>Kakigi, A., A. Kakigi, Y. Takubo, N. Egami, A. Kashio, M. Ushio, T. Sakamoto, S. Yamashita, and T. Yamasoba. Evaluation of the Internal Structure of Normal and Pathological Guinea Pig Cochleae Using Optical Coherence Tomography. *Audiol. Neurotol.* 18:335–343, 2013.
- <sup>12</sup>Landry, T., G. Earle, J. Brown, M. Bance. Real-time intracochlear imaging of automated cochlear implant insertions in whole decalcified cadaver cochleas using ultrasound. *Cochlear Implants Int.* 19:255-267, 2018.
- <sup>13</sup>Liu, W., F. Atturo, R. Aldaya, P. Santi, S. Cureoglu, S. Obwegeser, R. Glueckert, K. Pfaller, A. Schrott-Fischer, and H. Rask-Andersen. Macromolecular organization and fine structure of the human basilar membrane - Relevance for cochlear implantation. *Cell Tissue Res.* 360:245–262, 2015.
- <sup>14</sup>Markelj, P., D. Tomaževič, B. Likar, and F. Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Med. Image Anal.* 16:642–661, 2012.
- <sup>15</sup>Miracle, A. C., and S. K. Mukherji. Conebeam CT of the Head and Neck, Part 2: Clinical Applications. *Am. J. Neuroradiol.* 30:1285–1292, 2009.
- <sup>16</sup>Moghari, M. H., and P. Abolmaesumi. Point-Based Rigid-Body Registration Using an Unscented Kalman Filter. *IEEE Trans. Med. Imaging* 26:1708–1728, 2007.
- <sup>17</sup>Nguyen, Y., G. Kazmitcheff, D. De Seta, M. Miroir, E. Ferrary, and O. Sterkers. Definition of Metrics to Evaluate Cochlear Array Insertion Forces Performed with Forceps, Insertion Tool, or Motorized Tool in Temporal Bone Specimens. *Biomed Res Int.*, 2014.

- <sup>18</sup>Pandey, P., P. Guy, A. J. Hodgson, and R. Abugharbieh. Fast and automatic bone segmentation and registration of 3D ultrasound to CT for the full pelvic anatomy: a comparative study. *Int. J. Comput. Assist. Radiol. Surg.* 13:1515–1524, 2018.
- <sup>19</sup>Pau, P. D. >Hans W., E. Lankenau, T. Just, D. Behrend, and G. Hüttmann. Optical coherence tomography as an orientation guide in cochlear implant surgery? *Acta Otolaryngol. (Stockh.)* 127:907–913, 2007.
- <sup>20</sup>Powell, M. J. D. The BOBYQA algorithm for bound constrained optimization without derivatives. Report DAMTP, University of Cambridge, 2009.
- <sup>21</sup>Rasouljan, A., P. Mousavi, M. H. Moghari, P. Foroughi, and P. Abolmaesumi. Group-wise feature-based registration of CT and ultrasound images of spine. *Proc SPIE Int Soc Opt Eng.* 7625, 2010.
- <sup>22</sup>Shin, K.-J., J.-Y. Lee, J.-N. Kim, J.-Y. Yoo, C. Shin, W.-C. Song, and K.-S. Koh. Quantitative Analysis of the Cochlea using Three-Dimensional Reconstruction based on Microcomputed Tomographic Images. *Anat. Rec.* 296:1083–1088, 2013.
- <sup>23</sup>Snels, C., J. Inthout, E. Mylanus, W. Huinck, I. Dhooge. Hearing Preservation in Cochlear Implant Surgery: A Meta-Analysis. *Otol Neurotol.* 40:145-153, 2019.
- <sup>24</sup>Torbatian, Z., R. Adamson, R. van Wijhe, R. Pennings, M. Bance, and J. Brown. Imaging the auditory system: A new application of high-frequency ultrasound. *IEEE Int Ultrason Symp.* pp. 236-239, 2009.
- <sup>25</sup>Wein, W., S. Brunke, A. Khamene, M. R. Callstrom, and N. Navab. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Med. Image Anal.* 12:577–585, 2008.
- <sup>26</sup>Wein, W., A. Karamalis, A. Baumgartner, and N. Navab. Automatic bone detection and soft tissue aware ultrasound–CT registration for computer-aided orthopedic surgery. *Int. J. Comput. Assist. Radiol. Surg.* 10:971–979, 2015.
- <sup>27</sup>Winter, S., B. Brendel, I. Pechlivanis, K. Schmieder, and C. Igel. Registration of CT and Intraoperative 3-D Ultrasound Images of the Spine Using Evolutionary and Gradient-Based Methods. *IEEE Trans. Evol. Comput.* 12:284–296, 2008.
- <sup>28</sup>Yan, C. X. B., B. Goulet, J. Pelletier, S. J.-S. Chen, D. Tampieri, and D. L. Collins. Towards accurate, robust and practical ultrasound-CT registration of vertebrae for image-guided spine surgery. *Int. J. Comput. Assist. Radiol. Surg.* 6:523–537, 2011.
- <sup>29</sup>Zitová, B., and J. Flusser. Image registration methods: a survey. *Image and Vision Comput.* 21:977-1000, 2003.