Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Cochlear implant phantom for evaluating computed tomography acquisition parameters

Srijata Chakravorti
Brian J. Bussey
Yiyuan Zhao
Benoit M. Dawant
Robert F. Labadie
Jack H. Noble

Cochlear implant phantom for evaluating computed tomography acquisition parameters

Srijata Chakravorti,^{a,*} Brian J. Bussey,^b Yiyuan Zhao,^a Benoit M. Dawant,^a Robert F. Labadie,^c and Jack H. Noble^a

^aVanderbilt University, Department of Electrical Engineering and Computer Science, Nashville, Tennessee, United States

^bMedical Center North, Department of Radiology, Nashville, Tennessee, United States

^cVanderbilt University Medical Center, Department of Otolaryngology—Head and Neck Surgery, Nashville, Tennessee, United States

Abstract. Cochlear implants (CIs) are surgically implantable neuroprosthetic devices used to treat profound hearing loss. Recent literature indicates that there is a correlation between the final intracochlear positioning of the CI electrode arrays and the ultimate hearing outcome of the patient, indicating that further studies to better understand the relationship between electrode position and outcomes could have significant implications for future surgical techniques, array design, and processor programming methods. Postimplantation high-resolution computed tomography (CT) imaging is the best modality for localizing electrodes and provides the resolution necessary to visually identify electrode position, although with an unknown degree of accuracy depending on image acquisition parameters, like the hounsfield unit (HU) range of reconstruction, orientation, radiation dose, and image resolution. We report on the development of a phantom and on its use to study how four acquisition parameters, including image resolution and HU range of reconstruction, affect how accurately the true position of the electrodes can be found in a dataset of CT scans acquired from multiple helical and cone beam scanners. We also show how the phantom can be used to evaluate the effect of acquisition parameters on automatic electrode localization techniques. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.4.4.045002](https://doi.org/10.1117/1.JMI.4.4.045002)]

Keywords: cochlear implants; phantom; computed tomography; image quality.

Paper 17103PRR received Apr. 15, 2017; accepted for publication Oct. 27, 2017; published online Nov. 16, 2017.

1 Introduction

Cochlear implants (CIs) are neuroprostheses designed to restore hearing to patients with profound sensorineural hearing loss. They induce the sensation of sound by directly stimulating the surviving spiral ganglion nerve cells of the inner ear. An electrode array that is surgically implanted into the inner ear (Fig. 1) is connected by a wire leading to an internal receiver coil embedded under the skin at the surface of the skull. A CI processor worn behind the ear transforms sound detected by a microphone into electrical signals to be sent through a transmitter coil to the internal receiver and implanted electrodes. Outcomes are remarkable for most recipients, with average speech recognition approximately at 60% and 70% correct for consonant-nucleus-consonant word (CNC) and AzBio sentence recognition tests, respectively.^{1–5} Despite this success, outcomes are highly variable, and a significant number of recipients still experience poor speech recognition outcomes.

The position of the electrode array within the cochlea is not generally known intraoperatively because the surgeon must thread the array through a very small opening into the cochlea [usually the “round window (RW) membrane”], and it is not possible to visualize or estimate the final position of the array in the cochlea. However, recent studies have relied on postimplantation imaging studies to detect final electrode position and have indicated that there is a correlation between the position of the electrodes relative to intracochlear anatomy and the ultimate hearing outcome of the patient.^{6–13} For example, it is now well-known that arrays that cross the basilar membrane separating

the scala tympani and scala vestibuli (see Fig. 2) cavities of the cochlea are associated with poorer outcomes.^{10–13} However, further study of the relationship between position and outcomes is warranted because: (a) there are still conflicting conclusions among studies about which other specific factors of positioning affect hearing outcomes. For example, Holden et al.¹¹ show that a deeper insertion negatively affects outcome. This, however, contradicts previous studies^{8,12} that have reported a positive correlation between speech recognition and depth of insertion. And (b) most studies have not had access to a set of accurate automatic analysis tools to localize the anatomy and electrodes and thus have only analyzed small datasets, focusing only on the coarse measures of electrode position that can be manually quantified, such as whether or not an obvious scalar translocation occurs. Our long-term goal is to perform a large-scale cross-institutional study to more comprehensively and precisely analyze the relationship between electrode position and outcomes than has ever been done before. We will collect precise measures of electrode position, such as the distance from each electrode to the modiolus, the distance from each electrode to the basilar membrane, and the angular insertion depth of each electrode relative to the RW membrane (see Fig. 2). This project is enabled by algorithmic techniques we have developed that permit accurately identifying the anatomical structures that are not directly visible in Fig. 2 and methods to identify each of the electrodes,^{14,15} which would otherwise require substantial effort to manually identify. Because we aim to perform these analyses with thousands of images of subjects that have been acquired from a large number of institutions, it will involve processing

*Address all correspondence to: Srijata Chakravorti, E-mail: srijata.chakravorti@vanderbilt.edu

Fig. 1 CI array implanted in the left ear. In order, the yellow structure is the inner ear, the blue and red surfaces are the scala vestibuli and scala tympani intracochlear cavities, respectively, and the green structure is the modiolus. CI electrodes are shown in black, and the plastic carrier is shown in gray.

Fig. 2 Measurements of electrode position demonstrated on a specimen implanted with an Advanced Bionics Mid-Scale array. The red structure is the scala tympani, the blue structure is the scala vestibuli, and the green structure is the modiolus (following the color scheme of Fig. 1). The measures of interest to us are (1) angular insertion depth relative to the RW membrane, (2) distance of each electrode to the modiolus, and (3) distance of each electrode to the basilar membrane. Scalar translocation of the electrode array can also be observed as the array is initially in the scala tympani (red) but then crosses the basilar membrane into the scala vestibuli (blue) at greater insertion depths.

images acquired from different scanners and with different imaging protocols that we do not control. Since the accuracy of our algorithmic techniques will be highly dependent on the imaging protocol, we are currently interested in learning the relationship between common acquisition parameters and the accuracy of our algorithms. In this study, we have developed a phantom that will permit evaluating the relationship between acquisition parameters and the accuracy of our electrode localization techniques. Results from this phantom will permit controlling for expected electrode localization accuracy as a function of acquisition parameters in our dataset and to determine if images with certain acquisition parameters should be withheld from our analyses. In a future study, we plan to develop a phantom for our anatomy localization algorithms, which are also sensitive to acquisition parameters.

While the focus of our study is evaluating how algorithms perform on images that are acquired at many sites where we do not control acquisition parameters, it is important to note that many studies have been aimed at developing CI-specific image acquisition protocols. For example, it has been shown that CI electrodes can be localized even in simple two-dimensional radiographs of the skull, provided an appropriate projection is chosen. This is the principle of cochlear view,¹⁶ which allows localization of the individual electrodes to a good degree of accuracy, although it is limited in describing the positions of the electrodes in terms of the surrounding anatomy, which is critical for the analyses we would like to accomplish. Using multiple radiographs can improve the accuracy of electrode localization somewhat by adding a third dimension. For example, x-ray stereophotogrammetry¹⁷ can fuse information from a stereo

pair of radiographs to locate the positions of the electrodes in three-dimensions (3-D). Computed tomography (CT) imaging offers the best view of electrode position by providing a 3-D view of the temporal bone area on which the electrode positions can be estimated. Techniques for electrode localization have been explored on cone beam CTs¹⁸ as well as spiral CTs.^{6-8,19} Currently, postimplantation high-resolution CT imaging is the modality accepted as the standard for this application and provides the resolution necessary to visually identify electrode arrays. Standard CT protocols can be prone to beam hardening artifacts when radiodense materials like CI electrodes are present in the field-of-view. However, digital image processing methods exist to improve the image quality available from commercial scanners. For example, iterative reconstruction methods can reduce beam hardening effects and image deblurring^{20,21} can optimize the resolution of spiral CT slices retrospectively, thereby allowing more accurate electrode localization.

As discussed above, using high-quality CT images available today, it is possible for experienced surgeons to make rough estimations of angular depth, modiolar proximity, and scalar translocation or lack thereof. However, this manual approach is impractical for analyzing large datasets of CT images, and we aim to achieve more precise localization than is feasible manually and more than what would be possible by visual examination. Thus, our goal in this study is to develop and use a phantom to evaluate the relationship between CT acquisition parameters and the accuracy of our electrode localization algorithms because, while we have shown our methods work well with high-quality CT scans, the accuracy of our techniques with the range of CTs that have been acquired at many outside institutions is unknown.

The primary imaging parameters of interest are the CT image hounsfield unit (HU) reconstruction range, resolution (parameterized by the voxel diagonal length), CT dose (parameterized by the exposure per rotation), and orientation of the subject in the scanner. Reconstructions with extended HU have a greater intensity encoding range and are expected to provide greater contrast to the metallic electrodes against the surrounding high-intensity bony tissue or noise. Resolution is an important parameter given that the CI contacts typically occupy less than a voxel's volume, and the spacing between contacts is often close

to a voxel length even in high resolution scans. A lower dose CT image might have significant beam hardening artifacts that increase noise and make it difficult to distinguish electrodes from each other and from dense bone. Beam hardening is also influenced by the orientation of the head in the scanner because some orientations may lead to beams that pass through dense structures, such as the teeth or receiver coils in addition to the CI electrodes, which would worsen the artifact. By varying these four CT parameters and estimating the resulting localization errors on the phantom, we can obtain bounds on the expected error as a function of these CT parameters. While it is obvious that the highest quality parameters should permit the best possible localization, our aim is not to find an optimized protocol but rather to develop a system for calculating expected localization error bounds on any CT acquisition protocol for any given localization method. A secondary benefit of this phantom is that it could significantly aid the development and implementation of highly accurate automatic electrode localization methods, which would permit large-scale analysis of the relationship between electrode position and outcomes. A better understanding of this relationship could have implications for future hardware design and improvements in surgical techniques. This may also have important implications for clinical applications, such as techniques proposed for selecting personalized CI processor settings for patients based on electrode position to improve hearing outcomes.²²

2 Methods

A phantom was created using a skull implanted with CIs in its left and right ears. Several scans of this phantom were acquired from different scanners and using different acquisition parameters. Each image provides an estimate of the location of each electrode. Every electrode was localized in each image by averaging multiple manual localizations of the centroid of the electrode by experts. The different images were then coregistered to a chosen reference image, in which a ground truth was defined by experts, also through repeated electrode localization measurements. Localization errors were calculated with reference to this ground truth. These procedures, summarized in Fig. 3, are described in the following subsections.

Fig. 3 Method of processing the phantom images to obtain localization errors.

2.1 Construction of the Phantom

The phantom was created using a cadaveric skull implanted with CIs in both ears. The left ear of the skull was implanted with an Advanced Bionics (AB) (Valencia, California) 1J device (A1), whereas the right ear was implanted with an AB Mid-Scala device (A2). The skull was then enclosed in a plexiglass box (Fig. 4). The box and cranial cavity were subsequently filled with silicone (Ecoex 00-10 supersoft silicone rubber, Reynolds Advanced Materials) after removing air bubbles using a vacuum chamber, and the silicone was allowed to cure. Silicone rubber is used to mimic the properties of soft tissue in a CT image. Both A1 and A2 are arrays with 16 active contacts and 1 nonstimulating contact at the proximal end of the array that serves as a marker, which indicates that the full insertion depth of the array has been reached when the marker reaches the cochlea entry site. A1 has a larger electrode spacing between active contacts (1.1-mm pitch) than A2 (0.95-mm pitch), and this leads to different degrees of image contrast between the contacts in the phantom's images.

For registration purposes, eight fiducial marker spheres from Brainlab (Munich, Germany), each of diameter 8 mm, were embedded and glued into the skull and were spaced so that the centroid of the fiducial distribution approximately corresponds to the centroid of the skull. These markers can be seen in blue in Fig. 4 and have high intensity in CT images sufficiently contrasted against surrounding tissues to be reliably used for point-based registration.

2.2 Dataset

A dataset of 14 full head CT images of the phantom was created using different acquisition parameters in different scanners (Table 1), providing us with 28 images of CI arrays. Philips scanners are abbreviated as “PhCT,” Siemens scanners as “SiCT,” and Xoran scanners as “XCT.” All images marked with a superscript of (E) had extended HU reconstructions, and all other images have limited HU ranges. In every case, a filtered back-projection algorithm was used for the CT reconstruction. Since the layer of silicone rubber surrounding the skull is thicker than skin would be in a live patient, radioabsorption is greater than it would be for a patient, and the images contain more noise than

Fig. 4 CI phantom shown during construction with the plexiglass box only partly filled with silicone rubber. Most of the inferior half of the skull is still visible above the silicone.

standard clinical images. This ensures that the phantom represents the worst case scenario and obtains the upper bounds of localization error a particular localization algorithm can generate.

For the next sections, we group the images having resolution between 0.35 and 0.45 mm as “high res,” those in the range 0.45 to 0.55 mm as “mid res,” those at 0.5 to 0.65 mm as “low res,” and finally, the ones more than 0.65 mm as “very low res.” In a similar way, “low dose” refers to less than 250 mAs exposure value, “mid dose” refers to an exposure range of 250 to 375 mAs, and “high dose” refers to an exposure of 375 to 500 mAs.

2.3 Localization of Cochlear Implant Arrays

The centroid of each electrode in the CI arrays was localized multiple times by three expert raters on all the images in the dataset based on their intensity difference from surrounding bone and tissue. Each electrode roughly has the appearance of a high intensity blob in the image. The centroid of this blob can be visualized in three orthogonal views—axial, sagittal and coronal—to localize the electrode. This process is then repeated one to four times by each of the experts for each electrode for a total of 6 to 10 localizations of each electrode by all experts. It takes an expert about 2 to 3 min to complete a single localization of the electrode array. Thus, performing multiple localizations is a very time consuming process. The averages of the multiple sessions were used to represent the best estimation of electrode position for the respective images. Based on the image quality of $V_2^{(E)}$ and the fact that the raters had the best repeatability in determining the electrode position in this image, it was chosen

Table 1 Image dataset.

Number	Scanner	Image	Voxel diagonal (mm)	Exposure (mAs)	kVp
1	PhCT Brilliance 64	$V_1^{(E)}$	0.462	500	120
2	PhCT Brilliance 64	$V_2^{(E)}$	0.478	500	120
3	PhCT Brilliance 64	$V_3^{(E)}$	0.478	500	120
4	PhCT Brilliance 64	$V_4^{(E)}$	0.468	351	120
5	PhCT Brilliance 64	$V_5^{(E)}$	0.563	350	120
6	SiCT SOMATOM definition flash	$M_1^{(E)}$	0.581	395	120
7	XCT MiniCAT	$X_1^{(E)}$	0.346	58.8	125
8	XCT MiniCAT	$X_2^{(E)}$	0.346	58.8	125
9	XCT xCAT	$X_3^{(E)}$	0.693	57.6	120
10	PhCT iCT 128	V_1	0.446	500	120
11	PhCT iCT 128	V_2	0.460	280	120
12	PhCT iCT 128	V_3	0.591	280	120
13	PhCT iCT 128	V_4	0.540	280	120
14	SiCT SOMATOM definition flash	M_1	0.388	395	120

to serve as the ground truth of the dataset. $V_2^{(E)}$ will thus be the reference for registration (see Table 1) and will henceforth be referred to as V_{ref} . The ground truth of the CI arrays is defined as the average of 10 manual localizations by experts on this reference image.

2.4 Registration and Registration Accuracy

2.4.1 Point-based rigid body registration

Our first approach to coregister all the images was point registration using the fiducial markers embedded in the skull phantom. The eight markers were manually localized by an expert rater. Using their positions in the images, 10 out of the 14 CT scans are registered to the reference image V_{ref} by a point-based rigid body registration technique.²³ The other three images $[X_1^{(E)}, X_2^{(E)}, \text{ and } X_3^{(E)}]$ cannot be registered with this method because some of the markers are not visible in the field of view. This registration step ensures that the automatically localized CI arrays are effectively in the same image space as the ground truth arrays, allowing a direct comparison of the automatic localization procedure against the ground truth.

Because this registration is a crucial step for the performance evaluation, the errors in this step must also be taken into consideration. For unbiased results, the registration errors should be much smaller than the localization errors on average. The fiducial markers are 8 mm in diameter, which is 16 times larger in dimension than the slice width in the coarsest image, which is important because larger fiducials lead to lower fiducial localization error (FLE), leading to lower target registration errors (TRE). With fiducials of this size, it can be reasonably assumed that the localization of the marker centroids is robust against the image quality variations that affect the CI electrodes, which occupy less than a voxel in size. We, therefore, assume that the errors associated with locating the centroids of the fiducials, the FLE, in all the scans are isotropic, independent, and normally distributed. We use a standard registration error analysis approach²⁴ to calculate the overall fiducial registration error (FRE):

$$\langle \text{FLE}^2 \rangle \approx \frac{1}{M} \sum_{m=1}^M \frac{N}{N-2} \times \text{FRE}^2(m), \quad (1)$$

where M is the number of pairwise registrations and N is the number of fiducial markers, which remains constant for all registrations m . With this estimate of $\langle \text{FLE}^2 \rangle$, the TRE at the positions of the CI contacts, marked by the ground truth arrays, can be calculated:

$$\langle \text{TRE}^2(\mathbf{r}) \rangle \approx \frac{1}{N} \left(1 + \frac{1}{3} \sum_{k=1}^3 \frac{d_k^2}{f_k^2} \right) \times \langle \text{FLE}^2 \rangle, \quad (2)$$

where $\mathbf{r}$ is the position of the target in image space, d_k is the distance of the target calculated from the principal axis k , and f_k is the root mean square (rms) distance of all N fiducials from the same k 'th principal axis. The square root of $\langle \text{TRE}^2(\mathbf{r}) \rangle$, or the rms(TRE), serves as our estimate of registration accuracy at the position $\mathbf{r}$.

For all 34 possible locations of the contacts on the two sides of the skull, rms(TRE) is approximately 0.15 mm in every case.

2.4.2 Mutual information based rigid body local image registration

The point-based rigid body registration is guaranteed to provide a globally well registered image. However, our specific aim is to have good registration around each inner ear, which might be better obtained by local registration algorithms. Thus, we have tested an intensity-based local registration method in this work. All the images are first rigidly registered to V_{ref} using a mutual information (MI)²⁵ based image transformation, followed by localized rigid MI registration centered in a region of interest (RoI) of approximately $35 \times 35 \times 35 \text{ mm}^3$ around the inner ear. Once registered, the respective arrays are referred to with a subscript of (i, ref) , indicating registration of image i to V_{ref} .

Figure 5 provides a visual verification of the quality of this registration step. The difference image calculated between the point registered image and the reference image slice [Fig. 5(c)] shows larger areas of intensity disagreements (regions in shades of gray and white) compared to the difference image calculated between the locally registered image and the reference image slice [Fig. 5(e)]. The single large bright spot at the middle of the top half of Fig. 5(e) corresponds to an electrode in the image and is present in the difference image [Fig. 5(c)] as well. Even with a highly accurate registration, an electrode will have high intensity in a difference image because V_{ref} is an extended HU image and assigns much higher values of intensity to the metallic electrodes than the limited HU image V_4 . In Fig. 5(f), the locally registered array agrees more closely with the ground truth than the array obtained from point registration. It is visually evident that the local image registration yields better results than the global registration, and therefore, this is our choice of registration method. A disadvantage of using intensity-based methods is that there is no way to calculate TRE, which we must estimate to determine the effects of registration in our subsequent error analyses. Therefore, we will use the TRE estimates of Sec. 2.4.1 as an error bound to judge the quality of our intensity-based registration. Because our local image registration is more accurate than the point-based registration, the estimate of $\text{rms}(\text{TRE}) = 0.15 \text{ mm}$ obtained from Sec. 2.4.1 serves as an upper bound on our current registration error.

2.5 Electrode Localization Errors

The distance between an electrode localized in an image and its corresponding ground truth localization in the registered reference image V_{ref} defines what we call the “image-based” localization error for the electrode. Even though we calculate this distance on the basis of manual localizations, we use multiple averages from the manual results of different experts to virtually negate the procedural errors, so that the “image-based” localization error only reflects the error intrinsically due to image quality. The overall image-based localization errors are defined as the array of distances between each of the electrodes registered to the space of V_{ref} and the corresponding ground truth electrodes defined in V_{ref} . The array of distances is defined as $\{\Delta_{m(i,j)}^s\}_{(j=1)}^N$, where the distance for each electrode j in array s in an image i is

$$\Delta_{m(i,j)}^s = \|\text{Image}^s(j)_{(i,\text{ref})} - \text{Truth}^s(j)_{(\text{ref},\text{ref})}\|. \quad (3)$$

The subscript m in the equation is used to differentiate this measure of image-based localization error from the automatic algorithm localization errors that will be discussed below in

Fig. 5 Visual verification of registration accuracy in a representative scan (V_4). (a–e) The red isocontours were generated on the reference image (a) and indicate the edges of the bones in that panel. The isocontours were projected to the other panels to show that the locally registered version [(d) and (e)] closely matches (a), while the point registered image [(b) and (c)] has a slight but significant difference from the reference image. (a) Axial CT slice of V_{ref} with RoI marked. (b) Left ear of V_4 registered to V_{ref} using global point-based registration. (c) Difference image calculated between images (a) and (b). (d) Left ear of V_4 registered to V_{ref} using local image based registration. (e) Difference image calculated between images (a) and (d). (f) Registered arrays: the solid black line is the ground truth array A1, the dotted red line is the array of scan V_4 registered using the point-based registration, and the dashed blue line is the same array registered to V_{ref} using local image registration.

Sec. 3.7. For each scan, the errors can be considered to be drawn from a population that follows a continuous normal distribution with zero mean. Informal analysis of the correlations between the x , y , and z components of the errors of each electrode with other electrodes in the array confirmed the errors at each electrode are independent and statistical analysis can be carried out assuming independence accordingly. This also suggests that registration errors are negligible at the scale of the localization errors, since we expect that significant registration errors would have led to correlated directional errors. We conclude that bias due to the registration errors in measurement of localization errors is thereby not an issue here.

3 Results

Image-based localization errors averaged across all electrodes for both A1 and A2 for each CT image are shown in Fig. 6. The errors for $V_2^{(E)}$ are zero since it is the chosen reference image. In order to estimate the statistical significance of the effects of specific acquisition parameters on the image-based localization accuracy of the CI arrays, we carried out a series of statistical tests on these errors. Nine of the tests were families of two sample Welch's t -tests,²⁶ and one was a one way analysis of variance.²⁷ Welch's unequal variances t -tests were preferred to student's t -tests because we used unequal sample sizes in many of the tests, and it is reasonable to assume that the variances of errors from different images are different.

Along with the statistical significance of the effects, the effect sizes for each test are also reported in terms of the Cohen's d values²⁸ of the tests, which are just the ratio of the difference of group means to the pooled standard deviation, to

be better able to compare how much each parameter affects the localization accuracy.

The errors are grouped by the values of the three major acquisition parameters of the original scans—exposure, voxel diagonal lengths, and whether the reconstruction was carried out using an extended or a limited HU range—using the ranges of values in Sec. 2.2. Each of the test groups is designed in such a way that the only variable that changes between the chosen scans is the acquisition parameter whose effect we are testing for. Sections 3.1–3.5 exclusively deal with the outcomes of the tests on the image-based localization errors, with the results listed in the “image-based” sections of Tables 2–7. Later, in Sec. 3.7, we will look at the effects of the same acquisition parameters on the arrays localized by two automatic electrode localization algorithms, and the results are discussed in the “Auto I” and “Auto II” sections of Tables 2–7. Any “—” in Tables 2–7 signifies that sufficient data are not present to carry out the comparison. The null hypothesis is that the means of the different groups do not vary to within the significance level of 0.05. All the calculations are carried out in MATLAB (Mathworks, Natick, Massachusetts). Calculation of confidence intervals for measures of effect size was carried out using the measures of effect size²⁹ MATLAB toolbox.

3.1 Effect of Spacing Between Electrodes

In order to estimate the effect of electrode spacing, Bonferroni corrected t -tests were performed on the image-based localization errors, which showed that there is a significant effect of the array model on the accuracy of image-based localization

Fig. 6 Image-based localization errors across A1 and A2.

Table 2 Effect of electrode spacing on localization errors. The error subgroups being compared are mentioned in parentheses under the column “factor.” If a statistically significant effect ($p < 0.05$) is present, the group means are specified in mm, while “no” indicates the absence of a statistical effect.

Factor	Test conditions	Image-based	Auto I	Auto II
Spacing between electrodes (A1, A2)	All resolution, dose and HU ranges	(0.0835, 0.1442)	No	No

Table 3 Effect of HU range during reconstruction. Significant p -values are in bold text.

Factor	Test conditions	Localization method	Array type	t -statistic	p -value	Group means	Cohen's d
HU range (Ext HU, Lim HU)	High dose, mid res	Image-based	A1	$t(22.31) = -3.87$	< 0.001	(0.048, 0.081)	1.341
			A2	$t(19.01) = -4.62$	< 0.001	(0.054, 0.123)	1.745

(“image-based” column of Table 2). We can conclude from a one-tailed corrected t -test that the errors for array A2 are significantly worse than the errors for array A1. As shown in parentheses in the “image-based” column of Table 2, the mean of the errors for array A2 is around twice that of A1. This result is intuitive because the closer electrode spacing for A2 provides less image contrast between electrodes than A1.

3.2 HU Range of Reconstruction

Out of the 14 scans, 9 [$V_1^{(E)} - V_5^{(E)}$, $M_1^{(E)}$, $X_1^{(E)} - X_3^{(E)}$] were reconstructed with an extended HU (Ext HU) range, whereas the other 5 ($V_1 - V_4$, M_1) were reconstructed with a limited HU (Lim HU) range. The HU range of reconstruction dictates the visible contrast of the metallic electrodes against the surrounding high intensity bony tissue or noise. Welch's t -test was carried out on the image-based localization errors $\Delta_{m(i,j)}^s$ as per Table 3, and changing the HU range was found to affect determination of the correct CI electrode locations for both A1 and A2. The group mean of errors for a limited HU image is nearly

twice that of extended HU for A1 and more than twice for A2 at the same high dose and mid resolution.

3.3 Orientation

Images $V_1^{(E)}$, $V_2^{(E)}$, and $V_3^{(E)}$ were acquired with identical parameters but with different orientations of the phantom skull in the scanner. Orientation controls the severity of the beam hardening artifacts around the electrodes depending on whether dense structures like teeth or receiver coils lie in the path of the CT beams before striking the electrodes—severe beam hardening artifacts might form that could make it harder to localize the electrodes.

Welch's t -test carried out on the $\Delta_{m(i,j)}^s$ showed a significant error change only for A2, shown in Table 4, but this is a relatively small change in terms of actual group means.

3.4 Resolution

The length of a voxel diagonal in a CT acquisition is a direct measure of the resolution of an image. In most of the images

in our dataset, an electrode is physically smaller in volume than one voxel and generally appears to occupy no more than a few voxels in any image. Poorer resolutions cause partial volume effects to dominate, leading to more distortions in how the arrays are visualized in the final image.

Three Welch's t -tests on $\Delta_{m(i,j)}^s$ show that lower resolution is associated with poorer accuracy for localizing both the A1 and

A2 arrays under all conditions except one (A1 for limited HU, mid dose reconstructions). However, limited HU range can have somewhat similar effects on image quality as poor resolution. It is possible that in the case where change of resolution was not read as significantly affecting the localization accuracy, the stronger effects of a limited HU range occluded any effects due to low resolution.

Table 4 Effect of orientation. The “—” implies no single Cohen's d value that could be calculated since the test was an analysis of variance (ANOVA) between three groups. Significant p -values are in bold text.

Factor	Test conditions	Localization method	Array type	t -statistic or F -statistic	p -value	Group means	Cohen's d
Orientation	Ext HU, high dose, mid res	Image-based	A1	$t(31.91) = 0.58$	0.57	(0.050, 0.046)	-0.198
			A2	$t(31.93) = -2.78$	0.009	(0.044, 0.065)	0.954
		Auto I	A1	$F(2,48) = 0.01$	0.99	(0.147, 0.149, 0.150)	—
			A2	$F(2,48) = 0.87$	0.43	(0.181, 0.217, 0.157)	—
		Auto II	A1	$F(2,48) = 1.60$	0.21	(0.134, 0.176, 0.149)	—
			A2	$F(2,48) = 0.84$	0.44	(0.121, 0.148, 0.134)	—

Table 5 Effect of resolution. Incomplete data have been marked as “—,” implying that no test can be performed. Significant p -values are in bold text.

Factor	Test conditions	Localization method	Array type	t -statistic	p -value	Group means	Cohen's d
Resolution (Mid res, low res)	Lim HU, mid dose	Image-based	A1	$t(19.36) = -1.54$	0.139	(0.098, 0.131)	0.577
			A2	$t(18.82) = -3.30$	0.004	(0.154, 0.371)	1.252
		Auto I	A1	$t(24.75) = -41.54$	<0.001	(0.176, 4.357)	14.250
			A2	—	—	—	—
		Auto II	A1	—	—	—	—
			A2	—	—	—	—
Resolution (mid res, low res)	Ext HU, Mid dose	Image-based	A1	$t(30.79) = -3.92$	<0.001	(0.058, 0.096)	1.344
			A2	$t(16.21) = -6.02$	<0.001	(0.066, 0.344)	2.067
		Auto I	A1	$t(19.72) = -1.28$	0.22	(0.142, 0.186)	0.438
			A2	$t(30.37) = -2.34$	0.03	(0.147, 0.215)	0.806
		Auto II	A1	$t(18.77) = -4.90$	<0.001	(0.125, 0.276)	1.679
			A2	$t(20.07) = -2.75$	0.01	(0.128, 0.230)	0.942
		Image-based	A1	$t(19.56) = -5.84$	<0.001	(0.055, 0.109)	2.171
			A2	$t(20.26) = -5.07$	<0.001	(0.046, 0.106)	1.848
Resolution (high res, very low res)	Ext HU, Low dose	Auto I	A1	$t(24.69) = -0.50$	0.62	(0.090, 0.098)	0.164
			A2	$t(48.97) = -0.76$	0.45	(0.126, 0.144)	0.184
		Auto II	A1	$t(43.67) = 0.90$	0.37	(0.094, 0.084)	-0.238
			A2	$t(21.64) = -3.01$	<0.001	(0.079, 0.129)	1.057

Table 6 Effect of dose. Incomplete data have been marked as “—,” implying that no test can be performed. Significant p -values are in bold text.

Factor	Test conditions	Localization method	Array type	t -statistic	p -value	Group means	Cohen's d
Dose (high dose, mid dose)	Lim HU, mid res	Image-based	A1	$t(29.46) = -1.86$	0.07	(0.081, 0.105)	0.639
			A2	$t(22.59) = -1.26$	0.22	(0.123, 0.166)	0.432
		Auto I	A1	—	—	—	—
			A2	—	—	—	—
		Auto II	A1	—	—	—	—
			A2	—	—	—	—
Dose (high dose, mid dose)	Ext HU, Low res	Image-based	A1	$t(29.04) = -0.76$	0.45	(0.087, 0.096)	0.262
			A2	$t(24.94) = -3.81$	0.001	(0.143, 0.344)	1.307
		Auto I	A1	$t(29.26) = 0.71$	0.48	(0.225, 0.186)	-0.242
			A2	$t(21.18) = 1.47$	0.16	(0.305, 0.215)	-0.504
		Auto II	A1	$t(21.86) = -5.39$	<0.001	(0.102, 0.276)	1.848
			A2	$t(27.14) = -1.70$	0.1	(0.160, 0.230)	0.580
Dose (high dose, mid dose)	Ext HU, Mid res	Image-based	A1	$t(26.21) = -1.53$	0.14	(0.048, 0.058)	0.492
			A2	$t(47.03) = -1.96$	0.06	(0.054, 0.066)	0.498
		Auto I	A1	$t(31.76) = 0.49$	0.63	(0.149, 0.142)	-0.127
			A2	$t(49.43) = 1.47$	0.15	(0.185, 0.147)	-0.316
		Auto II	A1	$t(54.29) = 2.15$	0.04	(0.153, 0.125)	-0.444
			A2	$t(31.08) = 0.38$	0.70	(0.134, 0.128)	-0.1

Table 7 Standard linear model fit with intercept of average localization error values.

Variables	Estimated values of coefficients			
	Image-based (A1)	Image-based (A2)	Auto I	Auto II
(Intercept)	-0.00048	-0.1301	-3.6367	-0.036273
Resolution	0.23973	1.040524	11.079	0.49949
Dose	—	-0.0004	—	-0.000099
HU	-0.058276	-1.7414	-2.122	—

3.5 Dose

We use exposure values as indicative of the radiation dose since all images were acquired at similar tube voltages. Several different exposure levels were used for the acquisitions in our dataset (Table 1). A lower CT dose can lead to higher image noise, which in turn could introduce spurious noisy voxels that appear like electrodes and reduce the accuracy of localization.

Three Welch's t -tests for comparing the effects of dose on $\Delta_{m(i,j)}^s$ show that the effect of dose on the image-based localization accuracy of A1 and A2 is found to be significant in only one case out of six, and even then, it does not change the mean errors

as much as resolution. It is likely that the dose is not measured as having a significant effect in the first case because the poorer accuracy in that test can also be driven by the effects of limited HU. In the third case, the t values show that the effect of dose was sizeable but not enough to be significant at our sample sizes.

Figure 7 maps the effect sizes in terms of Cohen's d values of the different factors on image-based localization along with their 95% confidence intervals. The horizontal axis denotes each of the different tests carried out as listed in Tables 3–6. It can be concluded that while all factors affect localization errors to some degree, resolution and HU range have more impact in general than the other two factors.

Fig. 7 Effect size values for image-based localization errors across A1 and A2.

3.6 Modeling Acquisition Parameters versus Image-Based Localization Errors

Table 7 lists the estimated values of the intercept and coefficients for the significant factors when fitting a linear regression model to our average error data of the form $\text{error} = \alpha + \beta * \text{resolution} + \gamma * \text{dose} + \delta * [\text{HU}]$. In this model, “resolution” and “dose” are numerical variables, “HU” is a Boolean indicator, α is the estimated intercept, and β , γ , and δ are the estimated model coefficients. According to our fit, if we aim to only admit images where the image-based localization error of array A1 is 0.1 mm or less, and the image is a limited HU image with a mid-range dose of 350 mAs, the resolution must not exceed 0.42 mm in voxel diagonal length, whereas the resolution limit for an array A2 in a similar image would be 0.35 mm. Such a model can define a range of image parameters that are good enough for the purpose of limiting the error arising out of image quality.

3.7 Evaluation of Automatic Cochlear Implant Array Localization Using the Phantom

As an example of the phantom’s use as a validation tool, we compare the performance of two automatic electrode localization techniques on the same image set. The first method (Auto I) is a graph-based technique,¹⁵ whereas the second (Auto II) is a preliminary extension to this approach that we are currently developing. It fine-tunes the processing steps and optimization cost functions and is currently under evaluation. More detailed discussion of these two techniques would be out of scope for this paper since our aim is to develop the phantom as a generalized validation tool irrespective of how a particular image analysis technique actually works. In every CT scan, the electrode arrays were automatically localized using these two techniques. On average, it takes only about 10 s to automatically localize an array with either method. Also, automatic techniques can be executed without human intervention on large-scale datasets. Thus, the automatic methods are substantially faster and require much less effort than a manual segmentation approach. The designs of Auto I and II are such that they expect the electrodes to be brighter than bone and thus, they work best with extended HU images. Of the 28 images of the CI arrays, Auto I failed to produce any result only in one limited HU case (A2 of V_2). There were four other cases where the results were considered as “failures” by expert raters (A1 of V_1 and V_2 ; A2 of V_1 and V_4). Auto II failed on all limited HU images ($V_1 - V_4$, M_1), though this is unsurprising since this algorithm was fine-tuned for extended HU images and needs modification to handle limited HU images. Examples of results considered “good” and “failed” are demonstrated in Fig. 8. The failed results are not used in subsequent analyses. Overall, we obtain 23 automatic localization results from Auto I and 18 from Auto II.

Fig. 8 Good results (a and c) versus failed results (b and d). The solid black line is the ground truth array of V_{ref} , whereas the dashed red line is the automatically localized array registered to V_{ref} .

The error of automatic localization method “a” for every electrode j in each array s in scan i was defined similarly as for image-based localization errors:

$$\Delta_{a(i,j)}^s = \|\text{Automatic}^s(j)_{(i,\text{ref})} - \text{Truth}^s(j)_{(\text{ref},\text{ref})}\|.$$

Since image-based localization errors measure the error in the average of multiple expert localizations, they define the electrode localization error due to the imaging technique alone, whereas the automatic localization error also includes errors made by the algorithmic electrode localization process.

3.7.1 Automatic localization errors

Figures 9(a) and 9(b) show the rms errors over all electrodes for A1 and A2, respectively, in each image. Unsurprisingly,

Bonferroni corrected one-tailed t -tests statistically confirm that the automatic localization techniques add error to the image-based localization error. This error, contributed by the localization procedure, increases the net error in almost every case. The exceptions are the A2 array in the images $V_5^{(E)}$ and M_1 , where the random errors in the automatic localization tended to cancel out some of the image-based localization errors. However, the automatic methods provide very good results on the extended HU images, as rated by experts.

3.7.2 Effect of acquisition parameters

The effects of electrode spacing on automatic localization errors (using Bonferroni corrected t -tests) are shown in Table 2. It is not possible to test the effects of the HU range of reconstruction

Fig. 9 Automatic versus image-based localization errors. (a) $\Delta_{a(i,j)}^s$ and $\Delta_{m(i,j)}^s$ on A_1 and (b) $\Delta_{a(i,j)}^s$ and $\Delta_{m(i,j)}^s$ on A_2 .

Fig. 10 (a) Automatic localization errors for A₁ and A₂ (Auto I). (b) Automatic localization errors for A₁ and A₂ (Auto II).

on automatic localization techniques since the methods used did not work on enough limited HU images. Figures 10(a) and 10(b) show the errors of localization on the two types of arrays using the two methods, respectively. A Bonferroni corrected *t*-test shows no significant effect of the array model used for either method on the errors. In case of orientation, one way ANOVA carried out on the automatic localization errors $\Delta_{a(i,j)}^s$ showed no significant effects. The results are shown in Table 4. The sizes of the effects were not large enough to be considered significant.

T-tests on $\Delta_{a(i,j)}^s$ for Auto I show statistically significant effects of change of resolution on only two out of five of the conditions tested for, and three out of four test groups exhibited significant effects of resolution change for Auto II, with the lower resolution images producing the worse results in every

case (Table 5). Two *t*-tests on $\Delta_{a(i,j)}^s$ from both Auto I and Auto II show a significant effect of change of dose only for array A1 localized by Auto II. In every other case, the automatic localization results were unaffected by the dose changes tested (Table 6).

Figures 11 and 12 map the effect sizes of the tests on automatic localization errors in terms of their Cohen's *d* values and the 95% confidence intervals. The horizontal axis marks each of the different tests carried out as listed in Tables 5 and 6. The effect size of the first test on resolution in Fig. 11 is a clear outlier, which is obtained from the test of errors from Auto I on limited HU, mid dose images (Table 5). This was possibly due to quite a poor result on the limited HU low res image V_4 , which was nevertheless included because it was not a complete

Fig. 11 Effect size values for automatic localization errors across A1 and A2 (Auto I).

Fig. 12 Effect size values for automatic localization errors across A1 and A2 (Auto II).

failure. Although this was to be expected since Auto I was not designed to work for limited HU images, this effect of poor localization increases the apparent effect of resolution in this one test group. Even without counting this particular outlier,

resolution has a large and significant effect in at least one set of images for Auto I, whereas dose has no significant effect at all. In case of Auto II, the effect of resolution is much more pronounced, with three out of four test groups exhibiting large and significant effects. By comparison, only in one case, dose does have a large effect. We can thereby conclude that Auto I is, in general, robust to changes of dose and resolution, and the localization accuracy of Auto II is significantly affected by image resolution.

Standard linear regression models can be fitted to the observed errors for both Auto I and Auto II as previously done for image-based errors in Sec. 3.6. Unlike in the case of image-based errors, the automatic localization errors from arrays A1 and A2 can be pooled together because there is no significant effect of the array model on the errors. Auto I has the same model specifications as the image-based errors, whereas Auto II is fitted without the HU indicator. The coefficients are listed in Table 7, and we can determine acceptable ranges of parameters to limit the expected localization errors from either of these localization methods. For example, if we want to select images, which would have no more than 0.2-mm localization error from method Auto I, we would only accept extended HU images of resolution 0.54 mm or better and limited HU images of resolution 0.35 mm or better.

4 Conclusion

In this paper, we have described the development of and initial results from what is to the best of our knowledge the first CI phantom for evaluating the sensitivity of image analysis techniques to acquisition parameters. Using a dataset obtained by imaging the phantom in different scanners, we have determined the limit to how accurately electrodes can be localized in images with various acquisition parameters. The summary of our tests shown in Table 8 shows that changes of the HU range and resolution of reconstruction always affect the determination of the correct CI array locations. Dose and orientation affects localization accuracy as well, but their effects are not as strong

Table 8 Summary of effects of CT parameters on localization accuracy. The error subgroups being compared are mentioned in parentheses under the column “factor.” Incomplete data have been marked as —, implying that no test can be performed. If a statistically significant effect ($p < 0.05$) is present, the group means are specified in mm, while “no” indicates the absence of a statistical effect.

Factor	Test conditions	A1			A2		
		Image-based	Auto I	Auto II	Image-based	Auto I	Auto II
HU range (Ext HU, Lim HU)	High dose, mid res	(0.048, 0.081)	—	—	(0.054, 0.123)	—	—
Orientation	High dose, mid res	No	No	No	(0.044, 0.065)	No	No
Resolution (mid res, low res)	Lim HU, mid dose	No	(0.176, 4.357)	—	(0.154, 0.371)	—	—
	Ext HU, mid dose	(0.058, 0.096)	No	(0.125, 0.276)	(0.066, 0.344)	(0.147, 0.215)	(0.128, 0.230)
Resolution (high res, very low res)	Ext HU, low dose	(0.055, 0.109)	No	No	(0.046, 0.106)	No	(0.079, 0.129)
Dose (high dose, mid dose)	Lim HU, mid res	No	—	—	No	—	—
	Ext HU, low res	No	No	(0.102, 0.276)	(0.144, 0.344)	No	No
	Ext HU, mid res	No	No	(0.153, 0.125)	No	No	No

(Fig. 7). On our tested images, image-based localization of A1 is insensitive to changes of orientation (for high dose, mid resolution, and extended HU reconstructions), but this factor significantly affects the localization of A2. Dose has only minimal effect, with only the errors for array A2 being affected in one out of three test groups.

Our work identifies how the variation of four major CT acquisition parameters across different scanners affects the accuracy of localization of implanted arrays using the images. This provides a baseline of expected errors that can be used while evaluating the sensitivity of any automatic image analysis method on CI arrays on the images of this phantom. We need highly accurate electrode localization to investigate how electrode positions affect hearing outcome, and our phantom would allow us to estimate the bounds on localization error given a CT protocol and a localization method. We have demonstrated the use of the phantom for evaluating segmentation methods by also evaluating two automatic localization techniques. For these two methods, we confirmed that the major factor affecting their accuracy is resolution. Furthermore, our results can be used to determine minimum resolution, dose, and other protocol requirements for images that will be used in our study to evaluate electrode position, as demonstrated by fitting a general linear model.

One limitation of the current work is that the accuracy of the chosen ground truth is determined by the precision of the reference CT volume. Without a better option for a ground truth, we used the best electrode position that we could define by repeated measurements on our best CT image. A potentially better approach would be to use a micro-CT to define the ground truth.³⁰ Size limitations of scanners that are currently available do not allow us to micro-CT something as large as our phantom. Once micro-CT technology that allows for imaging of larger samples becomes more widely available, such images could provide a better gold standard for electrode position in the phantom.

Even with high-resolution CT, critical small-scale intracochlear structures are still invisible, and thus, measuring electrode position in the context of the cochlea remains difficult. The most accurate approach has been to use statistical shape models of the cochlea, which rely on the visible external walls of the cochlea as a landmark to estimate the location of invisible intracochlear structures.^{31,32} Thus, in future work, we will also use the phantom to evaluate how accurately intracochlear anatomy structures can be localized when using different image acquisition parameters.

Disclosures

The authors have no conflicts of interest to declare.

Acknowledgments

This research has been supported in part by Grant Nos. R01DC014037, R01DC008408, and R01DC014462 from the National Institute on Deafness and other communication disorders. The content is solely the responsibility of the authors and does not necessarily represent the official views of these institutes. The authors would also like to thank Advanced Bionics for providing the CI devices for this study, and the Biostatistics Clinics hosted by the Department of Biostatistics, Vanderbilt University for methodological guidance regarding the statistical calculations.

References

1. E. Buss et al., "Multicenter US bilateral MED-EL cochlear implantation study: speech perception over the first year of use," *Ear Hear.* **29**(1), 20–32 (2008).
2. M. F. Dorman et al., "Speech perception and sound localization by adults with bilateral cochlear implants," *Semin. Hear.* **32**, 73–89 (2011).
3. R. H. Gifford, J. K. Shalloo, and A. M. Peterson, "Speech recognition materials and ceiling effects: considerations for cochlear implant programs," *Audiol. Neurotol.* **13**(3), 193–205 (2008).
4. R. H. Gifford et al., "Availability of binaural cues for bilateral implant recipients and bimodal listeners with and without preserved hearing in the implanted ear," *Audiol. Neurotol.* **19**(1), 57–71 (2013).
5. R. Litovsky et al., "Simultaneous bilateral cochlear implantation in adults: a multicenter clinical study," *Ear Hear.* **27**(6), 714–731 (2006).
6. B. M. Verbist et al., "Multisection CT as a valuable tool in the postoperative assessment of cochlear implant patients," *Am. J. Neuroradiol.* **26**(2), 424–429 (2005).
7. A. Aschendorff et al., "Quality control after cochlear implant surgery by means of rotational tomography," *Otol. Neurotol.* **26**(1), 34–37 (2005).
8. M. W. Skinner et al., "In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea," *Ann. Otol. Rhinol. Laryngol.* **116**(4), 2–24 (2007).
9. G. B. Wanner et al., "Assessment of electrode placement and audiologic outcomes in bilateral cochlear implantation," *Otol. Neurotol.* **32**(3), 428–432 (2011).
10. G. B. Wanner et al., "Impact of electrode design and surgical approach on scalar location and cochlear implant outcomes," *Laryngoscope* **124**, S1–S7 (2014).
11. L. K. Holden et al., "Factors affecting open-set word recognition in adults with cochlear implants," *Ear Hear.* **34**(3), 342–360 (2013).
12. K. Yukawa et al., "Effects of insertion depth of cochlear implant electrodes upon speech perception," *Audiol. Neurotol.* **9**(3), 163–172 (2004).
13. C. C. Finley et al., "Role of electrode placement as a contributor to variability in cochlear implant outcomes," *Otol. Neurotol.* **29**(7), 920–928 (2008).
14. Y. Zhao et al., "Automatic localization of cochlear implant electrodes in CT," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 331–338, Springer International Publishing (2014).
15. J. H. Noble and B. M. Dawant, "Automatic graph-based localization of cochlear implant electrodes in CT," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 152–159, Springer International Publishing (2015).
16. J. Xu et al., "Cochlear view: postoperative radiography for cochlear implantation," *Am. J. Otol.* **21**(1), 49–56 (2000).
17. G. Wang et al., "Digital x-ray stereophotogrammetry for cochlear implantation," *IEEE Trans. Biomed. Eng.* **47**(8), 1120–1130 (2000).
18. H. W. Husstedt et al., "Nondestructive three-dimensional analysis of electrode to modiolus proximity," *Otol. Neurotol.* **23**(1), 49–52 (2002).
19. D. R. Ketten et al., "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays," *Ann. Otol. Rhinol. Laryngol.* **175**(11), 1–16 (1998).
20. G. Wang et al., "Spiral CT image deblurring for cochlear implantation," *IEEE Trans. Biomed. Eng.* **17**(2), 251–262 (1998).
21. M. Jiang et al., "Blind deblurring of spiral CT images," *IEEE Trans. Biomed. Eng.* **22**(7), 837–845 (2003).
22. J. H. Noble et al., "Image-guidance enables new methods for customizing cochlear implant stimulation strategies," *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**(5), 820–829 (2013).
23. J. Farrell et al., "Problem 65-1: a least squares estimate of satellite attitude," *Siam Rev.* **8**(3), 384–386 (1966).
24. J. M. Fitzpatrick, J. B. West, and C. R. Maurer Jr., "Predicting error in rigid-body point-based registration," *IEEE Trans. Biomed. Eng.* **17**(5), 694–702 (1998).
25. F. Maes et al., "Multimodality image registration by maximization of mutual information," *IEEE Trans. Biomed. Eng.* **16**(2), 187–198 (1997).
26. B. L. Welch, "The generalization of 'student's problem' when several different population variances are involved," *Biometrika* **34**(1/2), 28–35 (1947).

27. G. Casella and R. L. Berger, *Statistical Inference*, Vol. 2, Duxbury, Pacific Grove, California (2002).
28. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Academic Press, New York (1988).
29. H. Hentschke and M. C. Stüttgen, "Computation of measures of effect size for neuroscience data sets," *Eur. J. Neurosci.* **34**(12), 1887–1894 (2011).
30. G. Wang et al., "Design, analysis and simulation for development of the first clinical micro-CT scanner," *Acad. Radiol.* **12**(4), 511–525 (2005).
31. J. H. Noble et al., "Automatic segmentation of intracochlear anatomy in conventional CT," *IEEE Trans. Biomed. Eng.* **58**(9), 2625–2632 (2011).
32. J. H. Noble et al., "Statistical shape model segmentation and frequency mapping of cochlear implant stimulation targets in CT," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 421–428, Springer (2012).

Srijata Chakravorti received her BE degree in electrical engineering from Jadavpur University, Kolkata, West Bengal, India, in 2015. Currently, she is a PhD student of electrical engineering at Vanderbilt University, Nashville, Tennessee. Her research interests include medical image processing, image segmentation and statistical modeling for image-guided surgery.

Brian J. Bussey received his BS degree from Pennsylvania State University and MD degree from the University of Cincinnati. Currently, he is a chief resident in the Department of Radiology at Vanderbilt University Medical Center. His academic interests include emergency radiology and the business of radiology.

Yiyuan Zhao received his BS degree in telecommunication engineering from Xidian University, Xi'an, Shaanxi, China, in 2011. Currently,

he is a PhD student of electrical engineering from Vanderbilt University, Nashville, Tennessee. His current research project is automatic techniques for cochlear implant CT image analysis. His research interests include medical image processing, image segmentation and registration, and computer vision.

Benoit M. Dawant received his MSEE degree from the University of Louvain, Leuven, Belgium, in 1983, and his PhD from the University of Houston, Houston, Texas, in 1988. Currently, he is a professor at the Department of Electrical Engineering and Computer Science, Vanderbilt University. His primary research areas include medical image processing, segmentation, and registration. He currently focuses on applying these methods to the development of novel techniques for guiding surgical and interventional procedures.

Robert F. Labadie received his BS in mechanical engineering from the University of Notre Dame, Notre Dame, Indiana, in 1988, his PhD in bioengineering, and MD from the University of Pittsburgh, Pennsylvania, in 1995 and 1996. Currently, he is a tenured professor in the Department of Otolaryngology/Head and Neck Surgery with a joint appointment in the Department of Biomedical Engineering at Vanderbilt University. His clinical specialty is otology and surgical rehabilitation of the hearing impaired.

Jack H. Noble received his BE, MS, and PhD degrees in electrical engineering from Vanderbilt University, Nashville, Tennessee, in 2007, 2008, and 2011, respectively. Currently, he is an assistant professor at the Department of Electrical Engineering and Computer Science, Vanderbilt University. His primary research interests include medical image processing, image segmentation, registration, statistical modeling, and image-guided surgery techniques.