Imaging the inner ear in fossil mammals: High-resolution CT scanning and 3-D virtual reconstructions

Xijun Ni, John J. Flynn, and André R. Wyss

ABSTRACT

The bony labyrinth of mammals, a delicate and complex cavity within the petrosal, houses the organs of hearing and equilibrium of the inner ear. Because this region is typically lodged deep within the skull, there have been few morphological studies of the bony labyrinth in fossils—where it frequently is completely enveloped by surrounding bone and sediment matrix. The recent development of high-resolution X-ray Computed Tomography (CT) scanning provides a powerful new tool for investigating such tiny, often inaccessible structures. Here we introduce a protocol for producing three-dimensional (3-D) virtual visualizations of the bony labyrinth from high-resolution CT images. As a case study, we scanned the skull of a basal platyrrhine primate, *Chilecebus carraacoensis*, using the high-resolution CT facility at Pennsylvania State University, reconstructing the endocast of the bony labyrinth from the resulting data. Segmenting the original CT images is a vital step in producing accurate 3-D virtual reconstructions. We failed in efforts to isolate the bony labyrinth endocast through automated means, owing to the similar densities of the matrix filling the sinuses and spongy bone cavities of the specimen, and the bony labyrinth itself. Differing density contrasts across the fossil/matrix interface and overlapping grayscales of the fossil and matrix, required Half Maximum Height thresholds to be measured dynamically. Multiple thresholds are advantageous for processing the CT data of fossils that are inherently heterogeneous in material properties and densities. The iterative interaction between an operator and a computer offers the only means presently available for reliably discriminating the endocast of the bony labyrinth from the remainder of the specimen.

Xijun Ni. Division of Paleontology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024. nixijun@ivpp.ac.cn
Key Laboratory of Evolutionary Systematics of Vertebrates, Institute of Vertebrate Paleontology and Paleoanthropology, Xi Zhi Men Wai Street 142, Beijing, 100044
John J. Flynn. Division of Paleontology and Richard Gilder Graduate School, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024. jflynn@amnh.org
André R. Wyss. Department of Earth Science, University of California- Santa Barbara, Santa Barbara, CA 93106. wyss@geol.ucsb.edu

Key Words: Bony labyrinth; high-resolution X-ray CT; segmentation; three-dimensional (3-D) virtual reconstruction; *Chilecebus*; platyrrhine primate
INTRODUCTION

The inner ear labyrinth of mammals, a complex, delicate neurosensory system, is enclosed within the petrosal bone, forming an intricate system of osseous canals. Including the bony labyrinth itself, and the membranous labyrinth within, it consists of three parts: the cochlea at the rostral end, the vestibule in the middle, and the semicircular canals caudally. The cochlea houses the hearing organ, and it forms a canal with a varying degree of curvature in early-diverging fossil mammaliaforms, and a coiled, snail-like structure in fossil and extant therian mammals. Together, the three semicircular canals form the motion detection system or organ of balance, monitoring and transmitting information about head movements in three-dimensional space.

Comparative study of the bony labyrinth in living vertebrates dates back more than 150 years (Hyrtl, 1845). Gray (1907, 1908a, 1908b) provided detailed descriptions, illustrations, and comparisons of a wide range of vertebrate labyrinths in his milestone two-volume monograph. Subsequent investigators analyzed various aspects of the structural diversity, morphological-physical relationships, and evolution of the labyrinth (Werner, 1933; Turkewitsch, 1935; Gray, 1955; Jones and Spells, 1963; Fleischer, 1973, 1976; Takahashi, 1976; Curthoys et al., 1977a, 1977b; Ramprashad et al., 1984; Blanks et al., 1985; Reisine et al., 1988; Spoor and Zonneveld, 1998; Spoor et al., 2007; Walker et al., 2008; Macrini et al., 2010; Ni et al., 2010).

Research on the labyrinth in fossils is much more limited, mainly because this organ historically could only be readily studied through rarely preserved natural fossil endocasts of the labyrinth. In such exceptional preservation instances, however, natural endocasts of the bony labyrinth can provide a rich information source for understanding motion sensitivity and locomotor patterns of the taxa involved (Case, 1928; Cox, 1962; Thoss and Schwartz, 1975; Madsen, 1976; Meng and Wyss, 1995; Clarke, 2005).

Classical techniques of studying bony labyrinth morphology include direct dissection (Retzius, 1881, 1884; Turkewitsch, 1935; Blanks et al., 1972, 1985; Fleischer, 1973, 1976; Matano et al., 1985, 1986; Ghanem et al., 1998), casting the inner space of the bony labyrinth (Hyrtl, 1845; Takahashi, 1976), decalcification and disintegration of the petrosal bone (Gray, 1903, 1906, 1907, 1908a, 1908b), and serial sectioning (Igarashi, 1967; Curthoys et al., 1977a, 1977b; Ramprashad et al., 1984). In recent years, medical diagnostic tools such as X-ray radiography, CT, and MRI have been increasingly employed in studies of the labyrinth (Habersetzer and Storch, 1992; Spoor and Zonneveld, 1995, 1998; Held et al., 1997; Thorne et al., 1999).

Because of their destructiveness, classical techniques of studying the bony labyrinth have been applied in vertebrate paleontological research only sparingly, given the limited material from which many fossil taxa are known. Successful applications of such methods include the mechanical removal of bone to expose the fossilized endocast in Devonian jawless vertebrates (Stensiö, 1927), and the construction of models through serial sectioning in Devonian osteostracans (Stensiö, 1927) and in Late Cretaceous multituberculates (Kielen-Jaworowska et al., 1986; Hurum, 1998). Modern X-radiographic and Computed Tomography (CT) imaging techniques have been widely used to investigate the labyrinth in extant vertebrates, but to a much lesser extent in fossils. Paleontological examples include analyses of the inner ear of Late Cretaceous and Early Paleocene multituberculates (Luo and Ketten, 1991), the evolutionary development of the promontorium and cochlea of early-diverging mammals plus proximal outgroups (Luo, 1995; Luo et al., 1995; Ruf et al., 2009; Luo et al., 2011; Ekdale and Rowe, 2011), and the inner ear of Paleogene notoungulates (Macrini et al., 2010). Other recent investigations have centered on the vestibular apparatus in dinosaurs (Rogers, 1998), auditory region changes during the aquatic transition of early whales (Luo and Eastman, 1995; Geisler and Luo, 1996; Luo and Marsh, 1996; Spoor et al., 2002), and neurological features of the brain and vestibular apparatus of pterosaurs (Witmer et al., 2003) and Archaeopteryx (Alonso et al., 2004). Recent studies also have documented the utility of high-resolution CT for elucidating anatomical details of the labyrinth and for inferring the locomotor behavior of fossil anthropoids (Spoor et al., 1994, 1996, 2003; Spoor, 2003; Rook et al., 2004; Ni et al., 2010), subfossil lemuroids of Madagascar (Spoor et al., 2007; Walker et al., 2008), and plesiadapiforms and other putative or bona fide basal primates (Silcox et al., 2009). Previous studies have been hindered, however, by a lack of consistent CT-image “cleaning” and 3-D reconstruction protocols. The piecemeal approach that such studies have taken until now,
therefore, severely limits the direct comparison of results. Our goal in proposing a standardized methodology here is to facilitate comparisons between different studies. An important component of the proposed protocol is the Half Maximum Height (HMH) method (Baxter and Sorenson, 1981; Spoor et al., 1993); as it has not been applied consistently in previous studies, the identification of boundaries was often subjective, and not based on repeatable measurements.

Here we present a novel protocol for segmenting high-resolution CT images and producing 3-D virtual reconstructions of the bony labyrinth endocast in fossil mammals. As a demonstration case, we analyzed a high-resolution scan of the only known specimen of the basal platyrrhine primate, *Chilecebus carrascoensis*, represented by a nearly complete skull. Segmentation of this imagery and subsequent processing permitted reconstruction of an endocast of the bony labyrinth. Morphological and functional interpretations based on the digital reconstruction of the bony labyrinth of *Chilecebus* are detailed elsewhere (Ni et al., 2010). Our objective here is to more fully describe the new methodologies employed in generating these digital endocasts, as they are more generally applicable to the 3-D reconstruction of anatomical structures from high-resolution CT data.

**MATERIALS AND METHODS**

**Fossil Specimen**

The specimen analyzed in this study, a nearly complete skull of *Chilecebus carrascoensis*, is one of the best-preserved early platyrrhines known. The only known specimen of this taxon, SGOPV 3213, was recovered from volcanoclastic deposits of the Abanico (=Coya-Machalí) Formation in the Andes of central Chile (Flynn et al., 1995). An 40Ar/39Ar date of 20.09 ± 0.27 Ma is directly associated with the fossil (Flynn et al., 1995).

**High Resolution X-ray CT-scan**

High-resolution X-ray CT scanning was performed at the Center for Quantitative Imaging at Pennsylvania State University, USA, using an X-TEX 225kV microfocus X-ray source. The focal spot size was approximately 15 microns at a load of 130 kV and 0.110 mA (~14 watts). The detector used has a 1024 x 1024-pixel area with an image intensifier. The specimen was scanned in the coronal plane. Scans were collected in 2400 views, four samples per view, with a field of view of 40.96 mm. Interslice spacing was 0.04641 mm.

**Image Segmentation**

Through image segmentation, the CT image stack was divided into multiple segment stacks, thereby isolating one or more regions of interest (ROI)—here, the bony labyrinth. We used ImageJ (Abramoff et al., 2004), a public domain, Java-based image processing program developed at the U.S. National Institutes of Health (http://rsbweb.nih.gov/ij/download.html), to distinguish the matrix infilling the cavity of the bony labyrinth from other matrix and the bones of the skull. The HMH method (Baxter and Sorenson, 1981; Spoor et al., 1993) was used to assist in identifying the bone-matrix interface. Measurements of structures on CT images via this method are consistent with those generated by conventional means (Spoor et al., 1993). Because the densities of the fossil bones and the surrounding matrix vary across sub-areas of the CT images, no single threshold separates bone from matrix throughout the whole image stack. Consequently, we measured the HMH value dynamically, setting different thresholds for various regions. To ensure that anatomical structures were identified consistently, images were traced and separated repeatedly, not only in the original x-y plane, but also in the re-sliced z-x and y-z planes. Finally, different segments of the CT data, which include the various anatomical details of the skull, were saved as separate stacks. These stacks can be restored to the original CT images without losing or altering pixel grayscale values. In other words, although certain portions of the original images were analyzed and saved independently (highlighting the bony labyrinth, for example), no primary information was lost or added. It is important to emphasize this point because when small, delicate, and complex anatomical structures are reconstructed from CT data, artifacts generated during the segmentation process can introduce errors yielding spurious final results.

**Three-dimensional Visualization**

Following image segmentation, image data subsets of the bony labyrinth were imported into VGStudio Max 1.1 and 2.1 for 3-D visualization. The volume rendering technique was used to create 3-D virtual models from the CT data. Although an isocontouring (or surface) method is desirable for producing crisp and smooth 3-D reconstructions, volume rendering is better suited to CT imagery because it allows a greater range of primary voxel information to be analyzed for both surface and volumetric reconstruction in any given image (Ketcham and Carlson, 2001).
Because the skull of *Chilecebus* was slightly distorted by tectonic forces and sediment compaction, the left and right bony labyrinths are not completely symmetrically positioned about the midline in the transverse plane (parallel to the palate). The angular offset of landmarks on the left and right sides of the specimen were measured multiple times, yielding an average offset of 7.7 degrees. Bilateral symmetry was restored during 3-D visualization using the occipital condyles, foramen magnum, and posterior edges of the maxillary alveoli as points of reference. Remarkably, this minor correction in a single plane achieved symmetry in the full restoration, indicating that there is only minor and easily corrected distortion in this fossil.

**Skeletonization**

In shape analysis, the topological skeleton of a shape is the central path of that shape, equidistant to its boundaries. Skeletonization, the process through which the topological skeleton of any structure (such as the bony labyrinth) is identified, is a common preprocessing operation in digital raster-to-vector conversion or pattern recognition. It is also an essential step for the rotations required to determine the true planar orientations of the semicircular canals (David et al., 2010).

We have developed a new method of skeletonizing the central paths of the semicircular canals using the software ImageJ, an essential step in determining the rotations required to define the true planar orientations of the semicircular canals. The segmented image stack of the bony labyrinth in the x-y plane was re-sliced and transformed into stacks in the y-z and z-x planes, all three stacks— in the x-y, y-z, and z-x planes—then being skeletonized using the "skeletonize" command of ImageJ. Next, the skeletonized y-z and z-x plane stacks were re-sliced and transformed back into the x-y plane. This process projects the skeletonized image stacks in the y-z and z-x planes into the x-y plane, resulting in three stacks all in the x-y plane, each skeletonized from a different source. From these three skeletonized stacks we calculated a new logical conjunction image stack: only the pixels having the same none-zero value at the same position in all three stacks were preserved. This logical conjunction calculation ensures that only points on the central path in a 3-D space will be preserved. The 3-D reconstruction based on these logical conjunction images yielded a skeletonized bony labyrinth. The VGStudio Max 1.1 and 2.1 platform was used to then rotate the bony labyrinth into a position such that one or two of the semicircular canals were perpendicular to the observing plane, permitting angles between canals to be measured.

**RESULTS**

The initial CT data set obtained for *Chilecebus* consists of a 16-bit grayscale image stack in TIFF format. This stack includes 1148 images, each of which consists of 1024 x 1024 pixels. Each voxel in the stack has a resolution of 0.04 x 0.04 x 0.04641 mm.

The histogram of the grayscale distribution of the entire CT data set of *Chilecebus* shows two peaks (Figure 1, red line), one near 2000-3000 levels (of 65,535 levels spanning 16-bit grayscales), and the other near 6000-8000 levels. The lower grayscale peak (2000-3000) is composed of two sub-peaks, and represents the scattering background. The higher peak (6000-8000) arises from the fossil bone and the matrix. It is also composed of two closely spaced sub-peaks, a higher one (near 6000) reflecting matrix, and a lower one (near 7000) reflecting fossilized bone, as indicated by the grayscale distribution spectrum and interactive sampling of the CT images.

The two sub-peaks in grayscale distribution on each CT slice represent fossil bone and surrounding matrix; these peaks resemble those for the whole image stack (e.g., Figure 2.1 and 2.4). Distinguishing fossil material from the matrix infill-
ing the cavity of the bony labyrinth on each CT slice was accomplished manually. Images of the fossil material, and those of the matrix infilling the cavity of the bony labyrinth, were saved in different image stacks. The peak of the grayscale distribution of images containing fossil bone falls slightly to the right of the x-axis, but otherwise is very similar to the total distribution (e.g., Figure 2.2 and 2.5). The position of this peak reflects the similar density of the matrix filling the sinuses and cavities of the bones, the matrix filling the endocranial cavity, and the matrix filling the bony labyrinth. The grayscale distribution of the segmented bony labyrinth exhibits a single-peak (e.g., Figure 2.3 and 2.6), corresponding to the sub-peak of the matrix (as in Figure 2.4). No sub-peak reflecting fossilized bone persists after segmentation.

The histogram of the grayscale distribution of the entire segmented bony labyrinth image stack forms a simple, bell-shaped curve (Figure 1, green line), corresponding to the matrix sub-peak from the original data set. The histogram for the images in which the brain endocast, bony labyrinth, and part of the attached matrix have been “subtracted,” shows two main peaks (Figure 1, blue line). The left peak and its two sub-peaks are almost identical.

FIGURE 2. An example of the CT image segmentation process. The Half Maximum Height method (Baxter and Sorenson, 1981; Spoor et al., 1993) is used to assist in identifying the bone-matrix interface. 1, original CT images across the right ear region in coronal plane; 2, endocasts of the bony labyrinth and cranial cavity are segmented; 3, endocast of the bony labyrinth. 4, 5, and 6 indicate the grayscale distribution of 1, 2, and 3, respectively.
to the peak of the original CT images (Figure 1, red line); this peak represents the background values for the scans. The main peak in the right part of the curve (Figure 1, blue line) is also composed of two sub-peaks, just as in the original data (Figure 1, red line). These two sub-peaks represent the lower-density matrix and the higher-density bone. These histograms indicate that the tracing, re-slicing, and separation processes on the CT images successfully distinguished bone from matrix, particularly for the segmented internal ear region image subset.

Three-dimensional virtual reconstructions based on segmented CT images of the bony labyrinth are shown in Figures 3 and 4. Prior to symmetry restoration, the opposing bony labyrinths are askew, and each is deformed (Figure 3.1). Restoration in the transverse plane makes the bony labyrinths bilaterally symmetric (Figure 3.2). Many intricate details of the bony labyrinth, including the spiral turns of the cochlea, the oval and round windows, the vestibular aqueduct, the semicircular canals, and the ampullae, are evident on the 3-D virtual reconstruction (Figure 4.1).

Skeletonization of the bony labyrinth yielded a series of isolated points along centerlines rather than continuous centerlines (Figure 4.2).

DISCUSSION AND COMMENTS

The grayscale of each voxel in a CT image volume reflects attenuation of X-rays as they are scattered or absorbed in passing through corresponding points on the specimen. Because fossilized bones and adhering matrix are composed of different materials, presumably of at least subtly different density and atomic number, their grayscale distributions should theoretically display multiple peaks. The HMH method is a generally accepted means of identifying the thresholds needed to accurately distinguish bone from matrix within a CT image, without artificially changing the size of the ROI (Baxter and Sorenson, 1981; Spoor et al., 1993). Unfortunately the uneven distribution of high and low-density materials across the fossil/matrix interface often creates significant overlap of the fossil and matrix grayscale distributions. The resulting curve, having a single or two closely spaced peaks, makes it difficult to differentiate bone from matrix. In such instances, therefore, no single threshold will reliably separate bone from matrix in a CT image stack in an automated fashion. Using single thresholds results in the loss of non-identification of low-density bone, or the spurious inclusion of extraneous high-density matrix, in the final image. By employing multiple thresholds...
Across different regions of the CT images these pitfalls can be avoided. Indeed, multiple thresholds are advantageous not just for segmenting CT images of the bony labyrinth, but also in showing promise for the processing of CT images of fossils in general.

As in many fossils, the sinuses and spongy bone cavities surrounding the bony labyrinth in SGOPV 3213 are filled with matrix indistinguishable from that within the labyrinth itself, one reason that it is impossible to resolve the labyrinth without manual intervention. Similarly, the endocast of the bony labyrinth cannot be isolated through a simple automated process of segmentation using single or multiple threshold settings, given density variations across the specimen and the large volume of matrix filled sinuses, spongy bone, and other spaces surrounding it. Reliably discriminating anatomical structures of the bony labyrinth from its enveloping matrix can currently be accomplished only through an iterative process between an investigator and their analytical software. The iterative discrimination process is likely to be necessary in many other specimens with low density differentials.

Recently developed commercial voxel data visualization programs, such as VGStudio Max, Mimics, and Amira, offer sophisticated segmentation capabilities utilizing grayscale-values or geometry, and thus may be used in the iterative segmentation process described here. Dynamically measuring and setting multiple thresholds using the HMH criterion is far more cumbersome in these commercial applications, however, than in the freeware ImageJ. Another drawback of the proprietary software is that it is extremely hardware intensive, making segmentation not only expensive but also very slow. In contrast, the hardware requirements for ImageJ are modest. Because digital CT images each consist of a finite set of grayscale values arranged in rows and columns, they can be readily processed mathematically. Straightforward batch command functions enable the algorithms in ImageJ to segment images conveniently, thereby greatly reducing image-processing time.

The fossil analyzed in this demonstration study, *Chilecebus*, is preserved in extremely hard and brittle volcanioclastic rock. The matrix filling the endocranial cavity and sinuses of the skull contains minerals denser than the fossilized bones themselves. These high-density minerals are distributed randomly and in CT imaging appear as artifactual “starry spots” within the skull, sometimes even crossing the bone-matrix interface. Our results indicate that the 3-D image processing protocol introduced here is capable of discerning very fine-scale anatomical structures even against extremely complicated backgrounds, artifactual signals, or “noise.” Dynamic threshold measuring, coupled with manual human-computer interactive segmentation on x-y, y-z, and x-z planes, is essential for producing smooth, anatomically accurate, 3-D virtual reconstructions from CT imagery, and is applicable to a wide variety of fossils across a spectrum of sediment matrices and preservation styles.

Three-dimensional visualization of CT images is a powerful tool for morphological study, useful not merely for producing attractive graphics, but also for revealing anatomical details, such as the complex and delicate bony labyrinth of mammals, details that are otherwise wholly unobtainable. A complete description of the morphology of the skull and bony labyrinth of *Chilecebus*, partially detailed elsewhere (Ni et al., 2010), is beyond the scope of this paper. Nevertheless, we emphasize that the dimensions, proportions, shape indices, and orientations of the bony labyrinth defined by Spoor and...
Zonneveld (Spoor and Zonneveld, 1995, 1998) can all be determined from the 3-D virtual reconstructions produced by all three of the visualization programs used in our analyses, VGStudio Max (1.1, 2.1), and ImageJ. The skeletonized virtual model of the bony labyrinth is particularly useful for establishing these parameters.

ACKNOWLEDGMENTS

This research was supported by our home institutions, The Field Museum, research grants from the U.S. National Science Foundation (DEB-9317943, DEB-0317014 and DEB-0513476 to JF; DEB-9020213 and DEB-9318126 to ARW; and 0333415 (NYCEP) to XN), a John S. Guggenheim Memorial Foundation Fellowship (to JJF), and the Chinese National Science Foundation (40672009, 40872032 to XN), a John S. Guggenheim Memorial Foundation Fellowship (to JF), and the Chinese National Science Foundation (40672009, 40872032 to XN). For the CT scanning of Chileneus carrascoensis we thank Drs. T. Ryan and A. Walker of Penn State University, and the PSU Center for Quantitative Imaging. We are grateful for the long-term support of the Museo Nacional de Historia Natural (Santiago, Chile) and the Consejo de Monumentos Naturales de Chile. We greatly appreciate the constructive comments and helpful suggestions of Dr. Z. Luo and another anonymous reviewer.

REFERENCES


