

Virtual Labyrinthoscopy: Visualization of the Inner Ear with Interactive Direct Volume Rendering¹

Bernd F. Tomandl, MD • Peter Hastreiter, Dipl.-Ing. • Knut E. W. Eberhardt, MD • Christof Rezk-Salama, Dipl.-Inf. • Ramin Naraghi, MD
Holger Greess, MD • Urs Nissen, MD • Walter J. Huk, MD

Computed tomography (CT) is the modality of choice for detailed imaging of the bony labyrinth. Usually, information about the complex three-dimensional anatomic structures of the inner ear is presented as two-dimensional section images. Interactive direct volume rendering is a powerful method for visualization of the labyrinth. Unlike other visualization methods, direct volume rendering enables direct visualization of the bony labyrinth without explicit segmentation prior to the visualization process. Direct volume rendering was applied to visualization of the structures of the temporal bone in five patients without pathologic conditions and four patients with pathologic conditions. In all cases, clear representations of the bony labyrinth and the facial canal were provided. Because standard CT examinations combined with interactive visualization based on direct volume rendering are used, the method is fast and flexible. Therefore, this approach is applicable in routine clinical work. Problems occur in patients with effusion in the temporal bone because adjustment of imaging parameters for proper delineation of the target structures is difficult in this situation. However, direct volume rendering can produce meaningful images of high quality even in these problematic cases. The term *virtual labyrinthoscopy* is suggested for visualization of the labyrinth by using direct volume rendering.

Abbreviations: FOV = field of view, MIP = maximum intensity projection, SSD = shaded surface display, 3D = three-dimensional

Index terms: Computed tomography (CT), image processing • Computed tomography (CT), three-dimensional • Computed tomography (CT), volume rendering • Data fusion • Ear, labyrinth, 213.12117

RadioGraphics 2000; 20:547–558

¹From the Division of Neuroradiology (B.F.T., K.E.W.E., W.J.H.), Computer Graphics Group (P.H., C.R.S.), Department of Neurosurgery (R.N., U.N.), and Institute of Diagnostic Radiology (H.G.), University of Erlangen-Nuremberg, Schwabachanlage 6, D-91054 Erlangen, Germany. Presented as a scientific exhibit at the 1998 RSNA scientific assembly. Received March 8, 1999; revision requested April 28 and received June 2; accepted June 11. **Address reprint requests to** B.F.T. (e-mail: bernd.tomandl@stud.uni-erlangen.de).

Introduction

For detailed imaging of the temporal bone in patients with fractures, congenital malformations, or osteodystrophies involving the inner ear, computed tomography (CT) is still the modality of choice (1,2). Typically, information about the complex three-dimensional (3D) anatomic structures is presented as two-dimensional section images. Multiplanar reconstruction is helpful for further evaluation but still provides only two-dimensional images. In the past, it was difficult and time-consuming to perform 3D visualization of the tiny structure of the inner ear on the basis of high-resolution CT data (3,4).

Recently, direct volume rendering has received increasing attention within the medical community (5–7) because it includes all of the image data in the visualization process. Thereby, it provides semitransparent views, thus giving a good impression of the spatial relationships of all structures. So far, direct volume rendering has been applied to visualization of the trachea, the colon, and various blood vessels. Depending on the target structures, names like *virtual bronchoscopy* (7) and *virtual colonoscopy* (8) have been suggested. According to this nomenclature, we suggest the term *virtual labyrinthoscopy* for visualization of the labyrinth by using direct volume rendering. Owing to the tiny and complex structures of the inner ear, a fast and convenient analysis requires high-quality 3D representations and real-time manipulation of the viewing direction and the color and opacity values at any stage of the visualization process. These features are provided by interactive direct volume rendering.

Materials and Methods

Patients

Image data from five patients (three men, two women; mean age, 55 years) examined with spiral CT angiography for evaluation of an incidentally discovered aneurysm were used to reconstruct high-resolution images of the temporal bone (see the "CT Protocol" section). These patients had no known affliction of the temporal bone and no history of hearing problems, dizziness, or other symptoms related to the vestibulocochlear sys-

tem. Two female patients with tumors in the region of the temporal bone (one chondrosarcoma, one schwannoma of the glossopharyngeal nerve) were also included in the study. For the investigation of congenital malformations, we used image data from a female patient with an enlarged vestibular aqueduct and a male patient with a complex, Mondini dysplasia-like malformation of the inner ear.

CT Protocol

All examinations were performed with a spiral CT scanner (Somatom Plus 4; Siemens Medical Systems, Erlangen, Germany). In the five patients without pathologic conditions of the temporal bone, contrast medium was administered for CT angiography. The section thickness was 1 mm and the table speed was 1 mm/sec, resulting in a pitch of 1. The axial images were reconstructed with a high-resolution algorithm in steps of 0.5 mm and a field of view (FOV) of 60 mm². By using a 512² matrix, the resulting voxel size was 0.5 × 0.12 × 0.12 mm. We used the same protocol to evaluate the patient with chondrosarcoma.

The image data of the patient with schwannoma of the glossopharyngeal nerve were used for preoperative therapy planning. Because fusion with magnetic resonance (MR) images was envisaged, reconstruction of the section images was performed with an FOV of 180 mm². For image registration, a two-step procedure based on anatomic landmarks and consecutive automatic adjustment was applied.

The two patients with congenital anomalies were scanned with a section thickness of 1 mm and a table speed of 1.5 mm/sec (pitch = 1.5). Reconstruction was performed with steps of 1 mm and an FOV of 150 mm² to allow comparison of both sides.

Postprocessing

After transferring the image data to a high-end graphics workstation (Onyx Reality Engine II; Silicon Graphics, Mountain View, Calif), 3D visualization based on interactive direct volume rendering was performed with a tool developed by the Computer Graphics Group of the University of Erlangen-Nuremberg (9–11). By using the hardware capabilities of high-end graphics workstations, semitransparent representations of the

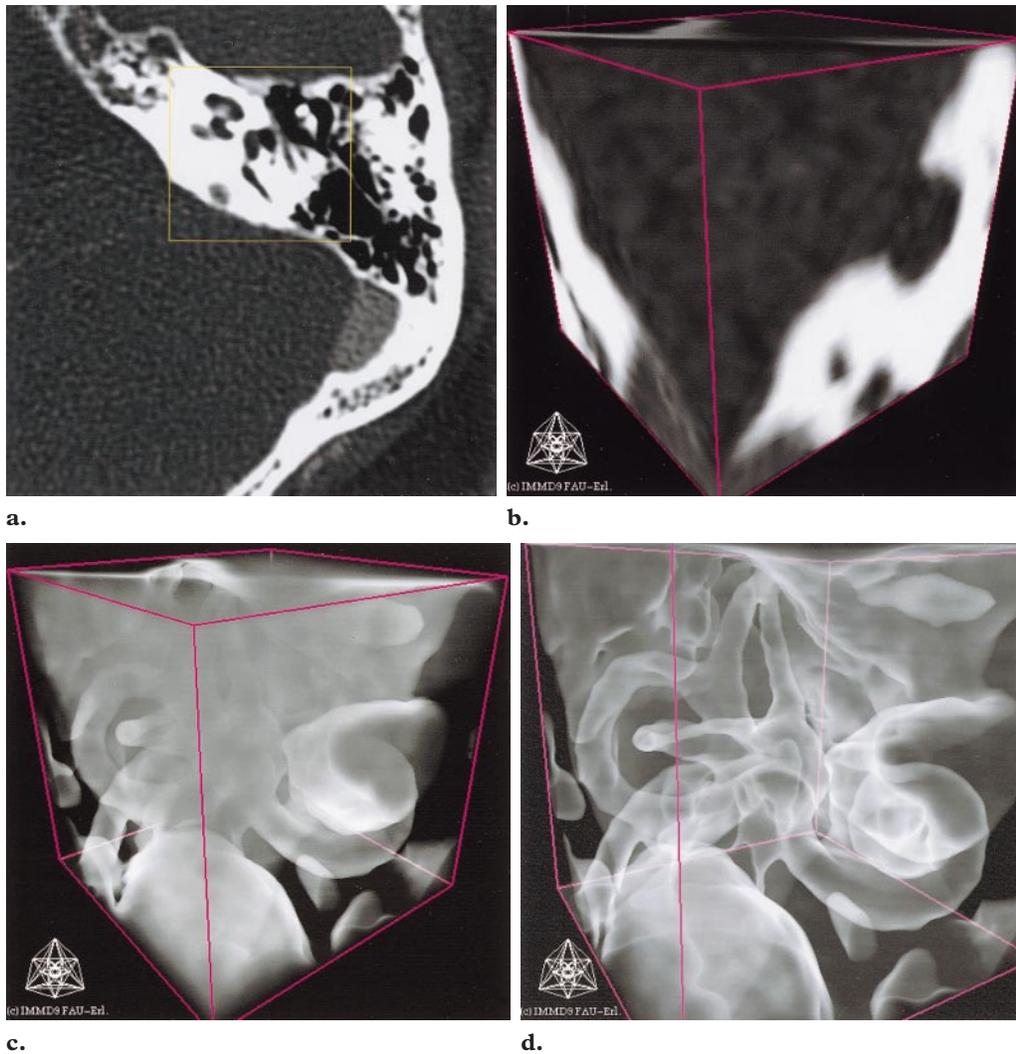


Figure 1. Visualization process based on high-resolution spiral CT data on the structures of the temporal bone. **(a)** Section image shows a volume of interest selected for 3D visualization (yellow square). **(b)** Anterolateral opaque 3D representation based on direct volume rendering shows the selected volume of interest. **(c)** Anterolateral view of the right labyrinth shows an intermediate step, during which color and opacity values are adjusted interactively. **(d)** Anterolateral semitransparent view shows the bony labyrinth.

structures of the inner ear were calculated. Because direct volume rendering considers all of the image data, no explicit segmentation prior to the visualization process was required. As a result of the hardware assistance, both color and opacity values were adjusted interactively to delineate all structures related to the inner ear in real time.

After an appropriate setting for optimal delineation of the target structures is defined, the color and opacity table can be stored and used for further studies. Owing to the tremendous acceleration of the visualization process (Fig 1), the whole

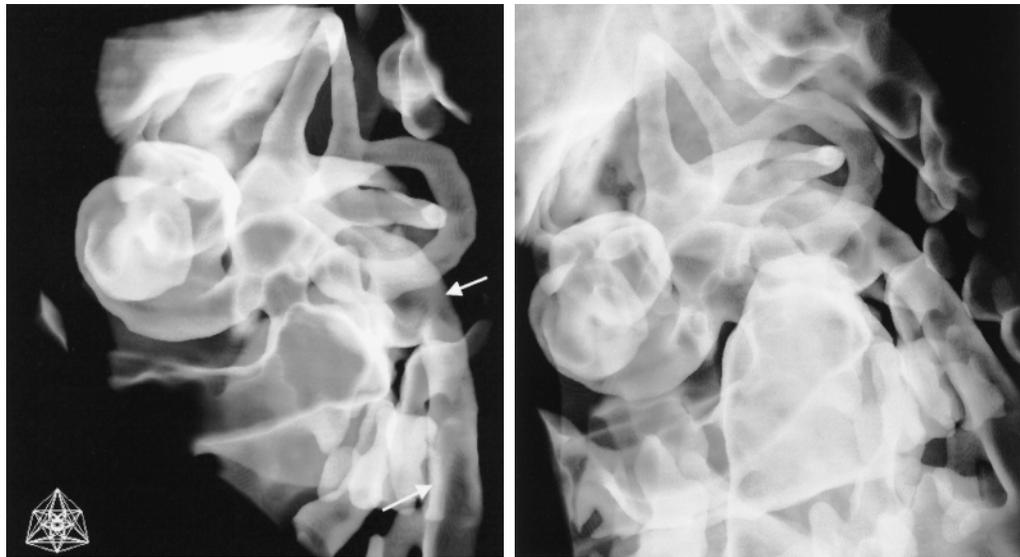


Figure 2. Direct volume rendering of a normal left labyrinth and facial canal. **(a)** Anterolateral view shows the relationship of the mastoid and labyrinthine segments of the facial nerve (arrows) to the bony labyrinth. **(b)** Anterolaterocaudal view shows the relationship of the bony labyrinth to the facial canal.

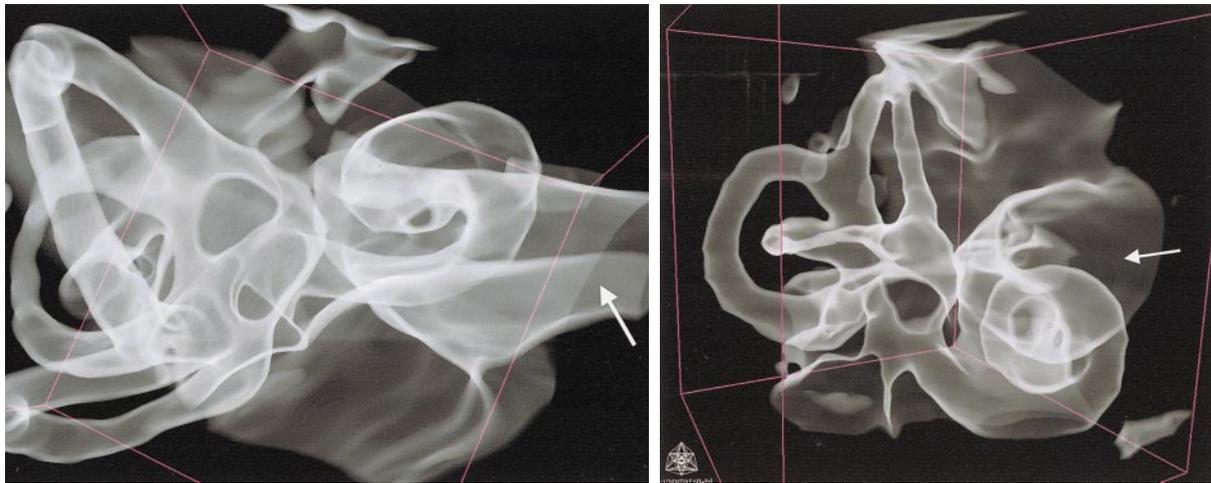
procedure, including the transfer of the data, was performed in less than 20 minutes. The standardized user interface ensured intuitive manipulation of any object in real time. This manipulation includes so-called clip planes, which can be applied to cut off disturbing structures virtually within the 3D representation. The software also allows distance measurements directly within the 3D scene.

For documentation of the visualization results, high-resolution screen images of selected directions of view were stored. The main target structures were (*a*) the semicircular canals, (*b*) the ampullae and vestibule, (*c*) the cochlea, (*d*) the inner

auditory canal, and (*e*) the mastoid and labyrinthine segments of the facial canal. In the cases of pathologic conditions, additional screen images that showed the best possible representation of the pathologic condition were stored.

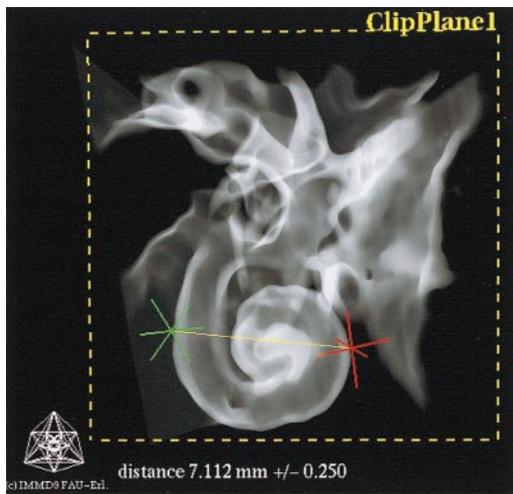
Results

In the five patients without pathologic conditions of the temporal bone, high-quality 3D visualization of the main target structures was possible (Fig 2). The selected volume of interest was large enough to show the entire facial canal as it runs through the temporal bone. For optimal delineation of the bony labyrinth, structures surrounding the otic capsule were suppressed from visualization with clip planes (Fig 3).



a.

b.



c.

Figure 3. Direct volume rendering of a normal left labyrinth including clip planes, which suppress disturbing structures within the 3D representation. **(a)** Superior view clearly shows the semicircular canals, ampullae, vestibule, cochlea, and inner auditory canal (arrow). **(b)** Anterolateral view shows the turns of the cochlea and the orthogonal semilunar canals. Arrow = inner auditory canal. **(c)** Anterolateral view shows the cochlea in detail and allows measurement of its diameter in a given clip plane.

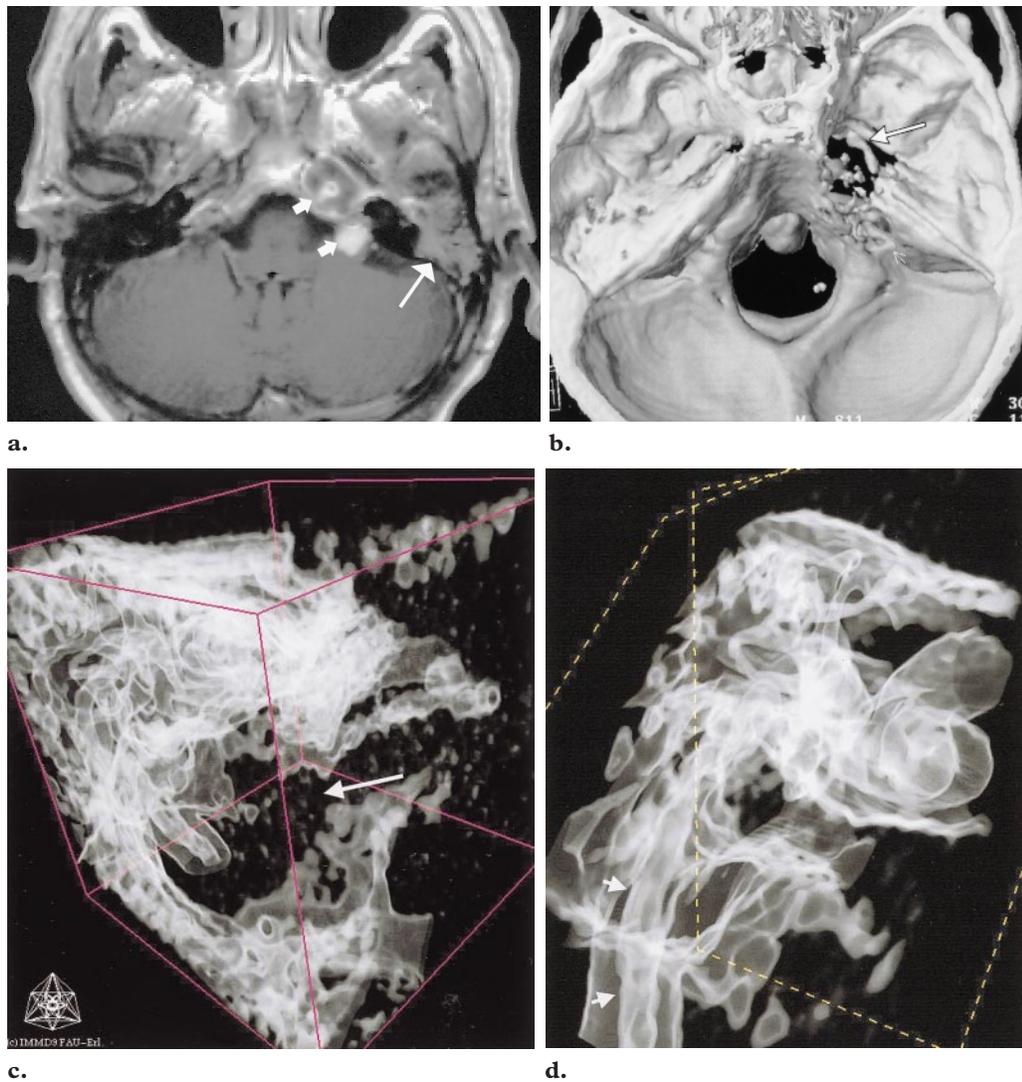


Figure 4. Chondrosarcoma of the temporal bone. **(a)** Gadolinium-enhanced axial T1-weighted MR image shows a partly cystic, enhancing tumor (short arrows). There is effusion in the temporal bone (long arrow). **(b)** Shaded surface display (SSD) image shows bone destruction in the apex of the pyramid (arrow). **(c)** Anterolateral view obtained with direct volume rendering shows the destruction caused by the tumor (arrow). **(d)** Anterosuperior view obtained with direct volume rendering does not show the inner ear clearly due to effusion within the temporal bone. Still, it is possible to see the mastoid segment of the facial canal (arrows).

In the patient with chondrosarcoma, there was an effusion of the temporal bone, making identification of structures difficult (Fig 4). In this case, a large volume of interest was required for in-

spection of the tumor and its relationship to surrounding structures. To overcome this problem, clip planes are a helpful tool for suppression of disturbing structures.

In the patient with schwannoma of the glossopharyngeal nerve, data from MR imaging, which

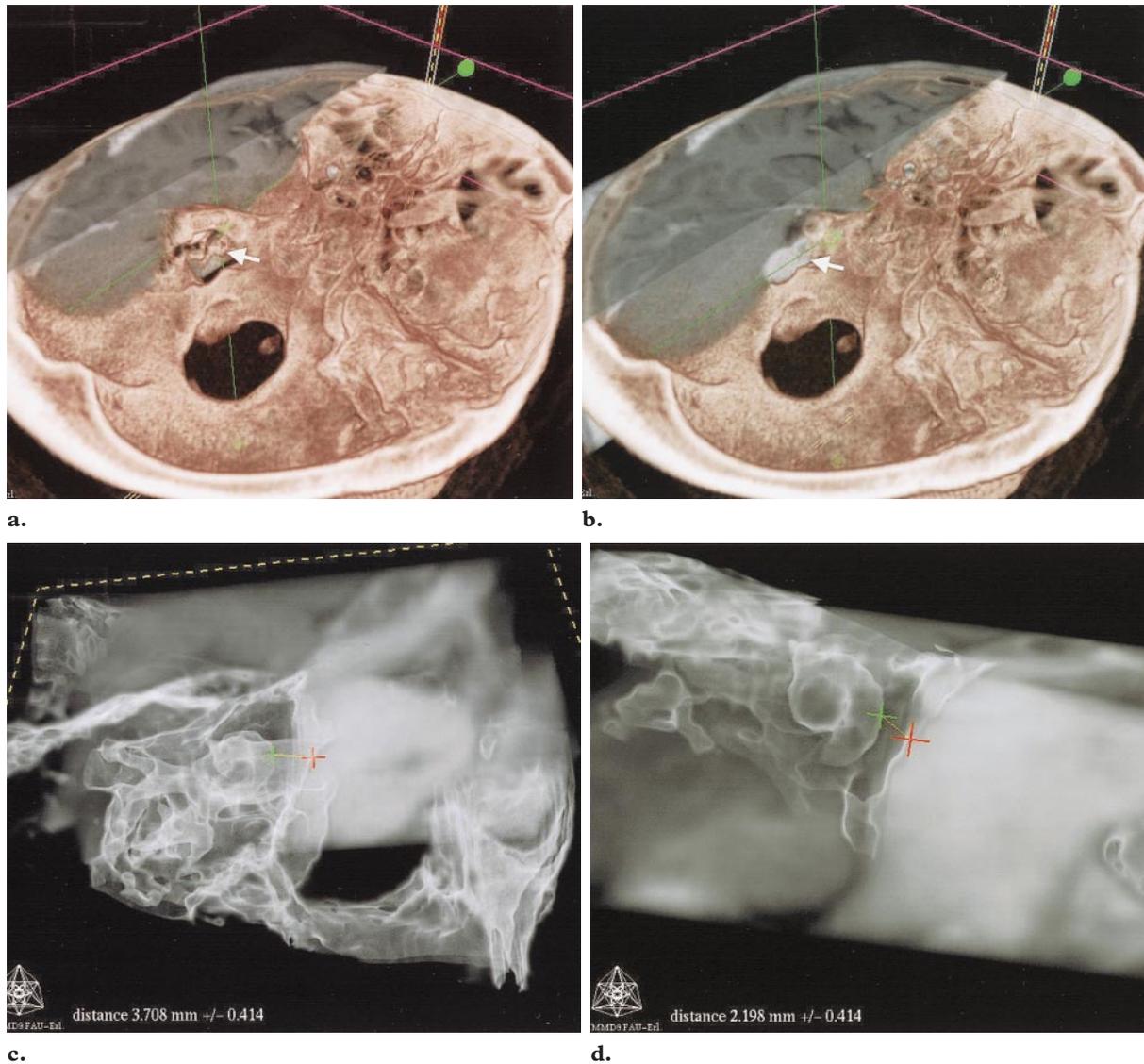


Figure 5. Schwannoma of the glossopharyngeal nerve inside the jugular foramen. **(a)** Fusion of MR and CT images shows a contrast-enhanced tumor (arrow) in relation to the skull base. **(b)** Fusion image after application of clip planes shows that the tumor has been eliminated to provide a view of the enlarged jugular foramen (arrow). **(c, d)** Superoposterior **(c)** and posteromedial **(d)** semitransparent views show the spatial relationship of the inner ear and the tumor. Interactive measurement of distances assists in therapy planning.

shows the contrast material-enhanced tumor, were fused with data from spiral CT, which shows the osseous structures. Thereby, the spatial relationship of the tumor and the temporal bone including the inner ear was clearly demonstrated (Fig 5). Despite use of a wide FOV (180 mm²), the 3D visualization based on direct volume rendering was of good diagnostic quality.

The images of the two patients with congenital malformations, which were reconstructed with a wide FOV (150 mm²), resulted in 3D visualizations of worse quality compared with those of the other patients. However, the reconstructed images were still of diagnostic value (Figs 6, 7).

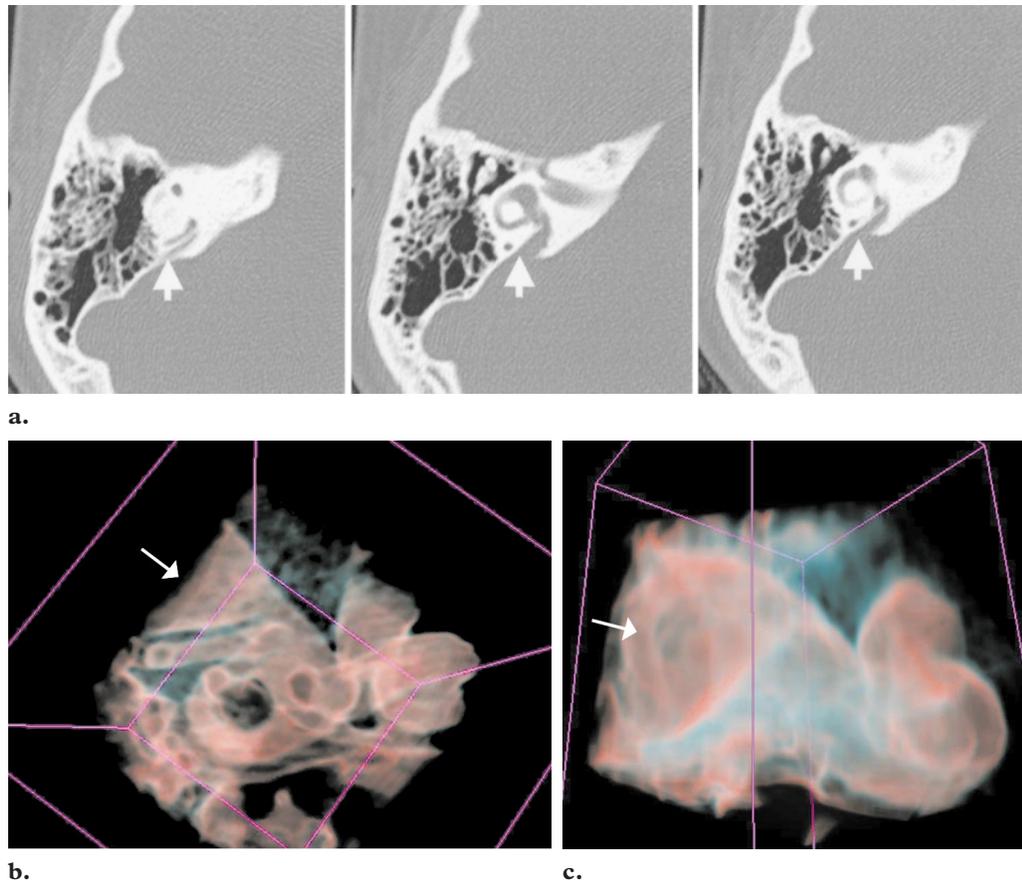


Figure 6. Enlarged vestibular aqueduct. **(a)** Section images (presented from superior [left] to inferior [right]) show an enlarged vestibular aqueduct (arrows). **(b)** Superior view obtained with direct volume rendering shows the spatial relationship of the funnel-like enlarged vestibular aqueduct (arrow) to the other structures of the inner ear. **(c)** Posterior view obtained with direct volume rendering shows the inside of the “funnel” (arrow).

Discussion

In 1993, Casselman et al (12) described use of a high-resolution MR imaging sequence for imaging of the inner ear. By using a 3D constructive interference in the steady state sequence and maximum intensity projection (MIP), 3D visualization of the membranous labyrinth in living subjects was easily performed.

Before this development, it was very difficult to produce 3D images of the inner ear from imaging studies in living subjects. For spatial understanding of the complex anatomy of the temporal bone, including the inner ear, drawings in anatomy books were the only source of information (13).

However, 3D visualization of the bony labyrinth from CT data is difficult due to the tiny size of the structures and their location within the temporal bone, which consists of many air-filled chambers. Therefore, it is almost impossible to visualize the structures of the inner ear with any commonly used technique without time-consuming explicit segmentation.

Volume Rendering versus Other Commonly Used Visualization Techniques

Various visualization techniques have been applied to produce 3D representations of the inner ear based on CT data. Reisser et al (4) and Frankenthaler et al (14) used automatic segmentation tools to extract the target structures from the im-

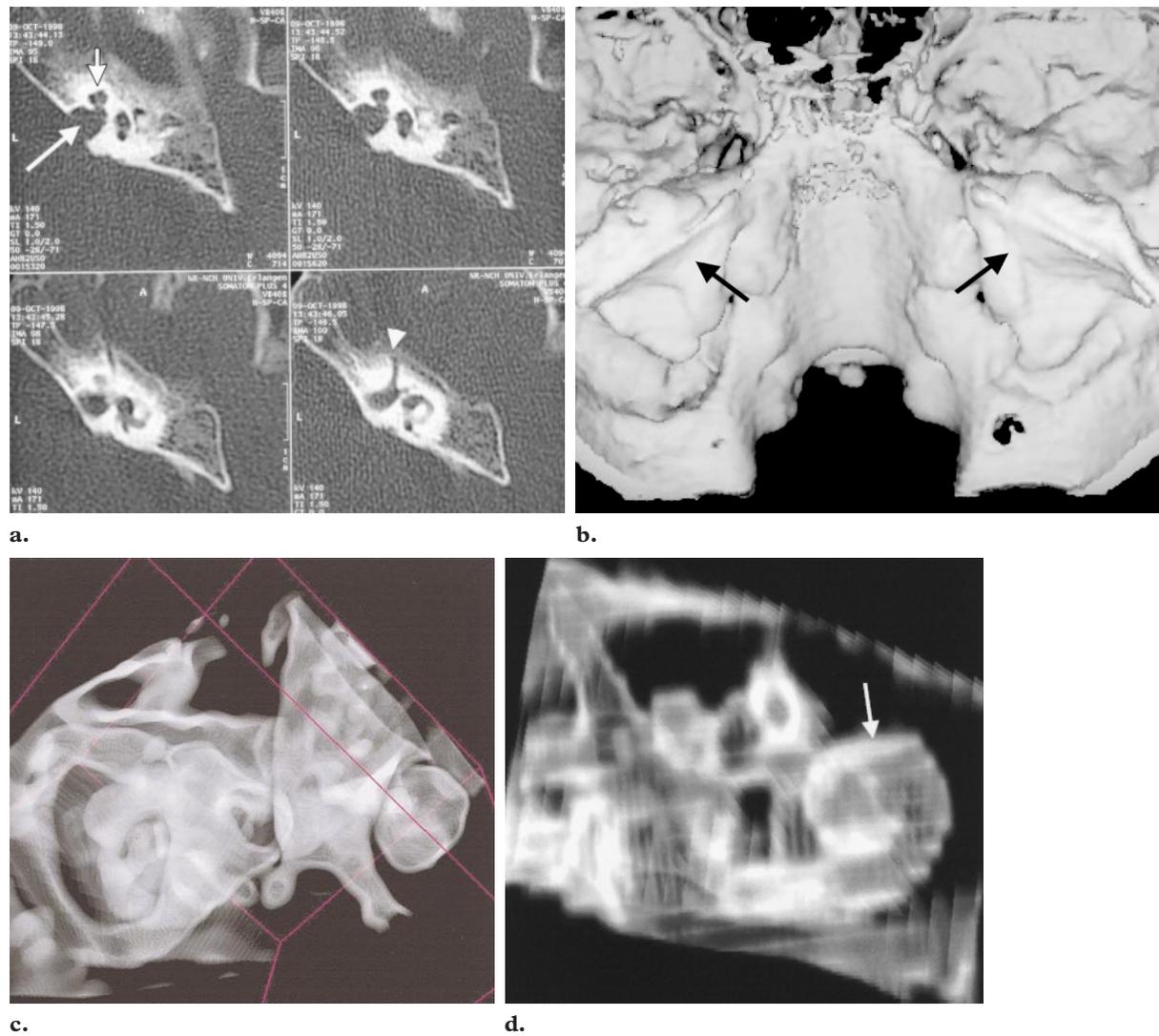


Figure 7. Complex deformity of inner ear structures. **(a)** Section images (presented from inferior [top left] to superior [bottom right]) show a small, irregular cochlea (short arrow); a short inner auditory canal (long arrow); a long labyrinthine segment of the facial nerve (arrowhead); and deformity of the temporal bone. **(b)** SSD image (posterosuperior view) of the skull base shows winglike deformity of the petrous parts of the temporal bones (arrows). **(c)** Superior view obtained with direct volume rendering shows the deformities. **(d)** Anterior view obtained with direct volume rendering shows only one turn of the cochlea (arrow).

age data set. The method used by Frankenthaler et al (14) produced beautiful 3D images of the labyrinth, the ossicles, and even vessels. However, a time-consuming process of several hours is required for segmentation of the image data and reconstruction of the geometric model. Therefore, it is difficult to apply this approach in routine clinical work. Conversely, direct volume render-

ing takes into account all of the information inherent in the CT data and provides a semitransparent representation of all structures. Because time-consuming preprocessing is not required, the analysis focuses on the visualization process exclusively.

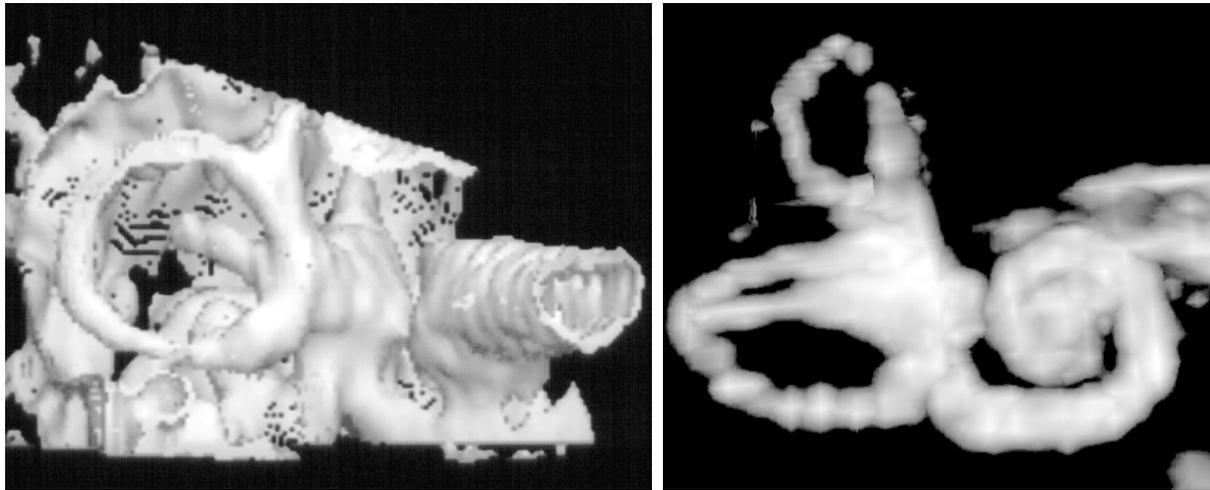


Figure 8. Commonly used visualization techniques. **(a)** SSD image (anteromedial view) shows the inner ear. Only parts of the structures are seen. **(b)** MIP image (anterolateral view) from high-resolution MR imaging data shows the inner ear.

For a better understanding of our approach, direct volume rendering is compared with SSD and MIP, which are commonly used for 3D visualization.

Shaded Surface Display.—In SSD, user-selected upper and lower thresholds are used to define a specific range of Hounsfield units to be displayed. Owing to the nature of the CT data, the result is opaque surface representations. To enhance the impression of depth, light effects producing shadows are included in the visualization. The drawback is that only the first voxel within the selected range of Hounsfield units along every ray of sight is displayed. Therefore, all structures behind that voxel are invisible (15). In theory, it should be possible to depict the structures of the inner ear just by selecting thresholds between the low value of air and the high value of bone. However, owing to partial volume effects, it is impossible to suppress the bone structures surrounding the inner ear. With a lot of patience, it is possible to create 3D representations (Fig 8a) that require time-consuming work but are of little diagnostic value.

Maximum Intensity Projection.—The other common approach, which is based on MIP, does not require selection of any thresholds. MIP displays only the brightest voxel along every ray of sight. As a result, all darker voxels in front of or behind a brighter voxel are not displayed. Thus, the depth information is lost, but some of the density information is preserved (15). Ambiguities easily occur because the displayed voxels are not necessarily part of the target structure, and important parts of a structure may disappear with only slight changes in viewing direction.

Obviously, the method is of no value for displaying hypoattenuating structures like the labyrinth due to the distribution of the data values within the structures of the temporal bone. However, the method is applicable for 3D display of the inner ear on the basis of high-resolution MR images (Fig 8b) (12).

Interactive Direct Volume Rendering.—In contrast to SSD and MIP, direct volume rendering uses all of the information contained inside a volume, thus allowing production of more meaningful images. By assigning a specific color and opacity value to every attenuation value of the

CT data, groups of voxels are selected for display. Depending on the selected opacity (high opacity produces low transparency, low opacity produces high transparency), nontransparent images similar to SSD images or semitransparent 3D representations providing views of all interior structures of the temporal bone are possible (5–7).

The software developed by the Computer Graphics Group of the University of Erlangen-Nuremberg is accelerated by hardware, thus taking advantage of the blending and interpolation capabilities of high-end graphics workstations. This feature allows real-time manipulation of any visualization parameter and guarantees high-quality 3D visualization of the inner ear (9–11). By interactively adjusting color and opacity values, the transition between soft tissue and bone structures is clearly displayed within the semitransparent representation, thus taking advantage of the partial volume effect. The use of clip planes allows one to cut off parts of the volume virtually, making visualization of the target structures more convenient. In addition, interactive measurement of point distances directly within the 3D scene assists in therapy planning (Figs 3c, 5c, 5d).

For high-resolution visualization, as shown in our examples, sophisticated software in combination with high-end graphics workstations is mandatory. Most of the volume rendering packages commercially available today do not provide the required high image quality.

Problems and Limitations

A major problem in 3D visualization is effusion within the temporal bone. To delineate the structures of the inner ear and the facial canal, it is important to have no objects of similar attenuation in the vicinity. Otherwise, it is difficult to sufficiently differentiate the target structures (Fig 4). Choosing a smaller volume of interest and applying clip planes improve the results achieved with direct volume rendering. In comparison, methods based on automatic segmentation perform poorly in this situation.

Virtual labyrinthoscopy can be performed with data from routine spiral CT examinations. After selection of an appropriate volume of interest within the section images (Fig 1), the whole visualization process is performed in less than 20 minutes without any explicit segmentation. To produce high-quality images, it is mandatory to use high-resolution CT images with the smallest possible section thickness, reconstruction of overlapping sections, and a narrow FOV mainly containing the target structures.

The relationship between the FOV and image quality is demonstrated in two cases of congenital malformations (Figs 6, 7). The section images in these cases were reconstructed with an increment of 1 mm and an FOV of 150 mm². The result was decreased quality of the 3D visualization.

If comparison of both sides is necessary, two or more sets of section images with a different FOV have to be reconstructed from the raw data. As a rule of thumb, the quality of the 3D images increases with decreasing size of the voxels of CT data.

Clinical Use of Virtual Labyrinthoscopy

In our preliminary study, we demonstrated the usefulness of virtual labyrinthoscopy in cases of congenital malformations. By using interactive direct volume rendering, the relationship between the structures of the inner ear is understood more easily. The method is also useful for preoperative therapy planning in patients with tumors of the temporal bone. In these cases, the location of the tumor relative to other structures inside the temporal bone is conveniently analyzed within the semitransparent representation. In the future, it should be possible to integrate this visualization approach into surgical navigation systems, thus providing the surgeon with additional orientation during surgery in the area of the inner ear.

Conclusions

Virtual labyrinthoscopy is a new approach to interactive 3D analysis of the inner ear. Major advantages of this technique are that it is reliably performed with standard spiral CT examinations and does not require explicit segmentation of the image data. Because the whole procedure is performed within a very short time, it is suitable for application in routine clinical work. Problems occur in patients with effusion in the temporal bone because adjustment of imaging parameters for proper delineation of the target structures is difficult in this situation. However, direct volume rendering is the only technique to produce meaningful images of high quality even in these problematic cases. The fast availability of a 3D representation as additional information after routine examinations with spiral CT makes this method a useful tool for further analysis of anomalies and tumors within the temporal bone.

References

1. Swartz JD, Harnsberger HR. Imaging of the temporal bone. New York, NY: Thieme Medical, 1998; 1–13.
2. Casselman JW, Bensimon JL. Imaging of the inner ear. *Radiologie* 1997; 37:954–963.
3. Ali QM, Ulrich C, Becker H. Three-dimensional CT of the middle ear and adjacent structures. *Neuroradiology* 1993; 35:238–241.
4. Reisser C, Schubert O, Forsting M, Sartor K. Anatomy of the temporal bone: detailed three-dimensional display based on image data from high-resolution helical CT—a preliminary report. *Am J Otol* 1996; 17:473–479.
5. Rubin GD, Beaulieu CF, Argiro V, et al. Perspective volume rendering of CT and MR images: applications for endoscopic imaging. *Radiology* 1996; 199:321–330.
6. Johnson PT, Heath GH, Bliss DF, Cabral B, Fishman EK. Three-dimensional CT: real-time interactive volume rendering. *AJR Am J Roentgenol* 1996; 167:581–583.
7. Remy-Jardin M, Remy J, Artaud D, Fribourg M, Duhamel A. Volume rendering of the tracheobronchial tree: clinical evaluation of bronchographic images. *Radiology* 1998; 208:761–770.
8. You S, Hong L, Wan M, et al. Interactive volume rendering for virtual colonoscopy. In: Proceedings of the Visualization Conference. Piscataway, NJ: Institute of Electrical and Electronics Engineers, 1997; 433–436.
9. Sommer O, Dietz A, Westermann R, Ertl T. An interactive visualization and navigation tool for medical volume data. In: Skala V, ed. Proceedings of the Sixth International Conference in Central Europe on Computer Graphics and Visualization '98 (WSCG). Plzeň, Czech Republic: University of West Bohemia, 1998; 362–371.
10. Hastreiter P, Rezk-Salama C, Tomandl BF, Eberhardt KE, Ertl T. Fast analysis of intracranial aneurysms based on interactive direct volume rendering and CT angiography. In: Proceedings of the Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI). New York, NY: Springer-Verlag, 1998; 660–669.
11. Hastreiter P, Tomandl BF, Eberhardt KE, Ertl T. Interactive and intuitive visualization of small and complex vascular structures in MR and CT. In: Proceedings of the Engineering in Medicine and Biology Society (EMBS), Institute of Electrical and Electronics Engineers. Vol 20. New York, NY: Springer-Verlag, 1998; 532–535.
12. Casselman JW, Kuhweide R, Deimling M, Ampe W, Dehaene I, Meeus L. Constructive interference in steady state-3DFT MR imaging of the inner ear and cerebellopontine angle. *AJNR Am J Neuroradiol* 1993; 14:47–57.
13. Netter FH. Nervous system, part I. The Ciba Collection of Medical Illustrations. Summit, NJ: Ciba-Geigy, 1983; 87–93.
14. Frankenthaler RP, Moharir V, Kikinis R, et al. Virtual otoscopy. *Otolaryngol Clin North Am* 1998; 31:383–392.
15. Regn J, Ezrielev J, Hupke R, Kalender WA, Maatsch K. Basics of CT angiography. In: Pokies H, Lechner G, eds. *Advances in CT III*. New York, NY: Springer-Verlag, 1994; 3–14.