ELSEVIER

Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric models

Yuvi Kahana*, Philip A. Nelson

Institute of Sound and Vibration Research, University of Southampton, Highfield SO17 1BJ, UK

Abstract

This paper investigates various aspects of numerically modelled individualised head-related (and pinna-related) transfer functions (HRTF). The computer simulations are based on the exact solution of the wave equation using the boundary element method (BEM). The basic features of the HRTF are investigated with accurate geometric models of two heads and six pinnae which are captured by using state-of-the-art three-dimensional (3-D) laser scanners and digitisers. These computer models are converted to valid BEM models and their frequency response is simulated. We present the results of simulated HRTFs, and show the inter-variability of the response among six baffled pinnae modelled in identical conditions. With current computing hardware power, and vigilant optimisation of the manipulated mesh models and the solving procedures, heads with pinnae (but without torso) can be investigated at least up to 10 kHz, and baffled pinnae can be investigated up to 20 kHz. We conclude that it is possible to implement individualised HRTFs in a 3-D sound system or an auditory display, without the need for measurements in an anechoic chamber, if highly accurate 3-D images of the head and pinnae are captured and modelled with BEM.
© 2006 Elsevier Ltd. All rights reserved.

1. Introduction

The localisation of the direction and distance of sound sources by humans can be accomplished by the use of acoustical cues only. For each listener, these are filtered by his or her own head-related transfer functions (HRTFs) [1–4]. For virtual acoustic imaging systems, these functions are often arranged in a database format in the time domain (termed head-related impulse responses—HRIRs) and consist of the impulse responses measured at each ear due to a spherical source positioned at a constant radius in a specified direction. The success of the production of virtual acoustic images in three dimensions around a listener is primarily based on the use of the "correct" HRTFs. It has been shown that with production over headphones, listeners localised virtual sounds with equal precision to sounds in real life, when binaural recordings were made individually [5], and with less precision when the recordings made in other ears, or the synthesised signals were produced with a generalised database of HRTFs [6]. Similarly, with production over loudspeakers listeners could discriminate successfully front and back confusion when monophonic sounds were synthesised with their own HRTFs [7].

---

*Corresponding author.

E-mail address: Yuvi@optoacoustics.com (Y. Kahana).

Since most three-dimensional (3-D) sound reproduction systems cannot use the "individualised" HRTFs of the listener, their databases are compromised and are based either on the HRTFs of a different listener (e.g. a good localiser), or on a dummy-head database. These are generally referred to as "non-individualised" HRTFs.

High-fidelity HRTFs are currently required by both the research community and the designers of virtual auditory displays. Traditionally, these databases are acquired by measurements using a rotating frame or inside a "loudspeaker cage" in an anechoic chamber [2]. The HRTF measurement procedure is very time consuming, and expensive. These are currently limited to well-equipped acoustic laboratories only, and as a result, many HRTFs are either confidential or restricted to research purposes. There are also many problems encountered when these functions are measured, analysed and compared between different studies. For example it is difficult to define the point at which the microphone should be positioned in the ear canal, the type of transducers used, equalisation techniques, dealing with signal to noise ratio problems, etc. In addition, HRTFs are generally measured only at discrete points with a low directional resolution. As a result, any real-time virtual auditory display would need to make use of interpolated functions. In this study, we suggest an alternative approach to acquire individualised HRTFs, by using computer simulation techniques rather than measurements. The conversion of optically acquired geometric data into its acoustical response should be achieved, in principle, by solving the wave equation.

In the last two decades, various computer simulation techniques have been suggested to model the modification of sound impinging on the human head or parts of the external ear. Weinrich [8,9] was the first to attempt modelling the response around a head. He also used analytical and numerical techniques in analysing the response of a very simple geometric model of the pinna with a mesh of only 20 elements resembling the shape of the concha. His solution was based on a finite difference approximation method, and the results approximated only roughly the dependence of the frequency response' first notch with elevation. Our current work was inspired by Katz [10,11] who acquired individualised HRTFs using the boundary element method (BEM) [12–14]. Although the idea was similar to the work of Weinrich, the use of BEM models converted from accurate laser-scanned models suggested that although the BEM is associated with 'low-frequency' modelling, we can now solve tens of thousands of simultaneous equations, and this predict the response of complex shapes such as the human head. Owing to limited computing power, his work was restricted to frequencies below 5 kHz; as a result, his simulations could not be validated in the high-frequency range where pinna resonance and anti-resonance affect the pressure variations.

Although the binaural technique is based on reproducing *exactly* the amplitude and phase information of the original sound in the ear-drums of the listener, it has been found that the blocked entrance is the most suitable point for binaural recordings and the measurement of HRTFs. This is mainly because the sound propagation in the ear-canal is independent of direction, and as a result, the complete spatial information can be detected at the entrance to the ear-canal and even a few millimetres outside it [15]. As a result, the measurement procedures of HRTFs have become more simplified, and in the case of BEM simulation, accurate models do not need to include an accurate description of the ear-canal.

In this research [16–19] (see Ref. [20] for full details), emphasis has been placed on the investigation of 'real' problems that include both large and arbitrary complex models. Although the BEM is a very powerful way of predicting accurately the vibro-acoustic characteristics of arbitrary structures, it is computationally expensive due to the inherent global connectivity of the equations needed to be solved. As a result, model solving optimisation was required both in obtaining the BEM mesh models and in simulation techniques and advanced BEM formulation.

Although the long-term goal of such a research project could be the acquisition of individualised HRTFs by means of an optical device for everyone, currently this numerical study can provide an accurate simulation tool which can replace complex measurement apparatus and procedures. Once the results were validated by measurements, as presented in Ref. [21], it was possible to investigate pinna modes and mode shapes [19,22], analyse the sound pressure anywhere on the surface of the head and the pinnae and also the scattered sound field around it. Finally, by integrating into the simulation the digital filters used to operate on the input to each source for virtual imaging systems, it was possible to visualise the resultant sound field either in the time, or frequency domains [20].

The paper first describes the infrastructure established for this study, namely the BEM method and the acquisition of accurate and usable mesh models. We then study individualised HRTFs and PRTFs and compare some of their characteristics with the literature. Validation of these results is presented in Ref. [21].

2. The boundary element method

In this section, we briefly review BEM formulations used in this study. The direct BEM (DBEM) and the indirect BEM (IBEM) can be used to solve different types of radiation and scattering problems. In some cases symmetry can be used to reduce calculation time, and in some cases special formulation of acoustic transparency in IBEM can be used in simulation with baffled pinnae with ear canal.

The numerical simulations were undertaken using the SYSNOISE software package [12], which uses the BEM in order to compute numerically the solution of the homogeneous Helmholtz equation. The complex pressure computed by the DBEM is given by

$$p(\mathbf{r}) = \int_s \left[p(\mathbf{r}_0) \frac{\partial g(\mathbf{r}|\mathbf{r}_0)}{\partial n} - g(\mathbf{r}|\mathbf{r}_0) \frac{\partial p(\mathbf{r}_0)}{\partial n} \right] \mathrm{d}S(\mathbf{r}_0), \tag{1}$$

which is Kirchhoff–Helmholtz integral equation where $g(\mathbf{r}|\mathbf{r}_0)$ is the free space Green function. SYSNOISE first solves the integral equation for the pressure $p(\mathbf{r}_0)$ on the surface S. Full details on further derivations of BEM which were used in this study are given in Refs. [12,20].

3. Mesh and BEM models

3.1. Acquiring accurate computer models

The initial assumption made during this work was that the highest resolution possible is required for the mesh models of the pinnae. There are currently a number of techniques available to obtain a computer model by scanning a physical model. These include computed tomography (CT), magnetic resonance imaging (MRI), and 3-D ultrasonic imaging. These are generally used for internal scanning for medical purposes. The main advantage of the 3-D laser scanner technique used in this research is that it can quickly produce an accurate mesh of the surface composed of triangles.[1] Two types of scanners were used: the Cyberware 3030RGB/ HIREZ which was used to capture the geometry of the heads, and the Cyberware 'Mini model' 3-D laser scanner which is ideal for small objects such as the pinna. With this model even the ear canal geometry can be obtained through multiple scans of the object and its cross-sections.

3.2. Mesh decimation

The original scan produced a polygonal mesh (see Ref. [23] for details) that describes the surface geometry of the pinna. Since the CPU time of the BEM increases drastically with the number of nodes (vertices), it is crucial to optimise the size of the mesh. It is well known that the maximum frequency in the BEM is proportional to the longest edge in the mesh. Any alteration to this global limit will distort the overall results. Therefore, a homogeneous distribution of the nodes and elements is required (see Ref. [24] for a thorough survey on mesh decimation techniques).

The main algorithm used in this research has been developed by Johnson and Hebert [25] and its main advantage is in successfully handling the two forces in mesh decimation: preserving the shape by limiting a defined maximum 'global shape error' and distributing the vertices homogeneously by local operators. This algorithm and more mesh manipulation techniques required in this study are described in detail in Kahana [20].

---

[1]Note that in principle, the use of quadrilateral elements can produce a higher accuracy of the simulation compared with triangular elements, but due to the format of the original data, quadrilateral elements were not used for scanned models.

3.3. Accurate BEM models

3.3.1. Artificial head

The KEMAR [26] artificial head (with pinnae and without the torso) was scanned. Two types of decimated KEMAR models are presented in Fig. 1. In both cases the original data included more than 400 000 triangle elements (around 200 000 vertices). The target was to obtain a suitable BEM mesh that could be used to modelling at the maximum frequency possible with the IBEM in-core solver.[2] The mesh on the left (Fig. 1a–c showing the vertices, elements, and rendered model, respectively) was decimated using our proposed algorithm that produces homogeneously distributed vertices, thus optimising its size, geometry, and maximum frequency. This resulted in approximately 23 000 elements that can be used up to 15 kHz if four elements per wavelength are assumed (for the ipsilateral ear), and 10 kHz if six elements per wavelength are assumed (for the contralateral ear).

When a conventional mesh decimation algorithm was used (the mesh on the right, optimised for computer graphics applications), and the number of elements was limited to 23 000 triangle elements, a mesh with non-uniform distribution of vertices was obtained (Fig. 1d), where planar areas were described with less triangles (Fig. 1e), and complex areas retained a higher density of triangles to preserve the geometry. Note that with both decimation algorithms the rendered images (Figs. 1c and f) are very similar. This mesh (on the right column) could be investigated using the BEM reliably only up to 1 kHz. This emphasises the significance of optimising the mesh distribution while retaining the accuracy of complex shapes such as the pinna. The pinna shown in these figures were scanned separately (see Section 3.3.2 for details) and was composed into this model. The decimation process resulted in geometrical errors of less than 1 mm for a given scanned vertex of the head, and less than 0.1 mm for a given vertex of the pinna.

In addition to the decimation algorithm applied on our models, it was found very useful to reorient the head in space so its ears are at the same height (both entrances to the ear canal are adjusted to be at $y = 0$). The interaural axis was defined on the line connecting these two points. Its mid-point was defined as the centre of the head which is the point of origin of the coordinate axis. Fig. 2 illustrates the head and the coordinate system. The azimuthal angles are defined as $0° \leqslant \phi < 360°$, $\theta = 0°$. At $\phi = 0°$, $\theta = 0°$ the source is in front of the listener, at $\phi = 90°$, $\theta = 0°$, the source is to the right ear, $\phi = 180°$, $\theta = 0°$ the source is in the rear, and at $\phi = 270°$, $\theta = 0°$ the source is on the left. The elevation angles are defined as $\phi = 0°$, $0° \leqslant \theta < 360°$, where $\theta = 0°$ is the horizontal plane, $\theta = 90°$ is above, and $\theta = 270°$ or $\theta = -90°$ is below (for convenience, in the figures which present median and lateral vertical planes, the elevation axis starts at $\theta = -90°$ and increases upwards).

The head was then divided into two identical parts through a vertical axis that intersects the head along the nose. The coordinate system is oriented such that the symmetry line that divides the head into two symmetrical parts, left and right, is located at $x = 0$. When a head was intersected at $x = 0$, some of the element vertices were not exactly located at $x = 0$ but a few millimetres to one or other side of the plane. These were snapped to $x = 0$ and the y and z coordinates were adapted in order not to change the element position and orientation. The new model was then decimated to a few mesh resolutions. Fig. 3 shows four different model resolutions of a half-model of KEMAR.[3]

When analysis at low frequencies is required, there is no need to use a very detailed model, so in practice the frequency range of interest is divided into 'band-pass' regions to optimise the size of the model and consequently its maximum frequency. This solving procedure, with mesh hierarchies, also reduces significantly the need for treatment of 'irregular frequencies' [27].

---

[2]SYSNOISE 5.4 is run with a 32-bit compiler. This limits the maximum RAM to be used with the in-core solver to approximately 1.2 Gb. Although more memory was available, it could not be used for larger models.

[3]These figures show the decimated KEMAR with the 'low-resolution' pinnae to emphasise the concept of 'mesh hierarchies'. In modelling, these pinnae were replaced with decimated high resolution pinnae.
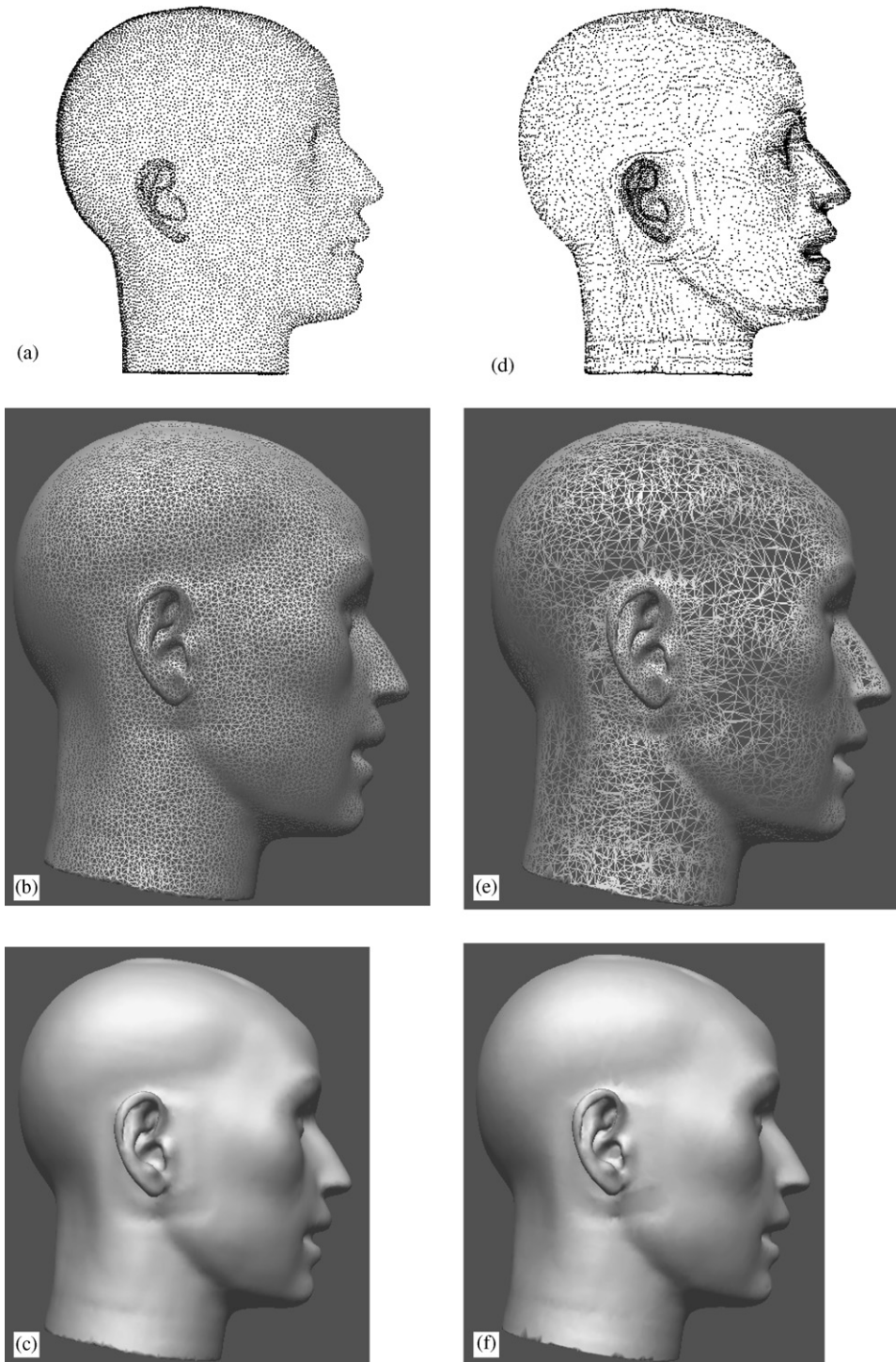
Fig. 1. Decimated mesh with vertices (top), elements (middle) and rendered models (bottom) of KEMAR. In both cases, the model comprises 23 000 elements and 11 500 vertices: (a–c) homogeneous mesh decimation optimised for the BEM. (d–f) conventional mesh decimation algorithm optimised for computer graphics.
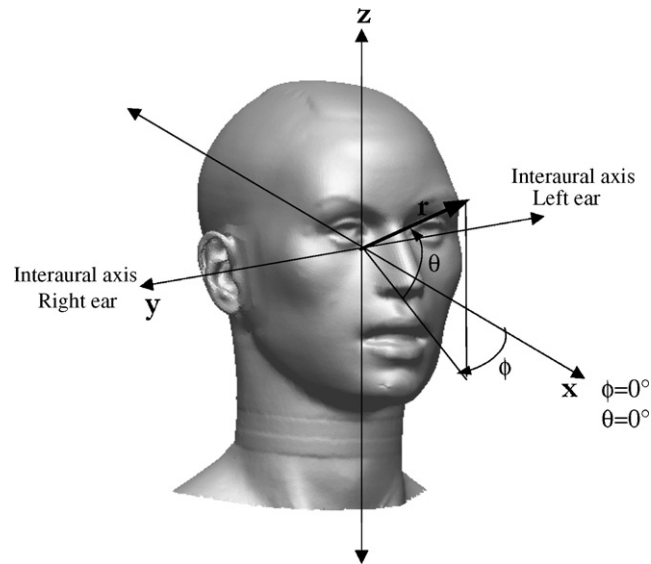
Fig. 2. Coordinate systems for simulation of HRTFs of human and artificial heads. The azimuthal angles are defined as $0° \leqslant \phi < 180°$, $\theta = 0°$. The elevation angles are defined as $\phi = 0°$, $0° \leqslant \theta \leqslant 360°$.
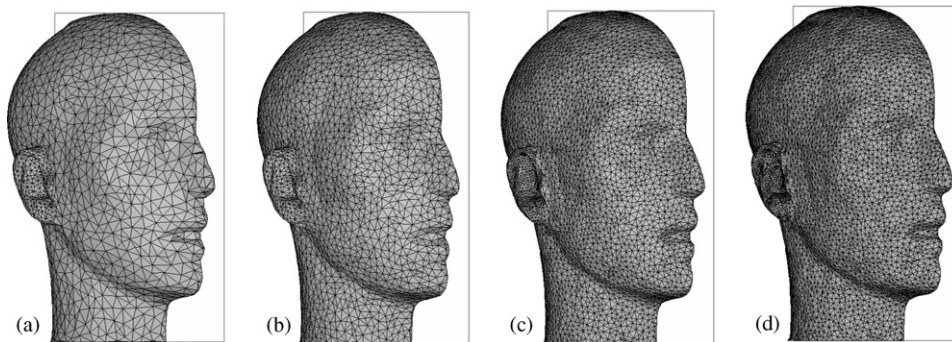

Fig. 3. Mesh hierarchy of half-models of KEMAR with the following approximate number of elements: (a) 2500, (b) 5000, (c) 10 000, and (d) 15 000. The pinna of the models is decimated separately with a higher resolution to enable the positioning of the source close to the entrance of the blocked ear canal, when the simulation is undertaken with the principle of reciprocity. Also shown is the plane of symmetry for each model.

3.3.2. Artificial pinnae

Seven pinnae were scanned and investigated. Four pinnae of KEMAR (the right pinnae: DB60[26], DB65[28], DB90 and DB95[4]), CORTEX [29,30], B&K [31], and the pinna of the first author (referred to as 'YK'). Since DB95 is not used to represent a typical human ear, and includes a large fitting space in the cavum concha for containing hearing aids, it was found that the resulting response was not representative of a typical ear, and it was decided not to include it in the following analysis of the results.

Fig. 4 illustrates the baffled pinna and the coordinate system, which is similar to the one used in Fig. 2. The origin of the coordinate system is located at the entrance to the blocked ear canal. All pinnae were scanned with the 'high-resolution' scanner. The original models included approximately 150 000 triangles (75 000

---

[4]Note that DB90 and DB95 have larger ear canal openings. Their geometry is mainly suited for the use of earmoulds.
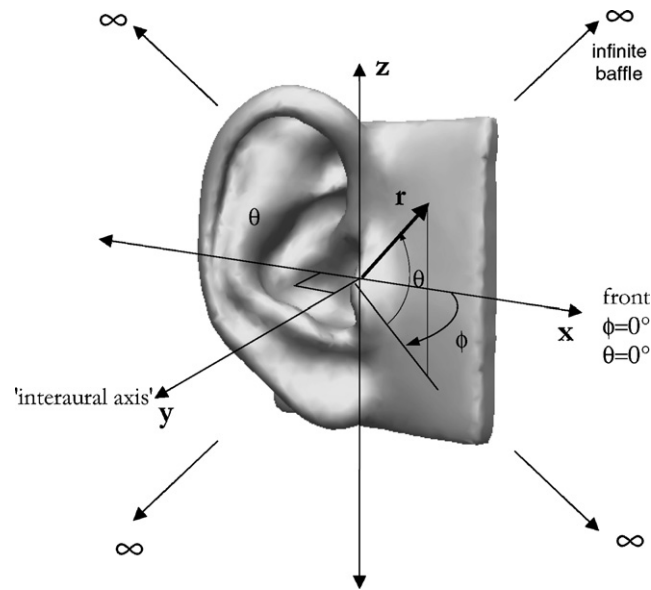
Fig. 4. Coordinate systems for simulation of baffled pinnae. Its centre is defined at the blocked entrance to the ear canal. 'Grazing incidence' angles are defined for $0° \leqslant \theta < 360°$, $\phi = 0°$. The 'Horizontal (azimuthal) Plane' for $0° \leqslant \phi \leqslant 180°$, $\theta = 0°$, and the 'Lateral Vertical (Frontal) Plane' for $-90° \leqslant \theta < 90°$, $\phi = 90°$.

vertices), and include 3-D information relating to the pinna (including its frame and base). Two strategies of model manipulation were used as follows: pinnae were fitted into the head (DB60 and YK pinna only) or pinnae were fitted into an infinite baffle by smoothing their boundaries to a specific plane ($z = 0$). In the first case, the surrounding of the pinna was removed, and only the outer, thin, shell surface remained. This model was later adjusted further to match the curvature of its boundaries and the surrounding of the pinna in the head (the mesh presented in Fig. 1 is composed of a head, and just such a high resolution pinna).

For the latter case, since the upper boundaries of the pinna do not lie on a single plane, a gradual slope was added from the higher level to the lowest level (approximately a height of 7–10 mm, as in the sides of artificial head pinnae). Later, the bottom level of the pinna was snapped to a single plane at $z = 0$. Six smoothed pinnae are presented in Fig. 5. The final BEM mesh models used in this study included on average 7400 triangles and 3800 vertices. These were investigated with mesh resolutions corresponding to a maximum frequency of more than 20 kHz. All models include a refined area around the blocked entrance to the ear canal for using the principle of reciprocity, where a monopole source is positioned close to the mesh surface. Geometrical properties and statistical values of these pinnae are presented in Appendix A).

Although we concentrated in our study on the 'blocked meatus' response, the addition of an ear canal is not impossible. Stinson and Lawton [32] showed large variations in ear canal shapes and sizes: the ear canal cross section area can vary between 40 and 90 mm^2 with an average of 65 mm^2. Also due to eardrum impedance variations, the calculated transformation of the pressure from the ear canal entrance to the innermost point in the ear canal can vary substantially, where up to 8 kHz the variations can be within a 4 dB from the average, and above this frequency range the variations can be of 20 dB and more. Therefore, we chose to add a typical cylinder with the dimensions of 22.5 mm length and 7.5 mm in diameter (as used by Zwislocki [33]). The difficulty in the process of this mesh is the special treatment required with the 'IBEM transparency' module. Since the baffle lies above the ear canal and below the pinna, and sound waves are required to propagate 'through' the entrance to the ear canal, the pinna must be oriented such that the ear canal entrance is exactly at $z = 0$. Then the entrance must be closed with a planar surface with elements at $z = 0$ and the edges at the boundary of the circular surface should be connected to edges of the canal below and edges of elements of the pinna above. These added elements will be defined as 'transparent' elements in the solving procedure (see detailed formulation in Ref. [12]).
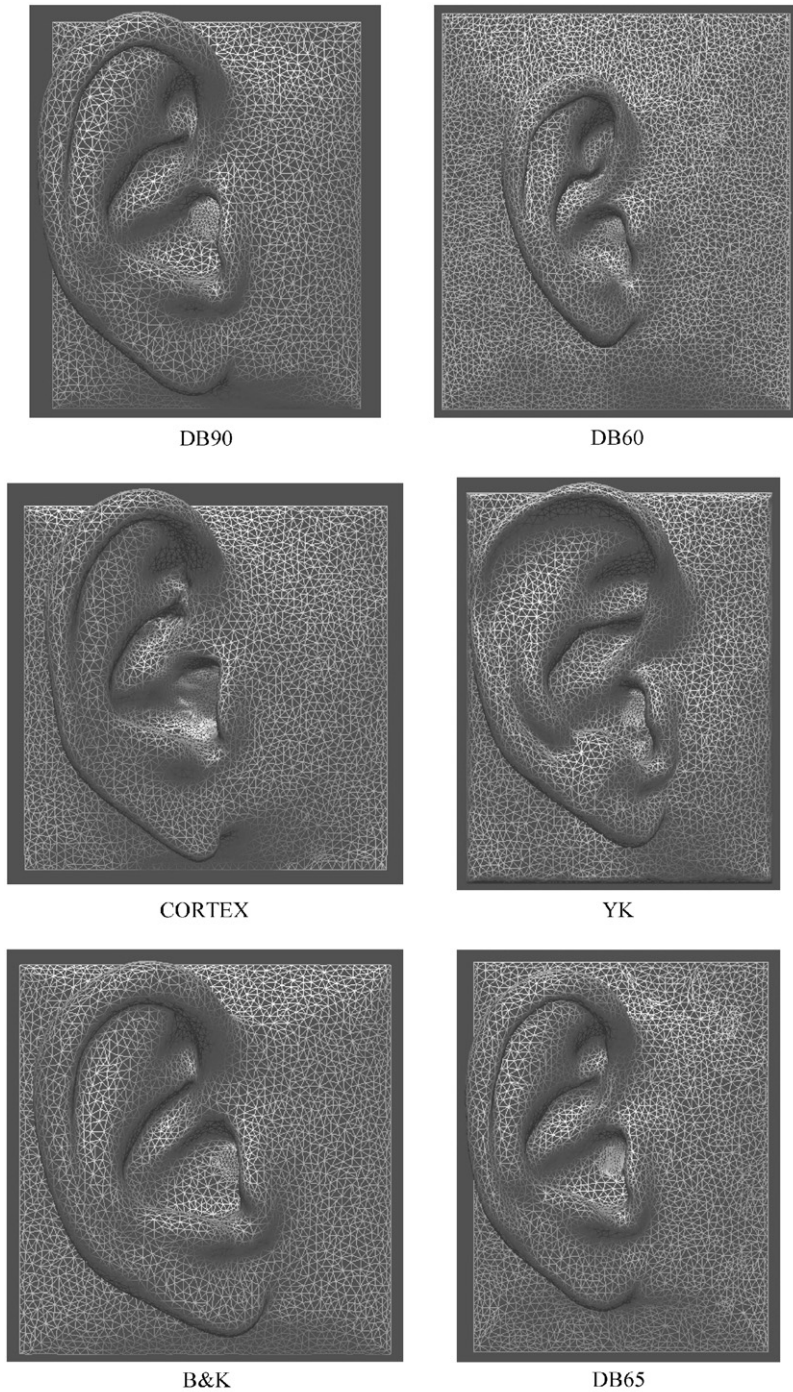
Fig. 5. BEM pinnae models: decimated, aligned and smoothed into a rectangular frame lying at $z = 0$. Distortion of the geometries of the pinnae was kept minimal.

An example of the cross-section of the baffled DB60 and the simplified ear canal (without the transparent elements) is given in Fig. 6. Note that in principle, the addition of the ear canal to the model of the *head* is a much simpler process, since the addition of transparent elements is not required, and the entire mesh model remains a closed volume.
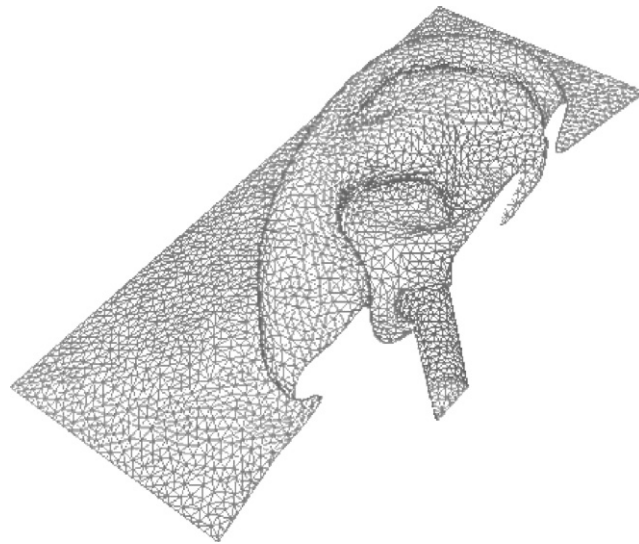
Fig. 6. Decimated BEM model of DB60 with the addition of a cylindrical ear canal. The entire model consists of 8189 nodes and 16 113 elements.
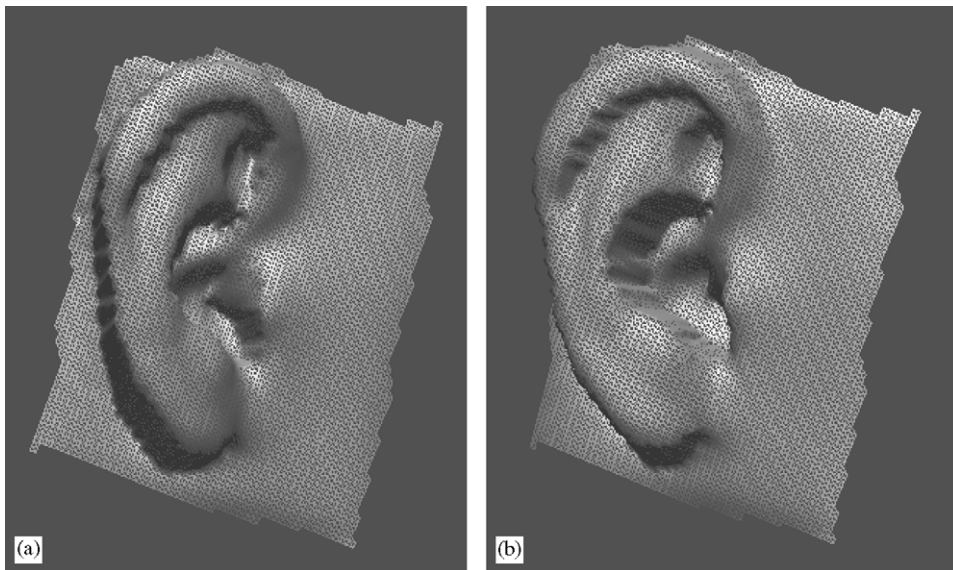

Fig. 7. Original 'low resolution' DB60 mesh model with approximately 6000 nodes and 12 000 elements. Two views are shown to demonstrate the coarse representation of the pinna (a) a 'solid' rear and shallow cavum concha (b) distorted antihelix and the posterior wall of the cavum concha.

Figs. 7a,b show two views of the DB-60 pinnae that were scanned with the head using the 'low-resolution' scanner. These cut-out models are shown before any decimation was applied. Cleary seen are the lack of details, the distortion (in the form of extrusion) in the back of the posterior wall of the concha, and the shallow cavity of the concha. This model will be investigated in Section 5.3.4 in order to study the degradation of the acoustical response when compared with the results obtained with the high accuracy mesh model of the DB60 pinna (this can establish whether a low resolution picture of a pinna, obtained with standard digital camera can provide adequate information for BEM modelling).
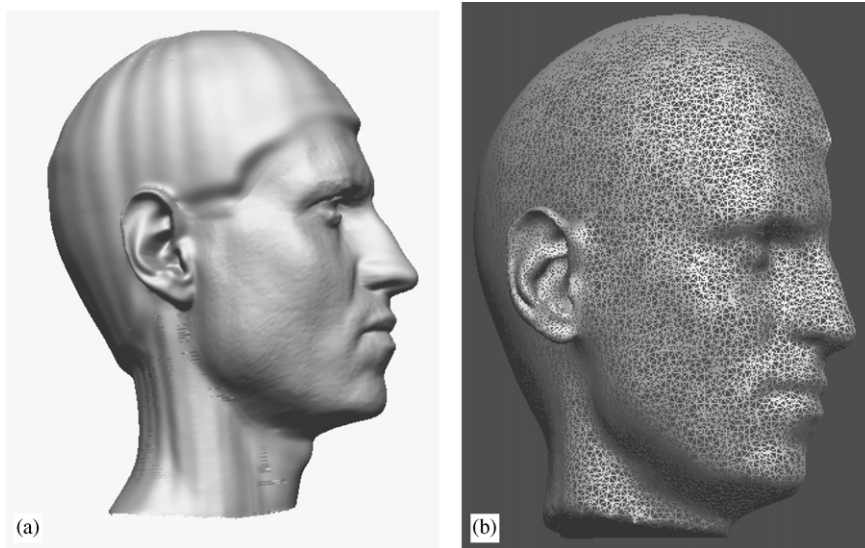
Fig. 8. YK head model: (a) Original rendered model with a 'shower cap' used to conceal the hair. The model consists of 210 000 vertices and 418 000 elements. (b) Hybrid model includes decimated head and pinnae captured with two types of scanners.

3.3.3. Human pinnae and head

Since we are interested eventually in modelling *individualised* HRTFs, the head of the first author (referred to as the 'YK head') and its moulded pinna were scanned. Since the high resolution scanner required a small model, two plaster moulds were produced. One was cut vertically, so that the scanner can accumulate 3-D surface geometrical data of the concha and the entrance to the ear canal.

The main problem when the head was scanned was the elimination of hair, which was achieved using a 'shower cap', as seen in Fig. 8a. The final manipulated BEM model is shown in Fig. 8b, where the shape of the scalp was smoothed, the original 'low-resolution' pinna was replaced with the accurate pinna, and the resolution optimised for maximum frequencies in the region of 10 kHz.[5]

4. Modelling individualised HRTFs

4.1. Limitations of the model

When HRTFs are measured with subjects in the anechoic chamber the sound transformation detected very close to the eardrum includes all the contributions form the body, torso, neck, head, pinnae, ear canals and eardrums, as well as the effects of clothing, and hair. In our simulations we limited our model to include only a rigid head and pinnae that are blocked at the entrance to the ear canals. We therefore expect the following discrepancies when our model is compared with an average measurement.

Blocked ear canal: The valuable positions of the 'microphone' at the blocked ear canal was suggested initially by Yamaguchi and Sushi [34], investigated extensively by Shaw [35], and evaluated with the binaural technology by Møller et al. [36] and Hammershøi and Møller [15]. It should be born in mind that the response detected at this position equalised with the free-field response (without the head) is the only directional component in the HRTF since the transformation in the ear canal is insensitive to direction up to around 12 kHz. The transformation therefore needs to be compensated for the ear canal response.

Missing torso: The effects of the neck and torso were investigated by Burkhard and Sachs [37], Preves (KEMAR [38]) and Kuhn [39]. Males and females have on average different neck lengths. These differences

---

[5]Due to larger dimensions of this head compared with KEMAR, for a given maximum number of elements that could be handled with the IBEM in-core solver, lower resolution with larger elements were obtained, hence slightly reduced frequency range compared to KEMAR.

can shift the frequency of the first pressure minimum (see Ref. [37]), detected at the eardrum for sources in the front from 1.4 to 1.2 kHz. As a result, KEMAR includes three neck sizes. The torso also contributes mainly up to 2–3 kHz. At 1.2 kHz, a clothed torso can alter the pressure by 3 dB compared with the head only.

Boundary conditions: The effects of simulating flesh, hair and clothing were also investigated by Shaw [35] and Burkhard and Sachs [37]. Hair in the form of a wig makes the pressure minima at the eardrum at around 10 kHz less deep but has very little effect at lower frequencies. The effect of neglecting the impedance of the skin is found to be negligible (< 1 dB difference below 8 kHz). Kuhn [39] found that ITD and ILD of KEMAR are different depending whether or not the torso is bare or clothed. Differences of ILD can be up to 3–5 dB up to 3 kHz, and ITD can be reduced significantly (from ~600 to ~500 μs) below 700 Hz if the torso is clothed. Modelling the impedance values of the hair was carried out by Katz [11]. He found that the hair contributes up to 6 dB at frequencies below 5 kHz. Shaw [35] stated that the contribution of hair is probably less than 3–5 dB (the effect of clothing) at very high frequencies.

4.2. Results

Visualisation of HRTFs could be presented in many ways, since the complex structure at high frequency varies significantly in both frequency and time domains, and with angular directions. In Fig. 9, 3-D colour maps show how the magnitude of the pressure in the near field (0.5 m) at the reciprocal points, changes with an angular position. The response is shown at discrete frequencies. As the frequency increases, the dynamic range obtained becomes higher. At 200 Hz, the variation is almost omnidirectional, and variations of ±3 dB are only due to proximity to the head, and equalisation with respect to its centre. At 1 kHz the head still has similar characteristics to a sphere, with equal contours of magnitude in both the ipsilateral and contralateral ears. Note that the minimum pressure occurs at angles that are not directly opposite the right ear, due to superposition of waves with equal path lengths. At 2 kHz, slight variation can be seen between positive and negative elevation angles. The pinna still has very little effect at this frequency. At 5 kHz the boost is mainly due to the first resonance of the concha (17 dB at this distance). We can expect even more complex variations in the contralateral ear as frequency increases.

These figures demonstrate the advantage of obtaining continuous maps of HRTFs. As frequency increases the complex patterns show that interpolation of HRTFs based on low-resolution sampling of measurements will produce large errors, especially for the contralateral ear. In the next sections, we investigate the variation of HRTFs with a higher frequency range at different planes.

In order to compare the results obtained with the numerical simulation and the literature, we present below the response in three different planes. Further simulation results with the KEMAR head are compared with measurements and presented in Ref. [21].

The results of the simulations are given in three planes as illustrated in Fig. 2: median vertical plane, horizontal (azimuthal) plane and lateral vertical (frontal) plane.

4.2.1. Median vertical plane

The modelled amplitude and magnitude of the HRTFs of the YK head are presented. In all cases, the response is detected at the blocked ear canal of the right ear, and the sources are positioned on a sphere with a radius of 2 m. Figs. 10a and b present the response due to sources in the median vertical plane. Since the torso is absent the angles below −40° and above 220° are meaningless and therefore are excluded from the figures. The figures are presented on both linear and logarithmic scales. The following observations can be made and compared with the literature.

The first resonance is clearly seen around 4 kHz. It almost does not change with the angle of elevation, but a maximum amplification is obtained when the sound is in the front and above, between $\theta = 0°$ and 40° (see Ref. [40] for similar trends although the pinna is baffled). At all angles the first resonance causes an amplification of 10–17 dB (as in Refs. [5,15,41]).

It has been noted by several researchers that a possible frontal cue could be the low pass notch that moves from 6 to 10 kHz as elevation increases from $\theta = -40°$ below the horizontal plane and up to $\theta = +60°$ above the plane (Shaw [35], Butler and Belendiuk [42] and Mehergart and Mellert [43]). The low-pass filter for these angles is clearly seen in the figures. This notch can be explained by the fact that for higher elevations, the
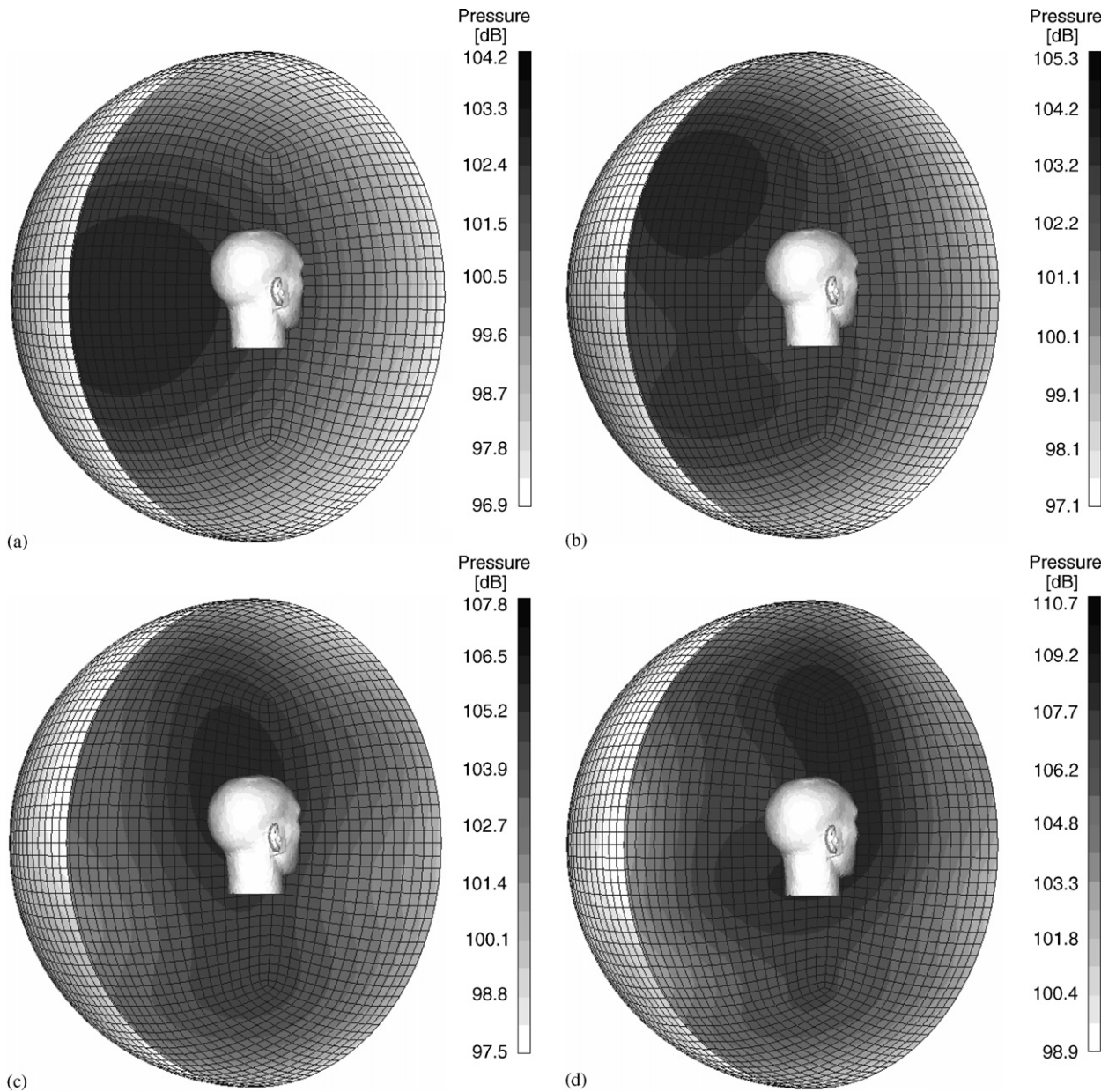
Fig. 9. Maps of HRTFs at single frequencies. The spherical source is positioned very close to the entrance of the left blocked ear-canal of KEMAR. The pinna resolution is refined. The grey scale represents the magnitude of the pressure at the left ear due to a source at that position: (a) 1000 Hz. (b) 2000 Hz. (c) 3000 Hz. (d) 4000 Hz. The response in the "cone of confusion" is clearly seen at low frequencies as the magnitude is the same in the front and in the rear. As frequency increases the diffraction around the pinna causes the response to be higher in the front, and lower in the rear.

reflected path length is shorter, so that the notch in the spectrum is at an increased frequency (Hebrank and Wright [44]).

Blauert [1] studied the perception of elevation in three sectors: frontal ($-15° < \theta < 45°$), above ($45° < \theta < 135°$), and rear ($135° < \theta < 200°$). A 'boosted band' appears around 8 kHz when the source is overhead (Blauert [1, pp. 110, Fig. 2.47]). Hebrank and Wright [44] also noted that sources above are characterised by a 1/4 octave peak between 7 and 9 kHz. In our case, the maximum amplification shown in Fig. 10b occurs at an angle of $\theta = 80–90°$ overhead and the frequency is 7.8 kHz.
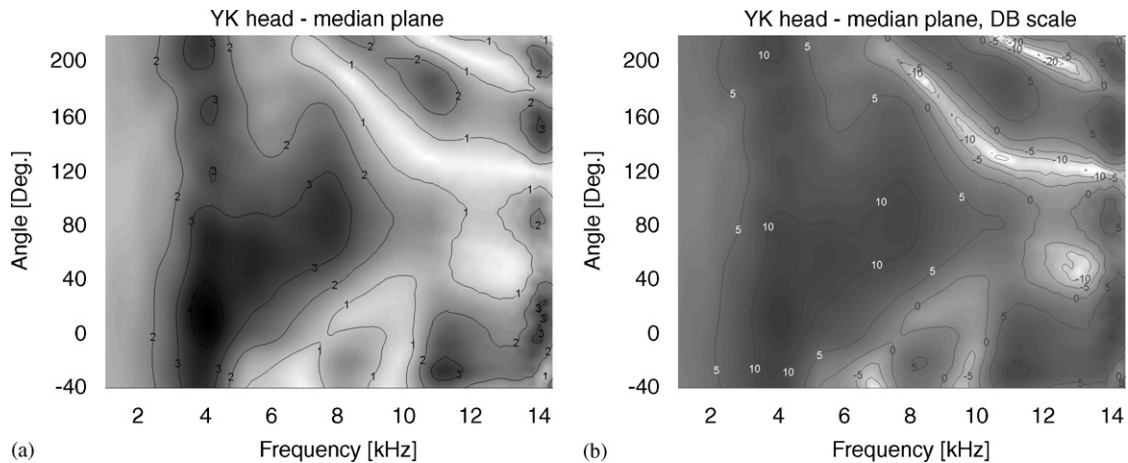
Fig. 10. Normalised median vertical plane HRTFs for the YK head (the pressure detected at the blocked ear canal and divided by the pressure detected at the centre of the head and the head was absent): (a) linear amplitude, (b) magnitude in dB. Simulation undertaken at a resolution of 1°, and steps of 200 Hz (71 frequencies).
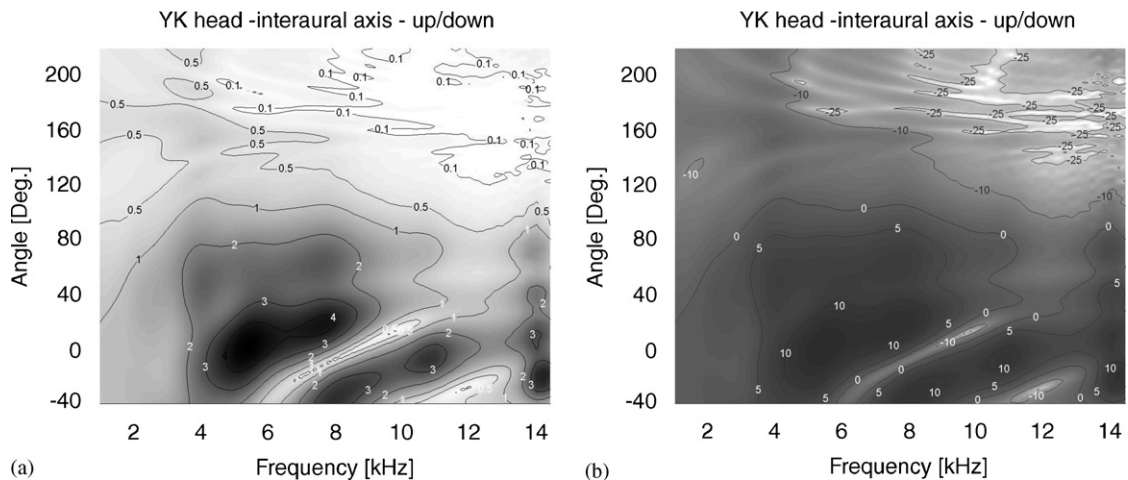

Fig. 11. Normalised lateral vertical plane HRTFs of the YK head (a) linear amplitude (b) magnitude in dB. Parameters as in Fig. 10.

4.2.2. Lateral vertical plane

With the simulation of HRTFs in the lateral vertical plane (presented in Figs. 11a and b) a larger dynamic range is noticed due to the shadowing of the head. The structure of the notches at contralateral angles is very complex. The general structure of this plane can be compared to the data measured on 40 subjects by Møller et al. [36] in this plane: the frequency of the notch at angle $\phi = 90°$, $\theta = -40°$ starts at 6 kHz and increases with angles up to 10 kHz at $\phi = 90°$, $\theta = +30°$. The same phenomenon appears in their results (Møller et al. [36, Fig. 15]), and also in Shaw and Teranishi [41, pp. 248].

The maximum pressure is obtained at $\phi = 90°$, $\theta = 0°$ at 5.5 kHz with a boost of 15 dB. The attenuation of angles in the shadow zone reaches levels of -33 dB, and the overall dynamic range of HRTFs in this defined plane is almost 50 dB. Therefore, the features of the HRTFs cannot be characterised on this scale. Also, the reliability of the large attenuation in the shadow zone is probably not very high due to lack of mesh resolution, and this is investigated for the case of KEMAR against measurements in Ref. [21].
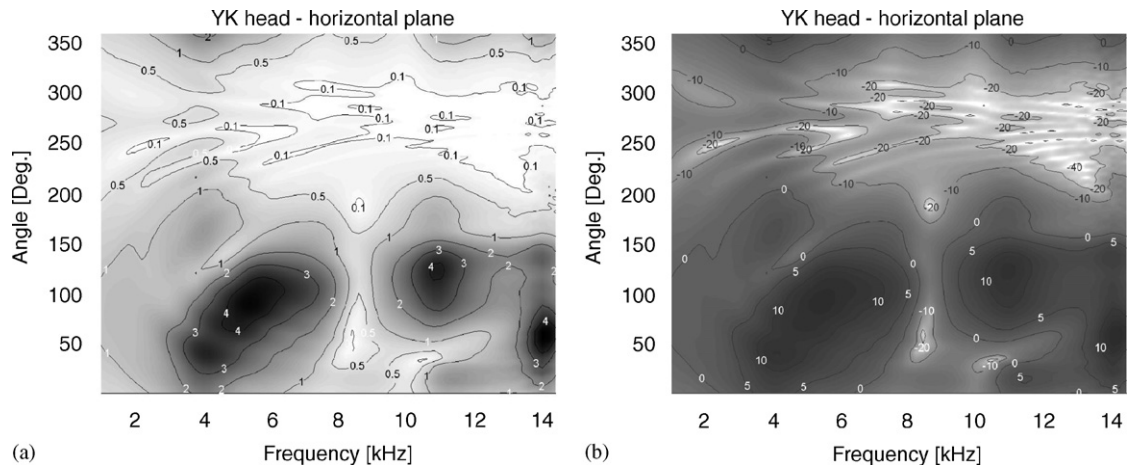
Fig. 12. Normalised horizontal plane HRTFs of the YK head: (a) linear amplitude (b) magnitude in dB. Parameters as in Fig. 10.

4.2.3. Horizontal plane

The variation of HRTFs in the horizontal plane (presented in Fig. 12) is mainly characterised by the deep notch whose frequency (8.5 kHz), is fairly constant. Similar features exist in Ref. [40, p. 29] (although the minima there occur at 9.5 kHz, see also Section 5.3.3). In his averaged data from 12 studies (more than 100 subjects), the measurement is made close to the eardrum. However, it was shown in Shaw and Teranishi [41], Figs. 4 and 10, that the minimum is present in both cases and that it is a result of an interference effect. They concluded that the minimum around 8 kHz is relatively independent of the angle of incidence in the horizontal plane as is shown in our case, although this minimum is not present in the contralateral ear response (as can be seen in Fig. 12).

5. Modelling the response of baffled pinnae

In this part of the study, we orient the manipulated pinnae models, shown in Fig. 5, so that their boundaries lie on the baffle plane, $z = 0$. The coordinate system is presented in Fig. 4. Its centre is defined at the blocked entrance to the ear canal. The 'Median Vertical Plane' is called also 'grazing incidence[6]' for $0° \leqslant \theta < 360°$, $\phi = 0°$. The 'Horizontal (azimuthal) Plane' for $0° \leqslant \phi \leqslant 180°$, $\theta = 0°$, and the 'Lateral Vertical (Frontal) Plane' for $-90° \leqslant \theta < 90°$, $\phi = 90°$.

5.1. Problem optimisation with the BEM

The principle of reciprocity is implemented here for computation efficiency. As can be seen in Fig. 5 the resolution of local elements in the vicinity of the entrance to the ear canal was increased, since a monopole source was positioned very close to this area. Fig. 13 shows the response at the blocked ear canal of a baffled pinna (DB60 with 6887 nodes and 13 488 elements). Two cases are investigated: the first, when the source is positioned in the far field, and the frequency response is detected 1 mm away from the surface of the blocked entrance of the ear canal and the second, when the source is positioned 1 mm away from the surface, and the pressure is detected in the far field at the position that was previously that of the source. It is shown that the results with the IBEM are nearly identical providing that enough over-determination points (30) are inserted in the cavity between the pinna and the baffle, to deal with the problem of irregular frequencies in the solution [26].

---

[6]Note that 'grazing incidence' was defined at a normal distance of 5 cm from the baffle, for compatibility with measurements since this was the minimal distance between the centre of the loudspeakers cone, and the baffle at $\theta = \pm 90°$.
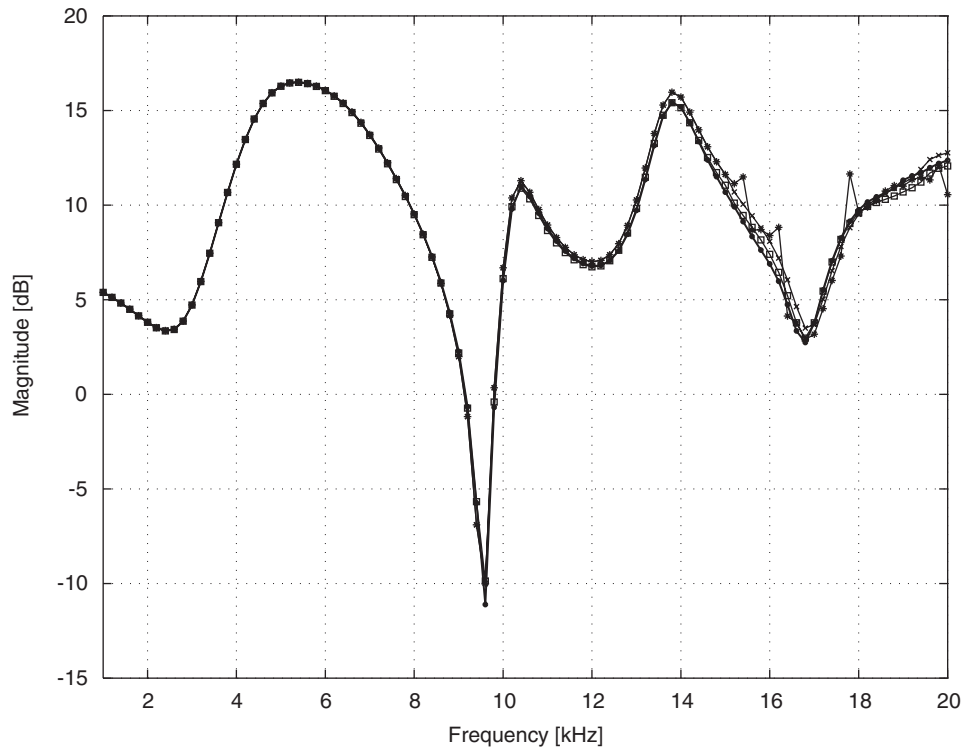
Fig. 13. Validation of the principle of reciprocity using a baffled DB60 pinna. The model consists of 6887 nodes and 13 488 elements. The curves shown are the frequency response at the entrance to the ear canal due to a source at normal angle, $\phi = 90°$. Reciprocity is checked when the pressure is investigated on the surface of the pinna due the source in the far field and also at the position of the source which is 1 mm away from the surface. —○— Source in the far field, pressure on the surface, with 30 overdetermination points —∗— reciprocity, without over-determination points, —□— reciprocity, with 30 overdetermination points, —×— reciprocity, pressure 1 mm away from surface, with 30 overdetermination points.

5.2. The effect of the baffle on the response

The pinna is attached to the side of the head at different positions and angles among individuals (see Refs. [37,45]). By attaching the pinna to a baffle we inevitably change the transformation of sound detected at the eardrum, or blocked entrance to the ear canal. However, since we investigate only the ipsilateral side, we assume that differences between the baffle and the head will not be significant, if we limit our expectations to the general structures of peaks and notches, varying across frequency and directions.

In Fig. 14(a–f), we compare the response of two pinnae (DB60 and YK pinna) for the baffled case and when the pinna is attached to the head. The positions of the sources in space were aligned approximately with the same angular directions since the coordinate systems are different in both cases. The plots on the left compare the responses at grazing incidence angles. These are very similar, and the slight differences can be attributed to the fact that the source and microphone arrangement in the two cases is different and therefore the angle of excitation is slightly misaligned. However, both the frequencies of the peaks and notches as well as the magnitudes represent the response obtained with the full head.

On the right, three source positions in the horizontal plane have been investigated with KEMAR and DB60. As the source moves to the rear, the error increases due to interaction with the baffle.[7] In this case, we can still study the general trend of the variation of the response but the accuracy is frequency and direction dependant, in particular if all pinnae under investigation have the same baffled conditions.

---

[7]In the figures on the right the response of KEMAR was 6 dB higher than presented due to free-field equalisation.
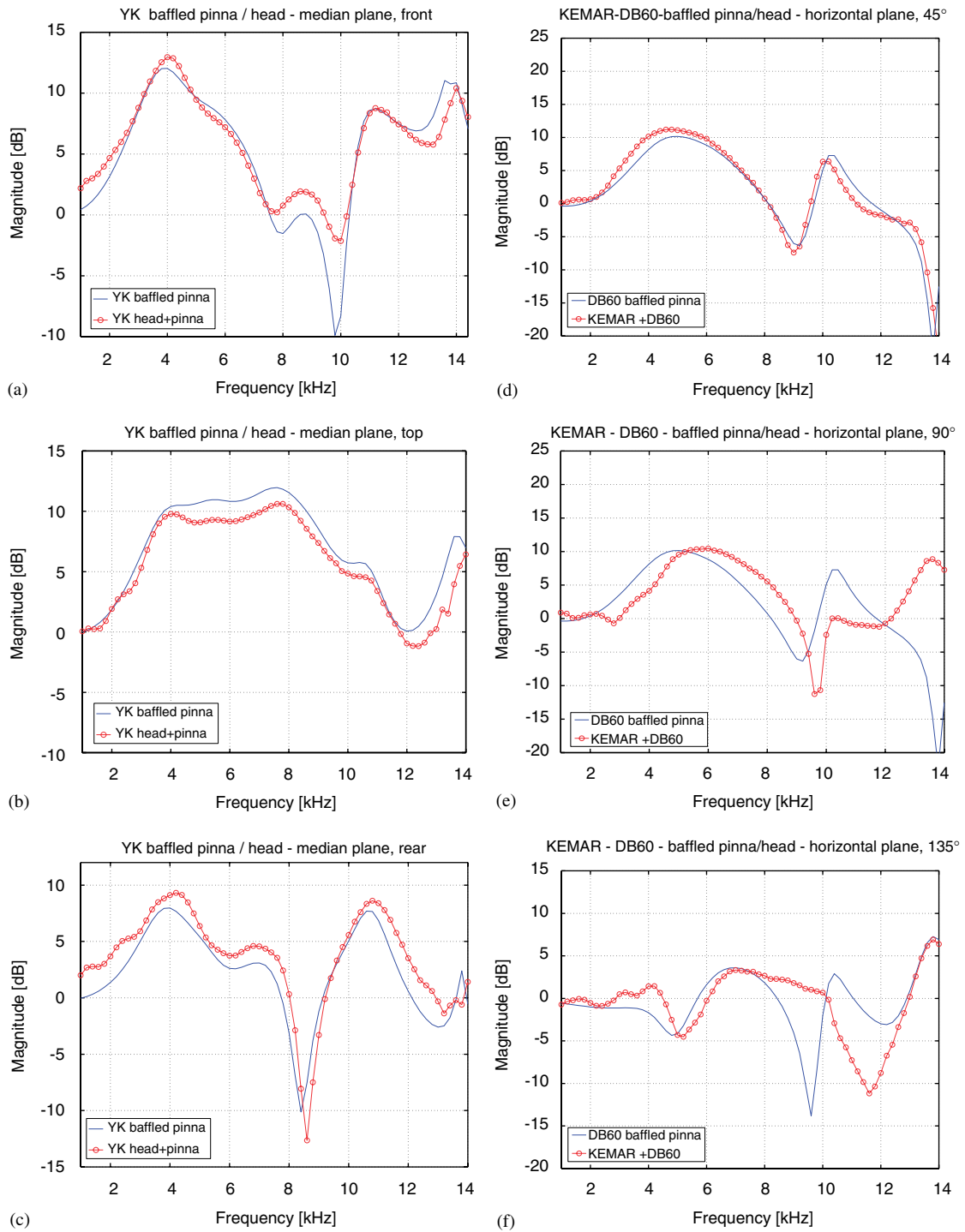
Fig. 14. Comparison of the simulated normalised frequency response of baffled pinnae, and HRTF (head and pinnae): (a) YK median plane: front $\phi = 0°$, $\theta = 0°$, (b) YK median plane: above $\phi = 0°$, $\theta = 90°$, (c) YK median plane: rear $\phi = 0°$, $\theta = 180°$, (d) KEMAR horizontal plane: above $\phi = 45°$, $\theta = 0°$, (e) KEMAR horizontal plane: normal angle $\phi = 90°$, $\theta = 0°$, (f) KEMAR horizontal plane: rear $\phi = 135°$, $\theta = 0°$.

5.3. Results

5.3.1. The response at grazing incidence

A few authors emphasised the significance of the notch that varies between 6 and 10 kHz according to the angle of sound source (Shaw and Teranishi [41], Blauert [1], Bloom [46,47] and Butler and Belendiuk [42]). The results presented in Fig. 15 can be compared with the results obtained by Shaw [40, pp. 34–35], since he investigated the response of baffled pinnae. Only the general trends should be analysed since probably the dimensions of the pinnae are different in these studies, and he used a point source positioned in the near-field (8 cm from the entrance to the ear canal), whereas, in this case a monopole was simulated in the far field using the principle of reciprocity. In his figures, the response was presented in a two-dimensional (2-D) format: magnitude versus frequency. In this form, the trends of the most significance frequencies are highlighted, and also certain features such the 'low-pass' observed with parallel sloping lines between 5 and 11 kHz. However, it is difficult to observe the variation of magnitude with angle, especially at high frequencies.

In Fig. 15 we present, on a linear scale, the response of six pinnae to excitation at grazing incidence. The response is normalised with the response detected at the centre of the baffle, without the pinna. At zero frequency, the amplitude of the response is unity (or 0 dB). In our cases, it is found that pinna amplification can be as high as 4.5 (increase of 13 dB). The figures show the variation of maxima and minima as a function of frequency and angle. We can conclude with the following observations. In a similar way to the case of the median plane angles to the full head (see Section 4.2.1) the notch for all six pinnae starts at around $\theta = -90°$ at 6 kHz and increases up to $\theta = +40°$ at 10 kHz.

In general, the peaks in the figures are arranged in vertical lines, i.e. at certain resonance frequencies, the pinna is excited with different efficiencies depending on the angle of excitation. There are no more than six resonance frequencies, as Shaw found, but in some pinnae only five resonance frequencies were found.

In the work of Shaw, maximum amplification is always observed at the first quarter wavelength (between 4 and 5 kHz). An almost similar amplification level can be found at higher frequencies as well (e.g. subjects A, D, and J, in his figures, Shaw [40, pp. 34–35]). This trend is evident in our case, with the exception that, with some ears, higher amplification at high frequencies is noticed compared with the first resonance. This can be explained by the fact that the source is positioned 5 cm away from the baffle, and amplification values at high frequencies are very sensitive to its position in the vicinity of the baffle. In addition, due to approximation made using the principle of reciprocity where the source is not positioned on the surface but 1–2 mm away, some variation of amplification can be obtained at frequencies above 10 kHz.

The range of frequencies around 4 kHz, corresponding to one of Blauert's [1] 'v'-bands, was found to play an important role in front-back discrimination. This frequency is attenuated for rear sources and boosted for frontal sources due to destructive and constructive interference, respectively, between direct and reflected sound. In all figures the frontal angles ($-30° \leqslant \theta \leqslant +30°$) are characterised by a boost, and angles at the rear (above and below) are characterised by smaller amplification values.

For grazing incidence excitation, a comparison of all resonance frequencies found by Shaw, and this work is given in Table 1. (A detailed study on pinna resonance and mode shapes is given by Kahana [19].) The resonance frequencies are found by identifying frequencies at which a boost at a certain angle is noticed (vertical lines in Fig. 15). The first resonance frequency is very similar in all cases except for a shift in the first resonance of DB60. This increase is expected due to smaller dimensions of this particular pinna (the larger version of KEMAR, the DB65 has a high level of agreement with the averages of Shaw). In two cases not all of the six resonance frequencies were found. The reliability of the last mode is doubtful due to the large sensitivity to source position, mesh geometry and alignment, although in general a peak was always found in this range.

5.3.2. The response in the lateral vertical plane

The main features appearing in Fig. 16 can be summarised as follows: two main peaks occur at different frequencies: at 4–5 kHz with the strongest excitation at normal direction ($\phi = 90°$, $\theta = 0°$), and around 10 kHz at upper angles ($\phi = 90°$, $40° \leqslant \theta \leqslant 80°$). Note the similarity with the resonance frequencies of the cylinder. The behaviour of the YK pinna is slightly different with the second resonance occurring at 8 kHz instead of 10 kHz. A progressive notch is observed from 6 kHz, down at $\phi = 90°$, $\theta = -90°$, and increases with

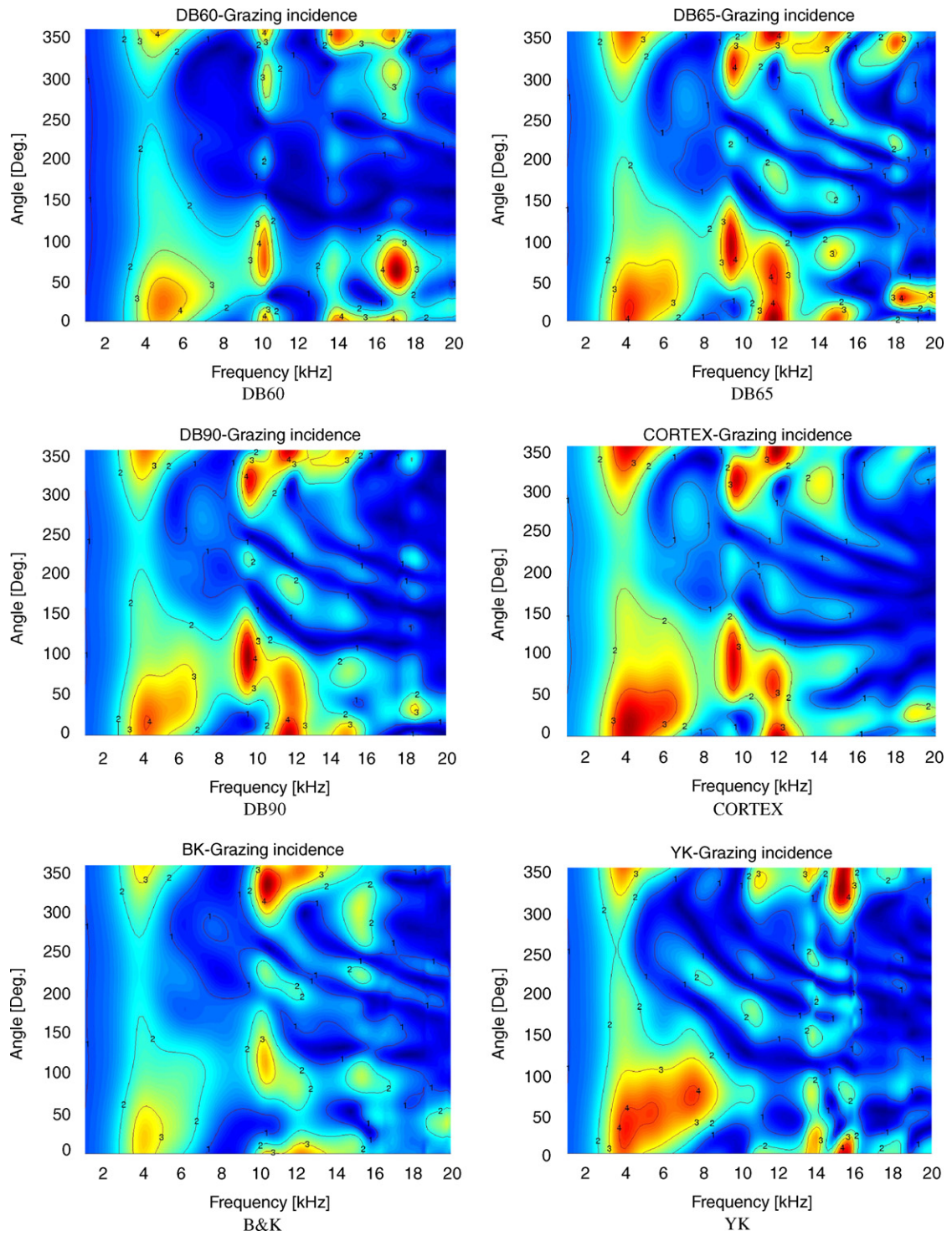Fig. 15. The normalised response of six baffled pinnae in grazing incidence at a resolution of 1° and steps of 200 Hz ($\phi = 0°$, $0° \leqslant \theta \leqslant 360°$).

Table 1
Resonance frequencies in kHz of six pinnae modelled with the BEM

DB60	DB65	DB90	YK	CORTEX	B&K	Shaw
4.9	4.2	4.2	4.1	4.2	4.1	4.2
7.8	7.2	7.2	7.6	7.2	7.7	7.1
10.3	9.5	9.6	—	9.6	10.5	9.6
—	11.6	11.8	11.2	11.8	12.2	12.2
14.0	14.8	14.7	14.0	14.1	15.3	14.4
17.0	18	18.4	17.8	17.4	18.0	16.7

The frequency corresponds to the frequency at which the maximum amplification is reached in the resonance frequency range. The results of Shaw [40] show the average of 10 pinnae.

frequencies up to approximately 12 kHz at approximately $\phi = 90°$, $\theta = +40°$. Variations above 10 kHz have no consistent structure.

5.3.3. The response in the horizontal plane

The most striking feature in the horizontal plane (Fig. 17) is the simpler variation of the first few peaks and notches. The first resonance shifts only slightly from 4 to 6 kHz for angular positions of sources that shift from the front to $\phi = 130°$ at the rear. In all pinnae, from $\phi = 130–180°$ the peak in this frequency range is substituted with a notch. In fact, this notch starts at frontal angles (from $\phi = 20°$ at around 1 kHz) and progress monotonically with frequency up to $\phi = 130°$ at 5 kHz, and then the attenuation is noticed at all angles. The deepest notch at around 9 kHz hardly changes as a function of the angle. This notch is found to be problematic in every HRTF database in the horizontal plane since it requires a large dynamic range in the equalisation process and also since it has a narrow bandwidth, equalisation may result in a boost at a slightly shifted frequency. As before, other spectral features do not repeat systematically among all pinnae above 10 kHz.

5.3.4. The response at grazing incidence of a 'low-resolution' DB60

In the initial stage of the work, the 'low-resolution' scanner was used. At this stage, the full HRTF database (of the head with the pinna) was simulated. Owing to the difficulties in obtaining reliable results at high frequencies, it was not clear to what extent the accuracy of the distorted geometry affects the acoustical response. In order to verify that the resonance frequencies are due to the geometry of the pinna, and not due to numerical errors that might arise with a large model, the original data of the pinna was 'cut' from the head and transformed to a baffled plane. The geometry of this pinna model is depicted in Figs. 7a and b. In this case, the original model comprises 5535 nodes and 10 844 elements. Virtually, the same acoustical characteristics appeared with the baffle compared with the full head for certain angles (see below). The mesh was not decimated to ensure that no additional geometrical distortions would occur, but only smoothed at its boundaries to the plane of the baffle. The response of the 'low-resolution' DB60 for grazing incidence is presented in Fig. 18a. The response of the accurate DB60 pinna for the same plane under similar conditions is presented in Fig. 18b.

It can clearly be seen that the visually distorted parts result in acoustical errors across angles and frequencies. The general trend that appears in both pinnae is two main resonance frequencies with large amplification at frontal angles, whereas rear and bottom angles have very little amplification compared with the front and top angles. However, two significant errors appear:

- Due to the shallow concha all resonance frequencies shift upwards. In this case, 6.5 and 11.5 kHz obtained with the 'low-resolution' pinna, compared with 4.9 and 10.3 kHz obtained with the accurate pinna.
- Although the magnitude of the response is of the same order the overall change of the response with angle and frequency is not as distinctive as with the accurate pinna. This is expected since the variation of the folded parts of the pinna is also not very detailed.
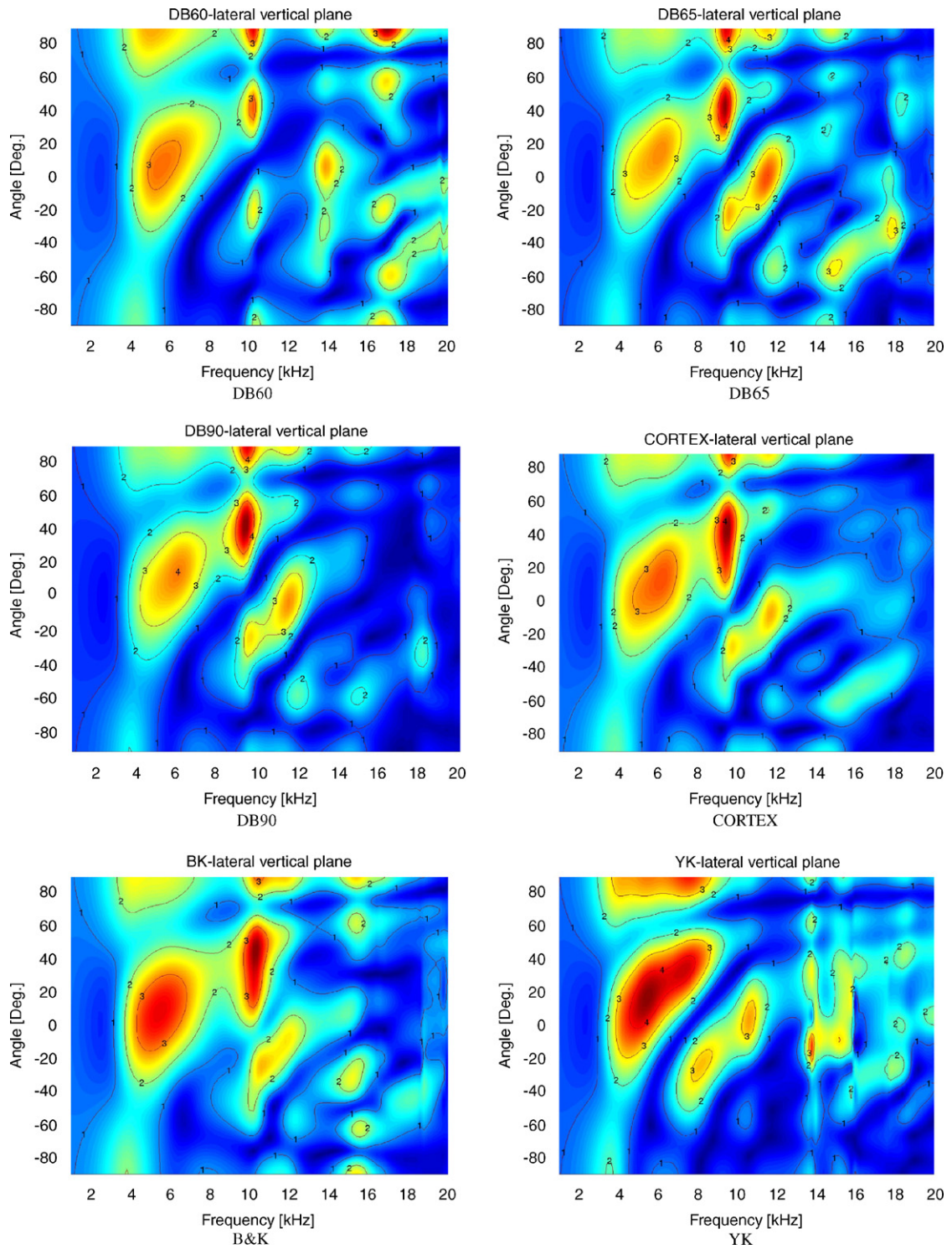
Fig. 16. The normalised response of six baffled pinnae in the lateral vertical plane at a resolution of $1°$ and steps of 200 Hz ($\phi = 90°$, $-90° \leqslant \theta \leqslant +90°$).
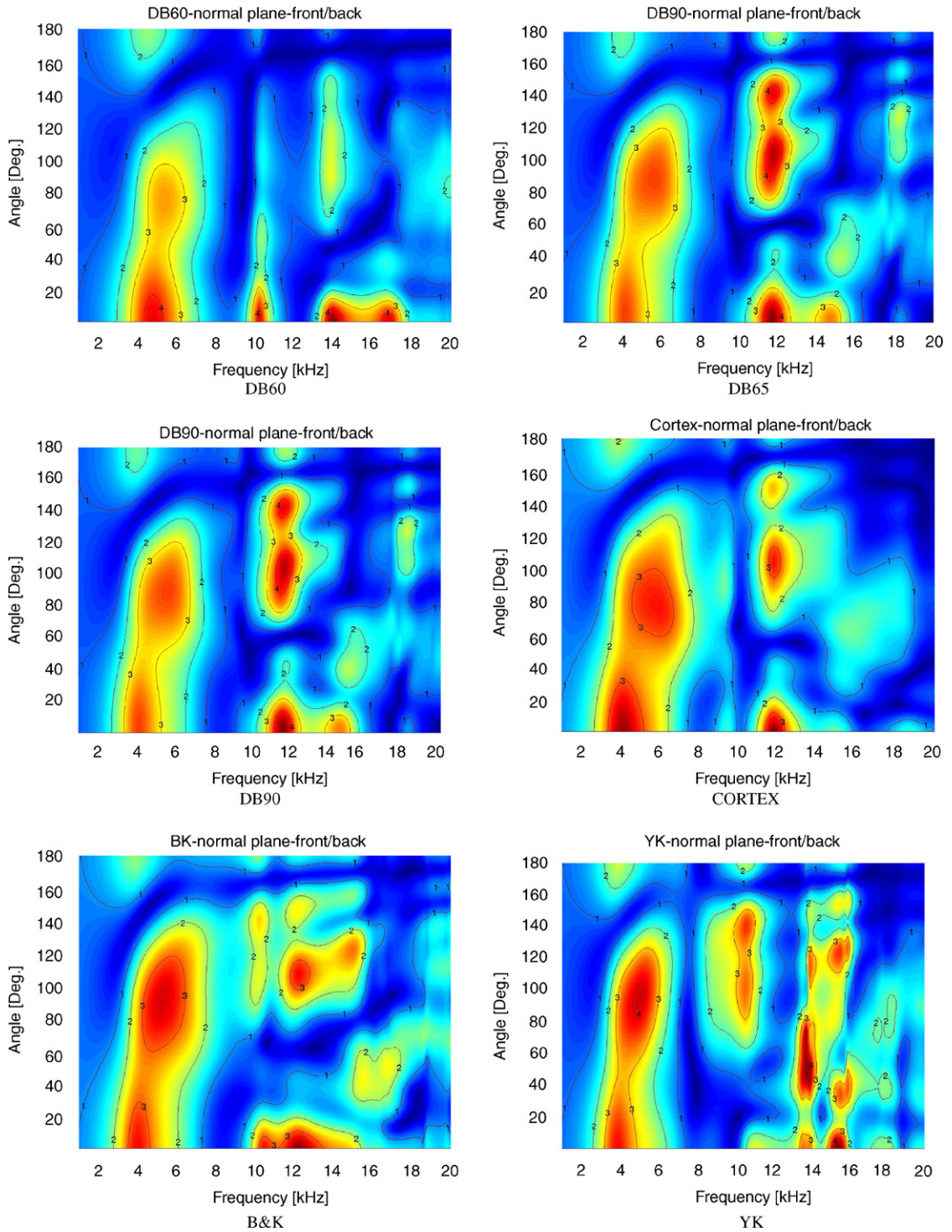
Fig. 17. The normalised response of six baffled pinnae in the horizontal plane at a resolution of $1°$ and steps of 200 Hz ($\theta = 0°$, $0° \leqslant \phi \leqslant 180°$).
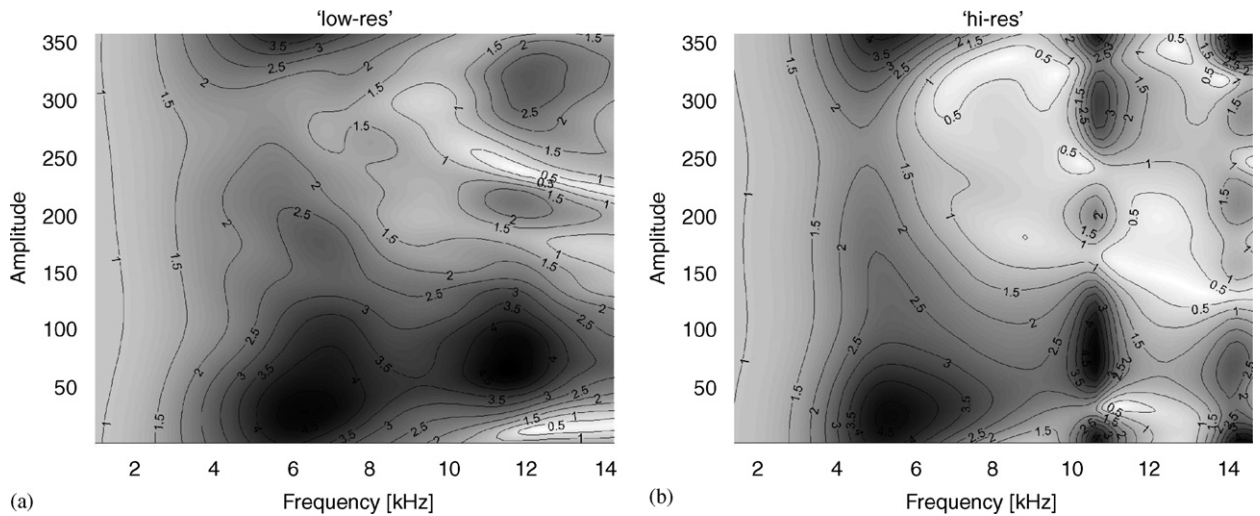
Fig. 18. Comparison of the normalised frequency response of (a) 'low resolution' DB60 and (b) accurate DB60 at grazing incidence. Contours represent amplification, after the response was equalised with the response detected at the centre of the baffle without the pinna.

It is concluded that the 3-D mesh with accurate geometry is essential for modelling the response at high frequencies.

5.4. Examples of modelled pinnae impulse responses

An analysis of measured impulse responses for baffled pinnae could not be found in the literature. The following impulse responses were calculated by applying the inverse fast Fourier transform (IFFT) to frequency responses calculated between 1 and 20 kHz, and therefore the sampling frequency is 40 kHz (as the baffled pinna does not contribute to the response below 1 kHz, the complex pressure was added at lower frequencies due to the contribution of the baffle alone). It should be noted that at the Nyquist frequency the imaginary part of the pressure should be zero. Therefore, a simple linear correction was applied to the complex pressure values at high frequencies (between 15 and 20 kHz, with gradual linear decrease of the phase down to $0°$), without changing the magnitude.

The impulse response at grazing incidence for DB60 and DB65 are given in Fig. 19a and b. These responses are characterised by two adjacent peaks for angles from $\phi = -50°$ up to $+30°$. The same trend can be seen in Hiranaka and Yamasaki [48, Figs. 3 and 4] although they investigated the impulse response with the effect of the head.

For the lateral vertical plane, the DB90 is investigated and the results are presented in Fig. 19c, with a clear trend of a delayed secondary reflection as the source is lowered. Hiranaka and Yamasaki [48] confirmed that major reflections occur within 350 μs after the first arriving sound, and that the delay increases as the source is lowered.

5.5. Baffled pinna with a cylindrical ear canal

In this part of the study, we demonstrate the effect of adding an average cylindrical ear canal using the 'acoustic transparency' formulation of BEM [12]. Møller [49] defined the following variables for sound transmission: p_1 is the pressure at centre position of head, p_2 is the pressure at entrance to blocked ear canal, p_3 is the pressure at entrance to the open ear canal, and p_4 is the pressure at the eardrum. Therefore, the free-field equalised HRTF can be expressed as

$$\frac{p_4}{p_1} = \frac{p_2}{p_1}\frac{p_3}{p_2}\frac{p_4}{p_3}. \tag{2}$$
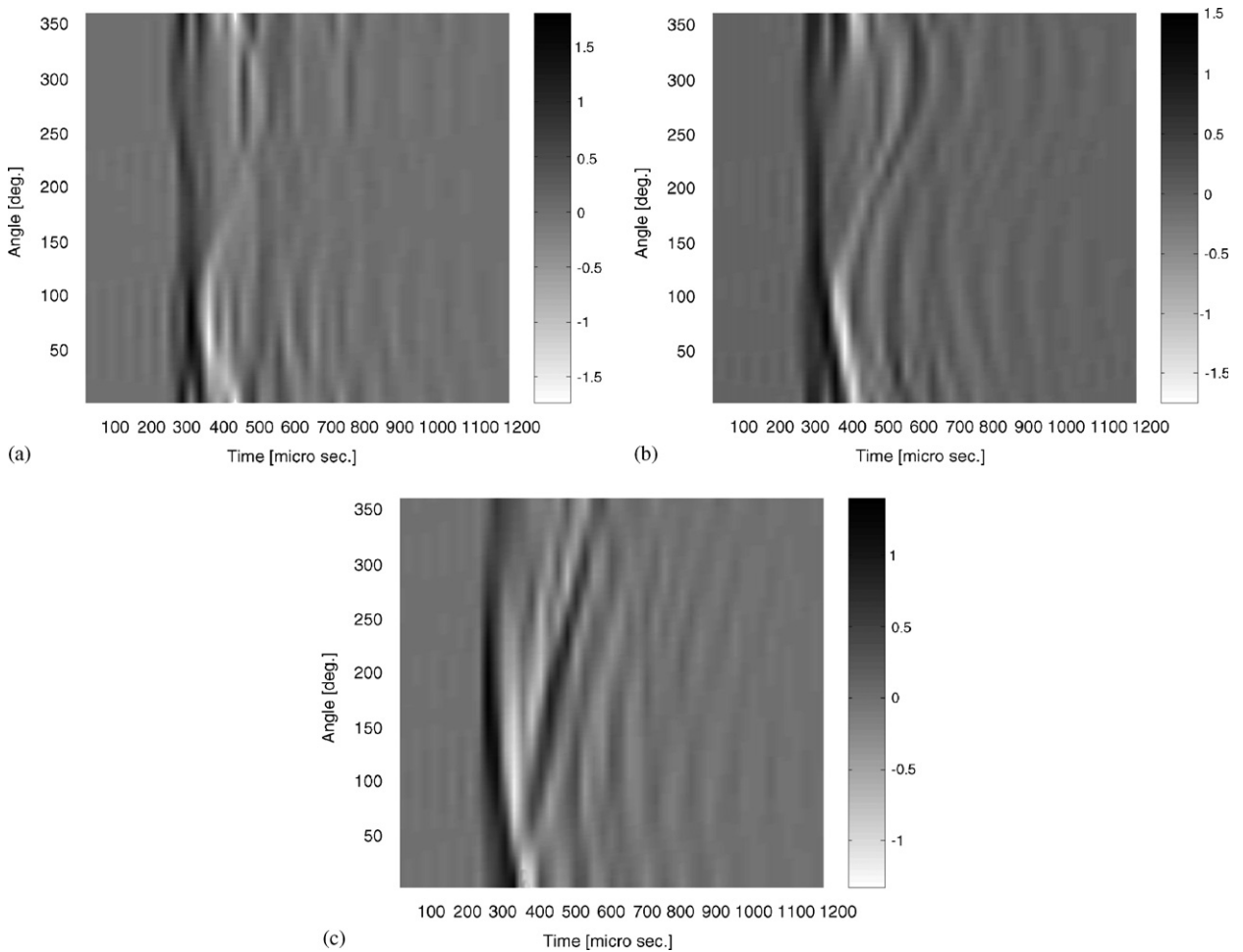
Fig. 19. Simulation of the impulse response of baffled pinnae: (a) DB60—grazing incidence, (b) DB90—grazing incidence, (c) DB90—lateral vertical plane. The Nyquist frequency is 20 kHz. Impulse responses were obtained by applying Inverse FFT to the responses presented in the previous figures, and applying linear phase correction at high frequencies.

It has been shown (Møller [49], Møller et al. [36], and Hammershøi and Møller [15]) that the ratio p_2/p_1 is direction dependant, whereas p_3/p_2 and p_4/p_3 are direction independent. It has also been shown that measurements at the blocked ear produce a lower standard deviation between subjects than measurements undertaken at the open entrance and the eardrum and this point is suitable for binaural recordings. The significance of this approach is the simplification of measuring HRTFs. However, the quality of the sound will be different due to the variation in resonance between blocked and open ear canal (see Ref. [41]).

The motivation of inclusion of a simple cylindrical canal (rather than a real shape) has been raised due to variations among individuals for the canal shape, size and eardrum boundary conditions. Therefore, we include an average ear canal with dimensions based on the Zwislocki coupler (Zwislocki [33]), and the boundary conditions based on average data published by Shaw [35].

5.5.1. Boundary conditions

Using the 'IBEM transparency' formulation, transparency conditions were applied for elements defined on the plane of the entrance to the ear canal, at $z = 0$. This is to ensure that waves can propagate through this 'opening', where the 3-D mesh is attached to an infinite baffle.

Using again the principle of reciprocity, the source was positioned 1 mm above the bottom of the canal (at–21.5 mm). The mesh includes 8203 nodes, 16 118 elements, of which 85 elements are transparent at $z = 0$, and 666 elements comprise the eardrum and are applied with the impedance boundary conditions.

Frequency dependant admittance values have been assigned to the elements of the circular area at the bottom of the cylinder. The values are based on the average impedance values (resistance and reactance) given by Shaw [40].

In the frequency range of up to 8 kHz these values are similar to the response of the Zwislocki ear canal simulator (Zwislocki [33]). The impedance values, measured in acoustic Ohms, were converted to specific impedance by multiplying with the area of the bottom of the canal and converted to specific admittance values in Rayls. It should be noted that this simulation provides only a rough approximation to the response that would have been detected by the Zwislocki simulator DB100. The impedance values were measured by Zwislocki at the four branches (see Ref. [38]), and not at the bottom. However, it is possible in principle to calibrate the response of the canal, by applying admittance values obtained from empirical data.

5.5.2. Results

The results presented in Fig. 20 include simulation with two models: blocked ear canal (see Fig. 5), and the current model with the inclusion of the ear canal. For the latter, two cases were investigated: rigid boundary conditions at the eardrum, and frequency dependant complex admittance values. The general characteristics of the three curves are in agreement with the results of Shaw and Teranishi [41]. A strong resonance appears at 3 kHz due to the canal resonance. Amplification is in the order of 25 dB compared to the response of the baffle. The admittance values reduce the amplification to 15–17 dB. The blocked ear canal creates a peak around 4.5–5.5 kHz depending on the angle of excitation. In grazing incidence the amplification is 10–12 dB (see also Ref. [40]). Note that the sharp minima around 8 kHz is present in both blocked pinnae types as stated by Shaw and Teranishi [41].

It should be noted that this feasibility study requires further investigation of the properties of boundary conditions at high frequencies, and more accurate modelling, for example of the four-branch Zwislocki coupler.

6. Conclusions

In this paper, the feasibility of modelling the HRTF and pinnae response with numerical techniques has been investigated. It has been recognised for some time that in principle, the solution of the wave equation could provide the ultimate solution when modelling the frequency response of the external ear, since this method relies on the *exact* solution for arbitrary, complex bodies. However, only very recently has the combination of improved integral equation techniques, advanced computing hardware, advanced visual capturing devices, and an understanding of the significance of HRTFs in psychoacoustic studies enabled us to carry out such research.

Different formulations of the DBEM and the IBEM have been used in the investigation of accurate geometrical models. It is clear that the efficiency of each numerical technique depending of the type of problem. With the closed models investigated, it was found that better control of accuracy and eliminating the problem of 'irregular frequencies' was achieved with the DBEM. However, the method is inefficient when the models become large.

The method of capturing and manipulating scanned surface models has been described. Since the BEM is very inefficient when large mesh models are investigated, the optimisation of various stages in the simulation is found to be crucial. The principle of reciprocity was used to demonstrate the spatial variation of the pressure at discrete frequencies. The simulation results of individualised HRTF and baffled pinnae were presented in the median, lateral vertical and horizontal planes. The results are compared to previously published observations of spectral features of the HRTF.
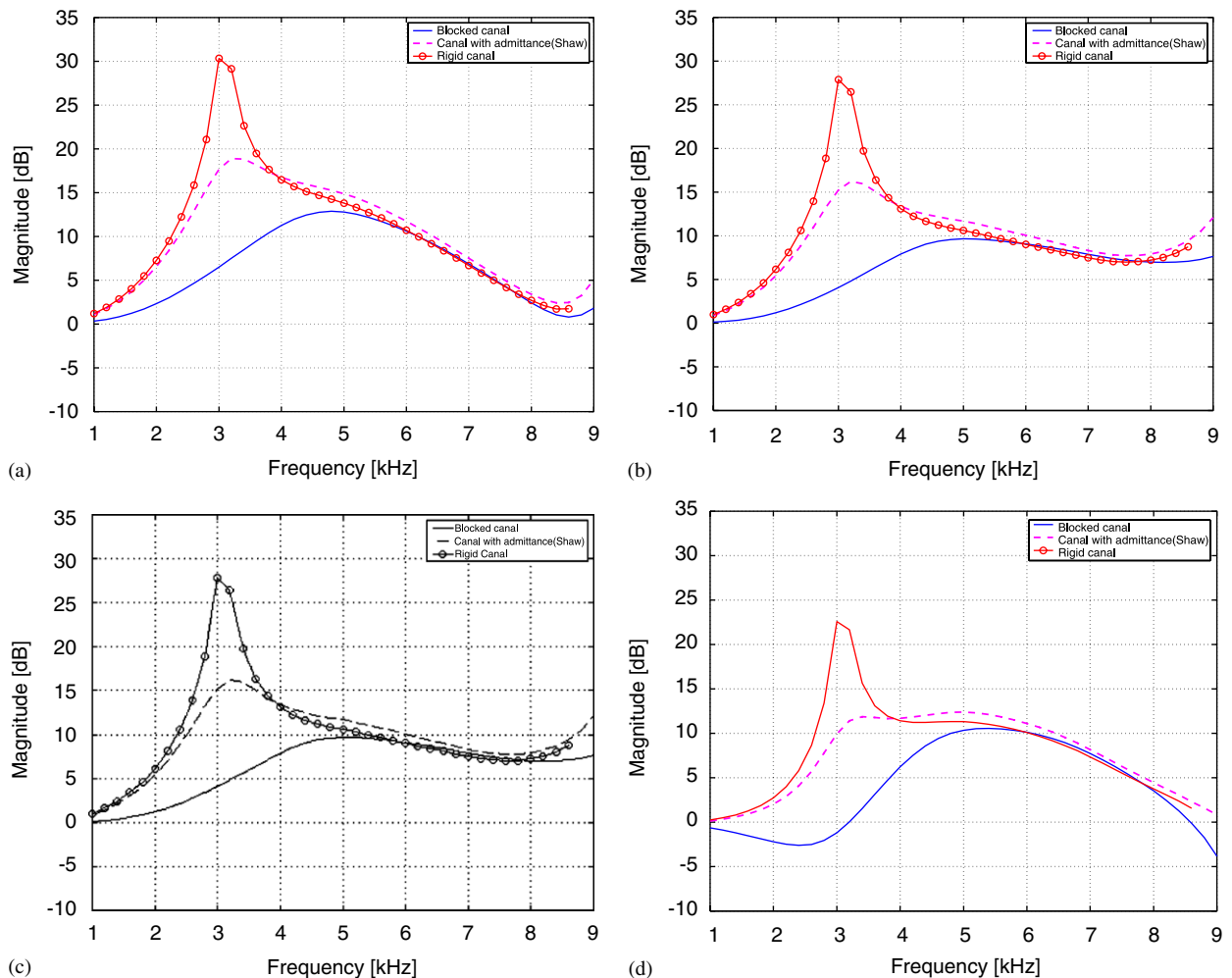
Fig. 20. The response of DB60 in three conditions: Blocked ear canal, with rigid eardrum, and eardrum with averaged impedance boundary conditions by Shaw (1974): (a) front $\phi = 0°$, $\theta = 0°$, (b) above $\phi = 0°$, $\theta = 90°$, (c) normal $\phi = 90°$, $\theta = 0°$ and (d) rear $\phi = 180°$, $\theta = 0°$.

We conclude that it is possible to implement individualised HRTFs in a 3-D sound system or an auditory display, without the need for measurements in an anechoic chamber, if highly accurate 3-D images of the head and pinnae are captured and modelled with BEM.

Acknowledgements

Appendix A. Computational cost

The main hurdle of modelling HRTFs with the BEM is the very high computational cost. At high frequencies, the size of the elements comprising the mesh model needs to be refined. As the dimensions of the model become larger and larger, and the requirements for accurate modelling at high frequency remains fixed, the efficiency of the BEM is dramatically reduced. This is the main reason why the BEM is mainly associated with the 'low' frequency range.

The main difficulties in using large models are summarised as follows:

- *Processing speed*: Although the majority of this work was undertaken on parallel computers, HRTFs can already be modelled on home PCs. The improvement of parallel computers and their interface made it possible to multiply performance depending on the number of processors available. More cost-effective parallel PCs can also be used. We therefore predict that the improvement in computational speed will make simulating HRTFs more accessible in the near future.
- *Physical memory size*: The performance of the BEM is optimised for the in-core solver. A few Gb of RAM are required for each processor (if a parallel computer is used). As before, this limitation is already diminishing.
- *Singularities*: Inherently part of the formulation, singularities are very difficult to remove. This is especially prominent when the cavity is large and many resonance exist (such as in the case of a head with torso).
- *Geometric accuracy*: Rapid improvement in accuracy and cost effectiveness of accurate scanning devices together with advances in computer graphics, make it possible to obtain excellent geometric description of the models. It is possible that even simpler capturing techniques such as digital cameras could be used in the future.
- *Mesh manipulation*: Currently. this is a complex and iterative process. It might be possible to develop an automatic procedure to convert digital images or scanned models directly to the BEM models.
- *Boundary conditions*: Approximate boundary conditions of the eardrum impedance, hair and clothing could be implemented, based on published average data or empirical data. Individualised data of this kind cannot be achieved without acoustical measurements.

The performance of the DBEM and the IBEM with various configurations of hardware is summarised and presented in Table 2 for modelling the response of KEMAR. The bench test results summarise the CPU time and space required to solve our HRTF problems with SGI Origin 2000 and HP exampler parallel computers.

Table 3 summarises the relations between the size of the pinnae mesh models, and the total CPU time.

Table 2
The relation between the size of the BEM head models and the maximum frequency, memory and CPU time requirements (the running time might be different for different platforms, and the maximum frequency assumes six elements per wavelength)

No. of nodes	No. of elements	Max. edge length (mm)	Approx.max. frequency (Hz) (6 e/w)	Memory (Gb)		CPU time—HP exampler (min)		CPU time—SGI origin (min)	
				DBEM	IBEM	DBEM	IBEM	DBEM	IBEM
202000	400000	3.1	18000	640	320	—	—	—	—
27000	50000	6.0	15000	12	6	—	—	—	—
14000	30000	7.2	10000	4	2	2800	4600	—	830
9900	20000	9.0	6400	1.8	0.9	540	900	—	281
5060	10000	11.5	5000	0.4	0.2	60	85	45	46
2600	5000	13.0	4500	0.12	0.06	20	40	18	10

All models are the left-half of the KEMAR head and modelled using the symmetry property in the BEM, and the time presented is the overall time calculation. Fields are blank whenever it is not possible to solve the problem with the in-core solver.

Table 3
The relation between the size of the BEM pinnae models and the maximum frequency, memory and CPU time requirements (the running time might be different for different platforms, and the maximum frequency assumes six elements per wavelength)

Pinna	Number of nodes	Number of elements	Max. edge length (mm)	Max. freq. (Hz)		CPU time—DBEM/SGI (min)
				6 e/w	4 e/w	
DB60	6887	13488	3.64	15560	23340	50
DB65	5199	10216	3.67	15420	23100	24
YK	6361	12523	2.96	19120	28600	41
CORTEX	5923	11632	2.99	18900	28360	34
B&K	5442	10898	3.42	16540	24810	28
DB90	4799	9409	3.12	18130	27200	20
DB95	4728	9267	3.20	17680	26520	19

References

[1] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, Cambridge, MA, 1983.
[2] D.R. Begault, *3-D Sound for Virtual Reality and Multimedia*, AP Professional, Cambridge, MA, 1994.
[3] S. Carlile, *Virtual Auditory Space: Generation and Applications*, Springer, Berlin, 1996.
[4] R.H. Gilkey, T.R. Anderson, *Binaural and Spatial Hearing in Real and Virtual Environments*, Lawrence Erlbaum Associates, Mahwah, NJ, 1997.
[5] H. Møller, M.F. Sørensen, C.B. Jensen, D. Hammershøi, Binaural technique: do we need individual recordings?, *Journal of the Audio Engineering Society* 44 (6) (1996) 451–469.
[6] H. Møller, C.B. Jensen, D. Hammershøi, M.F. Sørensen, *Evaluation of artificial heads in listening tests*, Audio Engineering Society 102nd Convention, pre-print 4404, 1997.
[7] T. Takeuchi, P.A. Nelson, O. Kirkeby, H. Hamada, *Influence of individual head-related transfer function on the performance of virtual acoustic imaging*, Audio Engineering Society, 104th Convention, pre-print 4700, 1998.
[8] S.G. Weinrich, Sound field calculations around the human head, Technical Report, No. 37, The Acoustics Laboratory, Technical university of Denmark, 1984.
[9] S.G. Weinrich, Deformation of a sound field caused by a manikin, *Journal of the Acoustical Society of America* 69 (3) (1981) 796–801.
[10] B.F.G. Katz, Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation, *Journal of the Acoustical Society of America* 110 (5) (2001) 2440–2448.
[11] B.F.G. Katz, Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparisons to real measurements, *Journal of the Acoustical Society of America* 110 (5) (2001) 2449–2455.
[12] LMS Numerical Integration Technologies, *SYSNOISE Theoretical Manual*, 1995.
[13] R.D. Ciscowski, C.A. Brebbia, *Boundary Element Methods in Acoustics*, Computational Mechanics Publications and Elsevier Applied Science, Southampton, 1991.
[14] W. Desmet, *Introduction to Numerical Acoustics, International Seminar on Applied Acoustics (ISAAC)*, University of Leuven, Belgium, 1997.
[15] D. Hammershøi, H. Møller, Sound transmission to and within the human ear-canal, *Journal of the Acoustical Society of America* 100 (1) (1996) 408–427.
[16] Y. Kahana, P.A. Nelson, M. Petyt, S. Choi, *Boundary Element simulation of HRTFs and sound fields produced by virtual acoustic imaging systems*, Audio Engineering Society, 105th convention, preprint 4817, San Francisco, 1998.
[17] Y. Kahana, P.A. Nelson, M. Petyt, S. Choi, Numerical modeling of the transfer functions of a dummy-head and of the external ear, Audio Engineering Society, *Sixteenth International Conference*, Rovaneimi, 1999, pp. 330–345.
[18] Y. Kahana, P.A. Nelson, Spatial acoustic mode shapes of the human pinna, Audio Engineering Society, 109th Convention, Los-Angeles, 2000.
[19] Y. Kahana, P.A. Nelson, Numerical modelling of the spatial acoustic responses of the human pinna, *Journal of Sound and Vibration* 292 (1–2) (2006) 148–178.
[20] Y. Kahana, Numerical Modelling of the Head Related Transfer Function, PhD Thesis, ISVR, University of Southampton, UK, 2000.
[21] Y. Kahana, P.A. Nelson, Experimental validation of boundary element simulations of the transfer function of human heads and baffled pinnae, *Journal of Sound and Vibration*, 2005, submitted for publication.
[22] P.A. Nelson, Y. Kahana, Spherical harmonics, singular value decomposition and the head related transfer function, *Journal of Sound and Vibration* 239 (4) (2001) (special ed.).
[23] J. Foley, A.V. Dam, S. Feiner, J. Hughes, *Computer Graphics: Principles and Practice*, Addison-Wesley, New York, 1990.
[24] P. Heckbert, M. Garland, Survey of Polygonal Surface Simplification Algorithms, Carnegie Mellon University, School of Computer Science, Technical Report (CMU-CS-97-194), 1995.
[25] A.E. Johnson, M. Hebert, Control of Polygonal Mesh Resolution for 3-D Computer Vision, Carnegie Mellon University, School of Computer Science, Technical Report (CMU-RI-TR-96-20), 1997.

[26] Knowles Electronics, KEMAR artificial head based on M.D. Burkhard and R.M. Sachs, Anthropometric Manikin for Acoustic Research, *Journal of the Acoustical Society of America* 58 (1975) 214–222.

[27] H.A. Schenck, Improved integral formulation for acoustic radiation problems, *Journal of the Acoustical Society of America* 44 (1968) 41–58.

[28] R.J. Maxwell, M.D. Burkhard, Larger ear replica for KEMAR manikin, *Journal of the Acoustical Society of America* 65 (1979) 1055–1058.

[29] CORTEX, Binaural recording head, MK1, Neutrik Cortex Instruments (NCI).

[30] IEC 959, Provisional head and torso simulator for acoustic measurements on air conduction hearing aids, International Electrotechnical Commission (IEC), International Standard IEC/TR0 60959 (1990–2005). ED 1.0. ICS code 17.140.50 Electroacoustics, 1990.

[31] Brüel and Kjær, *HATS Artificial Head*, Brüel and Kjær, Nærum, Denmark.

[32] M.R. Stinson, B.W. Lawton, Specification of the geometry of the human ear canal for the prediction of sound–pressure level distribution, *Journal of the Acoustical Society of America* 85 (1989) 2492–2503.

[33] J. Zwislocki, An acoustic coupler for earphone calibration, Special Report LCS-S-7, Syracuse University, New York, 1970.

[34] Z. Yamaguchi, N. Sushi, Real ear response of receivers, *Journal of the Acoustical Society of Japan* 12 (1) (1956) 8–13.

[35] E.A.G. Shaw, Transformation of sound pressure level from the free field to the eardrum in the horizontal plane, *Journal of the Acoustical Society of America* 56 (6) (1974) 1848–1861.

[36] H. Møller, M.F. Sørensen, D. Hammershøi, C.B. Jensen, Head-related transfer function of human subjects, *Journal of the Audio Engineering Society* 43 (5) (1995) 300–321.

[37] M.D. Burkhard, R.M. Sachs, Anthropometric manikin for acoustic research, *Journal of the Acoustical Society of America* 58 (1975) 214–222.

[38] KEMAR, Manikin measurement, *Proceedings of a Conference Organised by M.D. Burkhard*, Industrial Research Products, Inc., A Knowles Company, 1978.

[39] G.F. Kuhn, Model for the interaural time differences in the azimuthal plane, *Journal of the Acoustical Society of America* 62 (1) (1977) 157–167.

[40] E.A.G. Shaw, Acoustical features of the human external ear, in: R.H. Gilkey, T.R. Anderson (Eds.), *Binaural and Spatial Hearing in Real and Virtual Environments*, Lawrence Erlbaum Associates, Mahwah, NJ, 1997.

[41] E.A.G. Shaw, R. Teranishi, Sound pressure generated in an external-ear replica and real human ears by a nearby point source, *Journal of the Acoustical Society of America* 44 (1) (1968) 240–249.

[42] R.A. Butler, K. Belendiuk, Spectral cues utilized in the localization of sound in the median sagittal plane, *Journal of the Acoustical Society of America* 61 (1977) 1264–1269.

[43] S. Mehrgardt, V. Mellert, Transformation characteristics of the external human ear, *Journal of the Acoustical Society of America* 61 (6) (1977) 1567–1576.

[44] J. Hebrank, D. Wright, Spectral cues used in the localisation of sound sources on the median plane, *Journal of the Acoustical Society of America* 56 (6) (1974) 1829–1834.

[45] M. Alexander, L. Laubach, *Anthropometry of the Human Ear*, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, 1968.

[46] P.J. Bloom, Creating source elevation illusions by spectral manipulation, *Journal of the Audio Engineering Society* 25 (1977) 560–565.

[47] P.J. Bloom, Determination of monaural sensitivity changes due to the pinna by use of minimum audible-field measurements in the lateral vertical plane, *Journal of the Acoustical Society of America* 61 (1977) 820–828.

[48] Y. Hiranaka, H. Yamasaki, Envelope representations of pinna Impulse responses relating to three-dimensional localization of sound sources, *Journal of the Acoustical Society of America* 73 (1983) 291–296.

[49] H. Møller, Fundamentals of binaural technology, *Applied Acoustics* 36 (1992) 171–218.