Some observations on sensitivity to HRTF magnitude

Hoffmann, Pablo F.; Møller, Henrik

Published in:
Journal of the Audio Engineering Society

Publication date:
2008

Some Observations on Sensitivity to HRTF Magnitude*

PABLO F. HOFFMANN, *AES Associate Member,* **AND HENRIK MØLLER,** *AES Member*
(pfh@es.aau.dk) (hm@acoustics.aau.dk)

Acoustics, Department of Electronic Systems, Aalborg University, DK-9220, Aalborg Ø, Denmark

The spatial resolution at which head-related transfer functions (HRTFs) are available is an important aspect in the implementation of virtual spatial sound. How close HRTFs must be depends on how much their characteristics differ between adjacent directions and, most important, when these differences become audible. Thresholds for the audibility of differences in the spectral characteristics of HRTFs as a function of angular separation were measured. Listeners had to discriminate between stimuli spectrally shaped with different HRTFs but whose interaural time difference remained the same. Results showed that listeners were more sensitive to changes in the vertical position than to changes in the horizontal position. Results are discussed in connection with requirements for spatial resolution of HRTFs in the synthesis of three-dimensional sound.

0 INTRODUCTION

When sound reaches our ears many reflections from the torso, shoulders, head, and pinnae interact with the direct sound path. This interaction introduces patterns to the magnitude spectrum of the sound commonly characterized by peaks and notches. These patterns vary in a complex way as a function of the sound direction [1]–[3]. The head-related transfer function (HRTF) fully describes the sound transmission from a given direction, and thus HRTFs can be used as filters to synthesize virtual spatial sound. In this context, if the HRTF used to filter the sound is changed, the corresponding change in the output of the filtering is assumed to be perceived as a shift in the apparent location of that sound.

Because different HRTFs give different spectral shapes to the sound at the ear, differences between HRTFs may not only cause an apparent shift in direction but also a change in the perceived timbre. Evidence of this has been reported in a study conducted by Langendijk and Bronkhorst [4], who measured discrimination between interpolated and measured HRTFs. Their results showed that for broad-band stimuli presented with one-third-octave-band levels randomized within ±3 dB on every presentation, a spatial resolution of 10°–15° was required so that interpolated HRTFs generate the same spatial percept as measured HRTFs. It was argued that because of the randomization, changes in timbre were unlikely to be used, and actually, listeners reported that the only reliable cue was source location. If the one-third-octave levels were fixed, the required resolution was 6°, suggesting that for small spectral differences timbre-based cues appear to dominate spatial cues.

The ability of the auditory system to detect variations in the spectral pattern of sound, or profile analysis [5], has been studied in terms of the detectability of changes in the sign of spectral slopes [6], discrimination of broad-band noise shaped with different speech-like spectra [7], detection of peaks and notches from a flat spectrum [8], [9], and detection of a level increment in a single component relative to others in a multicomponent complex stimulus [10]. The motivation of this study is to explore the ability of listeners to discriminate changes in the spectral pattern of stimuli when HRTFs are used to produce these changes. Discrimination is compared in a selected number of spatial positions and for different directions of change. General results show that the sensitivity to HRTF magnitude is dependent on both spatial position and direction of change.

An important aspect of this experiment is that differences in the magnitude spectrum must provide the basis for discrimination, and thus we do not want any differences in phase to be audible. To this purpose HRTFs are implemented as minimum-phase filters. It has been shown that it is perceptually adequate to approximate HRTFs by

*Manuscript received 2007 August 9; revised 2008 August 25 and October 6.

a frequency-independent delay to control the interaural time difference (ITD) and minimum-phase filters to control the magnitude spectra [11].

1 METHOD

1.1 Subjects

Four paid subjects participated in the listening experiment, one female and three males. Their ages ranged from 23 to 28 years. Subjects had normal hearing and they were selected by means of a pure-tone audiometry screening at less than 10 dB HL from 250 to 4000 Hz in octave steps, and less than 15 dB HL for 8 kHz. All subjects had previous experience with listening experiments.

1.2 Stimuli and Apparatus

Broad-band pink noise (20–16 000 Hz) was used as the source signal. The simulation of directional sound was based on HRTFs measured with a resolution of 2° on an artificial head [12]. Eight directions were selected in the left half of the upper hemisphere. These directions are referred to as nominal directions. Directions are given as (lateral angle θ, polar angle ϕ) in a polar coordinate system with horizontal axis and left–right poles. This system is also referred to as the interaural–polar coordinate system. Here the lateral angle ranges from −90° to 90°, which corresponds to the left and right poles, respectively. Polar angles of 0° and 180° correspond to the frontal and rear portions of the horizontal plane, respectively, and 90° to the upper portion of the frontal plane. Fig. 1 shows the eight selected directions. Four directions were selected in the median plane (0° lateral angle; 0°, 44°, 136°, and 180° polar angle). Three directions were chosen on a cone of confusion [(58°, 0°), (46°, 90°), and (54°, 180°)], and they were selected to have the same ITD rather than being on the same geometrical cone, thus the lateral angle varies with the polar angle. The ITD for these directions corresponded to −437.5 μs and was calculated from (46°, 90°). Finally (90°, 0°) was also included, and its corresponding ITD was −625 μs. ITDs were derived from the interaural differences in group delay of the excess-phase components of the HRTFs evaluated at 0 Hz. This procedure has been shown to be adequate for the computation of ITDs [13].

The measured HRTFs—available as pairs of 256-coefficient impulse responses—were truncated using a 72-coefficient rectangular window. At a 48-kHz sampling rate the resulting impulse responses had a duration of 1.5 ms. This duration has been shown to be sufficient to avoid audible effects of the truncation when using noiselike stimuli [14]. To control the HRTFs at low frequencies the dc value of each impulse response was set to unity gain as described by Hammershøi and Møller [15, sec. 5.2]. Minimum-phase representations of the dc-corrected impulse responses were constructed using homomorphic filtering [16, ch. 12].

Stimuli were played back using a PC equipped with a professional audio card RME DIGI96/8 PST. The digital output of the audio card was connected to a 16-bit D/A converter (Big DAADi) set at a sampling rate of 48 kHz. From the D/A converter the signal went to a stereo amplifier (Pioneer A-616) modified to have a calibrated gain of 0 dB. To reduce the noise floor a custom-made 20-dB passive attenuator was connected to the output of the amplifier. The stereo output signal from the attenuator was delivered to the listener over equalized Beyerdynamic DT-990 circumaural headphones.

Two 256-coefficient minimum-phase FIR filters were employed in order to compensate for the left and right headphone transfer functions. The equalization filters were based on measurements made at the entrance to the blocked ear canal on 23 subjects (none of whom participated in this listening test). Five measurements were obtained from each ear and subject, and subjects were asked to reposition the headphones between measurements. Headphone responses were obtained using the maximum-length-sequence (MLS) technique [17], and the results

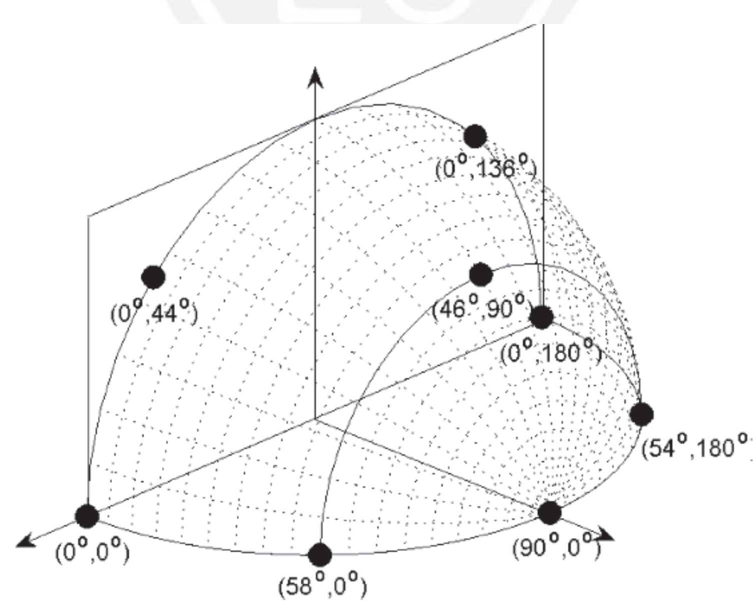

Fig. 1. Nominal directions selected in upper hemisphere, left half. Directions are specified in an interaural–polar coordinate system.

were in the form of 256-coefficient impulse responses for each ear. Impulse responses were then transformed to the frequency domain, and a representative transfer function was calculated by taking the mean of all transfer functions on a power basis. The equalization filter was designed based on the inverse of this mean response. To avoid excessive amplification at low frequencies due to the inversion, the dc value of the inverse was adjusted manually to unity gain. This value corresponded roughly to the observed gain at low frequencies. To reduce the amplification that also occurs at, and close to, the Nyquist frequency, the calculated inverse response was replaced by a low-pass filter for frequencies above 16 kHz. Fig. 2 shows the subjects' responses, the mean response, and the response of the equalization filter for the left ear. In order to obtain a time representation of the equalization filter, a minimum-phase approximation was computed for each ear using homomorphic filtering [16, ch. 12].

1.3 Psychophysical Procedure

Each nominal direction, with the exception of (0°, 0°), had an associated set of neighboring HRTFs spaced symmetrically about the nominal direction such that the absolute angular span was 2°, 4°, 16°, 24°, and 32°. For the forward direction the selected angles were 2°, 4°, 8°, 12°, 20°. Measured HRTFs were not available for the angular span of 2°, and therefore, these HRTFs were obtained by linear interpolation between the minimum-phase impulse responses of the closest measured directions. For any nominal direction there were two modes of angular changes in the HRTFs, namely, changes in lateral angle and changes in polar angle. In the remainder of this paper these modes of changes are referred to as directional modes. Fig. 3 shows a graphical representation of these directional modes when the selected nominal direction is (0°, 0°). For changes in lateral angle an arc is described along the hori-

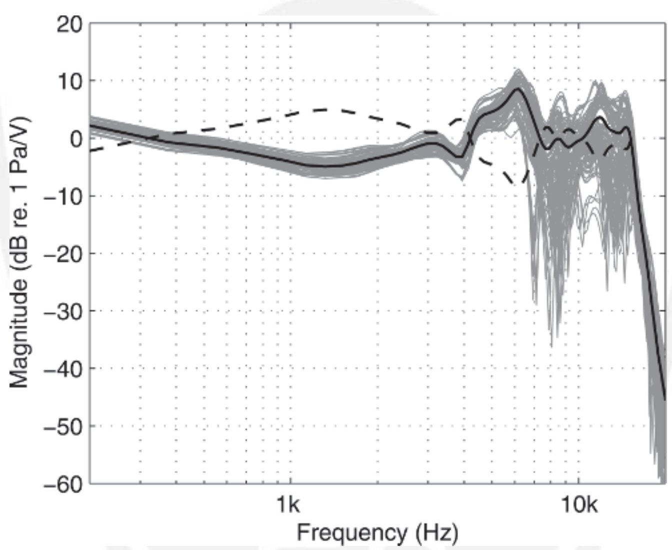

Fig. 2. HRTFs measured on left ears of 23 subjects (gray lines). —— mean response; – – – response of equalization filter, corresponding to inverse of mean response.

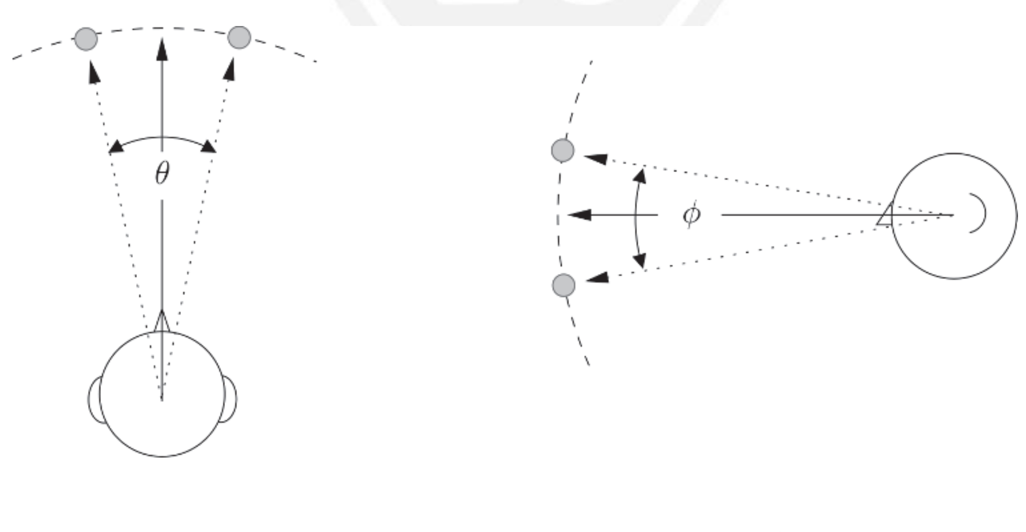

Fig. 3. Schemes of two directional modes used in experiment. Solid arrows—nominal direction [here (0°, 0°)]; dotted arrows—pairs of adjacent HRTFs describing arcs for changes in: (a) lateral angle with angular separation θ; (b) polar angle with angular separation ϕ.

zontal plane (angle θ). For changes in polar angle HRTFs along the median plane are used to describe the arc (angle ø). Note that the midpoints of the arcs always correspond to the nominal direction. For the particular case of (90°, 0°), where in the strict sense changes in polar angle cannot be applied, changes along the frontal plane spanning the angle vertically are used instead of changes in polar angle.

Discrimination of HRTF magnitude was estimated in a three-interval, two-alternative forced-choice (3I 2AFC) task using the method of constant stimuli. Stimulus and interstimulus intervals were of 300-ms duration. The stimulus presented on either the second or third interval differed from the other two. Subjects were asked to identify the interval that differed (that is, the target stimulus). They had to push one of two buttons to indicate a response. A feedback light was used to immediately show the subjects whether or not the response was correct. A 2-s silence interval was used between trials.

On a single trial a given nominal direction, directional mode, and angular separation were selected. The minimum-phase HRTF for the direction corresponding to one end of the arc (recall Fig. 3) was used to filter two of the three sound intervals (one corresponding to the first interval). This filter was regarded as the reference HRTF. The remaining sound interval was filtered with the minimum-phase HRTF for the direction corresponding to the other end of the arc, that is, the target HRTF. Note that the ITD of the nominal direction was used for both the reference and the target HRTFs. For changes in lateral angle, the spatial configuration of the pair reference–target HRTF was selected randomly to be either left–right or right–left from the nominal direction. Similarly, for changes in polar angle the pair reference–target HRTF could be either above–below or below–above the nominal direction.

1.4 Experimental Design

Subjects were in a sound-insulated cabin specially designed for psychoacoustical experiments. For familiarization and practice, blocks of sixteen trials were used, and only the largest angular separation was presented. As with the main experiment, feedback was provided to the subjects. Each subject completed three to four practice blocks. Since subjects had recently participated in similar listening tests and their performance was observed to be stable, no further practice was necessary.

In the main experiment each combination of nominal direction, directional mode, and angular separation was presented 30 times. Thus a total of 2400 responses were obtained per subject (8 nominal directions times 2 directional modes times 5 angular separations times 30 repetitions). In a block of trials all angular separations were repeated 15 times, and the order in which they were presented was random. The nominal direction and the directional mode were held constant within a block. Thus two blocks were necessary to collect all data for one nominal direction and one directional mode. Data were collected during three sessions that were held on different days. Blocks were distributed so that one session lasted from about one hour and a half to two hours.

1.5 Psychometric Functions

Psychometric functions, as used in this work, describe the total percentage of correct responses as a function of angular separation. The proportion of correct responses p obtained at each angular separation were used to estimate psychometric functions for each subject and each condition. We assumed a logistic form of the psychometric function. Its mathematical expression is given by

$$\hat{p} = \lambda + (1 - \lambda)(1 + e^{-(x-\alpha)/\beta})^{-1} \tag{1}$$

where $\hat{p}$ is the estimate of p, x is the angular separation, α is the threshold parameter, and β is the slope parameter (shallow slopes correspond to large values of β). The parameter λ corresponds to chance performance, and here is fixed to 0.5 (the reciprocal of the number of alternatives per trial). The threshold is defined as the angular separation that yields 75% of correct responses or, in other words, the value of α for which $\hat{p}$ equals 0.75. The slope parameter β provides information about the rate at which performance improves with increasing angular separation.

Psychometric functions were fitted using a least-square criterion based on the iterative Gauss–Newton algorithm. The fitting was performed on the log of the angular separation.

2 RESULTS

2.1 Analysis of Individual Results

Psychometric functions were obtained for the audibility of differences between minimum-phase HRTFs as a function of their angular separation. Fitted psychometric functions and the proportions of correct responses for each condition and subject are shown in Fig. 4. The abscissa specifies the angular separation in degrees, and the ordinate specifies the proportion of correct responses. Each panel shows proportions and fitted functions for each nominal direction. Results were generally consistent across subjects and performance improved with increasing angular separation.

Using the bootstrapping technique we computed confidence limits for the estimated parameters of each psychometric function. (We adopted this technique from the analysis done by Lutfi et al. [18] on psychometric functions for informational masking.) Fig. 5 shows the estimated confidence limits. The bootstrapping technique used to estimate confidence limits is described by Maloney [19]. From the psychometric function fitted to the empirical data we calculated the percent correct for each angular separation, and assuming they are binomially distributed, a simulated percent correct was randomly drawn for each angular separation, and a new fitting was performed on the simulated percent correct. This operation was repeated 10 000 times to provide 10 000 estimates of the threshold and slope parameters. Then the 2.5% and 97.5% quantiles were taken as the 95% confidence limits. The purpose of these calculations was to give the reader an idea of the uncertainty of the estimated parameters. We can observe from Fig. 5 that threshold and most slope estimates are reasonably good. Note that confidence limits for slope increase with increasing β. This may be due to the inverse relation between β and the slope of the psychometric function, which suggests that reliable estimates of β are difficult to obtain for shallow slopes.

2.2 Analysis of Mean Results

In order to observe group tendencies, the threshold parameter α and the slope parameter β were averaged across subjects. Geometric means were calculated for the threshold parameter and arithmetic means for the slope. Fig. 6 shows the calculated mean psychometric functions, and the mean parameters are summarized in Table 1.

A two-factor within-subject analysis of variance with nominal direction and directional mode as factors was conducted on the logs of the thresholds. The analysis revealed a highly significant main effect of nominal direction [$F(7, 21) = 12.9$, $p < 0.001$] and a significant main effect of directional mode [$F(1, 3) = 25.1$, $p < 0.05$]. That is, the thresholds for changes in polar angle were consistently lower than those for changes in lateral angle. The interaction between nominal direction and directional mode was also significant [$F(7, 21) = 3.2$, $p < 0.05$]. This can be attributed to the fact that for some directions the thresholds for changes in polar angle were slightly lower than for changes in lateral angle, whereas for other directions the difference was larger. A similar analysis on the logs of the slope revealed a significant main effect of nominal direction [$F(7, 21) = 4.7$, $p < 0.01$]. A post-hoc analysis (Tukey HSD) revealed that (46°, 90°) was sig-

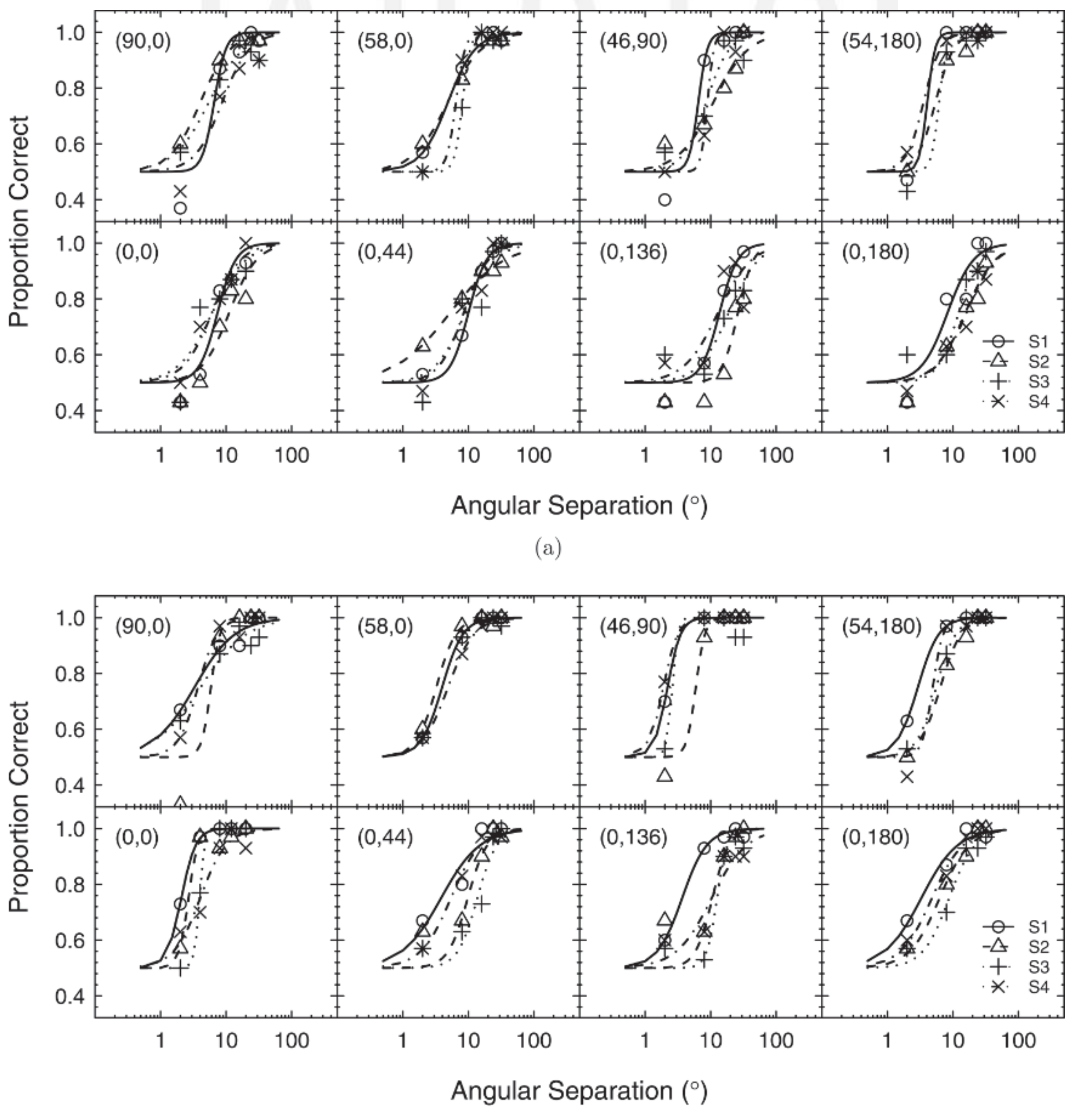

Fig. 4. Psychometric functions for discrimination of HRTF magnitude. (a) Results for each nominal direction and changes in lateral angle. (b) Results for changes in polar angle. Individual proportions are represented by different symbols and fitted psychometric functions by different lines.

nificantly different from (0°, 44°) and (0°, 180°), reflecting the very steep slope observed for changes in the polar angle around (46°, 90°) as compared to the others.

3 DISCUSSION

The audibility of spectral differences in HRTFs was estimated by measuring how well subjects could discriminate between changes in the minimum-phase HRTFs while the ITD remained constant. For the directions used in this study, the mean results were in the range of 2.8°–17.2°, depending on location and direction of change. This implies that different resolutions are required for minimum-phase HRTF filters according to the position of the virtual sources. In general, as the elevation increases in the median plane a lower resolution is required. In the horizontal plane and for lateral higher resolutions are required in more lateral positions. These results are similar in pattern to those reported by Minnaar et al. [20], who examined the required spatial resolution so that the error introduced by linear interpolation of minimum-phase HRTFs was inaudible.

An interesting agreement with the results from Minnaar et al. [20] is that for the directions (0°, 0°) and (0°, 180°) results showed a higher sensitivity to changes along the

Table 1. Estimated mean thresholds (α), in degrees, and slope (β) parameters across subjects for changes in lateral angle and polar angle for each nominal direction.

Nominal Direction	Lateral Angle α	Lateral Angle β	Polar Angle α	Polar Angle β
(90°, 0°)	6.0	0.20	4.1	0.22
(58°, 0°)	5.9	0.15	4.0	0.17
(46°, 90°)	8.6	0.14	2.8	0.08
(54°, 180°)	4.7	0.09	4.8	0.14
(0°, 0°)	7.4	0.21	3.1	0.11
(0°, 44°)	7.9	0.25	7.2	0.21
(0°, 136°)	17.2	0.19	8.4	0.17
(0°, 180°)	12.6	0.21	5.4	0.24

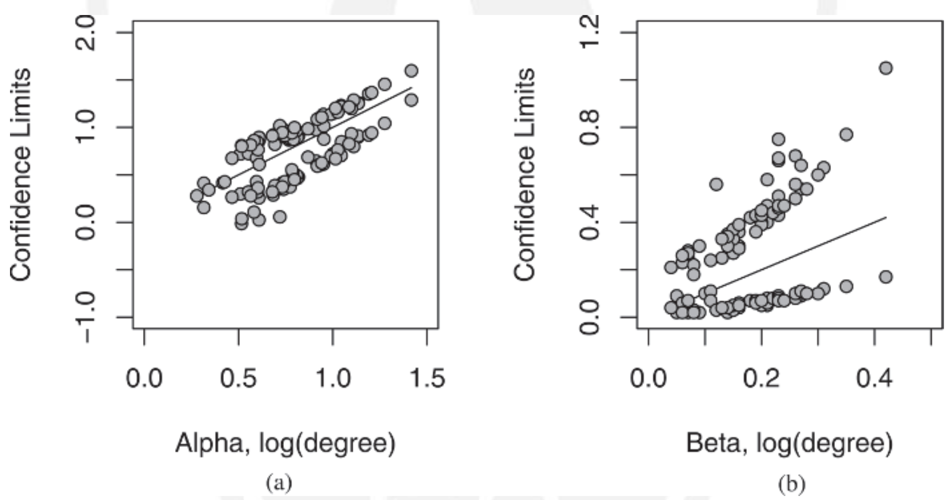

Fig. 5. Scatterplots of estimated confidence limits obtained by bootstrapping technique (see text for details). Lines—estimated parameters; pairs of points above and below lines—upper and lower limits for each listener and condition.

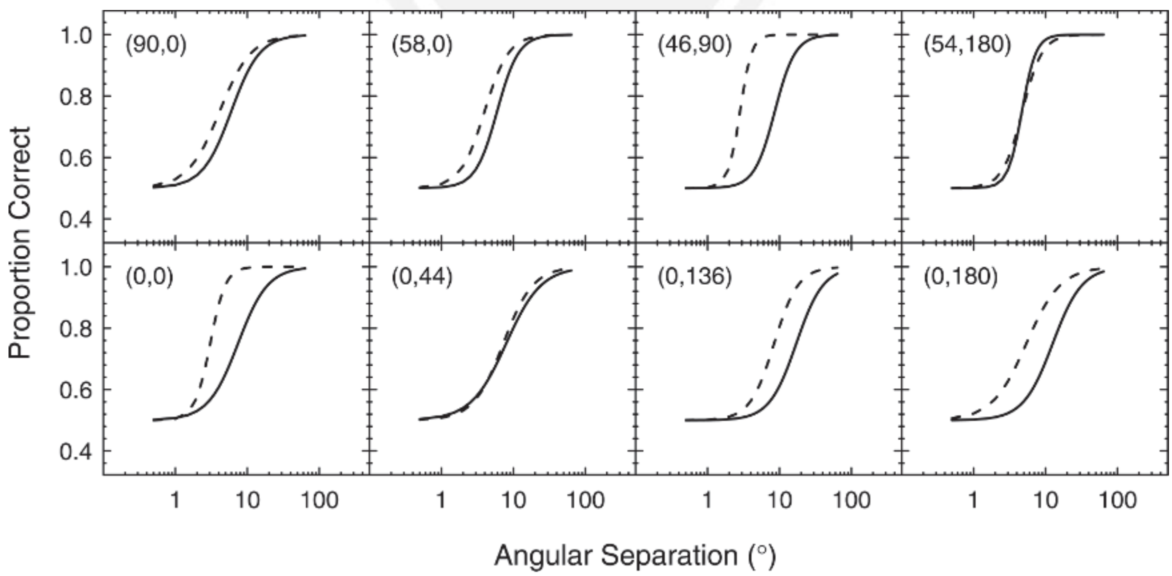

Fig. 6. Mean psychometric functions for discrimination of HRTF magnitude. —— changes in lateral angle; – – – changes in polar angle.

median plane than along the horizontal plane. This suggests that the rate of spectral change is higher for changes in polar angle than for changes in lateral angle. This is a striking observation considering that sound localization blur at these locations is smaller for horizontal displacement than for vertical displacement [21]. A possible reason for this discrepancy is that in both studies subjects were instructed to use any available cue for discrimination, and thus they may have effectively used cues that are not necessarily those related to localization. Because changing between HRTFs will change the tonal character of the stimuli it is conceivable that for small angular separations, a change in timbre may have been more dominant than a perceived shift in direction. Evidence of this has been shown in [4].

Though Minnaar et al. did not report specific thresholds, they suggested resolutions ranging from 4° to 24°, which is a larger range than that of our thresholds. A possible explanation for this difference is that Minnaar and colleagues measured the angular separation between minimum-phase HRTFs so that the interpolated HRTF, computed from the minimum-phase HRTFs (interpolators), was indistinguishable from a target empirical HRTF. The empirical target HRTF always corresponded to the direction in the center of the angle spanned by the interpolators. The fact that the interpolated HRTF cannot be distinguished from the empirical target HRTF does not imply that the difference between the interpolators is inaudible. It is the audibility of this difference that we have actually investigated in this study.

Results also indicate that the selection of representative locations for HRTF measurements need not be uniform. Some spatial regions need a more dense set of measuring locations than others. In addition, the fact that thresholds for changes in polar angle were consistently lower than thresholds for changes in lateral angle implies that synthesis of moving sound, in which HRTFs need to be constantly updated, requires higher spatial resolution for trajectories that incorporate changes in polar angle than for those where lateral-angle changes occur, that is, trajectories where ITDs also need to be updated.

To assess the possibility of specific frequency regions being more dominant as cues for discrimination, we computed spectral differences in one-third-octave bands between the HRTFs corresponding to the locations separated by the estimated thresholds. Fig. 7 shows these differences plotted for each nominal direction. They are given in absolute values in dB, meaning that higher values reflect larger differences for that frequency band. Fig. 7(a) shows differences at thresholds for changes in lateral angle, and Fig. 7(b) for changes in polar angle. It can be observed that for most of the directions off midline, differences in the contralateral component of the HRTF (gray lines) are larger than those in the ipsilateral component.

It is clear that larger differences occur at high frequencies, and this is somewhat expected considering that the contributions of spectral cues to sound localization are more prominent at high frequencies [3]. The fact that HRTFs differ in complex ways makes it difficult to confirm a specific criterion for discrimination. It seems more likely that discrimination could have been based on the most reliable cue available, whether this is a prominent difference in a particular narrow frequency region, or the integration of spectral differences over a wide frequency range. In this respect it is also important to note that for sufficiently large lateral changes, interaural level differences (ILDs) could have been used as the discrimination cues. It is, however, not clear at which point small spectral differences might have become large enough to be considered as ILDs. Though we acknowledge the importance of ILDs as localization cues [21], we made no attempt to differentiate between ILDs and small spectral differences because our goal was to measure audibility thresholds based on any available cue.

It might be argued that the audibility thresholds obtained in this study may differ from thresholds that would

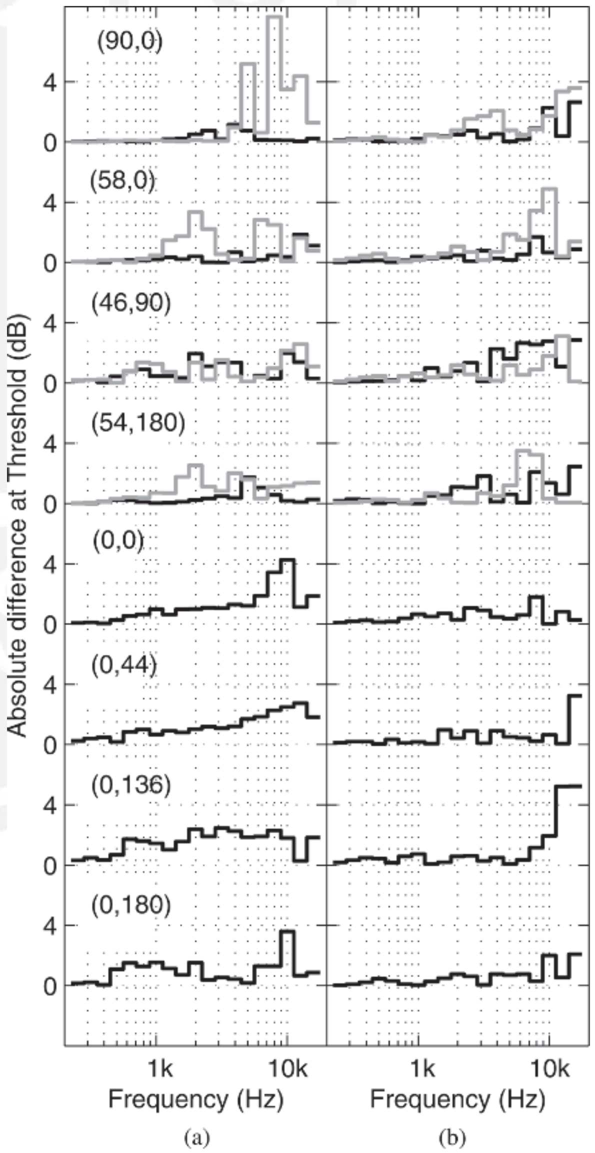

Fig. 7. Spectral differences at threshold in third-octave bands from 250 to 16 000 Hz. (a) Differences for changes in lateral angle. (b) Differences for changes in polar angle. Gray lines—differences between right-ear HRTFs; black lines—differences between left-ear HRTFs.

be obtained using individualized HRTFs. To measure just audible angular differences it is critical to have high spatial resolution in order to make sure that sensitivities are due to real differences in HRTFs and not to mismatches caused by interpolation. It is virtually impossible to measure HRTFs on human subjects with the same spatial resolution of the generic HRTFs used here. Involuntary movements can easily exceed the desired resolution between neighboring directions (2°), and thus corrupt the measurements. Although some devices may be used to fix the subject and subject's head, there are mainly two problems with this approach. First it is very difficult to get the correct posture of the subject. Second the size of the fixing devices can easily affect the sound field, especially for those frequencies whose wavelength is comparable to the dimensions of the devices. In addition, extra care is required to control for unwanted reflections in the HRTF impulse responses.

4 SPECTRAL DISTANCE MEASURE

In an attempt to represent the amount of spectral difference by a single number, we computed the standard deviation across frequency of the difference (in dB) as a function of angular separation. This metric has been employed in modeling the contribution of spectral cues to localization [3], and in minimizing intersubject differences in HRTFs [22]. One of the advantages of this metric is that overall level differences (differences that are constant across frequency) are eliminated, and thus only variations in spectral shape are emphasized.

Before computing the spectral differences the HRTFs were smoothed using a gammatone filter bank [23]. The motivation for using this smoothing technique is that these filters simulate the frequency analysis performed by the cochlea. Furthermore since the frequency resolution of the cochlea is poorer at high frequencies, this procedure effectively smooths out some of the spectral details that may wrongly inflate the estimation of spectral differences (such as spectral notches). For anechoic HRTFs it has been shown that by selecting an appropriate filter order for the gammatone filters (third or fourth), the smoothing results in imperceptible differences as compared to the original ones [24]. Here the order of the gammatone filters was set to four, and the smoothing procedure was based on the procedure used by Breebaart and Kohlrausch [24]. The mathematical equations used to derive the gammatone filters and to calculate the smoothed HRTFs are given in the Appendix.

Spectral differences were computed on HRTF differences produced by angular separations from 4° to 20° in steps of 4°. Recall that the resolution of the empirical HRTFs is 2°, and because in this experiment HRTFs were spaced symmetrically about nominal directions, then the smallest possible angular separation between measured HRTFs was 4°. To form a difference spectrum the difference between dB amplitudes was computed component by component in a frequency range of 1125–12 000 Hz in 187.5-Hz steps. Then the standard deviation across the components of the difference spectrum was computed.

Fig. 8 shows computed spectral differences for the directions in the median plane and separations spanned along the polar angle. Note that spectral differences increase with increasing angular separation, a pattern that is also observed for the other nominal directions. Furthermore note that spectral differences increase more rapidly for nominal directions whose estimated thresholds were lower. By doing a linear fitting to the data, we can use the slope of the fitted line to establish a relation between the rate at which spectral differences increase and the observed thresholds. A simple model that describes this relation can be expressed mathematically by

$$a = \frac{c}{b} \tag{2}$$

where a indicates a threshold estimate, b is the slope of the fitted curve, and c is a constant given in dB. To find a value for c we first computed four threshold estimates as the inverse of b for the same conditions presented in Fig. 8. By minimizing the sum of the square errors between the measured thresholds and the estimates we found a value of 0.4 for c. In this calculation only one ear was used because left and right HRTFs were identical in the median plane.

Fig. 9 shows thresholds and the approximations obtained from Eq. (2). The abscissa represents the polar angle of the nominal directions, which in turn are grouped by ITDs. In Fig. 9(a) thresholds for changes in polar angle are plotted together with the approximations for the left-ear HRTF and right-ear HRTF. The approximation derived from the left-ear HRTF is in good agreement with the thresholds obtained, with the exception of the nominal direction (54°, 180°). The approximation derived from the right ear, which corresponds to the contralateral component, is less accurate. In Fig. 9(b) we observe an almost identical pattern for changes in lateral angle. Here a good approximation is observed in the median plane. Approxi-

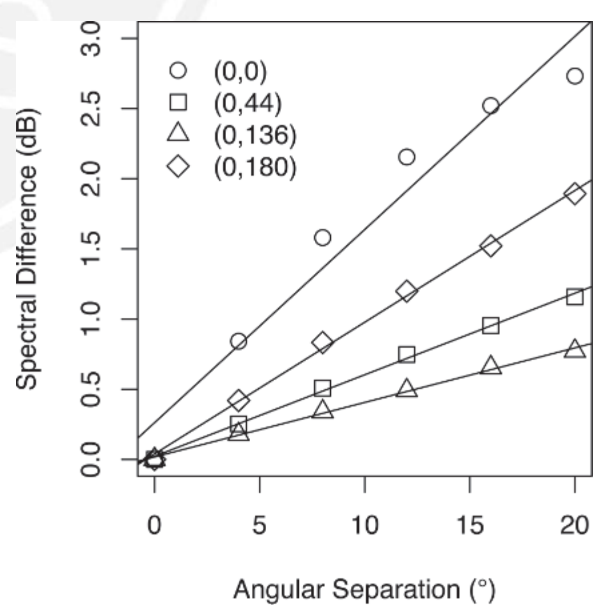

Fig. 8. Spectral differences between HRTFs separated symmetrically about nominal direction for directions in median plane (indicated by symbols). Angular separation is along polar angle, lines represent linear fittings for each nominal direction.

mations from the left ear are also good for (58°, 0°) and (90°, 0°). Right-ear approximations are generally less accurate, with the exception of (46°, 90°). These observations suggest that for the more lateral positions discrimination was mostly based on differences in the ipsilateral component. It is also possible that overall ILDs might have been used for discrimination. These changes in overall level are not included in this simple model. However, in terms of predicting the audibility of spectral differences along the median plane, the results of this analysis are encouraging.

5 CONCLUSIONS

Listeners were able to discriminate spectral differences in HRTFs for angular separations in a range of 2.8°–17.2° depending on direction. This relatively large range may be attributed to the fact that HRTFs change differently depending on the spatial location. Thresholds for changes in polar angle were consistently lower than for changes in lateral angle. A simple model for the discrimination of spectral differences was proposed based on the standard deviation across frequency of a difference spectrum. It was possible to account for thresholds measured for locations in the median plane. For lateral directions approximations were less accurate, probably because some other cues not included in the model, such as overall interaural level differences, may have been used. An important aspect of this study, which deserves further investigation, concerns the necessary limitation of using generic HRTFs. From our results it is not possible to know what would have been the sensitivity to HRTF changes had the subjects been presented with individualized HRTFs.

6 ACKNOWLEDGMENT

This work was supported by the Danish Technical Research Council and the Research Council for Technology and Production Science.

7 REFERENCES

[1] D. Wright, J. H. Hebrank, and B. Wilson, "Pinna Reflections as Cues for Localization," *J. Acoust. Soc. Am.,* vol. 56, pp. 957–962 (1974 Sept.).

[2] B. Rakerd, W. M. Hartmann, and T. L. McCaskey, "Identification and Localization of Sound Sources in the Median Plane," *J. Acoust. Soc. Am.,* vol. 106, pp. 2812–2820 (1999 Nov.).

[3] E. H. A. Langendijk and A. W. Bronkhorst, "Contribution of Spectral Cues to Human Sound Localization," *J. Acoust. Soc. Am.,* vol. 112, pp. 1583–1596 (2002 Oct.).

[4] E. H. A. Langendijk and A. W. Bronkhorst, "Fidelity of Three-Dimensional-Sound Reproduction Using a Virtual Auditory Display," *J. Acoust. Soc. Am.,* vol. 107, pp. 528–537 (2000 Jan.).

[5] D. M. Green, *Profile Analysis: Auditory Intensity Discrimination* (Oxford University Press, New York, 1988).

[6] N. J. Versfeld, "Discrimination of Changes in the Spectral Shape of Noise Bands," *J. Acoust. Soc. Am.,* vol. 102, pp. 2264–2275 (1997 Oct.).

[7] C. L. Farrar, C. M. Reed, Y. Ito, N. I. Durlach, L. A. Delhome, P. M. Zurek, and L. D. Braida, "Spectral-Shape Discrimination—I. Results from Normal-Hearing Listeners for Stationary Broadband Noises," *J. Acoust. Soc. Am.,* vol. 81, pp. 1085–1092 (1986 Apr.).

[8] B. C. J. Moore, S. R. Oldfield, and G. J. Dooley, "Detection and Discrimination of Spectral Peaks and Notches at 1 and 8 kHz," *J. Acoust. Soc. Am.,* vol. 85, pp. 820–836 (1989 Feb.).

[9] A. Alves-Pinto and E. A. Lopez-Poveda, "Detection of High-Frequency Spectral Notches as a Function of Level," *J. Acoust. Soc. Am.,* vol. 118, pp. 2458–2469 (2005 Oct.).

[10] D. M. Green, Z. A. Onsan, and T. G. Forrest, "Frequency Effects in Profile Analysis and Detecting Complex Spectral Changes," *J. Acoust. Soc. Am.,* vol. 81, pp. 692–699 (1987 Mar.).

[11] A. Kulkarni, S. K. Isabelle, and H. S. Colburn, "On

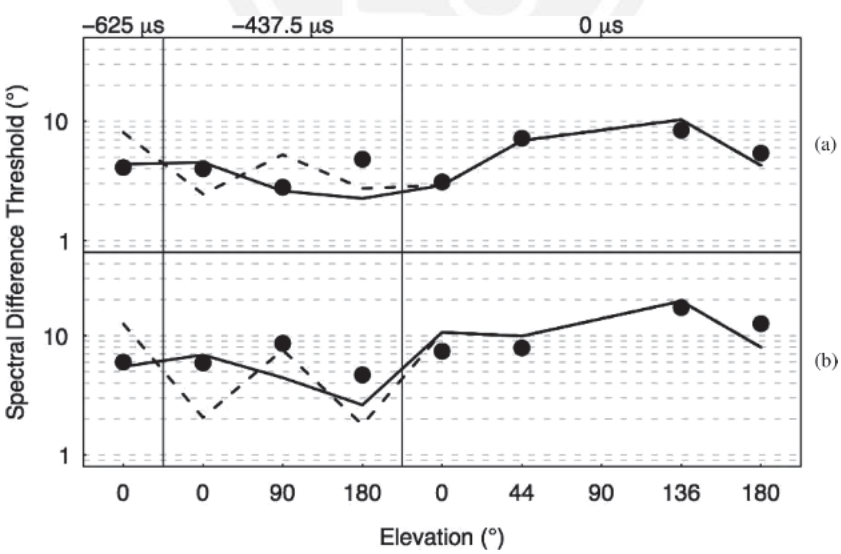

Fig. 9. Comparison between measured thresholds (circles) and predictions based on Eq. (2). (a) Changes in polar angle. (b) Changes in lateral angle. —— approximations using spectral differences for left-ear HRTFs (ipsilateral); – – – approximations for right-ear HRTFs (contralateral).

the Minimum-Phase Approximation of Head-Related Transfer Functions," in *Proc. ASSP (IEEE) Workshop on Applications of Signal Processing to Audio and Acoustics* (New Paltz, NY, 1995 Oct. 15–18), pp. 84–87.

[12] B. P. Bovbjerg, F. Christensen, P. Minnaar, and X. Chen, "Measuring the Head-Related Transfer Functions of an Artificial Head with a High Directional Resolution," presented at the 109th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 48, p. 1115 (2000 Nov.), preprint 5264.

[13] P. Minnaar, J. Plogsties, S. K. Olesen, F. Christensen, and H. Møller, "The Interaural Time Difference in Binaural Synthesis," presented at the 108th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 48, p. 359 (2000 Apr.), preprint 5133.

[14] J. Sandvad and D. Hammershøi, "What Is the Most Efficient Way of Representing HTF Filters?," in *Proc. Nordic Signal Processing Symp.* (*NORSIG '94*), (Lesund, Norway, 1994 June 2–4), pp. 174–178.

[15] D. Hammershøi and H. Møller, "Binaural Technique, Basic Methods for Recording, Synthesis and Reproduction," in *Communication Acoustics,* J. Blauert, Ed. (Springer, Berlin, Germany, 2005), pp. 223–254.

[16] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1989).

[17] D. D. Rife and J. Vanderkooy, "Transfer-Function Measurement with Maximum-Length Sequences," *J. Audio Eng. Soc.,* vol. 37, pp. 419–444 (1989 June).

[18] R. A. Lutfi, D. J. Kistler, M. R. Callahan, and F. L. Wightman, "Psychometric Functions for Informational Masking," *J. Acoust. Soc. Am.,* vol. 114, pp. 3273–3282 (2003 Dec.).

[19] L. T. Maloney, "Confidence Intervals for the Parameters of Psychometric Functions," *Percept Psychophys.,* vol. 47, pp. 127–134 (1990).

[20] P. Minnaar, J. Plogsties, and F. Christensen, "Directional Resolution of Head-Related Transfer Functions Required in Binaural Synthesis," *J. Audio Eng. Soc.,* vol. 53, pp. 919–929 (2005 Oct.).

[21] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA, 1999).

[22] J. C. Middlebrooks, "Individual Differences in External-Ear Transfer Functions Reduced by Scaling in Frequency," *J. Acoust. Soc. Am.,* vol. 106, pp. 1480–1492 (1999 Sept.).

[23] R. D. Patterson, M. H. Allerhand, and C. Giguère, "Time-Domain Modeling of Peripheral Auditory Processing: A Modular Architecture and a Software Platform," *J. Acoust. Soc. Am.,* vol. 98, pp. 1890–1894 (1995 Oct.).

[24] J. Breebaart and A. Kohlrausch, "The Perceptual (Ir)relevance of HRTF Magnitude and Phase Spectra," presented at the 110th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 49, p. 546 (2001 June), convention paper 5406.

[25] B. R. Glasberg and B. C. J. Moore, "Derivation of Auditory Filter Shapes from Notched-Noise Data," *Hear. Research,* vol. 47, pp. 103–138 (1990 Aug.).

APPENDIX

Formulas for the smoothing of HRTF spectra with gammatone filters

Following the procedure described by Breebaart and Kohlrausch [24], the smoothed magnitude $|Y(f_c)|$ of HRTF $X(f)$ was computed as

$$|Y(f_c)| = \sqrt{\frac{\int_0^\infty |X(f)|^2 |H(f,f_c)|^2 \, df}{\int_0^\infty |H(f,f_c)|^2 \, df}}$$

where $H(f, f_c)$ corresponds to the frequency response of the gammatone filter with center frequency f_c. This transfer function is given by

$$H(f,f_c) = \left[\frac{1}{1 + j(f - f_c)/b}\right]^n$$

where n is the filter order and b is the 3-dB bandwidth, which was set to the equivalent rectangular bandwidth (ERB) estimate of the human auditory system as derived in [25]. Its expression is given by

$$b(f_c) = \frac{24.7(0.00437\, f_c + 1)}{2\sqrt{2^{1/n} - 1}} .$$

THE AUTHORS

P. F. Hoffmann

H. Møller

Pablo Faundez Hoffmann was born in Chile in 1976. He received an acoustic engineering degree from Universidad Austral de Chile in 2000. Subsequently he went to Denmark to continue his studies at Aalborg University, where he received an M.Sc. degree in electrical engineering in 2003 and a Ph.D. degree in three-dimensional sound in 2008.

He is currently working as a postdoctoral fellow at Aalborg University, conducting research on technological and psychoacoustical aspects of sound in the use of multimodal interfaces for training applications. His main research interests are spatial sound perception, digital signal processing, and multimodal integration for virtual reality purposes.

Dr. Hoffmann is a member of the Audio Engineering Society and IEEE.

●

Henrik Møller was born in Århus, Denmark, in 1951. He studied electrical engineering and received a B.Sc. degree from the Danish Engineering Academy in 1974. He received a Ph.D. degree from Aalborg University, Denmark, in 1984.

From 1974 to 1976 he worked as a development engineer for Brüel & Kjær. Since then he has been working at Aalborg University. He became associate professor in 1980 and was appointed reader (docent) in 1988 and professor in 1996. From 1991 to 1994 he was partly on leave from the University to work as director for Perceptive Acoustics A/S, a research subsidiary company of Brüel & Kjær.

Dr. Møller's previous and current research reflect his long-time experience with sound, its influence on humans, acoustical measurement techniques, signal processing, hearing, and psychometric methods. His research areas include effects of infrasound and low-frequency noise on humans, investigations of hearing thresholds and loudness assessment, and exploitation of binaural techniques. He was responsible for the design and control of the work for the high-quality acoustical laboratories built at Aalborg University in 1987. Since then the activities of this section have increased and he is now manager of research and education in a wide range of areas such as human sound perception, audiology, psychometry, electroacoustics, recording and playback techniques, auralization in acoustic room modeling and virtual reality, acoustical measurement techniques, electronics, and signal processing.

He has organized conferences on Low Frequency Noise and Hearing (Aalborg 1980), general acoustics (Nordic Acoustical Meeting, Aalborg 1986), and Low Frequency Noise and Vibration (Aalborg 2000). He has also been deeply involved with standardization (ISO Technical Committee 43: *Acoustics*, and several of its working groups as member/convener). He is a member of the editorial board of the *Journal of Low Frequency Noise & Vibration* and serves as reviewer for a number of international journals and national and international research foundations. He is the author of numerous scientific publications and has written invited papers and conference papers. He holds several patents. Recent business relationships comprise the establishment of AM3D A/S and membership of its Business Board.

Dr. Møller is a Knight of the Royal Danish Order of Dannebrog. He is a member of the Danish Engineering Society, Audio Engineering Society, Acoustical Society of America, IEEE, Danish Acoustical Society, Danish Technical-Audiological Society, Danish Society for Applied Signal Processing, and Danish Standards (Board of Acoustics and Working Group of Audiometry and Hearing).