

Sound Field Analysis along a Circle and Its Applications to HRTF Interpolation*

THIBAUT AJDLER, *AES Student Member*, CHRISTOF FALLER, *AES Member*, LUCIANO SBAIZ, AND
 (ajdler@gmail.com) (christof.faller@epfl.ch) (luciano.sbaiz@epfl.ch)

MARTIN VETTERLI[†]
 (martin.vetterli@epfl.ch)

Audiovisual Communications Laboratory, EPFL, Lausanne, Switzerland

The sampling and interpolation of a sound field in two and three dimensions along a circle is discussed. The Fourier domain representation of the sound field is used, and an angular sampling theorem is developed for the sampling of the sound field along a circle. Based on these results, HRTF sampling and interpolation are discussed. This method achieves very precise interpolation in terms of mean square error. However, these results are only possible if very finely spaced HRTF measurements are available. A method is proposed to improve interpolation results when the HRTF measurements are more coarsely spaced than dictated by the angular Nyquist theorem. The proposed method interpolates the HRTFs in a subband domain. In subbands where small angular aliasing occurs, the previous method is applied. In the other subbands, interpolation is carried out in a complex temporal envelope domain to avoid aliasing. Simulations with models and measured data show that the proposed algorithm performs significantly better than previous methods in a mean square error sense.

0 INTRODUCTION

In this paper we explore the spatial evolution of the sound field along a circle. We study how many sensors need to be placed on a circular array to be able to recover the field at any position along the circle. Based on the results presented previously in [1], we present a novel angular sampling theorem of the sound field along a circle. This theorem describes quantitatively the tradeoff existing between the quality of reconstruction of the sound field and the angular spacing between sensors. Further, as an application of the theory presented, the sampling and interpolation of head-related transfer functions (HRTFs) is described. Results presented previously in [1] are first recalled. These results describe a technique to interpolate HRTFs precisely if finely spaced HRTF measurements are available. For example, at a sampling rate of 44.1 kHz,

HRTFs need to be available every 4.9° to avoid spatial aliasing during their interpolation. This interpolation method enables a very precise interpolation in a least mean square sense (much more precise than other methods), but its drawback is that it loses its advantage when applied to more coarsely spaced HRTFs than would be dictated by the angular Nyquist theorem. The theorem indeed often dictates very dense angular sampling, which makes the method difficult to apply in some practical situations.

To counteract the limitations of this first interpolation scheme, we are proposing here an HRTF interpolation algorithm that attempts to take advantage of spatial bandwidth considerations while improving its performance for the case when fewer HRTF measurements are present than required by the angular Nyquist theorem. In the low-frequency band, where very little spatial aliasing occurs, the proposed algorithm applies an angular interpolation method similar to the method proposed previously. At higher frequencies the interpolation is carried out in subbands. Angular aliasing is avoided by applying angular interpolation to the complex temporal envelope in each subband. Given the interpolated complex envelopes, the subband signals are regenerated by modulating them with a separately interpolated carrier signal. This technique is

*Manuscript received 2006 September 25; revised 2007 August 30 and 2008 January 6.

[†]Also with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA.

then compared with different existing techniques. It is shown that the new method still achieves good interpolation results in cases of angular undersampling by a factor of 2 to 4.

0.1 Related Work

Circular microphone or loudspeaker arrays, discussed in detail in this paper, are used in different techniques existing in the literature such as beamforming [2]–[4], wave field analysis and synthesis [5], [6], and HRTF interpolation [7]. Previously published work related to sound field sampling in general, sound field sampling on a circle, sampled sound field reconstruction, and HRTF interpolation is briefly reviewed in the following.

0.1.1 Sound Field Sampling

Related works exist on the topic of sound field sampling, but they have mostly considered the plane wave assumption [8]–[10]. In [6] Poletti has quantified the aliasing, noise, and transducer variability of circular microphone arrays under the far-field assumption. Note that there exists a major difference between plane waves studied along a line and on a circle. The linear array is the natural way of analyzing plane waves since one plane wave at one temporal frequency corresponds exactly to one point in the spectrum [11], [12]. In other words, one plane wave corresponds to one basis function of the Fourier decomposition. As known from the literature [3], when considering a plane wave along a line, the sound field can be shown to have a compact support in the frequency domain, and if the Nyquist sampling condition is fulfilled, perfect reconstruction can be achieved. When the same plane wave is studied along a circle, perfect reconstruction is no longer possible. It can be shown that the decomposition of a plane wave in cylindrical coordinates excites all possible modes [8]. Therefore, theoretically, sensors should be placed infinitely close to each other to be able to reconstruct the sound field perfectly. In practice, by neglecting the modes containing little energy, a very good reconstruction can still be achieved. Jones et al. [9] have derived bounds on the error made by reconstructing the field when only a limited number of those modes are considered.

0.1.2 Wave Field Synthesis

Substantial work has been done in the field of wave field synthesis (WFS) using circular arrays, and many existing systems make use of circular loudspeaker arrays [5]. Recently Spors et al. have investigated aliasing artifacts present when reproducing the sound field with WFS using linear and circular arrays of loudspeakers [13]. Also, wave field extrapolation algorithms can be adapted to describe the case of circular microphone arrays [4], [14], [15]. In [14] auralization of the sound field is described. The field is measured along a circular array and further extrapolated at other positions. Different possible geometries are described and compared for this purpose, and it is concluded that a circular array of microphones leads to the best extrapolation results.

0.1.3 HRTF Modeling and Interpolation

When an HRTF data set is measured, not all possible positions for the source can be considered, and angular sampling needs to be performed. A major question arises then about what this angular sampling should be and what should the method be to interpolate the HRTFs between the sampling positions considered. In this paper we answer these two questions. In the present section we present some existing techniques for HRTF modeling and interpolation.

One of the most simple and straightforward methods applies linear interpolation using the nearest neighbor measured HRTFs to obtain interpolated HRTFs at any angle in between the measurements [16].

A model for HRTFs has been proposed by Kulkarny et al. [17], which decomposes HRTFs as a minimum-phase and an all-pass function. This model is based on previous studies [18] showing that HRTFs can be approximated by a minimum-phase function up to 10 kHz. Minaar et al. [19] interpolated the minimum-phase components of HRTFs linearly in the time domain and claim to be able to predict the needed resolution of measurements without causing audible artifacts. Since the minimum-phase components have a minimum-phase lag, phase delay, and group delay for a given magnitude, they are optimally aligned in time.

This idea of alignment in time has been used in other papers to improve the quality of interpolation. It has indeed been shown that the performance of interpolation in the time or frequency domain can be improved by compensating the HRTFs prior to interpolation according to the time of arrival of sound [20], [21]. That is, the HRTFs are time aligned and interpolation is carried out on the time-aligned HRTFs. In addition the time of arrival is interpolated separately. In [20] linear, spline, and sinc interpolation is considered, and the best results are obtained using linear interpolation and time alignment of the HRTFs. Hartung et al. [22] have compared other methods, such as the inverse distance weighting method and the spherical splines method [23], [24]. As a result of psychoacoustic tests, they found out that the latter method performed slightly better.

Other methods used for HRTF modeling are their representation using a limited number of basis functions with corresponding weighting factors depending on the azimuthal and elevation angles. The basis functions can, for example, be obtained by principal-component analysis [18], independent-component analysis [25], or spatial feature extraction and regularization [26]. Obviously, the more basis functions are used, the better the modeling results are. Using measured HRTFs at known positions, the corresponding weighting factors are calculated. Interpolation of these weighting factors leads to the newly interpolated HRTFs.

Infinite impulse responses have also been considered in [27], [28] to model HRTFs. Two main models have been considered, the all-pole model and the pole-zero model. In [29] linear interpolation of the poles and zeros using a gradient search algorithm has been studied.

A very original and novel technique was described in [30], where HRTFs are obtained as the solutions of a scattering problem [31]. The solution obtained is expanded using multipole expansion [32]. Using this representation, regularized fitting of measured HRTFs is applied, and from there the parameters in the expansion can be found to describe the precise set of HRTFs. Then interpolation and range extrapolation can be achieved.

Recently, making use of a head tracker, an interesting approach to capturing, recording, and rendering spatial sound fields has been developed in [33]. There a few HRTF interpolation techniques making use of the Rayleigh duplex theory [7] are described and compared. These perceptually motivated techniques usually interpolate differently the high- and low-frequency contents of HRTFs.

0.2 Contributions

The contributions presented in this paper are the following:

- A thorough analysis of the sampling of the two-dimensional sound field is described together with an angular Nyquist theorem for circular arrays.
- Approximation formulas are given for the essential support of the spectrum of the three-dimensional sound field gathered on a circle.
- Best interpolation results for HRTFs in the horizontal plane are obtained when sampled with an angular spacing of 5° or less for reconstruction up to a bandwidth of 22.05 kHz.
- A novel technique is presented for undersampled HRTF data sets, where the HRTFs are decomposed into their envelope and carrier signals. This method achieves acceptable results in a mean square error sense for HRTFs measured with an angular undersampling factor of 2 to 4.

0.3 Outline

The paper is organized as follows. Section 1 studies the bandwidth limit of the sound field on a circle. It first considers a two-dimensional setup in Section 1.1, where an angular sampling theorem is developed, and results are generalized for the three-dimensional case in Section 1.2. Given this result, the sampling and interpolation of HRTFs are considered in Section 1.3. A new technique, presented in Section 2, is described in order to achieve good results in the case where HRTFs are sampled too coarsely along the angular dimension. It is based on the decomposition of HRTFs into their complex envelopes and carrier signals. Numerical simulations and experiments are presented in Section 3, together with a comparison with other existing methods. Conclusions are given in Section 4.

1 SOUND FIELD ANALYSIS ALONG A CIRCLE

In this section the angular bandwidth of the sound field along a circle is analyzed. In a first step, the study is performed in two dimensions in Section 1.1. A sampling theorem is presented to describe quantitatively the quality of reconstruction as a function of the angular sampling

frequency. The three-dimensional case is then discussed in Section 1.2. Based on these results, the sampling of HRTFs is discussed in Section 1.3 together with a first technique for their interpolation.

1.1 Angular Bandwidth of the Two-Dimensional Sound Field on a Circle

In this section two setups are explored. The first setup considers a plane wave arriving on a circle of microphones. The second setup considers a circular wave source. For the second setup a sampling theorem is presented.

1.1.1 Plane Wave Assumption

The first setup considered in this section consists of a circular microphone array of radius r . The coordinates of the different microphones are $\mathbf{r} = (r_x, r_y)$, with $r_x = r \cos \theta$ and $r_y = r \sin \theta$. A plane wave emitted by a plane source is incoming on the circle of microphones with a given angle θ_s , as shown in Fig. 1(a).

The sound pressure field along the circle is now defined as $p(\theta, t)$, where the sound field recorded at each angle θ is described as a function of time t . The sound field along the circle for a far-field source emitting a Dirac pulse is given by

$$p(\theta, t) = \delta \left[t - t_s - \frac{r \cos(\theta - \theta_s)}{c} \right] \quad (1)$$

with c being the speed of sound propagation and t_s the time taken by the wave to travel from the source position to the center of the microphone array. For simplicity of the calculations, we consider $t_s = 0$, meaning that for $t = 0$, the wave should pass through the center of the circular array. The sound field along a circle is shown in Fig. 1(b). Taking the Fourier transform of Eq. (1) along the temporal axis leads to

$$\tilde{p}(\theta, \omega) = e^{-j \frac{\omega}{c} r \cos(\theta - \theta_s)}. \quad (2)$$

In order to study the spatial bandwidth of the sound field recorded on a circle, the Fourier transform of Eq. (2) along the angular direction needs to be calculated. Denote the angular frequency by l_θ . Remark that $l_\theta \in \mathbb{Z}$ due to 2π periodicity of $\tilde{p}(\theta, \omega)$ in θ . We therefore have that

$$\hat{p}(l_\theta, \omega) = \frac{1}{2\pi} \int_0^{2\pi} \tilde{p}(\theta, \omega) e^{-j l_\theta \theta} d\theta \quad (3)$$

and

$$\tilde{p}(\theta, \omega) = \sum_{l_\theta=-\infty}^{\infty} \hat{p}(l_\theta, \omega) e^{j l_\theta \theta}. \quad (4)$$

To calculate $\hat{p}(l_\theta, \omega)$ in the present case, we make use of the Anger–Jacobi expansion [34], [9], [10]. It describes a plane wave as a function of a sum of Bessel functions,

$$e^{-j \frac{\omega}{c} r \cos(\theta - \theta_s)} = \sum_{l_\theta=-\infty}^{\infty} j^{l_\theta} J_{l_\theta} \left(\frac{\omega}{c} r \right) e^{j l_\theta (\theta - \theta_s)}. \quad (5)$$

Therefore by identification it can be observed from Eqs. (2), (4), and (5) that

$$\hat{p}(l_\theta, \omega) = j^{l_\theta} J_{l_\theta} \left(\frac{\omega}{c} r \right) e^{-j l_\theta \theta_s}. \quad (6)$$

The magnitude spectrum $|\hat{p}(l_\theta, \omega)|$ is plotted in Fig. 2 for a circle of 0.6-m radius. Note that most of the energy is present in a butterfly-shaped region. Outside of this region the energy present is decaying very fast. The intuition behind the shape of the spectrum is the following. For a low temporal frequency, because of the large wavelength, the angular frequency support is small. For increasing temporal frequency the butterfly is widening due to a larger angular frequency support. This support increases because of the smaller wavelength, indicating faster changes along the angular dimension. As can be observed from Fig. 2, the support of Eq. (6) can be approximated by the region satisfying

$$|l_\theta| \leq \frac{|\omega|}{c} r. \quad (7)$$

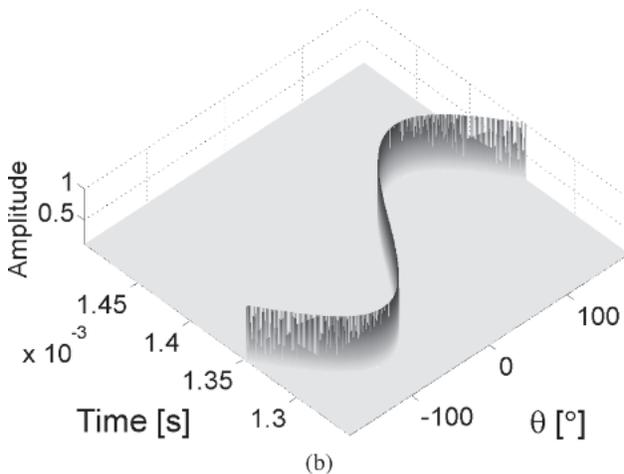
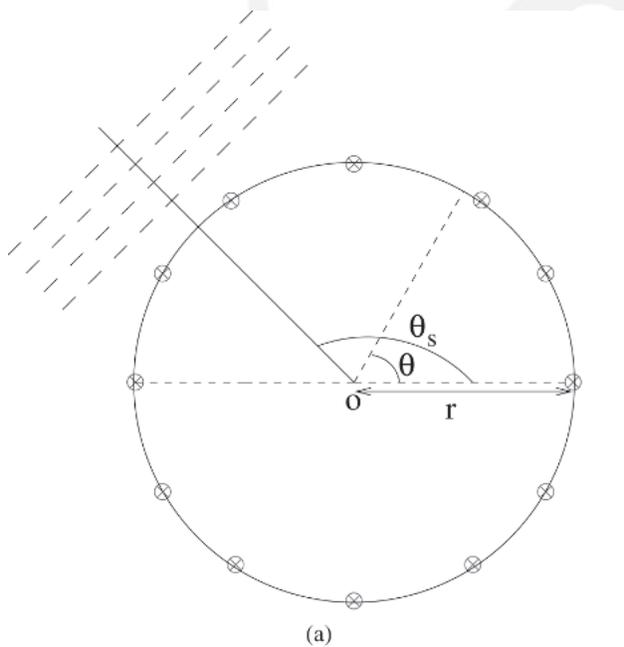


Fig. 1. Sound field analysis along a circle. (a) Plane wave incoming on a circle of radius r at angle θ_s . (b) Sound pressure field function studied along this circle.

Considering asymptotic expressions for the Bessel functions in [35], we can study how Eq. (6) behaves for large angular frequencies. For a given temporal frequency the spectrum behaves as (for large l_θ)

$$|\hat{p}(l_\theta, \omega)| \sim \frac{1}{\Gamma(l_\theta + 1)} \left(\frac{\omega}{2c} r \right)^{l_\theta} \quad (8)$$

where $\Gamma(l_\theta) = (l_\theta - 1)!$. Using Stirling's approximation, which represents an asymptotic formula for the Γ function, it can be shown that for large l_θ and a finite ω , the magnitude spectrum decays as

$$|\hat{p}(l_\theta, \omega)| \sim \frac{1}{\sqrt{2\pi l_\theta}} \left(\frac{re\omega}{2cl_\theta} \right)^{l_\theta} \sim \frac{c_1}{\sqrt{l_\theta}} \left(\frac{c_2}{l_\theta} \right)^{l_\theta} \quad (9)$$

with c_1 and c_2 two constants. This very fast decay can be observed in Fig. 2.

1.1.2 Circular Source

The second setup in this section explores the more realistic case, where the sound source is not considered as infinitely far away, but at a finite distance. The sound source is located at a distance s from the center of the array, and its coordinates are $s = (s_x, s_y)$, with $s_x = s \cos \theta_s$ and $s_y = s \sin \theta_s$. The setup is shown in Fig. 3. The same analysis is valid for distances s larger or smaller than the radius of the circle, as will be shown in the sequel. Consider the source to emit sound in the free field. The Green's function for the case of a circular sound source in two dimensions is given by [31], [35],

$$\tilde{p}(x, y, \omega) = \frac{-j}{4} H_0^{(2)} \left[\frac{\omega}{c} \sqrt{(x - s_x)^2 + (y - s_y)^2} \right] \quad (10)$$

with $H_0^{(2)}$ the zeroth-order Hankel function of the second kind. In our case we are interested in the field studied along a circular array. The time taken by sound to travel between the source and the different receivers located on the circle is

$$h(\theta) = \frac{\sqrt{(s_x - r \cos \theta)^2 + (s_y - r \sin \theta)^2}}{c}. \quad (11)$$

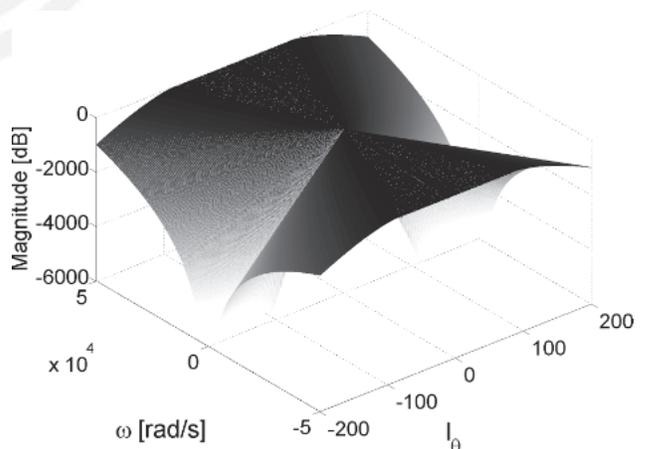


Fig. 2. Butterfly spectrum of sound pressure field along a circle considering the far-field assumption.

Using Eq. (11), the Green's function can be rewritten as

$$\begin{aligned} \tilde{p}(\theta, \omega) &= \frac{-j}{4} H_0^{(2)}[\omega h(\theta)] \\ &= \frac{-j}{4} H_0^{(2)} \left[\frac{\omega}{c} \sqrt{r^2 + s^2 - 2rs \cos(\theta - \theta_s)} \right]. \end{aligned} \tag{12}$$

Similarly to the Anger–Jacobi expansion described in Eq. (5), it is also possible to express a circular wave as a sum of Bessel functions. This is known as the addition theorem [34], [35],

$$H_0^{(2)} \left(\frac{\omega}{c} \|s - r\| \right) = \sum_{n=-\infty}^{\infty} J_n \left(\frac{\omega}{c} r_{<} \right) H_n^{(2)} \left(\frac{\omega}{c} r_{>} \right) e^{jn(\theta - \theta_s)} \tag{13}$$

with $\|s - r\|$ describing the distance between the source and the receivers, $r_{<}$ the lesser of r and s and $r_{>}$ the greater. In this section we consider $s \geq r$. The exact same results are obtained for the case where $s \leq r$, only s and r need to be swapped in the equations. This will be shown to be the case in the study of the angular bandwidth of HRTFs in Section 1.3.

Comparing the Fourier series in Eqs. (13) and (4) using Eq. (12) leads to

$$\hat{p}(l_\theta, \omega) = \frac{-j}{4} J_{l_\theta} \left(\frac{\omega}{c} r \right) H_{l_\theta}^{(2)} \left(\frac{\omega}{c} s \right) e^{-jl_\theta \theta_s}. \tag{14}$$

The spectrum obtained in Eq. (14) is shown in Fig. 4(a) for microphones located on a $0.6 = m$ radius and a source at $2 m$ from the center of the circle. When comparing the magnitude spectra of the far field and the circular source cases, we can observe that the circular case is obtained by multiplying the far-field case formula [Eq. (6)] with a zeroth-order Hankel function, as in Eq. (14). We show in Fig. 4(b) the effect of this multiplication for different ra-

tios $\rho = s/r$ in a slice of the magnitude spectrum at $\omega = 10\,000$ rad/s. It is shown that for a source located at three times the radius of the circle, the magnitude spectrum is very similar to the far-field case. When the source approaches the circle, the magnitude spectrum decays ever more slowly, as is shown for ρ , varying from 3 to 1.1.

A theorem quantifying the signal-to-noise ratio (SNR) of the reconstruction of the sound field is now presented for one source emitting in the free field. Consider a cut of the sound field spectrum at one temporal frequency. The spectrum only exists at integer values, namely, l_θ . This cut of the spectrum is shown schematically in Fig. 5(a). When sampling the sound field along the angular direction with an angular spacing of $\Delta\theta$, the angular sampling frequency l_{θ_s} is defined as

$$l_{\theta_s} = \frac{2\pi}{\Delta\theta}. \tag{15}$$

The spectrum of the sampled sound field is obtained by considering the original continuous angular two-

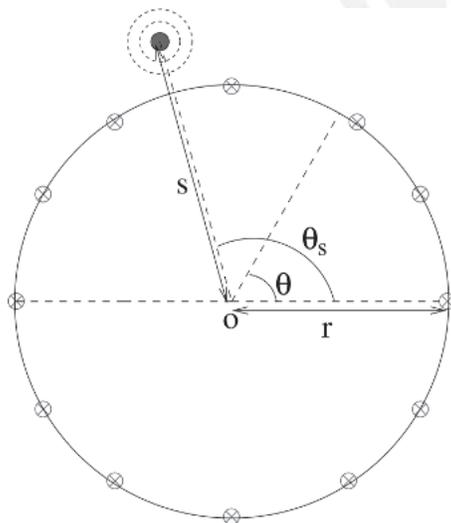


Fig. 3. Monopole source emitting sound in free field. Sound field is recorded along a circle of radius r with equally spaced microphones. Source is located at a finite distance s from the circle.

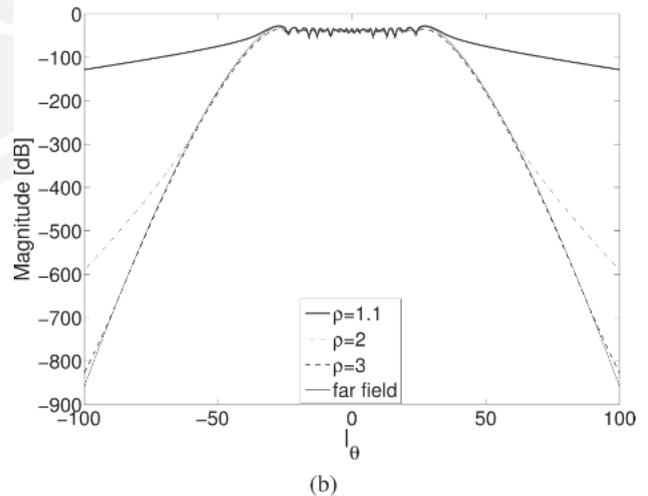
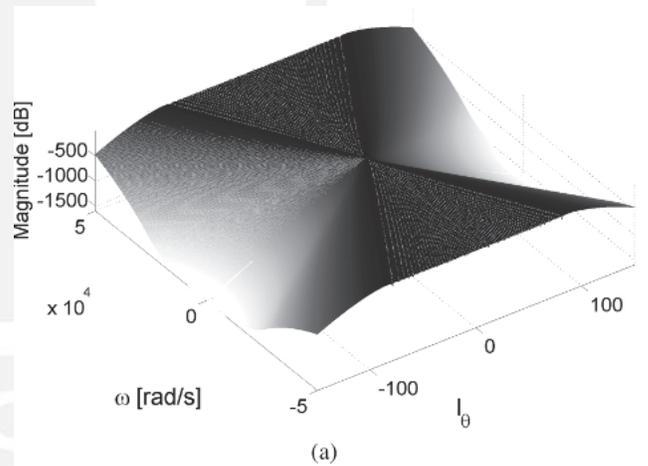


Fig. 4. Sound pressure field two-dimensional spectrum. (a) Butterfly spectrum of sound field along a circle for a circular source. (b) Comparison of sound field spectrum generated by a far-field source with spectrum generated by circular sources located at different distances from center of circle.

dimensional spectrum with all its spectral replicas, as shown in Fig. 5(b). It is given by

$$\hat{p}_S(l_\theta, \omega) = \frac{1}{\Delta\theta\Delta t} \sum_{k_1, k_2=-\infty}^{\infty} \hat{p}\left(l_\theta - \frac{2\pi k_1}{\Delta\theta}, \omega - \frac{2\pi k_2}{\Delta t}\right). \quad (16)$$

Due to the nonperfect bandlimitedness of the two dimensional spectrum, perfect reconstruction is not achieved. A small error remains on the reconstruction depending on the decay of the spectrum outside the butterfly region. Call $\text{SNR}(l_{\theta_s}, \omega_0)$ the SNR of the reconstruction for a sinusoid emitted at frequency $\omega = \omega_0$ with the microphones positioned with an angular sampling frequency l_{θ_s} . In the present case the SNR is defined as

$$\text{SNR}(l_{\theta_s}, \omega_0) = \frac{\sum_{l_\theta=-\infty}^{\infty} |\hat{p}(l_\theta, \omega_0)|^2}{4 \sum_{l_\theta \geq l_{\theta_s}/2} |\hat{p}(l_\theta, \omega_0)|^2}. \quad (17)$$

The numerator in Eq. (17) corresponds to the energy of the spectrum of the sound field at the temporal frequency ω_0 . The denominator in Eq. (17) corresponds to the energy contained in the spectral repetitions that will contaminate the reconstruction in the spectral domain of interest. Two different kinds of energy are present in this denominator, the “in-band” and the “out-of-band” energy. The in-band energy corresponds to the energy of all the spectral rep-

etitions in the domain of interest, namely, $[-l_{\theta_s}/2, l_{\theta_s}/2]$. The out-of-band energy is the energy present in the spectrum that is outside the domain of interest. It can be shown that the in-band and out-of-band energies are equal in the case of an infinite line of microphones. This explains the factor of 4 in the denominator of Eq. (17).

Theorem 2.1 In two dimensions, assume one single source emitting in the free field at a frequency $\omega = \omega_0$ located at a distance s from the center of a circle of radius r . When sampling the sound field on the circle with an angular sampling frequency of l_{θ_s} , for a particular $\omega = \omega_0$, and reconstructing it using an ideal interpolator, the SNR of the reconstructed signal in the band $[-l_{\theta_s}/2, l_{\theta_s}/2]$ can be expressed as

$$\text{SNR}(l_{\theta_s}, \omega_0) = \frac{\sum_{l_\theta=-\infty}^{\infty} \left| J_{l_\theta}\left(\frac{\omega_0}{c} r_{<}\right) H_{l_\theta}^{(2)}\left(\frac{\omega_0}{c} r_{>}\right) \right|^2}{4 \sum_{l_\theta \geq l_{\theta_s}/2} \left| J_{l_\theta}\left(\frac{\omega_0}{c} r_{<}\right) H_{l_\theta}^{(2)}\left(\frac{\omega_0}{c} r_{>}\right) \right|^2} \quad (18)$$

where $H_{l_\theta}^{(2)}$ is the Hankel function of order l_θ of the second kind, J_{l_θ} is the Bessel function of order l_θ , $r_{<}$ is the lesser of r and s and $r_{>}$ the greater.

Numerically Eq. (18) has been simulated, and the results are shown in Fig. 6 for temporal frequencies varying

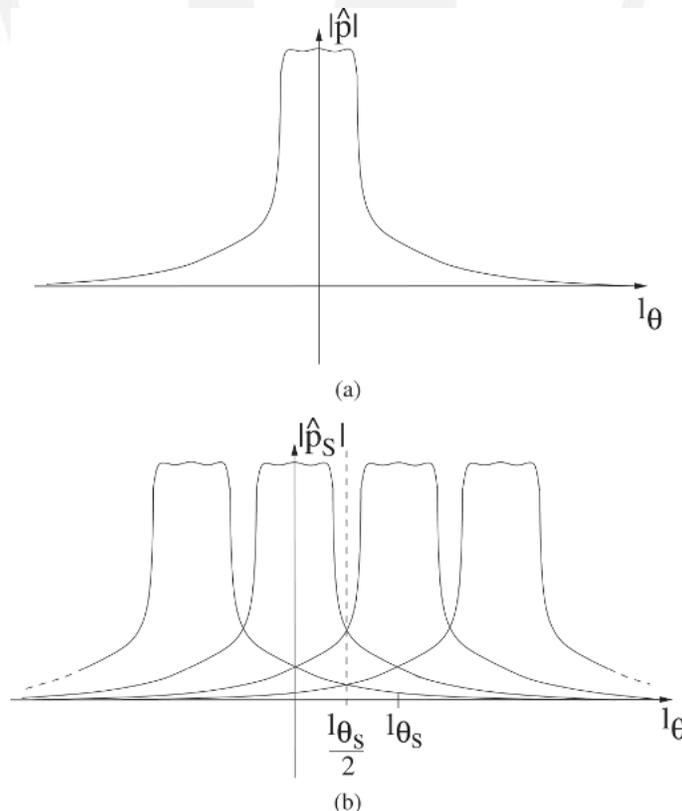


Fig. 5. Sampling theorem for sound field along a circle. (a) Cut at one specific temporal frequency of sound field spectrum. (b) Due to sound field sampling in angular direction, spectrum of sampled sound field contains multiple periodic repetitions of original spectrum.

from 1 to 8000 rad/s. The simulations have been performed for a source located at 2 m from a circle of radius 1 m. A numerical example is given, considering a circular wave emitting at 8000 rad/s (1.3 kHz). Using the approximation Eq. (7), one would require an angular spacing between consecutive microphones of 7.6° while from Eq. (18) and Fig. 6 it turns out that to obtain a mean SNR of 50 dB on the reconstruction, an angular spacing of 5.9° is necessary.

1.2 Angular Bandwidth of the Three-Dimensional Sound Field on a Circle

Up to now the sound field along a circular array has been studied for a two-dimensional plane wave and for a circular sound source. In a three-dimensional setup the sound source is a spherical source emitting in three dimensions. We therefore modify the setup slightly and introduce the third dimension. We consider a similar setup as in Fig. 3. The coordinates of the different microphones are now (r_x, r_y, r_z) , with $r_x = r \cos \theta$, $r_y = r \sin \theta$, and $r_z = 0$. The sound source is located at a distance s from the center of the array and with the coordinates (s_x, s_y, s_z) , where $s_x = s \cos \theta_s$ and $s_y = s \sin \theta_s$. The distance s can be decomposed into a distance $s_p = \sqrt{s_x^2 + s_y^2}$ in the plane of the microphones and s_z with

$$|s| = \sqrt{s_p^2 + s_z^2} \tag{19}$$

The different signals recorded at any angle on the circle are gathered in a function $p(\theta, t)$. When the sound source emits a Dirac signal, the microphones located on the circle measure

$$p(\theta, t) = \frac{\delta[t - h(\theta)]}{4\pi ch(\theta)} \tag{20}$$

where $h(\theta)$ is defined as

$$h(\theta) = \frac{\sqrt{(s_x - r \cos \theta)^2 + (s_y - r \sin \theta)^2 + s_z^2}}{c} = \frac{\sqrt{s^2 + r^2 - 2rs_p \cos(\theta - \theta_s)}}{c} \tag{21}$$

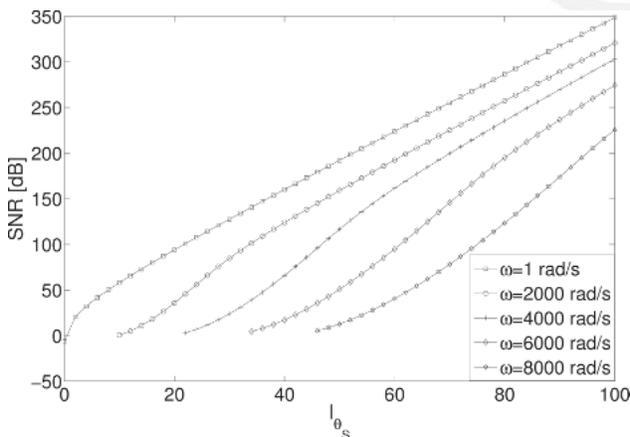


Fig. 6. Numerical simulation of SNR calculation for different ω_0 .

Similarly to the two-dimensional case, the function $h(\theta)$ represents, at each angular microphone position θ , the time of arrival taken by the sound to propagate from the source to the different microphones. An example of this function is given in Fig. 7. To study the angular bandwidth of the sound field recorded on a circle, the two-dimensional Fourier transform (2D-FT) of $p(\theta, t)$ is again studied,

$$\hat{p}(l_\theta, \omega) = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} p(\theta, t) e^{-j(l_\theta \theta + \omega t)} dt d\theta \tag{22}$$

$$= \frac{1}{8\pi^2 c} \int_0^{2\pi} \frac{e^{-j\omega h(\theta)}}{h(\theta)} e^{-jl_\theta \theta} d\theta. \tag{23}$$

A closed-form solution of this equation has not yet been obtained. Therefore, to estimate the spatial support of the spectrum, we present some approximation formulas in the sequel.

In the study of the spatial bandwidth of the sound field, the attenuation depending on the distance traveled has been shown to be negligible for the study of the essential support of the sound field spectrum [36]. It mostly affects the decay of the function. Therefore a good estimate of the angular bandwidth corresponds to the support of the function

$$\hat{p}(l_\theta, \omega) = \int_0^{2\pi} e^{-j\omega h(\theta)} e^{-jl_\theta \theta} d\theta. \tag{24}$$

Note that Eq. (24) corresponds to the Fourier transform of a phase modulation (PM) signal, where the carrier frequency would be equal to zero and the modulation function would be $h(\theta)$. The bandwidth of this signal can be estimated easily using Carson's rule [37], [38]. For a temporal frequency ω the bandwidth (BW) of $\hat{p}(l_\theta, \omega)$ along the angular frequency axis can be approximated by

$$BW_{l_\theta}[\hat{p}(l_\theta, \omega)] \approx \max_{\theta} \left[\frac{dh}{d\theta} \right] \omega + W \tag{25}$$

with W being the bandwidth of $h(\theta)$.¹

¹To apply Carson's rule, either $W \ll \max_{\theta} [dh/d\theta] \omega$ or $W \gg \max_{\theta} [dh/d\theta] \omega$ needs to be satisfied.

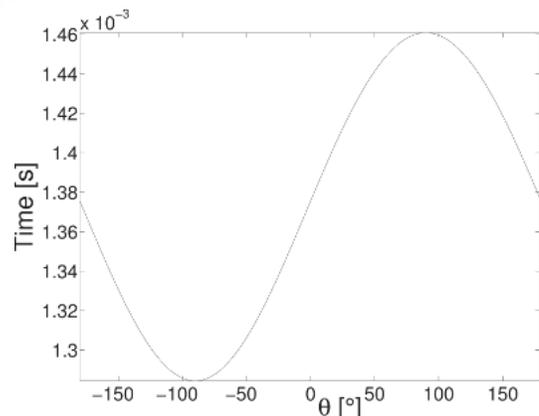


Fig. 7. Time of arrival from source to microphone at each angular position.

The maxima of the first derivative of $h(\theta)$ with respect to θ can be found to be [36]

$$\frac{dh}{d\theta} = \pm \frac{\sqrt{A + 2s_p r} - \sqrt{A - 2s_p r}}{2c} \quad (26)$$

with $A = s^2 + r^2$. For sources located in the same plane as the circular array, the expression of the derivative gets simpler.

- For a source located inside the circular array,

$$\max_{\theta} \left[\frac{dh}{d\theta} \right] = \pm \frac{s}{c}. \quad (27)$$

- For a source located outside the circular array

$$\max_{\theta} \left[\frac{dh}{d\theta} \right] = \pm \frac{r}{c}. \quad (28)$$

Further we can also show that the maximum derivative of $h(\theta)$ associated with a source outside the plane of the array is always smaller than the derivative for a source on the plane,

$$\left| \frac{\sqrt{A + 2s_p r} - \sqrt{A - 2s_p r}}{2c} \right| \leq \left| \frac{\min(s_p, r)}{c} \right|.$$

The maximum derivative is thus only dependent on the minimum between the distance from the center of the array to the source and the radius of the array.

$$\max_{\theta} \left[\frac{dh}{d\theta} \right] \leq \frac{\min(s_p, r)}{c}. \quad (29)$$

The signal $h(\theta)$ is a very smooth signal, and therefore its bandwidth W can be shown to be approximately zero (except for specific cases where the source is very close to the microphones, as has been discussed in Section 1.1). Therefore Eq. (25) becomes

$$\text{BW}_{l_\theta}[\hat{p}(l_\theta, \omega)] \approx \frac{\omega}{c} \min(s_p, r). \quad (30)$$

In most cases the source is considered to be in the same plane as the circular array, and we have then

$$\text{BW}_{l_\theta}[\hat{p}(l_\theta, \omega)] \approx \frac{\omega}{c} \min(s, r). \quad (31)$$

As indicated by Eq. (31), the angular frequency support increases linearly with the temporal frequency. This corresponds exactly to the butterfly support presented in Fig. 8(a), where most of the energy is present in the region satisfying

$$|l_\theta| \leq |\omega| \frac{r}{c}. \quad (32)$$

This spectrum was obtained by simulating the sound field measured on a circular array of 0.6-m radius with a source at 2 m from the center of the array. The same setup was followed for real measurements, where 1000 room impulse responses (RIRs) were measured in a sound-

insulated room. The measurements were performed with a microphone every 0.36° along a circle. These measurements were performed using an automatic pan/tilt unit *PTU-D46-70* rotating with a precision of 0.03° in the median plane. The RIRs were measured using a logarithmic sweep [39] of 2-second duration. The loudspeaker used was a Genelec 1029A, and the microphone Beyer-dynamic *MC-740*. The two-dimensional spectrum obtained from the real RIR measurements is shown in Fig. 8(b). The support of the spectrum for the real measurements corresponds well to the simulations. All these measurements are available online [40].

1.3 HRTF Sampling and Interpolation

In Sections 1.1 and 1.2 the spatial bandwidth of the sound field along a circle has been studied. The same theory can be applied for the study of HRTF interpolation. HRTF measurements are carried out in an anechoic chamber to characterize the effect of the pinnae, head, and torso of a person on the perceived sound [7]. A typical setup for HRTF measurements is shown in Fig. 9. The loudspeakers are located along a circle around the person. The microphone is located at the entrance of the ear canal of the listener to capture the sound. By constructing $p(\theta, t)$ as the HRIRs measured at every possible angle, the support of the spectrum $\hat{p}(l_\theta, \omega)$ can be estimated using the theory

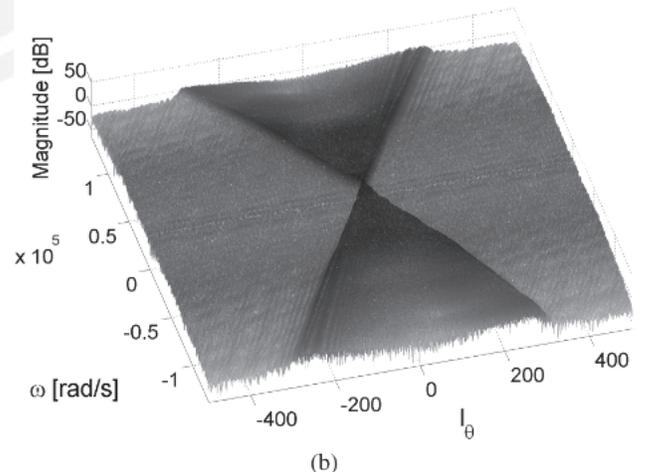
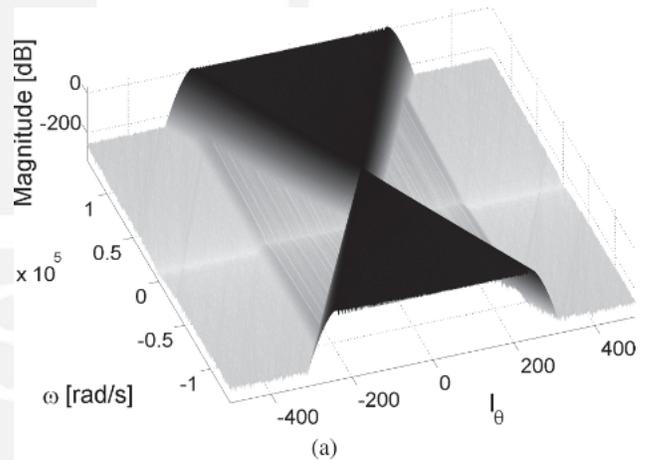


Fig. 8. 2D-FT of sound field recorded on a circular array. (a) Simulated case. (b) Real measurements.

presented in Section 1.2. With the head well centered in the middle of the loudspeaker array and the position of the microphone being $d/2 = 90$ mm away from the center of the circle (half the spacing between the two ears) [7] Eq. (32) is rewritten as

$$|l_\theta| \leq |\omega| \frac{d}{2c} \approx |\omega| \frac{0.09}{c}. \quad (33)$$

Therefore for any temporal frequency the necessary angular spacing between consecutive loudspeaker positions can be derived. Considering the energy outside the butterfly spectrum as negligible, the angular sampling frequency needs to satisfy

$$l_{\theta_s} > 2|\omega_{\max}| \frac{d}{2c} \approx 2|\omega_{\max}| \frac{0.09}{c} \quad (34)$$

with ω_{\max} the maximum temporal frequency present in the signal. Note that Eq. (34) can be seen as the usual half-wavelength rule. The distance along the circle between two measurement positions, denoted by D_m , can be written as

$$D_m = \frac{d\Delta\theta}{2} = \frac{d}{2} \frac{2\pi}{l_{\theta_s}}. \quad (35)$$

Further, using the relation

$$\lambda_{\min} = \frac{2\pi c}{\omega_{\max}} \quad (36)$$

with λ being the wavelength, we can rewrite Eq. (34) as

$$D_m < \frac{\lambda_{\min}}{2}. \quad (37)$$

In other words, the spacing between consecutive samples along the circle should be at least smaller than half the minimum wavelength present in the excitation signal.

In particular, Eq. (34) indicates that in order to sample HRTFs for an average adult human ($d \approx 0.18$ m) with a temporal sampling rate of 44.1 kHz, a spacing of at least 4.9° is necessary. Sampling the HRTFs with too large an angular spacing leads to spectral repetitions corrupting all temporal frequencies above ω_{\max} in Eq. (34).

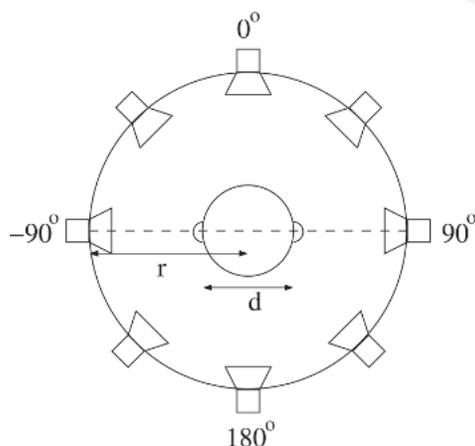


Fig. 9. Setup for recording of HRTFs.

1.3.1 HRTF Interpolation

Interpolation of the data set is done by observing that the spectrum of the HRTFs contains almost no energy outside the butterfly region. Therefore as the signal to be interpolated is almost bandlimited, good results are obtained using a sinc interpolator. The implementation of this interpolator happens most efficiently by zero padding in the frequency domain. This interpolation is very suitable in the case of a circular array since the Fourier transform is applied on an array that is 2π periodic. In the case of interpolation along nonperiodic arrays, border effects decrease the interpolation performance, as described in [4].

Depending on the angular sampling of the database of HRTFs considered, interpolation is only applied for frequencies satisfying Eq. (34). Higher frequencies will not be interpolated correctly because of the spectral repetitions leading to aliasing. For these higher frequencies other techniques need to be used, as will be shown in Section 2.

1.3.2 Head Shadowing

The theory expressed in the preceding section is only valid for the case of HRTFs where the effect of head shadowing is not considered. In practice, waves are diffracted by the head. This diffraction has to be taken into account. Diffraction has a large impact on the level of amplitude of the HRTF [7]. It also affects the shape of the function $h(\theta)$. In the model given in [41] the HRTFs are expressed as

$$\tilde{h}(\theta, \rho, \mu) = \frac{\rho}{\mu} e^{-i\mu\rho\Psi} \quad (38)$$

with

$$\Psi(\theta, \rho, \mu) = \sum_{m=0}^{\infty} (2m+1)P_m(\cos\theta) \frac{h_m(\mu\rho)}{h'_m(\mu)}$$

where μ is the normalized temporal frequency, P_m is a Legendre polynomial of degree m , and h_m is an m th-order spherical Hankel function.² Taking the Fourier transform of Eq. (38), a butterfly spectrum satisfying Eq. (33) is also observed, as shown in Fig. 10(a). The spectrum obtained considers HRTFs measured on a Kemar head [42] sampled every 5° in an anechoic chamber and is shown in Fig. 10(b). There Eq. (34) is also satisfied since a spacing of 5° results in almost no aliasing at 44.1 kHz.

2 HRTF INTERPOLATION IN A TEMPORAL ENVELOPE DOMAIN

With the analysis presented in Section 1.2, to reconstruct HRTFs at a sampling rate of 44.1 kHz, 72 HRTFs are necessary to achieve good interpolation results with little aliasing. This number is quite large and many avail-

²Note that this notation is different from the one used in this paper but has been kept as in the original publication [41].

able databases do not contain such a fine sampling. In this section a method will be presented to increase the performance of the interpolation in case of a coarser HRTF sampling. Considering a database containing two or four times fewer HRTFs than the number dictated by Eq. (34), it will be shown that satisfactory results in a least mean squared sense can still be achieved.

The technique presented in this section is based on a subband decomposition of the HRTFs. A scheme illustrating the method is shown in Fig. 11. As discussed in Section 1, the best interpolation is achieved using the sinc interpolation when no aliasing is present. Therefore this technique is still used for the range of temporal frequencies satisfying Eq. (34). This frequency range is denoted by a in Fig. 11. For the higher frequencies a new technique is presented. In this technique each subband is considered and analyzed separately. These subbands are denoted by b in Fig. 11. At each angular position θ_i , the subband signal is decomposed into its complex envelope signal and carrier signal. To apply interpolation to the complex envelope, the proposed algorithm aligns the carriers of the different angular positions to the carrier of a reference position. Therefore only the envelope signals are differing between the signals at various angular positions and interpolation is effectively applied to the envelopes. Due to the smaller bandwidth of the envelopes, spatial aliasing is reduced. The bottom two panels of Fig. 11 illustrate the

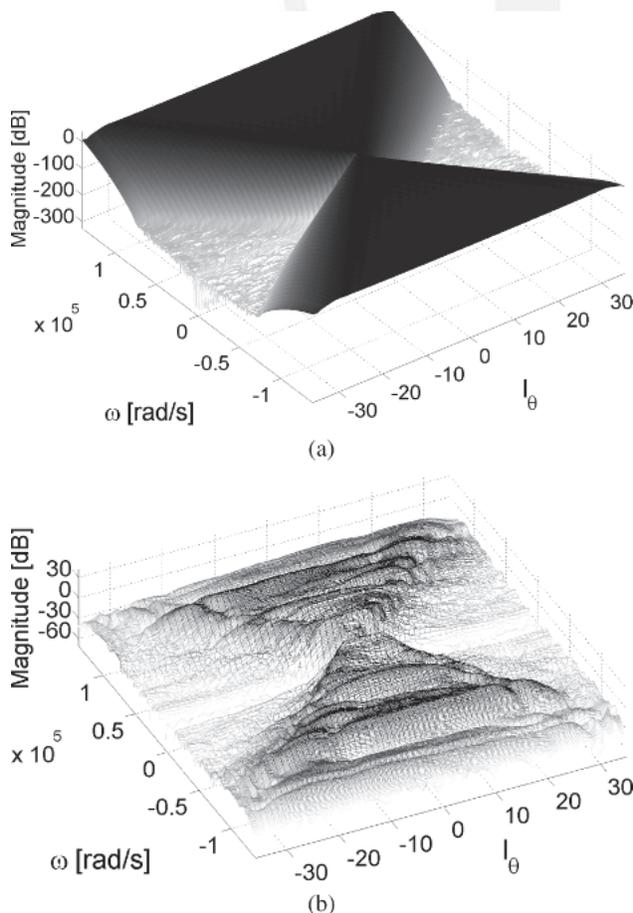


Fig. 10. Two-dimensional spectra of HRTFs. (a) Using a diffraction model. (b) Using measured data.

signals recorded at positions θ_1 and θ_2 . The envelopes of these two signals are shown as dotted lines. After carrier alignment of the signal at position θ_2 with the carrier of the reference signal at position θ_1 , the new aligned carrier at position θ_2 is shown in bold. It can be observed that the envelope of the signal at position θ_2 is kept unmodified by this alignment. Based on the idea that the envelope signals evolve more slowly along the angular dimension than the HRTFs themselves, satisfactory interpolation results for the envelope can be obtained even when the angular spacing is not satisfying Eq. (34).

The reason to use a subband decomposition is based on the fact that the carrier alignment factor needs to be estimated for each subband separately. In case the HRTFs would be exact delayed versions of one another, no subband decomposition would be necessary. But due to the fact that measured HRTFs do not present such a behavior, the proposed solution is to estimate a carrier alignment factor for each subband.

Note that this technique is suitable for HRTFs since HRTFs are largely determined by the direct sound. For HRTFs measured in a reverberant environment (binaural

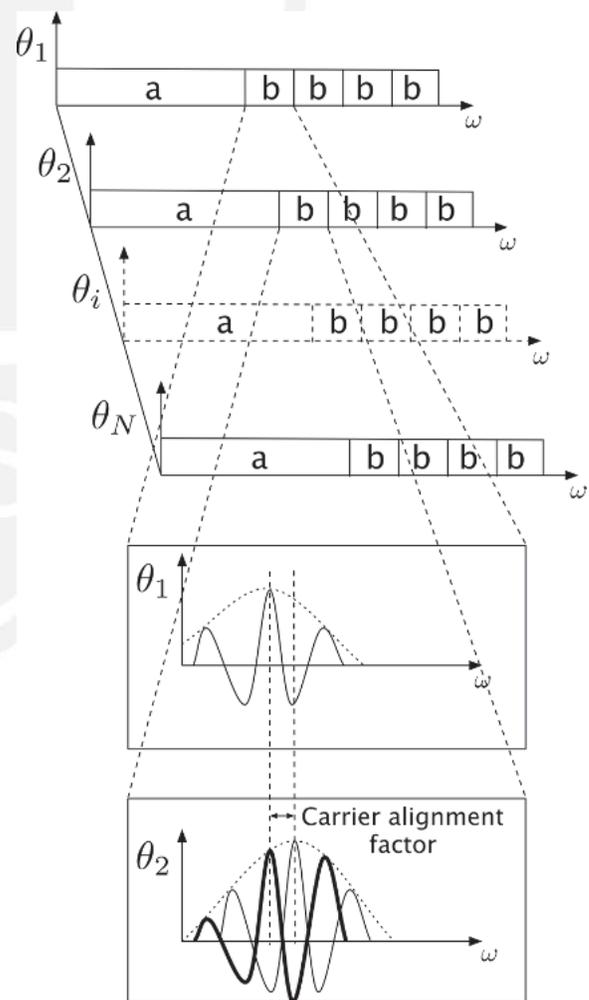


Fig. 11. Scheme of presented technique for HRTF interpolation. In frequency range a , interpolation happens with a sinc interpolator. In frequency bands b , technique presented in this section is used.

room impulse responses) or in case of multiple sources, the carrier alignment would not lead to a large reduction in the spatial bandwidth of the carrier-aligned HRTFs as is the case in our setup.

The technique can be summarized in four steps.

1) The first step consists in the estimation of the carrier alignment factors for each subband and at each angular position. The estimation of those factors is presented in Section 2.1. Section 2.2 shows further that the angular support of the carrier-aligned HRTFs is reduced when compared to the original one.

2) Once the HRTFs are aligned, the next step is to interpolate the carrier-aligned HRTFs. This is discussed in Section 2.3.

3) The third step is to interpolate the alignment factors used to align the different HRTFs, which is also discussed in Section 2.3.

4) In the last step all interpolated HRTFs are obtained by multiplying each interpolated carrier-aligned HRTF with its corresponding carrier alignment factor.

2.1 Computing the Carrier Alignment Factors in Subbands

Each subband signal is denoted by $h_i(\theta, t)$, where the index i corresponds to the subband index and θ stands for the azimuth angle characterizing the HRTF position. The Fourier transform with respect to time of $h_i(\theta, t)$ is denoted by $\tilde{h}_i(\theta, \omega)$. The basic idea behind the algorithm is that each subband signal is considered to be the product of a carrier and an envelope signal. For every subband the carriers of the different positions are aligned with the carrier of one reference position, called θ_r . Considering only the positive frequencies in $\tilde{h}_i(\theta, \omega)$, this alignment is achieved by multiplying each HRTF spectrum subband of position θ by a complex number $c_i(\theta) = e^{j\alpha_{\theta,i}}$. Denoted by $\tilde{k}_i(\theta, \omega)$ the aligned HRTF subband, we have that

$$\tilde{k}_i(\theta, \omega) = c_i(\theta)\tilde{h}_i(\theta, \omega). \quad (39)$$

The complex value $c_i(\theta)$ is obtained by minimizing the error J_i defined as

$$J_i = \int_{\omega_i}^{\omega_{i+1}} |\tilde{h}_i(\theta_r, \omega) - \tilde{h}_i(\theta, \omega)c_i(\theta)|^2 d\omega \quad (40)$$

with ω_i and ω_{i+1} being the start and stop frequencies of band i .

The value of $c_i(\theta)$ that minimizes J_i is obtained by calculating the first and second derivatives of J_i with respect to $\alpha_{\theta,i}$. The value of the minimum is

$$c_i(\theta) = \frac{\int_{\omega_i}^{\omega_{i+1}} \tilde{h}_i(\theta_r, \omega)\tilde{h}_i^*(\theta, \omega) d\omega}{\left| \int_{\omega_i}^{\omega_{i+1}} \tilde{h}_i(\theta_r, \omega)\tilde{h}_i^*(\theta, \omega) d\omega \right|}. \quad (41)$$

By calculating $c_i(\theta)$ as in Eq. (41) for each azimuthal angle and multiplying the corresponding subband signal with it as in Eq. (39), the carrier signals of all the HRTFs are aligned with the reference signal in a least mean square sense. Fig. 12 shows an example of this carrier alignment procedure. The signal measured at the reference position θ_r is shown by dashed lines together with its envelope as a dash-dotted line. The signal observed at position θ_2 is shown in full lines together with its envelope in dotted lines. Multiplying the signal with the carrier alignment factor yields the signal indicated as a bold solid line. It can be seen that both the signal at position θ_2 and its carrier-aligned version have the same envelope. The carrier of the newly obtained signal is now aligned with the reference signal carrier. The angular bandwidth occupied by this new HRTF data set is reduced, as will be shown in Section 2.2. This reduction of the bandwidth is the key factor in the possibility of achieving satisfactory results when interpolating signals whose angular spacing is not satisfying Eq. (34).

2.2 Support of Aligned HRTFs

For each subband HRTF $h_i(\theta, t)$ the analytical signal $h_{+i}(\theta, t)$ is given by

$$h_{+i}(\theta, t) = h_i(\theta, t) + j\check{h}_i(\theta, t) \quad (42)$$

where $\check{h}_i(\theta, t)$ stands for the Hilbert transform of $h_i(\theta, t)$. Recall that

$$\check{h}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{h(\tau)}{t - \tau} d\tau. \quad (43)$$

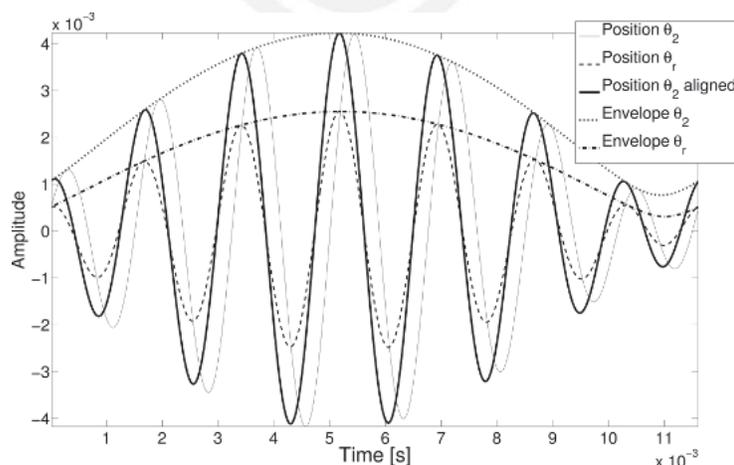


Fig. 12. Different subband signals.

The Fourier transform with respect to time of $h_{+i}(\theta, t)$ is denoted by $\tilde{h}_{+i}(\theta, \omega)$. For the positive frequencies it can be shown that

$$\tilde{h}_{+i}(\theta, \omega) = 2\tilde{h}_i(\theta, \omega). \quad (44)$$

For the negative frequencies $\tilde{h}_{+i}(\theta, \omega) = 0$.

The signal $h_{+i}(\theta, t)$ can be decomposed into a carrier $g_i(\theta, t)$ and a complex envelope signal $e_i(\theta, t)$. Writing the carrier signal as $g_i(\theta, t) = e^{j(\omega_i t - \alpha_{\theta,i})}$, $h_{+i}(\theta, t)$ can be written as

$$\begin{aligned} h_{+i}(\theta, t) &= e_i(\theta, t)g_i(\theta, t) = e_i(\theta, t)e^{j(\omega_i t - \alpha_{\theta,i})} \\ &= e_i(\theta, t)c_i^*(\theta)e^{j\omega_i t}. \end{aligned} \quad (45)$$

As discussed in Section 2.1, each carrier is aligned with the carrier of a reference HRTF. The direction of the reference HRTF is θ_r . After realignment of all HRTFs with this reference using Eq. (41),³ the new HRTF set is denoted by $k_{+i}(\theta, t)$, with

$$k_{+i}(\theta, t) = e_i(\theta, t)e^{j\omega_i t}. \quad (46)$$

Consider now the two-dimensional spectrum of the HRTF for one subband. Call $\hat{h}_{+i}(l_\theta, \omega)$ the 2D-FT of $h_{+i}(\theta, t)$ and $\hat{k}_{+i}(l_\theta, \omega)$ the 2D-FT of $k_{+i}(\theta, t)$. One can then write

$$\hat{k}_{+i}(l_\theta, \omega) = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} e_i(\theta, t) e^{-j(\omega - \omega_i)t} dt e^{-jl_\theta\theta} d\theta. \quad (47)$$

Denoting by $\hat{e}(l_\theta, \omega)$ the 2D-FT of the complex envelope signals $e_i(\theta, t)$, Eq. (47) can be rewritten as

$$\hat{k}_{+i}(l_\theta, \omega) = \hat{e}(l_\theta, \omega - \omega_i). \quad (48)$$

This shows that after carrier alignment, the two-dimensional spectrum of the aligned HRTFs is equivalent to the support of the complex envelopes of the HRTFs shifted to the temporal frequency of the subband considered.

As mentioned in Section 2.1, multiplying the signal by the alignment factor does not modify the real envelope of the signal. The real envelope of the signal $h_i(\theta, t)$ is obtained by calculating

$$\text{envelope}_i(\theta, t) = |h_{+i}(\theta, t)|. \quad (49)$$

In the present technique the signals $h_{+i}(\theta, t)$ are aligned by multiplication by a phase factor $e^{j\alpha_{\theta,i}}$. This multiplication does not modify the real envelope signal since

$$\text{envelope}_i(\theta, t) = |e_i(\theta, t)g_i(\theta, t)c_i(\theta)| = |e_i(\theta, t)|. \quad (50)$$

As an example of the efficiency of the carrier alignment technique, consider a simple model for HRTFs where only the delay and attenuation are considered [31]. Fig. 13(a) represents the 2D-FT of this database sampled every 10° . After carrier alignment the two-dimensional spectrum is much narrower, as can be observed in Fig. 13(b). Comparing Fig. 13(a) and (b) at the particular temporal frequency of 5 kHz ($10\,000\pi$ rad/s) yields Fig. 14. It can be

observed that the carrier-aligned HRTFs have a much narrower spectral content and can therefore be sampled with a larger spacing between consecutive angular positions.

2.3 Interpolation

The interpolation of the HRTFs is carried out in three steps, corresponding to steps 2) to 4) mentioned previ-

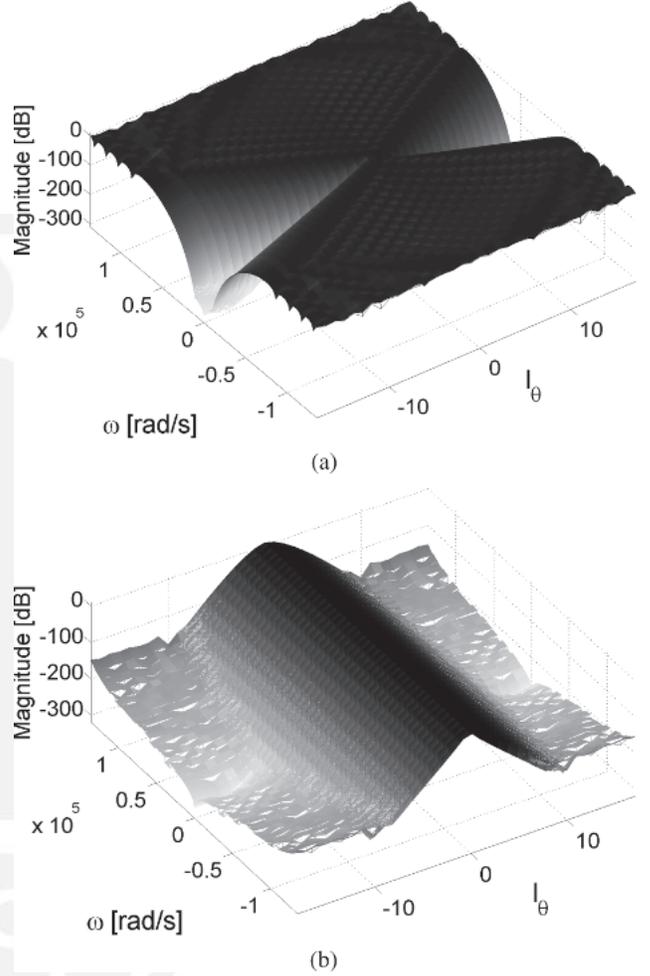


Fig. 13. Two-dimensional spectra of HRTF. (a) Original HRTF data set. (b) Carrier-aligned HRTFs.

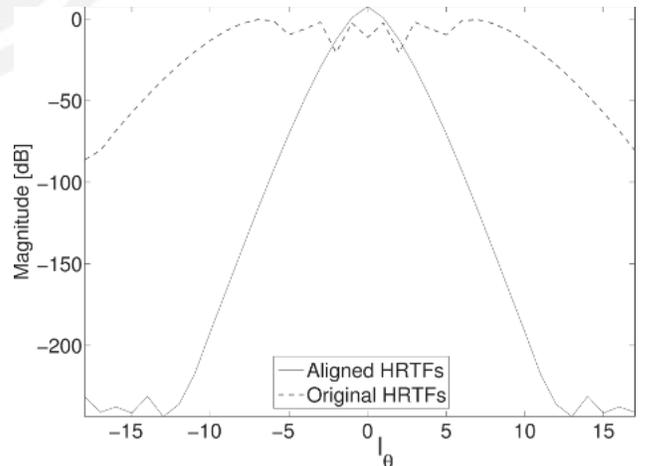


Fig. 14. At 5 kHz, angular frequency support of original HRTFs compared to support of carrier-aligned HRTFs.

³In this analysis we consider a perfect alignment of the different carriers with the reference one.

ously. The first step is to interpolate the carrier-aligned signals $k_{+i}(\theta, t)$. The second step is the interpolation of the alignment factors. The third step consists in multiplying the carrier-aligned HRTF at a specific angular position with the corresponding alignment factor. The interpolated HRIR is obtained by inverse Fourier transform of the spectrum where the negative frequencies have been added. The negative frequencies are the complex conjugate of the positive frequencies to ensure real HRIRs.

For the interpolation of $k_{+i}(\theta, t)$ a simple low-pass interpolation filter using the four nearest neighbors is used. Similarly to the analysis performed in Section 1.3, it would be interesting to know how many angular positions are necessary to interpolate the HRTFs at any angular position. Unfortunately this number cannot be calculated easily in a closed-form solution. This is due to the fact that the angular frequency support of the complex envelopes is not known in general. Nevertheless, as the angular support corresponding to the carrier-aligned signals corresponds to the one of the envelope signals, it can be concluded that the number of angular samples is reduced.

The second step is the interpolation of the carrier alignment factors $c_i(\theta)$ along the angular direction. To interpolate these complex numbers, only the phases $\alpha_{\theta,i}$ are interpolated. This interpolation is also achieved by a low-pass filter using the four nearest neighbors. Nevertheless, a difficulty due to unwrapping appears when getting the phase values from the complex numbers $c_i(\theta)$. The phase $\alpha_{\theta,i}$ is only known in the range from $-\pi$ to π . Denoting by $\hat{\alpha}_{\theta,i}$ the phase between $-\pi$ to π , the phase to be found satisfies

$$\alpha_{\theta,i} = \hat{\alpha}_{\theta,i} + 2k\pi \quad (51)$$

with $k \in \mathbb{Z}$. The unknown is here the value of k . Interpolation between wrong phases leads to wrong alignment factors and poor performance of the proposed algorithm. The unwrapping problem can be addressed by considering a recursive algorithm, where the subbands at lower temporal frequencies are used to estimate the range of values of the phase in the higher temporal frequency subband. The estimate of the new subband phase is obtained by assuming a linear phase. This estimation of the new phase allows us then to find the k in Eq. (51). Note that when the HRTF filter is very far from a linear-phase filter, some wrong evaluation of k can be found. Also, remark that as the algorithm is recursive, the higher the temporal frequency, the larger the error can become. With our algorithm the phases for the angular positions at the different subbands are shown in Fig. 15 for different HRTF databases. In the sequel the databases will be denoted by database 1, database 2, and database 3. Database 1 simply considers at each angular position a delay and an attenuation satisfying the wave equation. Database 2 uses the model of head shadowing discussed in Section 1.3.2 from [41]. Database 3 has been obtained by measurement of the HRTFs of a Kemar head [42].

In the technique presented one of the important parameters to be chosen is the number of subbands in which the

HRTFs are decomposed. Note that there exists a tradeoff in the choice of bandwidth and the number of subbands. Choosing too large a band makes it difficult to align one subband HRTF correctly at one position with its reference using only one value, the corresponding $c_i(\theta)$. In other words, the error J_i in Eq. (40) might remain non-negligible

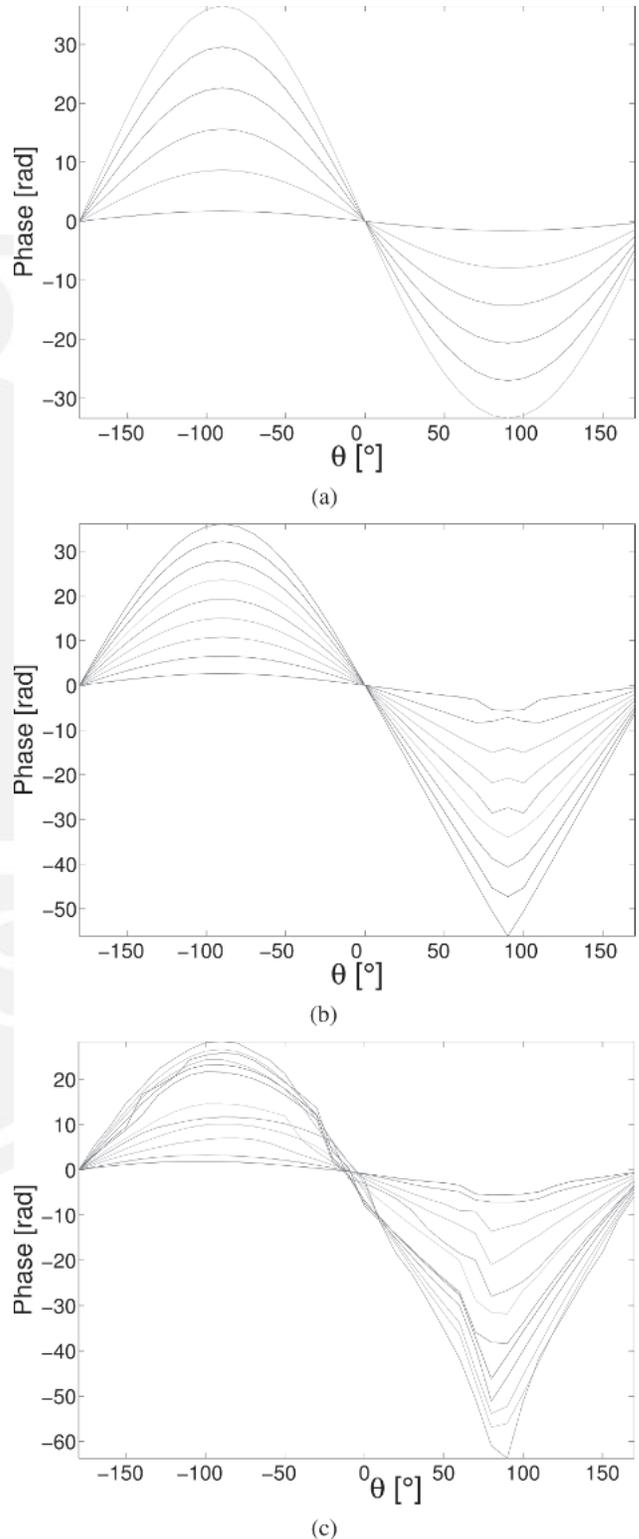


Fig. 15. Phase corresponding to carrier alignment factors for different frequency subband. (a) Database 1. (b) Database 2. (c) Database 3.

for large bands. Choosing the band too small leads to a very large number of bands with a more and more difficult unwrapping problem at high frequencies.

3 SIMULATIONS USING HRTF MODELS AND MEASURED DATA

In this section simulation results using HRTF models as well as experimental measurements are presented. The technique presented in Section 2 is compared with other interpolation techniques. The three databases described previously in Section 2.3 are used in this section. All HRTFs are sampled at a temporal sampling frequency of 44.1 kHz. The purpose of this section is not to provide a comparison of all existing techniques. Nevertheless, four methods are compared and further discussed. The typical setup for HRTF interpolation happens as follows. From N available HRTF measurements it is the purpose to obtain after interpolation mN HRTFs, where m stands for the interpolation factor. In other words, $m - 1$ new measurements are obtained between two consecutive original HRTFs.

- Method 1 considers an interpolation of each time sample using a periodic sinc interpolation along the spatial direction.
- Method 2 corresponds to the new technique presented in Section 2.
- Method 3 corresponds to the technique presented in [20]. There the HRTFs are first time-aligned. The alignment is obtained by finding the maximum of the cross-correlation function between the HRTFs considered. In order to have subsample precision in the time-delay estimation, the HRTFs are interpolated in the time domain prior to the cross-correlation estimation. Once aligned, the HRTFs are interpolated using standard interpolation. Similar results are obtained in the case of linear interpolation or low-pass interpolation using first neighbors.
- Method 4 is a new technique presented in this study, based on methods 1 and 3. The low-frequency content of the HRTFs satisfying Eq. (34) is interpolated using a sinc interpolator as in method 1. The higher frequency content is then interpolated using method 3.

To compare the performance of the different interpolation algorithms, the error obtained by interpolation of the HRTFs is studied either as a function of the different angular positions or as a function of the temporal frequency. The mean squared error (MSE) on an interpolated HRTF at one position is calculated as follows. Calling $h(\theta, n)$ the original discrete-time HRTF and $h_e(\theta, n)$ its estimated version, the MSE at an angular position θ_0 is defined as

$$\text{MSE}(\theta_0) = 10 \log_{10} \frac{\sum_{n=0}^T [h(\theta_0, n) - h_e(\theta_0, n)]^2}{\sum_{n=0}^T h^2(\theta_0, n)} \quad (52)$$

where T stands for the number of time samples of the HRTF. When considering the frequency-dependent error, the MSE is averaged over all spatial positions and studied as a function of the temporal frequency. We therefore introduce the Fourier transforms with respect to time of $h(\theta, n)$ and $h_e(\theta, n)$, denoted by $\tilde{h}(\theta, f)$ and $\tilde{h}_e(\theta, f)$, respectively. The variable f stands here for the Fourier bin. The frequency-dependent MSE is now defined at one temporal frequency bin f_0 as

$$\text{MSE}(f_0) = 10 \log_{10} \frac{\sum_{\theta=0}^N |\tilde{h}(\theta, f_0) - \tilde{h}_e(\theta, f_0)|^2}{\sum_{\theta=0}^N |\tilde{h}(\theta, f_0)|^2}. \quad (53)$$

In Section 1 considerable effort was spent to study the angular bandwidth of the sound field along a circular array. It was shown that using the sinc interpolation is recommended because of the almost band-limited character of the frequency support. Therefore the sinc interpolation is described in detail in Section 3.1.

3.1 Sinc Interpolation

In this section we describe results obtained using the sinc interpolator. Note that due to the periodic character of the sound field, the sinc is a “periodic” sinc [43]. The implementation of the sinc interpolator is done by zero padding in the frequency domain, as mentioned previously in Section 1.3.1.

As was explained in Section 1, applying sinc interpolation only makes sense when Eq. (34) is satisfied. This relation can be verified by observing the frequency-dependent error on the interpolated HRTFs. Fig. 16(a) presents the frequency-dependent error on HRTFs averaged over all interpolated angular positions. An interpolation of a factor of 2 was applied in the cases of 20° (curves A and B) and 10° (curves C and D) angular spacing between consecutive HRTFs. Curves A and C (solid lines) were obtained using database 1 and curves B and D (dotted lines) using database 2. It can be observed that the curves considering head shadowing are very close to the simpler model without considering head shadowing. Therefore it can be concluded that taking this effect into account does not modify the average error significantly. Fig. 16(b) presents the same results for database 3. The interpolation error when using HRTFs every 10° is shown as a solid line and as a dotted line for an angular spacing of 20° . In both charts the two vertical lines correspond to the maximum values of the temporal frequencies, corresponding to angular samplings of 10° and 20° as given by Eq. (34). Fig. 16 allows us to conclude that interpolating HRTFs for higher frequencies than the ones predicted by Eq. (34) leads to large errors, whereas the interpolation error stays limited when obeying Eq. (34).

Consider now 36 measurements spaced every 10° . These measurements are used to interpolate HRTFs every 5° . The interpolated HRTFs are then compared with the corresponding measurements, which are available in the

database, and the normalized MSE is calculated following Eq. (52).

Using a spacing of 10° , interpolation of the HRTFs can only be carried out correctly up to a maximum temporal frequency of 10.8 kHz, as follows from Eq. (34). Therefore prior to any interpolation the HRTFs (impulse responses) are low-pass filtered using a low-pass filter with a cutoff frequency corresponding to the maximum temporal frequency associated with the specific angular spacing.

The MSE of the interpolation applied to database 1 is shown by a solid line in Fig. 17. The same simulation for the MSE of database 2 is shown by a dashed line, and by a solid line with + signs for database 3. Databases 1 and 2 exhibit an MSE varying from -35 to -65 dB. The best interpolation is obtained at positions in the proximity of -90° and 90° whereas slightly worse interpolation is

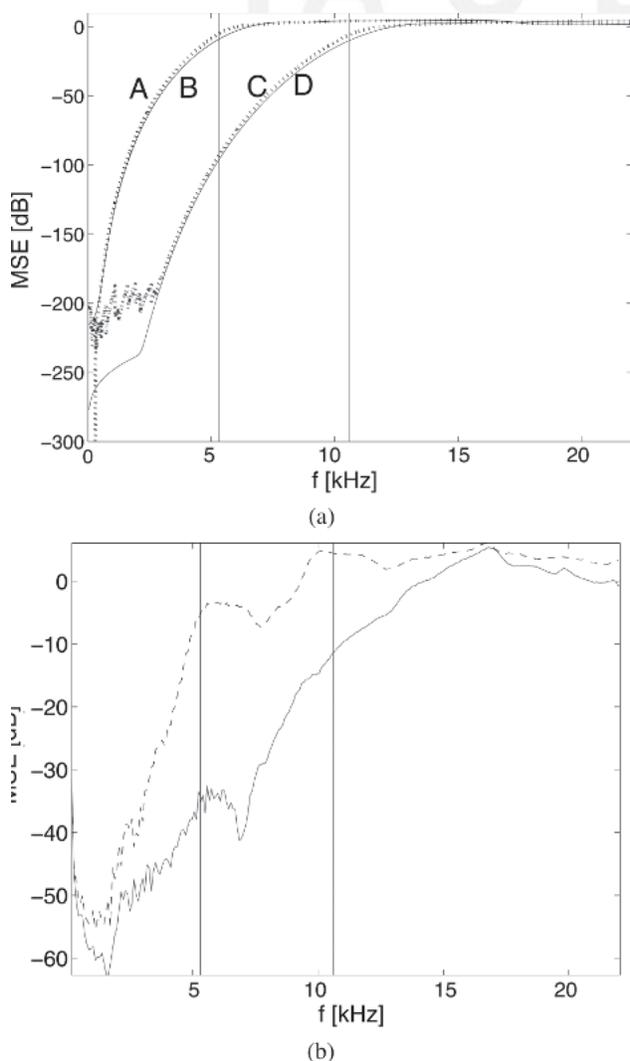


Fig. 16. Frequency-dependent MSE for interpolation of HRTF positions. (a) Simulations without considering head (\cdots , curves A and C) and with spherical head model (— , curves B and D). Curves A and B are obtained with an angular spacing of 20° , curves C and D with 10° . (b) Measured HRTFs for angular spacings of 10° (—) and 20° (---). Vertical lines correspond to maximum values of temporal frequencies corresponding to angular samplings of 10° and 20° as given by Eq. (34).

achieved in the proximity of 0 and 180° [the angles are referenced in Fig. 9(b)]. The MSE in the interpolation of database 3 was slightly higher than in the simulations, but still on the order of -40 dB.

These good interpolation results show that the sinc interpolation is very suitable to be applied when little spatial aliasing is present or when Eq. (34) is satisfied.

3.2 Comparison of Interpolation Methods

In this section methods 1 to 4 are compared and discussed. The different methods are compared in the MSE sense, and the frequency-dependent error is also shown for each. The databases used in this section contain HRTFs sampled every 5° . From these databases we use HRTFs measured every 10° and reconstruct the data set every 5° by means of interpolation. These interpolated HRTFs are then compared with the original HRTFs. Note that the following results are presented for the HRTFs recorded at the left ear. Therefore for sources located in the region around -90° the ear is the ipsilateral ear, whereas it is the contralateral ear for sources placed in the region around $+90^\circ$. The performance of the different algorithms is only shown for the left ear since very similar results are obtained when considering the right ear.

3.2.1 Database 1

Database 1 considers the solution of the wave equation in three dimensions along a circular array, as discussed in Section 1.2. The four methods have been tested and the results are presented in Fig. 18. In Fig. 18(a) it can be seen that method 1 only performs well around $\pm 90^\circ$. At the other positions, due to spatial aliasing, this technique leads to very poor results. Fig. 18(b) shows that most of the error resulting from interpolation is due to the bad interpolation of the high frequencies. Note that to improve the clarity of the results, the MSE is only shown here for values larger than -120 dB.

Method 2 performs best. The mean error is on the order of -48 dB, as can be seen from Fig. 18(a). Since for database 1 the HRTFs are modeled as linear-phase filters, the problem of unwrapping can be well addressed. There-

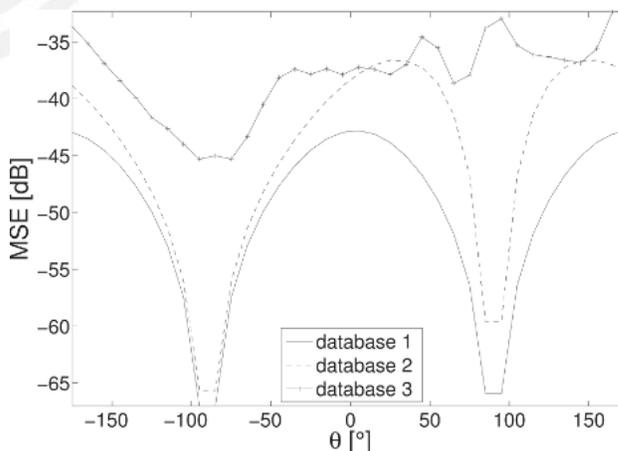


Fig. 17. MSE on reconstruction of HRTF databases 1, 2, and 3 when using sinc interpolation on low-pass signals.

fore we choose to work with narrow subbands (200 Hz). The results obtained show MSE below -40 dB at almost all angular positions. In Fig. 18(b) we can see that the low-frequency part of the spectrum is well reconstructed, but the main difference when compared to method 1 lies in the fact that the high-frequency content of the HRTFs is well interpolated. The mean error for the higher frequencies is below -60 dB.

Method 3 aligns the different signals before interpolation, using either a simple linear interpolator or a four-tap symmetric low-pass filter. In both cases the mean error is about -30 dB. The frequency-dependent error is shown in full lines in Fig. 18(b). It can be seen that the error is larger than the one presented in method 2, but it is nevertheless quite acceptable.

Method 4 considers the previous method, but the low-frequency part of the HRTFs is now interpolated using a sinc. It can be observed that this technique brings only little improvement. This is due to the fact that the main source of error is to be found in the higher frequencies, which are similarly interpolated for methods 3 and 4. In Fig. 18(b) the frequency-dependent error is presented by dashed bold lines.

To conclude this study for database 1, it can be observed that the best interpolation scheme is given by method 2.

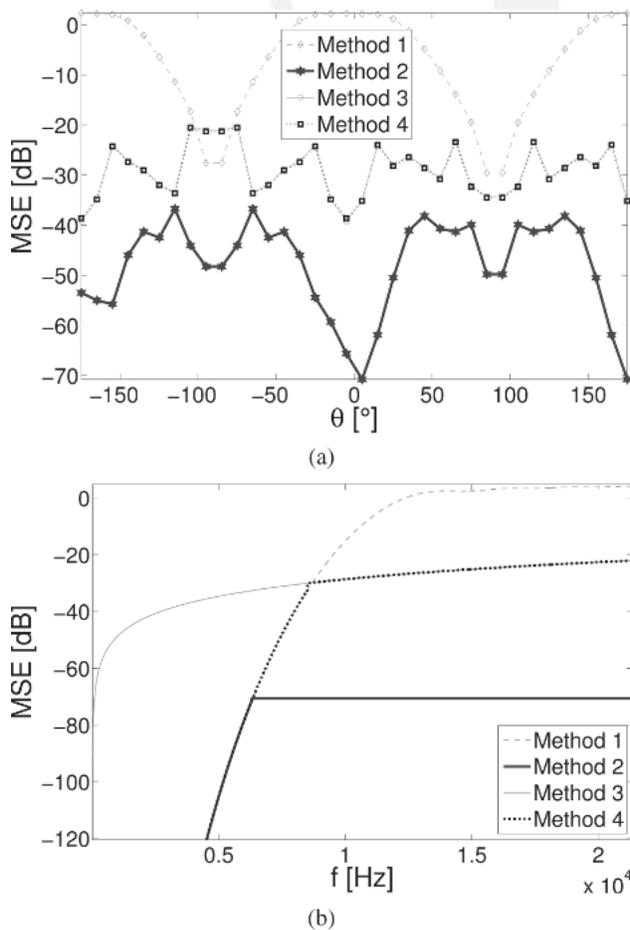


Fig. 18. Comparison of HRTF interpolation algorithms for database 1. (a) MSE for different methods. (b) Frequency-dependent error.

3.2.2 Database 2

Database 2 considers the model by Duda et. al. [41] described in Section 1.3.2. A comparison between the methods is shown in Fig. 19.

Method 1 gives similar results as for database 1. The interpolation results are only acceptable in the region of $\pm 90^\circ$. The frequency-dependent error is very large for increasing temporal frequencies. To improve the clarity of Fig. 19(b), only MSEs larger than -100 dB are presented.

Method 2 gives the best overall results with an MSE over the different positions of -33 dB. In the region around 90° the results are worse. This is due to the fact that this region puts in evidence the head-shadowing effect. The HRTFs are varying faster in this region, which makes them more difficult to interpolate. The frequency-dependent error shows an improvement compared to the sinc interpolation, and the mean error for high temporal frequency stays limited to -30 dB. For this method subbands of 500 Hz were chosen, as shown by vertical lines in Fig. 19(b). Note that the first subband is much larger since the low frequencies satisfying Eq. (34) are interpolated using the sinc interpolation.

Method 3 presents similar results as for database 1. The MSE averaged over all angular positions is about -20 dB.

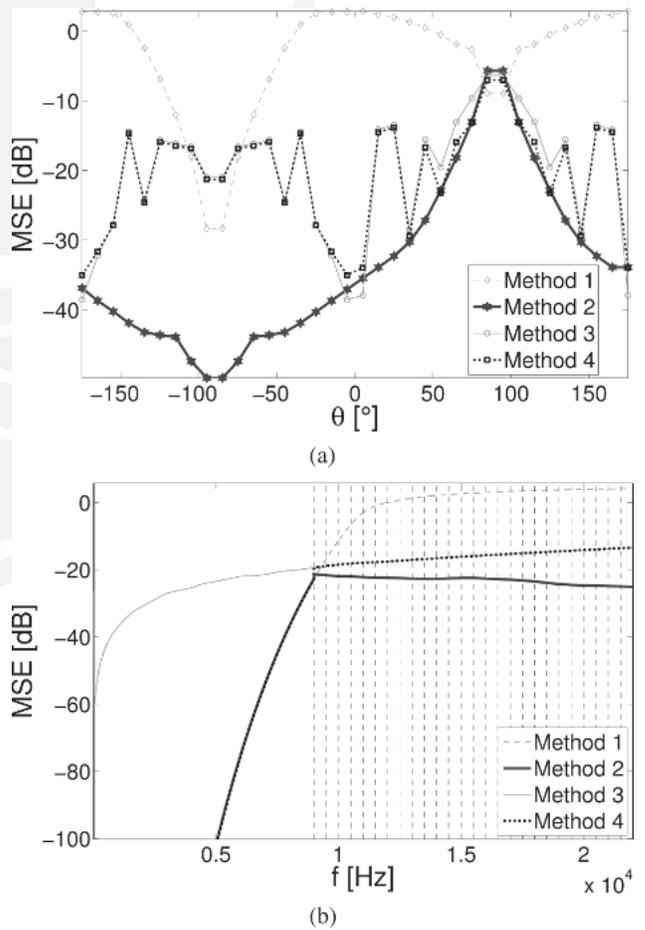


Fig. 19. Comparison of HRTF interpolation algorithms for database 2. (a) MSE for different methods. (b) Frequency-dependent error. Subbands used in method 2 are shown as vertical dashed lines.

Note that the MSE gets slightly worse in the region around -90° .

Method 4 is very comparable to method 3, and the overall MSE is about -21 dB when averaged over all angular positions.

Database 2 presents a more realistic model for HRTF interpolation. It is shown that method 2 delivers the best results. Around the region of -90° the HRTFs are interpolated with an MSE of less than -40 dB. Nevertheless, the region around 90° suffers poor interpolation, comparable to methods 3 and 4.

3.2.3 Database 3

Database 3 considers real measurements of HRTFs performed on a Kemar head. The results of the comparison are given in Fig. 20.

Method 1 presents a low MSE in the region of $\pm 90^\circ$. The other angular positions are more poorly interpolated. The average MSE over the different angular positions is on the order of -16 dB. The frequency-dependent error is shown in Fig. 20(b) to be large for high frequencies.

Method 2 achieves the lowest average MSE of -30 dB. The best interpolation results are to be found in the region of -90° . The region around 90° is less well interpolated

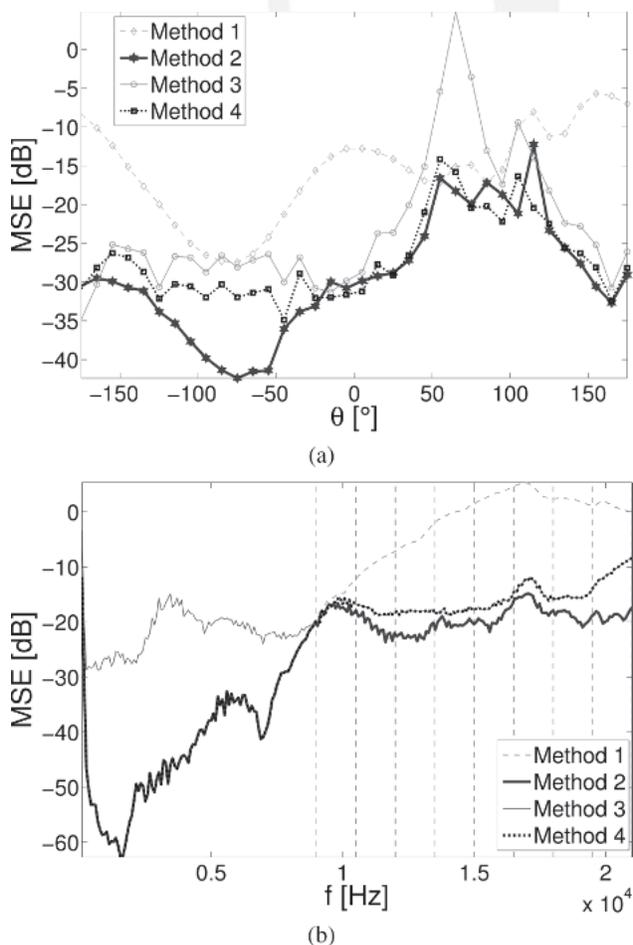


Fig. 20. Comparison of HRTF interpolation algorithms for database 3. (a) MSE for different methods. (b) Frequency-dependent error. Subbands used in method 2 are shown as vertical dashed lines.

with an MSE of about -15 dB. Nevertheless the frequency-dependent error is the lowest among the different methods, with a frequency-dependent error below -20 dB for higher temporal frequencies. For these measurements subbands of 1500 Hz were chosen, as shown by vertical dashed lines in Fig. 20(b). These subbands are wider than the ones used for the other databases. The real measurements used in database 3 lead to larger unwrapping errors when using a large number of small subbands. The choice of the subband width is based on a tradeoff between unwrapping error and correct alignment of the carriers, as described previously. In this case the optimum subband was obtained by sweeping over different widths, and an optimum situation was found for subbands of around 1.5 kHz. Note here also that the first subband is much larger due to the sinc interpolation applied for the low frequencies satisfying Eq. (34).

Method 3 delivers interpolation results that are good except in the region of the head shadowing, where the MSE increases a lot. As can be seen from Fig. 20(b), the frequency-dependent error is quite constant over all temporal frequencies. The low frequencies are quite poorly interpolated, which explains the large errors in the region of 90° .

In this scenario, using the sinc interpolation for the low frequencies allows method 4 to achieve quite good results, similar to those of method 2. The region around 90° is a bit less well interpolated, but overall the HRTFs are interpolated with an average MSE of -27 dB.

Comparing the different methods leads to the conclusion that method 2 performs slightly better than method 4, especially in the region around -90° . For the other regions similar results are obtained. Nevertheless, this is an interesting finding since method 4 is quite simple and still leads to good interpolation results.

For completeness we have considered the case of HRTF interpolation every 20° to obtain HRTFs every 10° . The MSE is given in Fig. 21 for the four methods considered. The four methods behave very similarly to the case where the HRTFs are interpolated from a spacing of 10° . The best total MSE is obtained for method 2 with an average

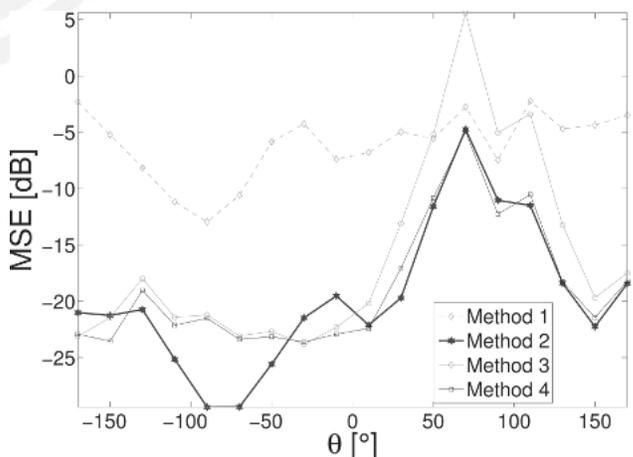


Fig. 21. Comparison of HRTF interpolation methods for spacing of 20° in database.

MSE of -20 dB. Method 4, which is simpler in processing, achieves a slightly worse MSE of -18.4 dB. Method 1 performs quite poorly, with an MSE not exceeding -7 dB. Method 3 shows a similar MSE than method 4, except for the region around 90° , where it performs badly.

3.3 Discussion

We have presented an overview of different databases tested with different interpolation methods. To summarize the results, Table 1 contains the MSE averaged over all positions interpolated for each database and each method. Note that database 3b considers database 3 where the HRTFs are present only every 20° , as described in Fig. 21.

Method 2 performs best of the different methods. In the measurements it has been shown that the simpler method 4 also performs quite well. The general trend present in these results shows that for sources that see the ear as the contralateral ear (source located on the other side of the ear, such as $+90^\circ$ for the left ear) interpolation is worse than in the other angular positions. On the other hand it is known that the head-shadowing effect creates an attenuation of approximately 20 dB on the perceived sound [7]. When localizing a sound, the brain uses the signals of the two ears. For a source located at $+90^\circ$, the HRTF from the source to the left ear has a low magnitude and contains the head-shadowing effect whereas the right ear gets a very direct HRTF with higher magnitude. When interpolation of HRTFs is applied, it is likely that the localization of the source in this scenario is still achievable, even when interpolation of the weak signal (left ear) is poor while being very precise for the strong signal (right ear). This topic is a matter of current research, and perceptual tests need to be conducted to study this scenario in more detail.

4 CONCLUSIONS

The sound field has been studied along a circle. The angular bandwidth of the sound has been studied in two and three dimensions and a sampling theorem has been presented to quantify the aliasing error as a function of the angular sampling frequency. Based on this study, an HRTF interpolation algorithm has been proposed. At low frequencies, where the spatial Nyquist theorem indicates that the given HRTFs can be interpolated very precisely, spatial interpolation on circles is applied. At higher frequencies, where the Nyquist theorem indicates that not enough information for precise interpolation is available, the spatial interpolation is carried out in the complex temporal envelope domain in subbands to avoid aliasing. The

Table 1. Summary of MSEs averaged over all interpolated positions for different methods and databases.

	Method 1	Method 2	Method 3	Method 4
Database 1	-5.5	-47.9	-29	-29.1
Database 2	-3.5	-33.64	-20.3	-20.7
Database 3	-15.65	-29.5	-23	-27.2
Database 3b	-6.2	-19.7	-16	-18.4

interpolated subbands are obtained by restoring the carrier after interpolating the complex envelopes. Numerical simulations carried out with HRTF models and measured data indicate that the proposed method performs better than previous HRTF interpolation methods in a mean square sense.

5 REFERENCES

- [1] T. Ajdler, C. Faller, L. Sbaiz, and M. Vetterli, "Interpolation of Head-Related Transfer Functions Considering Acoustics," presented at the 118th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 53, p. 662 (2005 July/Aug.), convention paper 6327.
- [2] C. A. Balanis, *Antenna Theory: Analysis and Designs* (Wiley, New York, 2005).
- [3] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
- [4] E. Hulsebos, T. Schuurmans, D. de Vries, and R. Boone, "Circular Microphone Array for Discrete Multi-channel Audio Recording," presented at the 114th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 51, p. 405 (2003 May), convention paper 5716.
- [5] S. Petrusch, S. Spors, and R. Rabenstein, "Simulation and Visualization of Room Compensation for Wave Field Synthesis with the Functional Transformation Method," presented at the 119th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 53, p. 1207 (2005 Dec.), convention paper 6547.
- [6] M. Poletti, "Effect of Noise and Transducer Variability on the Performance of Circular Microphone Arrays," *J. Audio Eng. Soc.*, vol. 53, pp. 371–384 (2005 May).
- [7] J. Blauert, *Spatial Hearing* (MIT Press, Cambridge, MA, 2001).
- [8] H. Jones, "On Multipath Spatial Diversity in Wireless Multiuser Communications," Ph.D. dissertation, Australian National University, Canberra, Australia (2001).
- [9] H. M. Jones, R. Kennedy, and T. Abhayapala, "On Dimensionality of Multipath Fields: Spatial Extent and Richness," in *Proc. Int. Conf. on Acoustics, Sensor and Signal Processing (ICASSP'02)*, (2002), pp. 2837–2840.
- [10] S. Spors, "Active Listening Room Compensation for Spatial Sound Reproduction Systems," Ph.D. dissertation, Universität Erlangen-Nürnberg, Erlangen, Germany (2005).
- [11] T. Ajdler, L. Sbaiz, and M. Vetterli, "The Plenacoustic Function and Its Sampling," *IEEE Trans. Signal Process.*, vol. 54, pp. 3790–3804 (2006).
- [12] M. Poletti, "A Unified Theory of Horizontal Holographic Sound Systems," *J. Audio Eng. Soc.*, vol. 48, pp. 1155–1182 (2000 Dec.).
- [13] S. Spors and R. Rabenstein, "Spatial Aliasing Artifacts Produced by Linear and Circular Loudspeaker Ar-

- rays Used for Wave Field Synthesis,” presented at the 120th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 54, p. 689 (2006 July/Aug.), convention paper 6716.
- [14] E. Hulsebos, D. de Vries, and E. Bourdillat, “Improved Microphone Array Configurations for Auralization of Sound Fields by Wave Field Synthesis,” presented at the 110th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 49, p. 532 (2001 June), convention paper 5337.
- [15] E. Hulsebos and D. de Vries, “Parameterization and Reproduction of Concert Hall Acoustics Measured with a Circular Microphone Array,” presented at the 112th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 50, p. 518 (2002 June), convention paper 5579.
- [16] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia* (Academic Press, Cambridge, MA, 1994).
- [17] A. Kulkarni, S. Isabelle, and H. Colburn, “On the Minimum-Phase Approximation of Head-Related Transfer Functions,” in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics* (New Paltz, NY, 1995).
- [18] D. Kistler and F. L. Wightman, “A Model of Head-Related Transfer Functions Based on Principal Components Analysis and Minimum-Phase Reconstruction,” *J. Acoust. Soc. Am.*, vol. 91, pp. 1637–1647 (1992).
- [19] P. Minaar, 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.).
- [20] M. Matsumoto, S. Yamanaka, M. Tohyama, and H. Nomura, “Effect of Time Arrival Correction on the Accuracy of Binaural Room Impulse Response Interpolation,” *J. Audio Eng. Soc. (Engineering Reports)*, vol. 52, pp. 56–61 (2004 Jan./Feb.).
- [21] F. Christensen, H. Møller, P. Minnaar, J. Plogsties, and S. K. Olesen, “Interpolating between Head-Related Transfer Functions Measured with Low-Directional Resolution,” presented at the 107th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 47, p. 1009 (1999 Nov.), preprint 5047.
- [22] K. Hartung, J. Braasch, and S. J. Sterbing, “Comparison of Different Methods for the Interpolation of Head-Related Transfer Functions,” in *Proc. AES 16th Int. Conf. on Spatial Sound Reproduction* (Rovaniemi, Finland, 1999), pp. 319–329.
- [23] S. M. Robeson, “Spherical Methods for Spatial Interpolation: Review and Evaluation,” *Cartogr. and Geogr. Inf. Sys.*, vol. 24, no. 1, pp. 3–20 (1997).
- [24] S. Carlile, C. Jin, and V. Van Raad, “Continuous Virtual Auditory Space Using HRTF Interpolation: Acoustic and Psychophysical Errors,” in *Proc. Int. Symp. on Multimedia Information Processing* (2000).
- [25] V. Larcher, J. M. Jot, J. Guyard, and O. Warusfel, “Study and Comparison of Efficient Methods for 3D Audio Spatialization Based on Linear Decomposition of HRTF Data,” presented at the 108th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 48, p. 351 (2000 Apr.), preprint 5097.
- [26] J. Chen, B. Van Veen, and K. Hecox, “A Spatial Feature Extraction and Regularization Model for the Head-Related Transfer Function,” *J. Acoust. Soc. Am.*, vol. 97, pp. 439–452 (1995).
- [27] M. Blommer and G. Wakefield, “Pole-Zero Approximations for HRTF Using a Logarithmic Error Criterion,” *IEEE Trans. Speech Audio Process.*, vol. 5, pp. 278–287 (1997).
- [28] A. Kulkarni and H. Colburn, “Infinite-Impulse Response Models of the Head-Related Transfer Function,” *J. Acoust. Soc. Am.*, vol. 115, pp. 1714–1728 (2004).
- [29] P. R. Runkle, M. A. Blommer, and G. H. Wakefield, “Comparison of Head-Related Transfer Function Interpolation Methods,” in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics* (New Paltz, NY, 1995).
- [30] R. Duraiswami, D. N. Zotkin, and N. Gumerov, “Interpolation and Range Extrapolation of HRTFs,” in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, (2004), pp. 45–48.
- [31] P. Morse and K. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
- [32] N. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions* (Elsevier Science, Amsterdam, Netherlands, 2005).
- [33] V. R. Algazi, R. O. Duda, and D. M. Thompson, “Motion-Tracked Binaural Sound,” *J. Audio Eng. Soc.*, vol. 52, pp. 1142–1156 (2004 Nov.).
- [34] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory* (Springer-Verlag, Berlin/Heidelberg, Germany, 1992).
- [35] E. G. Williams, *Fourier Acoustics* (Academic Press, San Diego, CA, 1999).
- [36] T. Ajdler, L. Sbaiz, and M. Vetterli, “The Plenacoustic Function on the Circle with Application to HRTF Interpolation,” in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, (2005), pp. 273–276.
- [37] A. Carlson, *Communication Systems* (McGraw-Hill, New York, 1986).
- [38] M. Do and M. Vetterli, “On the Bandlimitedness of the Plenoptic Function,” in *Proc. IEEE Int. Conf. on Image Processing* (Genoa, Italy, 2005).
- [39] S. Müller and P. Massarani, “Transfer-Function Measurement with Sweeps,” *J. Audio Eng. Soc.*, vol. 49, pp. 443–471 (2001 June).
- [40] <http://lcavwww.epfl.ch/~thibaut/Measurements/>; last viewed 2008 Jan. 30.
- [41] R. O. Duda and W. Martens, “Range Dependence of the Response of a Spherical Head Model,” *J. Acoust. Soc. Am.*, vol. 104, pp. 3048–3058 (1998).
- [42] B. Gardner and K. Martin, “HRTF Measurements of a Kemar Dummy-Head Microphone,” Tech. Rep., MIT Media Lab. (1994).
- [43] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1999).

THE AUTHORS



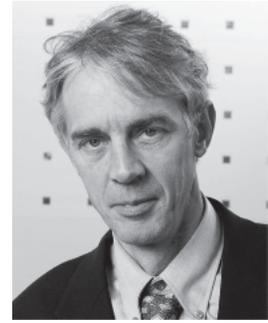
T. Ajdler



C. Faller



L. Sbaiz



M. Vetterli

Thibaut Ajdler received B.Sc. and M.Sc. degrees in electrical engineering from the Free University of Brussels (Vrije Universiteit Brussel), Belgium, in 2001 and a Ph.D. degree from Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 2006. His thesis, entitled “The Plenacoustic Function and Its Applications,” presents a study on the sampling of the sound field.

From 2000 July to September he was a research assistant at the Massachusetts Institute of Technology, Cambridge, MA, and from 2002 October to 2003 February he was an intern at Intel Corporation, Santa Clara, CA, working on acoustic source localization using distributed sensor networks. His research interests include audio signal processing, acoustics, and signal processing for communications.

Christof Faller received an M.S. (Ing.) degree in electrical engineering from ETH Zurich, Switzerland, in 2000, and a Ph.D. degree for his work on parametric multichannel audio coding from EPFL, Lausanne, Switzerland, in 2004.

From 2000 to 2004 he worked in the Speech and Acoustics Research Department at Bell Laboratories, Lucent Technologies and Agere Systems (a Lucent company), where he worked on audio coding for digital satellite radio, including parametric multichannel audio coding. He is currently a part-time postdoctoral employee at EPFL Lausanne. In 2006 he founded Illusonic LLC, an audio and acoustics research company.

Dr. Faller has won a number of awards for his contributions to spatial audio coding, MP3 Surround, and MPEG Surround. His main current research interests are spatial hearing and spatial sound capture, processing, and reproduction.

Luciano Sbaiz received a Laurea in Ingegneria degree in electronic engineering from the University of Padova,

Padova, Italy, in 1993 and a Ph.D. degree, also from the University of Padova, in 1998. Between 1998 and 1999 he was a postdoctoral researcher at the Audiovisual Communications Laboratory at EPFL, Lausanne, Switzerland, where he conducted research on the application of computer vision techniques to the creation of video special effects.

In 1999 he joined Dartfish Ltd., Fribourg, Switzerland, as project manager. Within the company he developed video special effects for television broadcasting and sport analysis. In 2004 he took a position as senior researcher at the Audiovisual Communications Laboratory at EPFL, where he conducts research on signal processing. His activities are in the fields of image and audio processing, superresolution techniques, and acoustics.

Martin Vetterli received his engineering degree from ETH, Zurich, Switzerland, an M.S. degree from Stanford University, Stanford, CA, and a Ph.D. degree from EPFL, Lausanne, Switzerland.

In 1986 he joined Columbia University, New York, first with the Center for Telecommunications Research and then with the Department of Electrical Engineering, where he was an associate professor of electrical engineering. In 1993 he joined the University of California at Berkeley, where he was a full professor until 1997. Since 1995 he has been a professor at EPFL, where he headed the Communication Systems Division (1996/1997) and now is heading the Audiovisual Communications Laboratory. From 2001 to 2004 he directed the National Competence Center in Research on mobile information and communication systems. He is also vice-president for International Affairs at EPFL. He has held visiting positions at ETHZ (1990) and Stanford (1998). His research interests are in the areas of applied mathematics, signal processing, and communications. He is the coauthor of the textbook *Wavelets and Subband Coding* and of over 100 journal papers.