

Research Article

Extraction of 3D Information from Circular Array Measurements for Auralization with Wave Field Synthesis

Diemer de Vries, Lars Hörchens, and Peter Grond

Laboratory of Acoustical Imaging and Sound Control, Department of Image Science and Technology, Faculty of Applied Sciences, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

Received 29 April 2006; Revised 3 October 2006; Accepted 8 February 2007

Recommended by Ville Pulkki

The state of the art of wave field synthesis (WFS) systems is that they can reproduce sound sources and secondary (mirror image) sources with natural spaciousness in a horizontal plane, and thus perform satisfactory 2D auralization of an enclosed space, based on multitrace impulse response data measured or simulated along a 2D microphone array. However, waves propagating with a nonzero elevation angle are also reproduced in the horizontal plane, which is neither physically nor perceptually correct. In most listening environments to be auralized, the floor is highly absorptive since it is covered with upholstered seats, occupied during performances by a well-dressed audience. A first-order ceiling reflection, reaching the floor directly or via a wall, will be severely damped and will not play a significant role in the room response anymore. This means that a spatially correct WFS reproduction of first-order ceiling reflections, by means of a loudspeaker array at the ceiling of the auralization reproduction room, is necessary and probably sufficient to create the desired 3D spatial perception. To determine the driving signals for the loudspeakers in the ceiling array, it is necessary to identify the relevant ceiling reflection(s) in the multichannel impulse response data and separate those events from the data set. Two methods are examined to identify, separate, and reproduce the relevant reflections: application of the Radon transform, and decomposition of the data into cylindrical harmonics. Application to synthesized and measured data shows that both methods in principle are able to identify, separate, and reproduce the relevant events.

Copyright © 2007 Diemer de Vries et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The traditional reproduction formats of audio reproduction, like two-channel stereo and Dolby surround [1], have the restriction that they provide correct spatial information only in a limited listening area, the so-called “sweet spot.” With the introduction of wave field synthesis (WFS) by Berkhout [2], it became possible to generate sound fields with natural spatial properties within an extended volume or area bounded by arrays of loudspeakers. The “ideal” WFS reproduction room would be a 3D space of proper dimensions, all boundaries of which are covered with closely sampled arrays of individually driven loudspeakers. Then, the acoustic wave field in any enclosed space could be simulated or reproduced and listeners within that space could walk around and perceive the acoustic conditions at any place with correct temporal and spatial properties. The reproduction of the acous-

tics of a hall in a different environment is called auralization in this context, as an extension of the definition of this term by Kleiner et al. [3].

However, from viewpoints such as accessibility and computational power, it is not realizable to cover all the boundaries with loudspeakers. Therefore, present WFS-based auralization systems consist of a loudspeaker array configuration in a horizontal plane, roughly at the elevation of the ears of the listeners. Typical input data for the auralization are multi-trace impulse responses measured [4] or simulated [5] along a 1D (linear) or 2D (cross-shaped circular) microphone array in a horizontal plane. By this way, wave components incident with an elevation angle unequal to zero are “projected” in the horizontal plane as will be further explained in the next section. The mirror image sources that represent the boundary reflections are reproduced in the horizontal plane according to the recorded arrival times, which

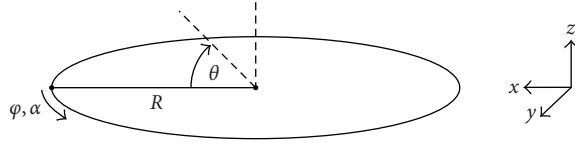


FIGURE 1: Geometry and variables of the circular array.

correspond to nonhorizontal travel paths. A 2D listening space is simulated with most boundaries at wrong positions, which might be perceived by the listener as missing the “hall volume impression.”

Since the sensitivity of the human ear to vertical localization is not very high [6, 7], some inaccuracy in the reproduction is permitted from a perceptual point of view. As the floor of a concert hall or theatre is usually covered with absorptive seats or well-dressed listeners, it can be assumed that sound waves, after being reflected by this floor surface, do not play a significant role in a measured impulse response anymore. Therefore, the first-order reflections from the ceiling and second-order reflections via side walls and ceiling should be the most important nonhorizontal contributions to the sound field. As a first attempt to create a realistic perceptual “volume” impression of a 3D hall, the TU Delft WFS system is being extended with a ceiling array to control the elevation angle of the significant ceiling reflections of the hall to be auralized. For that purpose, these reflections should be identified and selected from the dataset measured with (or simulated along) the microphone array in the horizontal plane. In this paper, data simulated and measured along a circular microphone array are considered, since such arrays have proven to be most efficient for auralization purposes [4].

Two different identification and selection methods are examined. The first method is based on application of the Radon transform. The second method decomposes the data into cylindrical harmonics. Both methods aim to transform the array data set from the “data space” to a “model space” in such a way that interfering events in the data space appear as distinct points in the model space. By this way, it will be easy to identify different events and separate them from each other. The last important step before the actual auralization is the reproduction of the separated events in data space, requiring a correct inverse transformation from model space to data space.

2. 3D INFORMATION IN CIRCULAR ARRAY DATA

The coordinates, variables, and geometry used in the following are specified in Figure 1. A circular microphone array with radius R is positioned in the horizontal (x, y) plane. Plane waves are considered incident on the array with azimuth angle φ and elevation angle θ (in degrees). Azimuth angle φ is relative to some reference point, chosen here ($x = -R, y = 0, z = 0$). Elevation angle θ is relative to the horizontal (x, y) plane. A microphone position on the array is specified by an angle α , also given in degrees relative to the reference point.

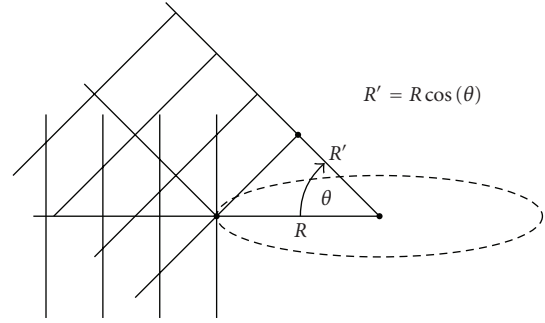


FIGURE 2: Geometry of plane waves incident on a circular array in the horizontal plane.

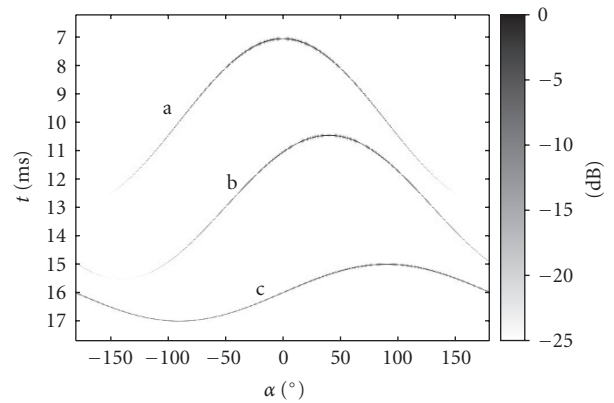


FIGURE 3: Arrival times as a function of microphone position on a circular array for three plane waves with (a) $\varphi = 0^\circ, \theta = 0^\circ, \tau = 10$ milliseconds, (b) $\varphi = 40^\circ, \theta = 30^\circ, \tau = 13$ milliseconds, and (c) $\varphi = 90^\circ, \theta = 70^\circ, \tau = 16$ milliseconds (c) for an array with radius $R = 1$ m.

If a plane wave travels in the horizontal plane, that is, with elevation angle $\theta = 0^\circ$, and azimuth angle $\varphi = 0^\circ$, it first arrives at microphone position $\alpha = 0^\circ$, see Figure 2. The arrival times of the wave on the circular microphone array as a function of microphone position α are described by a cosine-shaped curve:¹ the “maximum” is found for $\alpha = 0^\circ$, the “minimum” for $\alpha = \pm 180^\circ$, and the “zero value” for $\alpha = \pm 90^\circ$. In the following, the “average” arrival time at which the center of the array is reached is called the intercept time, denoted as τ . The geometrical amplitude of the cosine-shaped arrival time curve is R/c , where c represents the sound velocity. The arrival time curve for this wave is shown in Figure 3(a).

When a plane wave arrives at the array under a nonzero elevation angle θ and azimuth angle $\varphi = 0^\circ$, it again arrives

¹ For an arbitrary signal, “arrival time” should be more precisely defined. In the simulations and measurements considered in this paper, short broadband signals have been used which, in this context, can be considered as Dirac pulses with well-defined arrival times with respect to their time of generation $t = 0$.

first at $\alpha = 0^\circ$ and last at $\alpha = \pm 180^\circ$. The arrival time curve is still cosine-shaped, but now its geometrical amplitude with respect to the intercept time decreases to R'/c with $R' = R \cos \theta$ as illustrated in Figure 2. The arrival time differences between $\alpha = 0^\circ$ and $\alpha = \pm 90^\circ$, and $\alpha = 0^\circ$ and $\alpha = \pm 180^\circ$ are now reduced to R'/c and $2R'/c$, respectively. When, as extreme case, a plane wave arrives at the horizontal array with vertical incidence ($\theta = 90^\circ$), it reaches all microphones at the same time and the arrival time curve is a straight line. It can be concluded that the geometrical amplitudes of the arrival time curves in impulse response data measured along a circular array in the horizontal plane contain information about the elevation angles of incidence of the plane wave components. However, due to the symmetry of the setup with respect to the horizontal plane, it is not possible to distinguish between plane waves arriving from below or from above the array. Since ceiling reflections always arrive from above, the range of interest for the elevation angle θ is restricted to $0^\circ < \theta < 90^\circ$.

Variation of the azimuth angle of incidence φ of the plane wave leads to a spatial phase shift of the cosine-shaped curve: the first arrival at the array will occur for a nonzero microphone position α . The general expression of the arrival time curves is

$$t(\alpha, \varphi, \theta) = \tau - \frac{R \cos \theta}{c} \cos(\varphi - \alpha). \quad (1)$$

In the present work, a circular array using outward pointing cardioid microphones has been used for simulations and measurements, as is usually done now in array-based room acoustics analysis [4]. Therefore, the strength S of the signal received at a microphone position α is weighted by a factor according to the cardioid characteristics of the microphones, depending on the angles φ and θ of the incident plane wave:

$$S(\alpha, \varphi, \theta) = \frac{P_0}{2} [1 + \cos(\varphi - \alpha) \cos(\theta)], \quad (2)$$

where P_0 represents the pressure of the incident plane wave.

A simulation of the arrival time curves of three plane waves with elevation angles $\theta = 0^\circ, 30^\circ$, and 70° is shown in Figure 3. The waves reach the array with azimuth angles $\varphi = 0^\circ, 40^\circ$, and 90° , respectively.

For waves traveling in the horizontal plane, the strength of the signal at microphone positions $\alpha = \varphi \pm 180^\circ$, that is, the ‘‘rear’’ of the array for that particular wave, tends towards zero (see curve a in Figure 3(a)). The decrease in signal strength is less for elevated plane waves, as shown in curves b and c. For waves with vertical incidence ($\theta = 90^\circ$), the strength of the recorded signal corresponds to half the pressure of the incident plane wave at all microphone positions, as is easily seen from (2).

3. SEPARATION AND RECONSTRUCTION

In the previous section, it was shown that the arrival time curves of plane waves in 2D array data contain 3D information. This information will be used to identify first-order ceiling reflections in circular array data sets measured in a concert hall, using the realistic assumption that such reflections

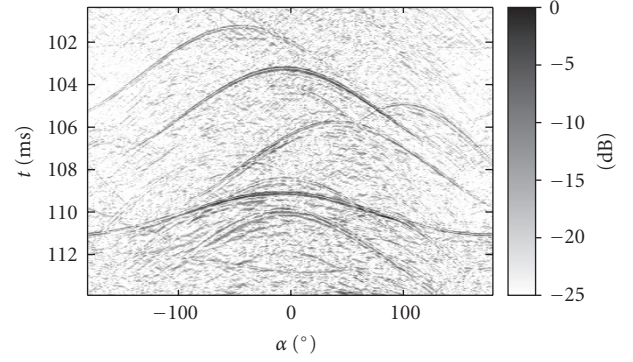


FIGURE 4: Multitrace impulse responses measured in the Frits Philips Concert Hall, Eindhoven, The Netherlands.

can be considered to be plane waves at the array position. Parts of such a data set, containing reflected waves arriving within 100 milliseconds and 114 milliseconds after generation of the impulsive test signal, are shown in Figure 4. The sound pressure of the reflected waves, measured with an outward pointing cardioid microphone traveling along a horizontal circle with a radius of 1 meter, is given as a function of travel time t (vertical axis (ms)) and microphone position α (horizontal axis, in degrees). This domain is called the *data space*. It is seen that the cosine-shaped curves representing the individual wave field components strongly interfere.

In order to make the identification of the ceiling reflection easier, the data is transformed to a *model space*, where each wave component is represented by a well-resolved event. After identification, the ceiling reflection is separated from the other components by filtering. Then it is reconstructed in the data space by inverse transformation, in a format that can be applied to drive the ceiling array of the loudspeaker configuration in the WFS reproduction room.

In the following, two methods will be investigated that might be applied in the above procedure: application of the Radon transform, and decomposition of the data into cylindrical harmonics, and from there, into plane waves.

4. THE RADON TRANSFORM

4.1. Identification

The Radon transform is widely used, especially in the field of seismic exploration, to detect components with a specific shape in complex 2D data sets [8, 9]. Data (in the present context: sound pressure data) are integrated along curves with that shape. When strong amplitudes are present along a certain curve, the integration will yield a high value, thus detecting that component and its position in the data space. The high integration value can be represented as a point in a 2D model space where the position parameters of the integration curve are the coordinates. Most integration curves can be written in the form

$$t = \tau + pf(\alpha), \quad (3)$$

where t is travel time, τ is the intercept time, that is, the time at which the center of the array is reached, $f(\alpha)$ is the function of the angular microphone position that defines the basic shape of the integration curve, and p is some slope or scaling parameter; the shape of the curve is fully specified by the product $pf(\alpha)$.

The Radon transform is usually given in terms of a forward transform from the data space $d(\alpha, t)$ to the model space $m(p, \tau)$ and the adjoint transform mapping from the model to the data space. These operators are not unitary in general. Therefore, the adjoint transform does not result in a perfect inversion of the forward transform: the successive application of the forward transform and its adjoint will not yield the original data in general. It is nevertheless possible to calculate an approximate inversion of either the forward or the adjoint transform. In the present work, the adjoint transform is chosen to be inverted. We therefore denote the result of the forward transform as approximation of the model space $\tilde{m}(p, \tau)$. In the final implementation, the simple forward transform is replaced by an inversion scheme, details of which are given further below.

The forward Radon transform describes a mapping from the data space $d(\alpha, t)$ to the model space $m(p, \tau)$:

$$\tilde{m}(p, \tau) = \int_{-\infty}^{\infty} d(\alpha, t = \tau + pf(\alpha)) d\alpha. \quad (4a)$$

The adjoint transformation is given by

$$d(\alpha, t) = \int_{-\infty}^{\infty} m(p, \tau = t - pf(\alpha)) dp. \quad (4b)$$

If, as in the context of this paper, the integration curves are time-invariant, the Radon transform can also be performed in the frequency domain, which appears to be faster than time-domain processing:

$$\tilde{M}(p, \omega) = \int_{-\infty}^{\infty} D(\alpha, \omega) e^{j\omega pf(\alpha)} d\alpha, \quad (5a)$$

$$D(\alpha, \omega) = \int_{-\infty}^{\infty} M(p, \omega) e^{-j\omega pf(\alpha)} dp. \quad (5b)$$

In the context of seismic exploration, the integration curves usually have the form of straight lines, parabolae or hyperbolae. In the present context, the wave components to be detected are found at curves in the data space which are shaped according to (1). We apply the transform for distinct values of the angle θ . Therefore, we obtain a Radon transform pair $D_j(\alpha, \omega)$, $M_j(\varphi, \omega)$ for each chosen elevation angle θ_j . We denote the effective radius of the array for the particular angle θ_j as $R_j = \cos \theta_j$.

Inserting (1) in (5a), (5b) then leads to

$$\tilde{M}_j(\varphi, \omega) = \int_{-\infty}^{\infty} D_j(\alpha, \omega) e^{-j\omega(R_j/c) \cos(\varphi - \alpha)} d\alpha, \quad (6a)$$

$$D_j(\alpha, \omega) = \int_{-\infty}^{\infty} M_j(\varphi, \omega) e^{j\omega(R_j/c) \cos(\varphi - \alpha)} d\varphi. \quad (6b)$$

These equations can be written in discretized form as

$$\tilde{M}_j(\varphi_m, \omega) = \sum_{i=0}^{N_\alpha} D_j(\alpha_i, \omega) e^{-j\omega(R_j/c) \cos(\varphi_m - \alpha_i)}, \quad (7a)$$

$$D_j(\alpha_i, \omega) = \sum_{m=0}^{N_\varphi} M_j(\varphi_m, \omega) e^{j\omega(R_j/c) \cos(\varphi_m - \alpha_i)}, \quad (7b)$$

or in vector/matrix form as

$$\tilde{\mathbf{m}}_j = \mathbf{L} \mathbf{d}_j, \quad (8a)$$

$$\mathbf{d}_j = \mathbf{L}^H \mathbf{m}_j. \quad (8b)$$

In order to optimize the correspondence between the result of the “inverse” Radon transform and the original data space, in the present work (8a) has been replaced by a high-resolution Radon transform as proposed by Sacchi and Ullrych [10]. This method employs a sparseness constraint on the model domain and uses a conjugate gradient algorithm in order to perform an approximate inversion of (8b) and to optimize the solution iteratively. By inverse Fourier transformation, the results of (8a), (8b) are brought back to the model space $m_j(\varphi, \tau)$ and the data space $d_j(\alpha, t)$, respectively.

As the Radon transform performs a summation along the curves in the data space, azimuthal strength information is lost in the model space. It is nevertheless possible to account for the angle-dependent signal strength due to the usage of cardioid microphones, by extension of the Radon transform presented above.

By integrating (2), it can easily be verified that the result of an integration of cardioid microphone signals along a circle is independent of the azimuth and elevation angle of an incident plane wave:

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{P_0}{2} [1 + \cos(\varphi - \alpha) \cos(\theta)] d\alpha = \frac{P_0}{2}. \quad (9)$$

The usage of cardioid microphones leads to a weighting of the model space by a factor 0.5. It is therefore sufficient to include the cardioid model when transforming from the model to the data space:

$$D_j(\alpha_i, \omega) = \sum_{m=0}^{N_\varphi} [1 + \cos(\varphi_m - \alpha_i) \cos(\theta_j)] \times M_j(\varphi_m, \omega) e^{j\omega(R_j/c) \cos(\varphi_m - \alpha_i)}. \quad (10)$$

Using a sparse inversion of (10) and a subsequent inverse Fourier transform, $m_j(\varphi_m, \tau)$ has been calculated for the simulated data set of Figure 3, for three discrete values of the elevation angle: 0° , 30° , 70° —corresponding to the elevation angles of incidence for the three waves in the data set. In addition, the transformation to the model space has been calculated for an elevation angle of 50° . Figure 5 shows the result. For each elevation angle considered, a 2D model space frame with (φ, τ) coordinates is given.

In the frames corresponding to plane waves with elevation angles actually present in the data set, the wave components with elevation angles $\theta = 0^\circ$, 30° , and 70° can be

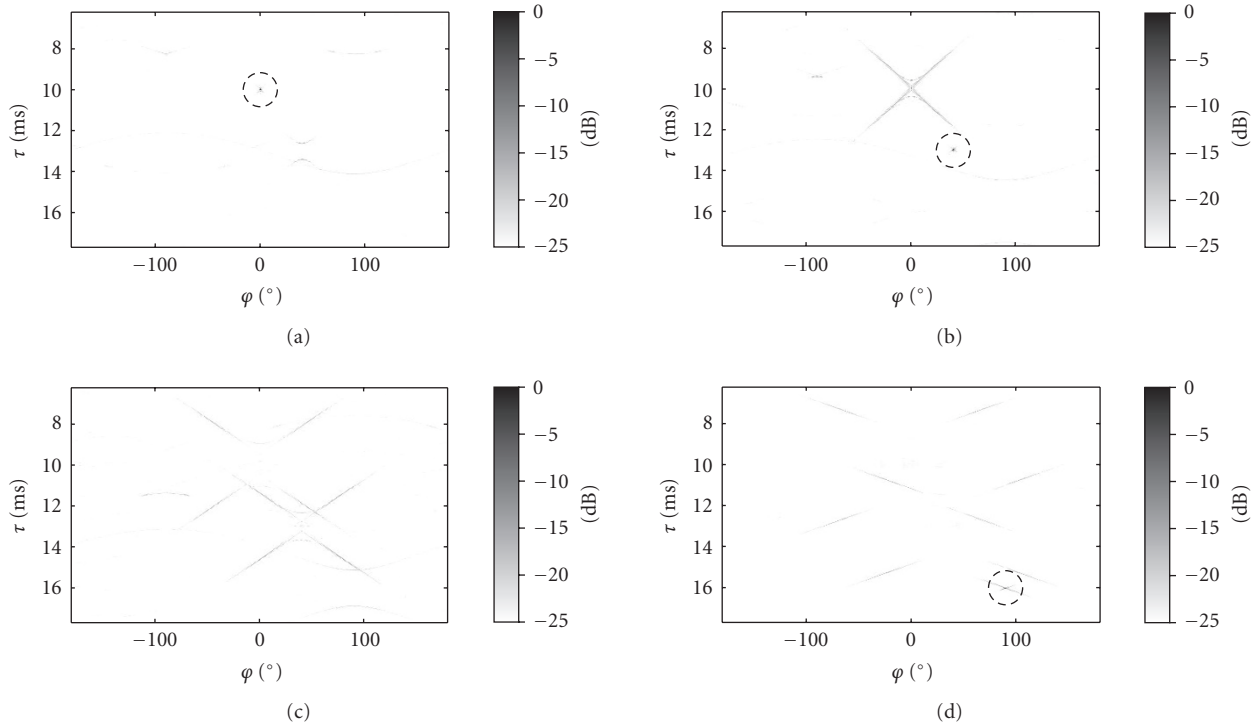


FIGURE 5: Result of Radon transform application to the simulated data of Figure 3. Model space frames are shown for elevation angles (a) $\theta = 0^\circ$, (b) 30° , (c) 50° , and (d) 70° . The representations of the three waves in the data set are marked with circles in the corresponding elevation angle frames.

found as well-resolved events, as expected (see circles). However, when considered in more detail, the events appear to be extended, due to the imperfect inversion of the Radon transform. Therefore, there is some “cross-talk” between the frames of the model space: in frames not corresponding to wave fronts in the data set, energy from other waves appears as cross- or bowtie-shaped shadows. In the frame for $\theta = 50^\circ$ (Figure 5(c)), only such shadows are present.

4.2. Separation and reconstruction

The wave component with an elevation angle $\theta = 70^\circ$ could be a first-order ceiling reflection. Therefore, it was attempted to separate this component from the model space and reconstruct it in the data space.

Figure 6 shows the model space frame for $\theta = 70^\circ$ (Figure 5(d)), zoomed in on the event representing the wave to be reconstructed. The cross-shaped extension of the representation is clearly seen. Top right, a “tail” of the representation of the wave with $\theta = 30^\circ$ is seen. Note that in the model space of a data set measured in a concert hall, many such representations interfere in such a way that even in that space, full separation of the individual waves is impossible. Therefore, a filtering window has to be chosen that will provide a compromise between losing energy belonging to the selected wave component and including energy of other components and noise.

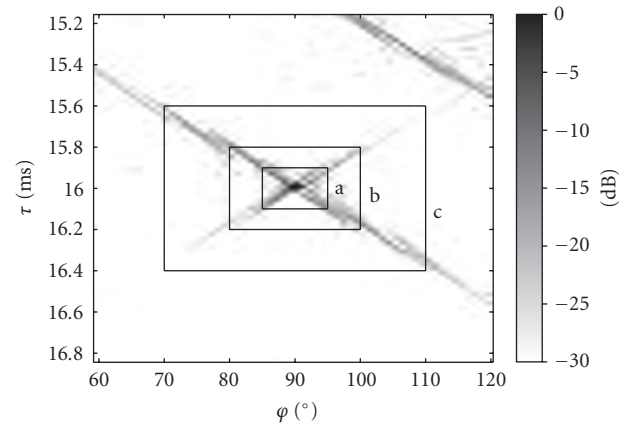


FIGURE 6: The model space frame for $\theta = 70^\circ$ (see Figure 5(d)), zoomed in on the representation of the wave incident under that elevation angle. Three filter windows are used for the selection of data for reconstruction in the data space.

As a first attempt, the rectangular window indicated as “b” in Figure 6 was chosen to select the data for inverse transformation to the data space. The reconstruction result is given in Figure 7(b). It clearly resembles the corresponding wave response in the original data set (see Figure 7(a)), although subtraction of the reconstructed wave component

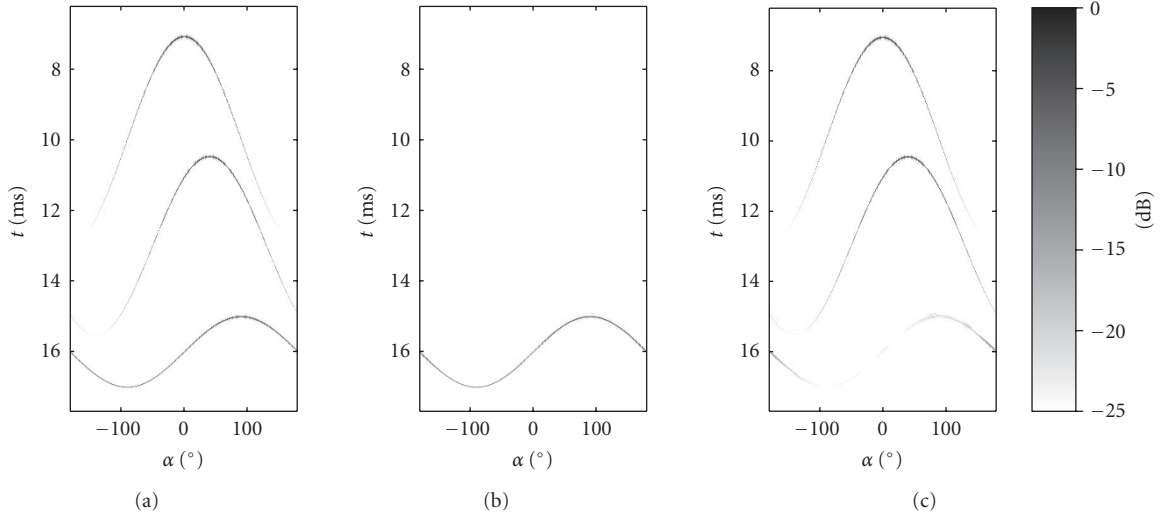


FIGURE 7: (a) The simulated input data set, (b) the reconstructed response of the wave with elevation angle $\theta = 70^\circ$, and (c) the result of subtracting the reconstructed response from the original data.

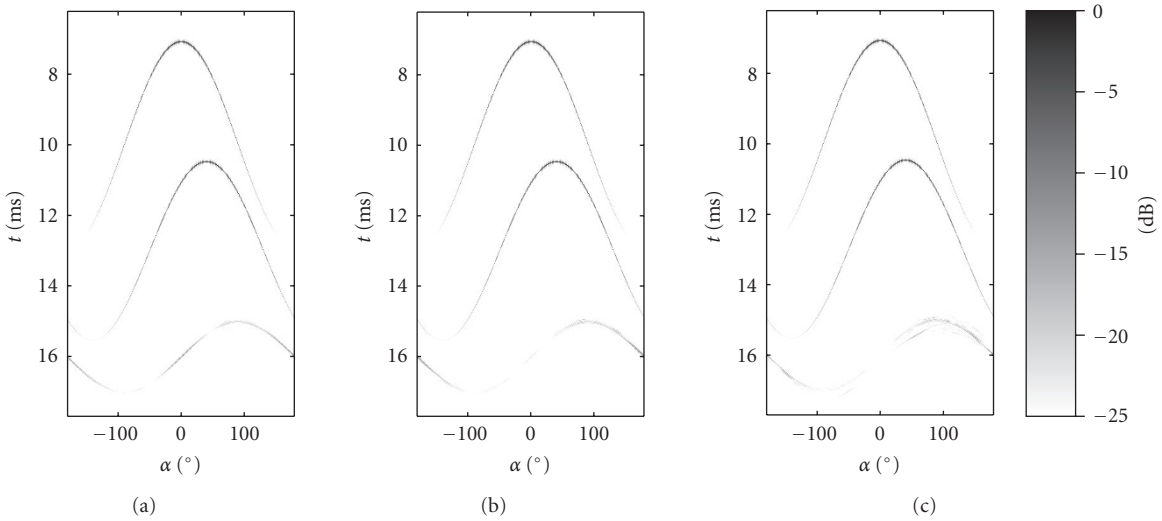


FIGURE 8: Results of subtracting the reconstructed response of the wave with elevation angle $\theta = 70^\circ$ from the data set of Figure 3, after selecting the data in the model space using filter windows “a,” “b,” and “c” (indicated in Figure 6), respectively.

from the original dataset does not yield perfect removal of the wave front (Figure 7(c)). This is mainly due to the imperfect inversion of the Radon transform and the loss in energy caused by the filtering window.

Two other filtering windows have been applied, indicated with “a” and “c” in Figure 6. Subtraction of the reconstructed response from the original data yields the data sets shown in Figures 8(a) and 8(c), respectively. In comparison with the results of window “b” (see Figure 8(b)), the smaller window “a” leaves some parts of the response uncompensated. The application of the bigger window “c” results in better suppression of the wave front, but also introduces some artifacts.

It can be concluded that the Radon transform method, at least for simple simulated data sets, allows us to perform the desired identification, separation, and reconstruction of nonhorizontal components in impulse responses measured

along a circular array of cardioid microphones in the horizontal plane. A more systematic study of the optimization of the filtering windows for the selection of model space data for reconstruction, including tapering and shaping according to the data configuration, could further improve the performance of the method.

5. CYLINDRICAL HARMONICS

Hulsebos et al. [4] have shown how a wave field can be decomposed into cylindrical harmonics,

$$\mathcal{P}_{k_\varphi}^{(1)}(r, \varphi, \omega) = H_{k_\varphi}^{(1)}(kr)e^{jk_\varphi\varphi}, \quad (11a)$$

$$\mathcal{P}_{k_\varphi}^{(2)}(r, \varphi, \omega) = H_{k_\varphi}^{(2)}(kr)e^{jk_\varphi\varphi} \quad (11b)$$

represent the pressure fields of the incoming and outgoing cylindrical harmonics, respectively. $H_{k_\varphi}^{(1)}(kr)$ and $H_{k_\varphi}^{(2)}(kr)$ are Hankel functions of the first and second kinds, respectively; k is the wave number. k_φ is an integer indicating the order of the cylindrical harmonic. It can be seen as an angular wave number, forming a Fourier pair with the azimuthal angle φ . Each cylindrical harmonic is the multiplication of a Hankel function with an orthogonal angular directivity function. The sound field of a monopole corresponds to $k_\varphi = 0$; a dipole field is obtained by taking a linear combination of $\beta_{-1}^{(2)}$ and $\beta_1^{(2)}$.

In [4], it is shown that after a double Fourier transform to the angular wave number and frequency domains, the measured data can be expressed in terms of the pressure P and the normal velocity V_n (i.e., the component perpendicular to the array) as

$$P(k_\varphi, \omega, R) = Q^{(1)}(k_\varphi, \omega)H_{k_\varphi}^{(1)}(kR) + Q^{(2)}(k_\varphi, \omega)H_{k_\varphi}^{(2)}(kR), \quad (12a)$$

$$j\rho c V_n(k_\varphi, \omega, R) = Q^{(1)}(k_\varphi, \omega)H_{k_\varphi}^{(1)'}(kR) + Q^{(2)}(k_\varphi, \omega)H_{k_\varphi}^{(2)'}(kR), \quad (12b)$$

where R is the array radius, and $H^{(1)'}$ and $H^{(2)'}$ are the derivatives of the Hankel functions with respect to kR . $Q^{(1)}$ and $Q^{(2)}$ are the expansion coefficients of the incoming and outgoing wave fields in terms of cylindrical harmonics, which can be found from (12a), (12b), if P and V_n are known on the array. If there are no sources located inside the array, the incoming and outgoing fields must be equal and it is possible to define a single set of expansion coefficients:

$$Q(k_\varphi, \omega) = \frac{1}{2}(Q^{(1)}(k_\varphi, \omega) + Q^{(2)}(k_\varphi, \omega)). \quad (13)$$

It is well known that the characteristics of a cardioid microphone can be obtained by combining a pressure-sensitive monopole microphone and a velocity-sensitive dipole microphone [11]:

$$S(k_\varphi, \omega, R) = \frac{1}{2}(P(k_\varphi, \omega, R) + j\rho c V_n(k_\varphi, \omega, R)). \quad (14)$$

Equations (12a) and (12b) can therefore be rewritten as

$$S(k_\varphi, \omega, R) = Q(k_\varphi, \omega)[H_{k_\varphi}^{(1)}(kR) + H_{k_\varphi}^{(2)}(kR) - jH_{k_\varphi}^{(1)'}(kR) - jH_{k_\varphi}^{(2)'}(kR)]. \quad (15)$$

The decomposition described above is only correct for waves propagating in the horizontal plane, that is, for elevation angle $\theta = 0^\circ$. From Figure 2, it follows that a plane wave with nonzero elevation “sees” a smaller circular array, with radius $R_j = R \cos \theta_j$. In order to determine the expansion coefficients and perform the plane wave decomposition, R_j should be inserted in (15) instead of R . Furthermore, the array receives only a projection of the normal velocity $V'_n = V_n \cos \theta_j$. Again, the transformation is carried

out only for specific values of θ_j . In order to account for plane waves which are not propagating in the horizontal plane, (15) therefore has to be rewritten as

$$\begin{aligned} S_j(k_\varphi, \omega, R_j) &= \frac{1}{2}(P(k_\varphi, \omega, R_j) + j\rho c \cos(\theta_j)V_n(k_\varphi, \omega, R_j)) \\ &= Q_j(k_\varphi, \omega)[H_{k_\varphi}^{(1)}(kR_j) + H_{k_\varphi}^{(2)}(kR_j) \\ &\quad - j \cos(\theta_j)(H_{k_\varphi}^{(1)'}(kR_j) + H_{k_\varphi}^{(2)'}(kR_j))]. \end{aligned} \quad (16)$$

Once the expansion coefficients have been found, the wave field can be calculated for each value of the radial coordinate r . It can be shown [4] that in a far-field approximation, the plane wave decomposition of the incoming and outgoing sound fields in terms of cylindrical harmonics is given by

$$s_\infty(\varphi, \omega) = \frac{1}{\pi} \sum_{k_\varphi} (-j)^{k_\varphi} Q_j(k_\varphi, \omega) e^{jk_\varphi \varphi}. \quad (17)$$

In the ideal case, a plane wave incident with a certain azimuth angle should appear as a point after decomposition in the model space if the elevation angle is chosen correctly for that particular wave. After separation, the reconstruction in the data space can simply be performed using (16) and (17) in reverse order.

To test this method, the simulated data set of Figure 3 was used again. As earlier for the Radon transform method, model space frames were constructed for elevation angles $\theta = 0^\circ, 30^\circ, 50^\circ$, and 70° , shown in Figure 9. Again, the three waves are represented by well-resolved “points” in the corresponding frames, whereas the frame for $\theta = 50^\circ$ only shows low-energy artifacts and noise. As before, the wave with elevation angle $\theta = 70^\circ$, representative for a ceiling reflection, is selected for separation and reconstruction.

Figure 10 shows a zoomed-in picture of its model space representation. In comparison with the Radon transform equivalent, Figure 6, a similar cross-shaped extension pattern of the “point” is seen, but now with lower energy. The same filtering windows “a,” “b,” and “c” as in Figure 6 have been chosen to select the data for reconstruction. Subtraction of the reconstructed responses from the original data set yields the results shown in Figures 11(a), 11(b), and 11(c) for the three windows, respectively. In all cases, the “ceiling reflection” response is reconstructed such that it can be removed almost completely from the data. The best results are now obtained for the larger windows “b” and “c.” Also here, a more systematic study on window optimization has to be carried out before real conclusions on the influence of its choice can be drawn.

It can be concluded that also the plane wave decomposition by means of cylindrical harmonics allows us to identify, separate, and reconstruct elevated components in impulse responses measured along a circular array in the horizontal plane, in a satisfactory manner.

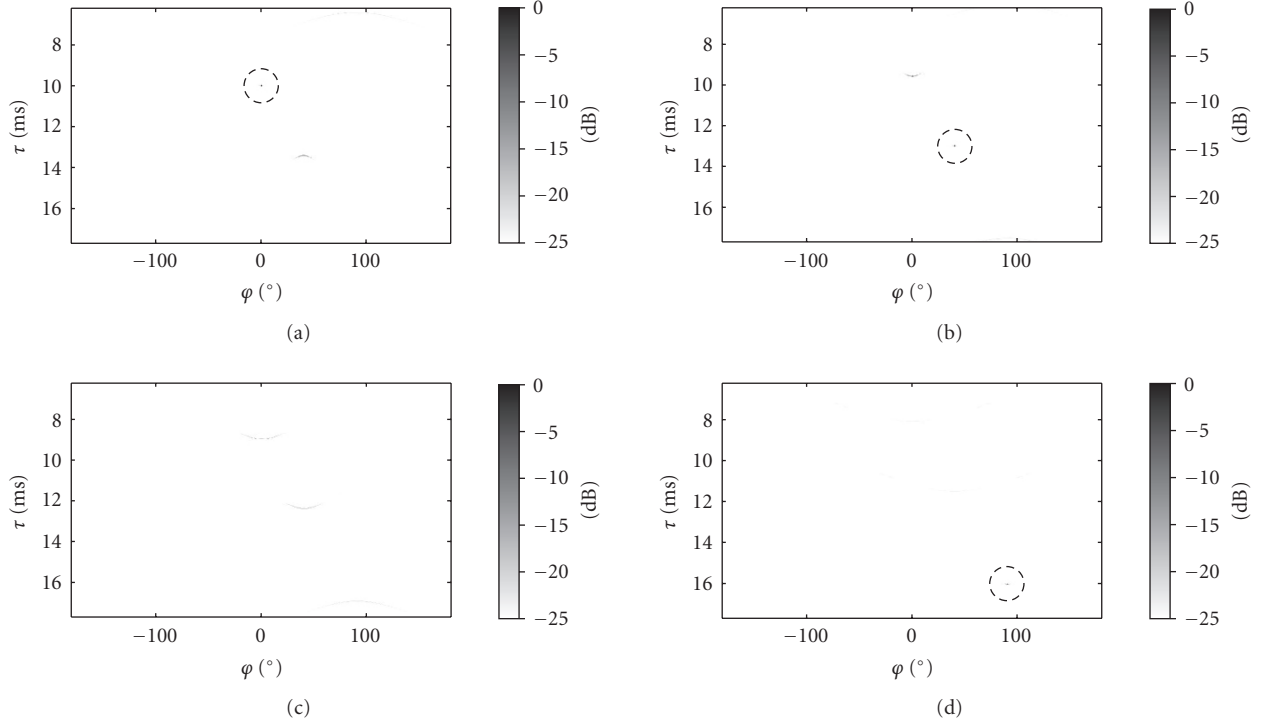


FIGURE 9: Result of cylindrical harmonic decomposition of the simulated data of Figure 3. Model space frames are shown for elevation angles: (a) $\theta = 0^\circ$, (b) 30° , (c) 50° , and (d) 70° . The representations of the three waves in the data set are marked with circles in the corresponding elevation angle frames.

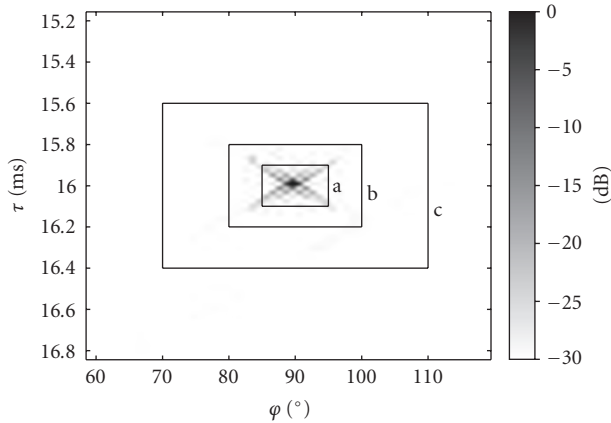


FIGURE 10: The model space frame for $\theta = 70^\circ$ (see Figure 9(d)), zoomed in on the representation of the wave incident under that elevation angle. Three filter windows are shown used for the selection of data for reconstruction in the data space.

6. APPLICATION TO MEASURED DATA

Both methods discussed above have been applied to the data set measured in the Frits Philips Concert Hall in Eindhoven, The Netherlands, shown in Figure 4. Figure 12(a) shows the measured data set, containing the first-order ceiling reflection around $t = 110$ milliseconds.

Note that since the signal source was positioned at the stage center, the azimuth angle of incidence of the ceiling re-

flexion is $\varphi = 0^\circ$. Figure 12(b) shows the ceiling reflection response after identification, separation, and reconstruction using the Radon transform method for an elevation angle $\theta = 70^\circ$. The result of subtracting the isolated ceiling reflection from the original data set is shown in Figure 12(c). The selected ceiling reflection is not sufficiently damped as can be seen from the remaining artifacts indicated by the arrow in Figure 12(c). This is mainly due to the imperfect inversion that is inherent to the Radon transform.

In Figure 13, similar results are presented for the cylindrical harmonic decomposition method. Figure 13(a) shows the measured input data, Figure 13(b) shows the response of the ceiling reflection after identification, separation, and reconstruction using the cylindrical harmonic decomposition method. The extracted ceiling reflection is about 5 dB stronger than in the Radon case. Therefore, better suppression can be achieved after subtraction from the original data set. The region indicated by the arrow in Figure 13(c) exhibits fewer artifacts than the corresponding region in Figure 12(c).

It can be concluded that both methods are able to identify, separate, and reconstruct specific reflections from a data set measured along a circular array in the horizontal plane.

7. CONCLUSIONS

In multitrace impulse responses measured in a hall, using a circular array with cardioid microphones in the horizontal plane, plane wave data appear on cosine-shaped arrival

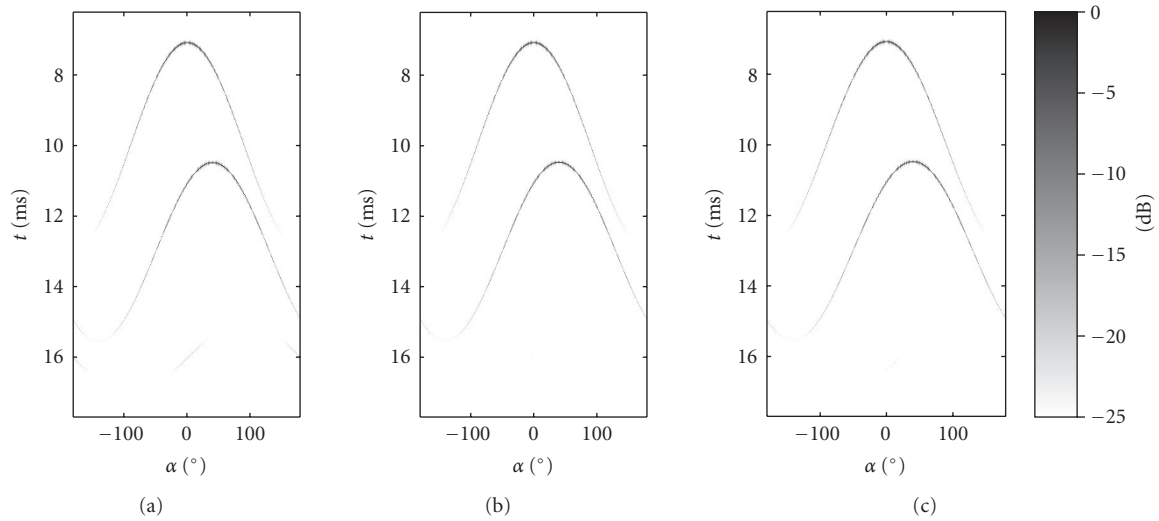


FIGURE 11: Results of subtracting the reconstructed response of the wave with elevation angle $\theta = 70^\circ$ from the data set of Figure 3, after selecting the data in the model space using filter windows “a,” “b,” and “c” (indicated in Figure 10), respectively.

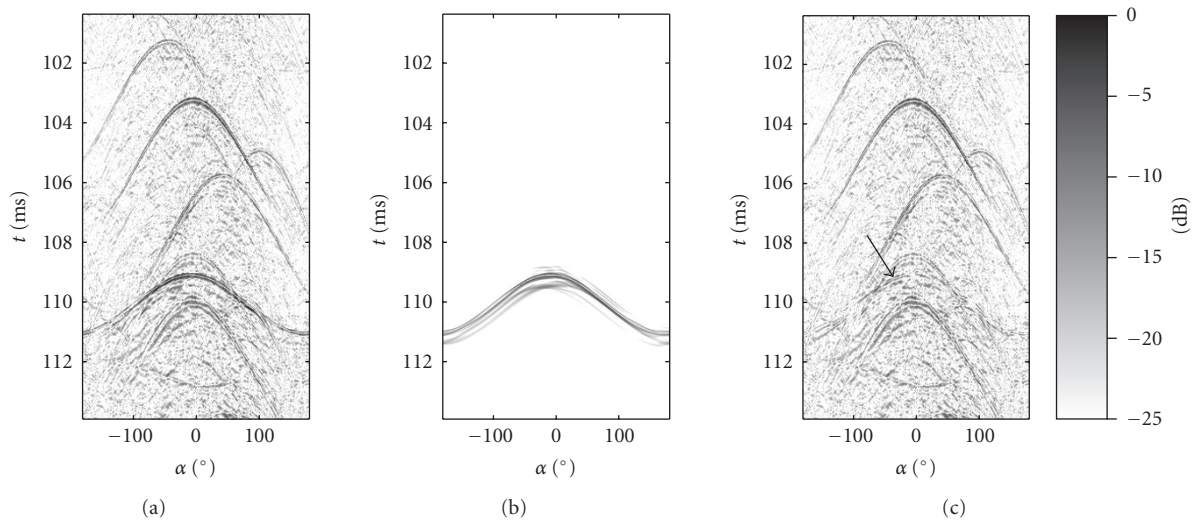


FIGURE 12: (a) Detail of the measured data set of Figure 4, (b) reconstruction of the response of the first-order ceiling reflection using the Radon transform method, (c) the result of subtracting (b) from (a).

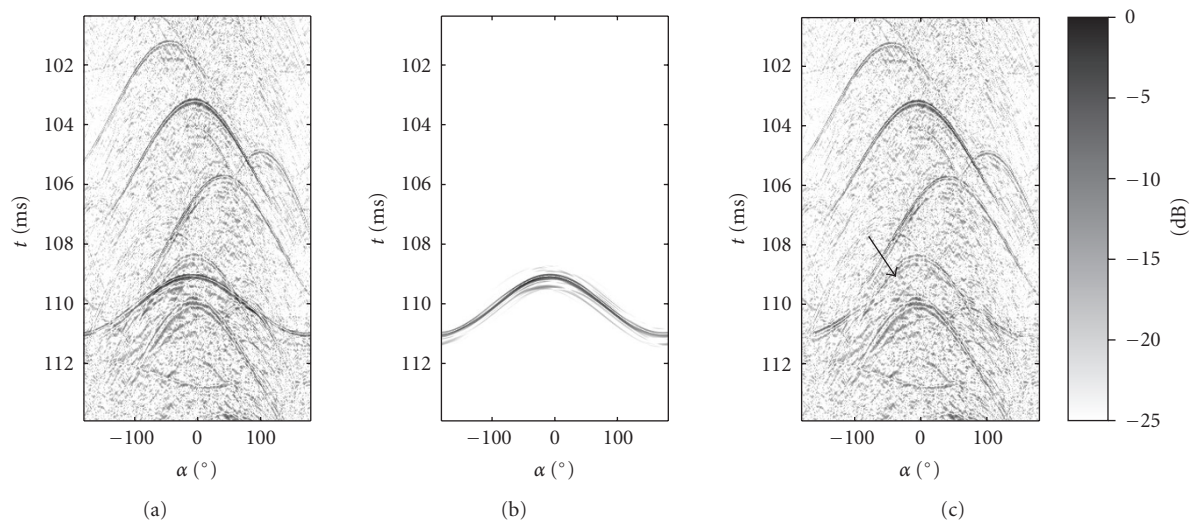


FIGURE 13: (a) The measured data set of Figure 4, (b) reconstruction of the response of the first-order ceiling reflection using the cylindrical harmonics decomposition method, (c) the result of subtracting (b) from (a).

time curves. The geometrical amplitude of the cosine is determined by the elevation angle under which the wave arrives at the array. The strength of the recorded signal at a certain position on the array is determined by both the azimuth and elevation angles of the incident wave.

By this way, a 2D array recording contains 3D information, which can be used to identify, separate, and reconstruct first-order ceiling reflections from the data for further processing in 3D auralization by wave field synthesis.

Two methods have been investigated for this purpose: application of a 2D Radon transform to the impulse response data, and decomposition of the data into cylindrical harmonics and from there into plane waves. Both methods transform the data to a so-called model space where the individual wave components are well resolved. Here, the relevant components are separated by filtering and are reconstructed in the data space by inverse transformation.

Both methods appear to be able to provide useful results. However, the cylindrical harmonics decomposition method outperforms the Radon transform due to the fact that the inversion of the Radon transform is only approximate.

Since the filtering windows used to separate the relevant data from the model space should have a significant influence on the performance of both methods, they should be further optimized.

REFERENCES

- [1] R. Dressler, *Dolby Pro Logic Surround Decoder Principles of Operation*, Dolby Laboratories Licensing, San Francisco, Calif, USA, 1993.
- [2] A. J. Berkhout, "A holographic approach to acoustic control," *Journal of the Audio Engineering Society*, vol. 36, no. 12, pp. 977–995, 1988.
- [3] M. Kleiner, B.-I. Dalenback, and P. Svensson, "Auralization—an overview," *Journal of the Audio Engineering Society*, vol. 41, no. 11, pp. 861–875, 1993.
- [4] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave-field synthesis," *Journal of the Audio Engineering Society*, vol. 50, no. 10, pp. 779–790, 2002.
- [5] A. J. Berkhout, D. de Vries, J. Baan, and B. W. Van den Oetelaar, "A wave field extrapolation approach to acoustical modeling in enclosed spaces," *Journal of the Acoustical Society of America*, vol. 105, no. 3, pp. 1725–1733, 1999.
- [6] J. Blauert, "Sound localization in the median plane," *Acustica*, vol. 22, no. 4, pp. 205–213, 1969.
- [7] J. Blauert, *Spatial Hearing*, MIT Press, Cambridge, Mass, USA, 1983.
- [8] S. R. Deans, *The Radon Transform and Some of Its Applications*, Jon Wiley & Sons, New York, NY, USA, 1983.
- [9] P. Toft, *The Radon Transform—Theory and Implementation*, Ph.D. thesis, Technical University of Denmark, Lyngby, Denmark, 1996.
- [10] M. D. Sacchi and T. J. Ulrych, "High-resolution velocity gathers and offset space reconstruction," *Geophysics*, vol. 60, no. 4, pp. 1169–1177, 1995.
- [11] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, chapter 14, John Wiley & Sons, New York, NY, USA, 2000.

Diemer de Vries was born on January 3, 1945, in Weststellingwerf, The Netherlands. He received his M.S. degree at Delft University of Technology in 1971, carrying out his graduate research at the Laboratory of Acoustical Imaging and Sound Control, which he afterwards joined as a member of staff. During his career as a university researcher, he worked on projects in room acoustics, building acoustics, and seismic signal processing. In 1984, he received the Ph.D. degree on a thesis in the latter field. He now coordinates, as an Associate Professor, the research on array technology-based wave field analysis and synthesis in room acoustics, building acoustics, and audio technology. Since 1981, he also teaches at the Royal Conservatory of Music in The Hague, at the Department of "Art of Sound." During the summer semester of 2001, he fulfilled the "Edgard Varese" guest professorship at TU Berlin. In 2004, he held a Guest Professor Chair at TU Ilmenau, Germany. Diemer de Vries is the Past Chairman of the Dutch Acoustical Association. He is a Fellow of the AES and a Member of the ASA. As a specific form of applied acoustics, he plays the double bass in several orchestras and chamber music ensembles.



Lars Hörchens was born in Mönchengladbach, Germany, in 1979. He studied media technology at Ilmenau Technical University with emphasis on audiovisual technology. After receiving his diploma in 2005, he joined the Laboratory of Acoustical Imaging and Sound Control at Delft University of Technology, where he is currently working on his Ph.D. thesis on the analysis of dispersive wave fields using array technology.



Peter Grond simultaneously started two studies after his secondary school: playing the violin at the Conservatory of Music in Rotterdam and Applied Physics at Delft University of Technology. In 2001, he graduated at the Conservatory. He decided to complete his studies in Delft before starting a musical career, which he did in 2005. The research for his M.S. thesis forms the nucleus of the work reported in this paper. At present, Peter is successful as a violinist in several ensembles, amongst which is his own formation "Free Impulse," for which he writes the arrangements.

