Acoustic imaging in enclosed spaces: Analysis of room geometry modifications on the impulse response

M. Kuster,[a] D. de Vries, E. M. Hulsebos, and A. Gisolf
Laboratory of Acoustic Imaging and Sound Control, Delft University of Technology, 2600 GA Delft, The Netherlands

(Received 28 November 2003; revised 19 May 2004; accepted 2 July 2004)

Sound propagation in enclosed spaces is characterized by reflections at the boundaries of the enclosure. Reflections can be wanted in the case when they support the direct sound or give a feeling of envelopment or they can be unwanted when they lead to echoes and colouration. When measuring multiple impulse responses in an enclosed space along an array the reflections can be mapped to the reflecting objects. Similar to seismic exploration, medical diagnostics, and underwater acoustics, an image of the reflecting objects is obtained in terms of reflected energy. The imaging process is based on inverse wave field extrapolation with the Kirchhoff–Helmholtz and Rayleigh integrals. The inverse of the imaging process recreates the measured impulse responses from the image and it allows one to remove or alter reflecting objects in the image and investigate their influence on the wave field in the enclosed space in a physically correct way. This can be verified by reimaging the altered wave field. Preliminary results from listening tests for the perceptual evaluation are presented. They indicate that the influence of a reflecting object can only be perceived in its close proximity.
[DOI: 10.1121/1.1785591]

PACS numbers: 43.55.Gx [MK] Pages: 2126–2137

I. INTRODUCTION

Traditional approaches to room acoustic modeling suffer from difficulty in the amount of details required to model. Usually, most objects are modeled as plane surfaces with a specified reflection and diffusion coefficient. With the method presented here it is possible to determine the influence of a small object on the acoustics and thus determine to what detail an enclosed space should be modeled. Further, in the case of complex structures such as rows of seats, one can perform an acoustic image of the seats and then ''paste'' that into the modeling software.

Another suggested application is the treatment of strong unwanted (early) reflections. The acoustic image of those will help to identify the reflecting object such that corrective measures can be taken.

The proposition of the work presented here is to map the reflections in the impulse response to the reflecting object resulting in an image of the enclosed space. It will be shown that it is then possible to remove an object from the image and quantify its influence on the acoustics in the enclosed space. The approach is essentially new though some effort to the same end has been made previously.[1]

It is expected that the proposed method will give a better understanding of the importance of individual objects on the acoustics of an enclosed space.

The principle of acoustic reflection imaging is well known in seismics where it is used to characterize the earth's subsurface.[2,3] The fundamental is the measurement of impulse responses along receiver arrays; more recently this has also been introduced in room acoustics.[4]

The acoustics of an enclosed space for a particular source and receiver position is completely described by the room impulse response (RIR) in the time domain or the room transfer function in the frequency domain. Mathematically the room can be interpreted as a linear time-invariant system leading to the following expression,

$$p(t)=\int_{-\infty}^{\infty} h(\tau)s(t-\tau)\mathrm{d}\tau, \tag{1}$$

where $p(t)$ is the pressure, $s(t)$ is the source function and $h(t)$ is the impulse response of the room. The problem is that $h(t)$ is space-variant such that there are infinitely many RIRs in an enclosed space, one for each source–receiver configuration. Traditionally, the acoustics of a hall was characterized, objectively and subjectively, by a small number of representative source and receiver positions.[5,6]

In Fig. 1 two RIRs are shown measured a distance of 0.5 m apart in a simple shoe-box-shaped enclosure. What can be seen is that within the first 22 ms the differences are rather small and for later times the structure becomes more complicated and there is no more similarity. When measuring not just a few impulse responses at arbitrary positions but many and along a straight line the offset-travel time registration shown in Fig. 2 is obtained. With this array measurement method discrete wavefronts caused by reflections can be identified and it is obvious that the individual impulse responses are correlated. The two RIRs of Fig. 1 are marked with arrows.

Array technology is widely used in many fields of science; the most well-known applications in acoustics are in the areas of seismic exploration and medical diagnostics. Maynard *et al.*[7] introduced a method to visualize the near-field radiation characteristics of a sound source. In more re-

[a]Electronic mail: martin@akst.tn.tudelft.nl

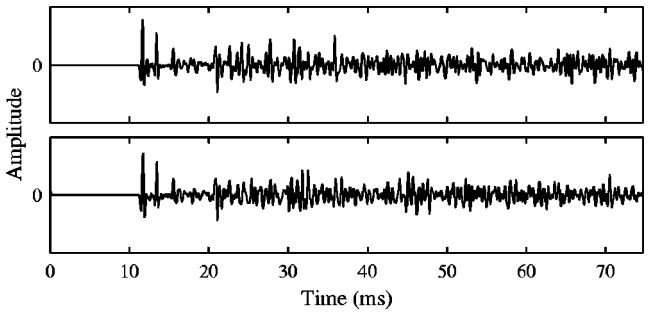

FIG. 1. Room impulse response (RIR) at two positions.

cent years Berkhout[8] proposed a method to reproduce sound fields with natural temporal and spatial properties from loudspeaker arrays.

II. THEORY

The Kirchhoff–Helmholtz integral states that the pressure at a point A in a source-free region can be expressed in terms of the pressure and normal component of the particle velocity on an enclosing surface S.[9] In the case that the surface S degenerates into an infinitely large plane the Kirchhoff–Helmholtz integral allows one to calculate the pressure for all points in the direction *away* from the source and this is known as forward wave field extrapolation. Strictly speaking, only knowledge of pressure or particle velocity is required, leading to the Rayleigh integrals. In the context of room acoustics where the observation plane S is embedded in the medium and waves consequently arrive from all directions, usage of both pressure and particle velocity allows one to distinguish between front and back incidence.

With the aid of inverse wave field extrapolation (or back propagation) it is possible to calculate the pressure at a point *A closer* to the source.[10] Due to the evanescent waves this operation is potentially unstable. To overcome this, the inverse wave field extrapolation operator is usually chosen such that it exponentially damps the evanescent waves. A natural choice is to use the complex conjugate of the forward operator which results in simple time reversal in the time domain.

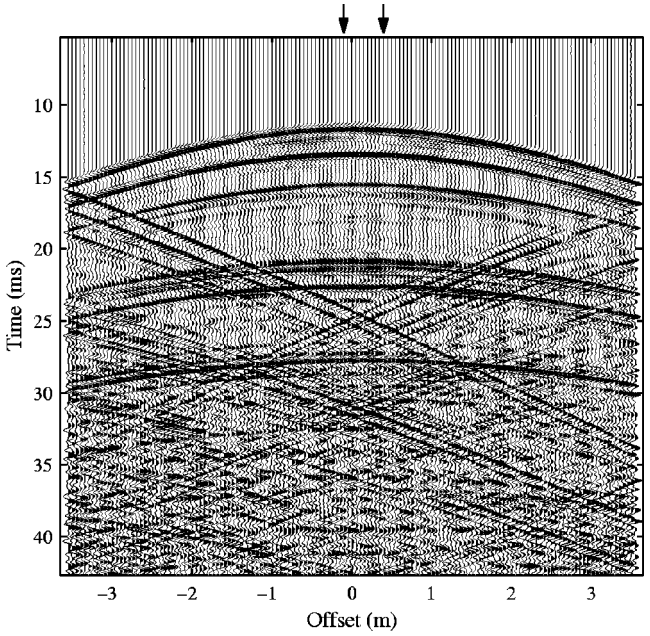

FIG. 2. Multiple room impulse responses measured along a linear array of microphones. The two RIRs of Fig. 1 are marked with arrows.

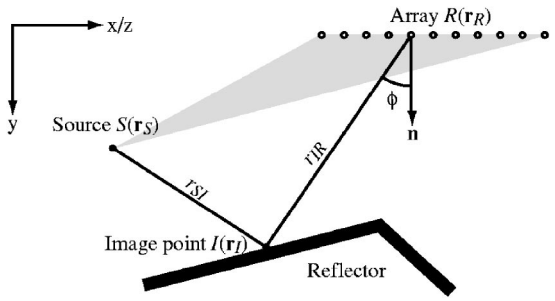

FIG. 3. Geometry for the imaging with source, image point, and receiver array position. The normal to the array points in the positive y direction. Imaging is not possible inside the gray-shaded area.

A. Imaging principle

The fundamental of acoustic reflection imaging is to treat each reflecting object as a secondary diffraction source and perform an inverse wave field extrapolation towards its position. Based on an expression for inverse wave field extrapolation with the Kirchhoff–Helmholtz integral in the time domain, the formulation reads mathematically as follows:[11]

$$\langle p_{\mathrm{Im}}(\mathbf{r}_I)\rangle = \int\int \mathrm{d}r_{R_x}\,\mathrm{d}r_{R_z}[w_{1I}(\mathbf{r}_I,\mathbf{r}_R,t)v_n(\mathbf{r}_R,t) + w_{2I}(\mathbf{r}_I,\mathbf{r}_R,t)p(\mathbf{r}_R,t)]\big|_{t=\tau(\mathbf{r}_S,\mathbf{r}_I,\mathbf{r}_R)}, \quad (2)$$

where $p(\mathbf{r}_R,t)$ and $v_n(\mathbf{r}_R,t)$ are measured pressure and normal component of the particle velocity, respectively, $p_{\mathrm{Im}}(\mathbf{r}_I)$ is the image at point I in terms of reflected pressure, the Green function kernels $w_{1I}(\mathbf{r}_I,\mathbf{r}_R,t)$ and $w_{2I}(\mathbf{r}_I,\mathbf{r}_R,t)$ are given by

$$w_{1I}(\mathbf{r}_I,\mathbf{r}_R,t) = \rho_0 \frac{1}{4\pi r_{IR}} \frac{\partial}{\partial t} \quad (3)$$

$$w_{2I}(\mathbf{r}_I,\mathbf{r}_R,t) = \frac{\cos\phi}{4\pi r_{IR}}\left(\frac{1}{r_{IR}} - \frac{1}{c}\frac{\partial}{\partial t}\right), \quad (4)$$

and

$$\tau(\mathbf{r}_S,\mathbf{r}_I,\mathbf{r}_R) = \frac{r_{SI}+r_{IR}}{c} \quad (5)$$

is the sum of travel times from the source S to the receivers R via the image point I. All other symbols have their usual meanings. The position vector $\mathbf{r}_R$ has the three Cartesian components $(r_{R_x}, r_{R_y}, r_{R_z})$. See Fig. 3 for a definition of the geometry.

Note that S is a constant, I is a parameter and R (more precisely r_{R_x} and r_{R_z}) is the integration variable in Eq. (2). The integration in Eq. (2) with the travel time condition of Eq. (5) is performed over the so-called Huygens surface.[3]

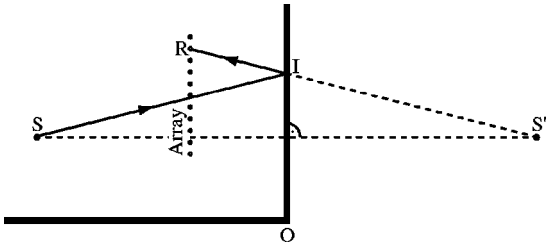

FIG. 4. Illustration of the imaging for one reflecting object.

In the case that only pressure information is available, imaging can be performed based on the Rayleigh II integral. Using the same notation as in Eq. (2), the formulation reads

$$\langle p_{\mathrm{Im}}(\mathbf{r}_I)\rangle = 2\int\int \mathrm{d}r_{R_x}\,\mathrm{d}r_{R_z}\, w_{2I}(\mathbf{r}_I,\mathbf{r}_R,t) \times p(\mathbf{r}_R,t)\big|_{t=\tau(\mathbf{r}_S,\mathbf{r}_I,\mathbf{r}_R)}. \quad (6)$$

The advantage of Eq. (2) over Eq. (6) is that the former allows us to distinguish between front and back by using both pressure and particle velocity. Equation (2) can be expressed in words as follows:

Step 1: Sample the wave field on the array, cf. $v_n(\mathbf{r}_R,t)$ and $p(\mathbf{r}_R,t)$ in Eq. (2).

Step 2: Perform an inverse wave field extrapolation to a point I, cf. $w_{1I}(\mathbf{r}_R,\mathbf{r}_I,t)$ and $w_{2I}(\mathbf{r}_R,\mathbf{r}_I,t)$ in Eq. (2) and $\tau = r_{IR}/c$ in Eq. (5).

Step 3: Calculate the travel time r_{SI}/c for the direct sound to arrive at I and use the extrapolated pressure amplitude at this time as image information, cf. p_{Im} in Eq. (2).

Step 4: Repeat steps 2 and 3 for all I in the region to be imaged.

The scheme has a potential problem when performing imaging on the side of the array containing the primary source. According to step 3 of the imaging process the pressure amplitude at the arrival time of the direct sound is used as image information. For image points in the area between the array and the primary source the inverse extrapolated field will not only contain the reflections but also the direct sound and the image will thus always show the direct sound. The conclusion is that imaging is not possible in the triangle defined by the array edges and the primary source position, corresponding to the gray-shaded area in Fig. 3.

Application of Eq. (2) requires a planar array setup. The formulation can still be employed with a linear array setup. However, since the linear array is rotationally invariant about its axis, the image of the three-dimensional enclosed space will be projected onto a plane (2-D Imaging).

B. Constructing the image of a reflecting object

The understanding of the imaging process is facilitated by the aid of geometrical room acoustics.

Consider the situation in Fig. 4 where a sound ray from the source S is travelling towards the vertical boundary and is reflected at the boundary in point I. For simplicity it is assumed that the reflection is specular and the ray will then be registered at the array in point R. The position of the boundary, however, is unknown in the imaging process. Instead, an inverse wave field extrapolation will be performed for many points between the array and the reflecting object and beyond. If the inverse wave field extrapolation is based on the Kirchhoff integral, both pressure and normal component of the particle velocity are available and consequently the extrapolated amplitude will only be nonzero (or above the noise floor) for points in the direction of the vector $\overrightarrow{RI}$. The second condition to be satisfied is the total travel time corresponding to the distance $\overline{SIR}$. The image point I can then only be placed at the unique position satisfying both conditions. Naturally, the distance $\overline{SI}$ is then equal to the distance $\overline{S'I}$ because that is how the image source was constructed in the first place.

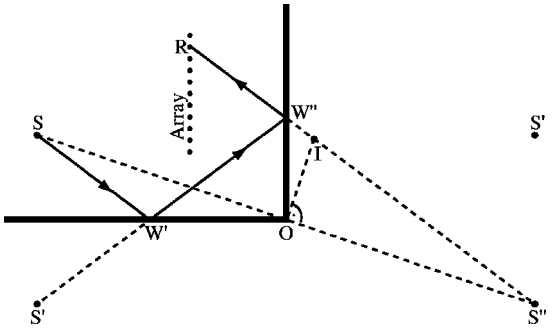

FIG. 5. Illustration of the imaging due to second-order reflections.

C. Constructing the second-order image

In practice wave propagation in enclosed spaces is not constrained to single reflections but subject to second- and higher-order reflections. Referring to Fig. 5, a sound ray travels from S towards the horizontal boundary where it is reflected at W' and then travels towards the vertical boundary. After reflection at point W'' the ray is registered at the array in point R. The imaging process, however, will assume that the ray was only reflected once. The construction of the image then follows along the same lines as in the case of a single reflection. This implies that the only point satisfying both ray direction and total travel time will be point I. The entire vertical boundary will then be imaged on the extension of the vector $\overrightarrow{OI}$. For this situation the distance $\overline{SI}$ is equal to the distance $\overline{S''I}$ for all image points. In the case that a sound ray is reflected first on the vertical and then the horizontal boundary, the image will be appear on the extension of the vector $\overrightarrow{OI}$ below point O.

The idea can also be extended to images from reflections at three objects and more. The image will always appear on the perpendicular bisection of the line from the primary source to the generating mirror image source. In the context of the work presented here, the nth-order image is referred to as the image caused by n reflections or equivalently by a mirror image source of order n.

One potential solution to the problem of higher-order images is to specify source position S' as the actual sound source. The second-order image would then appear at the actual position of the boundary with the drawback that the first-order image would be imaged at a wrong position inside

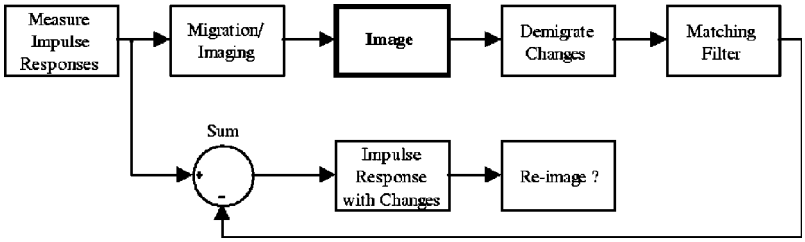

FIG. 6. Block diagram of the process to alter objects.

the enclosure. The effect of specifying an image source as actual sound source can be understood as a translation of the image from one reference frame to another. Rather than mirroring the source, it is also possible to mirror the array, an approach employed in time-reversal focusing.[12] Moreover, there are methods known in seismics to separate or at least attenuate multiple reflections from the RIR;[13] their implementation is the subject of further research.

As will become evident in later sections of this paper, the correct dealing of second- and higher-order images is important because (a) they allow one to add information where the first-order image is not visible and (b) in order to correctly remove the influence of an object from the RIR, not only the first-order reflection but all second- and higher-order reflections involving that object have to be removed.

D. Demigration principle

Demigration is understood as the inversion of the imaging (or migration) process. While imaging can be understood as a transformation of the measured RIRs into an image (from time to distance domain), the demigration transforms the image back into the RIRs (from distance to time domain).

The principle is to interpret each image point as a notional point source firing at the time corresponding to its distance from the primary source and with its strength given by the reflected pressure amplitude. The scheme can be interpreted as the inversion of Eq. (6). The integration is also performed over points in space and looks as follows:

$$\langle p(\mathbf{r}_R,t)\rangle = \int\int \mathrm{d}r_{I_x}\,\mathrm{d}r_{I_z}\, w_{2F}(\mathbf{r}_R,\mathbf{r}_I,t) \times p_{\mathrm{Im}}(\mathbf{r}_I)\big|_{r_{I_y}=\psi(\mathbf{r}_S,r_{I_x},r_{I_z},\mathbf{r}_R,t)}, \tag{7}$$

where

$$w_{2F}(\mathbf{r}_R,\mathbf{r}_I,t) = \frac{\cos\phi}{4\pi r_{IR}}\left(\frac{1}{r_{IR}}+\frac{1}{c}\frac{\partial}{\partial t}\right) \tag{8}$$

and the integration surface $r_{I_y}=\psi(\mathbf{r}_S,r_{I_x},r_{I_z},\mathbf{r}_R,t)$, the surface of equal reflection time or isochrone, is given implicitly by

$$t(\mathbf{r}_S,r_{I_x},r_{I_y}=\psi(\mathbf{r}_S,r_{I_x},r_{I_z},\mathbf{r}_R,t),r_{I_z},\mathbf{r}_R) = \frac{r_{SI}+r_{IR}}{c}, \tag{9}$$

in other words, by the same sum of travel times as in Eq. (5) except that in this case I rather than R varies. Equation (7) essentially describes forward wave field extrapolation with the Rayleigh II integral where the pressure at the image point I is only known at one instant in time (that is the travel time of the direct sound from S to I).

The fact that the notional sources have no directivity is maybe not too intuitive for the specular reflections but many notional sources of equal strength aligned on a plane will lead to a specular reflection.

When applied over all possible image points, application of Eq. (7) will potentially recreate the original RIR. In practice, where imaging is performed with band-limited signals and a finite aperture, it was found that there is a difference between the original and demigrated RIR in phase and amplitude and a matching filter was employed to minimize the difference in a least-square sense. Another point is that the image was sampled according to the minimum sampling criterion ($\Delta y=\lambda_{\min}/4$). Without oversampling in image space, a low-pass filter, analogous to the reconstruction filter in a digital-to-analogue conversion, has to be applied to the reconstructed impulse response in order to suppress the aliased high frequencies.

1. Alteration of objects

Using demigration it is possible to alter objects and investigate the changes in the acoustics by means of the impulse response. The scheme is shown in Fig. 6. The image of the object to be altered is demigrated. The demigration has to be applied only over those image points containing any changes. After applying the matching filter, the obtained impulse responses are subtracted from the original measured ones and the outcome will contain the impulse responses with the changes. The reason for this sequence is that the reflections due to the object to be altered first have to be removed. In order to take multiple reflections into account, the process has to be repeated with the images obtained from specifying the mirror image sources as primary source.

Examples of object alterations are as follows.

(i) Remove the object (this will result in a hole).
(ii) Replace the object with another object.
(iii) Alter the frequency-dependent reflection properties of the object by frequency filtering the demigrated impulse responses before subtraction.
(iv) Render the surface of a planar object more diffusive by displacing individual image points by a small random or periodic offset analogous to diffusors based on reflection gratings.

III. SIMULATIONS

The pressure and particle velocity impulse responses were generated using a mirror image source model of analytical monopoles with the second derivative of a Gaussian

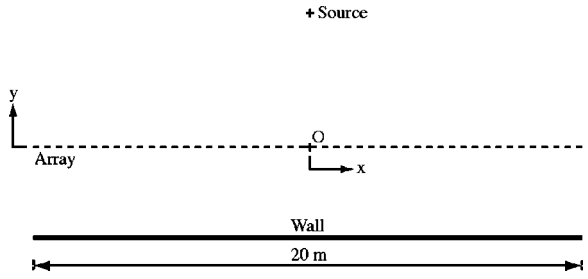

FIG. 7. Geometry of the single reflector imaging.

as wavelet. The array consisted of ideal pressure microphones spaced at 0.05 m. The sampling frequency was 16 kHz.

A. Illustration of the imaging principle

The workings of the imaging are illustrated for the case of one reflector as shown in Fig. 7. The reflector extends from $x=-10\,\mathrm{m}$ to $+10$ m and is at $y=-2.71\,\mathrm{m}$. A point source is placed at $y=4.10\,\mathrm{m}$. The linear array is at $y=0$ and has the same dimension as the reflector. Both, source and array, are placed symmetrically in front of it. In Fig. 8 the inverse extrapolated field is shown for four y positions. This corresponds to step 2 in the process of Sec. II A. The dashed hyperbola represents the travel time for the direct sound to arrive at the particular position and is the line along which the pressure will be used as image content for the given y position.

In Fig. 8(a) the y position is between the array and the reflector and the image content will be zero (or noise). In Fig. 8(b) the field shown is exactly at the reflector position and it can be seen that the hyperbola follows the wave front

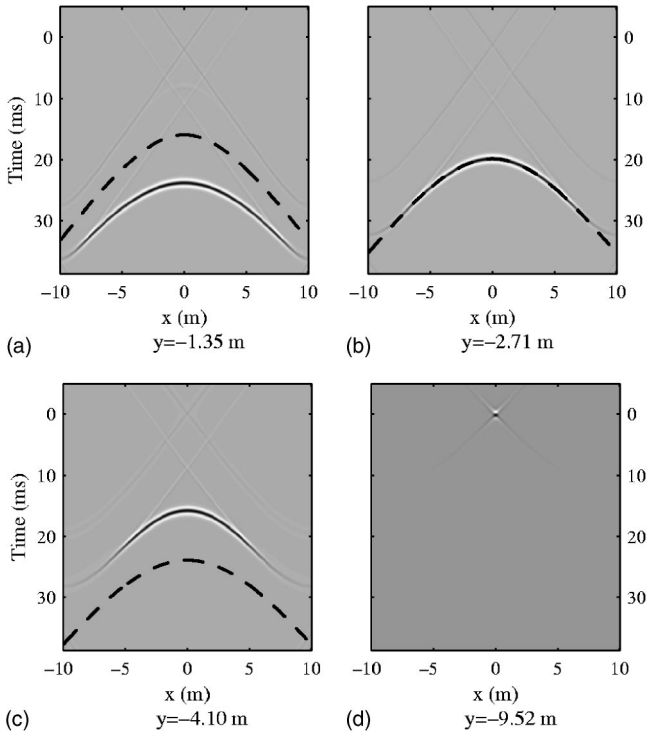

FIG. 8. Inverse extrapolated field due to one reflector at four y positions. The dashed hyperbola is the line along which the pressure will be used as image content.

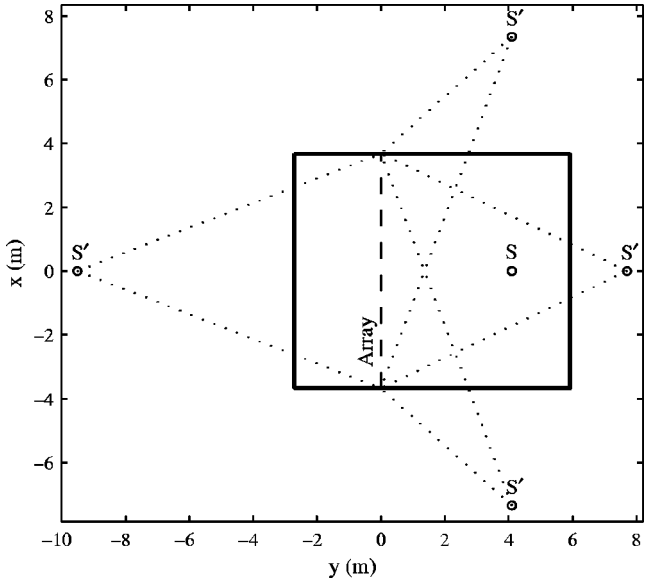

FIG. 9. Geometry of the shoe box enclosure with array and sound source S. First-order mirror image sources are indicated S′.

which will then constitute the image of the reflector. In Fig. 8(c) the y position is beyond the reflector and as in Fig. 8(a) the image will not contain significant amplitudes. The main characteristic when moving away from the array is that the extrapolated wavefront is shifted towards decreasing time and the hyperbola along which the pressure is used as image content is shifted towards increasing time. They only coincide at the position of the wall.

Since the wave front is not a delta function but spread out in time, the image amplitude just in front and behind the reflector is not zero but contains the amplitudes of the wavelet at either side of the main lobe. Imaging with band-limited wavelets will consequently always lead to images of reflectors with finite width.

In Fig. 8(d) the field is shown for the y position of the mirror image source caused by the reflector. The whole wave front has collapsed into a point at $t=0$ and $x=0$ and the mirror image source is in focus. However, the aim is to image the reflector and its secondary sources are in focus in Fig. 8(b).

B. 2-D imaging: Shoe box enclosure

A floor plan of the enclosure together with source and receiver geometry is shown in Fig. 9. The vertical distance from both the source and the array to the floor was 1.20 m and that to the ceiling 1.72 m. The reflection coefficient of all boundaries was set to 1. Since the simulations involved a linear array only 2-D imaging could be performed.

1. Enclosure image without floor and ceiling

In order to remove the floor and ceiling, the reflection coefficient of those two boundaries was set to zero. The obtained image of the enclosure is shown in Fig. 10 where for clarity the negative image amplitudes have been omitted. From the figure, it is immediately evident that only parts of the wall appear in the image. This is basically a visibility issue caused by the finite aperture of the array. In Fig. 9 the

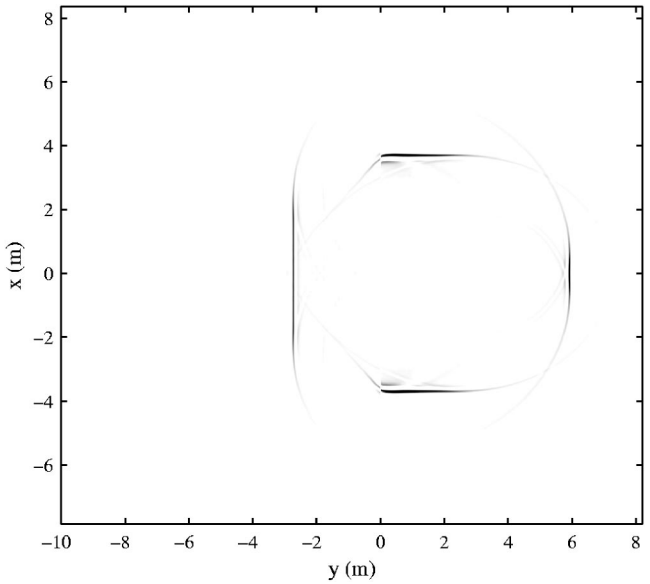

FIG. 10. Acoustic image of the shoe box enclosure from simulated data taking only the first-order image sources of the walls into account.

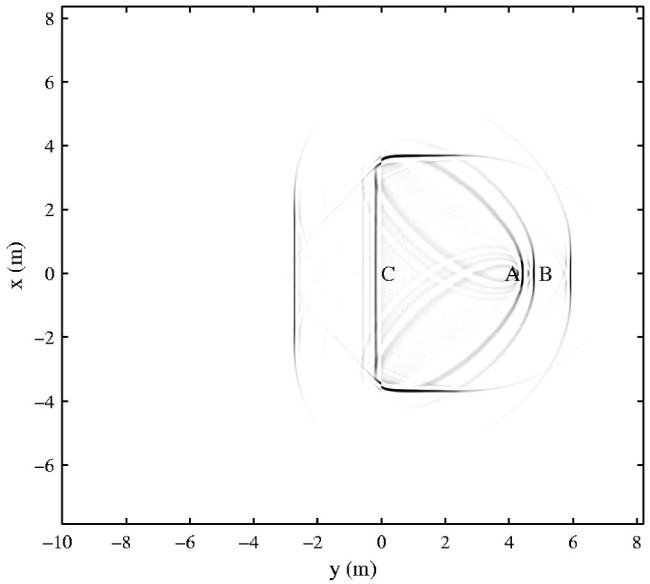

FIG. 12. Acoustic image of the shoe box enclosure from simulated data taking all first-order image sources into account.

two dotted lines from each image source to the array edges indicate which part of the wall will be seen on the image. This example illustrates nicely the helpfulness of the image source model to understand the imaging result.

2. Enclosure image without floor and ceiling from second-order reflections

Figure 11 shows the image of the enclosure when only second-order reflections are used in the simulation. The imaging has been performed specifying the first-order image sources as sound sources. For each one only those second-order reflections involving the particular image source have been included in the simulations. Figure 11 is thus a summation of the four images obtained from all four first-order image sources. From the figure it can be seen that the second-order images allow one to add information about the walls where there was no information in the first-order image, in particular at the enclosure corners.

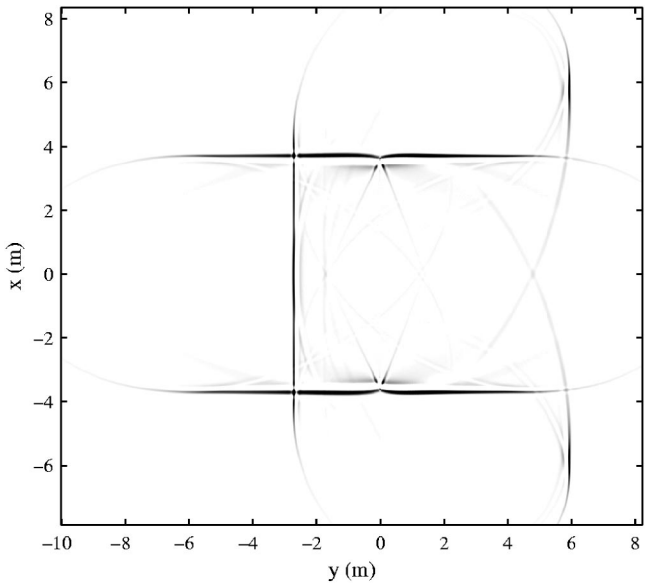

FIG. 11. Acoustic image of the shoe box enclosure from simulated data performing the imaging with respect to the first-order image source and considering only second-order reflections.

3. Enclosure image with floor and ceiling

The situation where floor and ceiling are present is now considered. The image obtained is shown in Fig. 12. Compared to Fig. 10, there are two more events marked ''A'' and ''B'' visible. ''A'' is the image of the floor and ''B'' is the image of the ceiling. The occurrence of these two images is a consequence of the projection resulting from the linear array setup as addressed in Sec. II A. Another issue is that the linear array relies on the polarity of pressure and particle velocity to distinguish between front and back. However, this works only correctly for mirror image sources in the $x-y$ plane and consequently the images of the floor and ceiling will appear twice. In Fig. 12 event ''C'' is the image of the floor on the incorrect side of the array while the one due to the ceiling is barely visible just to the left of it. Events ''A,'' ''B,'' and ''C'' are undesired artifacts purely due to 2-D imaging and will disappear in 3-D imaging.

4. Enclosure image incorporating higher-order image sources

Figures 10 and 12 were obtained by only including first-order mirror image source in the model. For a comparison with images obtained from measured data, the simulated data should incorporate all mirror image sources up to at least a specified distance. This leads to the image of the enclosure in Fig. 13 for the case that image sources within 20 m are included. From the figure, a whole series of additional images is visible. As discussed in Sec. II C, each mirror image source creates its own image. In the same section the position of those images was established. In the figure the three events above the letter ''D'' are the first- and second-order images of the side wall with negative x coordinate while the

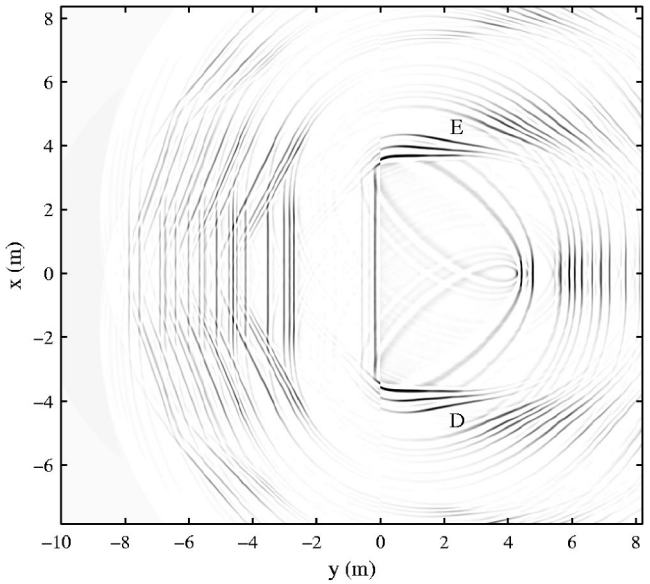

FIG. 13. Acoustic image of the shoe box enclosure from simulated data taking all image sources within a distance of 20 m from the array center into account.

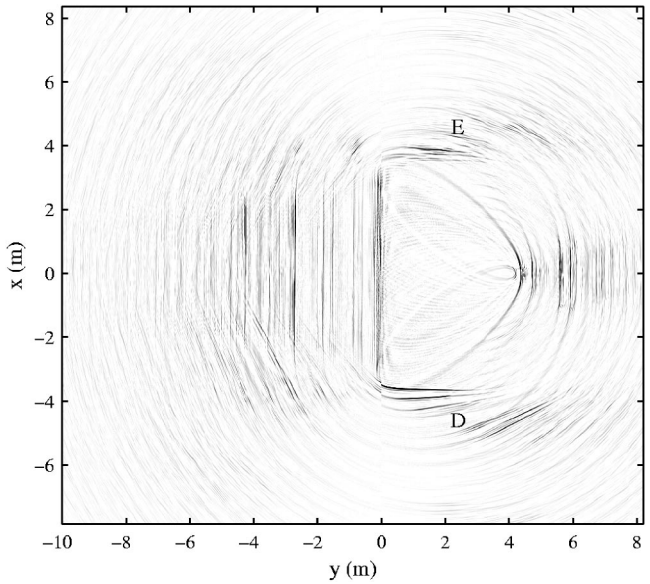

FIG. 14. Acoustic image of the shoe box enclosure from measured data.

events to the right are the higher-order images due to reflections between the side wall and the wall behind the source. The same structure can be seen around ''E'' for the side wall with positive *x* coordinate. The visibility of the higher-order images is defined in the same way as for the first-order images.

In comparison with Fig. 10, Fig. 13 is much more difficult to interpret. It is important to understand at this stage the difference between events due to 2-D imaging of a 3-D space and those due second- and higher-order reflections. The former's position in the image is a function of the room and measurement geometry and they can, as is the case here, appear *within* the physical boundaries of the enclosure. Images of second- and higher-order reflections, on the other hand, will always appear *outside* the physical boundaries of the enclosure. As Fig. 5 illustrates, enclosure corners are the only points where images of different order coincide. Neglecting the artifacts due to 2-D imaging, Fig. 10 can be retrieved from Fig. 13 by constructing radial traces from the source position in all directions and observing the position of the maximum pressure amplitude in those traces. The shortcoming of such an approach is that in the enclosure corners where the first-order image is not visible the algorithm would include the second-order image information. This problem can only be solved by interpreting the image, but in practice the enclosure geometry is more or less known as *a priori* information.

IV. MEASUREMENTS

Measurements have been performed using a source that was reasonably omni-directional within the area of interest. Pressure and particle velocity were measured using a Sound-Field MKV microphone, again with a spacing of 0.05 m. The sampling frequency was 16 kHz except for the shoe box enclosure where it was 14 980 Hz.

A. 2-D imaging: Shoe box enclosure

Figure 14 shows the acoustic image obtained from measurements in a lecture hall with the same dimensions as the enclosure of Sec. III B. When comparing to Fig. 13, it can be seen that all events discussed so far are reproduced in the image. Most of the discrepancies can be explained by the differences in reflection coefficients and the not modelled irregularity of some surfaces. The image of the side wall with positive *x* coordinate (below letter ''E''), for example, looks much more diffuse than the one at negative *x* coordinate (above letter ''D''). This is explained by the fact that the particular wall consisted of a window front covered by curtains rather than a hard wall.

B. 2-D imaging: Concert hall De Doelen

As a more practical case the imaging process is tested on measurements performed in the concert hall De Doelen, Rotterdam. A photograph of the hall is given in Fig. 15 and the floor plan is shown in Fig. 16. The source was situated on the stage approximately at the conductor's position. The linear array was positioned between the tiers of seating and spans the whole width of the inner seating area.

Figure 17 shows the acoustic image of the area within the dash-dotted rectangle of Fig. 16. In the figure the struc-

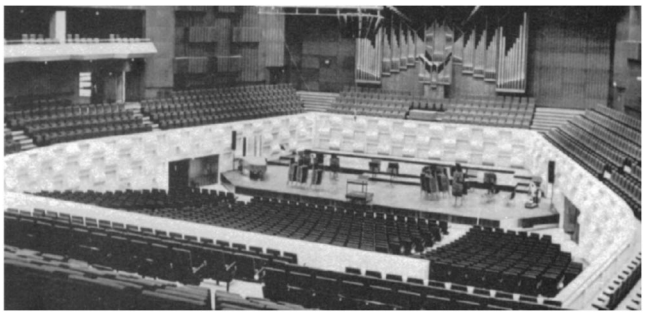

FIG. 15. De Doelen, Rotterdam (from M Barron. *Auditorium acoustics and architectural design*).

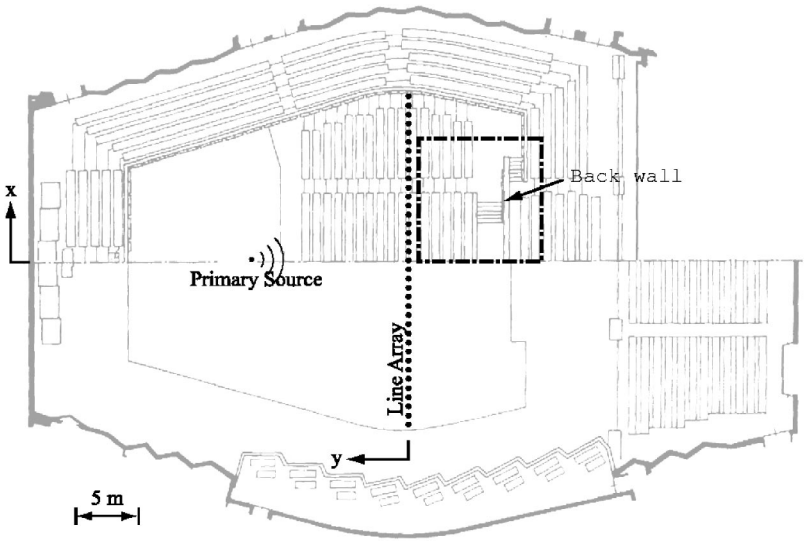

FIG. 16. Floorplan of De Doelen, Rotterdam with source and array position indicated. The dash-dotted rectangle is referred to in the text (from M Barron. *Auditorium acoustics and architectural design*).

ture enclosed in the black rectangle represents the image of the wall behind the last tier. The structure of this diffusive wall can be seen in Fig. 15 as the white checkerboard squares of convex curved shape surrounding the inner seating area. This sequence is very nicely reproduced in the image by segments above and below a median line. The vertical bands on the left side of the figure are the images of the tiers of seating and the dashed line indicates the position of the corridor between them. It is possible to recognize the top of the backrest of each individual seat in the row. Further, one can also observe that below the corridor seats from subsequent rows are offset by half the seat's width such as to improve visibility towards the stage whereas above the corridor this measure has been omitted since the audience already needs to look slightly sidewards to see towards the center of the stage. Also notable is that due to the mapping of the three-dimensional space onto a plane, the rows are not spaced equidistantly.

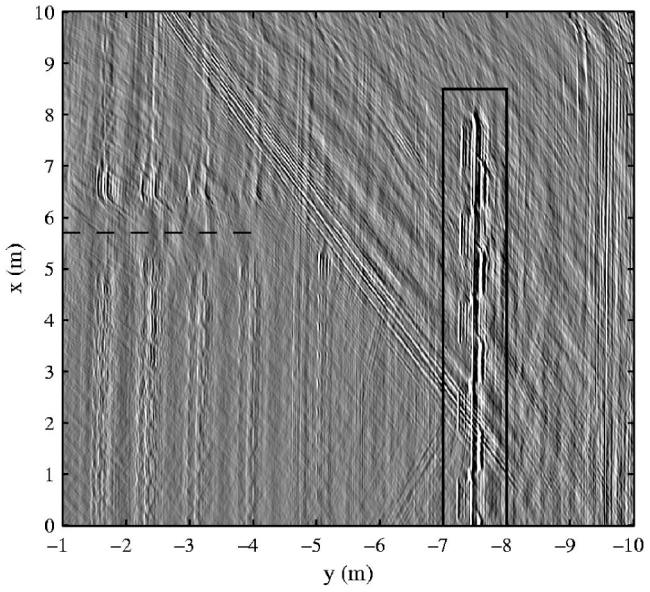

FIG. 17. Acoustic image of the area within the dash-dotted rectangle of Fig. 16.

C. 3-D imaging: Corridor wall

Planar array measurements have been carried out in a simple building corridor. The array runs parallel to the corridor wall at a distance of 0.88 m and over a length of 7.00 m. The height of the array was 2.50 m. The shortest distance between the source and the array was 1.68 m. A photograph of the corridor wall is shown in Fig. 18 while Fig. 19 shows the floor plan and elevation. The chair at the right of the left column is not present in the photograph. The origin of the coordinate system is at the center of the lower edge of the array aperture which is at a height of 0.35 m from the floor. The rather unusual axis orientation is due to the measurement setup.

The problem of displaying the three-dimensional image was approached by showing horizontal slices of the image at different heights and an attempt to visualize the wall surface in three dimensions.

1. Slices at different heights

In Fig. 20 horizontal slices are shown at four representative heights. With reference to Fig. 19 and the photograph in Fig. 18 most objects can be identified.

In all four images the vertical columns are seen at $x=-1$ m and $x=2$ m and also the wall in-between is imaged well. In Figs. 20(a)–(c) the front of the closet is seen between approximately $x=-3$ m and $x=-2$ m while in Fig.

FIG. 18. Photograph of the corridor wall.

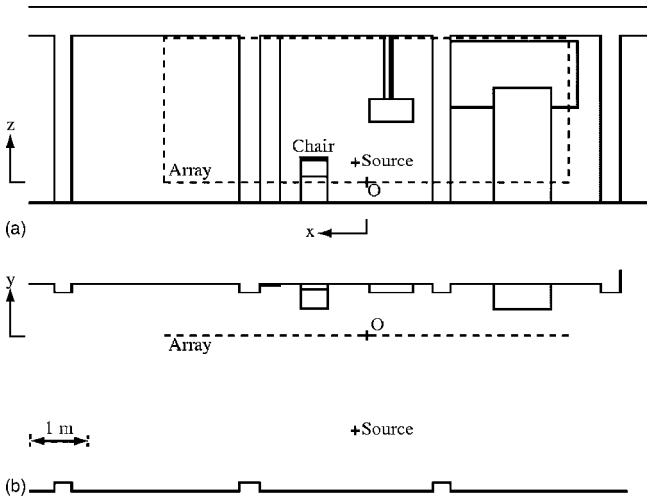

FIG. 19. Elevation (a) and floor plan (b) of the corridor.

20(d) it has disappeared since the slice is at a height above the upper edge of the closet. In Fig. 20(a) the backrest of the chair in front of the wall can be seen at $x=1$ m. In Fig. 20(c) the image of the wall between the columns is discontinued between $x=-0.5$ m and $x=0.4$ m by the electrical distribution box. Also seen, particularly well in Fig. 20(b), is the small protrusion, 0.33 m wide and constructed out of the same material as the columns, to the right of the left column.

The four slices shown were selected rather arbitrarily but showed the main objects present. Naturally, one could also display slices in the vertical plane. However, the visual perception of an object is limited in either case.

2. Surface representation

The second approach used to display the information is based on the fact that the wall and the objects in front form a surface. For each (x,z) pair of the volume image, the y position of the maximum pressure amplitude in the image is recorded. The motivation is that a ray towards the wall will be reflected at only one point (unless there are acoustically transparent materials). The surface formed is then passing through all those points. The result is shown in Fig. 21 for the part of the wall visible from the array. In comparison to the slices the interpretation is much more intuitive. The closet and the two vertical columns can be recognized easily. Also the structure of the distribution box can be identified. The lump at the bottom 1 m to the right of the left column is the surface due to the chair while the spikes to the left are due to the washbasin. The fine vertical line just left of the column is the image of the plastic drain pipe. The smallest object that can be positively identified is the light switch mounted at half-height on the left column.

V. DEMIGRATION

The images shown in Sec. IV were interesting to look at but compared to a photograph their creation is much more laborious and, moreover, due to the longer wavelength, they do not show the same amount of detail. The main difference, however, is that the image is obtained in terms of acoustic rather than electromagnetic reflection properties. The second

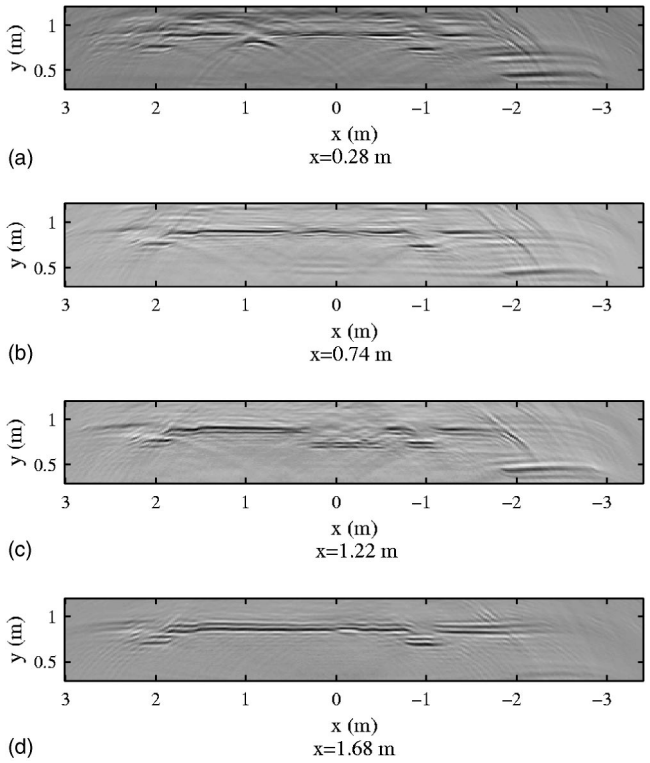

FIG. 20. Image slices at different heights.

important issue is that the room impulse responses used to obtain the image represent the transfer function between the source and the receiver and thus define the acoustic quality in an objective and subjective sense. The interesting part, introduced in this chapter, is the possibility to remove or alter objects in the image and investigate the influence on a particular impulse response either by objective parameters or perceptually.

A. Removal of objects

The process of removing objects was investigated for the distribution box and the closet in front of the corridor wall from Sec. IV C. In Fig. 22 a horizontal slice of the image is shown where the cross section of the volume over which the demigration was performed is indicated by the dashed rectangle. The volume extended over the whole height of the distribution box. The same volume was used in the demigration of the second-order images caused by ground and ceiling reflection, respectively.

For the closet the whole area of x values smaller than $x=-1.62$ m over its height was removed. The motivation was to include the second-order image caused by reflections from the side of the closet. For both objects the consequence of removal is that there is a hole in the wall. With the closet, the wall was effectively shortened.

1. Influence on the impulse responses

Another way to investigate the influence of the object on the impulse response is to calculate the energy contained in the reflections from the object after demigration. Figure 23 displays a view at the planar array aperture where the grayness is proportional to the reflected energy at that measure-

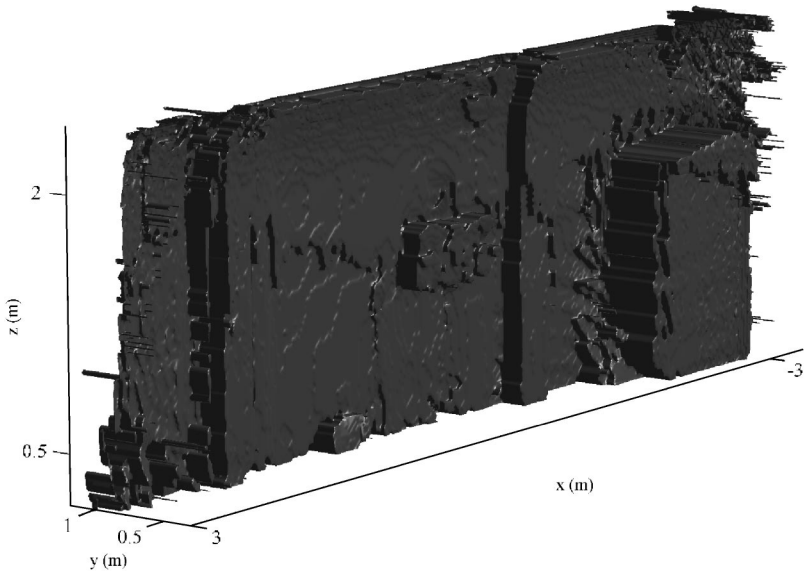

FIG. 21. Surface representation of the corridor wall.

ment position. The dotted rectangles mark the projection of the distribution box and the closet onto that plane.

The influence of the closet is generally confined within an area right in front of it though more concentrated on the right side because of the direction of the reflected waves from the source. There is some spread above the upper edge due to the second-order reflection via the floor. The gray area between $x=-1$ m and $x=-2$ m is caused by second-order reflections from the side of the closet and the wall. The influence of the distribution box is largest just above its position, again due to the direction of the reflected waves. The larger spread compared to the closet is partly due to second-order reflections via floor and ceiling and partly due to the fact that the distribution box is of irregular shape causing scattering. For better visibility the energy due to the distribution box has been amplified by a factor of 3.

B. Reimaging

The ultimate test to check whether the removal of the object was performed in a physically correct way is to reimage from the impulse responses obtained after demigration. This was performed for the distribution box. The result is shown in Fig. 24 for the same slice as in Fig. 22. As the figure shows, the object has really been removed almost completely while all other objects are still present. Presumably due to the matching filter, there are residues of the removed object and some minor differences in the fine structure of the image. When using the same surface representation as in Sec. IV C 2 in combination with a threshold to block small image amplitudes, Fig. 25 is obtained. From the figure it can be seen that, in fact, the object has been removed with the exceptions of some residues partly due to imperfections in the matching filter process and partly due to fact that the control volume did not extend deep enough in the z direction.

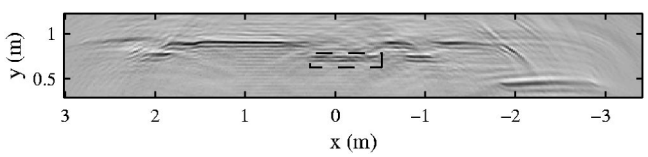

FIG. 22. Image slice at distribution box height ($z=1.16$ m).

VI. LISTENING TESTS

A complete psychoacoustical evaluation of the audibility of reflections from an indivual object is beyond the scope of this paper, Instead, a short initial study has been conducted with the aim to obtain a basic understanding to what extent the reflections from an individual object are perceptually relevant.

An extensive discussion about the subjective effects of reflections is discussed by Kuttruff.[14] In short, the main issues are that with a transient sound such as speech, the reflection will, depending on the delay time, either add to the loudness of the direct sound or be perceived as a distinctive echo. With a continuous sound such as noise the reflection will cause spectral colouration.

A. Experimental setup

The listening positions chosen were close to or right in front of the removed objects (closet and distribution box) and selected such that the energy ratios within the first 64 ms between altered and original RIR were distributed approximately even between 0.60 and 0.98. The selected positions

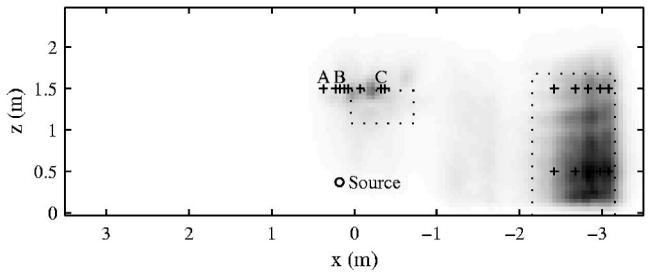

FIG. 23. View at array aperture. Grayness proportional to energy contained in the reflections from the closet and the distribution box. The dotted rectangles mark the projection of those onto the array plane. Positions included in the listening test are marked with crosses. Letters A-C are referred to in the text.

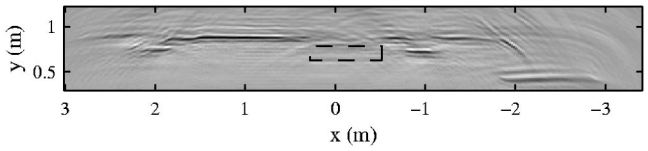

FIG. 24. Reimaged slice after removal of distribution box ($z = 1.16$ m).

used in the listening tests are indicated with crosses in Fig. 23. The shortest distance to the reflecting object was 0.45 m in the case of the closet and 0.75 m in the case of the distribution box.

The tests were performed binaurally with a closed headphone and selecting two RIRs at the approximate ear separation though no attempt has been made to correct for the head-related transfer function. The signals were fed to the headphone as if the subject was facing the reflecting wall. The RIRs were convolved with two monophonic signals, one being 4-s of noise shaped by the spectrum of male speech,[15] the other being a female speech sample of 2-s length.[16] The speech noise signals were normalized for equal energy of original and altered sample. There were a total of 13 listening positions with and 13 listening positions without direct sound. In the case without direct sound, the reflection from the object was the first wave front to arrive at the listening position.

A triple stimulus *AXB* listening test was employed where either of *A* and *B* were the samples with the original and altered RIR and *X* was the reference. The subjects were forced to choose whether *A* or *B* was equivalent to *X*. *X* was always the sample with the original RIR but this was not known by the test subjects.

The subjects were allowed and encouraged to take any perceived information such as loudness, spaciousness, coloration, and intelligibility into account. Because of the wide range of energy ratios, the difficulty of the task at hand varied by a great amount. The subjects were allowed to listen to a few randomly selected sets before the actual test for training purpose and also permitted to repeat any set during the test as many times as they wished to.

Fourteen subjects participated in the test, one of whom was later excluded due to inconsistent answers. All subjects stated to have normal hearing but their experience with listening tests varied by a great amount.

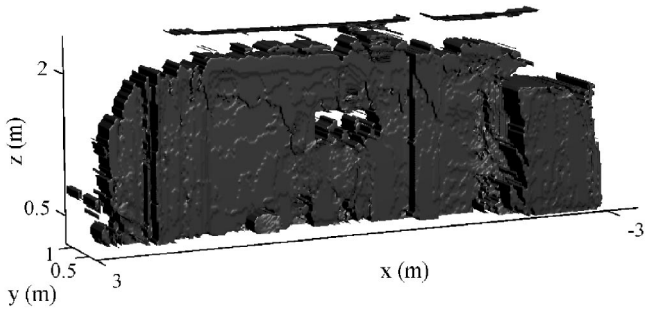

FIG. 25. Surface representation of the corridor wall without the distribution box.

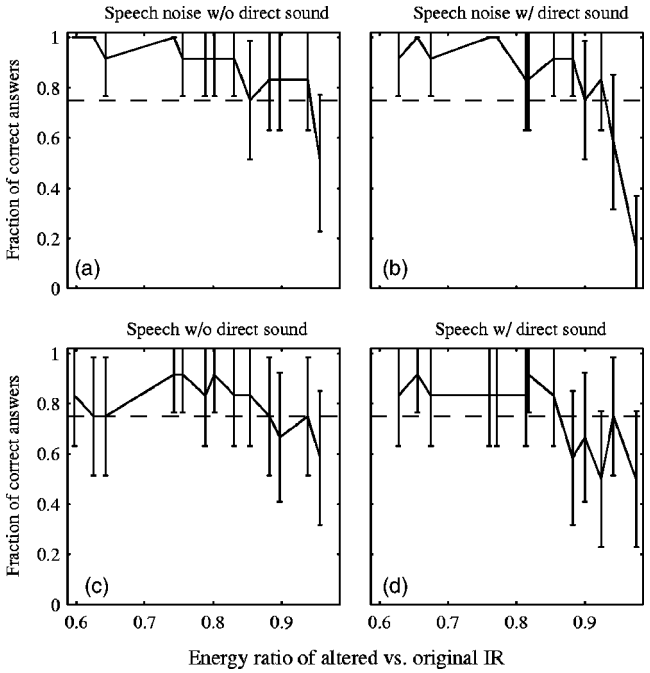

FIG. 26. Outcome of the listening tests expressed as mean between the subjects together with the 95% confidence bounds.

B. Results

The outcome of the listening tests is shown in Fig. 26 expressed as fraction of correct answers among the subjects versus energy ratio. The threshold, indicated by the dashed line, lies at 0.75 because statistically there is a 50% chance to guess the correct answer. The general trend observed in all graphs of Fig. 26 is a decrease from a fraction value close to unity towards a value of 0.5 for an increasing energy ratio. Values below 0.5 are a result of the small population size. Even though the confidence bounds are rather large, it appears safe to conclude that an energy ratio smaller than approximately 0.85 is required for a consistent detection of the objects' presence. Naturally, the energy in the RIR is a rather crude descriptor since the perception depends on the delay time and strength of individual reflections. Another observation is that with the direct sound it appears to be only slightly more difficult to detect the difference. It is possible that the comb filter effect occurring due to the short time delay between direct sound and reflection from the object facilitates the detection.

Looking at three positions from Fig. 23 in more detail, it was found that at position ''A'' none of the subjects were able to detect the difference. At position ''B'' the difference was detected reliably only in the case of speech noise without direct sound. At position ''C'' samples without direct sound allowed for the detection with both signals while with direct sound and speech noise the response was on the threshold. For the closet the differences were detected regardless whether with or without direct sound and speech or speech noise.

VII. CONCLUSIONS

A method has been proposed to map reflections in a room impulse response to the reflecting object resulting in an image of the enclosed space in terms of reflected pressure. The imaging process is based on inverse wave field extrapolation from the receiver to the object's position.

Simulations have been carried with the model of a simple shoe box enclosure. It was found that multiple reflections from more than one object lead to images of higher order. The position of those is generally outside the physical boundaries of the enclosed space. A comparison with measurements in a real enclosure of identical dimensions showed good agreement. Measurements performed in a corridor and a concert hall showed that considerable detail in the image can be obtained even at large distances.

By performing the inverse step of the imaging process it is possible to recreate the original impulse responses. This allows altering of objects in the image and their influence on the acoustics can (a) be calculated in a physically correct way and (b) be evaluated perceptually without actually altering the object. The process has been validated with two objects on one of the corridor's walls. Preliminary results from a listening test indicate that the influence of a reflecting object can only be perceived in its close proximity.

The applicability of the proposed method is limited because of the elaborate measurements associated with a planar array. Consequently, a way has to be found to deal with the inherent mapping of the three-dimensional space onto a plane due to linear array imaging. A possible approach is to exploit the fact that in a Cartesian system there are two normal particle velocity components on the array of which only one has currently been used. Further, the possibility of identifying and removing reflections involving a particular surface such as, e.g., the floor, would be highly beneficial.

ACKNOWLEDGMENTS

The authors would like to thank Eric Verschuur for his enthusiasm in helping them to apply the concepts of seismic imaging to room acoustics.

[1] J. Baan and D. de Vries, ''Auralization of sound fields by wave field synthesis,'' in *106th Convention* (Audio Engineering Society, Copenhagen, 1999).

[2] A. J. Berkhout, *Seismic Migration: Imaging of Acoustic Energy by Wave Field Extrapolation* (Elsevier, Amsterdam, 1984), Vol. 12.

[3] P. Hubral, J. Schleicher, and M. Tygel, ''A unified approach to 3-D seismic reflection imaging, Part I: Basic concepts,'' Geophysics **61**, 742–758 (1996).

[4] A. J. Berkhout, D. de Vries, and J. J. Sonke, ''Array technology for acoustic wave field analysis in enclosures,'' J. Acoust. Soc. Am. **102**, 2757–2770 (1997).

[5] J. S. Bradley, ''Comparison of concert hall measurements of spatial impression,'' J. Acoust. Soc. Am. **96**, 3525–3535 (1994).

[6] M. Barron and L. J. Lee, ''Energy relations in concert auditoriums. I,'' J. Acoust. Soc. Am. **84**, 618–628 (1988).

[7] J. D. Maynard, E. G. Williams, and Y. Lee, ''Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH,'' J. Acoust. Soc. Am. **78**, 1395–1412 (1985).

[8] A. J. Berkhout, D. de Vries, and P. Vogel, ''Acoustic control by wave field synthesis,'' J. Acoust. Soc. Am. **93**, 2764–2778 (1993).

[9] A. D. Pierce, in *Acoustics, an Introduction to its Physical Principles and Applications* (McGraw–Hill, New York, 1981), Chap. 4.

[10] A. J. Berkhout, in *Applied Seismic Wave Theory* (Elsevier, Amsterdam, 1987), Chap. 10.

[11] M. Tygel, J. Schleicher, and P. Hubral, ''A unified approach to 3-d seismic reflection imaging, part ii: Theory,'' Geophysics **61**, 759 (1996).

[12] S. Yon, M. Tanter, and M. Fink, ''Sound focusing in rooms: The time-reversal approach,'' J. Acoust. Soc. Am. **113**, 1533–1543 (2003).

[13] A. B. Weglein, ''Multiple attenuation: an overview of recent advances and the road ahead (1999),'' The Leading Edge **18**, 40–44 (1999).

[14] H. Kuttruff, in *Room Acoustics*, 4th ed. (Spon, London, 2000), Chap. 7.

[15] Instituut voor Zintuigfysiologie TNO, Spraakmateriaal behorende bij de test voor het meten van de spraakverstaanbaarheidsdrempel voor korte zinnen in stilte en in stationaire of fluctuerende ruis (English translation: Speech material accompanying a test for measuring the speech intelligibility threshold for short sentences in anechoic and stationary or fluctuating noise conditions), Audio CD, 1988.

[16] European Broadcasting Union, Sound Quality Assessment Material Recording for Subjective Tests, Audio CD, 1988, track 53.