3D localization of acoustic sources with a spherical array

Karim Haddad and Jørgen Hald

Brüel & Kjær Sound & Vibration Measurement A/S, Skodsborgvej 307, DK-2850 Nærum,
Denmark
khaddad@bksv.com

This paper describes a technique dedicated for the localization of acoustic sources in all directions and in the far-field. Classical beamforming techniques based on planar arrays provide an acoustic map restricted to a limited solid angle, but a spherical array does not have such a limitation since there is no preferential direction. In the processing called Spherical Harmonics Beamforming (SHB), the sound field on the sphere is decomposed with spherical harmonics functions, and then a corrected summation gives the acoustic contribution from a given direction. We have used a rigid spherical array, which has the advantage that cabling of microphones and integrated cameras can be hidden inside the sphere. A rigid surface also provides better numerical stability in connection with SHB. In this study, SHB is evaluated with respect to resolution and dynamic range. Simulated and experimental results are presented.

1   Introduction

Spherical arrays have recently attracted a lot of interests in many acoustic areas: room acoustics, sound field reproduction, noise source identification. In this last area, planar arrays are widely used for many purposes concerning outdoor or external measurements to localize acoustic sources. The main drawback of using such array shapes in confined spaces is the difficulty to separate sources coming from the front and from the rear of the microphone array. A back screen can help to solve this issue, but only for relatively high frequencies. Additionally the resolution of planar arrays decreases dramatically for important off-axis angles. Because the microphones are 'equally' distributed in all directions on a sphere, the spherical arrays do not suffer from such a variation of the resolution regarding the directions. The spherical array is therefore a very good tool to localize sources for interior noise, like in a car, an airplane or a room. It can be used as well for outdoor measurements.

We are considering here a processing technique called Spherical Harmonics Angularly Resolved Pressure (SHARP) based on the Spherical Harmonics Beamforming (SHB), [1,2]. The main benefit of the SHARP technique is that it is dedicated for source localization, and it provides a correct pressure contribution estimation for sources at the focus points, as is done also by traditional delay-and-sum beamforming. We apply this processing on signals delivered by microphones on a hard sphere, with integrated cameras to provide pictures which are overlaid with the acoustic map for a better understanding of the sources. Use of a hard sphere, instead of a transparent sphere, gives some advantages. Cabling of microphones and cameras are hidden inside the sphere. Additionally the hard sphere could represent a human head in a first approximation, creating the same level of scattering for example in a car cabin. Finally, in the context of the SHB, numerical instabilities appear in the calculation for particular frequencies when the sphere is transparent.

In this paper, we will introduce the SHB processing from which we will derive the SHARP technique. The main parameters to characterize an imaging device are the resolution and the dynamic range, which means, respectively, the capacity to separate two close sources and the level range for which we can detect the sources. We will describe these parameters in the context of the SHARP technique. Simulations and experiments will be finally presented to illustrate the performance of this technique.

2   The SHARP technique

2.1   Spherical Harmonics Beamforming

The basic idea of SHB is to map the sound field incident on the sphere thanks to the knowledge of the acoustic pressure on the spherical array. Consider a single monopole point source at the spherical coordinates (r_0, θ_0, ϕ_0) relative to the center of the sphere, producing the following pressure at the origin under free-field conditions (with no rigid sphere):

$$P_S = \frac{e^{j \cdot k \cdot r_0}}{k \cdot r_0} . \qquad (1)$$

The distribution of pressure produced by that point source on the rigid sphere is given by [2,3]:

$$p(a, \theta, \phi) = \sum_{n=0}^{\infty} R_n(kr_0, ka) \sum_{m=-n}^{n} Y_n^m(\theta_0, \phi_0)^* Y_n^m(\theta, \phi) \quad (2)$$

where k is the wavenumber and a is the sphere radius. The expressions $Y_n^m(\theta, \phi)$ are the spherical harmonics functions, and the symbol * represents the complex conjugate. The radial function $R_n(kr, ka)$ is given by the following expression, [4]:

$$R_n(kr, ka) = 4 \cdot \pi \cdot i \cdot h_n^{(1)}(kr) \cdot \left(j_n(ka) - d_n h_n^{(1)}(ka) \right)$$

Here j_n and h_n are the spherical Bessel and Hankel functions, and d_n is given by $j_n'(ka)/h_n^{(1)'}(ka)$, ' being the derivative symbol.

Eq.(2) can be re-written as:

$$p(a, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} p_n^m \cdot Y_n^m(\theta, \phi) \qquad (3)$$

where the coefficients p_n^m are the spherical Fourier transforms of the pressure $p(a, \theta, \phi)$ on the rigid sphere:

$$p_n^m = \int_0^{2\pi} \int_0^{\pi} p(a, \theta, \phi) \cdot Y_n^m(\theta, \phi)^* \cdot \sin(\theta) \cdot d\theta \cdot d\phi \quad (4)$$

Decomposing the sound pressure on the sphere with spherical harmonics is possible thanks to the following orthogonality relation:

$$\int_0^{2\pi} \int_0^{\pi} Y_\nu^\mu(\theta, \phi) \cdot Y_n^m(\theta, \phi)^* \cdot \sin(\theta) \cdot d\theta \cdot d\phi = \delta_{\nu,n} \cdot \delta_{\mu,m}$$

$$(5)$$

The comparison of the expressions Eq.(2) and Eq.(3) leads to:

$$p_n^m = R_n \cdot Y_n^m(\theta_0, \phi_0)^* \qquad (6)$$

The ideal beamforming output is a very sharp peak pointing to the direction of the source:

$$w(\theta, \phi, \theta_0, \phi_0) = \delta(\cos\theta - \cos\theta_0) \cdot \delta(\phi - \phi_0) \quad (7)$$

Or, in terms of spherical harmonics, thanks to the spherical Fourier transform (Eq.(3-4)):

$$w(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_n^m(\theta_0, \phi_0)^* \cdot Y_n^m(\theta, \phi) \quad (8)$$

Using Eq.(6) in this last expression leads to:

$$w(\theta, \phi, \theta_0, \phi_0) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{p_n^m}{R_n(kr_0, ka)} \cdot Y_n^m(\theta, \phi) \quad (9)$$

In practice, the pressure $p(a, \theta, \phi)$ on the sphere is known only at the microphone locations, which means that the spherical Fourier Transform Eq.(4) is approximated and that the summation in Eq.(9) is done until a degree N. The choice for N is discussed below.

Finally the spherical beamforming calculation is given by:

$$w_N(\theta, \phi) = \sum_{n=0}^{N} \sum_{m=-n}^{n} \frac{p_n^m}{R_n(kr_0, ka)} \cdot Y_n^m(\theta, \phi) \quad (10)$$

2.2 SHARP: pressure contribution determination

The spherical harmonics coefficients for the pressure Eq.(4) can be evaluated with an approximate Gaussian type of numerical integration:

$$p_n^m \approx \widetilde{p}_n^m = \sum_{i=1}^{Q} c_i \cdot p(a, \Omega_i) \cdot Y_n^m(\Omega_i)^* \qquad (11)$$

where Q is the number of sensors and $\Omega_i = (\theta_i, \phi_i)$. The microphone positions Ω_i and the associated weights c_i have been derived in such a way that they provide exact numerical integration over the full 4π solid angle of any angular function containing spherical harmonics of degree only up to N. It can be easily verified that this integration property can be true only if:

$$H_{nm\nu\mu} \equiv \sum_{i=1}^{Q} c_i \cdot Y_\nu^\mu(\Omega_i) \cdot Y_n^m(\Omega_i)^* = \delta_{\nu,n}\delta_{\mu,m} \quad (12)$$

for $\nu \leq N, n \leq N$.

Eq.(12) is a discrete orthogonality relation, analog to Eq.(5). Replacing the expression of p_n^m by its approximate version Eq.(11), the spherical beamforming expression Eq.(10) becomes:

$$w_N(\Omega) = A \cdot \sum_{n=0}^{N} \sum_{m=-n}^{n} \frac{Y_n^m(\Omega)}{R_n(kr_0, ka)} \cdot \sum_{i=1}^{Q} c_i \cdot p(a, \Omega_i) \cdot Y_n^m(\Omega_i)^* \qquad (13)$$

A is the correction factor to be determined in order to calculate the pressure contribution correctly. As the summation in the spherical beamforming calculation Eq.(13) is performed until a degree N, the calculated level for a source in a particular direction should be corrected accordingly. To determine the correction factor A, we consider again the monopole point source at the position (r_0, Ω_0) with $\Omega_0 = (\theta_0, \phi_0)$. This monopole produces a sound pressure at the microphone positions given by Eq.(2):

$$p(\Omega_i) = \sum_{n=0}^{\infty} R_n \sum_{m=-n}^{n} Y_n^m(\Omega_0)^* Y_n^m(\Omega_i) \qquad (14)$$

Here and below, the arguments of R_n are omitted for simplicity.

Substitution of Eq.(14) in Eq.(13) and making use of Eq.(12) leads to:

$$w_N(\Omega) = A \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} Y_\nu^\mu(\Omega) \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{R_n}{R_\nu} Y_n^m(\Omega_0)^* H_{nm\nu\mu} \,(15)$$

Averaging the beamforming output Eq.(15) with $\Omega = \Omega_0$ over all angles, and using the orthogonality relation (5) leads to:

$$\overline{w}_N = \frac{1}{4\pi} \iint_{4\pi} w_N(\Omega_0) d\Omega_0 = \frac{A}{4\pi} \sum_{\nu=0}^{N} \sum_{\mu=-\nu}^{\nu} H_{\nu\mu\nu\mu} \quad (16)$$

Using Eq.(12), we obtain:

$$\overline{w}_N = \frac{A}{4\pi}(N+1)^2 \qquad (17)$$

To get the desired pressure contribution, $\overline{w}_N = P_S$, from Eq.(1), the scaling factor A in Eq.(13) must have the following value:

$$A = \frac{4\pi \cdot e^{j \cdot k \cdot r_0}}{(N+1)^2 \cdot k \cdot r_0} \qquad (18)$$

It is also possible to perform sound intensity scaling on the spherical beamforming output in such away that area-integration provides a good estimate of the sub-area sound power. For that purpose we follow the procedure described in [4]. This calculation is carried out below, once we have defined the size of the main lobe in the next section.

3 Capabilities for Source localization

In this part, we discuss the capabilities of the SHB technique for the localization of noise sources in terms of resolution and dynamic range.

3.1 Spatial resolution performance

Substituting Eq.(6) in Eq.(10) and using some calculation properties of spherical harmonics [3] leads to:

$$w_N(\Omega) = \frac{N+1}{4\pi(\cos\Theta - 1)}[P_{N+1}(\cos\Theta) - P_N(\cos\Theta)]$$

$$(19)$$

This is the expression of the Point Spread Function (PSF) for a unit amplitude source at Ω_0, Θ being the angle between the vector directions Ω and Ω_0 [3]. Figure 1 shows this point spread function for different degrees N from 1 to 50.
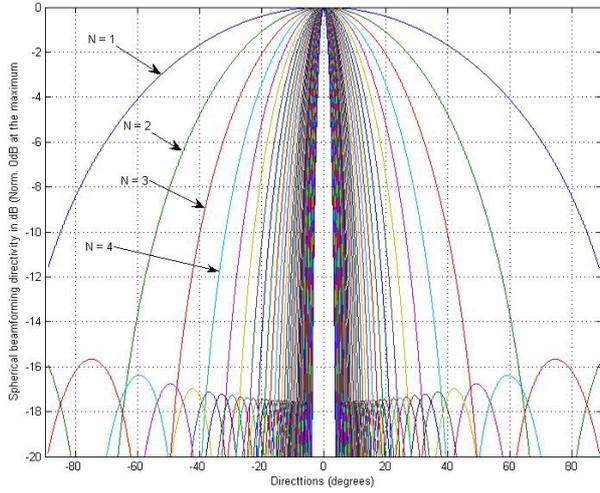
Fig. 1 Point spread functions for SHB ($|w_N|$ in dB) for different values N from 1 to 50. The levels are normalized to 0 dB for each *N*.

Figure 1 illustrates the fact that the resolution depends on the degree *N*: higher values *N* provide better resolution, since the main lobe representing a single source direction is narrower.

We have seen previously (paragraph 2.2) a limitation for *N* due to the number and the positions of the microphones. Another limitation for *N* is linked to the radial function R_n which drops down after *N = ka* [3], and then limiting the accuracy of the calculations. On the other hand, the Rayleigh resolution (full width of the main lobe) is well approximated by [3]:

$$2 \cdot \Theta_0 = \frac{2 \cdot \pi}{N} \qquad (20)$$

for *N* between 4 and 40.

Finally an approximation of the resolution is given by:

$$2 \cdot \Theta_0 = \frac{2 \cdot \pi}{ka} \qquad (21)$$

for $ka \geq 4$.

Below N = 4, the calculation of the main lobe width gives:

$2 \cdot \Theta_0 = 2 \cdot \pi/1.65$ for *N* =1, $2 \cdot \Theta_0 = 2 \cdot \pi/1.45$ for *N* = 2, and $2 \cdot \Theta_0 = 2 \cdot \pi/3.29$ for N = 3.

Because of the wavenumber *k*, the spherical beamforming resolution is in practice dependent on frequency: resolution improves with increasing frequency. We also notice that resolution improves with increasing radius of the sphere.

Now we know the width of the main lobe, therefore we can perform the sound intensity scaling, following the procedure described in [5]. The scaling factor α needed to obtain the intensity scaled spherical beamforming output is defined as:

$$w_{N,I}(\Omega) = \alpha \cdot |w_N(\Omega)|^2 \qquad (22)$$

This quantity $w_{N,I}$ is defined in such a way that its integral over the main lobe equals half of the radiated power P_a, the power radiated into the hemisphere containing the array:

$$\frac{1}{2}P_a = \int_{Mainlobe} w_{N,I} dS = \int_{Mainlobe} \alpha \cdot |w_N(\Omega)|^2 dS \qquad (23)$$

Considering the expression of the sound field Eq.(1) created by a monopole, we obtain for the half power:

$$\frac{1}{2}P_a = \frac{\pi}{\rho \cdot c \cdot k^2} \qquad (24)$$

where ρ is the fluid density, c is the speed of the sound.

The PSF Eq.(19) should take into account the scaling factor *A* given by Eq.(18):

$$w_N(\Omega) = \frac{e^{jk \cdot r_0}}{(N+1)kr_0(\cos\Theta - 1)}[P_{N+1}(\cos\Theta) - P_N(\cos\Theta)] \qquad (25)$$

So the relation Eq.(23) leads to:

$$\frac{1}{2}P_a = \alpha \frac{2\pi}{k^2} \cdot F(N) \qquad (26)$$

With

$$F(N) = \frac{1}{(N+1)^2} \int_{\varepsilon}^{\Theta_0} \frac{[P_{N+1}(\cos\Theta) - P_N(\cos\Theta)]^2}{(\cos\Theta - 1)^2} \sin\Theta \cdot d\Theta$$

ε is a non-zero very small value in order to avoid the division by 0 for the integrand. The part of this equation depending on N, *F(N)*, is calculated numerically and is represented on the figure 2 below. This function *F(N)* is fairly well approximated by a function $G(N) = 1.17 / N^2$ and also represented on the figure 2 in red. Substituting in Eq.(26) P_a by its expression Eq.(24), and replacing *F(N)* by *G(N)* lead to the following expression of the scaling factor:

$$\alpha = \frac{N^2}{1.17(2\rho \cdot c)} \qquad (27)$$

Considering the relation *N = ka*, we also obtain:

$$\alpha = \frac{2\pi^2}{1.17} \cdot \frac{1}{\rho \cdot c} \cdot \left(\frac{a}{\lambda}\right)^2 \qquad (28)$$

where λ is the wavelength.

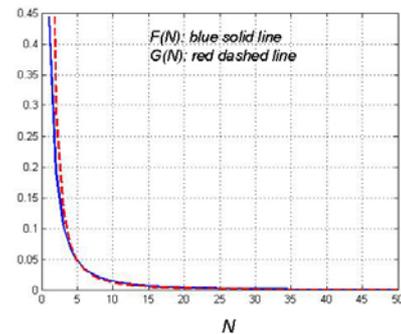

Fig. 2 Numerical integration *F(N)* (see Eq.(26)) for different values N (blue curve). The red curve, *G(N)*, is an approximation of the blue curve.

3.2 Dynamic range

Figure 1 shows also the presence of other lobes which are not representative of sources, but they are created by the processing. These side lobes may hide low level sources. That is why the Maximum Sidelobe Level (MSL) defines

the dynamic range of the system; the MSL should be as low as possible for a better dynamic range. On figure 1 we can see that the MSL is between -16 dB for the lower degrees (from $N = 2$) and -17.5 dB (until $N = 50$).

The PSF expression Eq.(19) is correct if the arrangement of the microphones ensures that the discrete orthogonality relation (12) holds up to the degree N. Nevertheless it is possible to get good results even if this condition is not fully satisfied. As an example, we consider on Figure 3 the localization of a source at 5 kHz with a 36-channels spherical array with radius 0.0975 m. We have $ka \approx 9$, and we apply the spherical beamforming calculation until degree 10, which is higher than the degree N for which the orthogonality relation is fulfilled (for this spherical array the orthogonality relation holds until $N = 5$).
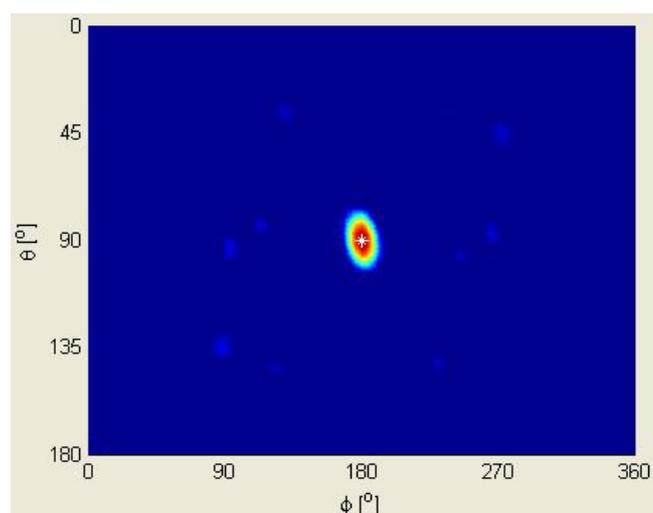

Fig. 3 2D plot of the point spread function for a source at 5 kHz. X-axis is the azimuth, Y-axis is the elevation. The processing is done until degree $N = 10$, dynamic scale: 8 dB

The simulation illustrated in figure 3 shows the main lobe at the center of the graph plus a few sidelobes. The MSL is -8.7 dB, which is higher than the value we get from figure 1 (about -17 dB), but still interesting for localization purpose.

The microphone distribution on the sphere plays a role for the dynamic range, but this is also the case for the degree N. To illustrate this point, we consider below the localization of a source at 2 kHz, and we apply the spherical beamforming processing with various maximum degrees N.
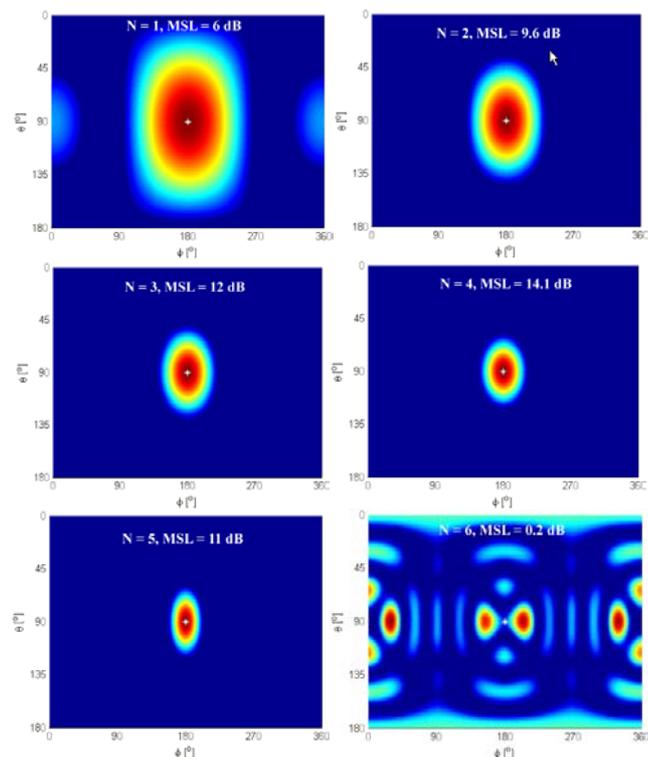

Fig. 4 2D plots of the point spread function for a source at 2 kHz for N = 1 to 6. X-axis is the azimuth, Y-axis is the elevation. Dynamic scale for all graphics: 10 dB

The figure 4 shows a PSF maps for different value N. We notice that there is an optimum value for the MSL at $N = 4$. Not surprisingly, it corresponds to $ka \approx 4$. Quickly after this optimum, the MSL decrease dramatically (N = 6), then the processing is no more able to localize the source. For the different microphone arrays we have tested, this optimum is always around $N = ka$, sometimes it is at $N = ka + 1$ or $ka - 1$.

These pictures illustrate also the evolution of the resolution with the degree N described in the previous section: the size of the main lobe is smaller with the increasing degrees.

4   Simulations and experiments

In this section we propose to compare simulated and experimental results for different frequencies: 500, 2000 and 6000 Hz. The experimental tests were performed in anechoic room with a spherical array including 36 microphones and 12 cameras (B&K type 8606).

Fig. 5 36 channels spherical array with 12 cameras

At 500 Hz (figure 5), the displayed range is 10 dB. We have very similar spot between the two pictures. For the frequencies 2 and 6 kHz (respectively figures 6 and 7), we have chosen the level range in order to show some side lobes. We can see in both cases that the size of the spot is the same between experimental and simulated results, but also the positions of the side lobes (except for one sidelobe at 2 kHz on the figure 6, but it is more probably a reflection from the ground).
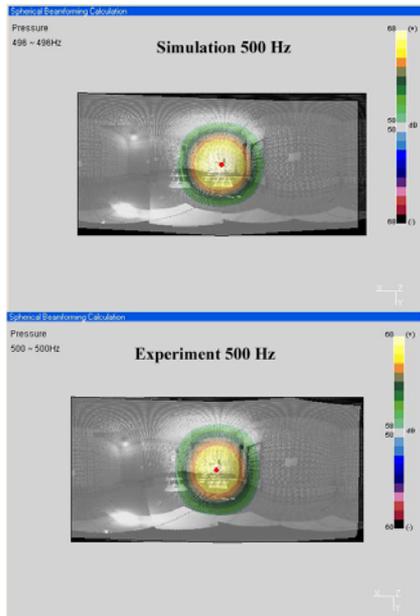

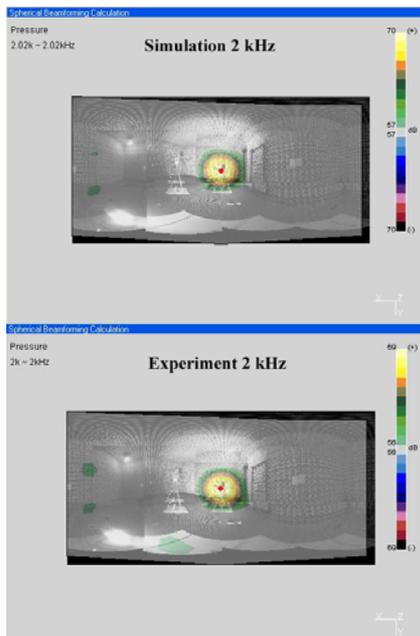Fig. 5 Acoustic maps at 500 Hz, simulation and experiment

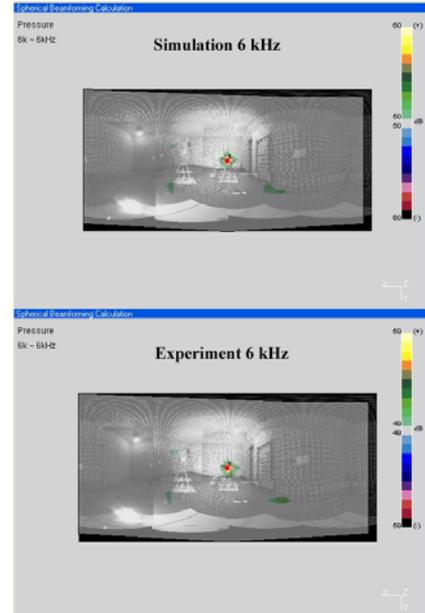Fig. 6 Acoustic maps at 2 kHz, simulation and experiment

Fig. 7 Acoustic maps at 6 kHz, simulation and experiment

5   Conclusion

In this article, we have presented an extension of the Spherical Harmonics Beamforming (SHB) called SHARP (Spherical Harmonics Angularly Resolved Pressure) dedicated to noise source identification and including a correction factor in order to get an estimate of the source levels. Also we have derived a sound intensity scaling from the SHARP output to provide a good estimate of the sub-area sound power. Finally the examples show experimental results very close to the simulations.

References

[1] P. M. Juhl, S. O. Petersen, J. Hald, "Localizing sound sources in 3-D space using spherical harmonic beamforming", *Inter-noise RIO*, Brazil (2005)

[2] J. Hald, J. Mørkholt and J. Gomes, "Efficient Interior NSI based on various Beamforming methods for overview and conformal mapping using SONAH holography for details on selected panels", *Proceedings of SAE 2007 Noise & Vibration Conference*, Paper 07NVC-334 (2007)

[3] B. Rafaely, "Plane-wave decomposition of the sound field on a sphere by spherical convolution", *J. Acoust. Soc. Am.* 116, 2149-2156 (2004)

[4] J. J Bowman, T. B. A. Senior, P. L. E. Uslenghi, "Electromagnetic and acoustic scattering by simple shapes", *Hemisphere Publishing Corporation*, 1987

[5] J. Hald, "Combined NAH and Beamforming Using the Same Array", *Brüel & Kjær Technical Review*, No 1-2005