

Spherical Beamforming with Enhanced Dynamic Range

Jorgen Hald

Brüel & Kjær Sound/Vib Meas A/S

ABSTRACT

Beamforming with an array of microphones on a sphere is an attractive tool for doing noise source localization in cabin environments. In order to achieve acceptable angular resolution, the array must have some minimum diameter, implying that many microphones are needed to obtain low sidelobe level over the frequency range of interest. For electric cars there is an increased need to cover high frequencies. The present paper describes a method to significantly reduce the sidelobe level over a broad frequency range relative to Spherical Harmonics Beamforming (SHB). For each focus point, a set of Finite Impulse Response (FIR) filters are optimized to minimize the highest sidelobe in a Filter And Sum (FAS) beamformer, while maintaining sensitivity at the focus point and limiting the White Noise Gain (WNG). Similar approaches have been published, but a main issue in connection with applications for noise source mapping is the computational load in performing the optimization of FIR filters for each focus point. The paper describes a method to handle that challenge. Basically, filter coefficient are pre-optimized for a pre-defined 3D mesh and stored in a database. When focusing has to be performed at an arbitrary point in space, an interpolation is performed, typically on the filter vectors related to a set of neighboring mesh points. Results are presented for simulated and real measurements. These show that the optimized FAS method provides very significant improvement relative to SHB on sidelobe suppression and thus on the dynamic range available for source identification.

CITATION: Hald, J., "Spherical Beamforming with Enhanced Dynamic Range," *SAE Int. J. Passeng. Cars - Mech. Syst.* 6(2):2013, doi:10.4271/2013-01-1977.

INTRODUCTION

Spherical microphone arrays are already widely used in the automotive industry for localizing sources of noise in cabin environments because of their ability to provide resolution in all directions. Typically, these systems perform so-called Spherical Harmonics Beamforming (SHB), where a spherical harmonics representation of the incident sound field is extracted from measured pressure data. The spherical harmonics representation is then used to achieve a directional resolution of the incident field, see for example references [1], [2], [3] and [4]. Use of a rigid spherical shell, with the microphones flush mounted, has the advantage of avoiding numerical instability at certain frequencies, [2], and of offering an interior space for arranging cables and electronics. The major limitations are 1) the fairly small dynamic range due to sidelobes, and 2) the poor angular resolution at low frequencies.

Resolution will be poor at low frequencies due to the fact that the array diameter will be small compared to wavelength. The angular resolution is approximately π/N , where N is highest spherical harmonics degree extracted and used in SHB, [2]. For N up to approximately ka , where k is

wavenumber and a is the radius of the sphere, the process is numerically stable, while for $N > ka$ the calculation will quickly involve extreme amplification of vanishing signal components, [2]. At the lowest frequencies, however, values of N up to $ka + 2$ can be used, which can significantly improve resolution relative to truncation around $N \sim ka$. Going for even higher resolution is an attempt to obtain super-resolution, which will be numerically unstable, unless some additional information (beyond the array pressure data) can be added about the source distribution.

The present paper focuses on the challenge of improving the sidelobe suppression in a robust and computationally efficient way. A broad range of publications exist on general methods to control the sidelobes, for example reference [5] dealing with adaptive beamforming and reference [6] describing an optimized broadband Filter And Sum (FAS) beamforming technique closely related to the method of the present paper. Reference [7] gives a good overview of methods for optimizing specifically the frequency-domain processing of spherical array measurement. But a main issue in connection with applications for broadband noise source mapping is the computational load in performing the optimization of the processing parameters for each focus

point and each frequency, in particular when a dense mesh of “sidelobe control positions” is needed during the optimization to make sure that the highest sidelobe is minimized. The present paper describes a method to handle that challenge. Results are presented for simulated measurements on monopole point sources and for real measurements in an anechoic room and in a wind-tunnel.

THEORY

The mathematical description will be entirely in the frequency domain, using implicit time dependence $e^{j\omega t}$ and wavenumber $k = \frac{\omega}{c} = \frac{2\pi f}{c}$, where c is the propagation speed of sound, and f is the frequency. The first two sections describe special implementations of SHB, with the output scaled to represent directional contributions to the free-field sound pressure at the center of the sphere. The first of these was presented in reference [4], while the second contains a modification to reduce sidelobes at low frequencies and to achieve smoother frequency responses of the beamformer. The smoother frequency responses of the second implementation makes the associated microphone weights better suited for use as starting point in the optimization to be presented in the third section.

Spherical Harmonics Angularly Resolved Pressure (SHARP)

The expression for the SHARP beamformer, truncated at maximum harmonics degree N , is given in reference [4] as:

$$B_N(r, \theta, \varphi) = \frac{4\pi e^{-jkr}}{(N+1)^2 kr} \cdot \sum_{n=0}^N \sum_{m=-n}^n \frac{Y_n^m(\theta, \varphi)}{R_n(ka, kr)} \sum_{i=1}^I w_i P_i Y_n^m(\theta_i, \varphi_i), \quad (1)$$

where a is the sphere radius, (θ_i, φ_i) are the angles of the I microphones with corresponding weights w_i and measured sound pressure values P_i , Y_n^m are the spherical harmonics functions, and the radial functions R_n are defined as

$$R_n(ka, kr) = -4\pi j h_n(kr) \left[j_n(ka) - \frac{j'_n(ka)}{h'_n(ka)} h_n(ka) \right]. \quad (2)$$

Here, h_n are the spherical Hankel functions of the second kind representing outwards propagating waves, j_n are the spherical Bessel functions, and the prime represents derivative with respect to the argument.

To derive a version of the [expression \(1\)](#) without the summation over spherical harmonics order we rotate the coordinate system to have $(\theta, \varphi) = (0, 0)$ in the focus direction. In that case we can use the following relation [8]

$$Y_n^m(0,0) = \begin{cases} \sqrt{\frac{2n+1}{4\pi}} Q_n(\cos(0)) = \sqrt{\frac{2n+1}{4\pi}} & \text{for } m = 0 \\ 0 & \text{for } m \neq 0, \end{cases} \quad (3)$$

Q_n being the Legendre polynomial of degree n . This identity introduces a factor of zero in all contributions with $m \neq 0$ in the summation of [Eq. \(1\)](#), which simplifies to:

$$B_N = \frac{e^{-jkr}}{(N+1)^2 kr} \cdot \sum_{n=0}^N \frac{2n+1}{R_n(ka, kr)} \sum_i w_i P_i Q_n(\cos(\psi_i)). \quad (4)$$

Here, the elevation angle ψ_i of microphone number i in the rotated coordinate system is equal to the angle from the focus direction (θ, φ) to the direction (θ_i, φ_i) of microphone number i in the original coordinate system. The cosine of that angle can be easily calculated from these original directions as, [2]:

$$\cos(\psi_i) = \cos(\theta) \cos(\theta_i) + \sin(\theta) \sin(\theta_i) \cos(\varphi - \varphi_i), \quad (5)$$

so in practice no coordinate rotation is needed, just the calculation of $\cos(\psi_i)$.

Notice that in [Eq. \(4\)](#) the factor $(2n+1)$ is just the number of spherical harmonics of degree n . Its appearance can probably be explained by the fact that all spherical harmonics components of degree n appear at $m=0$ after the coordinate system rotation. After the summation in [Eq. \(1\)](#), a division is performed by $(N+1)^2$, which is the total number of spherical harmonics of degree up to N . So the division by $(N+1)^2$ is a division by the sum of all the factors $(2n+1)$. This observation is utilized in the following section to replace the truncation in spherical harmonics degree with a smooth cut-off.

A remaining issue in connection with the SHARP beamformer of [Eq. \(4\)](#) is the definition of the upper limit N on the spherical harmonics degree used in the summation. To define that we need the highest spherical harmonics degree for which the microphone positions (θ_i, φ_i) and weights w_i provide accurate integration over the sphere. Typically, N_{design} will be the largest integer, such that the number $(N_{\text{design}} + 1)^2$ of spherical harmonics of degree up to N_{design} does not exceed the number I of microphones. This holds true for the two arrays to be dealt with in the present paper. At the frequency $f = f_{\text{design}}$ where $ka = N_{\text{design}}$, the average spacing between the microphones on the sphere is slightly smaller than half wavelength. The following value of N in [Eq. \(4\)](#) has been found to provide a good compromise between resolution on one side and sidelobe and noise suppression on the other side:

$$N = \begin{cases} [ka] + 1 & \text{for } [ka] + 1 \leq N_{\text{design}} \\ [ka] & \text{otherwise,} \end{cases} \quad (6)$$

$[\cdot]$ representing rounding of a floating point number to the nearest integer.

Smooth Spherical Harmonics Cutoff instead of Truncation in SHARP

Definition of the truncation function $F_{n,N}$ as

$$F_{n,N} \equiv \begin{cases} 1 & \text{for } n \leq N \\ 0 & \text{for } n > N. \end{cases} \quad (7)$$

allows Eq. (4) to be rewritten in the form:

$$B_N = \frac{e^{-jkr}}{kr} \frac{1}{\sum_n (2n+1) F_{n,N}} \cdot \sum_{n=0}^{\infty} \frac{(2n+1) F_{n,N}}{R_n(ka, kr)} \sum_i w_i P_i Q_n(\cos(\psi_i)) \quad (8)$$

Here, we can replace the truncation function $F_{n,N}$ by a smooth cut-off function $F_{n,\epsilon}$ defined in the following way:

$$F_{n,\epsilon} \equiv \frac{|R_n(ka, kr)|^2}{|R_n(ka, kr)|^2 + \epsilon^2 |R_0(ka, kr)|^2}. \quad (9)$$

The factor $F_{n,\epsilon}$ will serve as regularization in connection with the division in Eq. (8) by the radial functions R_n , which decay quickly towards zero with increasing n for $n > ka$. The small positive regularization constant ϵ defines indirectly the cutoff point in the spherical harmonics summation. Substitution of $F_{n,\epsilon}$ for $F_{n,N}$ in Eq. (8) leads to:

$$B_\epsilon = \frac{e^{-jkr}}{kr} \frac{1}{\sum_n (2n+1) F_{n,\epsilon}} \cdot \sum_{n=0}^{N_{\max}(ka)} \frac{(2n+1) F_{n,\epsilon}}{R_n(ka, kr)} \sum_i w_i P_i Q_n(\cos(\psi_i)) \quad (10)$$

where a frequency dependent maximum harmonics degree $N_{\max}(ka)$ has been introduced. Practical experience has shown the following choice to work well:

$$N_{\max}(ka) \equiv \begin{cases} N_{\text{design}} & \text{for } ka \leq N_{\text{design}} \\ [ka] & \text{for } ka > N_{\text{design}} \end{cases} \quad (11)$$

In all applications of the beamformer formulation (10) in the present paper, the regularization constant ϵ has been set to 0.1, corresponding to a dynamic range equal to 20 decibel.

Optimized Filter And Sum (FAS) Spherical Beamforming

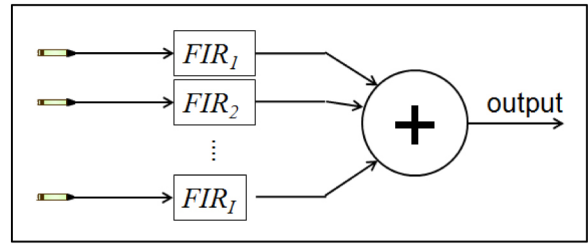


Figure 1. Principle of FAS spherical beamforming

Figure 1 illustrates the simple principle of the FAS beamformer for a single focus direction: The microphone signals are filtered through individual Finite Impulse Response (FIR) filters, and the output from these filters are added. The FIR filters are here optimized to meet the following requirements:

1. Minimize Maximum Sidelobe Level (MSL)
2. Retain sensitivity at the focus position
3. Limit the White Noise Gain (WNG).

A detailed description can be found in reference [9]; here only a rather short description will be given.

Denoting by $H_i(\mathbf{r}_0)$ the frequency response of the FIR filter applied to microphone number i in order to focus the beamformer at $\mathbf{r}_0 = (r, \theta_0, \varphi_0)$, the beamformer output $\mathbf{B}(\mathbf{r}_0)$ can be written simply as

$$B(\mathbf{r}_0) = \sum_{i=1}^I H_i(\mathbf{r}_0) P_i. \quad (12)$$

At the considered frequency we define a mesh of "sidelobe control positions" $\mathbf{r}_l = (r, \theta_l, \varphi_l)$ $l = 1, 2, \dots, L$, at the same radial coordinate r but covering with "sufficient density" all directions outside the mainlobe of the spherical harmonics beamformer of Eqs. (10-11). Some of these points must be close to the mainlobe in order to prevent a widening of the mainlobe during the optimization. For the derivation of the frequency responses H_i we consider monopole point sources with amplitudes S_l $l = 0, 1, 2, \dots, L$, at the focus position ($l = 0$) and at the sidelobe control positions ($l > 0$), and we define a set of transfer functions $G_i(r_l)$ from each one of these points to all microphones on the sphere. The point sources are scaled in such a way that S_l represents directly the free-field sound pressure contribution from source number l at the center of the sphere, i.e. the pressure without the rigid sphere in place. With this definition, the sound pressure at microphone number i can be expressed in the following way: