Comparison Between the Spherical Harmonics Beamforming and the Delay-and-Sum Beamforming

Frederico Heloui de ARAUJO¹; Fernando Augusto de Noronha Castro PINTO²

¹, ² Universidade Federal do Rio de Janeiro, Brazil

ABSTRACT

The beamforming algorithm using spherical microphones arrays are nowadays widely used to identify sound sources in several types of environments. The biggest advantages of spherical arrays in relation to plane arrays is the possibility to derive the same beampattern for every direction in space and the possibility to decompose the arriving sound wave in spherical harmonics instead of the use of the traditional beamforming algorithm, the delay-and-sum. This paper makes a brief introduction to spherical harmonics beamforming (SHB) and it makes the comparison between the SHB and the delay-and-sum algorithm, using a spherical array with 20 microphones. The measurements were made inside an isolated wooden box and it shows the frequency range where the superiority of the SHB is notable.

Keywords: Beamforming, Delay-and-Sum, Spherical Harmonics

1. INTRODUCTION

Noise is any unwanted sound, it can interfere with the communication or cause discomfort or health problems. It can be also an indication of malfunction of some mechanism. The first step to eliminate it is identify it, guaranteeing the well-being of people around and the proper operation of the equipment. The beamforming is one of the most used technique for such sound identification. It is based on microphone arrays allowing analyze waves coming from several directions independently.

The conventional algorithm, a.k.a. delay-and-sum, is used since the World War I (1) and have been quickly developed in the last decades boosted by the advance in the acquisition hardware and processing software. Using spherical microphones arrays, the Spherical Harmonic Beamforming (SHB) appears as an alternative for the delay-and-sum. The SHB is based on the spherical harmonic decomposition of the sound pressure along a spherical surface (2, 3, 4) taking into account the spatial sampling in the sphere (5, 6, 7).

In this paper, a brief review of the delay-and-sum and the SHB are provided, and then, is made a comparison between both techniques. For such, a spherical array with 20 microphones was used for measurements inside an isolated wooden box.

2. Spherical Harmonic Decomposition

The spherical harmonics are orthonormal base of functions that are obtained by solving the wave equation for spherical coordinates. The degree $n$ and order $m$ spherical harmonic is given by:

$$Y_n^m(\theta, \varphi) = \sqrt{\frac{2n+1}{4\pi} \frac{n-m}{n+m}} L_n^{|m|}(\cos\theta)e^{im\phi}$$  \hspace{1cm} (1)

where $\theta$ and $\varphi$ are the elevation and azimuth angles, respectively, and $L_n^{|m|}$ is the associated Legendre polynomial.

¹ freddearaujo@gmail.com
² fcpinto@ufrj.br
Considering a function $f(\theta, \phi)$ which is entirely defined and integrable on the surface, $\Omega$, of a unit sphere. It is possible to decompose that function in spherical harmonics in the following way:

$$f(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \overline{f}_{nm} Y_n^m(\theta,\phi)$$

(2)

The overline $f$ is the degree $n$ and order $m$ coefficient, it can be obtained by:

$$\overline{f}_{nm} = \frac{1}{\Omega} \int_{\Omega} f(\theta,\phi) Y_n^m(\theta,\phi) \ast d\Omega$$

(3)

where $\ast$ is the conjugated complex representation.

3. Beamforming

The beamforming is a spatial filter technique that uses microphones arrays, where is possible to control the sensitivity to the incidence of sound waves in a certain direction or point in space, i.e., it allows the spatial identification of the sound sources.

Each sensor in the array occupies a different position in space, then it has a different time signals that are a function of the distance between the sensor and the sound source. To drive the array for certain direction in space, where you want to search the source, different time delays is applied in each signal in order to amplify the signal coming from that direction. Therefore, when in the chosen direction there is an arriving sound source, all signals will be in phase, therefore will have a much stronger response than if the array had analyzed a region with no source.

3.1 Delay-and-sum

Modeling a plane wave travelling in the direction $\hat{s}$, the sound pressure $p$ will be a function of the position $\hat{x}$ and time $t$, i.e.

$$p(\hat{x},t) = f(\hat{x} \cdot \hat{s} - c_0 t)$$

(3)

where $c_0$ is the speed of sound. For each microphone $j$, where $j=1,2,\ldots,M$, the sound pressure will be only time-dependent. The beamforming response, $b(t)$, is given by the sum of the $M$ signals delayed to each other:

$$b(\hat{s},t) = \frac{1}{M} \sum_{j=1}^{M} p_j(t - \Delta_j(\hat{s}))$$

(3)

The symbol $\Delta_j$ represents the delay of the sensor $j$ compared with the reference point in the array. For plane waves, the delay is given by:

$$\Delta_j = \frac{\hat{x}_j \cdot \hat{s}}{c_0}$$

(3)

where $\hat{x}_j$ is the position of the microphone $j$ (Erro! Fonte de referência não encontrada.). The result is normalized by the factor $1/M$. When selecting adequate delays, the signal from the focused region is amplified while the signal from the other regions is diminished.
If the spherical waves model were assumed, it would be necessary to use a weight factor for each microphone signal to correct the decay of the amplitude.

### 3.2 Spherical Harmonic Beamforming

The advantage of the spherical arrays is the possibility of having the same beampattern for all the directions in space.

The ideal beampattern is a Dirac delta, where the direction of interest have a gain and all the others have none.

The generic spherical harmonic beamforming expression for a driven direction \((\theta_s, \phi_s)\) is given by:

\[
B(ka, \theta_s, \phi_s) = \int w(ka, \theta_s, \phi_s, \theta_s, \phi_s) P(\theta_s, \phi_s) d\Omega
\]

where \(w\) is a ponderation that drive the algorithm for the direction \((\theta_s, \phi_s)\), \(k\) is the wave number, \(a\) is the sphere radius and \(P\) is the Fourier Transform of the pressure over the sphere surface.

Using the spherical harmonic decomposition in the ideal beampattern can be expressed by the spherical harmonics:

\[
\delta(\theta - \theta_s)\delta(\phi - \phi_s) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \delta_{nm} Y_n^m(\theta, \phi)
\]
\[ \overline{\delta}_{nm} = Y_n^m(\theta, \varphi) \]  

Therefore, the ponderation is given by (2): 
\[ w(ka, \theta, \varphi, \theta_s, \varphi_s) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{\overline{\delta}_{nm}}{4 \pi^2 b_n(ka)} Y_n^m(\theta, \varphi) \]  

where \( b_n \) is mode strength, given by: 
\[ b_n(ka) = j_n(ka) \]  

for open spheres and 
\[ b_n(ka) = j_n(ka) - \frac{j'_n(ka)}{h'_n(ka)} h_n(ka) \]  

for rigid spheres. Where \( j_n \) and \( h_n \) is the Bessel spherical function and Henkel spherical function, respectively, \( j'_n \) and \( h'_n \) are their derivatives. The second term in Equation 9 is due the scattering around the sphere. The mode strength can be seen in Figure 3.

![Figure 3 - Mode strength for rigid and open sphere](image)

The disadvantage of a rigid sphere is that it interferes in the sound field. But this scattering can be easily handled considering not having further reflections caused by other objects of the scattered waves for the SHB. But this calculation is not so simple for the delay-and-sum algorithm. This effect may become more significant with large spheres.

The advantage of the rigid sphere is the numerical stability. Because the inversion of \( b_n \) must be taken in consideration, as seen further in this section. For the open sphere the mode strenght is zero or tends to zero for some values.

In practice, the pressure is not known along the entire sphere surface, only in discrete points where the microphones are positioned. For this reason, it is not possible expand the ideal beampattern to infinity orders. Generally, to ensure a truncation of the harmonic spherical decomposition in order N, it is necessary M microphones (6), where:
For greater N, the beampattern will be closer from the Dirac delta. The Figure 4 shows beampattern truncated for some N values.

\[ M \geq (N + 1)^2 \]  

(10)

Figure 4 – Beampattern truncated

The upper limit frequency is settled for the argument \( ka \) when \( b_{N+1}(ka) \) is no longer less then -10dB, as can be seen in Figure 3 (4).

4. Experimental Set

A spherical array with 20 microphones positioned on the vertices of a dodecahedron was built in PLA (Polylactic Acid) for this application (Figure 5). This array allows the spherical harmonics decomposition up to order 2. Therefore, the spatial resolution is about 105°, which means that the algorithm cannot separated two waves arriving from directions that form between themselves an angle less than 105°. This resolution is acceptable for mapping single sources. The theoretical frequency range for the SHB is from 550 up to 1000 Hz. For the delay-and-sum, it can be extend up to 2000 Hz, where the spatial aliasing appears.

Figure 5 – Spherical array with 20 microphones.

The Figures Figure 6 and Figure 7 shows the beampattern for the delay-and-sum and for the Spherical Harmonic Beamforming, respectively.
For the measurement, ¼” BSWA microphones, model MPA416, were chosen. For data acquisition, three NI 4472 boards were used. The boards were synchronized by the platform NI PXI 1041Q. Also from NI, the software LabVIEW were used for control the acquisition and for process the incoming data.

The measurements were made inside an isolated wooden box with dimensions 960 x 750 x 580 mm, approximately. All the experimental set can be seen in Figure 8.
5. Results

Five measurements were made inside the isolated wooden box, one in each third-octave center frequency between 500 and 1250 Hz. The sound source was far enough from the array for adoption of the plane wave model.

It was applied the delay-and-sum and the SHB in each tone, both techniques in the same signal.
previous recorded. The Figures Figure 9, Figure 10, Figure 11, Figure 12 and Figure 13 show the comparison for the 500, 630, 800, 1000 and 1250 Hz tone, respectively, with a flattened map and a spherical representation.

Figure 9 – Comparison between the techniques for a 500 Hz tone

Figure 10 - Comparison between the techniques for a 630 Hz tone
Figure 11 - Comparison between the techniques for an 800 Hz tone

Figure 12 - Comparison between the techniques for a 1000 Hz tone
As expected, the SHB for 500 Hz and 1250 Hz results many side lobes that contaminate the acoustic maps. For this reason, they cannot be used for sound source identification. The others maps show the efficiency of the SHB, it can be seen that the lobe width remains the same for all the frequencies, since that parameters only depends of the order N.

On the other hand, the delay-and-sum have the lobe width inversely proportional to the frequency and only for frequencies higher than 1000 Hz the width gets narrower than SHB. The delay-and-sum could be extend up to 2000 Hz, as mentioned before.

In this case, the delay-and-sum could be applied for the rigid sphere because, for this frequency range, the mode strength for both types of sphere are very similar in magnitude and behavior.

In this paper, it was not made a calibration for the lobe amplitudes since the objective is the sound source location. But it can be seen that the variation of lobes amplitudes between the techniques are nearly constant, about 4dB in the frequencies where the SHB can be employed.

6. Conclusions

This paper has shown a comparison between the delay-and-sum beamforming and the Spherical Harmonic Beamforming. The results showed the superiority of the SHB for low frequencies (up to 1000 Hz). Although this had been already proven before (2, 3, 4), these results are remarkable because the spherical array used had only 20 microphones that allowed the spherical harmonic expansion only for order 2, which is way below that the orders used in the early publications (3, 4).

These results indicate that is possible combined both the techniques explored in this paper for expand the frequency range of the spherical array based in the dodecahedron. The SHB could be used for low and medium frequencies and the delay-and-sum for higher frequencies.

REFERENCES


2. Rafaely, B., “Plane Wave Decomposition of the Sound Field on a Sphere
by Spherical Convolution”. Isvr technical memorandum no 910, University of Southampton, 2003


