PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Identification of the Room Characteristics Using a Spherical Microphone Array

Frederico ARAUJO[1]; Julio Cesar TORRES[2]; Fernando Augusto PINTO[3]

[1,2,3] COPPE/UFRJ, Brazil

ABSTRACT

This work proposes the use of a Spherical Microphone Array (SMA) aiming not only to localize sound sources, but also the characteristics of the environment, such as dimensions and baffles positions. This is accomplished by a spherical microphone array based on dodecahedron geometry. The processing consists in the spherical harmonics decomposition of the sound field created around the spherical surface by a sound source inside an enclosure environment. Only the sound source signal must be known, while its location and the reflections from the surrounding surfaces can be blindly detected by using the Wavelet Transform and Beamforming. The geometry of main room surfaces is obtained by identifying the time of arrival and direction of the direct sound and first reflections. Results obtained from ray-tracing simulation and *in situ* measurements are presented and compared, showing the applicability of the proposed method.

Keywords: Spherical Harmonic Beamforming, Room Identification, Wavelet Transform, Ray-tracing Simulation

1. INTRODUCTION

Microphone arrays have been used in the last decades to identify sound sources in several environments. One of the most used techniques is the beamforming, that can be understood as a spatial filter for selecting sound waves coming from the direction of interest. Nowadays, due to the acquisition hardware and the processing software capacity, it is already to possible apply the beamforming using higher number of microphones in the array. This condition allows, for example, the use of a Spherical Microphone Array (SMA) for monitoring all directions in space with high resolution. This is possible because SMA allows to apply the beamforming algorithm in the spherical harmonics domain.

In this work, it is proposed to extend the application of the Spherical Harmonic Beamforming (SHB) to identify blindly the room geometry where the system is located. This might be an interesting application to robots and drones, which need to recognize their surrounding environments. Using as input signals the impulse responses measured by each microphone in the array related with a single sound source,

The proposed method allows identifying the time and the directions of arrival of the direct sound and the first order reflections of the environment. The method is based on the analysis of the impulse responses for all array microphones related to a given sound source. Only the emitted signal from the sound source must be known, in order to calculate the impulse responses. The other environment such as main dimensions or baffles positions, are obtained from the array signal processing.

After a small revision of SHB techniques, it is presented the results using a SMA with 20 equally spaced microphones in a real classroom with measured signals and signals generated from a ray-tracing simulation.

2. SPHERICAL HARMONIC BEAMFORMING

Beamforming is a spatial filter technique that uses microphone arrays and signal processing to control computationally the directivity and the sensitivity of the array with respect to the incoming

[1] fh.araujo@mecanica.ufrj.coppe.br
[2] julio@poli.ufrj.br
[3] fcpinto@ufrj.br

wave. Each sensor in the microphone array is in a different spatial location, thus providing different time signals according to the relative position between sensor and sound source.

When SMA are used, the SHB appears as an alternative for the conventional algorithm, also known as Delay-and-Sum algorithm. The SHB is based on the spherical harmonic decomposition of the sound pressure along a spherical surface (1), taking into account the spatial sampling over the sphere, as shown in (2, 3). The utilization of spherical arrays in acoustics have been employed after (4, 5).

In the Delay-and-Sum method, the array response for a given direction is obtained by applying weights and delays to each sensor signal. The summation of these signals emphasizes the sound coming from that direction. If there is a single arriving sound wave from the focused direction (unique sound source), all signals should be in similar phase, with higher amplitude than those coming from other directions. In the SHB, considering that the sound pressure is known all over the spherical surface Ω, the output B is given by the weighted integral of the sound pressure P in this surface:

$$B(ka,\theta_s,\varphi_s)=\int_{\Omega} w(ka,\theta_s,\varphi_s,\theta,\varphi)P(ka,\theta,\varphi)d\Omega \tag{1}$$

where (θ,φ) are, respectively, the elevation and azimuth angle, k is the wave number, a is the radius of the SMA and the index s represents the direction focused by the algorithm. The weight w will depend on the beampattern intended. The advantage of the spherical arrays is the possibility of having the same beampattern for all directions in space. The ideal beampattern is a Dirac delta, where signals coming from the direction of interest are emphasized and from all other attenuated.

Therefore, the weighting for the desired beampattern, according to (6), is given by:

$$w(ka,\theta_s,\varphi_s,\theta,\varphi)=\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{Y_n^m(\theta_s,\varphi_s)}{4\pi i^n b_n(ka)}Y_n^m(\theta,\varphi)^* \tag{2}$$

where b_n is the mode strength, $i=\sqrt{-1}$ and the symbol * denotes the complex conjugated. Y_n^m is the n^{th} degree and m^{th} order spherical harmonic defined as

$$Y_n^m(\theta,\varphi)=\sqrt{\frac{2n+1}{4\pi}\frac{(n-m)!}{(n+m)!}}L_n^{|m|}(\cos\theta)e^{im\varphi} \tag{3}$$

where $L_n^{|m|}$ is the associated Legendre polynomial.
and for rigid sphere (5):

$$b_n(ka)=j_n(ka)-\frac{j'_n(ka)}{h'_n(ka)}h_n(ka) \tag{4}$$

where j_n and h_n are, respectively, Bessel and Hankel sphere functions and j'_n and h'_n their derivatives.

In practice, the pressure is not known along the entire sphere surface, only at discrete points where the microphones are located. For this reason, it is not possible to expand the ideal beampattern to infinity orders and the integration over the sphere surface becomes a sum:

$$B(ka,\theta_s,\varphi_s)=\sum_{j=1}^{M}\alpha_j w(ka,\theta_s,\varphi_s,\theta_j,\varphi_j)P(ka,\theta_j,\varphi_j) \tag{5}$$

where M is the number of microphones in the array and α_j are weights that depends on the microphones distribution around the sphere. This work uses an equally spaced microphone array, for which these weights are constant.

For this reason, the weight w, expressed in Eq. (2), will be truncated in order N. Generally, to ensure a truncation of the spherical harmonic decomposition to order N, M microphones are necessary (7), where:

$$M \geq (N+1)^2 \tag{6}$$

For greater N, the beampattern tends to the Dirac delta. Figure 1 shows beampattern truncated as function of N. The array spatial resolution is given by the angle Ψ_0 where the first zero in the beampattern occurs. Using the Legendre functions, the resolution can be approximate by:

$$2\Psi_0 \approx \frac{2\pi}{N} \quad (7)$$

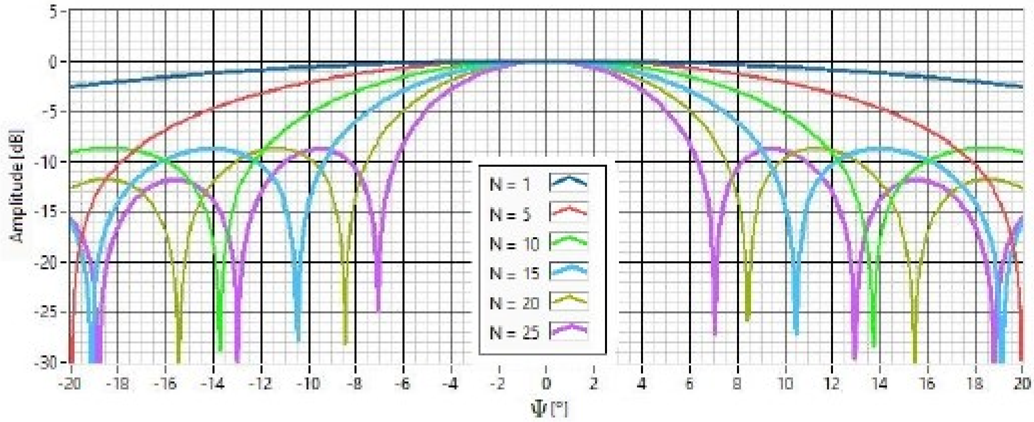

Figure 1 – SHB beampattern truncated to order N

For the truncated expansion of the spherical harmonics decomposition, the following restrictions must be satisfied (7):

$$b_n(ka) \neq 0 \mid n = 0...N, ka \in [ka_{\min}, ka_{\max}] \quad (8)$$

and

$$b_n(ka) = 0 \mid n > N, ka \in [ka_{\min}, ka_{\max}] \quad (9)$$

where $[ka_{\min}, ka_{\max}]$ is related to the lower and upper frequency of the array. In practice, upper limit frequency is set for the argument ka when $b_{N+1}(ka)$ is lower than -10dB.

More details about this procedure can be found in

2.1 Spherical Microphone Array

A spherical array with 100 mm radius and 20 microphones was printed in PLA (Polylactic Acid) for this application (Figure 2), where the microphones are located at vertices of a dodecahedron.

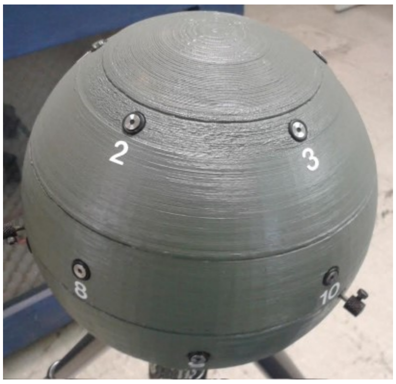

Figure 2 - SMA with 20 microphones

This array allows a spherical harmonic decomposition up to 3rd order (N = 3), due its geometry and the microphone distribution. The spatial resolution is about 120°. This means that the algorithm is not capable to distinguish two simultaneous incoming waves whose relative angle is smaller than this resolution angle. The theoretical frequency range for the SHB is between 550 and 1000Hz, according to Eqs. (9) and (10).

3. Method Description

3.1 Spherical Harmonic Beamforming Combined with Wavelet Transform

The Spherical Harmonic Beamforming is calculated using indirectly measured impulse responses of an environment for each microphone, by exciting a loudspeaker with a sine sweep with the same frequency band of the SHB with duration of 10s. A time window was applied in the impulse response signals in order to select the first 100 ms. This selection intends to reduce the processing time, focusing only in the early reflections.

After obtaining a signal relative to each direction, these outputs can be analyzed using the Wavelet Transforms (WT). Multi-resolution discretized maps (time x frequency) are then calculated, allowing the identification of the time and direction of arrival of the direct sound and the reflection in the room. The WT allows a time resolution of 0.2 ms.

The Morlet wavelet *g(t)* family was chosen for the decomposition, due to its exponential decay and frequency control:

$$g(t) = e^{-\frac{t^2}{2}} e^{-i\omega_0 t} \tag{10}$$

This wavelet family has the advantage that the dilatation and translation parameters, a and b respectively, can be easily related to time and frequency. The translated and dilated function are given by (8):

$$g_{a,b}(t) = g\left(\frac{t-b}{a}\right) \tag{11}$$

3.2 Environment Surface Identification

Using the arrival time and directional information from the direct sound and the first order reflections, it is possible to identify the surface responsible for each reflection. Figure 3 shows a geometry sketch of an environment with a surface. Point S represents the sound source location, S' the source image related to the surface and O is the observer (receiver) location, in this case, the microphone array. Using trigonometry the angle β can be determined by knowing the direction and propagation time of the direct sound (t_0) and of a given reflection (t_R). With the time of arrival it is possible to estimate the distances $c \cdot t_0$ and $c \cdot t_R$, where c is the sound speed. The reflection point location (R), as well as the surface normal vector **n**, can also be determined. Therefore, by knowing the directions of the reflected waves from the spherical beamforming and the time from a temporal energy analysis of the impulse responses, each plane defined by a given reflection point (R) and a surface normal **n** can be determined. Figure 4 shows how a wave is identified, using the direct sound as example.

3.3 Ray-tracing simulation

In order to evaluate the proposed method, results from simulated and measured scenarios are compared. The software BRASS (Binaural Room Acoustic Simulator), developed at UFRJ, was selected to simulate the environment. The software implements a modified version of the raytracing method proposed by (10) with source and receivers defined after (11). The modified ray-tracing implements a clustering algorithm applied on the first reflections up to a given order, usually 4th or 5th reflection order. With such approach, the energy spread by the ray divergence is clustered back into the direction of the nearest ray to the receiver center. Therefore, the energy can be focused in a specific time and direction, similar to the image-source method. This software was used in the first moment to create a reference list of reflections including the time and direction of arrival when simulating a single point located in the center of the SMA. Then, it was simulated 20 impulse responses, one for each measurement point the SMA. These IR's were used as input in the SHB and the results were compared with the measured ones.

Figure 5 shows the classroom in UFRJ where the measurements were carried out with a top view and Figure 6 presents the room model used in the simulation.

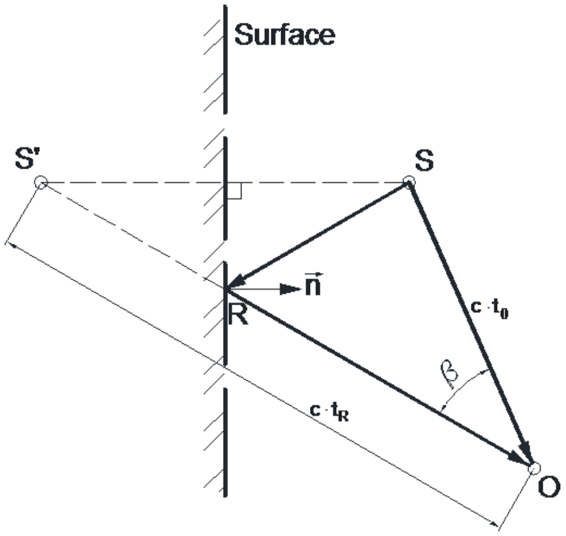

Figure 3 - Geometry sketch of a surface reflection

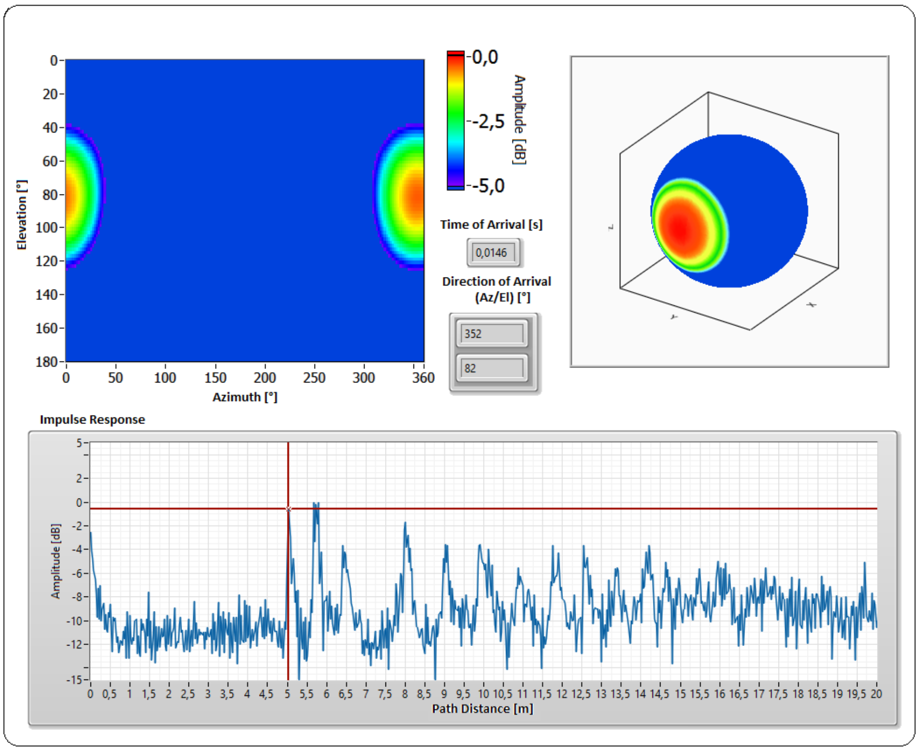

Figure 4 - Identification of the direct sound

4. Results

Table 1 shows the time and the direction of arrival of the direct sound and all the first order reflections for three situations: (a) Ray-Tracing Simulation (RTS) for a single receiver, at the spherical array center; (b) the SHB output from measurements and (c) the SHB output from ray-tracing simulation of 20 receivers at the microphone locations. The beamforming is calculated for an equiangular space for elevation and azimuth, both spacing was set to 2°. The Wavelet translation factor b was set to 0.2 ms, defining the time resolution.

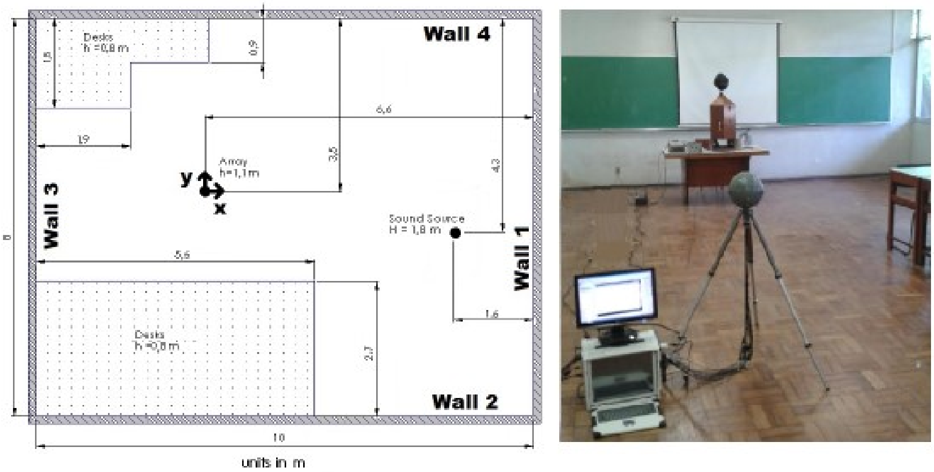

Figure 5 - Classroom: Top View (left) and Measurement setup (right).

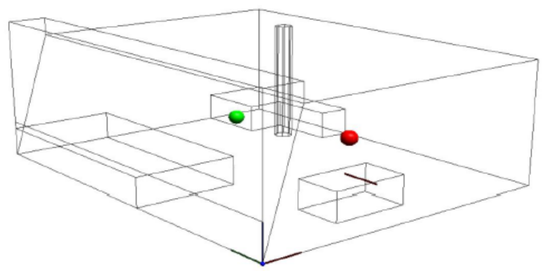

Figure 6 – Simplified room model for the ray-tracing simulation

Table 1 - Incoming reflections

	RTS for single receiver			SHB from measurements			SHB from simulation		
Reflection Type	Time [ms]	Az [°]	El [°]	Time [ms]	Az [°]	El [°]	Time [ms]	Az [°]	El [°]
Direct Sound	14.6	351	82	14.6	352	82	14.6	352	84
Floor	16.7	351	119	16.8	352	128	16.8	352	114
Ceiling	18.7	351	51	18.8	352	48	18.8	352	56
Wall 1	23.8	354	85	23.0	356	88	24.1	356	86
Wall 4	26.7	57	86	26.4	62	80	27.0	58	84
Wall 2	27.7	301	86	27.6	308	100	27.6	300	88
Wall 3	34.1	184	86	34.8	184	76	34.2	186	88

Figure 7 shows the room reconstruction from the ray-tracing for a single point and the SHB results, while Figure 8 and Figure 9 show, respectively, the reconstruction for the measured SHB and the simulated SHB. As expected, the reconstruction from the single-receiver was very close to a parallelepiped, the real shape from the classroom, while the ones for the SHB have deformed shapes. However, the calculated dimension, considering the center of the surfaces as reference, were very close to the real ones, as can be seen in

Table *2*.

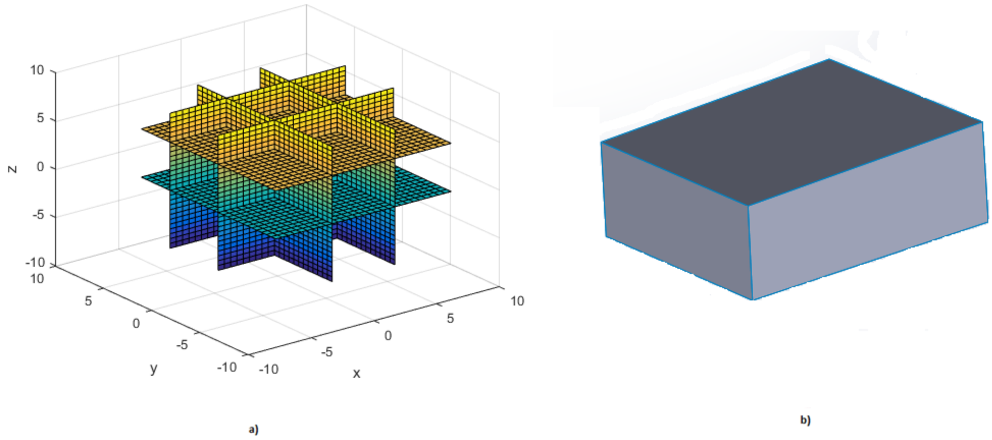

Figure 7 - Room reconstruction from the single-point ray-tracing simulation: a) plane reconstruction; b) solid reconstruction

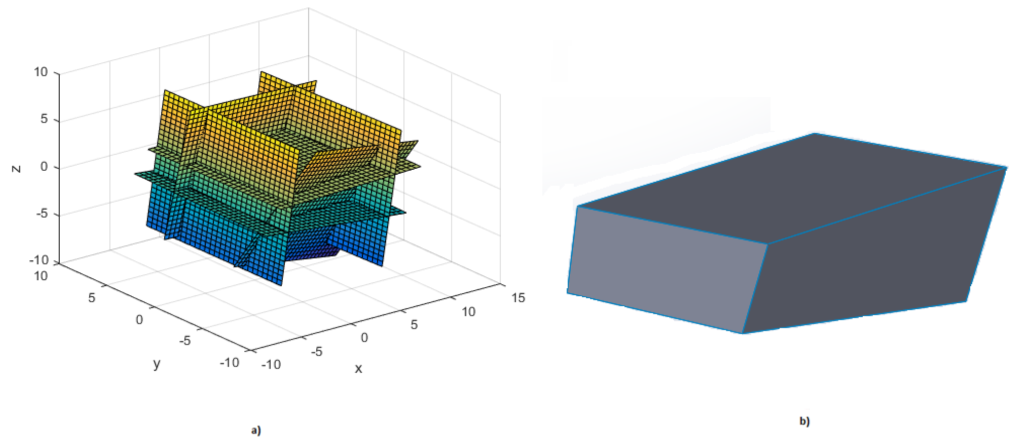

Figure 8 - Room reconstruction from the measured SHB: a) plane reconstruction; b) solid reconstruction

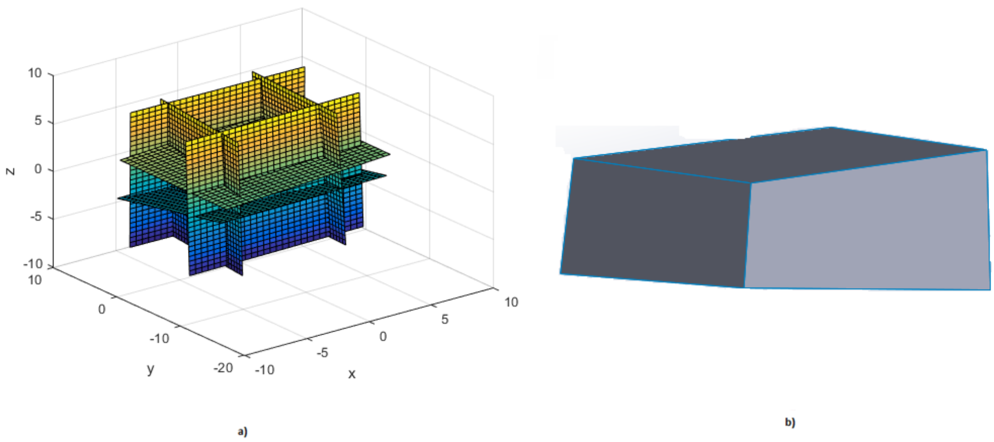

Figure 9 - Room reconstruction from the simulated SHB: a) plane reconstruction; b) solid reconstruction

Table 2 - Calculated room dimensions

Dimension [m]	X	Y	Z
Room	10.00	8.00	3.50
Single-point RTS	9.79	7.85	3.45
Measured SHB	9.79	8.25	3.54
Simulated SHB	8.90	7.81	3.68

5. CONCLUSIONS

This research presented a methodology to identify time and direction of arrival of reflections using a Spherical Microphone Array with a technique based on the combination of the Spherical Harmonic Beamforming with the Wavelet Transform. It was applied in a real classroom, where were possible to identify all the first order reflections, that were compared with a ray-tracing simulation method also presented here. With this information, it was possible to measure the dimensions of the room. The results were very close to the real dimensions, with an error smaller than 3% for the measured data and 11% for the simulated data.

The next steps, it is planned to apply this procedure in rooms with more complex geometry and with more surfaces to be identified. Also it is intending to include in this procedure an amplitude analysis. Thus it will be possible to identify some specific characteristics from the surfaces, for instance, absorption coefficients.

ACKNOWLEDGEMENTS

The authors would like to thank the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the financial support.

REFERENCES

1. B. Rafaely, Analysis and design of spherical microphone arrays, Speech and Audio Processing, IEEE Transactions on 13 (1) (2005) 135-143.
2. R. H. DuHamel, Pattern synthesis for antenna arrays on circular, elliptical, and spherical surfaces, Radio Direction Finding Section Elect. Eng. Res. Lab. Rep., Univ. of Illinois, Urbana, 1952
3. M. Hoffman, Conventions for the analysis of spherical arrays, Antennas and Propagation, IEEE Transactions on 11 (4) (1963) 390-393
4. T. D. Abhayapala, D. B. Ward, Theory and design of high order sound field microphones using spherical microphone array, (ICASSP 2002) IEEE International Conference on Signal Processing
5. J. Meyer, G. Elko, A highly scalable spherical microphone array based on an orthonormal decomposition of the sound field, Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on 2.
6. B. Rafaely, Phase-mode versus Delay-and-Sum spherical microphone array processing, Signal Processing Letters, IEEE 12 (10) (2005) 713-716.
7. T. D. Abhayapala, D. B. Ward, Theory and design of high order sound field microphones using spherical microphone array, (ICASSP 2002) IEEE International Conference on Signal Processing..
8. M. Vetterli, J. Kovacevic, Wavelets and Subband Coding, Prentice-Hall, Englewood Cli_s, New Jersey, 1995.
9. DE ARAUJO, Frederico Heloui; PINTO, Fernando Augusto de Noronha Castro; TORRES, Julio Cesar Boscher. Room reflections analysis with the use of spherical beamforming and wavelets. Applied Acoustics, 2018, 131: 192-202.
10. A. Krokstad, S. Strom, S. Sorsdal, Calculating the acoustical room response by use of a ray tracing technique, Journal of Sound and Vibration 8 (1) (1986) 118-125.
11. R. A. Tenenbaum, T. S. Camilo, J. C. B. Torres, S. N. Y. Gerges, Hybrid method for numerical simulation of room acoustics with auralization: part 1 - theoretical and numerical aspects, Journal of the Brazilian Society of Mechanical Sciences and Engineering 29 (2007) 211-221..