inter.noise
HAMBURG 2016

Comparison Between the Impulse Response Beamforming with the Ray Tracing Simulation of a Room

Frederico Heloui de ARAUJO[1]; Julio Cesar Boscher TORRES[2]; Fernando Augusto de Noronha Castro PINTO[3]

[1, 2, 3] Universidade Federal do Rio de Janeiro, Brazil

ABSTRACT

The beamforming algorithm is widely used nowadays to localize sound sources in several types of environments, however the results are expressed either in the time domain or the frequency domain. Therefore the analysis of these signals will have the intrinsic deficiencies of the time/frequency domain. For this reason, using the Wavelet Transform in the beamforming output, is possible to create acoustic maps discretized in time and frequency, then, the beamforming can be applied in non-stationary sound sources. Using a spherical array with 20 microphones, this paper applies the beamforming in the impulsive response of a room and the maps are compared with the ray tracing simulation of the same environment. The results shows the similarity between the measurement and the simulation.

Keywords: Beamforming, Ray Tracing Simulation, Sound Identification

1. INTRODUCTION

There are many techniques that use microphone arrays for sound source identification. The beamforming have been rapidly developed, driven by the signal acquisition hardware and the signal processing software evolution. However, there are situations where non-stationary sound still cannot be identified precisely. For example, inside automotive vehicles where intermittent noises are difficult to identify by human action and by conventional algorithms.

The conventional algorithm, a.k.a. delay-and-sum, is used since the World War I (1). 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).

This paper proposes a wavelet decomposition of the beamforming output signals in order to avoid such problem. The beamforming used is based on the spherical harmonics decomposition.

The proposed wavelet technique was tested using a spherical microphone array with 20 sensors to map the sound propagation, through its impulse response, in a classroom from UFRJ. In this text, it is called Impulse Response Beamforming (IRB). The beamforming reflection map is compared to a simulated impulse response, in order to evaluate the applicability of the proposed method to identify the time and direction of the room reflections.

2. Beamforming

The beamforming is a spatial filter technique that uses microphones arrays to control the sensitivity to the incidence of sound waves, for a given direction or point in space, allowing the spatial identification of the sound sources.

Each sensor in the array occupies a space location and has different time signals, according to the distance between the sensor and the sound source positions. To drive the array for certain direction in space, where the source is being searched, different weights and delays are applied to each signal, in order to amplify the sound coming from that direction. If there is an arriving sound wave from the

[1] freddearaujo@gmail.com
[2] julio@poli.ufrj.br
[3] fcpinto@ufrj.br

focused direction, all signals should be in phase, and there will have a much stronger response than when the array is pointed to another direction.

2.1 Spherical Harmonic Decomposition

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

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

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

Considering a function f(θ,φ) which is entirely defined and integrable on the surface Ω of a unit sphere, it is possible to decompose it in spherical harmonics, in the following way:

$$f(\theta,\varphi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \overline{f_{nm}}\, Y_n^m(\theta,\varphi) \tag{2}$$

The overline f is the n^{th} degree and m^{th} order coefficient and can be obtained by:

$$\overline{f_{nm}} = \int_{\Omega} f(\theta,\varphi) Y_n^m(\theta,\varphi) * d\Omega\,, \tag{3}$$

where * denotes the conjugated complex representation.

2.2 Spherical Harmonic Beamforming

The advantage of the spherical arrays is the possibility to have the same beampattern for all the directions in space. The ideal beampattern is a Dirac delta, where the direction of interest have an amplitude and all the others have none.

The generic spherical harmonic beamforming expression for a driven direction (θ_s,φ_s) is given by:

$$B(ka,\theta_s,\varphi_s) = \int_{\Omega} w(ka,\theta,\varphi,\theta_s,\varphi_s) P(\omega,\theta,\varphi) d\Omega\,, \tag{4}$$

where *w* is an weight that orientates the algorithm for the direction (θ_s,φ_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, the ideal beampattern can be expressed by the spherical harmonics:

$$\delta(\theta-\theta_s)\delta(\varphi-\varphi_s) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \overline{\delta_{nm}} Y_n^m(\theta,\varphi) \tag{5}$$

with

$$\overline{\delta_{nm}} = Y_n^m(\theta_s,\varphi_s)\,. \tag{6}$$

Therefore, the weighting is given by (3):

$$w(ka,\theta,\varphi,\theta_s,\varphi_s) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} \frac{\overline{\delta_{nm}}}{4\pi i^n b_n(ka)} Y_n^m(\theta,\varphi) \tag{7}$$

where b_n is given by:

$$b_n(ka) = j_n(ka) \tag{8}$$

for open spheres and

$$b_n(ka) = j_n(ka) - \frac{j'_n(ka)}{h'_n(ka)} h_n(ka) \tag{9}$$

for rigid spheres. Where j_n and h_n are the Bessel and Henkel spherical functions, respectively, and j'_n

and h'_n are their derivatives.

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 (5), where:

$$M \geq (N+1)^2 \tag{10}$$

For larger N, the beampattern will be closer to the Dirac delta. The Figure 1 shows beampattern truncated for some values of N.

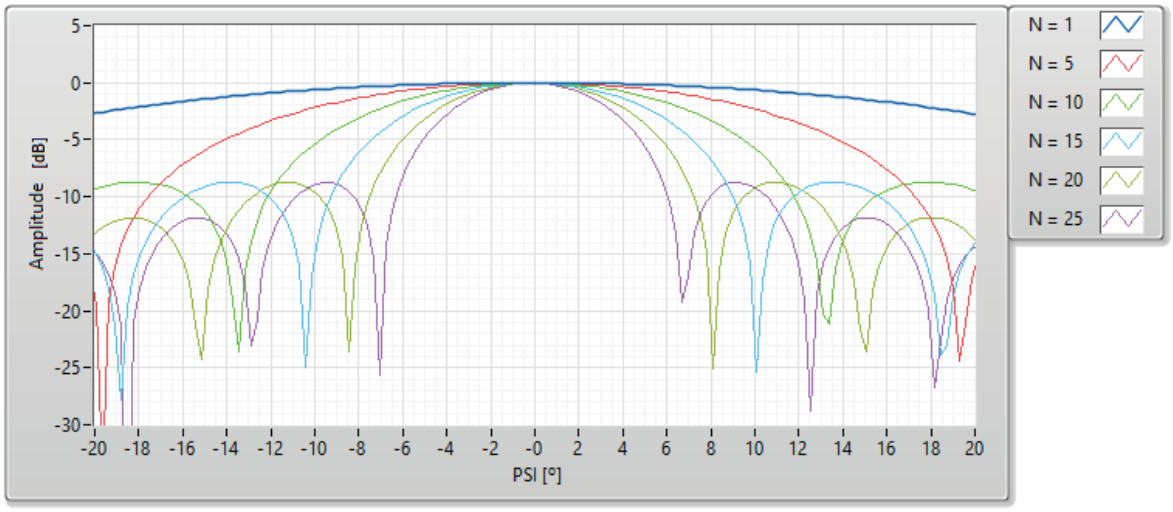

Figure 1 – Truncated beampatterns.

3. Mapping with the Wavelet Transform

In this paper, the beamforming is calculated using the impulse response of each microphone, measured indirectly by a sine sweep generation on the speaker. After applying the beamforming technique, a signal relative to each direction tuned by the algorithm is obtained. That signal can be analyzed in the time or frequency domain and subject to the intrinsic deficiencies of the time/frequency domains. In this paper, the Wavelet Transform is used in each directional signal to identify the arrival direction and time of each reflection in the room.

Figure 2 shows the procedure adopted for the creation of the impulsive response maps. First, the signal of all the microphones are acquired simultaneously with the reference signal on the speaker's input. Then, the beamforming is applied to the impulsive responses. Lastly, the Wavelet Transform is applied to the signals of every direction and works similarly to the Short Fourier Transform Analisys.

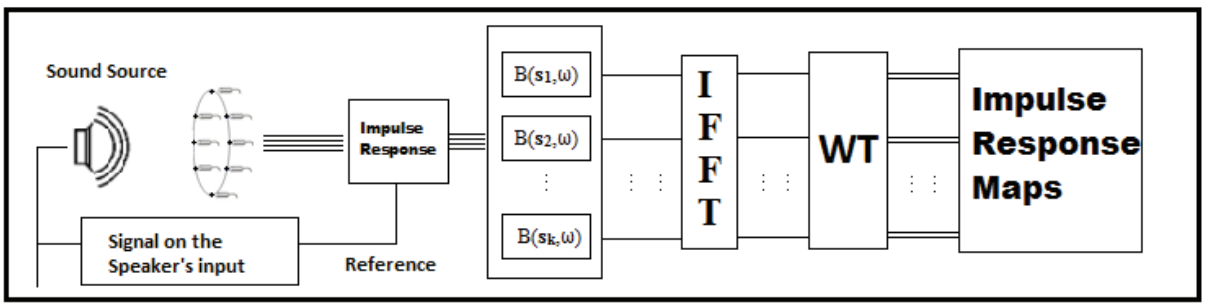

Figure 2 - Procedure adopted for the creation of the impulse response maps

The Morlet wavelet family was chosen for this situation. With this wavelet family, the translation and dilatation parameters can be easily related with time and frequency. This wavelet is given by an exponential decay of an exponential complex.

$$g(t) = e^{\frac{-t^2}{2}} e^{-i\omega_0 t} \tag{11}$$

4. The Classroom Impulse Response Beamforming

A spherical array with 20 microphones positioned on the vertices of a dodecahedron was built in PLA for this application (Figure 3). This array allows the spherical harmonics decomposition up to 2nd order. The spatial resolution is about 105°, which means that the algorithm cannot separate two waves arriving from directions that form between themselves an angle less than 105°. This resolution is only enough for mapping the early reflections. The frequency range of this microphone configuration ranges from 550 up to 1000 Hz, considering only the spherical harmonics decomposition.

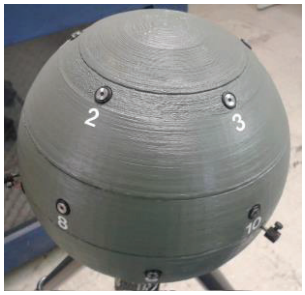

Figure 3 - Spherical array with 20 microphones.

The classroom used for the experiments is shown in Figure 4. All the students desks were joint, forming large blocks. An omnidirectional sound source with 12 speakers (dodecahedron) was used to fill the room with a sweep sine, with a background noise at least 20 dB lower than the source output.

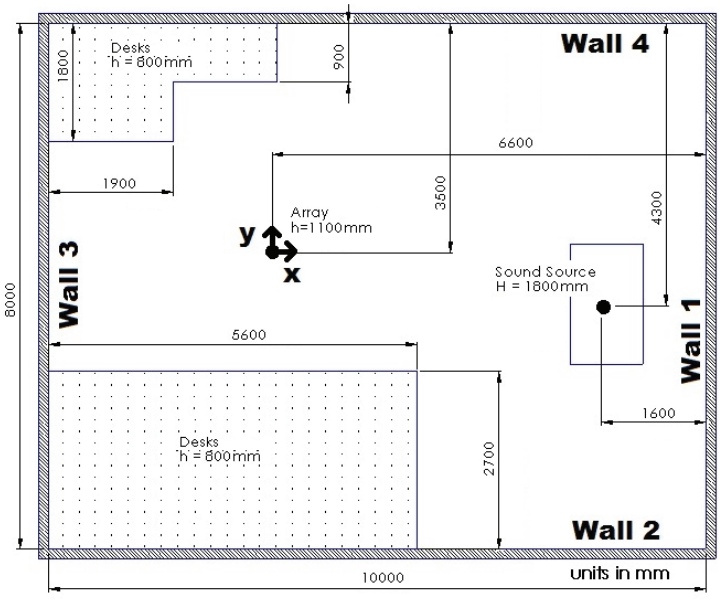

Figure 4 - Classroom: (a) Top View and (b) Measurement setup.

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 was used to control the acquisition and to process the incoming data.

5. Simulation

The simulation was performed using software BRASS – Binaural Room Acoustic Simulator, developed at UFRJ. It implements a modified version of the ray-tracing method proposed by Krokstad et al (8) with source and receivers defined after Tenenbaum et al (9). 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 strategy allowed to compare the simulated energy, in time and space, with the beamforming array measurements.

The room was modeled considering the simplified geometry, shown in Figure 5, where the floor, walls, ceiling, windows and desks were defined by planes. The area below the desks were considered was purely absorptive, while the remaining planes were defined according to Table 1. The omni-directional source and the receiver were simulated at the same measured locations, as stated in Figure 4(a).

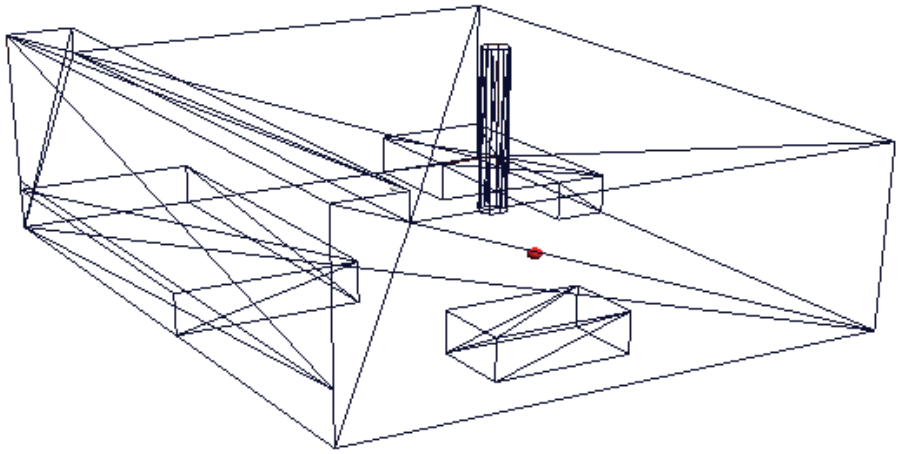

Figure 5 - Simplified room model for ray-tracing simulation.

Table 1 - Absorption coefficients used for simulation

Surface	Material	Absorption Coefficient per Frequency Band								
		62	125	250	500	1k	2k	4k	8k	16k
Floor	Parket	0.20	0.20	0.15	0.10	0.10	0.05	0.10	0.10	0.10
Ceiling and Walls	Concrete	0.07	0.07	0.08	0.09	0.09	0.10	0.10	0.10	0.10
Windows	Glass	0.08	0.08	0.04	0.03	0.03	0.02	0.02	0.02	0.02
Board and Desks	Wood	0.14	0.14	0.10	0.06	0.08	0.10	0.10	0.10	0.10

Although using average values of known materials, they are not exactly those present in the room. Therefore, it is expected to appear amplitude differences between measured and simulated peaks in the impulse response. For this reason, only the arrival time and direction of the reflections were analyzed.

6. Results

Table 2 shows the reflections up to order 1 with the simulated time and direction of arrival and the information obtained by the IRB. Comparing the values of Tab. 2, the small deviations in the arriving directions can be explained by the uncertainties in the measurements, such as the room dimensions, the array position and the desk locations used to build the classroom acoustic model. The time deviations are due the sound speed differences between simulation and measurements. Another issue is that the ray tracing simulation calculates the arrival time based on the distance from the source center to the microphone array center. On the other hand, the beamforming algorithm uses one of the microphones to calculate those times, instead of the array center, since there is no real microphone there.

Table 2 - Incoming reflections.

Order	Simulation			IRB			Figure
	Time	Azimuth	Elevation	Time	Azimuth	Elevation	
0	146 ms	351°	82°	146 ms	352°	82°	Figure 6
1	167 ms	351°	119°	168 ms	352°	128°	Figure 7
1	187 ms	351°	51°	188 ms	352°	48°	Figure 8
1	238 ms	354°	85°	230 ms	356°	88°	Figure 9
1	267 ms	57°	86°	264 ms	62°	80°	Figure 10
1	277 ms	301°	86°	277 ms	308°	100°	Figure 11
1	341 ms	184°	86°	348 ms	184°	76°	Figure 12

The beamforming mapping for the direct sound and the reflections of first order are presented from Figures 6 to 12.

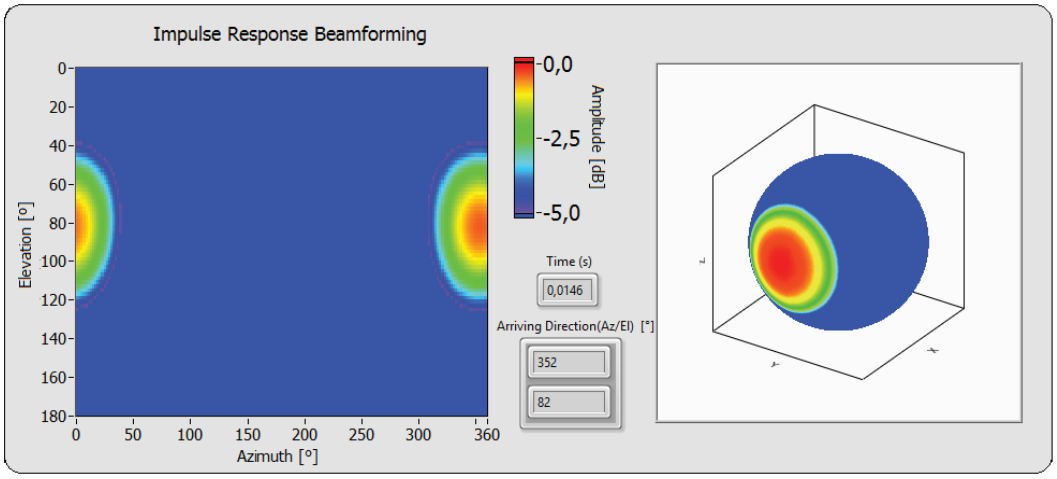

Figure 6 - Direct sound.

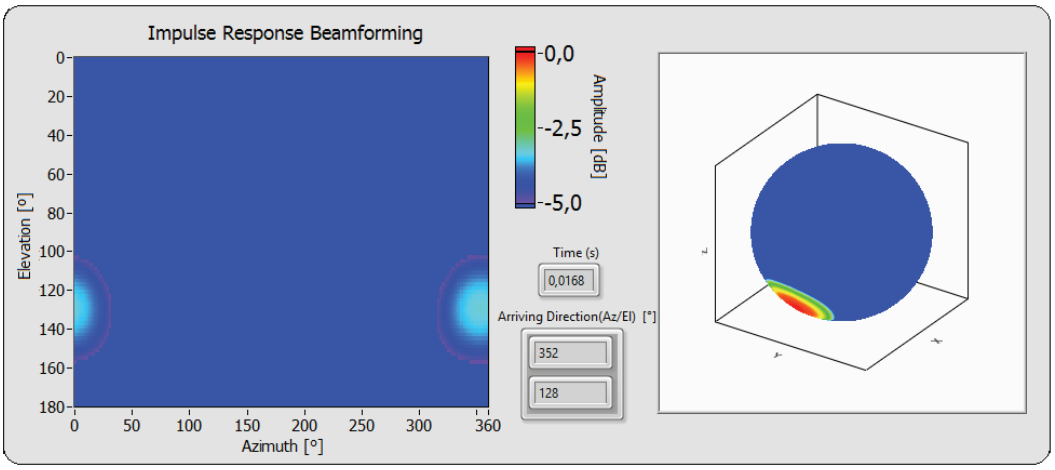

Figure 7 - Floor reflection.

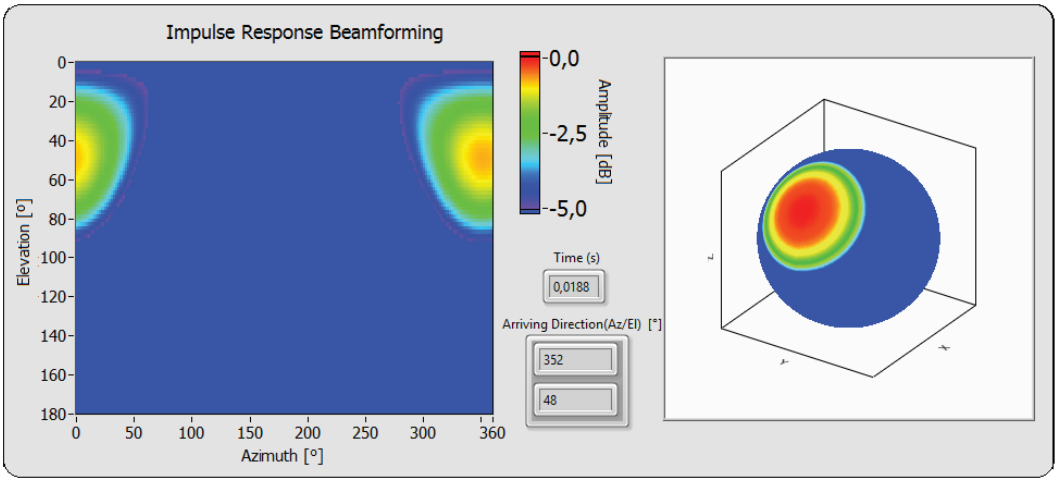

Figure 8 - Ceiling reflection.

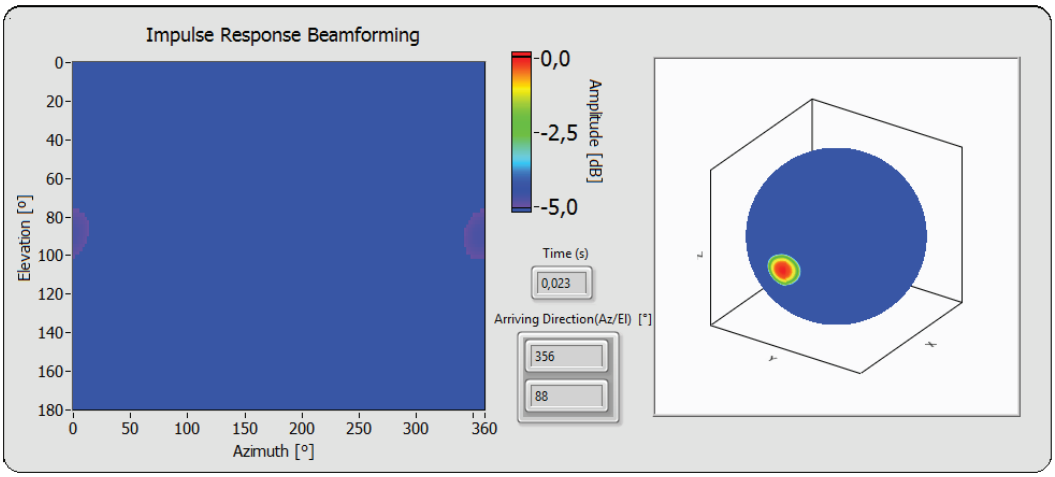

Figure 9 - Wall 1 reflection.

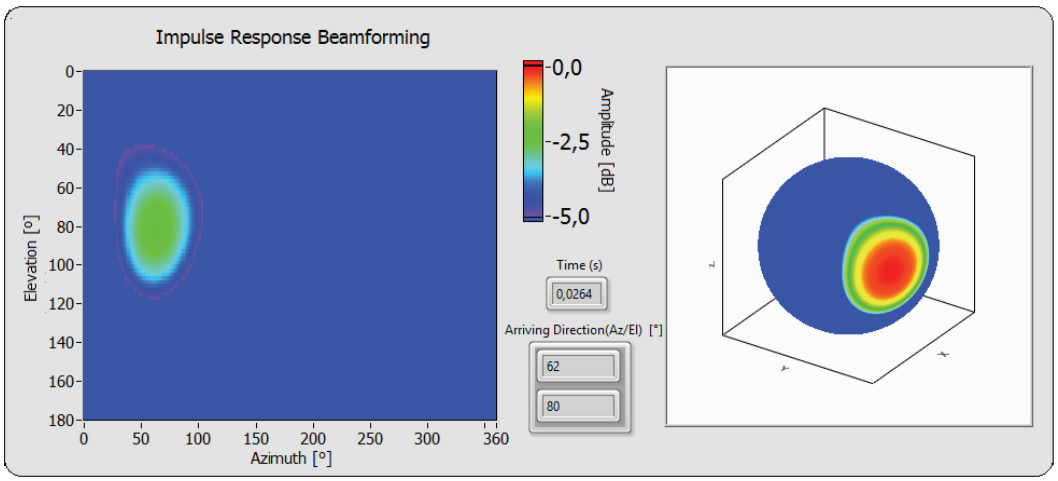

Figure 10 - Wall 4 reflection.

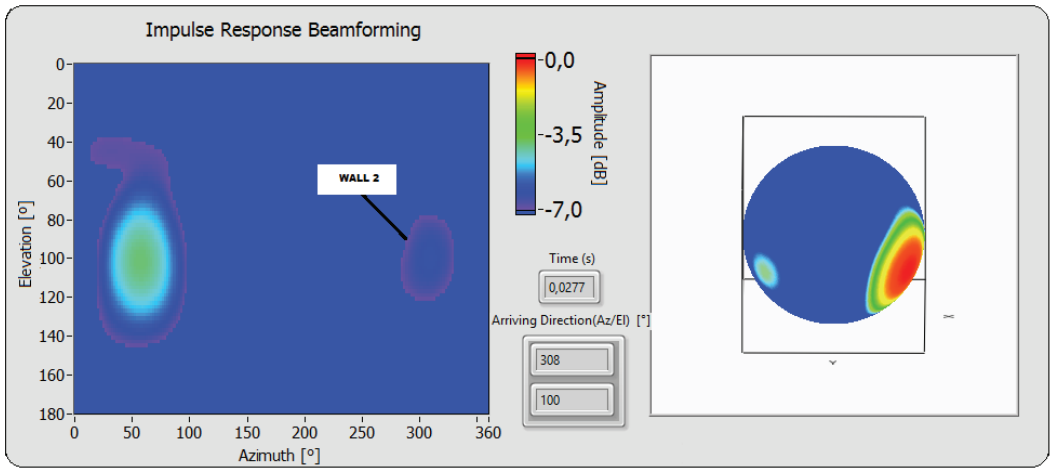

Figure 11 - Wall 2 reflection.

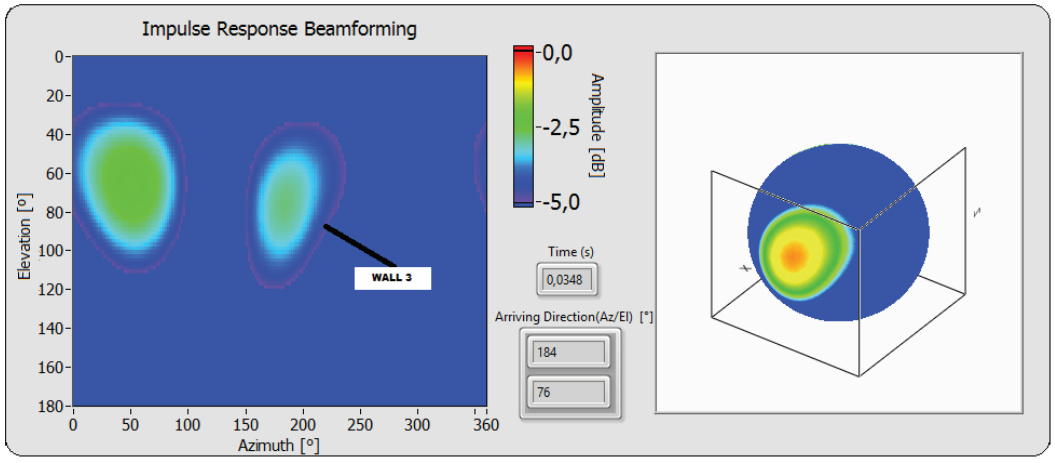

Figure 12 - Wall 3 reflection.

It can be seen in Figure 11 and Figure 12 a “time leakage” of some reflections. It is due to the time resolutions of the wavelets used. Hence, the some reflections are still visible, with smaller amplitudes, in time instances after the actual reflection arrival time.

Figure 13 shows a historic from the early 20 ms after the direct sound. Left plot shows all identified reflections during these 20 ms while at the right plot, the ray-tracing results are presented as function of the reflections order. For the IRB map, it is not possible identify all the incoming waves due the low spatial resolution of the microphone array. Therefore, some lobes were masked and only a few were

identified. Nevertheless, the similarity between the maps is noticeable over all the regions that concentrate the incoming waves.

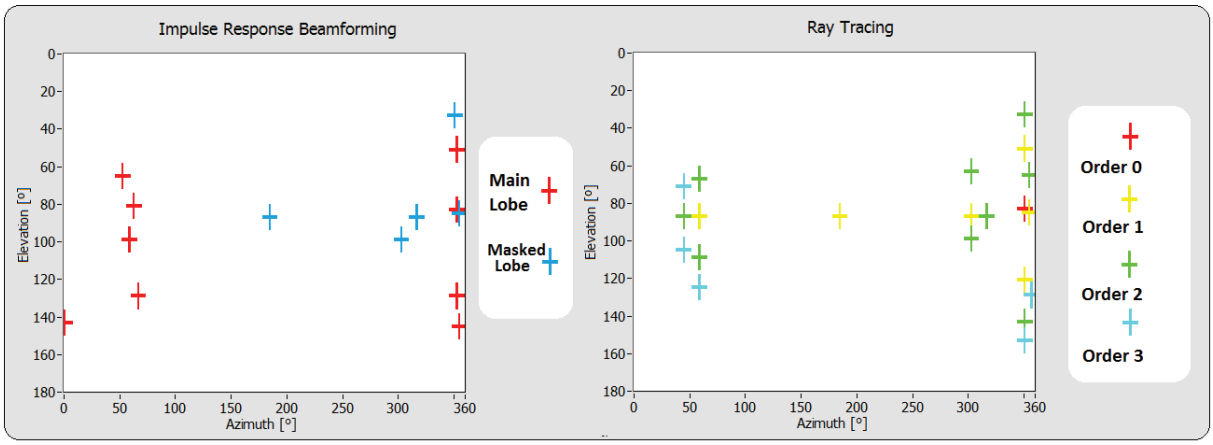

Figure 13 – Historic Maps

7. Conclusions

Despite of the low spatial resolution of the array, due the order limit N=2 of the spherical harmonic decomposition, the IRB allowed the identification of the early reflections of the impulse response. Even with a simple model of the classroom, the ray tracing simulation was extremely compatible with the beamforming for reflections up to order 1 and had coherence up to order 3.

Although this technique is still a conceptual idea, it could allow the identification of higher reflections order if employed more microphones in the array, i.e, having a higher truncated order N. This way, the wavelet decomposition can be used, not only in the IRB, but also for any non-stationary source identification, such as intermittent or moving sound sources.

Other issue to be settled is the "time leakage" (Figure 14), where the main lobe of a real reflection "leak" for to the instants of time near the moment of arrival with a lower amplitude. That could mask other incoming waves. As previously said, that is caused by the time resolution of the wavelet. The alternative would be use other wavelet family, but the correlation between dilatation and frequency would be compromised.

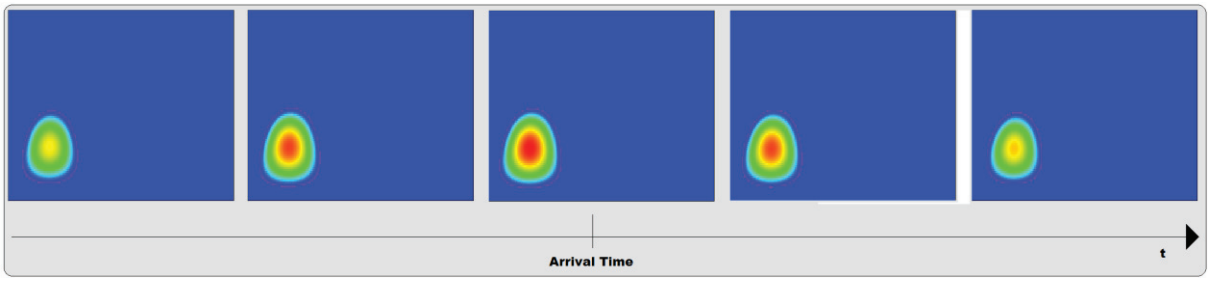

Figure 14 - Time leakage

REFERENCES

1. Johnson, D. H., Dudgeon D. E., "Array Signal Processing", Prentice Hall, Englewood Cliffs, NJ, USA, 1993

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

3. Meyer, J., Elko, G., "A highly scalable spherical microphone array based on an orthonormal decomposition of the sound field". In: Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on, v. 2, pp. II-1781. IEEE., 2002

4. Abhayapala, T. D., Ward, D. B., "Theory and design of high order sound field microphones using

spherical microphone array". In: Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on, v. 2, pp. II-1949. IEEE, 2002

5. Duhamel, R., “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

6. Hoffman, M., “Conventions for the analysis of spherical arrays", Antennas and Propagation, IEEE Transactions on, v. 11, n. 4, pp. 390-393, 1993

7. Chan, A. K., Ishimaru, A., “Equally Spaced Spherical Array. Technical report, DTIC Document, 1966

8. Krokstad A., Sorsdal S., "Calculating the acoustical room response by use of a ray tracing technique", Journal of Sound Vibration, Vol. 8, Pag. 118-125, 1986

9. Tenenbaum, R. A., Camilo, T. S., Torres J. C. B., and Gerges S. N. Y.” 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, 29:211–221, 2007.