Development of a 3D impulse response interpretation algorithm

Federico Nahuel Cacavelos\textsuperscript{(a)}, Augusto Bonelli Toro\textsuperscript{(b)}, Alejandro Bidondo\textsuperscript{(c)}

\textsuperscript{(a)}Universidad Nacional de Tres de Febrero, Argentina, fnahuelc@gmail.com
\textsuperscript{(b)}Universidad Nacional de Tres de Febrero, Argentina, abonellitoro@gmail.com
\textsuperscript{(c)}Universidad Nacional de Tres de Febrero, Argentina, abidondo@gmail.com

Abstract

In this paper, the development of software 3D interpretation impulse responses is presented. This new tool allows to identify the spatial provenance for each time window of an impulse response recorded in a soundfield microphone. In this way it is possible to discern between the direct sound and the different reflections both temporarily and spatially. To understand the basics of this method, a brief introduction to all concepts involved is presented. Alternatively, different types of procedures are shown for data evaluation. Finally system applications are shown in controlled acoustical environments so evaluating the precision and the potential of this tool.

Keywords: Impulse response 3D, Hedgehog, Spacialization, Soundfield, Ambisonics
Development of a 3D impulse response interpretation algorithm

1 Introduction

Recording of impulse responses has been widely studied in various areas of acoustics. The most notable use is given in the commercial production industry of audio, because the convolution of this type of impulse responses with anechoic signals allows to generate surround sound environments. In this work, alternative use for such responses is presented, with applications in the characterization and design parameters of acoustics enclosures. Methodology and development of this algorithm is presented. To achieve this task, the MATLAB platform was used as a programming environment. This platform allows visualization of current vectors in three dimensions. Thus, one can distinguish spatial and temporal form between the direct sound and its reflections from an impulse response recorded with an A-Format microphone. Various techniques of digital signal processing were addressed for data evaluation. Finally, the algorithm was tested with a series of measurements.

1.1 Applications of the 3D impulse response

The acoustic environments are commonly evaluated using omnidirectional impulse response measurement techniques. This techniques allow to obtain the time and amplitude of the signal, but they can’t bring information about the direction of this signal. In this sense, the advantage in the comprehension and characterization of the acoustic parameters with the 3D impulse response measurement is to get the direction which the signal comes from. Obtaining the signal direction allows to relate the sound to the physical characteristics of the room reflections, observe the directional distribution of early and late energy, and identify areas that generate unwanted reflections. Moreover, the comparison of the spatial components of impulse responses of various rooms can contribute to identifying differences and similarities in acoustic characteristics.

Although at present exists acoustic parameters describing spatial characteristics of the impulse response, it is possible to develop new parameters to consider the direction as an independent variable.

1.2 A – Format and B - Format

The A – Format consist in a recording of four audio signals, this requires using a microphone array which has four sub-cardioid capsules located on the faces of a tetrahedron.

From a theoretical aspect, capsules should coincide in one point in the space, but due to physical limitations this is impossible. This condition causes phase distortion, occurring in shorter wavelengths four times smaller than the distance between the capsules. To solve this
problem, filtering should be implemented in order to place the capsules virtually in the centre of the tetrahedron.

The B–Format contains four signals named W, X, Y and Z. The W signal represents the capture of the omnidirectional pressure component, while X, Y and Z represent the velocity vectors associated with the horizontal axis (X and Y) and the vertical axis (Z). The microphone B–format involves three figure of eight polar pattern microphone and one omnidirectional microphone. Another way to get this format is performing a combination between the A–format signals and applying an amplitude and phase compensation filter [1, 2].

1.3 Intensity vectors from Format – B signals

As Farina and Tronchin establish in their research at the University of Parma (2013), vectors intensity (corresponding to one or more sources) can be obtained by the product of the pressure (i.e. the W signal from the omnidirectional microphone) and velocity vector (X, Y or Z). Subsequently, the three cartesian components of the intensity values in each of its axes can calculate the spherical coordinates (modulus, azimuth and elevation). Trigonometric equations are used in the aforementioned work [3].

1.4 Log Sine Sweep y Kirkeby Algorithm

The measurement technique with Log Sine Sweep (LSS) is based on using as excitation signal a sine signal which varies exponentially between two frequencies. This technique takes advantage respect to a linear sweep because it has logarithmic spacing, giving more definition to low frequency tones. This way you can achieve a better Signal to Noise ratio even at low frequencies. In contrast to a linear to a linear sweep, the latter maintains the same duration for all the frequency range [4, 5], thus being necessary to apply greater sweep duration times. The application of the inverse filter from the source also produce an improvement in the quality of the impulse response measurement. Using this tool helps to reduce potential alterations that can produce the nonlinealirity of the source. For practical purposes, it is possible to achieve a good evaluation of the non-ideal acoustic conditions. The following section will develop some plausible tools. Complementary, Kirkeby algorithm is used for applying the inverse filter to the measured signals. Applying these filters improves the frequency content, eliminating the time dispersion caused by the nonlinealirity of the source [5].

2 Procedure

2.1 Acquisition of impulse responses

The technique developed by Farina with LSS was used to obtain the impulse responses. The Kirkeby algorithm in conjunction with the inverse filter allow to reduce alterations in the signal produced by the characteristics of the source.

Because of the difficulty in obtaining the impulse response of the omnidirectional source in anechoic conditions, a number of alternative methods were implemented.
Omnidirectional microphone placement was made away from a reflective surface in order to obtain the direct sound from the source with enough spacing respect the early reflections. The criterion to choose this distance was exceeding at least one wavelength of the lowest frequency being analyzed. In cases where the minimum distance was not achieved, another technique used was by placing the reflective surface and attached to the microphone. This technique modifies the microphone polar pattern to cardioid and precludes the acquisition of reflections from the surface.

It must be clarified that the application of the inverse filter response of the source does not make modification to the differences in response between the capsules. That's why in future measurement this must be taken into account in order to make a characterization of the capsules used.

2.2 Gain correction algorithm
One of the most important aspects to consider is related to the configuration of the interface gain in each channel. Inequity gain relationships between the different channels can cause erroneous results. That's the reason because a gain correction algorithm was developed. In this algorithm the user can start the measurement with any kind of gain relation between channels, preferably achieving the highest signal to noise ratio without distortion. Then, a white noise signal with the same length is inserted (can be generated with the same interface) in each of the channels used. The relative differences between energy amounts of the acquired signals represent the relationships between channels gain. Thus, amplitude corrections must be applied to obtain the same energetic sum in each channel.

2.3 Practical limitations and filters
As mentioned above, there is a restriction in frequency due to the distance between capsules. To compensate the intrinsic phase distortion at high frequency, is necessary to use a low pass filter with a cutoff frequency according to the format – A microphone used. A FIR linear phase filter with order 400 and a cutoff frequency of 5 kHz was used as seen in Figure 1.

![Figure 1: Linear phase FIR filter to avoid phase distortion.](image)

The need to implement a linear phase filter lies in the possible alteration in the interpretation of results when a phase shift that comes from the filter adds to the original signal if another type of filter is used.

2.4 Graphics presentation
The graphics are represented as vectors with the same origin and different direction. Each of the vectors represents each of the time windows. Its magnitude is given by the energy sum of
the intensity vectors. The magnitudes were normalized based on the direct sound. The color symbolizes the temporal distance to the direct sound (red to blue). In order to facilitate the understanding of the results, the graphs are presented in the three orthogonal planes including the origin point.

2.5 Temporary length of the plotted vectors

The most discussed issue in the development of the calculation procedure was the choice of the time windows representing each one of the plotted vectors. The minimum variation in information contained in each window influence the direction and magnitude of the vectors. That is why different selection methods were developed in order to compare and contrast the results.

The problem lies in the number of samples that are acquired from a time window. This modifies the frequency resolution obtained and thus, the lower frequency of analysis. Therefore, a time window with fewer samples implies a higher cutoff (low) frequency. On the other hand, taking a big time window doesn’t resolve this problem because it can contain energy from more than one reflection, resulting in constructive and destructive interference, and therefore the analysis of the results should be taken judiciously.

Differences in sensitivity at low frequencies capsules are considered a critical factor that can introduce errors in the data analysis.

![Figure 2: Two graphs obtained for different window sizes.](image)

2.6 Criteria to select the direct sound

Different methods were studied for the extraction of direct sound impulse response and reflections [7, 8]. The technique used in this study employs peaks extraction with certain parameters selected by the user. One involves that the distance between peaks should satisfy a specific ratio. From the obtained points, the one with the maximum level is taken by the center of the direct sound. A temporal window is applied with the aim of extracting the direct sound that comes from the source. In order to avoid errors, the user is forced to choose and confirm the center of the direct sound (Figure 3).
In case of having the impulse response of the source in anechoic condition or if the direct sound and reflections are sufficiently separated temporarily (one wavelength of the lowest frequency), it is possible to perform an additional analysis. This allows to obtain the specific position of each reflection even if exists overlapping positions. The reflectogram (Figure 4) makes use of the cross-correlation function by evaluating the direct sound window with the full vector of the impulse response.

2.7 Representative windows selection Criteria

Another important factor lies in the choice of the most representative time windows to plot. Vector density resulting from the number of analysis windows is such that the proper visualization of data is hampered. For this reason it is important factor to extract the most important magnitudes of intensity level. This selection also takes into account the proximity of the windows, avoiding overlap. This criteria is currently under discussion. When considering overlapping windows, it should be clarified that each sample of the impulse response will have an associated time windows, so regardless of the window length, there will be many windows as samples the audio file. Figure 5 shows the results for the three cases described above.
3 Implementation and results

With the aim to analyze the implemented methods, several known measurements were used. An example for St. Margaret’s Church from New York made by Aglaia Foteinouy and Simon Shelleyen in 2011 with a Soundfield SPS422B microphone and a Genelec S30D source is shown. LSS technique was implemented with the following characteristics:

<table>
<thead>
<tr>
<th>Table 1: LSS specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frequency range</strong></td>
</tr>
<tr>
<td><strong>Duration</strong></td>
</tr>
<tr>
<td><strong>Sampling frequency</strong></td>
</tr>
</tbody>
</table>

In figure 7 can be seen the top view with the measurement positions. Also, the results from the implementation of the algorithm are included.
In this measurement, different points located equidistantly around the room were selected to test the algorithm.

This geometric dispersion is favourable to assess the variation in the direction of the direct sound and reflections. It is noteworthy that the results were not oriented toward the source manually but is the result of the detection system.

Looking at the results that are in positions near the corners of the room it is possible to observe how the secondary components acquire a magnitude comparable with the direct sound. This is because the proximity to the reflective surfaces. By contrast, for positions P5 and P14 the direct sound has much more energy than secondary components. This tells that the results are consistent with the acoustics of the room.

Complementarily, complete system was evaluated performing a measurement on a real and known situation. To achieve this task, a conditioned room with FONAC class 1 material with a 50 mm thickness, a Soundfield SPS200 microphone, a KRK Rockit 8 source was used. As mentioned above, it is recommended to perform a measurement of the source in anechoic conditions (or almost). The techniques described previously were used with an 8 ms window (given the geometry of the room and the minimum distance to reflective surfaces), reaching a lower frequency analysis of 121 Hz. Using the LSS (50s duration between 80 and 15 kHz with a sampling frequency of 44100) and implementing Kirkeby8 technique, impulse response measurements were achieved. Subsequently, Format A to B conversion was performed using a plugin provided by the Soundfield Company.
Fig. 8. Measurement points and results achieved in an acoustically treated room

Results show that the direction of the vector corresponding to the direct sound maintains a correspondence with the location of the source. However, it is slightly oriented to the right (where there are energy components of the reflection). This may be due to overlapping of reflections in the same time window, since the measurement was performed in a room where the time difference between the sound paths is small. It can also be noticed from the reflective and diffusing surfaces.

4 Conclusions

An algorithm for obtaining impulses responses in three dimensions was developed and implemented. The results obtained prove the correspondence with the different acoustic situations evaluated, demonstrating the validity of the theory used. Thus, it can be noted that the method involves a useful and practical tool for the analysis of any type of acoustic environment. Its scope and applications should be studied in the future in order to build the basis for the description of acoustic problems. It is noted that like the impulse responses, other parameters can also be analyzed by interpreting intensity vectors, so different applications in acoustics may be benefited with the development of this tool.

It should be noted that the study should be continued with measurements to determine the precision in magnitude and direction of the vectors obtained by relating them to the acoustic situation.
Acknowledgments

The authors would like to acknowledge the University of San Buenaventura, Colombia for offer assistance and dedication in the development of the algorithm. Also a special thanks to the Universidad Nacional de Tres de Febrero multimedia staff for providing the equipment for measurement.

References


