On the 3D Acoustic Analysis in UNESCO Sites: The Example of San Vitale, Ravenna, Italy

Vincenzo Vodola

1 Department of Architecture (DA), University of Bologna, via dell'Università 50, 47521 Cesena, Italy
*vincenzo.vodola2@unibo.it

Abstract. The study and analysis of the sound field’s 3D properties has been strongly enhanced in recent years after spatial properties of sound propagation have been acknowledged to be important during the design or correction of theatres and auditorium. Besides, a proper assessment of spatial accuracy is requested for 3D sound reproduction systems, initially designed for acoustical virtual reality and now also employed in the entertainment/cinema industry. Often only monaural or binaural measurements are performed by means of omnidirectional microphones and dummy heads, although international standards like ISO 3382/1:2009 also define some “truly spatial” parameters such as JLF and JLFC. Even though the two latter parameters are derived from measurements made with a pressure velocity (p/v) microphone, they still represent a 2-channel measurement. 3D Impulse Responses (4-channel B-format) have for many years been measured and employed for sound reproduction. Recently, higher-order 3D Impulse Responses have been measurable thanks to the availability of compact microphone arrays employing a much larger number of transducers. In this paper, a procedure for measuring and analyzing the complete spatial sound information is presented, which is aimed to create easy-to-understand images and videos showing the direction-of-arrival of the room reflections. The description of this technique is emphasized and applied in the Basilica of San Vitale, Ravenna, Italy, which represents one UNESCO site.

1. Introduction
Since the burning of the Theatre La Fenice in Venice, occurred in 1996, it became essential gathering full information about the acoustics of historical buildings [1]. This attention followed the same care given to musical instruments [2,3] that are nowadays measured also by means of new methods [4], including not linear emulation [5, 6]. The same attention to historical building is already given in several field within the energy [7,8] decarbonization [9,10], microclimate [11], renewables [13,14] at different scales [15,16], including economic evaluation [17] and noise control [18,19]. In other words: global environmental quality in buildings [20, 21].

The acoustic measurements in theatres have been analysed and described by several papers [22,23], as well as in Churches [24]. The most important aspects have been the location of the sound source and of the microphones, as well as the conditions of the room. All these aspects have been taken into account during the virtual reconstruction of theatres [25]. However, only few papers have considered the analysis and the possibility to standardize the whole procedure, which particular attention to sound perception during playback in listening rooms [26]. It should also be considered that all these aspects are quite important when the purpose of the measurements of the Impulse Responses is to obtain 3D
auralisation of the theatre and not only calculating the acoustic parameters described in the ISO 3382 standard. The full spatial acoustic performance of a theatre, i.e. both knowledge of intensity, energy and position of early reflections in the theatre, is necessary to detect and correct some limits that of course cannot be detected considering only mono or binaural IRs.

A new measurement method was described in 2005 [22] and 2013 [23]. This new method incorporates the previously measurement techniques already known in a single approach. Three microphone systems were mounted on a rotating table (they were a binaural dummy head, two cardioid microphones in ORTF configuration, and a Soundfield microphone), as reported in Figure 1. Using this approach, eight Impulse Responses were measured at each angular position. The two microphones in the ORTF configuration follow the standard, French method for recording dual-channel audio signals, using the two microphones spaced 170 mm and oriented 110 degrees each other. The Soundfield microphone, which was firstly introduced by Gerzon, consisted of capsules that allow to measure 4-channel Impulse Responses, i.e. spatial properties of the sound field. The four channels recorded by means of the Soundfield microphone, are normally known as “B-format” signals. They consist of one omnidirectional (sound pressure) channel and three with a polar pattern called “figure-of-eight”, which follow the three Cartesian axes X, Y, Z. These three channels acquire a signal which is proportional to the Cartesian components of the particle velocity vector).

By combining the three different measurement methods, it is possible to obtain a new, general method that allow to obtain all the standard multichannel playback formats (i.e., 2.0, 5.1, 7.1, 10.2, etc.).

![Figure 1. Measurement method previously proposed in 2005.](image)

2. Materials and Methods

The method presented in [22,23] was very slow, because all the rotations and the procedure for setting the microphone system was particularly delicate. Moreover, the post processing of all the data from the measurements was particularly intricate, having three microphone systems employed together.

A much more powerful and simple measurement system was recently proposed [27]. It is based on a 32 capsules spherical microphone array, The array is mounted on a small sphere (80mm diameter), which contains the preamplifiers the A/D converters and a specific audio-over-ethernet chipset. The system, called Eigenmike™, makes it possible to measure 3D multichannel Impulse Responses, and at the same time provides a much finer spatial resolution with respect to the previous method.

2.1. Microphones

In order to measure the acoustic parameters in a room, the ISO 3382/2009 standard [28] specifies that the monophonic acoustic parameters must be measured with an omnidirectional, monoaural microphone. The standard only specifies the dimension of the microphones (normally, up to 13mm). In order to measure binaural Impulse Responses and IACC, the standard requires binaural
microphones (real heads or dummy heads) and describes their characteristics. Finally, the ISO 3382/2009 suggest using a figure-of-eight microphones to measure lateral-energy parameters, such as Lateral Fraction and LFC. Unfortunately, it does not give any technical specifications for these types of directive microphones.

Obviously, the microphones described in the standard could only be able to measure the acoustic parameters (even for spatial sound distribution), but neither the figure-of-eight microphones can completely describe spatial sound distribution in the room. This information could be obtained only using a multi-microphone system specifically designed.

2.1.1. **B-format microphones.** For many years, following the ISO 3382 standard, acoustic measurements in theatres and concert halls have been performed using monoaural microphones and dummy heads for calculating monoaural and binaural parameters. In order to measure 3D Impulse responses, the B-Format microphone (like Soundfield™) has been considered the optimal solution. This microphone could also be employed for calculating monoaural acoustic parameters, which could be obtained from the W channel. For calculating spatial parameters like LE, LF or LFC, the Y channel could be utilised to provides the necessary figure-of-eight signal. Moreover, the last two directive channels (X and Z) could be utilised to obtain the global 3D soundscape inside a listening room by means of the 1st order Ambisonics technology.

It should be considered that the 1st order Ambisonics playback system could not be considered able to reproduce spatial sound characteristics with enough precision for properly localise sound sources for a listener, because 1st order Ambisonics could not properly synthesize sound sources having high level of polarisation. In this case, the sound would be perceived as it was coming from some undefined large position. In order to overcome this limit, the high-order Ambisonics (HOA) systems could be used. Of course, HOA would require a multichannel microphone system which should correspond to spherical harmonics expansion of 3rd or 4th order at least. Unfortunately, it is well known that HOA is able to properly reproduce synthetic signals, in which the signals containing high-order spherical harmonics are computer-generated, whereas recording HOA signals is really difficult. This difficulty is caused by the complex microphone array composed by dozens of elements, and by the low value of S/N ratio, especially at low frequencies. Moreover, at high frequency the spatial resolution of HOA and of the microphone system are not matched. This very often provokes a reduction of the bandwidth one octave band or less.

One more possibility for properly reproducing spatial sound distribution could be a proper decoding method that could be applied to the B-format, 1st order signal. Following this possibility, it could be possible to make a “spatial analysis” of the signal, and afterwards to “steer” the incoming sound just in the very specific directions from where it arrives, for each instant and at each frequency. This possibility could be based on the Sound Intensity theory and on plane wave decomposition. This possibility has been recently proposed by Harpex by Berge and Barrett and Sirr/Dirac by Pullki [29].

2.1.2. **Large-number microphone array.** The Eigenmike™ probe, a 32-capsule microphone system, which was introduced a few years ago, represents a concrete possibility to synthesize “virtual microphones” having arbitrary directivity pattern, using a novel processing method which does not utilised High Order Ambisonics [23] This method has been utilised by some researchers, including some Authors of this paper.

2.2. **Spatial analysis of 3D Impulse Responses**

In order to take advantage of a great number of measurements performed in the last 10 years, a spatial analysis was developed and presented here. This method can analyse B-Format Impulse Response measured by means of a B-format microphone.

2.2.1. **Analysis from B-Format signals.** The new method starts from the characteristics of the B-Format signals, which are able to detect the direction-of-arrival of each arriving wave-front,
calculating the sound intensity vector \( I \) and the value of the energy ratio \( r_E \) for one specific instant. It is based on the same scheme initially proposed which uses vector decomposition, and it is also related to the later SIRR method [29]. The sound intensity vector can be simply obtained from the B-Format Impulse Response by means of the following equations, which give the three components of the vector:

\[
I_x = w \cdot x; \quad I_y = w \cdot y; \quad I_z = w \cdot z; \quad (1)
\]

Where \( w, x, y, \) and \( z \) are the four signals of the B-Format Impulse response. The total energy density could be calculated considering that these signals are directly proportional to sound pressure and particle velocity. The total energy density is therefore:

\[
E_D = \frac{\sqrt{w^2 + x^2 + y^2 + z^2}}{c} \quad (2)
\]

The modulus of the sound intensity vector is:

\[
|I| = \sqrt{I_x^2 + I_y^2 + I_z^2} \quad (3)
\]

The following equation gives the ratio between the active intensity and energy density:

\[
r_E = \frac{\frac{|I|}{E_D \cdot c}}{\sqrt{\frac{(w \cdot x)^2 + (w \cdot y)^2 + (w \cdot z)^2}{w^2 + x^2 + y^2 + z^2}}} \quad (4)
\]

The azimuth (horizontal) \( a \) and elevation (vertical) \( e \) angles could be obtained from basic trigonometric equations:

\[
a = \arctan \frac{I_x}{I_y}; \quad e = \arcsin \frac{I_z}{|I|} \quad (5)
\]

Applying these equations over 1 ms time interval, a sort of “time history” of the descriptors could be calculated for all the length of the Impulse Response measured in the room. Moreover, the results could be plotted dynamically by means of a tool which has been specifically developed, which is able to create for each frame a circle, located at the position \((a, e)\). The circle has a diameter which is proportional to the modulus \(|I|\) of the sound intensity, and the opacity which is proportional to \(r_E\). The circle is plotted on a panoramic image, 360 x 180 degrees, which represents a picture taken exactly at the microphone position. Moreover, there is a synchronized marker which could move over the energy graph of the Impulse Response, making it easy to check the arrival direction of each reflection. At the same time, the tool allows to see how much the corresponding wavefront is “polarised”.

When the sound energy coming from a surface is strongly concentrated along a specific direction, \(r_E\) is approaching to 1, which means traveling wave. On the other hand, when the sound energy is diffuse, \(r_E\) is approaching to 0, which means standing wave.

The figure reports the superposition of the energy part of the Impulse Response (i.e., ED expressed in dB) with the aligned dB scales of intensimetric IR (i.e., \(|I|\) expressed in dB) and not the “Impulse Response of \(r_E\)”. The alignment of both the dB scales have been calibrated in order to have for a perfectly plane and progressive wave (when \(r_E = 1\)) the two values for ED and \(|I|\) in dB which are the same.

It should be noted that the dynamic display of the duration of the Impulse Response of the spatial-temporal distribution of sound does contain not only the information of the trajectory of the reflected sound but also information about the degree of diffusion. When the level of ED is much larger than the level of \(|I|\) the sound is fully diffused (and hence \(r_E\) approaches 0). On the other hand, the sound is
strongly “polarised”, a propagative wave traveling in a precise direction when, instead, the two levels are almost equal (meaning that \( r_\text{E} = 1 \)).

This new method allows reconsidering the information of a large number of B-format IRs previously measured by the Authors in the last 15 years by means of Soundfield microphones or other similar tetrahedral probes, obtaining a lot of new information that goes beyond the traditional acoustic parameters.

2.3. Post processing

The measured 3D B Format Impulse Responses could be post processed in two manners:

- The spatial distribution of the energy reflected within the running time could be performed. This graphical processing allows to see from where the reflections of the walls are coming.
- The 3D rendering for a number of listeners could be realised, within a special listening room equipped with a suitable number of loudspeakers, able to include the listeners’ position within a sphere or a cube.

Two different tools could perform the graphical analysis:

- The first tool is able to report the “moving circle”, which represents the instantaneous point of arrival of the sound intensity, starting from the analysis of the corresponding B-format IR pleasureed according to the method described above.
- The second tool is able to create a video rendering of the sound map, which is over-plotted on the panoramic image. In this case, there is no need of any graphical algorithm, since the graphics library is employed to realise the colour map, which is based on the 32 “instantaneous” values of the sound level captured by the microphones (32 virtual channels).

By means of these two new tools, an animated video rendering of the instantaneous sound intensity vector, and of a colour map of sound distribution, could be calculated and plotted over the panoramic (360x180) image. Figure 2 report an example of a frame from this video. The 3D audio rendering is obtained by processing the original impulse response recording: the tool allows calculating a new set of virtual microphones, where each loudspeaker reproduce each channel of the virtual microphone. The processing is not the same both for B-format impulse response and for 32-channels impulse responses, even though a same methodological approach could be employed. A simple linear equation system allows calculating the set of filters employed for deriving the “playback” virtual microphones. I would be sufficient to set that the signals re-recorded placing the B-format microphone (i.e. the probe) in the centre of the playback system are maximally similar to the original waveforms recorded with the same probe in the theatre. This method, even if it is not Ambisonics-based, could corrects, both in terms of magnitude/phase response and in terms of placement/aiming/shielding, the playback system.

3. Results

The case study here considered consists of the post-processing of B-format IRs measurements made in the Basilica of San Vitale, Ravenna, Italy, and recently published [30]. In this example, the B-Format IRs were captured by a Soundfield™ MKV microphone. The input signal was an ESS signal ranging from 40 Hz to 20 kHz, played on a digitally pre-equalised omnidirectional loudspeaker (dodecahedron). During the measurement, the positions of the sound source and of the microphone were accurately monitored. Considering the position of both sound source and microphone, a panoramic picture (360x180) if the Basilica was taken, as reported in the Figure 2:

In this specific example, the 4-channel method allows to locate only very clear, early reflections, after which the value of \( r_\text{E} \) becomes too small and the sound field becomes completely diffuse. It means that the spatial analysis of late reflection become meaningless. This limit is caused by the high level of reverberation of the Church, where the late (reverberant) sound field in high, whereas the direct (early) sound field very short.
Even with this limit, the procedure here presented has demonstrated its feasibility to obtain spatial information about acoustics and sound propagation in rooms even from B-format measurements collected during past years. This of course might be possible only in case that the exact positions of sound source and microphone are very accurately measured. This offers the chance to address building physics comfort comprehensively [31,32].

4. Discussion
In this paper, two novel methods for analysing and measuring 3D Impulse Responses in concert halls, opera houses and auditorium have been presented. The two new methods could provide a spatial resolution much better than what is actually obtainable considering the existing technology (e.g. 1st order Ambisonics for B-format microphones). Moreover, the new methods allow to easily read and understand the spatial-temporal information obtained with B-format IRs measurements, by means of an animated graphical representation over a panoramic picture.

When the measurements in the auditorium are performed by a 32-channels microphone array, the first method allows displaying spatial information with high precision, being based on the synthesis of 32 highly-directive virtual microphones, which are calculated by means of a huge matrix of FIR filters.

In case the measurements are performed using 4-channel B-Format Impulse Responses, the second method could provide useful results, starting from the “instantaneous sound intensity” analysis. This method, even if less powerful if compared with the previous one, allows analysing the spatial sound propagation of auditorium measured in the last 15 years by means of B-format (normally, Soundfield™) microphones. However, the method requires that the exact position of both the sound source and the microphone are known and requires a panoramic (360x180) picture taken from each microphone position.

The new tools are able to let the acoustician to better know and understand the acoustics of the auditorium for one specific condition not only from the numerical values of the acoustic parameters but also finding out the directions from where the reflection are coming.

Moreover, using the 32-channel spherical microphonic array, the tool allows getting a useful colour map, which represents the full spatial sound distribution at a specific instant, thanks to the high number of virtual microphones, which can realise a spatial sampling of the whole sound field; On the
other hand, when the measurements are performed with a B-format microphone, the tool can display the direction of the instantaneous, total acoustic energy flux, which is provided by the sound intensity vector, together with a qualitative information of the degree of diffuseness of the sound field, shown by different colours, given by the ratio $r_E$.

This research, in very recent months, has moved to the possibility to use the database of B-Format IRs measurements with the new analysis recently proposed by some tools (e.g. Harpex, SIRR). This should be feasible, and it could allow obtaining high level of information in spatial sound distribution in auditorium simply using a cheaper microphone (Soundfield™, B-format) instead of an expensive one (Eigenmike™ probe, 32 channels).

Acknowledgments

The Authors wish to thank prof. Lamberto Tronchin and prof. Angelo Farina for their valuable help during the development of the research.

References

[1] Tronchin L and Farina A 1997 AES J. Audio Eng. Soc. 45(12) 1051
[2] Farina A, Langhoff A and Tronchin L 1998 J. New Music Research 27(4) 359
[3] Farina A and Tronchin L 2000 Acustica 86(4) 737
[4] Tronchin 2005 J. Acoust. Soc. Am. 117(2) 926
[5] Tronchin L 2012 AES J. Audio Eng. Soc. 60(12), 984
[6] Tronchin L and Coli V L 2015 AES J. Audio Eng. Soc. 63(9), 671
[7] Mancini F and Nastasi B 2019 Energies 12 2788
[8] Tronchin L, Manfren M and James P A 2018 Energy 165, 26
[9] Tagliabue L C, Manfren M, Cribini A L C and De Angelis E 2016 Energy Build. 128 119
[10] Tronchin L, Manfren M and Nastasi B 2019 Energy Procedia 157 1486
[11] Tronchin L and Fabbri K 2017 Energies 10 1621
[12] Castellani B, Morini E, Nastasi B, Nicolini A and Rossi F 2018 Energies 11(7) 1921
[13] Tronchin L and Fabbri K 2008 Int. J. of Exergy 5(5-6) 605
[14] Nastasi B 2019 Int. J. Hydrogen Energy 44 9485
[15] Tronchin L, Manfren M and Tagliabue L C 2016 Sust. Cities and Society 27 296
[16] Tronchin L Tommasino M C and Fabbri K 2014 Int. J. of Sustainable Energy Planning and Management 349
[17] Tronchin L 2013 Int. J. Mechanics 7(3) 318
[18] Piana E A, Petrogalli C, Paderno D and Carlsson U 2018 Appl. Sci 8
[19] Fabbri K and Tronchin L 2015 Energy Procedia 78 2778
[20] Castellani B, Rinaldi S, Morini E, Nastasi B and Rossi F 2018 Energy Conv. Manag. 171 626
[21] Farina A and Tronchin L 2005 Ac. Science and Technology 26(2) 193
[22] Farina A and Tronchin L 2013 Acta Acustica United Acustica 99(1) 118
[23] Martellotta F, Cirillo E, Carbonari A and Ricciardi P 2009 Appl. Ac. 70(2) 378
[24] Tronchin L 2013 Acta Acustica United Acustica 99(1) 91
[25] Shimokura R, Tronchin L, Cocchi A and Soeta Y 2011 J. of Sound Vibration 330 3526
[26] Farina A, Campanini S, Chiesi L, Amendola A and Ebri L 2014 AES 55th Int. Conference 113245
[27] ISO 3382-1 2009 — “Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces”
[28] Pulikki V 2007 AES J. Audio Eng. Soc. 55(6) 503
[29] Tronchin L and Knight D J 2016 Int. J. of Historical Archaeology 20(1) 127
[30] Manfren M, Nastasi B, Piana E and Tronchin L 2019 AIP Conference Proceedings 2123 020066
[31] Manfren M and Nastasi B 2019 AIP Conference Proceedings 2123 020100