ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Development of MIMO technique for 3D Auralization

This is the final peer-reviewed author’s accepted manuscript (postprint) of the following publication:

Published Version:
Bevilacqua A., Merli F., Tronchin L. (2021). Development of MIMO technique for 3D Auralization. Institute of Electrical and Electronics Engineers Inc. [10.1109/I3DA48870.2021.9610972].

Availability:
This version is available at: https://hdl.handle.net/11585/852888 since: 2022-08-30

Published:
DOI: http://doi.org/10.1109/I3DA48870.2021.9610972

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

20 December 2024

This is the final peer-reviewed accepted manuscript of:

A. Bevilacqua, F. Merli and L. Tronchin, "Development of MIMO technique for 3D Auralization," *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2021, pp. 1-5.

final published version is available online at: *https://dx.doi.org/10.1109/I3DA48870.2021.9610972*

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

Development of MIMO technique for 3D Auralisation

Antonella Bevilacqua
Department of Engineering and Architecture
University of Parma
Parma, Italy
antonella.bevilacqua@unipr.it

Francesca Merli
CIRI Edilizia e Costruzioni
University of Bologna
Cesena, Italy
francesca.merli8@unibo.it

Lamberto Tronchin
Department of Architecture
University of Bologna
Cesena, Italy
lamberto.tronchin@unibo.it

***Abstract* - Within the frame of SIPARIO, a project recently funded by the Italian region Emilia Romagna, it aims to virtually reconstruct both 3D audio and 360° video of real performances by recording and undertaking acoustic measurements inside the most historical theatres and concert halls spread all over the Europe. One of the most relevant topics of this project consists of the acoustic implementation of a new prototype of loudspeaker that will be necessary to properly determine the acoustic characteristics of these performing arts places. This loudspeaker can be employed in 3 different ways: one of the added values compared to a standard loudspeaker would be the simulation of a dynamic directivity pattern of any type of sound source that would be found in a performing arts place. This paper deals with the explanation of the other 2 types of techniques that the aforementioned prototype loudspeaker is able to perform. The loudspeaker can work like a traditional omnidirectional sound source by feeding the Exponential Sine Sweep (ESS) signal to all the 32 drivers in order to measure the room impulse response (RIR). Furthermore, it could have a complete spatial control of the sound propagation through the Multiple Input Multiple Output (MIMO) approach, owing to the spherical array of 32 channels for both source and receiver. In this paper the mathematical framework of previous articles has been applied to extract the arbitrary directivity of virtual microphones from real microphone array and for a virtual reality source from the spherical speaker array. Moreover, future developments of the new sound source are briefly described, including the applications for the specific project.**

I. Introduction

In room acoustics very often the main purpose of the experimental analysis is based on the measurements of RIR. These kinds of measurements employ a sound source which, even if the wavefront is not perfectly omnidirectional, is good enough to obtain acoustic information of a room. However, when the aim of the acoustic measurements attempts to rebuild complex polar patterns of sound source (e.g. human voice or specific musical instruments), the standard equipment (i.e. dodecahedron speaker) is limited. Moreover, in relation to performing arts places, where the radiation of polar patterns of singers and musicians changes continuously over time, the acoustic data obtained by measuring a mono signal recorded with a single microphone is absolutely not suitable for this purpose. As such, the MIMO approach is one of the most feasible methods to virtually reconstruct any real-time sound source with arbitrary directivity. Thus, a multichannel spherical loudspeaker array has been employed as a sound source for the MIMO room impulse response measurements or alternatively for tracking all the possible trajectories between the sound source and the receiver.

II. The Sipario project

Despite the important progression of the modern industry into the fields of science and technology during the last few decades, current research and investment have been moved towards digitalization and augmented reality. In this environment, the SIPARIO project has developed the idea of utilizing musical and theatrical products in a digital format through immersive reality of artistic performances executed in historical theatres. Considering halls and auditoria as an extension of musical instruments, SIPARIO has understood that all concerts and shows should be recorded and reproduced in such a way as to appreciate both the artistic experience and the acoustic information of historical architecture, whose access is often limited to preserve their integrity. One of the objectives of the project is to virtually create a 3D sound field and a 360° video where the audience can be approximated to an invisible point able to move all around the space (e.g. standing onto the stage, sitting in one of the balconies or from privileged positions). The facilitation of immersive listening experiences that respond to the listener's movement would be a different way to attract a wide audience, not only limited to experts, musicians or acousticians but also to involve tourists and amateurs of arts. This plan can easily network international theatres in partnership to exchange technical information and competencies in addition to building up collaboration with operators having common issues to solve. As such, the necessity to develop new techniques of RIR measurements for an advanced auralization justifies the introduction of a new generation of spherical microphone and speaker arrays.

III. Microphone array

The process technique to derive sound signals of any array of *M* "real microphones" and *V* "virtual microphones" with specific directivity is called beamforming. The following figure and equation show a generic signal processor where filters' coefficients are determined.

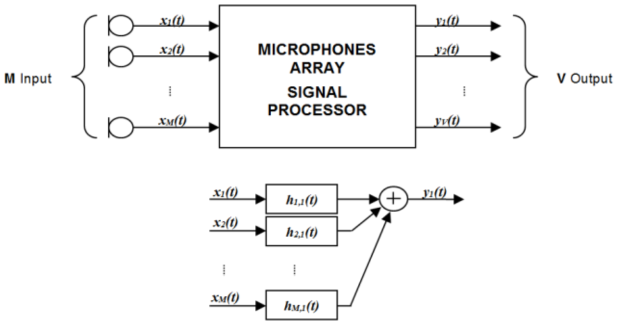

Figure 1: Representation of the microphone array signal processor [1]

$$y_v(t) = \sum_{m=1}^{M} x_m(t) * h_{m,v}(t) \qquad (1)$$

A technique to find the filtering coefficients, $h_{m,v}$, was described extensively in [1]. This technique is applicable to any microphone array by using the numeric inversion of the characterization impulse response of the microphones measured in an anechoic chamber or simulated by a variety of methods (Closed form mathematics, FEM, etc.).

In particular, the impulse response of the *M*-channel microphone array has been captured using incident plane waves from a large number of characterization directions *D* (usually 362 directions with nearly uniform distribution that constitutes an adequate spatial oversampling for a 32 channel microphone array like the em32 Eigenmike®, a product by mh-acoustics). Each measurement determines a row of a matrix $c[n]$ of size $D\times M$ (see Fig. 4).

Figure 2: Microphone em32 Eigenmike® (mh-Acoustics product)

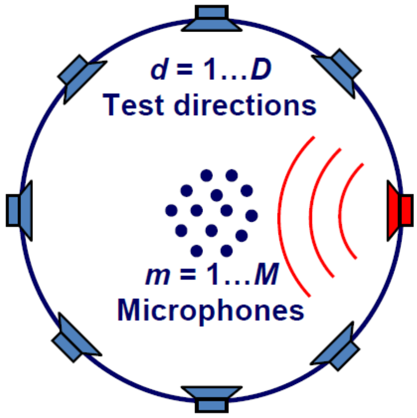

Figure 3: Microphone array characterization [1]

$$c[n] = \begin{bmatrix} \boxed{c_{1,1} & c_{1,2} & \cdots & c_{1,m} & \cdots & c_{1,M}} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,m} & \cdots & c_{2,M} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{d,1} & c_{d,2} & \cdots & c_{d,m} & \cdots & c_{d,M} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{D,1} & c_{D,2} & \cdots & c_{D,m} & \cdots & c_{D,M} \end{bmatrix}$$

Figure 4: Matrix representation of a microphone array [1]

Therefore, it is necessary to define the target directivity *Q* having dimensions $D\times V$ which synthetizes, for each of the *V* virtual microphones, the target directivity in each direction *D* of the array characterization.
These directivity amplitude gains could be arbitrarily defined; however, it is convenient to use a mathematical definition as for the directivity of a cardioid of n^{th} order.

$$Q_n(\vartheta,\varphi) = [0.5 + 0.5\cdot\cos(\vartheta)\cdot\cos(\varphi)]^n \qquad (2)$$

Because the set of filters *h* can be derived directly from a set of measurements, the matrix has to be numerically inverted by using a technique called Least Squares (LS) plus regularization. In this way the outputs of the microphone array are very close to the number of responses.
It is necessary to solve this matrix using a Discrete Fourier Transform (DFT) to pass into the frequency domain. In this way the overdetermined linear system can be evaluated separately for each frequency index *k*.

$$\overset{D\times M}{\|C[n]\|} * \overset{M\times V}{\|H[n]\|} = \overset{D\times V}{\|Q\|} \cdot \delta[n] \overset{DFT}{\Rightarrow} \overset{D\times M}{\|C[k]\|} \cdot \overset{M\times V}{\|H[k]\|} = \overset{D\times V}{\|Q\|} \qquad (3)$$

The solution could be obtained as reported in the following equation, which is the solution using the Least Squares method. A casualization term has been added to the numerator as well as a regularization term to the denominator.

$$\overset{M\times V}{\|H[k]\|} = \frac{\overset{M\times D}{\|C[k]\|^*} \cdot \overset{D\times V}{\|Q\|} \cdot e^{-j\pi k}}{\underset{M\times D}{\|C[k]\|^*} \cdot \underset{D\times M}{\|C[k]\|} + \beta[k] \cdot \underset{M\times M}{\|I\|}} \tag{4}$$

To find H as the Hermitian operator (conjugate transpose) a casualization term $e^{-j\pi k}$ was necessary to be added to the numerator as a pure delay. It is equivalent to the half filter length in order to synchronize the time zero with the central sample of the sequence. The same concept was applied to the denominator by adding a frequency-dependent regularization term $\beta[k]$. β has been kept high at the boundaries of the frequency band of the transducers in order to promote the filters' convergence and it has been reduced at the center of the frequency band in order to promote a more precise inversion.

Finally, the filters $\overset{M\times V}{\|H[k]\|}$ determined in the frequency domain could be reversely transformed into time-domain through the Inverse Fast Fourier Transform (IFFT) in order to obtain a matrix that represents the FIR filters coefficients, useful for the virtual microphone *V*.

$$\overset{M\times V}{\|h_{MA}[n]\|} = IDFT\left[\overset{M\times V}{\|H[k]\|}\right] \tag{5}$$

The choice of the samples number should be made in order to avoid the phenomenon of circular rotation of the signals owed to the introduction of the casualization term.

IV. SPEAKER ARRAY

The first prototype of spherical array loudspeaker realized in 2015, as a product of the research studies conducted by the University of Parma, consists of a 200mm diameter wooden sphere equipped with 32 2" drivers (i.e. RCF MB2N101), as shown in Figure 5.

Figure 5: Prototype spherical array loudspeaker [3]

The necessity of creating this new loudspeaker was given by the reproduction of the same number of channels as equipped the microphone been employed (i.e. EM32 Eigenmike®) for the measurement tests, and also by the independency of all the 32 channels that can be fed by the signal individually.

Likewise, in the previous discussion related to the microphone array, it is possible to create radiation patterns of arbitrary shape in relation to the speakers. It is necessary to define a processing system in order to generate driving signals for *S* individual transducers in order to create a similar matrix of FIR filters to create *W* virtual sources having arbitrarily directivity defined in advance.

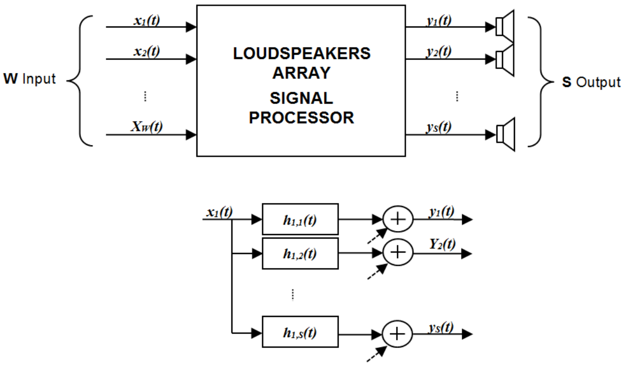

Figure 6: Representation of the speaker array signal processor [3]

Also in this case the signal processing is obtained through filters operations, but the complexity dwells into the determination of filters coefficients necessary to accomplish the directivity of the array requirements. The formula (1) is therefore modified into formula (6) below.

$$y_s(t) = \sum_{w=1}^{W} x_w(t) * h_{w,s}(t) \tag{6}$$

The impulse response obtained by test measurement of each transducer should be adjusted to the array in compliance with each direction *D* characterizing the system. The result would be the determination of the columns of a matrix $C[n]$ having dimensions $D\times S$. Considering that the prototype speaker has 32 transducers installed as uniformly as possible over a surface of a sphere, directions *D* equal to 362 is the number to be applied to acquire a nearly uniform distribution.

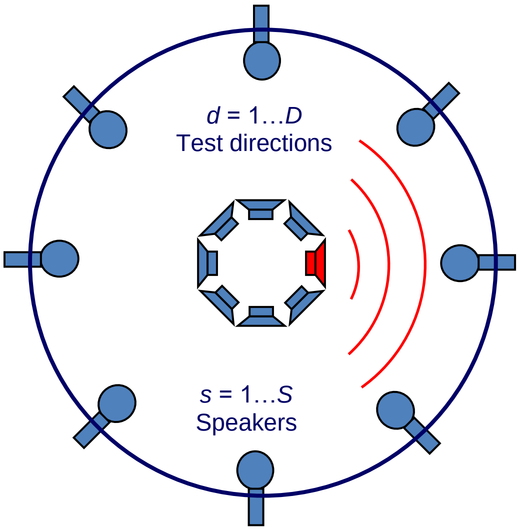

Figure 7: Speakers array characterization [3]

$$C[n] = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,s} & \cdots & c_{1,S} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,s} & \cdots & c_{2,S} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{d,1} & c_{d,2} & \cdots & c_{d,s} & \cdots & c_{d,S} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{D,1} & c_{D,2} & \cdots & c_{D,s} & \cdots & c_{D,S} \end{bmatrix}$$

Figure 8: Matrix representation of the speaker array [3]

For every W virtual source the following step is to determine the matrix giving the target directivity Q with dimension $D\times W$. Likewise, for the virtual microphones, it is therefore possible to determine the following system passed to the frequency domain, in order to simplify the convolutions.

$$\overset{D\times S}{\|C[n]\|} * \overset{S\times W}{\|H[n]\|} = \overset{D\times W}{\|Q\|} \cdot \delta[n] \overset{DFT}{\Rightarrow} \overset{D\times M}{\|C[k]\|} \cdot \overset{M\times V}{\|H[k]\|} = \overset{D\times V}{\|Q\|} \qquad (7)$$

The equation (7) should be adjusted by adding a Kirkeby regularization factor for MIMO systems resulting in

$$\overset{S\times W}{\|H[k]\|} = \frac{\overset{S\times D}{\|C[k]\|^{*}} \cdot \overset{D\times W}{\|Q\|} \cdot e^{-j\pi k}}{\overset{S\times D}{\|C[k]\|^{*}} \cdot \overset{D\times S}{\|C[k]\|} + \beta[k] \cdot \overset{S\times S}{\|I\|}} \qquad (8)$$

The result should be converted back to the time domain with the addition of the transposition of the filter resulting from the synthesis process, in order to preserve the conventional operation direction (from input space of W directive sources to the output space of S speaker feeds). The following equation summarizes this result:

$$\overset{W\times S}{\|h_{SA}[n]\|} = IDFT\left[\overset{S\times W}{\|H[k]\|}\right]^{T} \qquad (9)$$

V. Procedure of MIMO impulse responses analysis

In accordance with the ISO 3382:2012, measuring techniques for the impulse responses of musical performance rooms (i.e. theatres, auditoria, concert halls, TV studios) can be undertaken with both traditional tools (i.e. balloons, firecrackers, pistol) and modern electrical instrumentations (i.e. dodecahedron), as long as the sound source reaches a power energy at least 45 dB above the background level. Because the purpose of the impulse response measurement is to calculate objective acoustic parameters (mainly the reverberation time), the operational procedure indicates the sound source to be placed in one or more points on the stage (where normally the actors stand) whilst an omnidirectional microphone is capturing the sound signal in selected points, conventionally in correspondence of one or more seats of the audience. For each combination (source vs microphone position) a single channel impulse response is gathered.
By developing this concept, the result can be obtained in a similar way by employing the proposed spherical speaker array and the spherical 32-channel microphone array (i.e. em32 Eigenmike®) with a sound signal of an Exponential Sine Sweep (ESS).

An explanation of how the signal has been generated and processed is briefly introduced. The core of the setup is the audio interface (ORION 32, *Antelope Audio*) which synchronizes the reproduction and recording of the audio test signal ESS via a multichannel audio digital interface (MADI).

Figure 9: Audio interface ORION 32, by Antelope Audio

The audio interface ORION 32 is the instrumentation through which the computer sends the EES signal to the amplifier of the spherical speaker array. The process of information acquisition starts with the em32 Eigenmike® microphone that acquires the room responses and sends them to the ORION 32 through its Eigenmike® Microphone Interface Box (EMIB) and, consequently, to the PC for recording and storage.

For each pair of source-receiver positions, the measuring process should consider only one speaker a time having a signal direction coming from one of the 32 transducers, as previously discussed.
To be noted that in large halls variations of temperature and hence of the sound velocity with time and position cannot be entirely avoided. In addition, because of the air conditioning, the air inside such halls is not completely at rest. The propagation of sound wave certainly depends on variable parameters (i.e. air density ρ, temperature T, pressure p, angular frequency ω) and a slight attenuation at high frequencies is inevitable but if it is considered that such procedure and equipment have been developed for Italian historical theatres, having similar volume size (with

a capacity ranged between 1100 and 1600 seats, except for Teatro alla Scala of Milan having 2000 seats), the MIMO measurements would be undertaken in a limited short time period that these inhomogeneities become so small that they can be neglected.

The convolution of the recorded signals, coming from every speaker's direction individually, with the inverse sweep determines the matrix IR_{MIMO} having size $S \times M$, where S = 32 speakers and M = 32 microphones.

$$\|\overset{S\times M}{IR_{MIMO}}[n]\| = \begin{bmatrix} ir_{1,1} & ir_{1,2} & \dots & ir_{1,m} & \dots & ir_{1,M} \\ ir_{2,1} & ir_{2,2} & \dots & ir_{2,m} & \dots & ir_{2,M} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ ir_{s,1} & ir_{s,2} & \dots & ir_{s,m} & \dots & ir_{s,M} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ ir_{S,1} & ir_{S,2} & \dots & ir_{S,m} & \dots & ir_{S,M} \end{bmatrix} \quad (10)$$

Appropriate beamforming filters for both source and receiver could now be applied to the matrix in (10). In this way a new beam-formed matrix should be able to provide a specific directivity in relation to the pattern that is planned to be achieved, as shown in equation (11).

$$\left\|\overset{W\times V}{IR_{MIMO-BF}}[n]\right\| = \left\|\overset{W\times S}{h_{SA}}[n]\right\| * \left\|\overset{S\times M}{IR_{MIMO}}[n]\right\| * \left\|\overset{M\times V}{h_{MA}}[n]\right\| \quad (11)$$

In particular, this powerful matrix is able to create different types of beamforming, for both the source and the receiver. Some of the applications are given below:

- Omnidirectional beamformer with a single channel;
- SPS (Spatial PCM Sampling) beamformer composed of 122 virtual microphones/speakers with 8th order cardioid directivity in order to uniformly cover the whole solid angle [2].

In other words, the matrix previously discussed allows one to perform simple calculations, like the standard single-channel RIR used for the determination of the acoustic parameters as recommended by the ISO 3382:2012, [12], and more complex auralization design. A few more challenging applications are advanced High Order Ambisonics (HOA) or the synthesis of 122 very directive patterns as indicated in (13), for both virtual source and receiver.

$$IR_{OMNI}[n] = \left\|\overset{1\times S}{h_{SA-OMNI}}[n]\right\| * \left\|\overset{S\times M}{IR_{MIMO}}[n]\right\| * \left\|\overset{M\times 1}{h_{MA-OMNI}}[n]\right\| \quad (12)$$

Equation (12) is a classical single-channel impulse response with omnidirectional directivity, obtained by a simple transfer function where details in relation to the spatial information are not required [4].

$$\left\|\overset{122\times 122}{IR_{MIMO-SPS}}[n]\right\| = \left\|\overset{122\times S}{h_{SA}}[n]\right\| * \left\|\overset{S\times M}{IR_{MIMO}}[n]\right\| * \left\|\overset{M\times 122}{h_{MA}}[n]\right\| \quad (13)$$

Equation (13) represents the development of an SPS methodology given by the synthesis of very directive patterns for both source and receiver [4]. In this way the spatial energy distribution can be encoded as a square matrix of impulse response measured between a large number of highly directive virtual sources and a large number of highly directive virtual receivers.

The multitasking matrix can be employed indeed for different scenarios of real-time auralization in order to have a complete spatial control of the surrounding sound field.

VI. IMPACT ON AUDIO-VIDEO PROCESSING AND ON SIMILAR AREAS

The possibility to enjoy musical performances through virtual reality systems assumed recently a considerable importance given the global conditions of COVID-19 pandemic that forces people to stay home, missing the plan of assisting personally to concerts and other public events. Other than reproducing faithfully the physical and acoustics parameter of a historical architecture, SIPARIO project aims to approach people that never attended such live performances by using inexpensive devices (e.g. smartphone, head-mounted displays plus headphones) in order to perform the auralization of these venues. As such, a comparison between the innovative ideas proposed by SIPARIO and the traditional ways in reproducing live events as attended throughout the last decades would be much more intriguing. The passage to an immersive 3D virtual video-sound would give the same feedback as when actors and musicians play and/or sing live, with the sensation to be in a real theatre. Furthermore, a high-quality listening combined with the possibility to explore unique architecture throughout a 360° video gives the choice to be in any part of the world just by staying in a common room.

It has been thought that robots became more pervasive in modern age and the progress compared to the previous century, when the machines heavily contributed to support hard and physical labor of human workers, nowadays requires to move towards the field of machine learning, artificial intelligence (AI) and speech recognition. It would be worthwhile if the explained methodology used by SIPARIO would be applied to teach robots how to listen, isolate and recognize the location of a sound in any environment. The artificial learning presented with example input and giving an optimized output based on the interaction with the environment aims to a dynamic and active outcome. This idea can be developed even to free learning algorithms in a way that the robot would be completely free to find the structure of the input, interacting on its own. The audio input library for robots, mainly based on isolated sounds, should be implemented with overlapped sounds as it is in the real world. Technically, the convolution of audio processing should mime the resemblance of the sparse coding sounds to the hearing filters of the auditory cortex.

VII. CONCLUSIONS AND FURTHER RESEARCH

Italian opera theatres belong to an important and specific category of buildings and they are subject to continuous and intense architectural studies across the world. Although most of them are protected by international organizations (e.g. UNESCO) that limit the easy access in agreement with the local authorities, the enthusiasm of research studies in relation to the acoustics of these unique premises is persistent, especially throughout the last decades. Alongside the aim of faithfully reproducing the physical and acoustic characteristics, one of the scopes of SIPARIO project is to realize an immersive listening room that gives the possibility to virtually live the experience of freely walking around inside these historical buildings during a musical or art performance.

To accomplish this target, the technology has the primary aim of providing these unique experiences to the large percentage of people that have never attended such live performances. As such, the 3D audio reproduction with wavefield analysis and/or Ambisonics together with the 360° spatial navigation is one of the methods to make such experiences more attractive and approachable.

The MIMO auralization is necessary for this project to reconstruct the virtual reality [13, 14]. By known input data (i.e. geometrical positions of source and receiver, path distance of the sound ray, direction of emission from the speaker, direction of reception at the microphone position) it is possible to determine the trajectories of all the reflections and potentially also the intersection point between the direction of the ray coming from the source and the direction of the ray arriving into the microphone after hitting the finish surfaces of boundaries. For each theatre that has been tested already, massive data of large RIR matrices have been managed up to 3^{rd} order reflections, including for tracing the trajectories of single rays. It would be challenging to push beyond these current capabilities in order to simulate a real sound decay but could be a future extension of this study.

The other matters that can be objects of further development are the following topics:

- Frequency response;
- *n*-channel microphone array.

In [4], a considerable number of tests have been carried out in relation to the polar patterns of the prototype loudspeaker used like an omnidirectional source. The results, after simulating n^{th} order cardioid, showed that the new speaker array has a very good performance of a frequency response between 125 Hz and approximately 4-5 kHz considering a synthesis of a 4^{th}-order cardioid. Above this upper limit the spatial aliasing effects disrupt the capability of getting a correct spatial control of the radiated sound. Nevertheless, the prototype speaker under these circumstances can emulate a virtual source with an arbitrary directivity pattern typically met in a performance arts room (e.g. human voice, musical instruments). Definitively a wider frequency range response is able to amplify the depth of the virtual audio and, hence, more realistic

In Section 3 it has been explored about how to compute the beamforming of any array of virtual microphones. The microphone array used for MIMO measurements of SIPARIO project is a spherical 32-channel produced by the mh-acoustics, called em32 Eigenmike®, available on the market. The em32 Eigenmike® is going to be replaced by a new spherical microphone array made by the same factory, called em64 Eigenmike®.

Certainly, a double number of channels (i.e. 64) will require a massive convolution processor able to calculate a real-time synthesis of the filtering coefficients. The authors already identified an innovative support like Dante Brooklyn II by Audinate, an interface audio device able to module up to 64 channels.

Despite not yet being available on the market and hence to be tested, the expected results of the em64 Eigenmike® would include a creation of more highly directive virtual microphones, with a beam width sharper than what obtained with 32 channels.

VIII. ACKNOWLEDGMENTS

The authors wish to thank Lorenzo Chiesi and Angelo Farina, for their constant effort in supporting the realization of this study. This research was partially funded by the project “SIPARIO - Il Suono: arte Intangibile delle Performing Arts – Ricerca su teatri italiani per l’Opera POR-FESR 2014-20”.

IX. REFERENCES

[1] A. Farina, A. Capra, L. Chiesi, L. Scopece: “A Spherical Microphone Array for Synthesizing Virtual Directive Microphones in Live Broadcasting and in Post Production”. *AES 40th International Conference*, 2010 October 8-10, Tokyo, Japan.

[2] A. Farina, A. Amendola, L. Chiesi, A. Capra, S. Campanini: “Spatial PCM Sampling: A New Method For Sound Recording And Playback”. *AES 52nd International Conference*, 2013 September 2-4, Guildford, UK.

[3] A. Farina, L. Chiesi: “A novel 32-speakers spherical source,” *AES 140th Convention Program, e-Brief 258*, 2016 June 4-7, Paris, France.

[4] A. Farina, L. Chiesi: “Measuring Spatial MIMO Impulse Reponses in Rooms Employing Spherical Transducer Arrays”, *AES Conference on Sound Field Control*, 2016 July 18-20, Guilford, UK.

[5] L. Gomez-Agustina, J. Bernard: “Practical and technical suitability perceptions of sound sources and test signals used in room acoustic testing”, *Internoise*, 2019 June 16-19, Madrid, Spain.

[6] L. Tronchin: “The emulation of nonlinear time-invariant audio systems with memory by means of Volterra series”, J. Audio Eng. Soc., 60 (12), 984.996 2012

[7] L. Tronchin, V.L. Coli: “Further investigations in the emulation of nonlinear systems with Volterra series”. J. Audio Eng. Soc. 63, 671–683 2015

[8] C. Nachbar, F. Zotter, E. Deleflie, A. Sontacchi: “ambiX - A Suggested Ambisonics Format,” in Proceedings of the 3rd Ambisonics Symposium, Lexington, KY June 2-3, 2011

[9] O. Kirkeby, P.A. Nelson, H. Hamada, F. Orduna-Bustamante: “Fast deconvolution of multichannel systems using regularization”, IEEE Transactions on Speech and Audio Processing Volume 6, Issue: 2, Mar 1998.

[10] A. Farina: “Simultaneous measurement of impulse response and distortion with a swept-sine technique” - 108th AES Convention, Paris 18-22 February 2000

[11] A. Farina: “Advancements in impulse response measurements by sine sweeps” - 122th AES Convention, Vienna, Austria, 5-8 May 2007

[12] International Organization for Standardization. *ISO 3382-2:2008 Acoustic-Measurement of room acoustic parameters.* Geneva, Switzerland 2008.

[13] L. Reed, P. Phelps, “Audio reproduction in virtual reality cinemas - Position paper”, 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 – n. 8797904, 1513-1516, 2019,

[14] J. Herre, J. Hilpert, A. Kuntz and J. Plogsties, "MPEG-H 3D Audio-The New Standard for Coding of Immersive Spatial Audio", IEEE Journal Of Selected Topics In Signal Processing, vol. 9, no. 5,. 770-779, August 2015.