Cinema Sound and audio 3D: The case study of Cinema Lux, Rome, Italy

Published Version:
Dolci M., Tronchin L., Merli F., Bevilacqua A. (2021). Cinema Sound and audio 3D: The case study of Cinema Lux, Rome, Italy. Institute of Electrical and Electronics Engineers Inc. [10.1109/I3DA48870.2021.9610881].

Availability:
This version is available at: https://hdl.handle.net/11585/852883 since: 2022-02-04

Published:
DOI: http://doi.org/10.1109/I3DA48870.2021.9610881

Cinema Sound and audio 3D: the case study of Cinema Lux, Rome, Italy

Marco Dolci
Gruppo CSA S.p.A.
Rimini, Italy
mdolci@csaricerche.com

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

Francesca Merli
CIRI - Interdepartmental Centre for Industrial Research in Building and Construction,
University of Bologna
Bologna, Italy
francesca.merli8@unibo.it

Antonella Bevilacqua
Department of Industrial Engineering
University of Parma
Area delle Scenze, Parma, Italy
antonella.bevilacqua@unipr.it

***Abstract*— The cinema "Lux" is one of the most relevant cinema in Rome. It is located in the historical city centre of the city, close to the Tevere river. The acoustics of the cinema Lux has been analysed considering different approaches. First of all, monoaural acoustic parameters have been obtained following the ISO 3382 Standard. The binaural acoustic parameters have been obtained employing a dummy head (namely Neumann KU 100) and the B-Format impulse responses (utilized for calculating Lateral Fraction and Lateral Efficiency). Moreover, the 3D acoustic information has been obtained employing a 32 channels Eigenmike Microphone, equipped with a 360 degrees camera, in order to reconstruct the reflection paths from the cinema. This paper will present some of the outcomes of the measurements, together with an analysis and a comparison with other similar cinemas.**

***Keywords*— acoustic parameters, spherical array microphone, cinema, spatial PCM sampling, dubbing room**

I. Introduction

As Moore predicted in 1965, the evolution of technology has taken on an exponential trend in recent decades and this phenomenon has obviously also affected the audio department, allowing wider use of cheap but good quality materials. New streaming platforms provide good-quality entertainment material for both video and audio. Film proprietors should confront the approach of new diversion stages, for example, Netflix and Amazon prime video that keep individuals at home. Further developing sound quality with more vivid frameworks that are now accessible available can be the way. Likewise, they should forsake or if nothing else overhaul the old adjustment strategies like RTA Analysis through evening out, being this procedure lacking logical help and if not utilized as expected can demolish sound insight. Specifically, the utilization of a framework that can all the while record sound toward each path, like Eigenmike™, can be a right apparatus for imagine the wellspring of aggravation in an exhibition room. With the information on how the sound is spatially circulated in the room, one can follow up on the components that trouble the discernment subsequently working on the acoustic nature of the space. In order not to lose its effectiveness, the cinema environment must commit itself, despite technological growth, to offer the user an experience that he will not be able to find elsewhere. In particular, the ideal objective is to ensure a reproduction in the performing room identical to that in the Dubbing Room (Cinema Mix). Already in the past efforts have been made in this direction looking for standards that allow a faithful reproduction in the path of sound that goes from the Mix Room to the final destination of the cinema and the multiplex.

The standardization operations used today are often based on a system that lacks valid scientific support. One of the calibration systems of cinemas that is still implemented is based on RTA analysis via pink noise. According to THX and many others, this system would allow the room to be 'equalized' according to the form that results from frequency analysis of pink noise recording at 2/3 of the room, flattening it if there are peaks or valleys going to adjust the gains before the sound is reproduced by the electroacoustic system. This measurement system might justify equalization to get a good flat response for a specific location but forget about the other seats in the room in the listening area. The cinema being a vast environment, the answer can change tremendously, because of the ways, as soon as we move to measure a new location and then already from this fact you can see a lack of scientific coherence. As we will see it is not possible to equalize a room and so if we want to improve the quality of entertainment this is definitely not the way to go. To understand the acoustic differences between the mix room and the cinema, we conducted a campaign of measures in Rome studying both environments on site.

II. Cinema Standards

A. RTA Method

In this chapter we are going to focus more on the acoustical part of Cinema Environment describing in detail how the measurements were made in the past and how they should be done today.

Fig. 1. RTA Pc interface

The recommended method in “X-Curve” document suggests the use of RTA analysis to equalize the room under study. The RTA system well served the cinema industry for more than thirty years and now it’s becoming obsolete.[1] Its assumption is: “a steady state measurement of the frequency response in the far field of a cinema when adjusted to a specific characteristic (X-curve) will ensure a flat frequency response in the near field”. RTA is a time blind measurement so it: cannot provide the complex response of the device under test which should include both magnitude and phase response, cannot shows the way the total system (Loudspeaker + room), frequency response evolves over time at a specific location, cannot isolates direct sound and later sound so cannot shows how the reproduced sound would be perceived by our hearing system which is dependent on the direct sound.

B. *Critics to RTA Measurement method*

In RTA if we measure the signal in one location, from the frequency analysis we distinguish peaks and valley and what everybody was doing in film industry was attempting to fill in those pinnacles and valley through balance. Through equalization they were able to fit the frequency response into the X-Curve so they could respect the standard. It must be said that the equalization step happens before the signal reaches the reproduction system so the myth that one can “equalize” a room is wrong because the coloration added by the environment (last filtering process before the sound arrives at the ears) is still present. A consequence of this kind of approach was that a technician had to calibrate the movie theatres frequently because something drifted in the frequency response (changing the microphone position of some cm results in a completely different frequency response). Another wrong element present in the X-Curve standard is that the equalization takes place by studying a frequency response that is an average of five microphones measurements. The average has sense if we take it in the same measurement location, but has no purpose if it is made on five recording positions. Moreover, we can have five bad frequency response, but their average is flat so one is brought to assume that the quality target is reached when it’s not. All these scientific misunderstanding lead to the statement from SMPTE technicians that sometimes the equalizer was useless, and many skilled listeners appreciated more the soundtrack without the equalization, judging it to be cleaner and more enjoyable. In the SMPTE report of 2014 the technicians made the measurements of the calibrated rooms with FFT analysis of the impulse responses and discovered the sorry state of cinema sound [1]. There was no match whatsoever between the frequency responses of the various measured cinemas and dubbing stages. The chain that brings the soundtrack from the re-recording mixer ears to the spectator in the performance theatre is broken and this is mainly due to the implementation over the years of the wrong calibration method. So it’s impossible to equalize a room and the only way we could fix this chain is to act directly on the acoustics of the room.

III. MEASUREMENTS

Measurements are performed in two rooms representing the starting point (dubbing stage room) and the destination (cinema performance room) of the soundtrack of a commercial movie. The dubbing stage room is a simple rectangular room of 5,75 mt x 7.55 mt x 3,15 mt. A big mixing-desk, two pc screens, and two rows of 4 padded chairs in the back are the main elements present in the room.

The geometry of room 10 of cinema Lux is more complex than the simple rectangular one of the Mixing Room, with its four angles are cut with oblique walls with different lengths.

The room dimensions are approximately 11 mt x 9 mt.

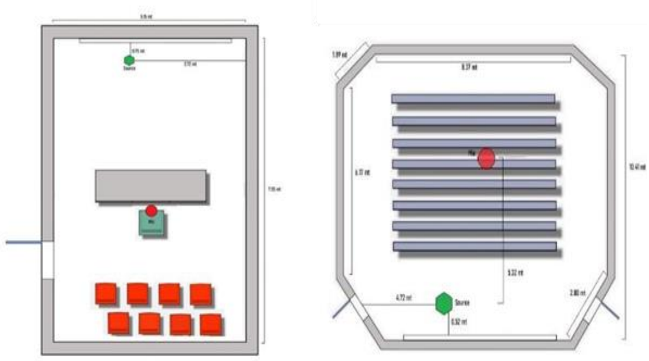

Fig. 2. Dubbing stage Room on the left and Cinema Lux on the right, with indication of measurement layout. The two green marks indicate the position of the sound sources, instead the reds marks indicate the position of the microphones.

The acoustic comparison between the two venue was carried out by utilising two different measurement techniques. The first technique is the traditional one and consisted in measuring the monaural and binaural parameters in line with the ISO 3382-1 [2] and provide quantitative parameters to compare dubbing stage room and room 10 of cinema Lux.

The second one is a qualitative comparison of 3D recording of sound and intended to capture all the effects necessary for an acoustic rendering suitable for virtual reality applications with the support multichannel spherical array microphone [6].

The acoustic survey was performed with the following instrumentation:

A. *Traditional measurements*

- Neumann KU Dummy Head
- Sennheiser Ambeo Soundfield microphone
- Behringer ECM 8000 Omnidirectional microphone
- Omnidirectional sound source S103AC Lookline dodecahedron
- Personal Computer connected to the loudspeaker and all the receivers.

B. *Measurements with sperical microphone array*

- 32-channel spherical array (Mh Acoustic em32 Eigenmike®);
- Audio interface EMIB

- Personal Computer connected to the loudspeaker and all the receivers.
- 360° camera (Rico Teta V).

Fig. 3. Eigenmike system layout

In the dubbing room the omnidirectional sound source for the Sweep test have been placed asymmetrically at 2,72 mt from the right wall and the recordings have been done at the re-recording mixer's location in as shown in Fig. 2. In room 10 of Cinema Lux the source for the Sweep test have been set asymmetrically at 4,72 mt from the right oblique wall and at 0.52 mt from the screen and the recordings have been done in an asymmetric point at row n.5.

The excitation signal emitted by the sound source was the Exponential Sine Sweep (ESS) having a duration of 15 s in a uniform sound pressure level for the range between 40 Hz and 20 kHz [3]. The measurements have been undertaken in unoccupied condition.

IV. RESULTS

In the following section, are presented the results obtained with the two implemented measurement systems.

A. Traditional parameters

The recorded ESS signals have been processed by using the plugin Aurora suitable for Audition 3.0 [4] Different acoustic parameters defined by the international standards ISO 3382-1[5] have been examined. The energetic and spatial parameters like: T30 (Reverberation Time), EDT (Early Decay Time), C50 , C80 (Clarity), D50 (Definition) and the spatial parameters like IACC (Interaural Cross Correlation) and LF (Lateral Fraction) have been analyzed.

The figures below report the main acoustic parameters in the octave bands between 125 Hz and 16 kHz.

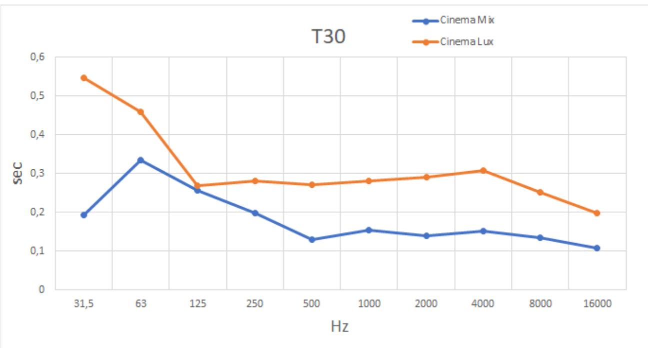

Fig. 4. Comparison between T30 measured in Cinema Mix and Cinema Lux

Cinemas and Conference rooms request a value that goes from 0,5 s to 1,5 s and in our measurements in cinema Lux we found values standing below this range.

The Dubbing stage Room instead can be considered a television studio with Reverberation Time values ranging from 0,25 s to 0,75 s and the value we found falls just at the beginning of this interval. The optimal values of these parameters are reported in Fig. 5. [8]. The Orange curve representing the values of T30 of Cinema Lux is always above the Dubbing stage blue curve except for the 125 Hz octave where they match.

As expected for both rooms we noticed a general higher T30 in the low frequencies and a gentle roll off towards mid-high frequencies.

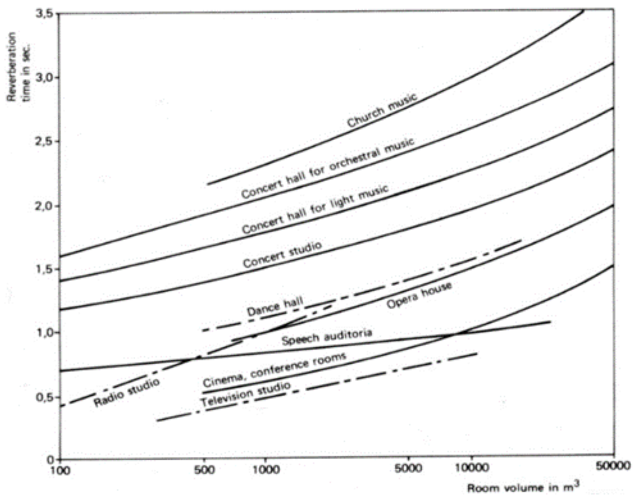

Fig. 5. Ideal RT values for different performance spaces.

Typically, the parameters T_{20}, T_{30} and EDT in a room where the sound field is strongly diffuse should be identical. The ideal value for EDT stands in the range 0.75 T_{mid} and 0.9 T_{mid} where T_{mid} is the average value of T30. T_{mid} is the average of T30: for Dubbing stage is 0.22 s and for Cinema Lux is 0.28 s. From EDT analysis it can be immediately notice that for 125 Hz octave band the Dubbing stage room shows a higher value than Cinema room and this is probably due to the interaction of a mode in this octave range that creates a resonance so that the sound at that frequency is slower to decay.

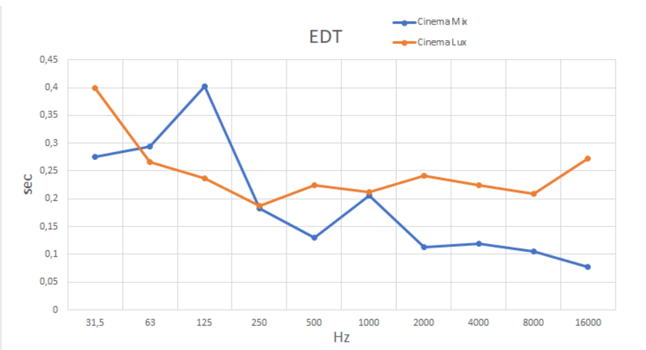

Fig. 6. Comparison between EDT measured in Cinema Mix and Cinema Lux

The parameter D50 is related to the intelligibility of syllables in speech and by literature [9], a good speech definition is defined to have values higher than 0.5 (50%). In both venues analyzed we have found higher values for this parameter like shown in Fig. 7.

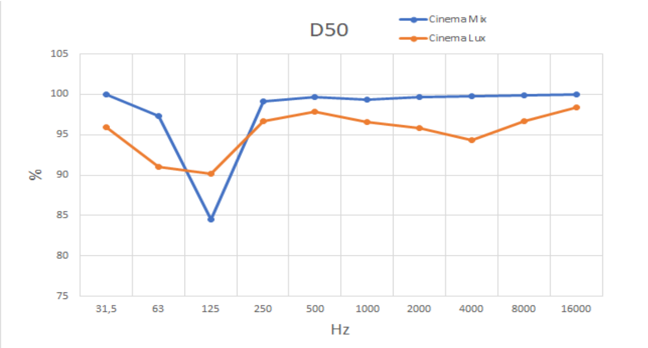

Fig. 7. Comparison between D50 measured in Cinema Mix and Cinema Lux

As regards the spatial parameters that have been taken in to account were Inter aural cross correlation (IACC) and lateral fraction (LF). The IACC values range from -1 to +1, a value of -1 means the signals are identical, but completely out of phase, a +1 means they are identical and 0 means they have no correlation at all. IACC exploits a cross-correlation operation between the two signals recorded at the Dummy ears and it reveals the spatial degree of the information. For both Dubbing Stage and Cinema Lux we have, in the low frequencies, high values of this parameter. This comes from the omnidirectional nature of the radiation at low frequencies and in fact, as we move towards higher frequencies, we have approximately decreasing correlation values.

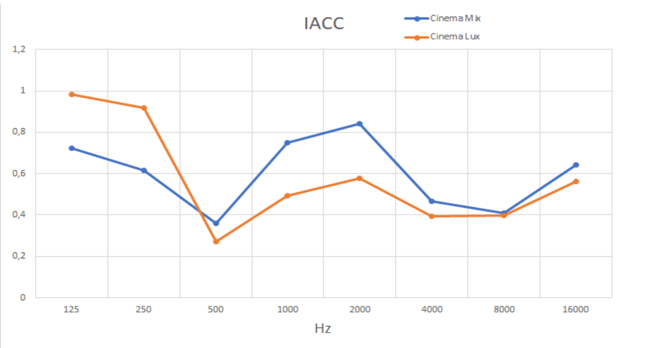

Fig. 8. Comparison between IACC (Interaural Cross Correlation) measured in Cinema Mix and Cinema Lux

The parameter lateral fraction calculates the ratio between the lateral sound energy and the total energy giving a measure of the apparent source width. In the Cinema Mix room, we can see two peaks in the graph of LF meaning that in those octaves (63 Hz and 500 Hz) the energy that comes from the lateral reflections is high. In Cinema Lux room instead, we have a good balance between lateral energy and the total energy.

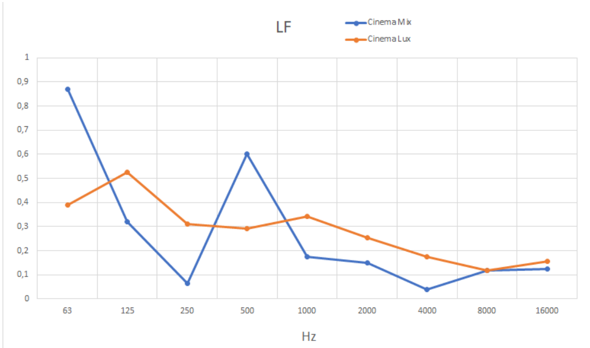

Fig. 9. Comparison between LF (Lateral Fraction) measured in Cinema Mix and Cinema Lux

B. *Acoustic analysis of 3D sound maps*

By taking advantage of the em32 Eigenmike® microphone's capabilities [6], sound immersive maps were obtained for each source-receiver position. Such maps have been elaborated to visualize the architectural elements interacting with the sound waves hitting the boundaries of the room. This outcome is an additional understanding of the trajectories run by the late reflections that have been analyzed in combination with the relative sound intensity. This different data analysis methodology is obtained by the employment of the last generation of equipment, to be included an omnidirectional sound source, a multichannel microphone (i.e. em32 Eigenmike®) and a panoramic camera. The 32 microphone signals have been processed by extracting 122 high directivity virtual microphones (with 8^{th} order cardioid setup) encoded by a Spatial PCM Sampling (SPS)[10]; the elaboration of the output sound signal has been overlapped with a 360° image represented in an equirectangular view. The result obtained by this process is a color map overlay showing the soundwaves arriving at the receiver from all the possible directions based on the spherical array of the microphone. The video has been realized by processing 1024 samples at 48 kHz sampling rate. A convolution of 32 input channels with 32 FIR filters has been found to be necessary to elaborate the matrix obtained by 32 virtual microphone outputs. The map overlay shows the sound pressure levels represented by contour levels. The color scale is comprised between red colors indicating a high level of energy and blue shades meaning a poor energy sound.

Here below are shown the color maps obtained for the dubbing room (Fig. 10, Fig. 11. Fig. 12.) and those obtained for the lux cinema (Fig. 13., Fig. 14., Fig. 15.)

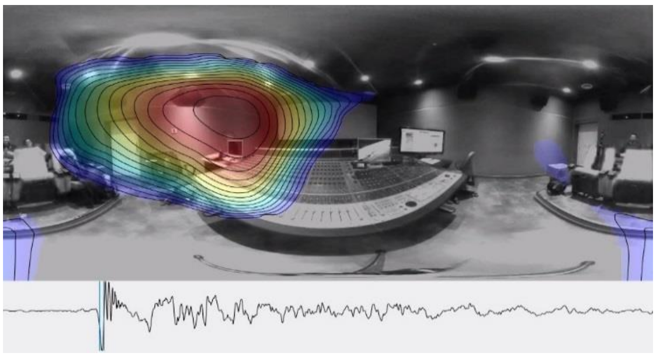

Fig. 10 Acoustical map of the cinema mix showing the direct sound arriving at the receiver placed in behind the mixing table.

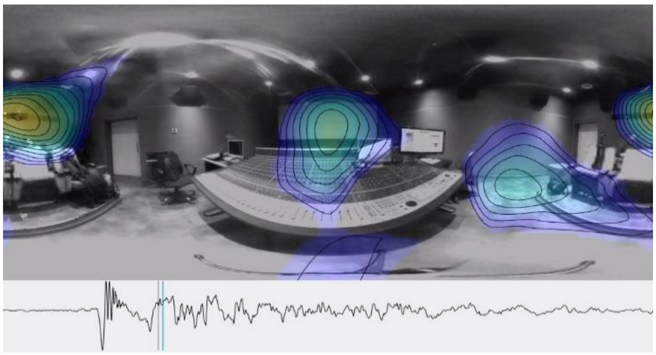

Fig. 11. Shows the firsts reflections that come from projector glass, floor and screen

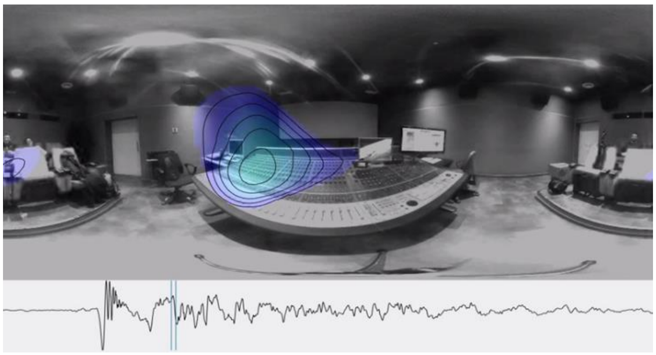

Fig. 12. Shows the seconds reflections

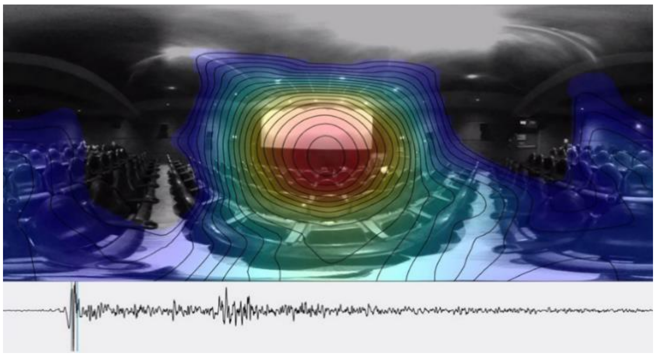

Fig. 13. Acoustical map of the cinema Luxshowing the direct sound arriving at the receiver placed in the fifth row of the venue.

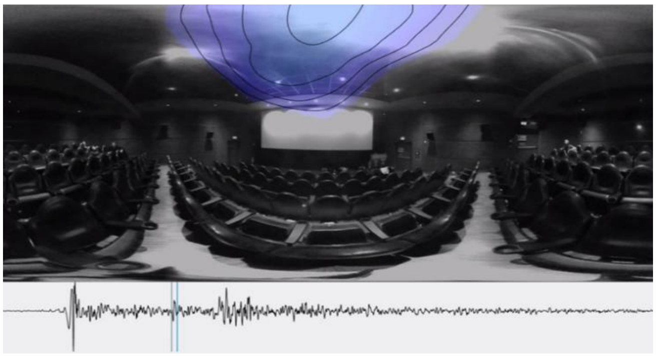

Fig. 14. Shows the firsts reflections that come from the ceiling.

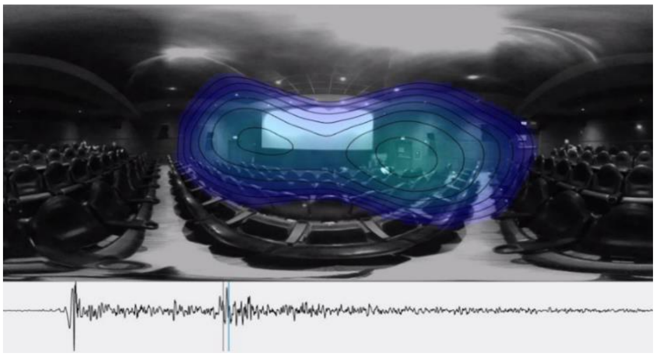

Fig. 15. Shows a group of late reflection bouncing from the right door.

In Cinema Mix room we have a higher risk of reflections due to the smaller volume (sound reaches reflective elements with more intensity) and to the presence of more reflective elements like glass of projector screen and mixing table. In Cinema Lux the acoustic is well managed except from some reflections from the entrance doors.

V. CONCLUSIONS

This paper deals with the different perception of a movie sound track in dubbing rooms and in performance cinema rooms. Observing the results, it can be concluded that the reproduction of the sound track in dubbing rooms and in performance cinema give two different percepitions. The two venues have different acoustic fingerprint and colors the sound in its own peculiar way depending on the geometry, the furniture and the disposition of absorbing materials. RTA estimation strategy for the alignment of sound framework through leveling should be abandoned because it is based on wrong scientific assumptions, so it is pointless in improving the overall quality of the sound. To enhance the acoustic of a room at a good degree, one could instead measure the spatial distribution of sound, to see which elements produce excessive reflections and to act on them, for example, by putting absorbing or diffusive materials. The implementation of array microphones (such as Eigenmike 32 capsules or similar), for the recordings in 3D space, and 3D cameras for shooting panoramic images, allow the operator to precisely identify the source of disturbances in the room under study. If the technician acoustically treats the space in a proper way, considering the positioning of loudspeakers and other issues, there should be no need to equalize the signal, and the sound system should behave in an optimal way without the need to frequently “re-calibrate” the room.

REFERENCES

[1] Murphy, David J, “Electro-acoustic measurements on cinema B chains in Australia”, http://www.aes.org/e-lib/browse.cfm?elib=17607.*)*

[2] ISO 3382-1: Acoustics - Measurement of Room Acoustic Parameters; Part 1: Performance Spaces; ISO: Geneva, Switzerland, 2009.

[3] A. Farina “Simultaneous measurement of impulse response and distortion with a swept-sine technique”. 108th AES Conv. 2000, France, Paris, February 19-22.

[4] A. Farina, “Aurora listens to the traces of pyramid power”. Noise & Vibration Worldwide, Vol. 26, 6, pp. 6-9, 1995.

[5] J. S. Bradley, “Using ISO 3382 measures, and their extensions, to evaluate acoustical conditions in concert halls." Acoustical Science and Technology, Vol 26, 2, pp. 170-178, 2005

[6] A. Farina, L. Tronchin, “3D sound characterisation in theatres employing microphone arrays”. Acta acustica united with Acustica, 99, 1, pp. 118-125, 2013.

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

[8] P. Fausti, R. Pompoli and N. Prodi, “Acoustics of opera houses: A cultural heritage”. J. of Acoustics Society of America, Vol. 105 (929), 1999.

[9] R. Thiele, “Richtungsverteilungs und zeitfolge del schallruckewurfe in raumen”. Acta Acust. 3, pp. 291-302, 1953.

[10] Tronchin, F. Merli, A. Bevilacqua, “MIMO Auralization: an example of loudspeaker”, Proc. 24th Intern. Conf. Circ. Syst. Comm. Comp. (CSCC), 9402332, pp. 159-164, 2020.