Audio Engineering Society

Convention Paper

Presented at the 130th Convention
2011 May 13–16 London, UK

Spatial analysis of room impulse responses captured with a 32-capsules microphone array

Angelo Farina[1], Alberto Amendola[1], Andrea Capra[1] and Christian Varani[1]

[1] Industrial Engineering Department, Università di Parma, Via G.P.Usberti 180/A, 43100 Parma, Italy
farina@unipr.it

ABSTRACT

The authors developed a new measurement system, which captures 32-channels impulse responses by means of a spherical microphone array and a matrix of FIR filters, capable of proving frequency-independent directivity patterns. This allows for spatial analysis with resolution much higher than what was possible with obsolete sum-and-delay beamforming.

The software developed for this application creates a false-colour video of the spatial distribution of energy, changing with running time along the impulse response duration. A virtual microphone probe allows to extract the sound coming from any specific direction. The method was successfully employed in three concert halls, providing guidance for correcting some acoustical problems (echo, focusing) and for placing sound reinforcement loudspeakers in optimal positions.

1. INTRODUCTION

The main goal of this research is for providing assistance to the acoustic consultant, when he is required to evaluate the acoustical behaviour of an existing concert hall or theatre, and to design either sound treatment or sound reinforcement for that room. The new approach here described provides easy-to-understand visual display of the spatial properties of the sound inside the room: the examples presented in this paper show how easy is to find the origin of unwanted reflections (echoes), to evaluate the spectral content of these reflections (for choosing the optimal sound absorbing or sound scattering material), and to check if sound reinforcement loudspeakers are correctly located and aimed. For reaching the desired results two sessions of measurements were performed: the first session employed a single microphone rotated and tilted in the space for capturing the sound at 360°, the second session employed an advanced probe made of 32 capsules, capable of a quick recording of the impulse responses. In this second case we developed a filtering technique based on "Kirkeby multichannel inversion" for creating 32 highly directive microphones (4th order cardioids). This virtual set of microphones is necessary for a proper spatial sampling of the sound around the probe.

2. THE FIRST ATTEMPT: A ROTATING SHOTGUN MICROPHONE

A first experiment of a 3D impulse response was made in 2008 in the opera houses of Modena and Como. The goal was the dynamic view of the changes of the sound pressure level in the measure's points by means of a movie, useful for detecting unwanted reflections: dynamic 3D view permits to discriminate the sound intensity due to direct sound or to reflections.

2.1. The measuring setup

The measuring setup was made of a Sennheiser ME66 shotgun microphone mounted on a rotating table, a dodecahedron as test signal source and a laptop connected with an Edirol FA101 soundcard, employed as digital recorder of the sine-sweep signal. The microphone was rotated for 18 steps in the azimuthal plane (20 degrees for each step) and for 8 steps in elevation, covering the upper hemisphere (22.5 degrees for each step). In order to obtain the impulse responses, the sine-sweeps recorded were processed by using Aurora Plugins [9] and Adobe Audition [10]. Matlab [11] software was prepared for designing dynamic polar plots of the impulse responses for the section of the upper hemisphere and for the azimuthal plane; in both cases the analysis was made in octave bands.

2.2. Some results

2.2.1. "Teatro Comunale" – Modena

As shown in Figure 1, the omnidirectional source was placed in the orchestral pit and the microphone system was placed in three positions (A, B and C). The temporal analysis of the impulse responses provides useful information about the sound inside the theatre. For example, if the source is in the orchestra pit, the direct sound that arrives to the performer on the stage is rich of low frequencies for the diffraction effect and very poor of high frequencies (Figure 2).

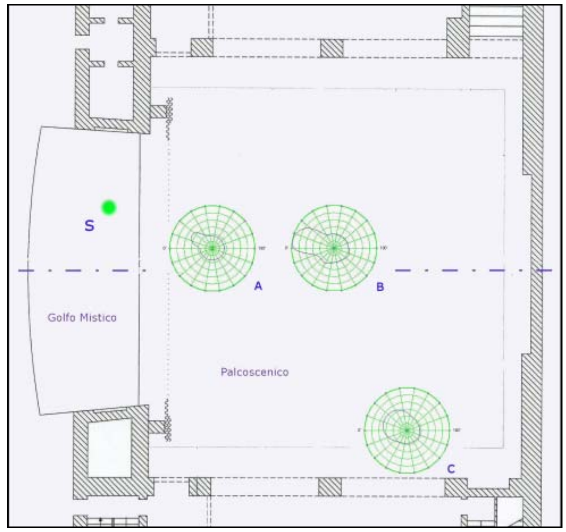

Figure 1 Map of Teatro Comunale (Modena)

The high frequencies, as shown in Figure 3, arrive to the performer only 16 ms after the direct sound, thanks to reflections caused by the wall of the theatre.

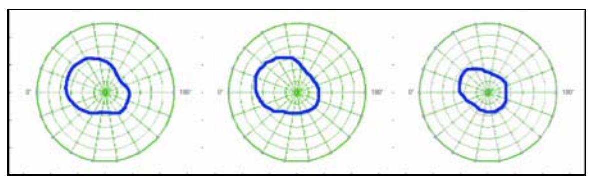

Figure 2 source on pit, Point A, 125 Hz, 1000 Hz, 4000 Hz at 24ms (direct sound)

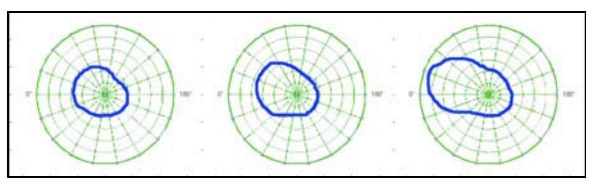

Figure 3 source on pit, Point A, 125 Hz, 1000 Hz, 4000 Hz at 40ms (first reflection)

2.2.2. "Teatro Sociale" – Como

In Como Theater we investigated the behavior of the sound coming from a source in the orchestral pit and arriving to a receiver in the stalls. As we expected, the direct sound is weaker than the reflections coming from the proscenium arch (Figure 4).

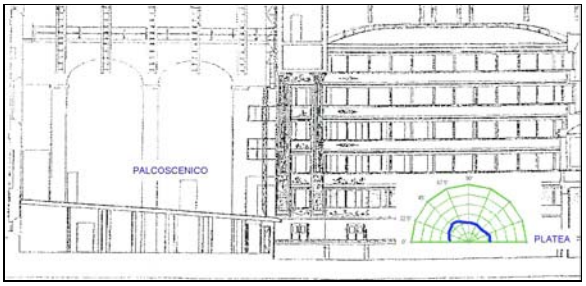

Figure 4 section of Como theatre

It was also possible to compare the temporal development of sound in the stage with or without loudspeakers from pit. If the source is placed on the stage (Figure 5) the polar plots show that the first reflection comes from the back of the stage after 40 ms from the direct sound; if the source is in the pit (Figure 6) the direct sound at 4000 Hz is weaker than the first reflection coming after 56 ms from the proscenium arch.

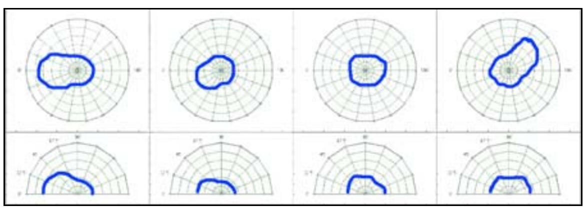

Figure 5 Direct sound in the 4000 Hz band at 55ms, 67ms, 91ms and 94ms

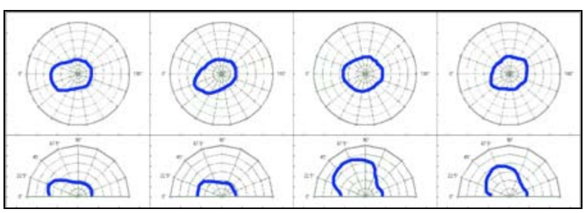

Figure 6 Source in orchestral pit, 4000Hz at 46 ms, 55 ms, 102 ms and 106 ms.

3. THE NEW APPROACH

The use of a single microphone, mounted on a rotating system, causes a significant waste of time. The measure is quicker if a single probe can capture all the directional impulse responses at once. This new approach is described in the paragraph below.

3.1. The microphone array

The experimentation described in this paper was realized using the Eigenmike^{TM} microphone array produced by MH acoustics [4]. As shown in Figure 7, the Eigenmike^{TM} is a sphere of aluminium (the radius is 42 mm) with 32 high quality capsules placed on its surface; microphones, pre-amplifiers and A/D converters are packed inside the sphere and all the signals are delivered to the audio interface through a digital CAT-6 cable, employing the A-net protocol.
The audio interface is an EMIB Firewire interface; being based on the TCAT DICE II chip, it works with any OS (Windows, OSX and Linux through FFADO). It provides to the user two analogue headphones outputs, one 8-channels ADAT digital output and the word clock ports for syncing with external hardware.
The preamplifier's gain control is operated through MIDI control; we developed a GUI (in Python) for making easy to control the gain in real-time with no latency and no glitches.

Figure 7 Eigenmike^{TM} probe

Our idea is to derive a set of 32 directive virtual microphones in the directions of the capsules employing a set of digital filters. In our case the *M=32* signals coming from the capsules need to be converted in *V=32* signals yielding the desired virtual directive microphones: so we need a bank of $M \times V$ filters. As always, we prefer FIR filters.

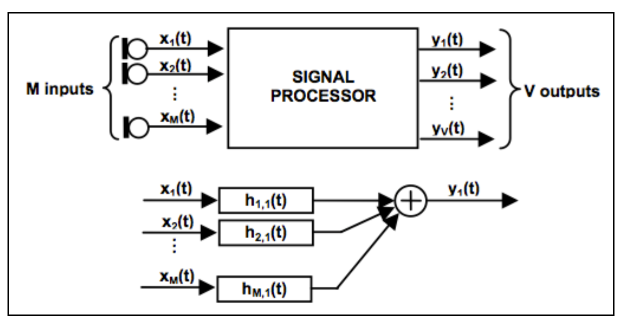

Figure 8 Scheme of the signal processing

Assuming $\boldsymbol{x_m}$ as the input signals of M microphones, $\boldsymbol{y_v}$ as the output signals of V virtual microphones and $\boldsymbol{h_{m,v}}$ the matrix of filters, the processed signals can be expressed as:

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

Where * denotes convolution, and hence each virtual microphone signal is obtained summing the results of the convolutions of the M inputs with a set of M proper FIR filters.

Most of the techniques employed for deriving filter banks for generating virtual microphones usually operate following one of several complex mathematical theories, based on the solution of the wave equation, often under certain simplifications, assuming that the microphones are ideal and identical. In some implementations the signal of each microphone is processed through a digital filter for compensating its deviation, with a heavier computational load.

In this novel approach no theory is assumed: the set of filters h are derived directly from a set of measurements, made inside an anechoic room. A matrix of measured impulse response coefficients $\boldsymbol{c}$ is formed and the matrix has to be numerically inverted (usually employing some approximate techniques, such as Least Squares plus regularization) [5,6].; in this way the outputs of the microphone array are maximally close to the ideal responses prescribed. This method also inherently corrects for transducer deviations and acoustical artefacts (shielding, diffraction, reflection, etc.).

In order to derive the matrix of filters, a Matlab script was produced. This script employs 2048 samples of each impulse response and it needs as inputs the number of virtual microphones to synthesize, their directivity, their azimuth and elevation. From these inputs, according with the theory and the procedure previously described [8], it is possible to invert the matrix of impulse responses obtaining the matrix of filters to be applied to the capsule signals. The convolution of the FIRs matrix with the 32 signals coming from the capsules of the array should give as outputs the signals of virtual microphones with the desired characteristics. Some experimental results are shown in one of our previous paper [8], showing some of the different directivity patterns obtained. For this research only the 4th order cardioid was used and the polar pattern of its directivity is shown in Figure 9.

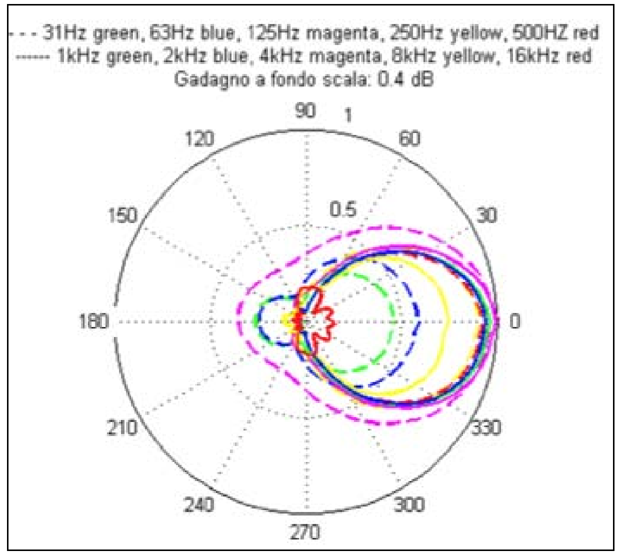

Figure 9 plot of the 4th order cardioid polar pattern obtained with the multichannel inversion technique

3.2. The new session of measurements

The second session of measurements was performed inside the Opera House "La Scala" (Milan – Italy) and the auditorium "Sala dei Concerti" of Casa della Musica (Parma – Italy).

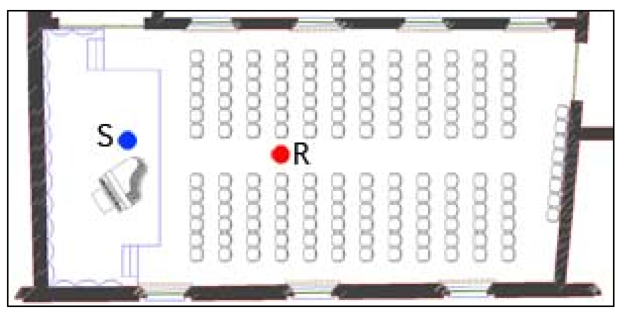

Figure 10 map of "Sala dei Concerti" (Parma)

The Eigenmike^{TM} probe, connected to a laptop, was used for recording the sine-sweeps generated by a dodecahedron; all the convolutions for obtaining the impulse responses were made using Aurora Convolver [10]. In "Sala dei Concerti" (Figure 10) we used as test signal source a Lookline dodecahedron, placed 1 m off the center of the stage; the receiver was placed as shown in Figure 10.

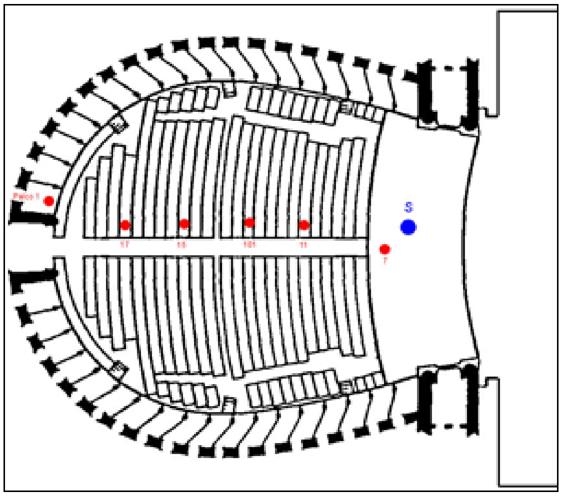

Figure 11 Map of “La Scala” (Milan)

In “La Scala Theatre” the dodecahedron source was placed 1m off the center of the stage and a set of points was chosen for the receiver (Figure 11). For both the theatres, in correspondence of the receivers, a panoramic 360° picture was taken with the aim of creating on it the dynamic map of the sound levels.

3.3. The post-processing software

All the processing useful for deriving the behavior of the sound inside the theatres is implemented in Matlab [11]. The GUI is made of two main parts: a polar plot of the sound levels in the horizontal plane and a mesh map of the levels plotted on a 360°x180° picture of the theater. They are both dynamic plots, a movie of the sound at the receiver. The processing is done in three steps:

- Import of WAV files (multiple files or one single multichannel file) and selection of the right portion
- FFT processing and choice of the number of frames in which perform the calculation
- Dynamic plot of the results

In the second step the user can choose the frequency band of investigation. For a dynamic representation of the sound at the receiver, multiple, overlapped FFTs are performed. In particular for making a slow-motion analysis of an impulse response, we have to increase the number of time steps by increasing the overlapping of the FFT blocks. In this way a 3D matrix (microphone, frequency band, time step) is created. The user can also choose the type of values being charted: sound pressure, SPL in dB, squared pressure (energy) are the possible choices for visualizing all ranges of values. The plot of the values is performed considering the level derived from every single virtual microphone and representing it on a polar or a map in the corrsponding coordinates. A cubic interpolation is necessary for giving continuity to the colour maps, whilst a simple linear interpolation is employed for polar plots.

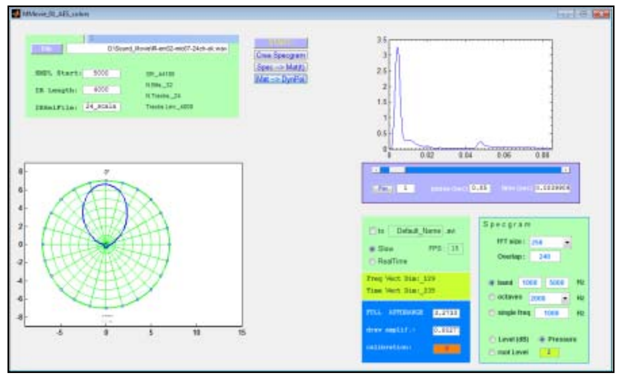

Figure 12 GUI of the post-processing Matlab software

At the end of the processing the user can export a movie (an “.AVI” file), which includes a two-channels impulse response referred to that specific measurement’s point.

3.4. Analysis of the results

We did not made the calibration of the probe: for this reason every point of measure has a its own level normalization and consequently an its own color scale. Nevertheless we can notice as, growing the distance between source and receiver, the lobes of the reflections becomes more relevant in comparison with the direct sound and they permit a mental reconstruction of the sequence of the walls beaten by the sound.
We choose to synthesize 24 virtual microphones regularly spaced on the azimuthal plane and 32 virtual microphones in the direction if the capsules: the first set of microphones is used for a polar plot of the sound at elevation 0°, the second one is useful for creating a color map of the sound on 360° photos.

3.4.1. “Teatro La Scala” – Milan

We can compare the results obtained with the two set of virtual microphones only in the case of measures made with the receiver in the position number 7 (Figure 13) because of the absence of 360°-photos in the other points.

Figure 13 La Scala: receiver in position n.7

As we expected, the direct sound produces in the map a blur that grows in parallel with the increasing of the frequency band of analysis (Figure 14 and Figure 15). Is interesting to notice that in the 500 Hz band, two blurs, one on the stage and one on the back wall, appear before the arrival of the direct sound: this is due to the radiation of the wooden stage under the excitation of the subwoofer while this is playing very low frequencies. The source is close to the receiver so the direct sound covers the first reflection on the floor masking it on the map. The first visible reflection appears after 4 ms coming from the column behind the source, followed by a second reflection on the floor and by a reflection coming from the opposite direction of the first reflection. An interesting behavior is revealed at 57 ms where the sound bounces between the columns of the proscenium arch (Figure 17). Only after these reflections arrives the one coming from the ceiling.
From this dynamic view we can remark some observations:

- The dynamic view of the impulse responses shows as the shape of the theatre refocus the sound in the point of provenience.
- As we expected in the 500 Hz band lobes appears on the map: they are the effect of the diffraction not present at higher frequencies.
- The reflections coming from the walls of the room are displayed on the map with cold colors and this chromatic difference from the direct sound reveals a big amount of absorption in that area of the theatre.

Figure 14 Direct sound in the 500 Hz band

Figure 15 Direct sound in the 4000 Hz band

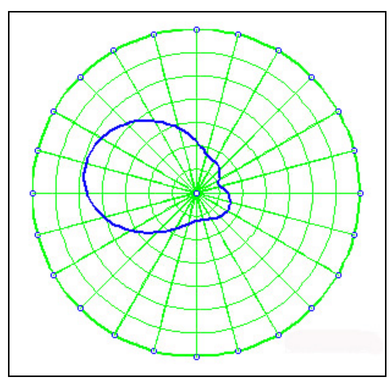

Figure 16 Polar plot of the sound in the 4000 Hz band

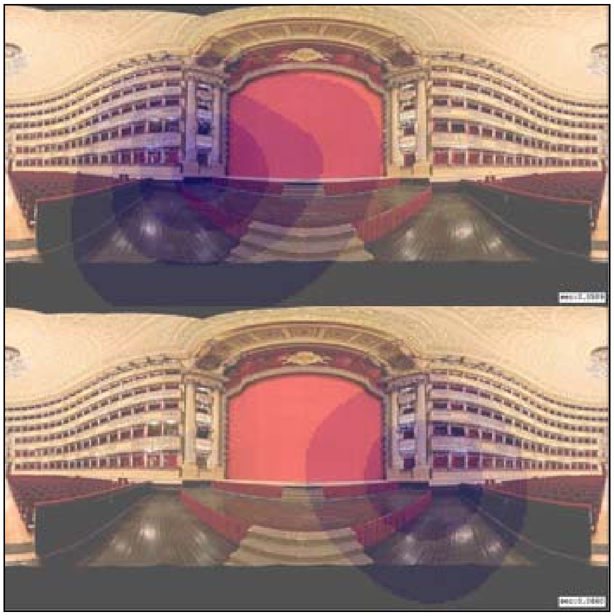

Figure 17 Bounce of the sound between the columns of the proscenium arch

3.4.2. "Sala dei Concerti" – Parma

Figure 18 Sala dei Concerti (Parma)

In this case the first reflection on the floor is clearly visible, followed by the diffuse sound coming from the wooden diffusors on the stage. After 21 ms is visible a strong lateral reflection at 90° due to a plane even surface (Figure 19); on the opposite side the reflection is not present because of the presence of absorbent curtains.

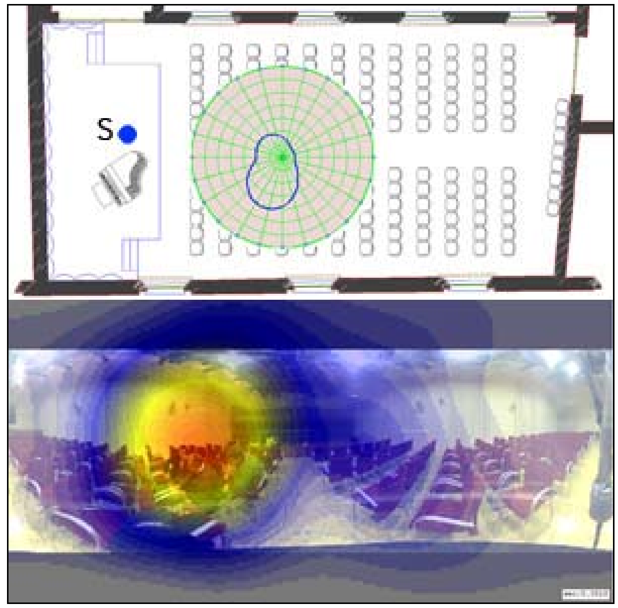

Figure 19 Polar and map of the lateral reflection in the 4000 Hz band

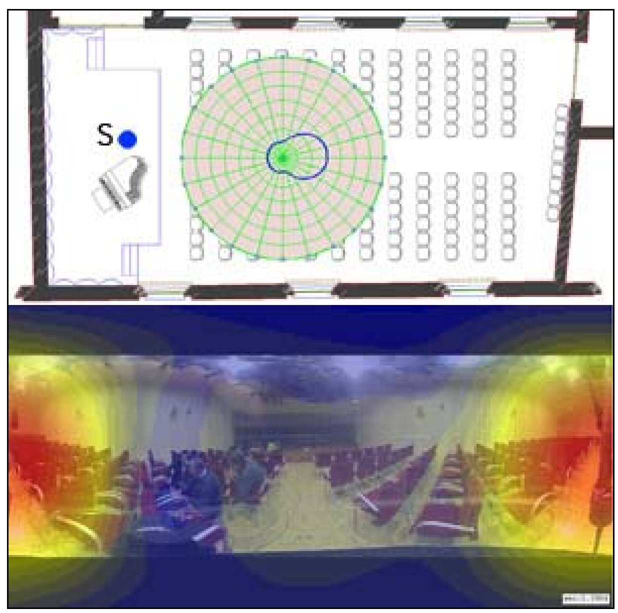

Figure 20 Polar and map of the reflection coming from the back wall in the 4000 Hz band

After these reflections the sound appears quite diffuse in the room, a part from an expected bounce between the lateral parallel walls. At 97 ms from the direct sound a strong reflection arrives to the receiver from the back of the room (Figure 20): this effect is audible also from the stage and causes problems to the performers with high repeated notes. Observing the 500 Hz band, also in this case we can notice that low frequencies anticipate the direct sound for the resonance due to the wooden structure of the stage. A comparison of the direct sound at 500 Hz and 4000 Hz is shown in Figure 21 and, as we expected, the wave front at low frequencies is considerably larger than the one at high frequencies.

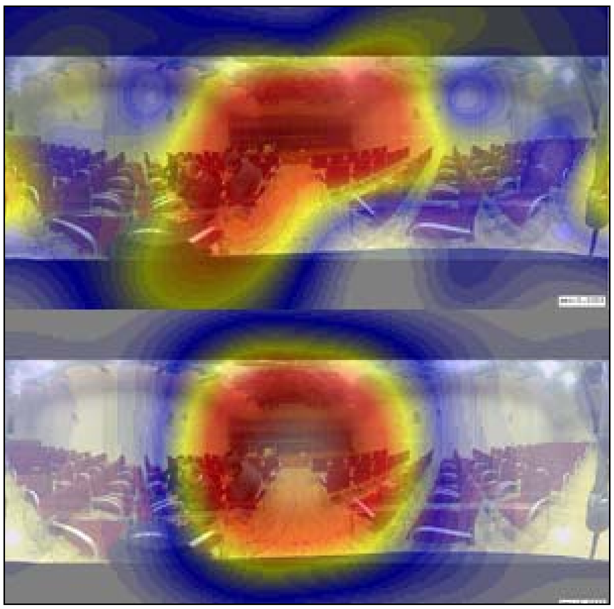

Figure 21 Maps of the direct sound in the 500 Hz and 4000 Hz bands

4. CONCLUSIONS AND FUTURE WORK

In this paper a new approach to impulse response measurements is presented. It permits the dynamic view of the impulse responses plotted on a 360°x180° picture of the space under inspection by using 32 high directive virtual microphones. This way of analyzing the acoustical behavior of theatres, spaces, rooms, permits an easy visualization of undesired reflections, offering to the acoustical designer a simple way for locating from where they are coming and for correcting them. This is a first step in the development of the method but the way is long, a lot of work has to be done for optimizing the measure system and the measure technique:

- All the results are derived from a not-calibrated probe. The comparison among different measurements is currently possible only with regard to the temporal distribution of the sound, not to the absolute levels. For these reason a next step will be the calibration of the probe.

- All the processing is done through a Matlab script but the goal is the creation of a plugin (VST or Audacity format), significantly faster and more useful thanks of its integration with an audio host.

- The probe has not its own panoramic camera for taking 360° pictures: in the future an IP panoramic surveillance camera will be attached to the probe, avoiding to the user the waste of time of taking dozens of photos, which are later merged together in a single panoramic image by means of a photo-stitching software.

5. ACKNOWLEDGEMENTS

The authors are thankful to "RAI - Centro Ricerche e Innovazione Tecnologica", Turin, Italy, for allowing to use one of their Eigenmike™ probes.

6. REFERENCES

[1] A.J. Berkhout, D. de Vries, and P. Vogel, *Acoustic control by wave field synthesis*, Journal of the Acoustic Society of America, 93(5):2764– 2778, May 1993

[2] S. Moreau, J. Daniel, S. Bertet, *3D sound field recording with High Order Ambisonics - objective measurements and validation of a 4th order spherical microphone,* 120th AES Convention, Paris, France - May 20-23, 2006.

[3] F. M. Fazi, P. A. Nelson, *The ill-conditioning problem in Sound Field Reconstruction*, 123rd AES Convention, New York, NY, USA - October 5/8, 2007.

[4] http://www.mhacoustics.com

[5] O. Kirkeby, P. A. Nelson, *Digital Filter Design for Inversion Problems in Sound Reproduction*, AES, vol. 47, no. 7/8 (1999 July/August).

[6] O. Kirkeby, P. Rubak, A. Farina, *Analysis of ill-conditioning of multi-channel deconvolution problems*, 106th AES Convention, Munich, Germany - May 8-11, 1999.

[7] A. Farina, *Simultaneous measurement of impulse response and distortion with a swept-sine technique*, 110th AES Convention, Paris 18-22 February 2000.

[8] A. Farina, A. Capra, L. Chiesi, L. Scopece, *A spherical microphone array for synthesizing virtual directive microphones in live broadcasting and in post production*, 40th AES Conference, Tokio, Japan, 8-10 October 2010.

[9] http://www.aurora-plugins.com

[10] http://www.adobe.com/it/products/audition/

[11] http://www.mathworks.com