VALENCIA 24-25 November 2011
ISVA International Seminar on Virtual Acoustics

# RECORDING, SIMULATION AND REPRODUCTION OF SPATIAL SOUNDFIELDS BY SPATIAL PCM SAMPLING (SPS)

Reference PACS: 43.60.Fg. 43.38.Hz
.

Farina, Angelo; Binelli, Marco; Capra, Andrea; Campanini, Simone; Amendola, Alberto.

University of Parma, Industrial Eng. Dept.
Parco Area delle Scienze 181/A
43124 PARMA
ITALY
e-mail: farina@unipr.it

**ABSTRACT.**

A new method for recording the spatial properties of a soundfield, or for generating a synthetic three-dimensional soundfield, is described. The spatial distribution of sound waves passing at a point in space is sampled by means of a number of virtual super-directive microphones, covering almost uniformly the surface of a sphere. This corresponds to a discretization of the spatial information, which is exactly the spatial equivalent of the PCM sampling of a waveform.
Continuing the comparison, traditional High Order Ambisonics can instead be seen as the correspondent of the Fourier transform of a periodic waveform, representing the spatial distribution of sound as the superposition of a number of spatial frequencies (the spherical harmonics).
As with waveforms, the PCM model is generally more versatile and less constrained than the corresponding frequency-domain representation. Although some operations are easier in frequency domain, others are easier in the original PCM domain.
When dealing with spatial information, modifications of the sound field, such as rotation, stretching (zooming), and emulating sound effects such as reverb, shielding, etc., are easily performed on the spatial PCM signals.
And when the spatial information must be reconstructed, by means of a suitable 3D loudspeaker array, it is possible to derive the speaker feeds employing a proper set of filters (which are, in practice, a “spatial FIR filter”).
The paper shows how it is possible to record a 32-channels spatial PCM waveform employing a spherical microphone array, how a corresponding synthetic signal can be generated on a computer by “spatial encoding” of the signals corresponding to virtual sources located in space, and how a suitable matrix of “decoding filters” can be generated for recreating faithfully the original sound field inside a loudspeaker array.
In the end, spatial PCM revealed to be a simpler and more general way of capturing, synthesizing, modifying, transmitting and reconstructing a 3D spatial sound field.

## 1. INTRODUCTION

The first comprehensive method for recording, synthesis, modification and reproduction of a truly three-dimensional sound field was Ambisonics [1], which describes the complete 3D information of the soundfield in a point by means of 4 signal, the sound pressure (W) and the three Cartesian components of the article velocity (X, Y and Z). The 4-channels signal according to this definition is called “B-format”.

However, extracting the complete and detailed spatial information form this B-format stream revealed to be quit hard. The traditional Ambisonics method provides very poor spatial resolution, resulting in “virtual microphone” signals that can replicate the polar pattern of 1$^{st}$-order microphones (from omni to figure-of-eight, passing through cardioid and hypercardioid). The spatial reconstruction, obtainable inside loudspeaker rigs, provides very little spatial separation of sound sources, and a lot of sound comes from everywhere.

Recently, more advanced processing methods for B-format signals were presented, such as DIRAC by Ville Pulki [2], or Harpex-B by Svein Berge [3].

These methods are indeed tricks based on “intelligent steering”, they cannot really retrieve the complete spatial information form such a small number of channels.

The only robust approach relies on the usage of more channels for sampling the spatial information, and so High Order Ambisonics (HOA) was developed, in which the number of “spherical harmonics” signals employed for encoding the spatial information is much larger.

HOA relies on the similitude among the polar patterns of B-format microphones (spherical and figure of eight) with the shape of the molecular orbitals for electrons around the nucleus of an atom. Here, after the orbitals “s” and “p”, corresponding to the spherical harmonics of order 0 and 1, we have a number of higher-order spherical harmonics: 5 2$^{nd}$-order harmonics, 7 3$^{rd}$-order harmonics, 9 4$^{th}$-order harmonics, and so on.

So, for example, the 3 $^{th}$-order HOA representation of a sound field requires 16 channels, corresponding to all the spherical harmonics signals up to 3$^{rd}$ order.

HOA revealed to work very well for creating synthetic spatial soundfields, for performing manipulations over them (rotation, zooming, etc.) and for replaying them inside loudspeaker rigs [4]. The possibility to employ HOA for recording live sound was very limited, and only in 2010 the Eigenmike™ microphone system was launched on the market. This system does not provide explicitly the HOA signals as output, but the raw signals coming from the 32 capsules can be processed with a software tool, which internally employs 3$^{rd}$ order HOA, for extracting a number of virtual microphones which can provide directivity polar patterns much sharper than those obtained form a 1$^{st}$-order B-format signal.

Jerome Daniel [5] developed the first complete HOA solution, capable of transforming the raw signals coming form a spherical microphone array into the HOA components up to 4$^{th}$ order, to manipulate them by means of suitable VST plugins, and finally to render the results over a loudspeaker rig.

However, recording high-order signals revealed to be problematic, as the corresponding polar patterns have very complex shape, and it is very difficult that these shapes are accurately obtained. Furthermore, at low frequency the high-order signals exhibit a lot of noise, and at high frequency artefacts due to “spatial aliasing” colour significantly the sound.

## 2. SPATIAL PCM SAMPLING

We recently developed an alternative technology, which does not rely anymore on spherical harmonics, but instead employs simply a number of highly-directive cardioids, covering almost uniformly the surface of a sphere, for capturing the complete spatial information [6].

In practice, this is equivalent in space, to the representation of a waveform (in time) as a sequence of impulses (PCM, pulse code modulation), instead as the superposition of a number of sinusoids (Fourier transform).

We call this new approach Spatial PCM Sampling (SPS); for implementing it in practice, we start with the signal coming from a spherical microphone array (actually, an Eigenmike™). The signals coming form the 32 capsules are filtered by means of a matrix of 32x32 FIR filters, which synthesize 32 virtual high order cardioids, pointing in the same direction as the 32 capsules. So the Eigenmike is employed as a superdirective beamformer.

The 32 superdirective virtual microphones perform a spatial PCM sampling, as each of them can be thought as having a directivity pattern approximating a “spatial Dirac’s Delta function”.

Fig. 1 compares the standard PCM representation of a waveform in time with the “spatial PCM” representation of a directivity balloon in space.

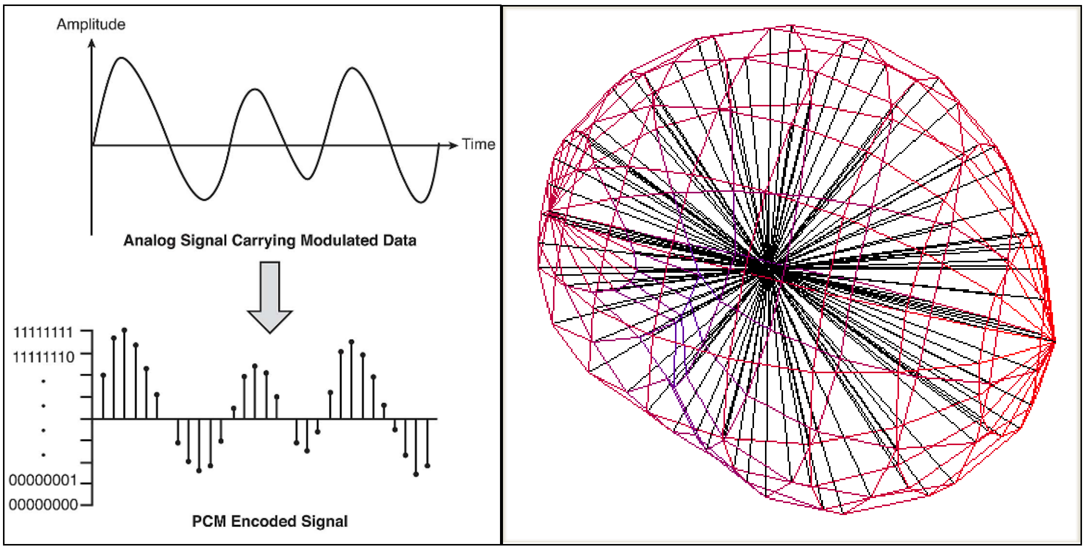


Figure 1: PCM sampling of a waveform in time (left) and of a balloon in space (right)

Fig.2, instead, shows the reconstruction of a waveform (in time domain) or of a spatial directivity balloon by means of the Fourier principle, that is, the superposition of a number of sinusoids (in time) or of spherical harmonics (in space), each with proper gain.

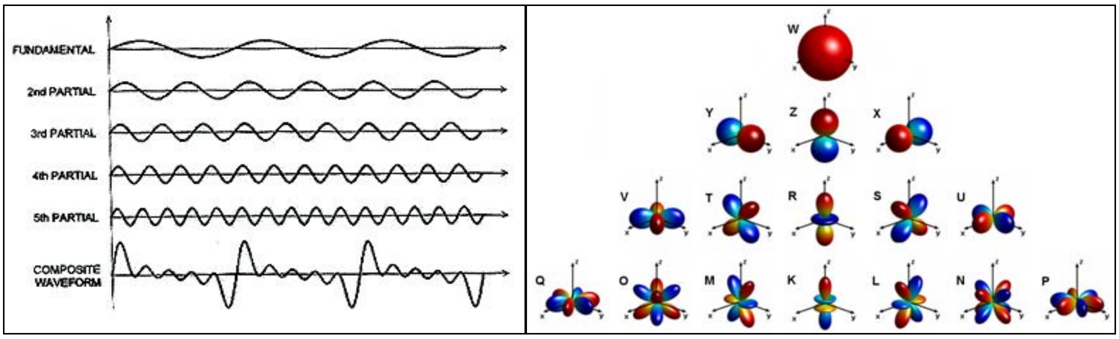


Figure 2: Fourier Analysis (left) and Spherical Harmonics (right)

Once the SPS signals have been obtained (either by recording or by synthesis), it is possible to manipulate them quite easily, performing standard operations such as rotation, stretching, zooming, etc.
Instead of performing these operations in the spherical harmonics domain, we can now perform them directly in the spatial PCM domain: hence, the most general transformation assumes the form of a spatial FIR filter...
Finally, it is possible to employ the SPS signals for deriving the speaker feeds, to be employed in a playback system. As it will be explained at the end of this paper, the math for designing these filters is substantially identical to the math employed for encoding the SPS signals.
We compared side-by-side the usage of the HOA method and of SPS, starting with the same signals captured by a spherical microphone array, and playing back the recording inside our listening room equipped with 16 loudspeakers. SPS outperformed HOA, for his better spatial resolution, wider frequency range and lower noise.

### 3. 32-CAPSULES SPHERICAL MICROPHONE ARRAY

For realtime recording/broadcasting applications, a new microphone system was recently developed by the RAI Research Center in Turin and by AIDA, a spinoff of the University of Parma. It is based on a 32-capsules spherical microphone array (Eigenmike™), and a realtime filtering software that is capable to synthesize up to 7 virtual microphones, which can be moved in real-time, and with variable directivity (zooming) capability. The “virtual” microphones, are controlled by mouse/joystick gestures in order to follow actors on the stage in realtime, and to zoom in or out, by changing the sharpness of the directivity pattern. The pattern is chosen among a family of cardioid microphones of various orders, according to this formula:

$$Q_n(\vartheta,\varphi) = \left[0.5 + 0.5 \cdot \cos(\vartheta) \cdot \cos(\varphi)\right]^n \quad (1)$$

where n is the directivity order of the microphone.

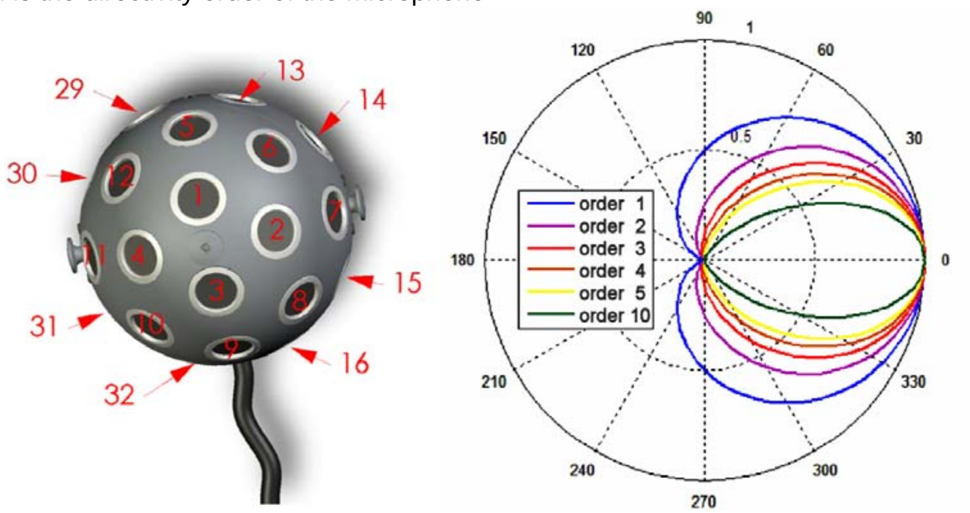



Figure 3 Capsule positions and directivity patterns of the 3D virtual microphone

When this system is employed for deriving the SPS signals, indeed, instead of synthesizing just 7 virtual microphones with variable aiming and directivity, a set of 32 static filters is employed, creating 32 signals corresponding to 32 $4^{th}$ order cardioids, pointing in the same direction of the capsules, as shown in fig. 3.

If the aiming directions of the 32 virtual microphones are overplotted over a panoramic 360°x180° image taken from the microphone position, one can “see” where the 32 microphones are pointing, a shown in figure 4:

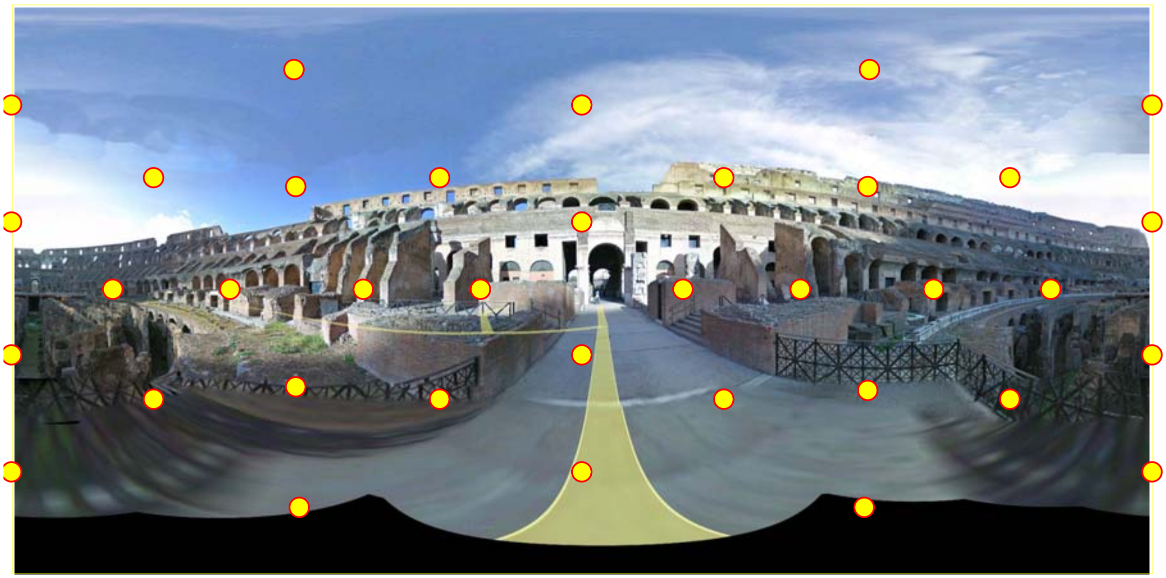


Figure 4 The 32 virtual microphones pointing all around inside the Colosseum in Rome

As shown in Figures 5 and 6, the Eigenmike™ 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-5 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 5 – The Eigenmike™ probe

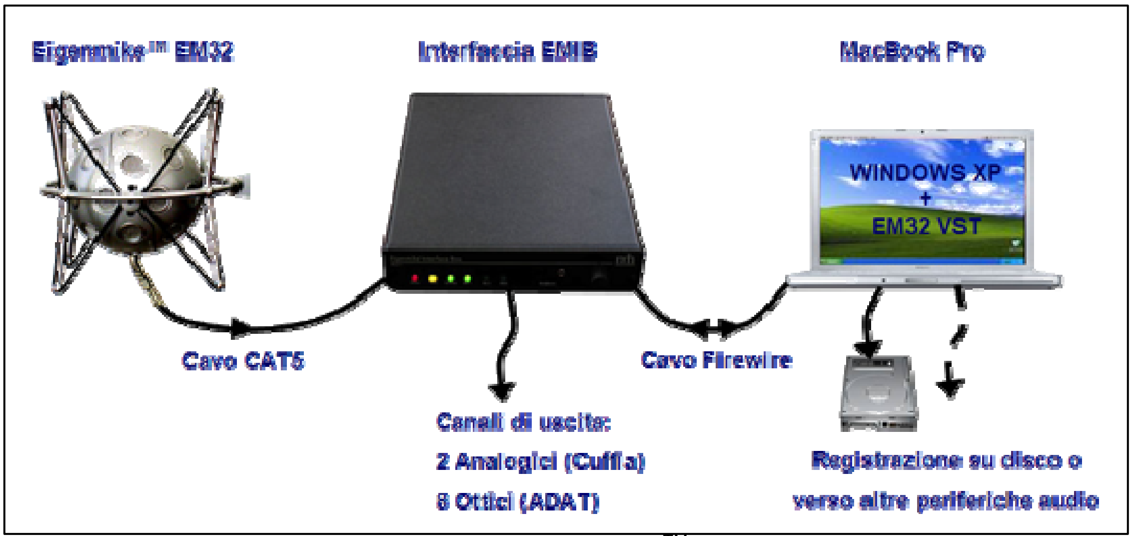



Figure 6 – The complete Eigenmike™ recording system

## 4. SYNTHESIS OF THE ENCODING FILTERS

Given an array of transducers, a set of digital filters can be employed for creating the output signals. In our case the *M* signals coming from the capsules need to be converted in *V* signals yielding the desired virtual directive microphones: so we need a bank of *M*×*V* filters. As always, we prefer FIR filters.

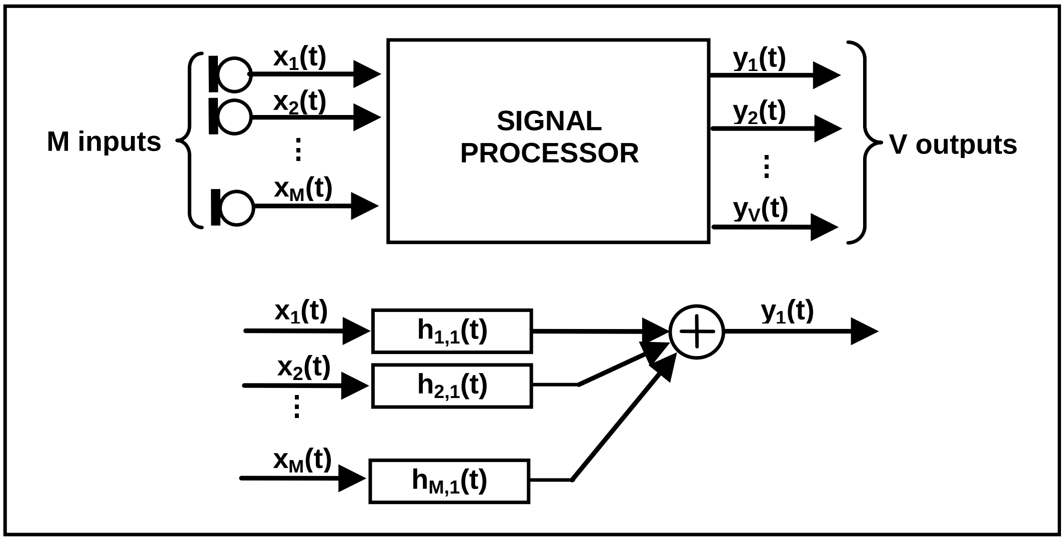


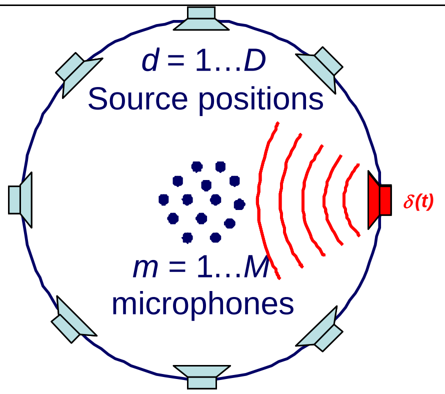

Figure 7: Scheme of the signal processing

Assuming $\{\mathbf{x}_m\}$ as the input signals of *M* microphones, $\{\mathbf{y}_v\}$ as the output signals of *V* virtual microphones and $[\mathbf{h}_{m,v}]$ the matrix of filters, the processed signals can be expressed as:

$$y_v(t) = \sum_{m=1}^{M} x_m(t) * h_{m,v}(t) \qquad (2)$$

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.
In our approach, the outputs of the processing system are directly the result of the digital filtering of the input signals, with a different set of filters for every virtual microphone.
In principle this allows for synthesizing virtual microphones having an arbitrary directivity pattern. In practice we decided, for now, to synthesize frequency-independent high-order cardioid virtual microphones, having the polar pattern described by eq. (1).

We are not assuming any theory for computing the filters *h*: they are derived directly from a set of measurements, made inside an anechoic room. A matrix of measured impulse response coefficients ***c*** is formed and the matrix has to be numerically inverted (usually employing some approximate techniques, such as Least Squares plus regularization); in this way the outputs of the virtual microphone is maximally close to the ideal response prescribed. This method also inherently corrects for transducer deviations and acoustical artefacts (shielding, diffraction, reflection, etc.).
The characterization of the array is based on a matrix of measured anechoic impulse responses, obtained with the sound source placed at a large number *D* of positions all around the probe, as shown in Figure 8.




$$c = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,d} & \dots & c_{1,D} \\ c_{2,1} & c_{2,2} & \dots & c_{2,d} & \dots & c_{2,D} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{m,1} & c_{m,2} & \dots & c_{m,d} & \dots & c_{m,D} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{M,1} & c_{M,2} & \dots & c_{M,d} & \dots & c_{M,D} \end{bmatrix}$$

Figure 8: impulse response measurements from D source positions to the M microphones

The processing filters ***h*** should transform the measured impulse responses ***c*** into the prescribed theoretical impulse responses ***p***:

$$\sum_{m=1}^{M} c_{m,d} * h_m \Rightarrow p_d \qquad d = 1..D \qquad (3)$$

Please notice that in practice the target impulse responses $\boldsymbol{p_d}$ are simply obtained applying a direction-dependent gain *Q*, given by eq. 1, to a delayed unit-amplitude Dirac's delta function $\delta$:

$$p_d = Q_d \cdot \delta. \qquad (4)$$

Computation is easier in frequency domain (that is, computing the complex spectra, by applying the FFT algorithm to the *N*-points-long impulse responses ***c***, ***h*** and ***p***). Let's call ***C***, ***H*** and ***P*** the resulting complex spectra. This way, the convolution reduces to simple multiplication between the corresponding spectral lines, performed at every frequency index *k*:

$$\sum_{m=1}^{M} C_{m,d,k} \cdot H_{m,k} \Rightarrow P_d \qquad \begin{cases} d = 1..D \\ k = 0..N/2 \end{cases} \qquad (5)$$

Now we pack the values of ***C***, ***H*** and ***P*** in proper matrixes, taking into account all the *M* input microphones, all the measured directions *D* and all the *V* outputs to create:

$$[H_k]_{MxV} = \frac{[P]_{DxV}}{[C_k]_{DxM}} \qquad (6)$$

This over-determined system doesn't admit an exact solution, but it is possible to find an approximated solution with the Least Squares method, employing a regularization technique for avoiding instabilities and excessive signal boost [7]. The block diagram of the least-squares method is shown in 9:

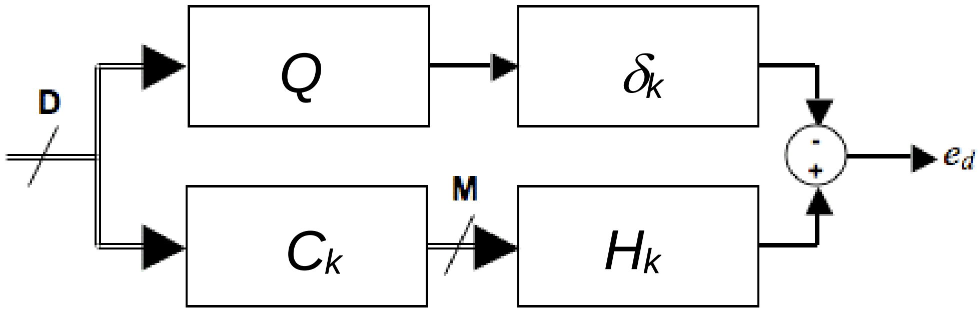



Figure 9: scheme of the Least Squared method with a delay in the upper branch

In this scheme we observe the delay block $\delta_k$ required for producing causal filters, and the resulting total modelling error $e_d$, which is being minimized by the least-squares approach
In general, the frequency-domain representation of a Dirac's delta delayed by $n_0$ samples is given by:

$$\delta_k = e^{-j2\pi k \frac{n_0}{N}} \tag{7}$$

Albeit various theories have been proposed for defining the optimal value of the causalisation delay $n_0$, we did take the easy approach, setting $n_0=N/2$. Choosing $N/2$ samples is a safe choice, which creates inverse filters with their "main peak" close to their centre, and going smoothly to zero at both ends.
Furthermore, a regularization parameter is required in the denominator of the matrix computation formula, to avoid excessive emphasis at frequencies where the signal is very low.
So the solution formula, which was first proposed in Kirkeby et al. [7], becomes:

$$[H_k]_{MxV} = \frac{[C_k]^*_{MxD} \cdot [Q]_{DxV} \cdot e^{-j\pi k}}{[C_k]^*_{MxD} \cdot [C_k]_{DxM} + \beta_k \cdot [I]_{MxM}} \tag{8}$$

As shown in the image below, the regularization parameter $\beta$ should depend on frequency [8]. A common choice for the spectral shape of the regularization parameter is to specify it as a small, constant value inside the frequency range where the probe is designed to work optimally, and as much larger values at very low and very high frequencies, where conditioning problems are prone to cause numerical instability of the solution.

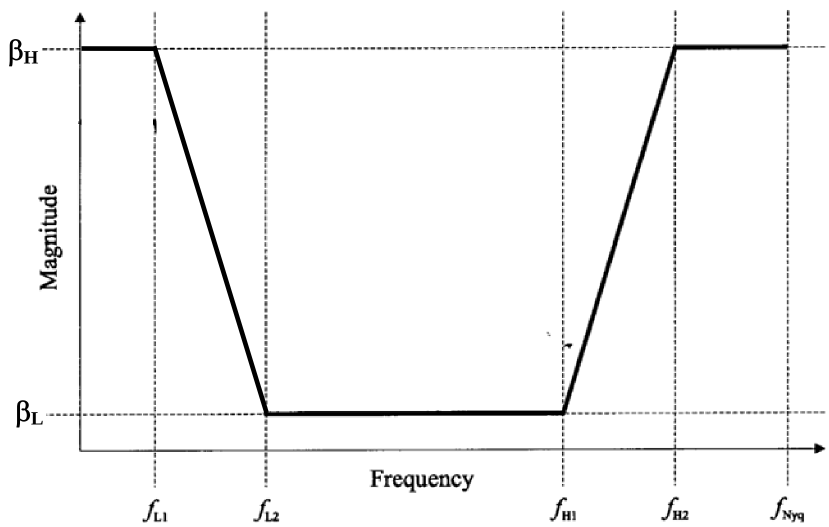



Figure 10: regularization parameter in dependence of the frequency

In our case, the number of virtual microphones being synthesized is 32, and their directivity and aiming are those defined in chapter 3 (Fig. 3). Typically, each filter is 2048 samples long (at 48 kHz sampling rate). Each virtual microphone, thus, requires to sum the results of the convolution of 32 input channels with “his” 32 FIR filters. And for getting all the required 32 virtual microphone outputs, we need to convolve-and-sum over a matrix of 32x32 FIR filters, each of 2048 samples.
For performing these massive multichannel filtering operations, a special VST plugin was developed, called X-volver, and running either on Mac or Win32 platforms; this plugin is freely available in [9]. Fig. 11 shows the X-volver plugin being used inside Audio Mulch, a multichannel VST host program: a 32x32 filter matrix is being employed for converting the signal coming form the 32-capsules spherical microphone array to the 32 SPS signals.
A modern laptop, equipped with at least an Intel i5 processor, can easily perform such filtering in realtime, during the recording.
Nevertheless, we usually prefer to always record the “raw” 32-channels coming from the capsules, for being able subsequently to reprocess them with different sets of SPS-encoding filters, or for deriving directly other types of virtual microphones.

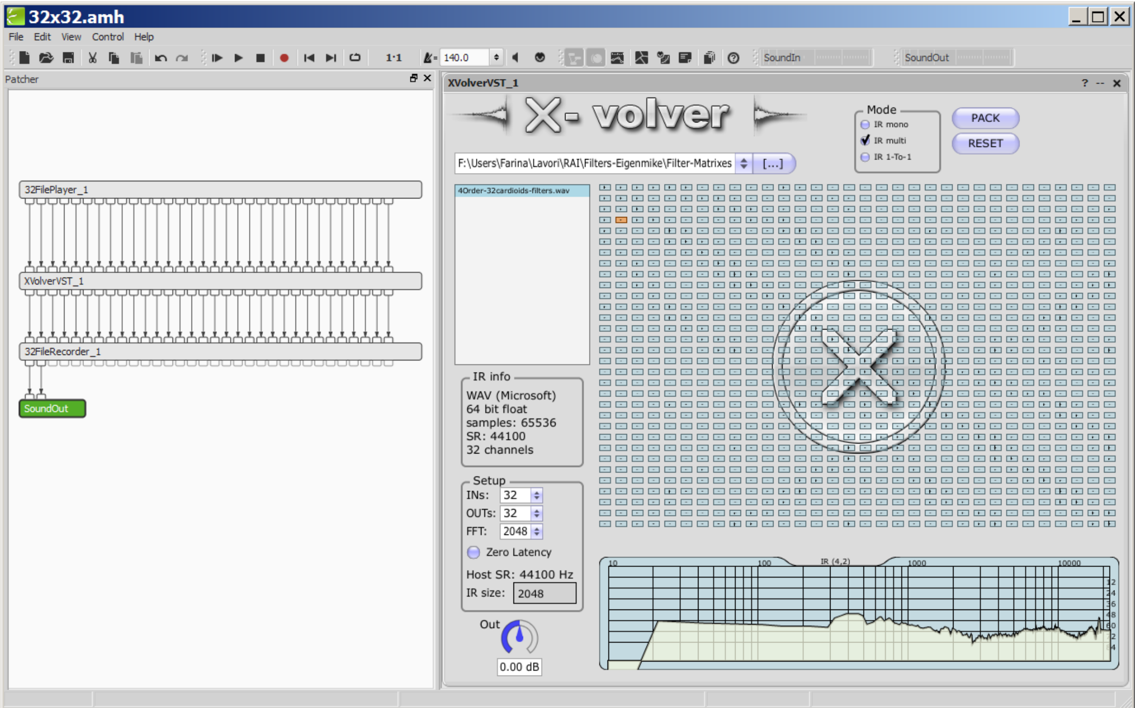


Figure 11: Graphical User’s Interface of X-volver, inside the Audio Mulch host program

## 5. POST PROCESSING

After the SPS recording has been made, it is possible to post-process the results in two ways:

- A graphical analysis can be performed, showing the spatial distribution of the incoming energy along the running time – this allows to “see” from where the sound is coming
- An audible rendering can be presented to a group of listeners, inside a special room equipped with a suitable array of loudspeaker, surrounding completely the listening area around a sphere

The graphical analysis is performed thanks to a Matlab program, which creates an animated colour video rendering of the sound map, overplotted over the 360°x180° panoramic image. A frame of such video rendering is shown in figure 6.
The audible rendering is obtained by reprocessing the SPS recording: a new set of virtual microphones is extracted, one feeding each loudspeaker of the playback array. The directivity and aiming of each of these virtual microphones is obtained by solving a linear equation system, imposing that the signals re-recorded placing the Eigenmike™ probe at the centre of the playback system are maximally similar to the original recorded signals. This approach, which is NOT Ambisonics-based, also corrects inherently for deviations from ideality of the loudspeakers employed, both in terms of magnitude/phase response, and in terms of placement/aiming/shielding.

### 5.1 Mapping room reflections inside La Scala theater

A first "real-world" experiment was conducted inside La Scala theatre in Milan. A set of 32-channels SPS impulse responses has been measured in 10 positions, along the stalls and the boxes, as shown in fig. 12.

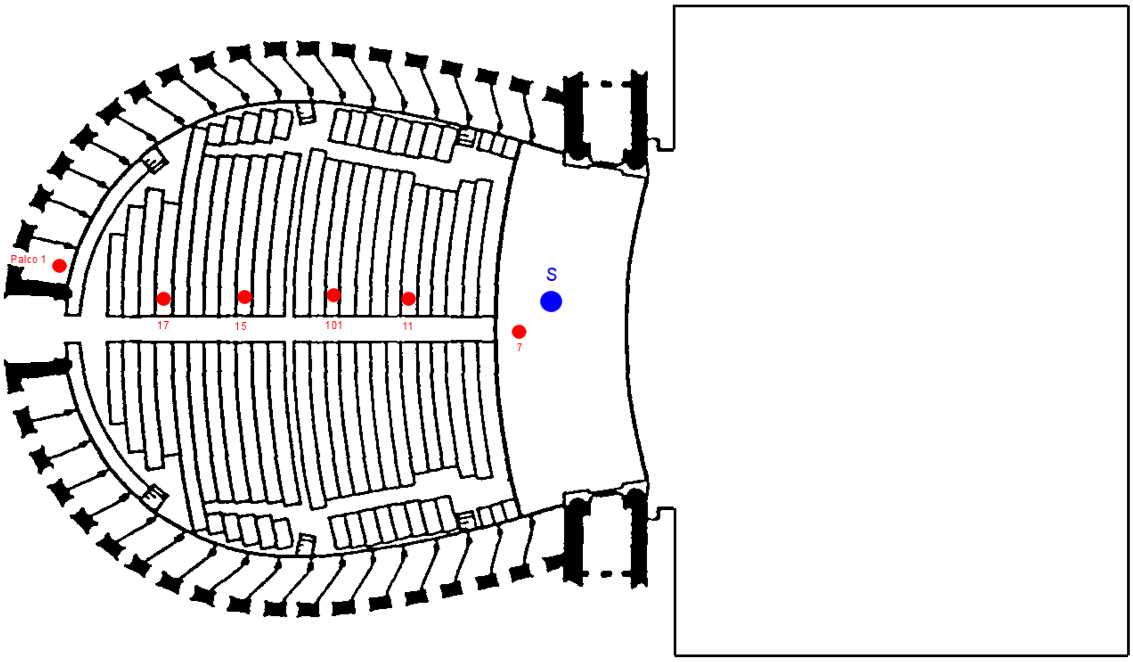



Fig. 12 – measurement positions inside La Scala theater

At each recording positions, a set of photos have been taken, pointing the compact wide-angle camera in every direction. These photos are thereafter stitched together, thanks to a free software tool made available by the Microsoft Research Center, and called ICE (Image Composite Editor).

This software performs a geometrical transformation, known as the Mercator's Projection, as shown in fig. 13.

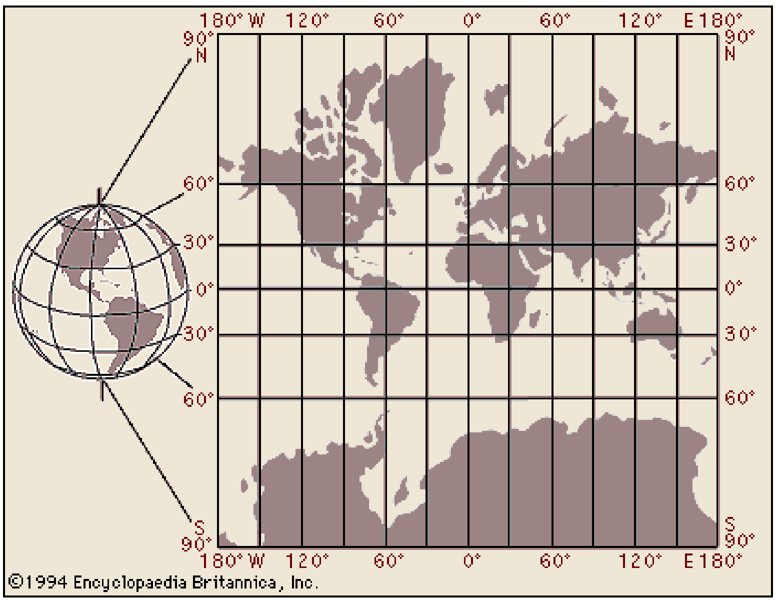



Figure 13: Mercator's Projection

Figure 14 shows Microsoft ICE in action, while creating the panoramic image of a box in La Scala theater:

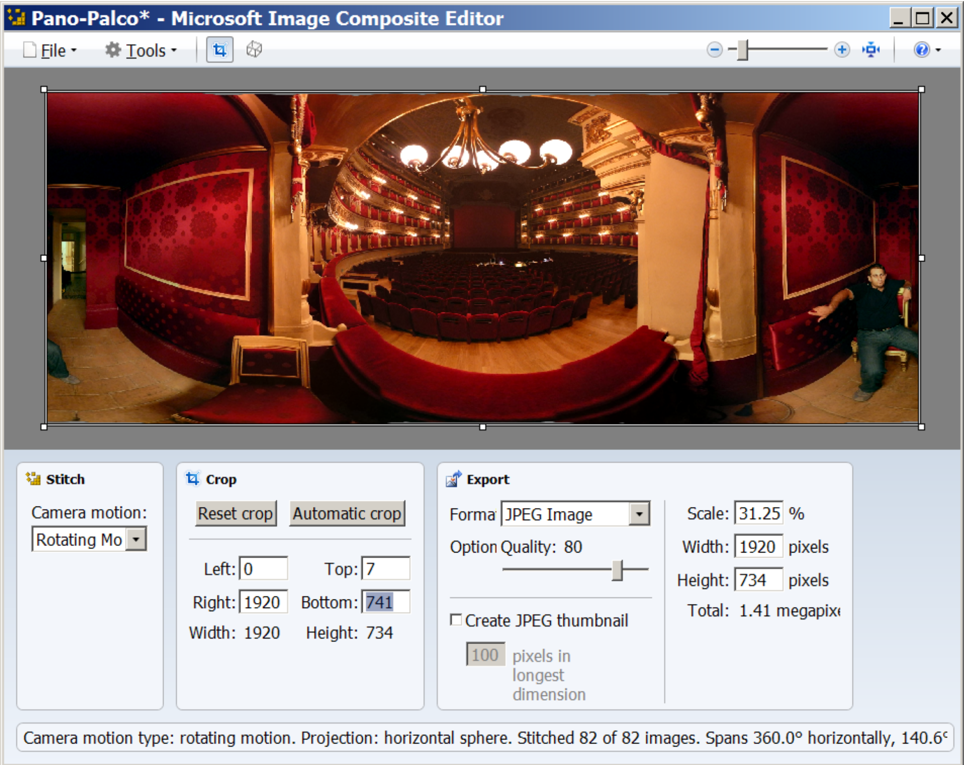



Figure 14: Microsoft ICE

By means of a Matlab program, the amplitude of the signals recorded by the 32 virtual microphones have been overplotted over these panoramic images, in the form a pseudo-color map. The map is recomputed every 5 ms, creating a slow-motion video of the sound distribution during the playback of the measured impulse response. This makes it possible to localize the direction-of-arrival of the most evident reflections.
Figure 15 shows a “frame” of the video rendering performed at the Director’s position, showing a strong reflection bouncing back from the side wall. The 3D impulse responses measured at La Scala have been employed also for high-quality spatial sound processing.

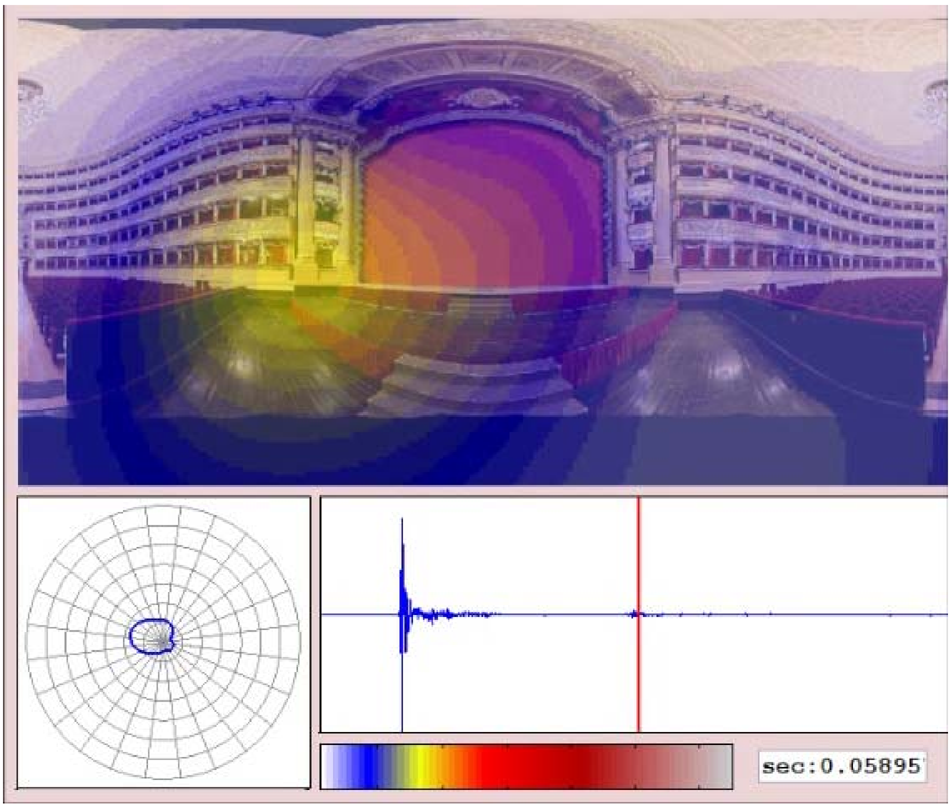



Figure 15 - Snapshot of the video processing, a strong reflection from the side wall

## 5.2 Sound reproduction over a 3D loudspeaker array

The SPS signals can be reproduced employing a suitable loudspeaker rig. This approach shares with Ambisonics the capability of rendering the signals over a generic loudspeaker array, in principle composed of an arbitrary number of transducer, and in arbitrary positions, as the SPS signals being transferred are not "speaker feeds", such as in 5.1, 7.1, 10.2, 22.2, etc. Instead, the 32 signals of the SPS signal are a "spatial kernel", codifying the whole spatial information, exactly as the Ambisonics signals. With the difference that the SPS signals are "PCM encoded", whilst the Ambisonics signals are in the domain of "spatial frequency".
So let's assume that we have a suitable listening room, equipped with a reasonable number of loudspeakers, more-or-less uniformly covering the whole sphere, as shown in fig. 16.

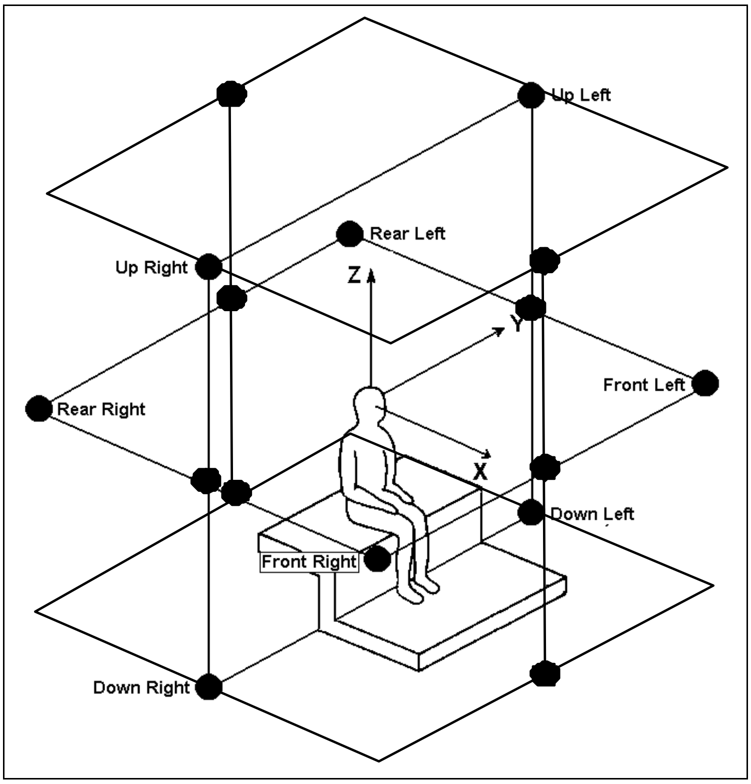



Fig. 16 – loudspeaker array with 16 loudspeakers and a listener at the centre

In our approach, there is no requirement for the loudspeakers to be equidistant from the listener, so they can be conveniently placed along the walls and in the corners of the room.
For feeding our 16-loudspeakers array with our 32-channels SPS signals, we need to create a "decoding matrix" of 32x16 FIR filters, with substantially the same mathematical approach employed for deriving the "encoding matrix" of 32x32 FIR filters, already described in chapter 4.
In practice the 32 SPS signals {y} must be convolved with the matrix of filters [f], yielding the required speaker feeds {s}:

$$s_r(t) = \sum_{i=1}^{32} y_i(t) * f_{i,r}(t) \quad \Rightarrow \quad \{s\} = \{y\} * [f] \tag{9}$$

For determining the filters [f], we start form a set of measurements of the loudspeaker's impulse responses, performed placing our 32-capsules microphone array at the centre of the listening room (in the "sweep spot position", where the head of the listener should be). Let's call [k] the matrix of these measured impulse responses.
The conditions to be imposed for finding the values of [f] are that the signals captured by the microphone array, if placed in the centre of the listening room, are identical to the "original" SPS signals {y}:

$$\{y_{out}\} = \{s\} * [k] = \{y\} * [f] * [k] \tag{10}$$

Of course, the recovered signals $\{y_{out}\}$ will never be really identical to the original ones $\{y\}$, some error will always occur, as shown in fig. 17.

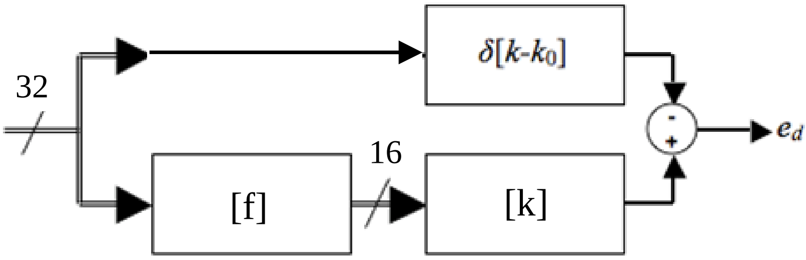



Fig. 17 – block diagram of the playback system

As we did for computing the encoding filters [h], we now set up a least-squares approach for finding the matrix of decoding filters [f], operating in frequency domain and employing a frequency-dependent regularization parameter $\beta$, and setting up a "modelization delay" $\delta$ of N/2 samples:

$$[F]_{16x32} = \frac{[K]^*_{16x32} \cdot e^{-j\pi k}}{[K]^*_{16x32} \times [K]_{32x16} + \beta_f \cdot [I]_{32x32}} \tag{11}$$

Again, the frequency dependence of $\beta$ is as shown in fig. 10, with frequency limits generally narrower than those used for encoding (typically loudspeakers have a more limited usable frequency range than microphones).
The creation of a pseudo-inverse of the reproduction matrix [k] is much more difficult than the inversion of the microphone matrix [c]: the inversion is optimal only if the loudspeakers are all identical, placed on a perfect sphere, as shown in fig. 18. This is the playback system employed by Nelson and Fazi [10] at ISVR, in Southampton, UK.

Fig. 18 – ISVR's spherical playback system

In our case, we employ a much worst playback system, as shown in fig. 19 (panoramic image): as the room is not really anechoic, and the loudspeakers are not located at the same distance form the center, the matrix becomes more tricky to invert, and the resulting filters need to be much longer, typically 4096 or even 8192 samples.

Fig. 19 – Listening room of Casa della Musica, University of Parma, ITALY

Despite the acoustical and geometrical deficiencies of such a listening room, the matrix of inverse filters does the magic, and listening to SPS recordings inside the playback system provides an experience really very similar to listening inside the original concert hall.
At the moment, we successfully performed this experiment just once: we recorder a trio of classical instruments (arpa, spinetta and trumpet) inside the concert hall of La Casa della Musica, in Parma, as shown in fig. 20, and performed listening tests in the playback room.

Fig. 20 – Recording of a trio at Casa della Musica, Parma, ITALY

## 6. CONCLUSIONS

This paper has described a new method for recording 3D sound signals, providing on one side a spatial resolution significantly better than what was obtainable with current technology, and, on the other side, allowing a very simple post-processing of the results, which allows both for an easy-to-understand graphical representation of the spatial-temporal information, and to replay the recorded signals inside a 3D surround system.

As the method is very new, we managed just to perform a single experiment for each of these two possible applications: indeed they confirmed that the method work significantly better than High Order Ambisonics, which is the alternative approach usable with the same microphonic probe and the same 3D sound playback system.

In the next future, other experiments will be conducted, employing this technology for different applications, such as analyzing the noise inside factories or outdoors, tracing the position of a moving sound source, extracting the voice of a single talker in the middle of a crowd, etc.

## 7. REFERENCES


[1] M.A. Gerzon, *Ambisonics in Multichannel Broadcasting and Video*, J. Audio Eng. Soc., vol. 33 no. 11, pp. 859-871 (1985 Nov.)

[2] Vilkamo J, Lokki T, and Pulkki V., *Directional audio coding: Virtual microphone-based synthesis and subjective evaluation,* J. Audio Eng. Soc., 57(9):709–724, September 2009.

[3] Svein Berge, Natasha Barrett, *A new method for B-format to binaural transcoding*, AES 40th international conference, Tokyo, Japan, October 8–10, 2010

[4] Sebastien Moreau, Jerome Daniel, Stephanie Bertet, *3D Sound Field Recording with Higher Order Ambisonics – Objective Measurements and Validation of Spherical Microphone* in 120th Convention of the Audio Engineering Society (2006)

[5] Daniel, J., Nicol, R., Moreau, S., (2003) “Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging” Presented at the 114$^{th}$ AES Convention, Amsterdam, 2003 March

[6] Angelo Farina, Alberto Amendola, Andrea Capra, Christian Varani *Spatial analysis of room impulse responses captured with a 32-capsules microphone array*, 130th AES Conference, London, 13-16 May 2011

[7] Kirkeby, O., Nelson, P.A., Hamada, H., Orduna-Bustamante, F., *Fast deconvolution of multichannel systems using regularization*, IEEE Transactions on Speech and Audio, **6,** (1998).

[8] O.Kirkeby, P.A. Nelson, P. Rubak, A. Farina , *Design of Cross-talk Cancellation Networks by using Fast Deconvolution*, 106th AES Convention, Munich, 8-11 may 1999.

[9] http://pcfarina.eng.unipr.it/Public/Xvolver/

[10] Poletti, M., Fazi, F.M. and Nelson, P.A. *Sound-field reproduction systems using fixed-directivity loudspeakers*, Journal of the Acoustical Society of America, 127, (6), 3590-3601, (2010).