19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

MEASURING SPATIAL IMPULSE RESPONSES IN CONCERT HALLS AND OPERA HOUSES EMPLOYING A SPHERICAL MICROPHONE ARRAY

PACS: 43.55.Cs

Angelo, Farina[1]; Andrea, Capra[1]; Lorenzo, Conti[1]; Paolo, Martignon[1]; Filippo Fazi[2]
[1]Laboratorio di Acustica ed Elettroacustica, Università di Parma; Via G.Sicuri n.60/A, 43100 Parma, Italy; www.laegroup.org
[2]Institute of Sound & Vibration Research, Southampton University SO171BJ, United Kingdom.

ABSTRACT

Traditional methods for measuring impulse responses in rooms provide detailed time and frequency resolution, but very poor spatial information.
On the other hand, when a disturbing echo or discrete reflection is found, it would be very useful to be able to understand its exact direction-of-arrival. And detailed spatial information is also required when the measured impulse responses are employed as digital filters for virtual listening tests.
The method described here is based on the usage of a small spherical microphone array, equipped with 32 subminiature capsules (Knowles Electronics) mounted flush on the surface of a rigid sphere (70mm diameter). A set of 32x25 digital FIR filters are employed for processing the signals coming from the capsules, and deriving the 25 channels corresponding to a spherical harmonics decomposition of the sound field, up to 4th order. Proper postprocessing tools allow for plotting the spatial information contained in the measured set of 25 impulse responses, making it easy to visualize the spatial distribution of sound during the propagation.
This set can also be employed inside a special listening room, equipped with a spherical array of loudspeakers, for reconstruction faithfully the original soundfield, providing a realistic and immersive listening experience.

INTRODUCTION

A 3D microphone could take an important role for the design of new theatres, music hall and spaces in which the “good sound” is the main goal. Having detailed spatial Impulse responses we will be able to determine the arrival direction of non-wanted reflections and echoes. Recording sounds with such a microphone also a surround listening could be more involving and realistic.

The theory on which we invested is the well know Ambisonics theory: it is based on the decomposition of the sound field in spherical harmonics of different order and, consequently, of different accuracy in the description of the field. Microphones that use 1st order spherical harmonics are quite diffuse (Soundfield, DPA-4, CoreSound TetraMic) but their sampling of the sound field doesn’t permit an accurate reproduction and description.

The research is growing up and we just have some examples of High Order Ambisonic microphones, such as the spherical one of France Telecom R&D [1] or the randomised-layout microphone of Trinnov [2], all made for having a 3rd order ambisonic microphone. They tend to the same goal but with different approach: the first is based principally on a heavy mathematical theory in which the regularization of the capsules’ layout and the polar pattern of the capsule have a fundamental role, the second one is based on a random disposition of the capsules and on filters “empirically” calculated. Our research tries to mix the two different approaches in order to have a 3rd order microphone.

COMPUTATION OF THE FILTERS

The theory

A set of digital filters can be employed for synthesizing the required spatial patterns, the spherical harmonics, either when dealing with a microphone array or when dealing with a loudspeaker array. In this case the N signals coming from the capsules need to be converted in M ambisonic signals, where the number M depends from the considered order: for the processing of the N inputs we need a bank of filters.

Assuming x_i as the signals of N microphones, y_j as the M ambisonic signals and h_{ij} the calculated filters the precedent scheme is translated in the formula below:

$$y_j = \sum_{i=1}^{N} h_{ij} \otimes x_i \quad \text{(Eq. 1)}$$

y_j represent the time-domain sampled waveform of a wave with well defined spatial characteristics, such as a spherical wave centred in a precise emission point, a plane wave with a certain direction and a spherical harmonic spectrum referred to a receiver point.

The processing filters h_{ij} are usually computed following one of several complex mathematical theories, based on the solution of the wave equation, often under certain simplifications, assuming 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 h_{ij} filters are derived directly from a set of impulse response measurements, designed according to a least-squares principle. A matrix of filtering coefficients is formed and the matrix has to be numerically inverted (usually employing some regularization technique); 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, diffractions, reflections, etc.).

Let we take, for example, the synthesis of a 0-order shape. The microphone array impulse responses $c_{k,j}$ are measured for a number of P incoming directions (Figure 1).

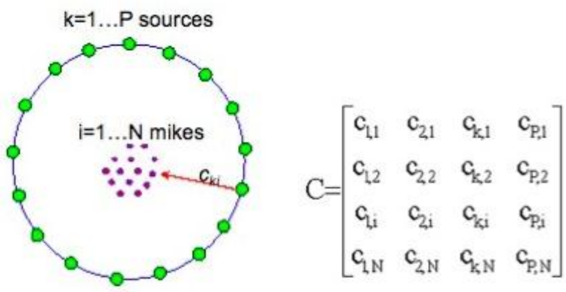


Figure 1.- Measurement of microphone array impulse responses

Processing the matrix C we can get the filters to be applied to the microphone signals. Filters are designed in such a way that the response of the system is the prescribed theoretical function v_k for the k-th source (an unit-amplitude Dirac's Delta function in the case of the example, as the 0-order function is omni directional). So a linear equation system of P equations can be set up imposing that the convolution of the h-matrix with the c-matirx should be a vector of Dirac's deltas (v_k). Once this matrix of N inverse filters are computed (for example employing the Nelson/Kirkeby method), the output of the array, synthesising the prescribed 0-th order shape, will again be simply:

$$y_0 = \sum_{i=1}^{N} x_i \otimes h_{i,0} \quad \text{(Eq.2)}$$

For computing the matrix of N filtering coefficients h_{i0} a least-squared method is employed. A “total squared error” is defined as:

$$\varepsilon_{tot} = \sum_{k=1}^{P}\left[\sum_{i=1}^{N}\left(h_{i0} \otimes c_{k,i}\right) - v_k\right]^2 \quad \text{(Eq.3)}$$

A set of N linear equations is formed by minimising ε_{tot}, imposing that for i=1 to N:

$$\frac{\partial \varepsilon_{tot}}{\partial h_{i0}} = 0 \quad \text{(Eq. 4)}$$

The development for a 4-channel microphone array

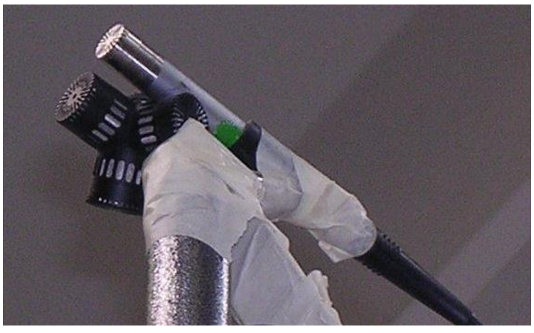
For a better comprehension of the operating procedure let we see the example of the design of filters for the DPA 4: it is a microphone with 4 capsules placed at 4 of the 8 vertexes of a cube (Front Left Down, Front Right Up, Back Left Up and Back Right Down). The goal is to derive, by proper processing, the 4 channels of a standard B-format stream (named W, X, Y, and Z) recording the 4 channels coming from the capsules and performing a proper matrixing of them in according with these formulas:

```
W'=+FLD+FRU+BLU+BRD
X'=+FLD+FRU-BLU-BRD
Y'=+FLD-FRU+BLU-BRD
Z'=-FLD+FRU+BLU-BRD
```

The problem was solved numerically, starting from anechoic measurements performed on the real microphone. These measurements can be carried out also without the availability of an anechoic room, as modern impulse response measurement techniques are now available, allowing us to easily "window out" unwanted room reflections. You need just to have a room big enough, so that the reflections arrive with a substantial delay after the direct sound.

Two sets of measurements are done. In the first set each of the 4 capsules is compared side by side wit a reference measurement microphone, as shown in the Figure 5a. The result of such a measurement is a stereo IR, containing the response of the capsule-under-test in the left channel, and the response of the reference microphone in the right channel. The transfer function between the capsule-under-test and the reference microphone is computed, and employed as a per-capsule equalizing filter. This ensures that all 4 capsules are perfectly identical, producing flat frequency response and linear phase response. The Aurora plug-in [3] named “Cross Functions” was used for performing the computation of the transfer function, saving the result as a time-domain waveform (an impulse response).

A second set of measurements is made processing the signals of the capsules through their equalizing pre-filters, and matrixing the resulting filtered signals with the formulas shown above. The results are of course W', X', Y' and Z'. The measurement is repeated three times, aligning the probe with the source along the three Cartesian axes, so that for each of the three channels X, Y and Z it is possible to get an impulse response which should, in principle, equate that of the reference microphone. The Figure 2b shows this second type of measurements (for Z).



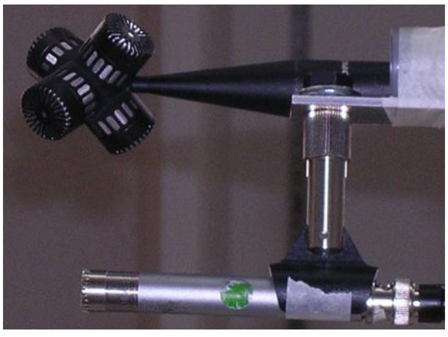


Figure 2a.- Comparison of the DPA-4 capsules with a reference microphone
Figure 2b.- Measure of the post-filter for Z axis

In this way, it is possible to measure the transfer function between each of the 4 processed signals (W', X', Y' and Z') and the reference microphone. These transfer functions are treated as the post-filters, performing the required equalization and conversion to the "correct" signals W, X, Y and Z. The post-filter for the W channel is scaled down by 3 dB (the coefficients are multiplied by 0.707), so that the resulting W channel will be gain-reduced, as the standard B-format signal is expected to be. At the end, we have 8 numerical filters, which are in the form of FIR filters (impulse responses) of suitable length, typically between 2048 and 8192 samples (depending on the low.-frequency limit of the analysis and on the sampling rate). These 8 filters (4 pre-filters, each per capsule, and 4 post-filters, each per output channel) should be used as shown in Figure 3.

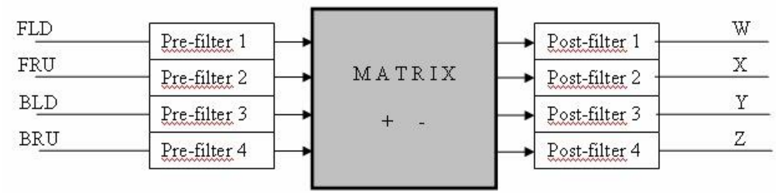


Figure 3.- Conversion from DPA signals to B-format signals

For performing this filtering/matrixing/filtering in real time, while recording with an A-format microphone, a fast multi channel convolver is needed. A first solution consists in using the Convolver-VST open-source program [4]. This plug-in can be loaded with an 8-channels WAV file, containing the 8 filters obtained from the measurements. Of course, after having performed the first 4 pre-filters, the signals must be matrixed before being passed through the second group of 4 post-filters. AudioMulch [5] or Plogue Bidule [6] can be used as host programs. The second solution is to use, in the precedent hosts, the VST plug-ins “DPA-4 Console”, made with SonicBirth [8], in which the conversion shown in Figure 4 is implemented using our pre-calculated filters.

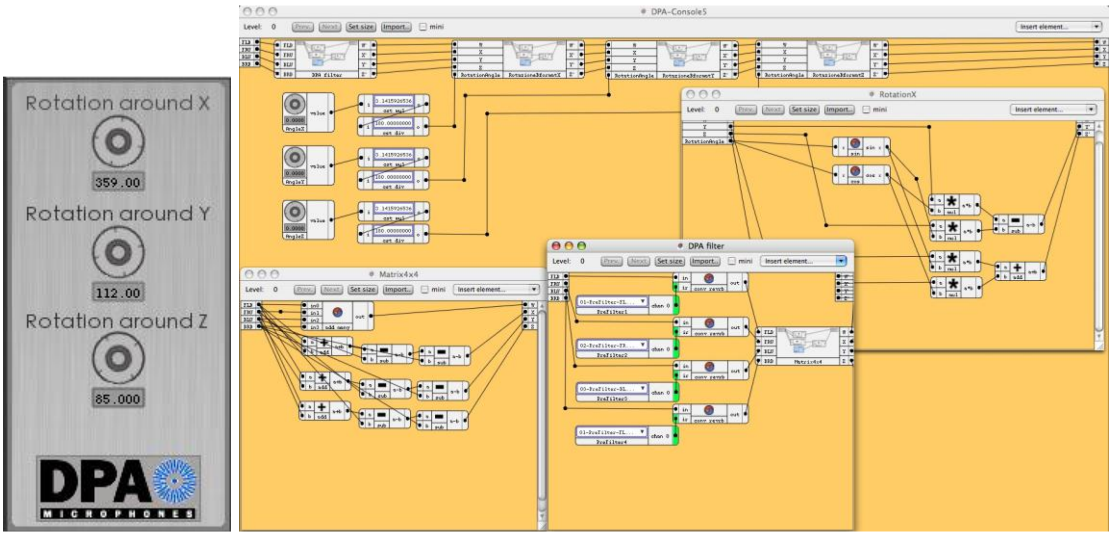


Figure 4.- Skin of “DPA-4 Console” and patch of SonicBirth

THE 32-CAPSULES MICROPHONE

The design

According with the prototype of France Telecom [1], we chose 32 capsules placed on the 12 vertices of a dodecahedron and 20 vertices of an icosahedron: they make a sampling of the sound filed in which they are immersed. In analogy with the temporal sampling we can consider a sort of “Shannon theorem” for spatial sampling [7]. In order to limit the spatial aliasing the series of Fourier-Bessel is truncated at a certain order L that mainly depends on the number of sampling points (capsules). So, the maximum order L available using N capsules, is obtained from the formula:

$$(L+1)^2 \leq N \quad \text{(Eq. 5)}$$

With 32 capsules the maximum order it should be the 4^{th}.

The sphere in which the capsules are mounted has a diameter of 7 cm, a compromise between the reduction of the spatial aliasing at high frequencies and the reliability for the bass frequency’s sampling at higher orders. With this kind of layout there are two different angular gaps between the vertices and, in a first optimistic view, they permit to have spatial aliasing between 6 to 7 KHz [1].

For a more accurate study of the influence of spatial aliasing depending on the order of the spherical harmonics, we must refer to the spherical Bessel and Hankel functions. We studied the influence of each order of harmonics on the spherical spectrum of the sound field as a function of the distance from the origin (r) and of the frequency (f). From the Matlab simulations (Figure 5a) it is possible to see that the contribution of 0^{th} order, at 1 m and 180 Hz, is null whilst at the same frequency the contribution of 1^{st} and 2^{nd} orders are equal.

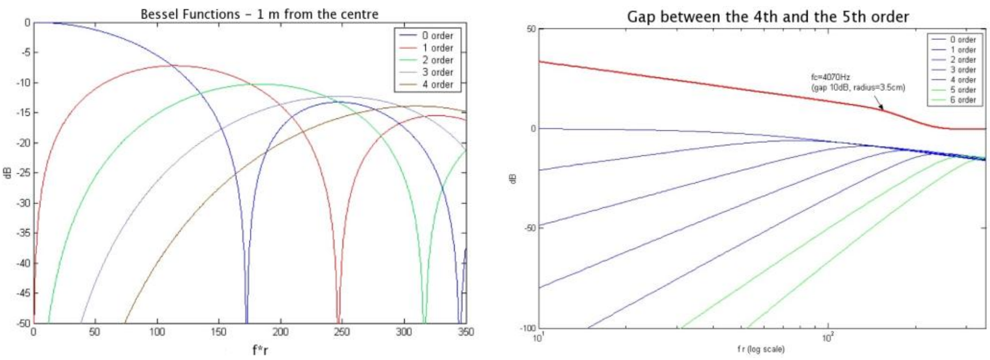


Figure 5a.-Spherical Bessel functions at 1 meter from the centre of the microphone
Figure 5b.-Spherical Hankel functions: gap between the 4^{th} and the 5^{th} order harmonics

To prevent spatial aliasing between 4^{th} and 5^{th} order when measuring the sound field diffracted by a rigid sphere located at the origin, we need to maximize the gap between the two radial functions corresponding to the spherical Hankel functions (Figure 5b): if the level of the 5^{th} order’s harmonics, compared with the one of the 4^{th} order, is low then it should be non-influential. Unfortunately we have the maximum of this gap for very small values of the product f*r and we need a compromise between that value of the product and the value of the gap (Figure 5b). Assuming a gap between the 5^{th} order harmonics and the 4^{th} order harmonics of 10 dB, the critical frequency, limit between paltry and relevant aliasing, is around the 4070 Hz. To prevent the effects of the aliasing on the 4^{th} order component it would be reasonable to make a smoothing filter (low pass filter) for that component with a cut off frequency at 4 KHz.

The realization

The sphere, designed with CAD, is made of expanded-polyurethane and was realized using a milling cutter controlled by a computer; the sphere was divided into two parts in order to simplify the mounting of the capsules and then glued with silicon.

The 32 sub miniature capsules (Knowles Electronics) require a power supply of 1.5 volts, provided by a self made power pack that is connected with all the cables of the capsule through a multi-core cable and Edac connectors.

Figure 6.-The CAD drawing of the sphere (a), the two parts of the sphere (b), the self made power pack (c)

The sphere is mounted on a steel tube (60 cm) with a thread at the bottom in order to mount it on a standard microphone support. A steel tubular cover, designed with a non influential reflectivity angle, is used to protect all the connections between the wires of the capsules and the multi core cable.

Figure 7 -The microphone array

The multi core cable arriving from the power pack is connected through XLR connectors to 4 converters (Behringer ADA8000, 8 channels, 48 kHz, 24 bit), controlled through ADAT by a Firewire M-Audio ProFire sound card. With a Mac laptop we can use the sound card giving directly the power supply through the firewire port.

CONCLUSIONS

The same procedure descript here for a 4-capsules microphone will be implemented for our new 32-capsules-microphone, obtaining a filtering matrix derived directly from a set of impulse response measurements. This method, used instead a completely theoretical one, corrects for transducer deviations or acoustical artefacts and gives a microphone response close to the ideal responses. The dimension and the disposition of the capsules give the capability of synthesising spherical harmonics up to the 4th order, with good rejection of 5th order harmonic's spatial aliasing.

References: [1] S. Moreau, J. Daniel, S. Bertet: 3D soundfiled recording with high order ambisonics – objective measurements and validation of a 4th order spherical microphone. 120th AES Convention, Paris

[2] http://www.trinnov.com

[3] http://www.ramsete.com/aurora/

[4] http://convolver.sourceforge.net/vst.html

[5] http://www.audiomulch.com

[6] http://www.plogue.com

[7] A. Laborie, R. Bruno, S. Montoya: A new comprehensive approach of surround sound recording. 114th AES Convention, Amsterdam

[8] http://www.sonicbirth.sourceforge.net/