Acoust. Sci. & Tech. **26**, 2 (2005)

TECHNICAL REPORT

Measurements and reproduction of spatial sound characteristics of auditoria

Angelo Farina[1,*] and Lamberto Tronchin[2,†]

[1]*Industrial Engineering Department University of Parma, via delle Scienze 181/A, 43100 Parma, Italy*

[2]*Department of Energetic, Nuclear and Environmental Control Engineering, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy*

(*Received 17 July 2004, Accepted for publication 17 January 2005*)

Abstract: The definition and measurement of sound spatialisation have been strongly enhanced in last years, as nowadays spatial properties of sound propagation are considered quite important during design of auditoria. Besides, a proper description of spatiality is requested during virtual audio reproduction of sound quality in dedicated listening rooms for 3D reproduction purposes. Normally, only binaural measurements are performed, by means of a dummy head, even though international standards like ISO 3382 require measuring some spatial parameters (i.e. *LE*, *LF*, *IACC*). 3D impulse responses are rarely measured and utilised for sound reproduction. In this paper, an innovative procedure of measuring spatial sound characteristics is presented. The application of this new technique in virtual 3D sound reconstruction is emphasized. Furthermore, the methodology is compared with other techniques of 3D sound reproduction. Finally, the results of a wide campaign of measurements of spatial parameters among different auditoria all over the world, and conducted with the novel methodology, are compared with the results of standard binaural and 3D measurements. The possibility to enhance the spatial reproduction of sound quality in real spaces and the comprehensibility of spatial parameters is then considered and presented in different cases.

Keywords: Measurements, Auralization, Spatial sound properties

PACS number: 43.55.Ka, 43.55.Hy [DOI: 10.1250/ast.26.193]

1. INTRODUCTION

The night of 29 January 1996 the famous and renowned Gran Teatro La Fenice in Venice burned completely. One of the best sounding opera houses in the world suddenly disappeared. However, its sonic behavior was at least partially saved, since several acoustical measurements had been performed just two months before the burning, employing the binaural impulse response technique [1].

M. Gerzon [2] first proposed to systematically collect 3D impulse responses measured in ancient theatres and concert halls, to assess their acoustical behavior and preserve them for posterity. His proposal found sympathetic response only very recently, with the publication of the "Charta of Ferrara" [3] and the birth of an international group of researchers who agreed on the experimental methodology for collecting these measurements.

In order to obtain a complete description of sound spatialisation in the auditorium, a new measurement method is described [4], which incorporates all the previously known measurement techniques in a single, coherent approach: three different microphones are mounted on a rotating beam (a binaural dummy head, a pair of cardioids in ORTF (Office de Radiodiffusion Télévision Française) configuration, and a Soundfield microphone), and a set of impulse responses are measured at each angular position. The ORTF configuration represents a standard method (adopted by French Television) for recording dual-channel signals, in which the cardioid microphones are spaced 170 mm and are diverging by 110 degrees. The Soundfield microphone, introduced by M. Gerzon [2], allows measuring 4-channels impulse responses, and therefore spatial properties of sound field. A Soundfield microphone captures 4 signals, known as "B-format" signals: one omnidirectional (pressure) and three with a polar pattern called "figure of eight," oriented along the three cartesian axes X, Y, Z (these three channels capture a signal proportional to the Cartesian components of the particle velocity vector).

*e-mail: angelo.farina@unipr.it
†e-mail: lamberto.tronchin@unibo.it

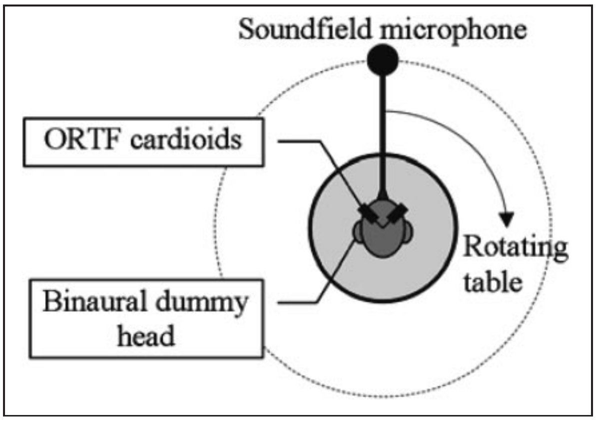


Fig. 1 Scheme of microphones.

Figure 1 shows a schematic of the microphone setup.

The results obtained from the measurements are compatible with the already proposed methods for measurements in concert halls (Binaural, B-format), but allow also to derive standard surround formats, as described in section 2.4, such as OCT (Optimized Cardioid Triangle) and INA (Institut National de l'Audiovisuel) [5], as well as advanced three-dimensional reproduction techniques, such as the novel Binaural Room Scanning method [6], the Poletti high-order circular microphones [7], the Wave Field Synthesis (WFS) approach [8] and the Ambiophonics hybrid approach [9].

The paper briefly describes the measurement results, and more in detail the process of employing the experimental impulse responses for advanced method of auralization, aimed to the realistic reproduction of the spatial properties of the original soundfield.

2. MEASUREMENT OF MULTICHANNEL IMPULSE RESPONSES

The combination of three different measurement methods, properly combined, provides a general method from which all known multi-channel formats could be derived.

2.1. Test Signal and Deconvolution

The excitation-deconvolution technique employed for the measurement of the impulse response is the log sine sweep method, as initially suggested by one of the authors [10]. A good compromise between measured frequency range, length of the sweep and signal-to-noise ratio has been reached, by choosing the parameters shown in the following Table 1.

The traveling time of the rotating table requires a longer silence between sweeps. The rotation is triggered by a proper impulsive signal, automatically generated in the middle of the silence gap on the second channel of the sound card.

Table 1 Characteristics of log sine sweep utilized in the measurements.

Start frequency	22 Hz
End frequency	22 kHz
Length of the sweep	15 s
Silence between sweeps	10 s
Sweep type	Exponential

The choice of the above parameters allows the measurement of impulse responses, which have wide frequency span, good dynamic range (approximately 90 dB) and are substantially not influenced by background noise (which could be found in the rooms during the measurements).

2.2. The Sound Source

An omnidirectional sound source is usually preferred for measurements of room impulse responses. Albeit the dodecahedron does not correspond to the effective directivity pattern of real-world sound sources (such as musical instruments or human talkers and singers), the usage of an omnidirectional sound source is predicated by current standards (i.e. ISO 3382), and avoids exploiting strange room effects (abnormal concentration of energy and focalizations for selected orientations of the source), as can happen employing highly directive loudspeakers.

A special, ultra-compact dodecahedron loudspeaker was specifically built for measurements presented here, employing 12 full-range drivers installed on a small size enclosure (the diameter is about 200 mm). As this unit is not capable of producing significant acoustical power under 120 Hz, a subwoofer was added, in order to extend the low frequency range. It is incorporated inside the cylindrical transportation case, which also contains the power amplifier (300 W RMS) and serves as supporting base for the dodechaedron. A small Italian company, LookLine, specialized in high quality, custom-built loudspeaker systems, built the doedechaedron.

2.3. The Microphones

Three different microphonic probes were employed:

- a pair of high quality cardioids in ORTF configuration (Neumann K-140, spaced 180 mm and diverging by 110°);
- a binaural dummy head (Neumann KU-100);
- a B-format 4-channels pressure-velocity probe (Soundfield ST-250).

All these microphones were installed over a rotating table, in such a way that the rotation center passed through the center of the dummy head, and through the point at the intersection of the axes of the two cardioids (which were mounted just above the dummy head). Alternatively the Soundfield microphone was displaced exactly 1m from the

Table 2 List of Auditoria measured.

N.	Theatre	N. sources/receivers
1	Uhara Hall, Kobe, Japan	2/2
2	Noh Drama Theater, Kobe, Japan	2/2
3	Kirishima Concert Hall, Kirishima, Japan	3/3
4	Greek Theater in Syracuse, Italy	2/1
5	Greek-Roman Theater in Taormina, Italy	3/2
6	Auditorium of Parma, Italy	3/3
7	Auditorium of Rome (Sala 700), Italy	3/2
8	Auditorium of Rome (Sala 1200), Italy	3/3
9	Auditorium of Rome (Sala 2700), Italy	3/5
10	Sydney Opera House (opera theater), Australia	4/3
11	Sydney Opera House (concert hall), Australia	3/3
12	Sydney Opera House (the studio), Australia	3/1

rotation axis, in front of the dummy head. The angular step was 10° (36 measurements for the complete rotation).

2.4. Measured Data

Twelve famous theaters in Japan, Italy and Australia were measured with the previously described method; the list of theatres and the number of source/receiver positions are reported in Table 2.

A very different sound quality characterized the theatres. The theatres ranged from a very reverberant hall, like Auditorium of Parma or Auditorium of Rome (Sala 2700), to very dry theatres, as Greek theatre in Syracuse (open space). From the impulse responses a wide set of acoustical parameters have been calculated, ranging from reverberation time to IACC and LE, and others. These data are not presented here: when the study will be completed (the program is to measure approximately 50 halls), a detailed comparative analysis of the measured data will be presented. The focus here is instead on the usage of the measured impulse responses as digital filter for auralization experiments.

3. METHODS FOR AURALIZATION

The basic method for auralization is convolution: the impulse responses are employed as very long FIR filters, applied to dry (anechoic) recordings of music or speech. Convolution is a very efficient filtering technique, particularly if implemented with proper (old) algorithms on fast (new) processors: as clearly demonstrated in [11].

In the following it is described how to create sets of impulse responses suitable for being employed by these software convolvers, creating the results in any of the currently available formats suitable for multichannel reproduction, and attempting to recreate as faithfully as possible the spatial attributes of the original soundfield.

3.1. ORTF-Stereo Impulse Responses

This is the most basic processing, aimed at the creation of a "standard" stereo presentation of the results of the auralization. The process is based on the availability of a number of dry mono recordings, one for each section of the orchestra or for each singer.

Each mono recording has to be convolved with a specific stereo impulse response, obtained by the pair of cardioid microphones in ORTF configuration. In principle, each of these impulse responses should be measured with the proper position of the sound source. Finally, the results of the convolution of all the dry recordings are summed in a single stereo output file, which is suitable for reproduction in a normal stereo system (2-loudspeakers).

3.2. Binaural Impulse Responses (Binaural Room Scanning)

The basic binaural approach is substantially the same as for the previous ORTF-based method, but employing the binaural IRs. This way, the result of the convolution is a 2-channels file, suitable for headphone reproduction.

However, two methods can be employed for substantially improving the surround effect obtained: for loudspeaker reproduction a proper cross-talk cancellation must be added, and for headphone reproduction an head-tracking sensor can drive a real-time convolver, switching the impulse responses being convolved as the listener rotates his head.

3.3. B-Format Impulse Responses (Ambisonics)

In this case, each dry mono source is convolved with the proper B-format impulse response. So, after the mixing of all these convolutions, a 4-channels B-format output is obtained.

The reproduction of a B-format signal over a suitable array of loudspeakers requires an Ambisonics decoder, for computing the proper feed for each speaker.

The creation of a software-based decoder has been pioneered by one of the authors [12], and has been further perfected by colleagues at the University of York, who recently released for free a suite of VST plugins [13], allowing for manipulation and decoding of B-format signals over various loudspeaker rigs.

3.4. ITU 5.1 Surround (from Selected B-Format Impulse Responses)

The basic approach for ITU (International Telecommunications Union) 5.1 rendering is to first select a configuration of microphones to be employed, for driving the 5 main loudspeakers [5]. Many of these microphone arrangements have been proposed, and in a recent round-robin project, called the Verdi project, most of them were comparatively evaluated [14].

The systems differ each other for the directivity, position and orientation of the microphones. Here we consider just two of them, which got good results in the aforementioned comparative test: Williams MMA (Multichannel Microphone Array) [15] and OCT [5]. Figure 2 shows the microphone configurations for these two setups.

For each of the above setups, it is possible to select a subset of 5 of the 36 positions where the Soundfield microphone was displaced, corresponding as close as possible to the intended positions of the chosen setup. Then, from the B-format impulse response measured in each of these 5 selected positions, a single (mono) impulse response is extracted, thanks to the program Visual Virtual Microphone, developed by David McGriffy and freely available on the Internet [16].

Visual Virtual microphone computes the response of a microphone which can have any directivity pattern (from omni to figure of eight, passing through subcardioid, cardioid and hypercardioid); this virtual microphone can be aimed anywhere, specifying the azimuth and elevation.

Finally, each mono dry source is convolved with the 5-channels impulse response derived from the corresponding sound source position over the stage, and the results of all these convolutions are mixed in a single final 5-channels track, which is suitable for reproduction over a standard ITU loudspeaker rig.

3.5. Mark Poletti's High-Directivity Virtual Microphones

During the rotation of the microphonic assembly, the two cardioids employed for ORTF recordings also describe a small circumference, with a radius of approximately 110 mm, Looking for simplicity to just one of the two microphones, it samples 36 impulse responses during its complete rotation. From this set of data, it is possible to derive the responses of a set of various-orders coincident microphones, ideally placed in the center of rotation, making use of a modified version of the Poletti's theory [7].

The basis of this method is to define a class of multileaf-shaped horizontal directivity patterns of various orders. The order 0 is an omnidirectional, order 1 are two crossed figure-of-eight microphones (as in horizontal-only Ambisonics); then order 2 and 3 are added, with directivity patterns corresponding respectively to the cosine of twice and three times the angle. Figure 3 shows the shape of

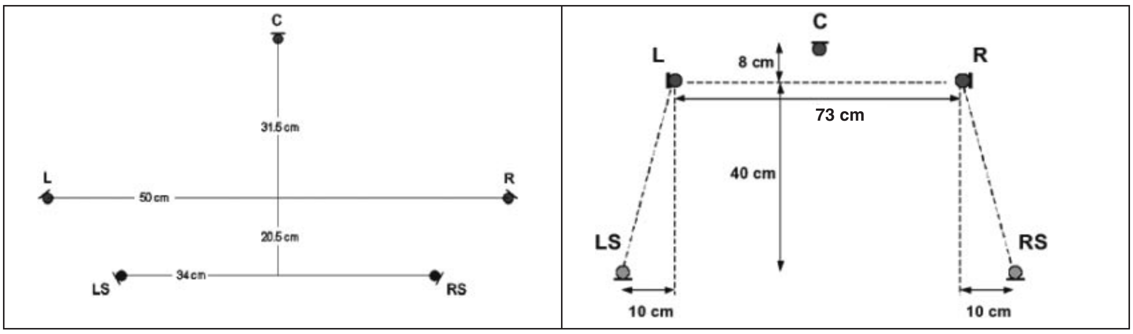
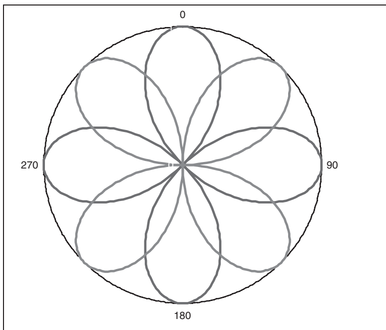


Fig. 2 Layout of 5 microphones. Left: Williams MMA; right: OCT.



$$D_0 = 1$$

$$D_{1,n} = \cos\left(\vartheta + n \cdot \frac{\pi}{2}\right) \qquad n = 0,1$$

$$D_{2,n} = \cos\left(2 \cdot \vartheta + n \cdot \frac{\pi}{4}\right) \qquad n = 0,1$$

$$D_{3,n} = \cos\left(3 \cdot \vartheta + n \cdot \frac{\pi}{6}\right) \qquad n = 0,1$$

Fig. 3 Poletti's theory: concept (left) and formulas (right).

these high-order patterns, and the corresponding formulas.

Form these, an advanced high-order Ambisonics decoder (horizontal-only) can be fed. A second possible way of employing these high-order signals is to drive a standard 5.1 ITU array, by synthesizing 5 proper asymmetrical directivity patterns, as suggested in [17].

3.6. Circular WFS Approach

The 36 B-format measurements made along the 1m-radius circumference are exactly the set of data required for employing the WFS method described in [6], and outlined in the following Fig. 4.

The basis of this method is the Huygens principle. The WFS is a 2D reduction of this general theory, where the microphones are placed along a closed curve around the listening area, and consequently the expansion/shrinking can only be done in the horizontal plane. Starting with a 1m-radius array, it is quite easy to derive the feeds for a loudspeaker array suitable for a medium-sized listening room, and to "stretch" the array so that the loudspeakers are arranged in 4 linear arrays instead of in a circular array [8].

3.7. Hybrid Methods (Ambiophonics)

The Ambiophonics method is a hybrid solution, aimed to mask the defects of two basic systems: cross-talk cancelled reproduction of binaural material over closely-spaced loudspeakers (Stereo Dipole) and 3D surround driven by convolution of corresponding oriented virtual microphones.

The following Fig. 5 shows a typical Ambiophonics array, (frontal stereo dipole, plus 8-loudspeakers surround rig).

The theory for deriving the signals for these loudspeakers has been already presented above, and the assembly of the whole system has been thoroughly described in [4].

4. LISTENING ROOMS

Two listening rooms suitable for three-dimensional auralization have been designed and built: the "ASK"

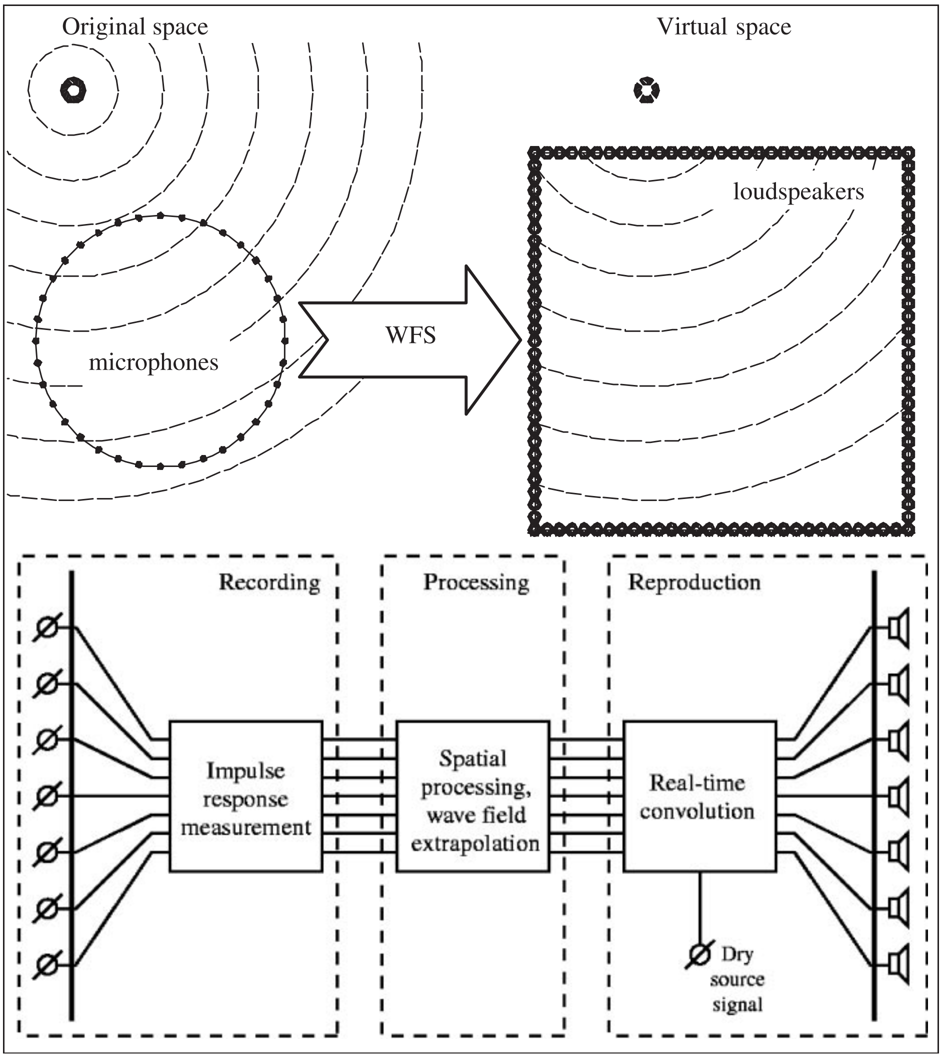


Fig. 4 WFS processing scheme: concept (above) and implementation (below).

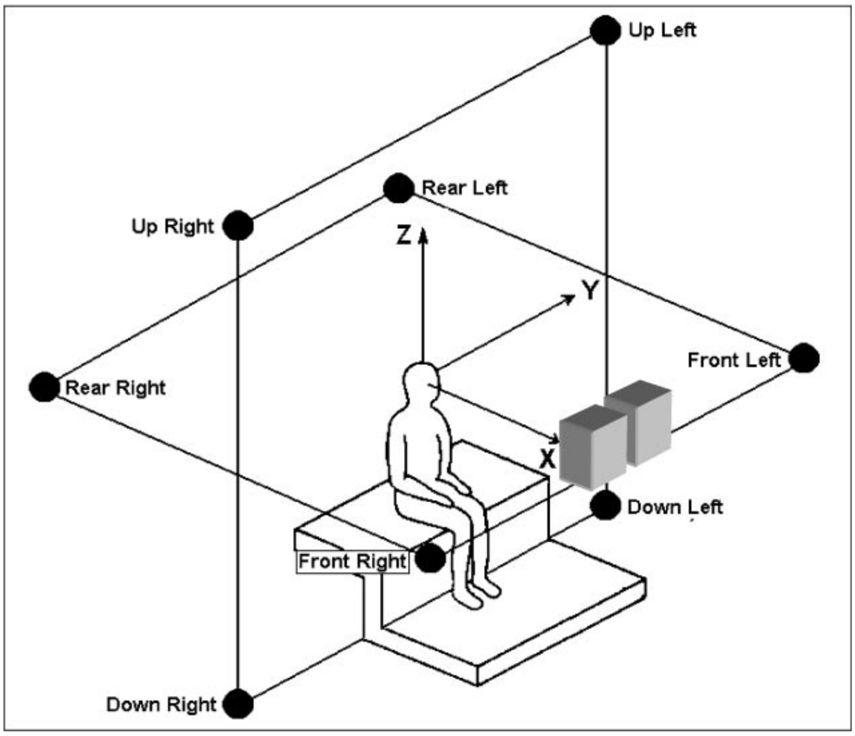


Fig. 5 Ambiophonics array.

room in Reggio Emilia and the "Arlecchino" room in Bologna [18]. Both rooms are equipped with an 8-loudspeakers Ambisonics array, plus a frontal and rear Stereo-Dipole pair. The "ASK" room is also equipped with a subwoofer.

In both cases the room is provided with a multi-channel sound board (12 channels) directly connected to a PC in which special cross-talk filters have been designed for reproduction over front and rear Stereo-Dipole techniques and a dedicated software for Ambisonic decoding has been utilized. During the calculation of the filters, frequency responses of the loudspeakers have been equalized, and the minor time delay differences have been compensated for.

A special care was reserved to the architectural design of the rooms. Reverberation time was designed ranging from 1.3 s at 63 Hz to 0.3 s at 8 kHz, since preliminary psycho-acoustical experiments in listening rooms suggested to avoid an excess of absorption in the room, and therefore improve spaciousness in the room.

5. CONCLUSIONS

This paper has described how it is possible to derive subsets of impulse responses suitable for the reproduction of virtual acoustic spaces, following the currently available reproduction technologies. Referring in particular to the reproduction of the spatial properties of the sound field, it is noticeable that the data measured with advanced techniques allow for the auralization of the results employing:

– Standard stereo reproduction over a pair of loudspeakers;
– Binaural reproduction over headphones, with head tracking;
– Reproduction over closely-spaced loudspeakers (by means of cross-talk cancelling filters);
– Ambisonics reproduction over a 2D or 3D regular array of loudspeakers
– ITU 5.1 "surround" reproduction conforming to "standard" microphonic setups (OCT, INA, etc.)
– High directivity, multichannel reproduction by means of Mark Poletti's circular-array method.
– Wide-area auralization by means of the Wave Field Synthesis approach (WFS)
– Any combination of the above methods, resulting in hybrid, higher-level surround methods (Ambiophonics and derivations).

Consequently, this method provides the best available approach for storing the acoustical properties of famous and valuable rooms, such as concert halls and theatres, and preserving them for the posterity. The resulting data can be used for audible reconstructions (auralization) by means of today's surround systems, without limiting the future usage by sticking to the limited reproduction technology currently available.

REFERENCES

[1] L. Tronchin and A. Farina, "The acoustics of the former Teatro "La Fenice," Venice," *J. Audio Eng. Soc.*, **45**, 1051–1062 (1997).
[2] M. Gerzon, "Recording concert hall acoustics for posterity," *J. Audio Eng. Soc.*, **23**, 569–571 (1975).
[3] "Carta di Ferrara," *CIARM*, http://acustica.ing.unife.it/ciarm/

Carta.htm

[4] A. Farina and R. Ayalon, "Recording concert hall acoustics for posterity," *24th AES Conf. Surround Sound, Techniques, Technology and Perception*, Banff, Canada 26–28 June (2003).

[5] G. Theile, "Multichannel natural music recording based on psychoacoustic principles," *AES 19th Int. Conf.*, May (2001).

[6] A. Karamustafaoglu, U. Horbach, R. Pellegrini, P. Mackensen and G. Theile, "Design and applications of a data-based auralisation system for surround sound," *106th AES Convention*, Preprint 4976 (1999).

[7] M. A. Poletti, "A unified theory of horizontal holographic sound systems," *J. Audio Eng. Soc.*, **48**, 1155–1182 (2000).

[8] E. Hulsebos, D. de Vries and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave-field synthesis," *J. Audio Eng. Soc.*, **50**, 779–790 (2002).

[9] A. Farina, R. Glasgal, E. Armelloni and A. Torger, "Ambiophonic principles for the recording and reproduction of surround sound for music," *19th AES Conf. Surround Sound, Techniques, Technology and Perception*, Schloss Elmau, Germany, 21–24 June (2001).

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

[11] A. Torger and A. Farina, "Real-time partitioned convolution for Ambiophonics surround sound," *2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House New Paltz, New York, October 21–24 (2001).

[12] A. Farina and E. Ugolotti, "Software implementation of B-format encoding and decoding," *104th AES Convention*, Amsterdam, 15–20 May (1998).

[13] A. Field, "B-dec high resolution first order ambisonic B-format decoder," University of York, http://www.york.ac.uk/inst/mustech/3d_audio

[14] J. Roland, MultiMedia Projekt VERDI, TU Ilmenau Laboratory, Germany (2002), http://www.stud.tu-ilmenau.de/˜proverdi/daten/um1en.html

[15] M. Williams and G. Le Du, "Multichannel microphone array design," *108th AES Convention*, Preprint 5157 (2000).

[16] D. McGriffy, "Visual Virtual Microphone," http://mcgriffy.com/audio/ambisonic/vvmic

[17] E. Hulsebos, T. Schuurmans, D. de Vries and R. Boone, "Circular microphone array for discrete multichannel audio recording," *114th AES Convention*, Amsterdam, 22–25 March, Preprint 5716 (2003).

[18] L. Tronchin, V. Tarabusi and A. Giusto, "The realization of ambisonics and ambiophonics listening room "Arlecchino" for car sound systems evaluation," *Proc. 21st AES Conf.*, St. Petersburg, Russia (2002).