Audio Engineering Society

Convention Paper

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA



# Capturing Manipulation and Reproduction of Sampled Acoustic Impulse Responses

Ronen Ben-Hador[1] and Itae Neoran[1]

[1] Waves Audio Ltd., Tel-Aviv, Israel
ronenb@waves.com and itai@waves.com

**ABSTRACT**

**We discuss the capturing manipulation and reproduction of impulse responses (IRs) of acoustic spaces. While trying to maintain the accuracy of an IR, other factors such as sound quality and musical character of sound, should also be considered. Furthermore, IRs are not limited to preserving the sound of venues but also as a tool in music production. Therefore, the IRs are converted to standard multi-channel reproduction formats, such as stereo and ITU 5.0. In order to obtain a flexible reverb tool, the IRs are manipulated to modify acoustic properties such as reverb time and inter-channel de-correlation. A new real-time audio plug-in was developed for which IRs of venues and devices were recorded worldwide. The IRs are convolved with dry audio. The plug-in supports mono, stereo, and surround, at sample-rates up to 96kHz**

## *1.1. Introduction*

Recently the growing power of CPUs and DSPs has driven new reverberation tools for music production, where venue impulse responses (IRs) are used via convolution. These tools are chiefly designed to preserve the character of a certain venue when convolved with dry music or sound, hence they mainly consist of convolution engines. The procedure of capturing and processing the IRs is a subject for continuous research effort. Gerzon [1] first proposed to start a collection of 3D impulse responses measured in ancient venues, for preserving it for posterity. Since, new methods for recording halls have been developed which also support the possibility of reproducing various reproduction formats such as stereo and surround. Blesser. B [9] gives a detailed description of IR production in terms of accuracy, noise, and perceptual criteria.

Farina [3] presents an apparatus designed to capture an IR, where he uses a log sweep technique. He also gives an extensive description of a fast convolution algorithm (Torger [13]). In another paper he describes the speaker microphone array and describes the methods to record mono stereo and surround from the recorded output of the array [4].

We have used this apparatus to record IRs in venues around the world. The IRs need to be processed so they are not dominated by background noise and so they sound natural and pleasing to the listener. This process is actually being directed by a group of 6 listeners, all of them experienced sound mixing engineers in music production, with an aim to achieve the most satisfying reproduction system. Many possibilities of reproduction exist (microphone and speaker types and geometrical configuration) especially for surround with competing methods. Therefore many configurations need to be evaluated where some dictate the structure of the recording system and the processing of the IRs.

The main surround reproduction methods are Ambisonics([2]), Ambiophonics([7]) and ITU 5.0 discrete surround (Holman [5]) and combinations (Holman [5]).

Another issue which should be taken care of is the ability to manipulate an IR, so as to control its acoustical properties.

In music production, and especially for studio recordings, the reverberation process is used in a creative way in order to naturalize, enhance, emphasize, or hide certain instruments or vocals. Thus, the reverberation character needs to be matched to the audio musical track. In these applications, sound engineers are looking for ways to manipulate the IR sound rather than preserve it. For example lengthening or shortening the IR or changing relative level of the IR parts: direct sound, early reflections, late reflections (tail). While these manipulations change the IR, the user still expects to perceive the original venue with added features.

In section 1 of this paper we describe the capturing system in section 2 we describe the pre-processing applied to the recorded IRs. In section 3 we present IR manipulations, and in section 4 we describe the reproduction of IRs and comparison between reproduction methods.

## 1. Capturing

The considerations when recording IRs are mainly: the reproduction format, the recording procedure, and the noise.

### 1.1 Microphone considerations

Several microphones types may be considered for capturing the response of a room. They mainly differ in their radiation pattern and therefore produce different sound. Directional microphones such as cardioids are designed to capture the maximum of acoustic energy in a certain direction. The main advantage of such microphone is the fact that it is gives a better sense of directionality. Omni microphones on the other hand capture all the energy at the vicinity of the microphone they are do not give as good sensation of direction as directional microphone, but they give better sense of the venue since they could record energy coming from the sides and the rear. Soundfield microphones produce four components which are X, Y, Z and an omni component W. Their main advantages are for surround reproduction (Ambisonics) and in addition they can simulate different radiation pattern microphones by a simple transformation. Binaural microphones are microphone located in a dummy head ears. They are mainly used to simulate the ears and head which is the main advantage of these microphone. Since each microphone type has its advantage The recording are done with all the types above.

### 1.2 Excitation signal

There mainly two types of source used for capturing IR. MLS (Maximum Length Sequence) and Chirps. Following Farina [3] we are using a log sweep:

$$x(t) = \sin\left[\frac{w_1 T}{\ln\left(\frac{w_2}{w_1}\right)}\left(e^{\frac{t\cdot\ln\left(\frac{w_2}{w_1}\right)}{T}} - 1\right)\right]$$

where T is the total time of the sweep and w1 , w2 are the frequencies at time zero and time T. We have used this source because of its advantages which are: the inverse is the same as the sweep with reverse order. As shown by Polletti [9] and Farina [3] that the non linear distortion are pushed to be in front of the linear responses and therefore will appear before the direct arrival which makes it very easy to cut them later in the processing stage. We are using a 15 seconds sweep which results in a noise floor of around -100 dB. By using a longer sweep we could theoretically get a better signal to noise ratio, but in fact, for longer sweep durations the IR starts to smear due to slight air pressure variations in the acoustic space. As a consequence, not only the IR is distorted, but also due to this smearing, as our experiments show, the actual improvement of signal to noise is only by 2-3 dB for a 4 times longer sweep tone. Therefore, we found that 15 seconds sweep is the optimal duration. The frequency range we use for the sweep is 22Hz to 32 kHz.

**1.3 Recording Setup**

We used a specially designed rotating table to record the sweep responses of acoustic spaces. **Figure 1** depicts the apparatus used. It consists of a vertical bar with a dummy head ((Neuman KU-100) placed at the top and just below two ORTF (Neuman K-140) directionamicrophones. Soundfield microphone are placed on the edge of a horizontal bar attached to the vertical bar, the length of the horizontal bar is adjustable, and was set to 0.5 meter and 1 meter in length.

On the bottom there is a rotating table that rotates the vertical bar, controlled by an electrical engine. This setup is necessary since reproduction of stereo and surround format demands recording at different angles. Furthermore, for supporting future reproduction setups which use more speakers it is advisable to record as many angles as possible. The emerging Wave Field Synthesis (Hulesebos et al [6]) also requires such a rotation. We took measurements at angular steps of 10, 15 and 22.5 degrees. In addition, we placed two widely spaced omni-directional microphones (Earthworks STC) in a setup known as A-B method. Although an omni pattern could also be reproduced from the Soundfield microphone, the maximum distance between recordings depends on horizontal bar so there is a need for another pair.

**1.4 Loudspeaker**

We used Genelec S30D loudspeakers which have an almost flat response over the audible spectrum. For each angle of rotation the source is recorded 3 times: in front of the apparatus, in the left-front of the apparatus, and in the right-front, with equal spacing.

### 1.5 Reproduction set-ups

Figures 2 to 5 depict the speaker-microphone configurations designed for different reproduction formats.

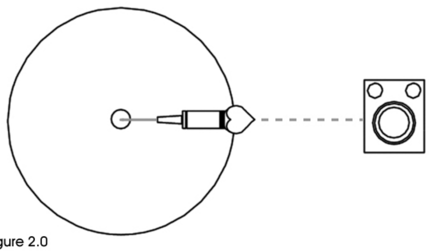

Figure 2.0
Mono to Mono

**Figure 2** is a mono to mono configuration, where the speaker is directly in front of the source. In reproduction one convolution is needed.

**Figure 3** is a mono to stereo configuration, where the speaker is at the center and the microphones used are at angles of -55 and 55 degrees. In reproduction two convolutions are needed.

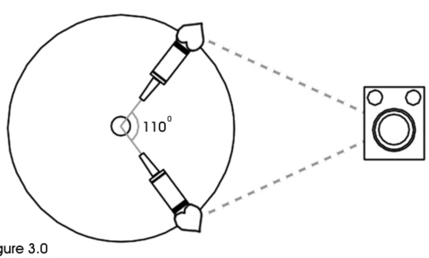


Figure 3.0
Mono to Stereo

**Figure 4** is a stereo to stereo configuration, where first a left speaker is applied and then the right speaker.

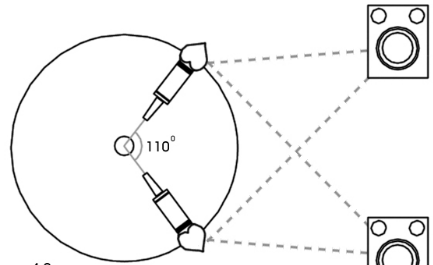


Figure 4.0
Stereo to Stereo

The microphones in stereo to stereo are placed the same as in mono to stereo. It takes 4 convolutions: Left speaker to left microphone left speaker to right microphone, right speaker to left microphone and right speaker to right microphone.

The Omni recording is depicted in **figure 5**. The mono to mono is produced by taking the left microphone excited from the left speaker. The omnis do not revolve and are not attached to the revolving table such that the distance between them can be altered.

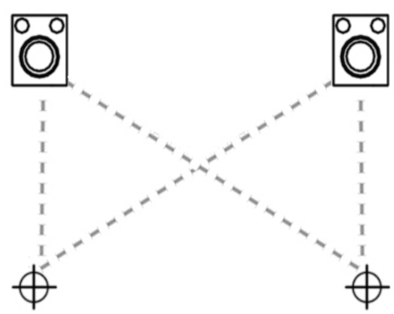

Figure 5.0
Omnis

For surround recording the source receiver configuration depends on the reproduction formats, where the main being: Ambisonics, Ambiophonics, ITU 5.0 discrete surround (also 7.0 and 10.0), and WFS. For Ambisonics we use the SoundField microphone. We also decode, in real-time, the output of the convolution to the ITU 5.0 reproduction setup. Gerzon[2], and Malham[12] decode the music to five or more speakers, based on reconstructing wave fronts that are represented

by spherical harmonics . Ambiophonics (Farina[7] Glasgal [11]) may be reproduced from our recording setup using various angles of cardioids along with the binaural recordings.
ITU 5.0 discrete surround may be produced by either discrete microphone positions, or by pre-decoding the output of the SoundField microphone. The locations of the microphones in a circular array do not correspond to the 5.0 microphoning standards (Williams, OCT, INA). However, a close location to the standards may be produced. The WFS which uses Huygens principle to exactly reconstruct the wave field via Kirchhoff-Helmholtz integral. The circular array fulfills the requirement for WFS. For more details on surround see section 4.

## 2. Pre-processing

### 2.1 De-convolution

The recorded sweep-responses are subject to several processing stages. First they are convolved with an anti-sweep. As described by Farina [4], the de-convolution is obtained by linear convolution with the time reversal of the excitation signal, amplitude equalized if needed.

### 2.2 Removal of pre-responses

When using the log-sweep signal, since most of the non-linear parts (harmonics from loudspeakers and microphones) are pushed to earlier times than the linear parts, they can be safely removed from the de-convolved IR. This procedure is done by locating the direct arrival and cutting in a smooth manner the earlier part of the IR.

### 2.3 Fade-out

The de-convolved IR is then cut in length. This operation is needed since after a certain amount of acoustic decay, the noise dominates the tail of the IRs. In order to cut an IR in a natural way, an exponential regression to the RT60 curve is calculated, and the IR is faded in a natural manner that is an extrapolation of the RT60 curve, down to the requested dynamic range, and then another fast fade-out is applied. This operation is also needed since convolution is a time consuming process, and for efficient real-time processing the IR should be as short as possible. Obviously this process attenuates the noise which dominates the later parts of the IR tail. The noise floor is estimated from a silence-recorded channel, and then the RT60 is calculated. It is important to emphasize that due to algorithm by Schroeder [10] the RT60 should be calculated after removing the noise floor.

### 2.4 Equalization

Since most speakers tend not to have an exact flat response at all frequencies contained in the log sweep excitation signal, the impulse response has to be properly equalized. This operation does not always turn to be trivial since it is desired to avoid notches. In addition, as the Genelec loudspeaker has a directional radiation pattern, different reflections are subject to different frequency responses depending on the direction in which they left the source. Thus, the equalization used needs to be an average over the complete radiation pattern of the loudspeaker. We have computed such an average for the Genelec speakers and provided a matched equalizer.

### 2.5 Virtual microphone

IRs are produced for cardio, SoundField, and omni microphones, which provide a large variety of IRs to choose from, since no unique microphone setup has been shown to have the "best" sound but rather different setups have been shown to be pleasing for different applications. Different microphone radiation patterns are preferred in different applications. SoundField microphones enable us to produce Cardio, Omni, ms stereo, and other radiation patterns, by simple transformations. David Mccgriffy [20] wrote a nice interface to calculate virtual microphone.

### 2.6 Normalization

In reproduction, IRs from different venue location are used. Moreover, some reproduction methods require different microphones types as an input. This may lead to distortion as the different recorded channels are not balanced for amplitude. Normalization is also required to preserve the correct distance effect when replacing the recorded direct sound with a dry signal, a common practice is many studio applications. The normalization procedure should use the closest microphone to the speaker, also, it should have the biggest amplitude response from all the other microphones type if more than one microphone type is used. For the rotating table we used cardio at angle zero (right in front of speaker) as a reference for normalizing.

## 3. Manipulations

A major use of the IRs is for music production, so it is worthwhile to manipulate the IR to achieve certain qualities. In this section we describe some of the user-controlled manipulations which are to the recorded IR.

For certain types of manipulations, such as distance control or coloration control, it is best to process only the early-reflections part of the IR. For other manipulations such as reverb-time modification or density modification, it is preferred not to manipulate the direct and the early reflection part of the IR, but rather to apply manipulations to the late reflections since the later exhibit statistical qualities and manipulation of the earlier part pf the IR could result in an unexpected outcome.

### 3.1 Detection of direct part of IRs and its removal

In most cases the direct part of the IR presents itself by it's high amplitude. Nevertheless an early reflection may posses higher amplitude when the microphone is not at angle 0. In addition, the duration of direct is not always obvious, since a very early reflection may interfere with the long direct response, resulting from the combined responses of the loudspeaker, the air, and the equalization filters. Therefore, an algorithm to detect the direct arrival and to determine its length was applied, so the direct portion can be further manipulated (removed, scaled). The algorithm is based on finding the first significant local minima. Removing the direct is an operation needed for replacing it with a digital dry signal. The latter is needed mainly in music production for minimum alteration of the input sound, when only a small amount of reverb is added to a musical track.
Note also that for the stereo-to-stereo configuration, the direct path consists of arrivals from two loudspeakers to two microphones (4 paths), resulting in coloration of the summed response if the stereo input of the convolution is panned to the center.

When the direct part is removed, a cross-fade is applied, and the pre-roll of the IR is cut away so that the peak of the original (missing) direct is aligned with the dry signal.

### 3.2 Separable gain and delay to direct, early reflections and late reflections

Many times it is of a musical value to scale different parts of the IR as it grants new qualities to the IR (see Gerzon [14]) although engineers used partitioning long before Gerzon [14] based on practical experience. The task of partitioning the IR to its components is done by:

a) Detect the direct (see section 3.1 above).
b) From first reflection take the minimum between RT10 time (the time the Shroeder integral decays by 10 dB) and 80 milisec, to be the early reflections zone.
c) Later part of the IR is the 'tail' (late reflections).

Via the above partition it is then possible to manipulate each of the IR components separately. The conditions to determine the length of the early reflections is based upon subjective listening criteria and include directional cues, distance cues and coloration cues (see Begault [16] p.101).

### 3.3 Reverb time modification

Reverb time is measured by RT60 which is the length of time the energy decays to –60 dB (Schroeder [10]).
To lengthen or shorten effective reverberation time, a time-stretching algorithm is used. Many polyphonic time-stretching algorithms are available, see for example Zolzer [15]. The choice of time-stretching algorithm is such that preserves the overall average statistics and does not destroy transients.
Gerzon [14] (section 9.5 p.24) explored the possibility of connecting reverb time of early reflections with room distance information, through the use of Craven hypothesis.

### 3.4 De-correlation

Lowering the correlation between different recorded channels can enhance the perceived spaciousness. This can be done in 3 ways: (1) By increasing the order of the Ambisonics recording (for example a high-order Ambisonics microphone, see 4.4.4), (2) By using an alternative psycho-acoustic decoder from Ambisonics B-format to stereo or 5.0 (as suggested in the SIRR algorithm [18], see 4.3), and (3) artificially by processing the IR to compensate for the inherent correlation existing in the recording set-up.

In (2) and (3), we make sure the de-correlation process does not cause undesired artifacts such as coloration. Another requirement is to maintain compatibility to later down-mixing of the channels, from stereo to mono, or from surround to stereo.

The artificial de-correlation algorithm (3) is based on random time shifts of parts of the IR in the late reflections (time stretch). Since time shift may cause unwanted panning and comb-filter effects, the time shifts should be random (adding or reducing time). The shifts need to be large enough to reduce the chance of reflections interfering and thus avoid comb effects, other algorithms for de-correlation such as pitch shift were tried (see Holman [17]) but were found unsuitable to IRs due to lack of mono down-mixing compatibility. Listening tests on the recorded IRs and on the reproduced tracks showed that artificial de-correlation can be applied successfully to the late reflections of the IR without corrupting the sound quality.

### 3.5 Envelope and other manipulations

Changing the IR envelope can be achieved by applying a user-controlled sequence of exponential curves connecting fixed gain points. Other manipulations (not in the scope of this paper) include: equalizing the IR, damping filters, gating, reverse IR shape, manipulation of perceived room-size, and manipulations of room resonances.

## 4. Reproduction

The processed IR is applied to a dry source in real-time via convolution. In this section we describe the process of applying the IR to the source.

### 4.1 Convolution Engine

The convolution of the IR with the dry track is time consuming. We used a fast convolution algorithm called equal time slice, which is based on partitioning the time series into equal partitioned sections and applying convolution through FFT algorithm. This algorithm has a trade-off between the time cost (latency) and number of partitions (total CPU). The size of each section is determined by the required latency, but also by other considerations such as the dictated buffer sizes from the host of the plug-in, and other constrains of the environment (like free memory available). The convolution engine accepts any IR in mono, mono to stereo, stereo to stereo, mono to surround, stereo to surround, and surround to surround.

### 4.2 Efficient stereo

The efficient stereo component designed to save computer resources if limited. Instead of using 4 convolutions as required by the stereo to stereo component. It uses only 2 convolutions by assigning the left source to be convolved with left microphone IR and the right source to the right hand side microphone. To create an artificial cross-talk between opposite channels the user is allowed to pre-mix some of the left channel into the right and vise-versa.

### 4.3 Spatial reproduction of room acoustics (SIRR)

A technique for spatial reproduction of room acoustics, Spatial Impulse Response Rendering (SIRR), has been recently proposed (pulkki [18]). In the method, a multichannel impulse response of a room is measured, and responses for loudspeakers in an arbitrary multichannel listening setup are computed. When the responses are loaded to a convolving reverberator, they will create a perception of space corresponding to the measured room. The method is based on measuring with a sound field microphone or a comparable system, and on analyzing direction-of-arrival and diffuseness at frequency bands. An omnidirectional response is then positioned to a loudspeaker system according to analyzed directions and diffuseness. In this paper the SIRR method is reviewed and refined. The reproduction quality of SIRR and some other systems is evaluated with listening tests, and it is found that

SIRR yields a natural spatial reproduction of the acoustics of a measured room.

### 4.4 Surround

We describe two surround reproduction configurations which we thought to be reprehensive of surround formats. Ambisonics (within ITU 5.0) and Discrete (ITU 5.0 or Quad within ITU 5.0), where the audio inputs are mono tracks or stereo tracks.

#### 4.4.1 Ambisonics (Horizontal B-format decoded to ITU 5.0)

Horrizontal Ambisonics is based on the existence of 3 sound field component (X,Y,W). It uses a decoding formula to produce 5 loudspeaker feeds (see **Figure 6**), usually Front Left, Front Center, Front Right, Rear Left and Rear right.. Reproducing surround in horizontal B-format requires 3 convolutions for a mono track and 6 for a stereo track. If an IR was not recorded with a SoundField microphone it should be pre-encoded to B-format. More on Ambisonics encoders and decoders formulas are available at http://www.york.ac.uk/inst/mustech/3d_audio/ambis2.htm.

The Ambisonics surround IRs were reported by our group of mixing engineers to give the best balanced surround with a good sense of direction.

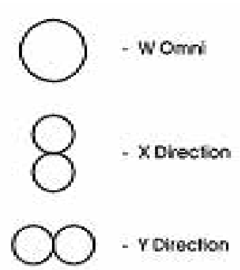


Figure 6.0 A
Soundfield First Order Sphreical Harmonics Components: W,X,Y

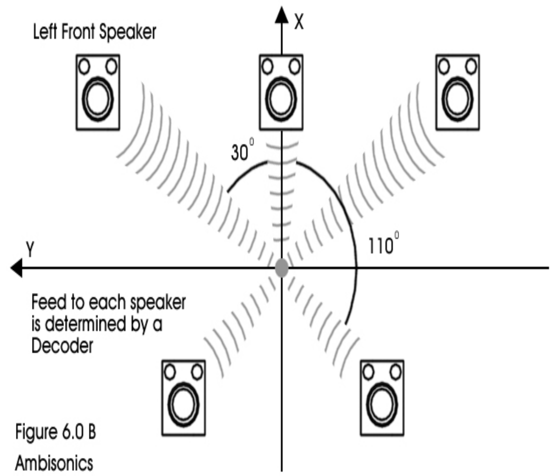


Figure 6.0 B
Ambisonics

#### 4.4.2 Discrete surround (ITU 5.0 and quad within ITU 5.0)

Quad discrete surround, is reproduced within ITU 5.0, where the center channel is used only for the dry signal (when replacing the direct). It requires an IR of 4 channels recorded at angles corresponding to frontal stereo pair and rear omni stereo pair (see **Figure 7** ). The front ones are taken from the ORTF setup, and the rear are taken from Omni microphones at 90 and –90 degrees, this combination was preferred by our listeners (the group of mixing engineers), as it minimizes the cross-channel correlation while still preserving the sense of direction. Quad format require 4 convolutions for a mono track and 8 for stereo tracks.

A complete ITU 5.0 5-channel discrete configutaion may also be achieved by using 5 convolutions for a mono source. However using 3 frontal microhpones is problematic in that it would cause increased coloration on the direct sound and on major early-reflections from the front. Thus, alternatives for the center channel are either up-matrixing 2 recorded microphones to the 3 ITU frontal channels, or the SIRR approach [18].

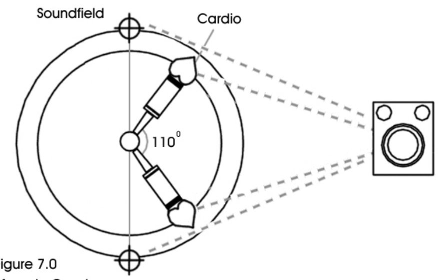


Figure 7.0
Mono to Quad

### 4.4.3 Comparing surround configurations.

We tested several configurations trying to reproduce the most pleasing results to be chosen by the mixing engineers listening group. The group found that first order Ambisonics gave better sensation of direction and ambience, but had a smaller sweet-spot than the discrete IRs.

In the discrete surround configurations, the experiment included mainly a pair of ORTF for the front since they were shown to give the most pleasing stereo. As rear channels we have tried pair of ORTF's in a rear stereo configuration (135 and 225 degrees) we also tried virtual cardioid derived from soundfield mikes at a different diameter from the center. Also other configurations which included a central cardioid a frontal stereo pair and rear stereo pair. The group of listeners clearly voted for the ORTF as frontal center and a pair of Omni's in 90 and 270 degrees as rear channels at a larger diameter than the ORTF's from center. The omni's perceived a better ambience of the room. This result might be attributed to the fact that the omni's gave better sensation of ambience, since they record all direction as opposed to directional microphones.

### 4.4.4 High order microphone

In a first order horizontal soundfield only three components are recorded (W-omni,X and Y) In the second order horizontal case two additional channels U and V are produced and for a third order another two channels R and T should be produced. U,V,R,T are spherical components of the soundfield and their production requires a large number of microphones to be placed. Theoretically high order should give a better notion of directionality which is missing in the first order, and would also yield a better channel separation that is essential for reducing the cross-channel correlation. The high order microphone although technologically challenging seem to be vital to decrease the correlation of the channels. Also an added value of more accurate soundfield and enlargement of sweet spot area which tend to be narrow with first order is achieved in high order. Therefore our future vision for a better surround IR and general music capturing is in the development of a high order microphone.

## References

**1**. Michael Gerzon-"Recording Concert Hall Acoustics for Posterity", JAES vol. 23, Number 7 p.569 (1975).

**2**. Michel Gerzon- "Hierarchical System for Surround Sound Transmission for HDTV" 92 AES convention. Viena.

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

**4**. A Farina. R Ayalon-"Recording Concert Hall Acoustics for posterity". AES 24-th international Conference on Multichannel Audio. Banff Canada 26-28 June (2003).

**5.** T. Holman "5.1 Surround Sound Up and Running". Focal Press ISBN 0-240-80383-3.

**6.** E. Hulsebos D.de Vries and E. Bourdillat- "Improved Microphone Array Configurations for Auralization of Sound Fields by Wave-Field Synthesis", JAES Vol. 50 Number 10 P.779 (2002).

**7.** A. Farina R. Glasgal, E, Amelloni, A. Troger- "Ambiophonic Principles for the recording and Reproduction of Surround Sound for Music" 19th AES Conference on Surround Sound, Techniques, Technology and Preception- Schloss Elmau, Germany 21-24 June (2001).

**8.** B. Blesser " An Interdisciplinary Synthesis of Reverberation Viewpoints. J. Eng. Soc. Vol 49. No. 10 p.867 (2001).

**9.** M. Polleti-"Linearly swept frequency measurements, time delay spectrometry, and the Wigner distribution", JAES vol. 36 Number. 6 P457 June (1988).

**10.** M Schroeder –"New method for measuring reverberation time". J. Acoust. Soc. Am. Vol .37 P 409 (1965).

**11.** R. Glasgal –" Ambiophonics, The science of domestic concert Hall design", http://www.ambiophonics.org/.

**12.** D. Malham. "3-D sound for virtual reality systems using Ambisonic techniques". http://www.york.ac.uk/inst/mustech/3d_audio/vr93papr.htm [1993].
**13.** A. Torger, A. Farina-"Real-Time partitioned convolution for ambiophonics surround sound". 20001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics-Mohonk Mountain House New Paltz, New York October, p.21 (2001).
**14.** M. Gerzon.-"The Design of Distance Panpots". AES 92nd convention Vienna. March 24-27 (1992).
**15.** U. Zolzer –"Dafx Digital Audio Effects"- Wiley, ISBN 0 471 490784 (2003)
**16.** Duran Beagault-"3-D sound for virtual reality and multimedia" AP professional ISBN 0-12-084735-3 (1957).
**17.** T. Holman –" New factors in sound for cinema and television" 89 AES convention, September 21-25, Los Angeles,(1990).
**18**. Ville Pulkki, Juha Merimaa, Tapio Lokki. Reproduction of Reverberation with Spatial Impulse Response Rendering. 116AES Convention, Berlin 8-11 may (2004).
**19** T. Chen, Working in B+ format**.**
http://www.ambisonic.net/chenbplus.html
**20**. D. Mccgriffy. Visual Virtual Microphone.

HTTP://MCGRIFFY.COM/AUDIO/AMBISONIC/VVMIC/