Ambiophonic Principles for the Recording and Reproduction of Surround Sound for Music

Angelo Farina (1), Ralph Glasgal (2), Enrico Armelloni (1), Anders Torger (1)

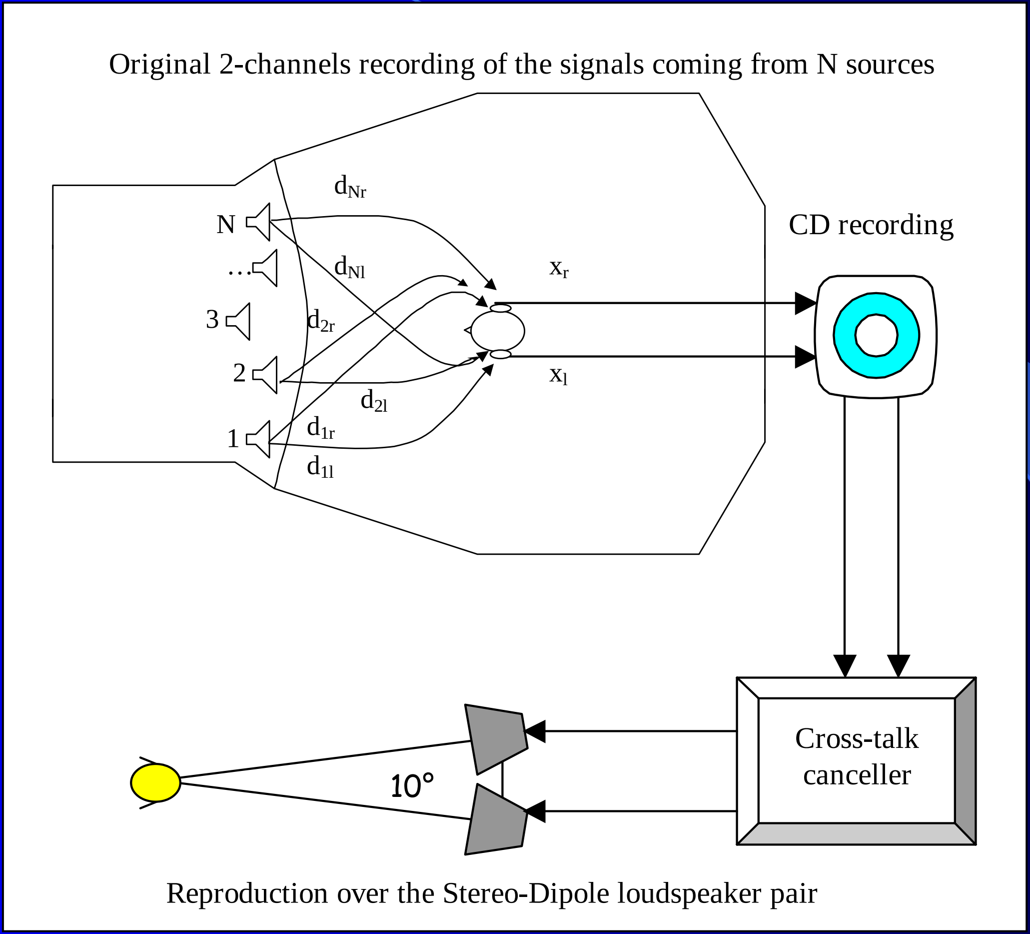
(1) Industrial Engineering Dept., University of Parma, Via delle Scienze 181/A Parma, 43100 ITALY – **HTTP://pcfarina.eng.unipr.it**

(2) Ambiophonics Institute, 4 Piermont Road, Rockleigh, New Jersey 07647, USA **HTTP://www.ambiophonics.org**

The Ambiophonics method

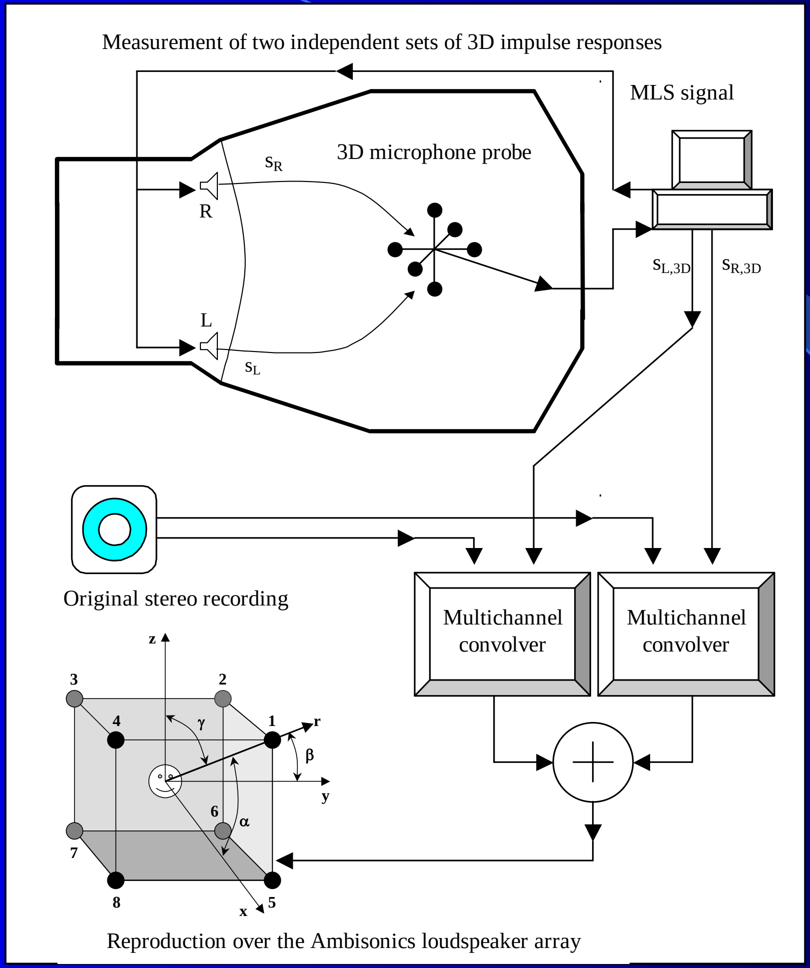
- Ambiophonics is an hybrid method for creating a realistic spatial reproduction of staged music, starting from two-channel recordings, but extensible to various kinds of microphonic arrangements up to discrete multichannel
- The system is based on two indipendently designed groups of loudspeakers: a Stereo Dipole, responsible for the reproduction of the direct sound and early reflections coming from the stage, asnd a surround periphonic array, driven by real-time convolution with room impulse responses

The Stereo Dipole



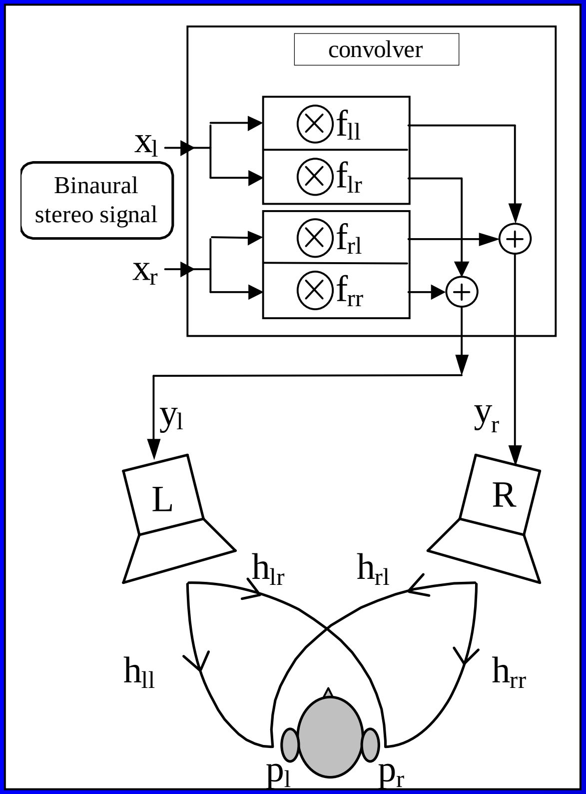
- The cross-talk cancellation allows for the replica of the recorded signals at the ears of the listener

The surround convolution



- The cross-talk cancellation allows for the replica of the recorded signals at the ears of the listener

Design of cross-talk canceling filters



- First, a binaural measurement is made in front of the Stereo Dipole loudspeakers
- Then, the cross-talk cancelling filters are computed, so that their convolution with the measured impulse responses reduces to the identity matrix

Theory of cross-talk canceling filters

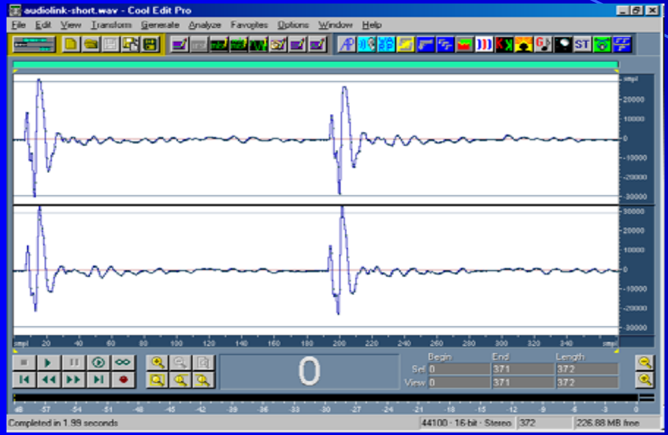
$$\begin{cases} f_{ll} = (h_{rr}) \otimes InvDen \\ f_{lr} = (-h_{lr}) \otimes InvDen \\ f_{rl} = (-h_{rl}) \otimes InvDen \\ f_{rr} = (h_{ll}) \otimes InvDen \\ InvDen = InvFilter(h_{ll} \otimes h_{rr} - h_{lr} \otimes h_{rl}) \end{cases}$$

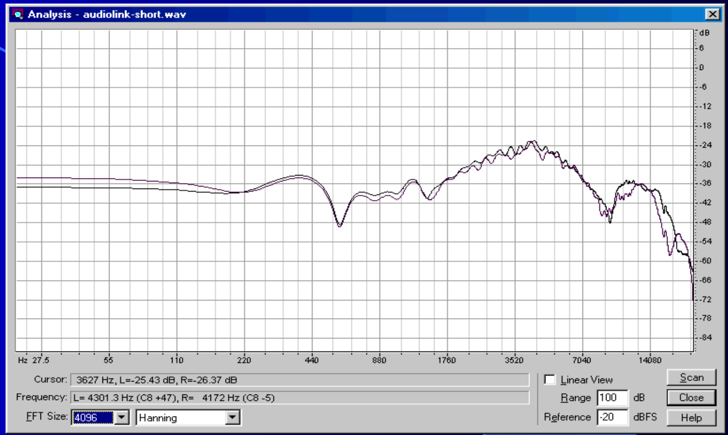
$$C(\omega) = FFT(h_{ll}) \cdot FFT(h_{rr}) - FFT(h_{lr}) \cdot FFT(h_{rl})$$

$$InvDen(\omega) = \frac{Conj[C(\omega)]}{Conj[C(\omega)] \cdot C(\omega) + \varepsilon(\omega)}$$

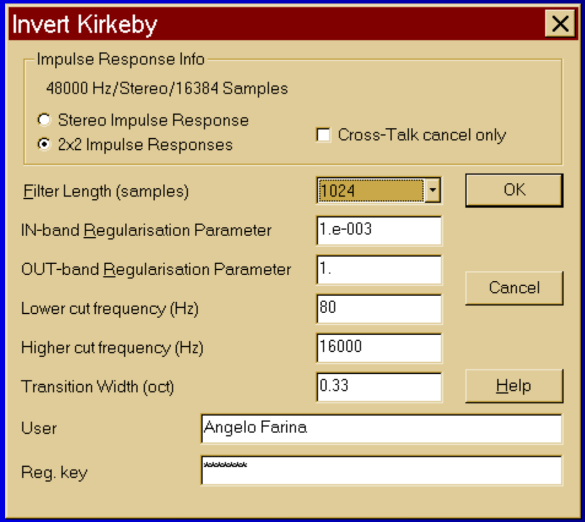
- The regularization parameter, ε, has to be adjusted by trials

Example



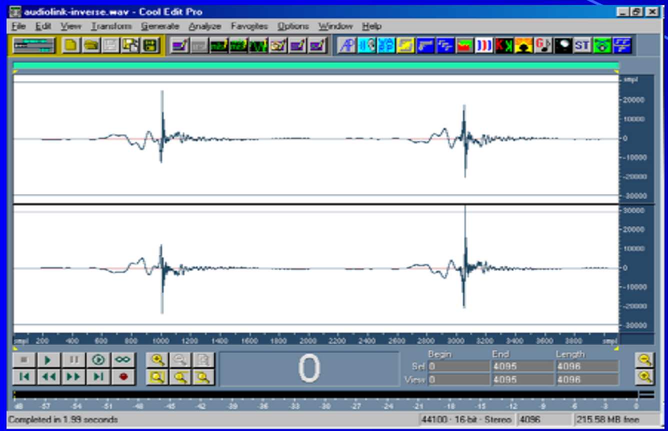


Measured impulse responses h

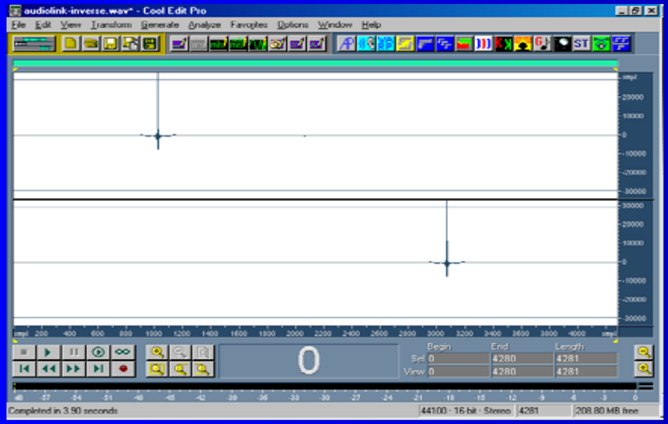


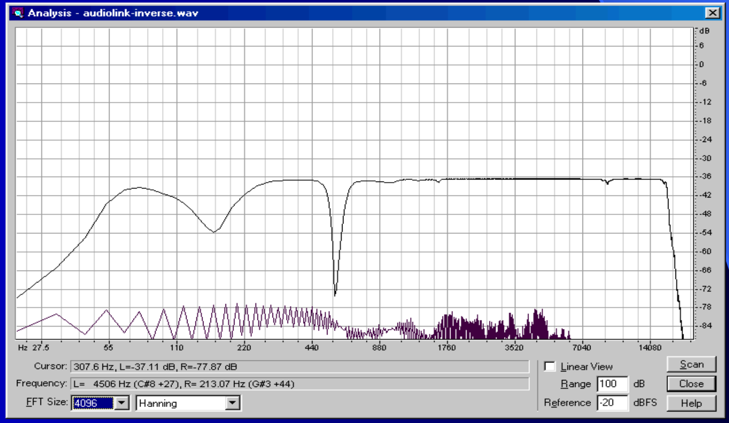
Plugin for CoolEdit which computes the inverse filters f

Example



Computed long-FIR inverse filters f

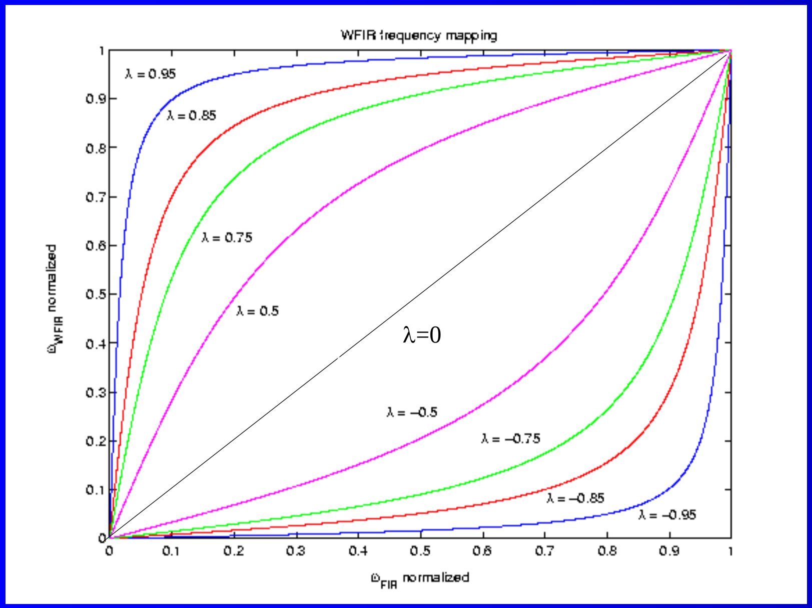




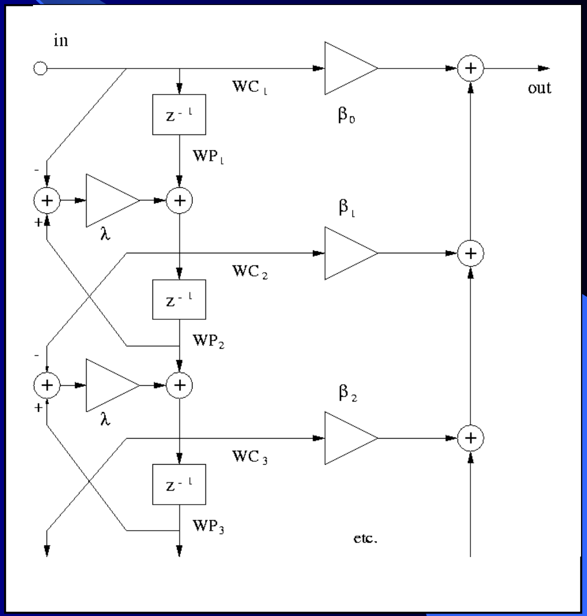
Verification of the cross-talk cancellation

Warped FIR cross-talk cancellation

- Todays DSP boards are not powerful enough for convolving long inverse FIR filters
- Warping can be used for concentrating the computing power in the frequency range where it is most needed



$$z = A_\lambda(\zeta) = \frac{\zeta + \lambda}{1 + \zeta \cdot \lambda}$$



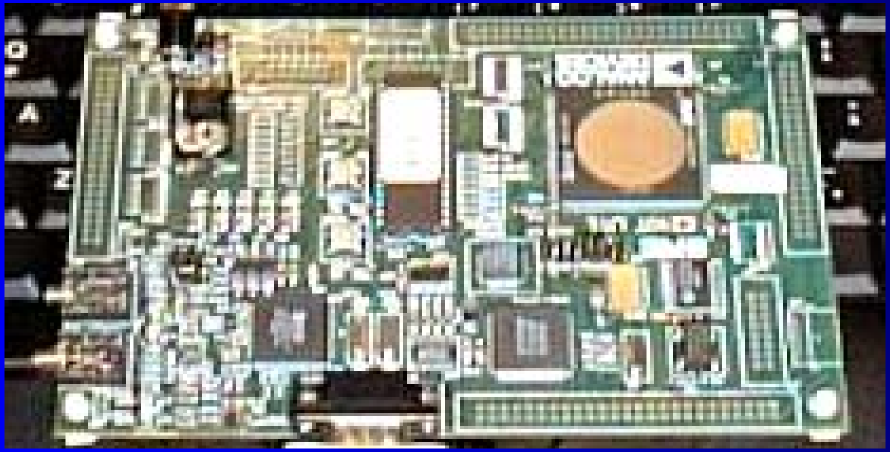
As the processing network is more involved than a traditional FIR, the number of taps which can be used is lower. Nevertheless, the perceived performances, on a given DSP board, are usually better than with a normal FIR

Warped FIR implementation on a SHARC

- The WFIR structure was coded in assembly on the AD21061 and on the AD21065L processors: here the assembly code of the main cycle is shown:

```
LCNTR=Wfilter_taps-1 , DO wmac_rr UNTIL LCE;
F12=F2*F4, F9=dm(I5,M7), F4=pm(I9,M8);
F10=F2*F5, F8=F8+F12, F9=dm(I5,M6);
F1=F9-F10, F9=dm(I5,0);
F10=F1*F7, dm(I5,M7)=F2;
wmac_rr:F2=F9+F10;
```

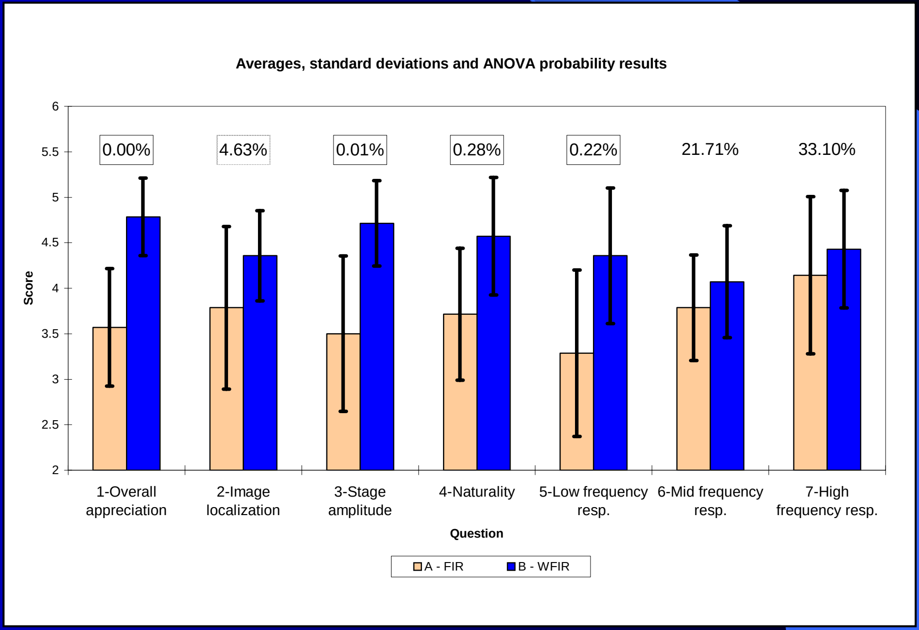
It takes 5 lines (CPU cycles), whilst the normal FIR is just one line – thus a SHARC can run a set of 4 FIRs of more than 200 taps each, but only 42 taps with WFIR.



Subjective blind comparison: FIR vs. WFIR

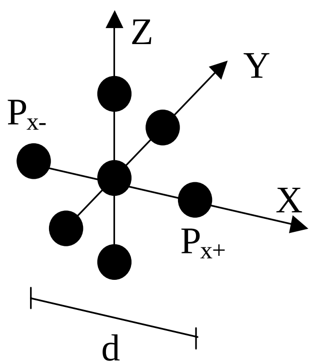
- 14 normal-hearing subjects (6 females, 8 males)
- Two sound samples: binaural recording of natural sounds and a piece of pop music (Elton John)
- 5-levels scale (insufficient, mediocre, sufficient, fair, good)
- The listener was free to switch at will between the two processing algorithms, denoted simply as A and B
- Classic ANOVA analysis of the subjective response

Question	Avg. A	Avg. B	Anova's F factor	Prob.
Overall appreciation	3.57	4.79	**34.47**	0.00%
Image localization	3.79	4.36	**4.38**	4.63%
Stage width	3.50	4.71	**21.72**	0.01%
Naturality	3.71	4.57	**10.88**	0.28%
Low frequency resp.	3.29	4.36	**11.56**	0.22%
Mid frequency resp.	3.79	4.07	1.60	21.7%
Hi frequency resp.	4.14	4.43	0.98	33.1%



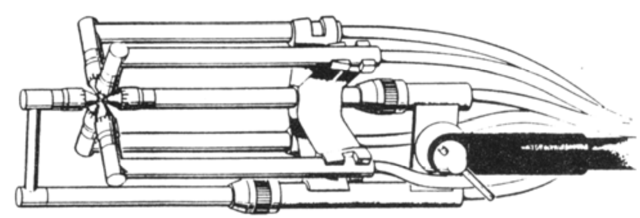
"Virtual Ambisonics" surround

- Measurement of 3D (B-format) impulse responses in theatres, with two source positions on the stage
- The IRs are processed, deriving the responses of several directive microphones
- Each soundtrack of the original stereo recording is convolved with the corresponding IR
- For each loudspeaker, the results of the two convolutions are mixed



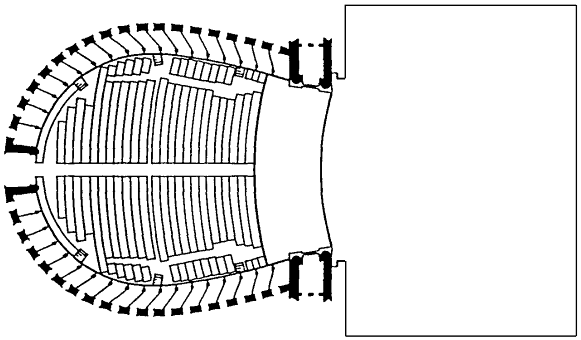
7 spaced omni

Soundfield MK-V

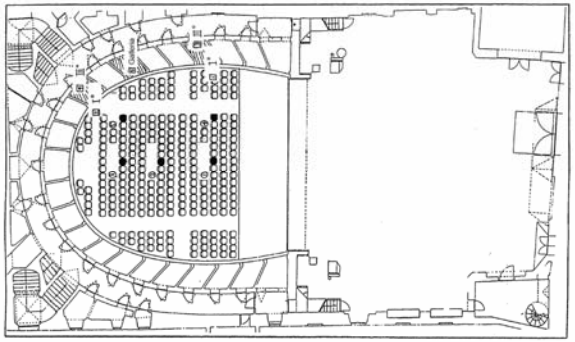


B&K WA0447

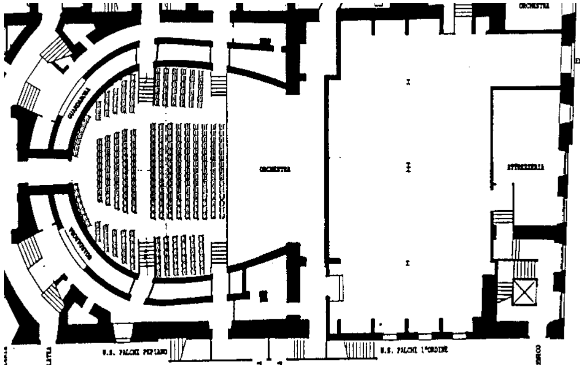
Measurements in 3 Italian theatres



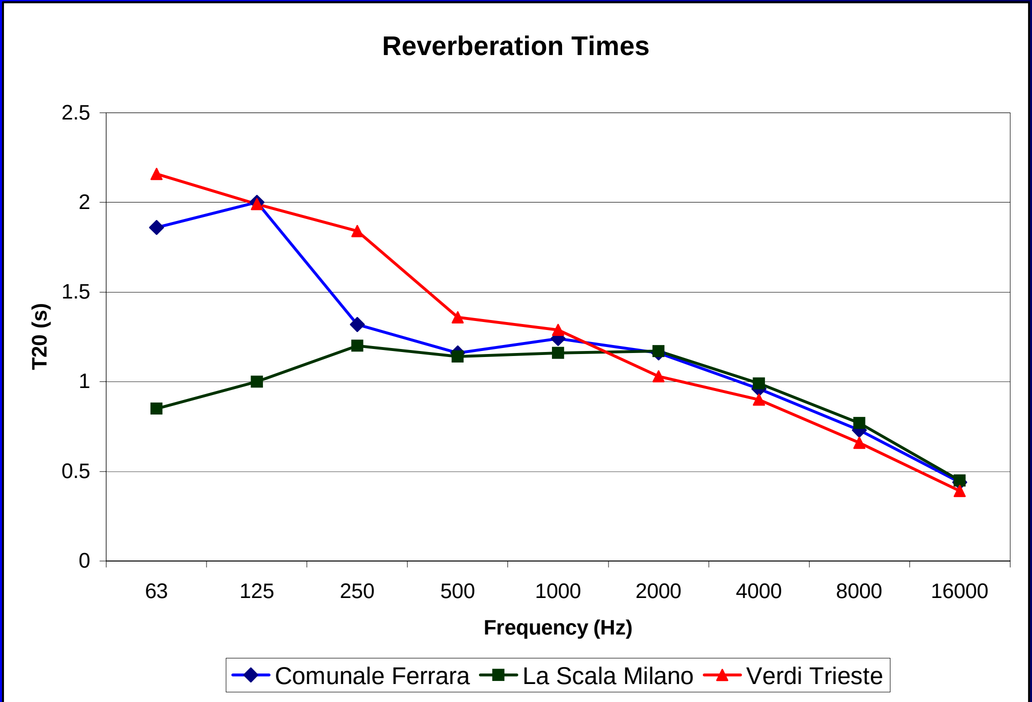
La Scala, Milan



T. Comunale, Ferrara



T. Verdi, Trieste



Synthesis of directive microphones

- The WXYZ channels of a B-format IR can be processed, extracting a single (mono) response of a virtual microphone pointing along a given versor ***r*** *(r_x, r_y, r_z)*:

$$V(\vec{r}) = \frac{1}{2} \cdot \left[(2 - D) \cdot W + D \cdot \left(r_x \cdot X + r_y \cdot Y + r_z \cdot Z \right) \right]$$

The directivity factor D can assume the following values:

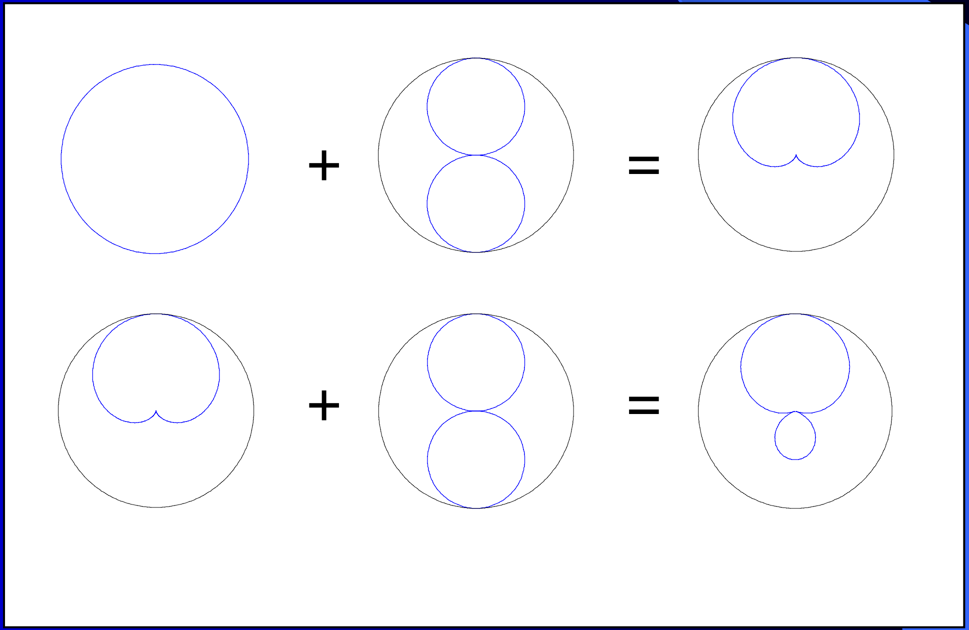
D=0.0 ➔ omnidirectional

D=0.5 ➔ subcardioid

D=1.0 ➔ cardioid

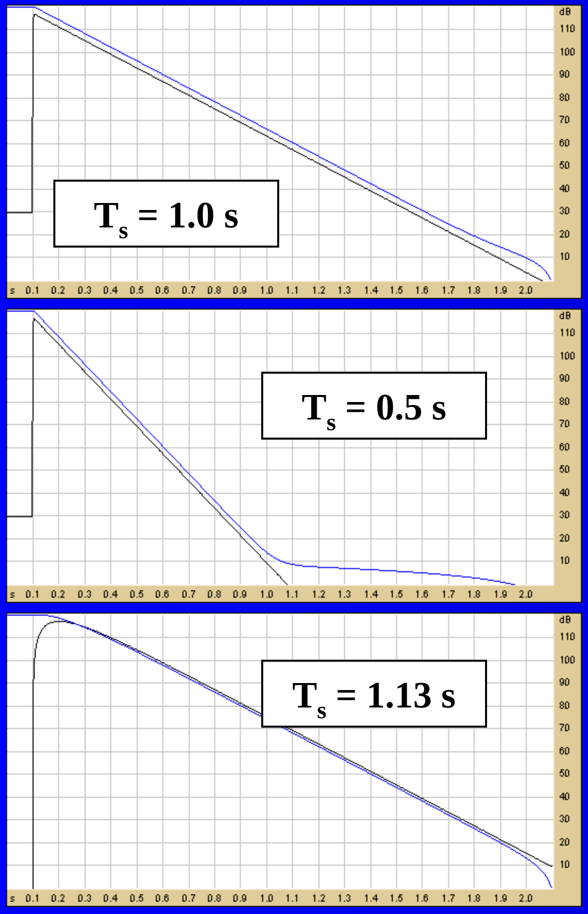
D=1.5 ➔ hypercardioid

D=2.0 ➔ figure-of-eight

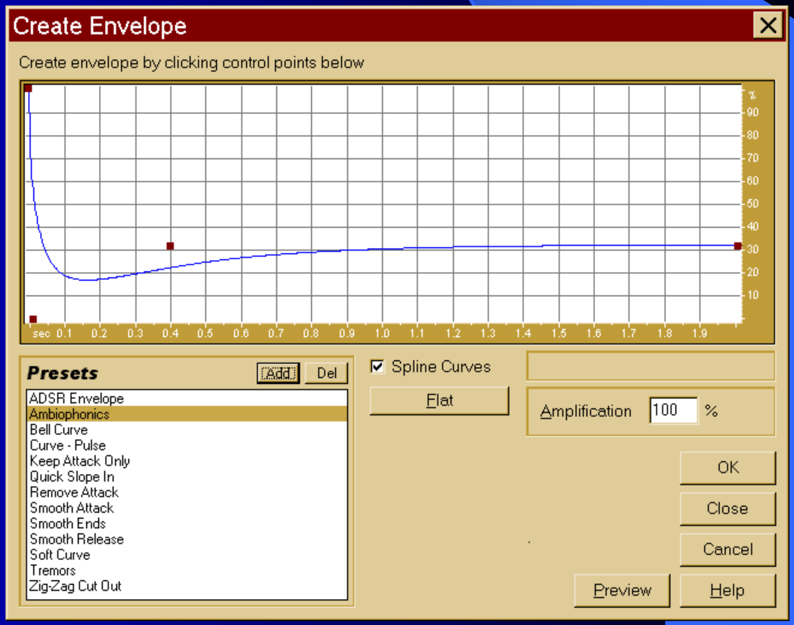


The Double-reverberation problem

- When an impulse response is reproduced in another reverberant space, the resulting reverberant tail is the convolution of the two reverberant tails



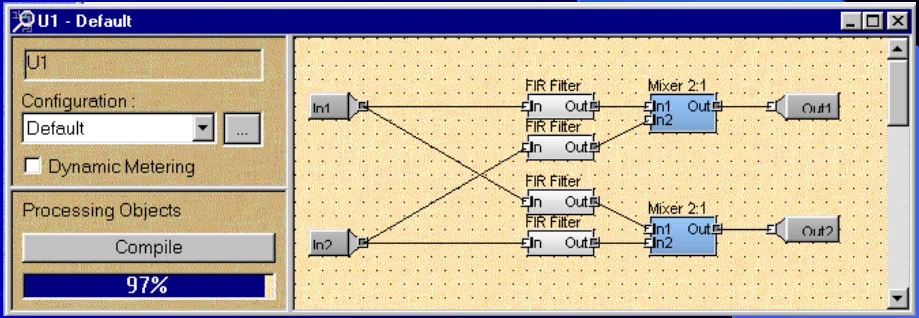
The problem can be compensated for, by manual editing the first impulse response, applying (with CoolEdit) a time.varying amplitude shaping:



Hardware implementation

- A complete Ambiophonics system can be implemented, nowadays, coupling a general-purpose DSP unit (cross-talk cancellation) and convolution-based reverberators

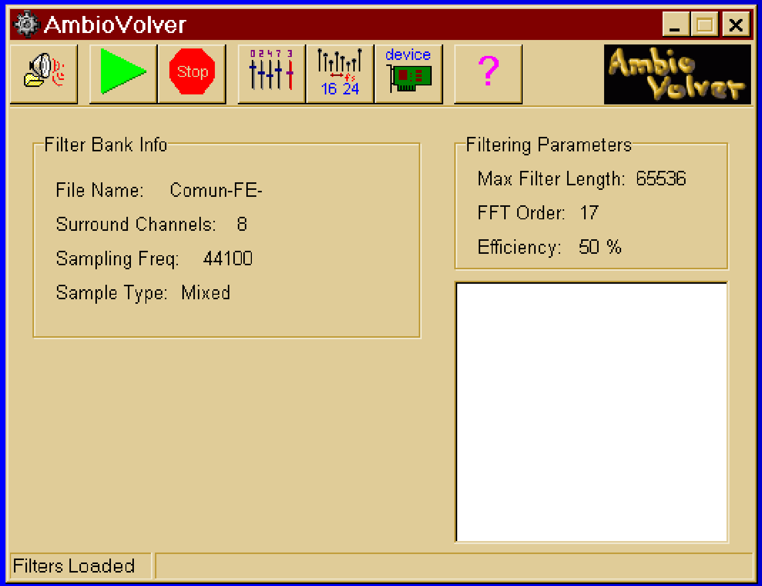
Here the cross-talk cancellation network implemented on the Soundweb is shown:



The limit of this system is in the number of coefficients of both the cross-talk canceler and the room convolver

Software implementation

- The preferred implementation is by means of a simple software convolver and a cheap, modern PC. Two solutions are currently available:



Ambiovolver is a program running on any Windows PC equipped with a multichannel soundboard

BruteFIR is an highly optimized Linux program running on a customized, noiseless computer with a digital RME audio interface

Both systems are being demonstrated in room 22

Latency vs. performance

- The software implementation is based on frequency-domain convolution (overlap-and-save), which inherently introduces some latency.
- Furthermore, the audio stream I/O on a PC is always buffered, so an intrinisic latency is caused by the buffer size
- BruteFIR distinguishes himself from other convolvers by the fact that it implements **partitioned convolution**: the impulse response is subdivided in many segments of equal length, and this reduces the latency to twice the length of a segment, instead of twice the length of the whole IR.
- On modern CPUs, the partitioned convolution is more efficient than traditional unpartitioned overlap-and-save, with a reduction of CPU load of 20-50%, and can reduce the overall latency to less than 100 ms.
- Very efficient FFT implementations are freely available (Intel NSP, FTTW), and thus the computing power of a PC is enough for real-time convolution of **20 IRs, at 44.1 KHz, 32 bits, each being 65,536 points long**. The demonstration machine, installed in room 22, is an old Pentium-II 400 MHz.

Subjective comparative experiment

- 9 normal-hearing subjects (males)
- Three sound samples:
- Simple ranking test between three systems: Stereo-Dipole, Virtual Ambisonics, complete Ambiophonics
- Each listener can switch freely among the three systems during the playback

Music piece	Theatre	Cross-talk filters
Mozart, Te Deum K141, Sennheiser MKE2002 ("Mozart Sacro", n. 1)	La Scala	Binaural
Buxtehude KFM -6 (Ambiopole demo 1, n.13)	Teatro Comunale	Sphere
Mozart, Overture "Le nozze di Figaro", bars 1 -50, ORTF (Denon PG 6006, n. 37)	Teatro Verdi	ORTF

Results

Method	Stereo Dipole	Virtual Ambisonics	Ambiophonics
Avg. Score	1.99	2.77	1.24

Conclusions

- Ambiophonics revealed to give significant advantages over the two surround systems which constitutes it.
- It recreates a realistic virtual acoustic space by means of convolution with proper digital filters
- The computational power required can be obtained cheaply by means of a modern PC
- The system can be configured for different number and position of the loudspeakers
- The “sweet spot” can easily accomodate three persons, and also far from this area, the overall acoustic impression remains that of being in a concert hall.

Internet Links

- The CoolEdit plugins, employed for measuring and processing the impulse responses, can be downloaded freely from **HTTP://www.ramsete.com/aurora**

- The sets of impulse responses, and the sound samples employed for the subjective tests are available for download at **HTTP://pcangelo.eng.unipr.it/public/AES19**

- The programs for computing a B-format IR from 7 spaced omnis, and for deriving the virtual directive microphones, can be freely downloaded from **HTTP://pcangelo.eng.unipr.it/public/B-format/software**

- The BruteFIR convolver for Linux can be freely downloaded, with its source code, from **HTTP://www.ludd.luth.se/~torger/brutefir.html**