LISTENING TESTS PERFORMED INSIDE A VIRTUAL ROOM ACOUSTIC SIMULATOR

A. Farina, P. Martignon, A. Azzali, A. Capra
Industrial Engineering Department – University of Parma – Italy

Abstract: Auralization is a method for recreating the aural impression of a room. Usually it is implemented by convolution, employing dry (anechoic) music as the source signal, which is passed trough a very long FIR filter, loaded with coefficients taken from the impulse response of the room to be simulated. This impulse response can be either measured or synthetic (obtained by a simulation performed through a room acoustics numerical solver, typically based, nowadays, on the beam-tracing computational scheme).
A software and hardware tool for real time processing and routing of the signal, allowing to switch among four different two-channels reproduction systems, was set up inside a listening room in the "Casa della Musica", in Parma. The reproduction systems available are stereo dipole, double stereo dipole, "normal" stereo (ORTF) and headphones. The whole system is described, paying attention to room response, and principles and methods of the systems implemented are explained as well. A little discussion about pro and cons of different reproduction systems is proposed, to be checked in the future based on results of the currently going on comparative subjective tests.

1. INTRODUCTION

Many subjective tests have been carried out in the past by us and others to investigate correlation between subjective and objective parameters [1], mostly using headphones and more rarely other methods.

So, we first describe a set of preliminary listening tests performed during the spring 2004, employing a headphones-based binaural system, in which we did compare the acoustic behavior of 5 theatres, employing 6 different sound samples, and a questionnaire to be filled during the listening.

Here we describe the system employed for performing the listening tests, the creation of suitable inverse filters for making it perfectly "transparent", and finally the software tool employed for controlling the playback of the sound samples and simultaneously for collecting the questionnaires.

A basic statistical analysis of the results did show quite bad correlation between objective parameters and subjective responses.

However, it resulted that the headphones system has some inherent weak points, so we started an investigation about "alternative" reproduction systems for stereo (2-channels) soundtracks, capable of conveying a better spatial impression and a more faithful enveloping.

A special listening room was set up, equipped with three additional systems:

- Stereo Dipole
- Dual Stereo Dipole
- Normal Stereo

In particular this multi-listening setup for two-channels systems arises from a research project, which sees the partnership of UNIPR and University of Sidney; it consists of subjective tests for investigating how much different two-channel systems can reproduce the perception of basic spatial characteristics of the virtual acoustic space, like source distance and room size. The results of these tests are yet to be completed and published, here we just describe and discuss the system implemented and the underlying technology.

As in the case of the preliminary test, here we describe the systems, the creation of suitable inverse filters for making them "transparent", and the software tool employed for controlling the playback of the sound samples and for collecting the questionnaires.

Finally, a quick foreword of the forthcoming research is given, which will also include Ambisonics-based multichannel systems, planned to be implemented in our listening rooms during the next months.

2. PRELIMINARY SUBJECTIVE LISTENING TEST BY HEADPHONES

Aim of this preliminary session of listening test is the correlation between objective and subjective parameters. The knowledge of this relationship is the base for interventions on existent theatres or design of new ones.

2.1. Preparation

This preliminary test consisted in listening to several anechoic musical tracks auralized with the impulse response of some Italian theatres measured with a binaural microphone, and compiling at the same time a questionnaire.

The rooms object of the tests were: Auditorium Paganini (Parma), Auditorium Sala700 (Roma), Teatro Valli (Reggio Emilia), Teatro Regio (Parma), Teatro Olimpico (Vicenza). For the auralization, we did choose binaural impulse responses recorded with the Neumann KU-100 dummy head in the central position of every room between the 5th and the 6th row of seats.

The tracks used for the listening are divided in two categories: orchestrals and vocals. We choose this differentiation for the different function of theatres and auditorium.

There are three purely-orchestral tracks:

- Mozart, Overture of "*Le nozze di Figaro*";
- Strauss, "Pizzicate Polka";
- Verdi, Prelude at first act of "*La Traviata*"

and three vocal tracks:

- "*My funny Valentine*";
- Mozart, aria from *"Cosi' fan tutte"* vocal and piano;
- Tosti, "Non t'amo più"

The orchestral tracks were auralized making a convolution between binaural tracks recorded using omnidirectional speakers placed on the left and on the right of the stage. The pieces of Tosti and Mozart are made of a Left track, containing a piano, and a Right track containing the voice. We convolved the Left one with the recording of an omnidirectional source placed on the left of the stage and the right one with a directive source placed in the centre of the stage in order to recreate the concert configuration (piano on a side and the singer in the centre).

After this, the two resulting stereo tracks are mixed (left summed with left, right summed with right) and reproduced over headphones (Sennheiser HD580).

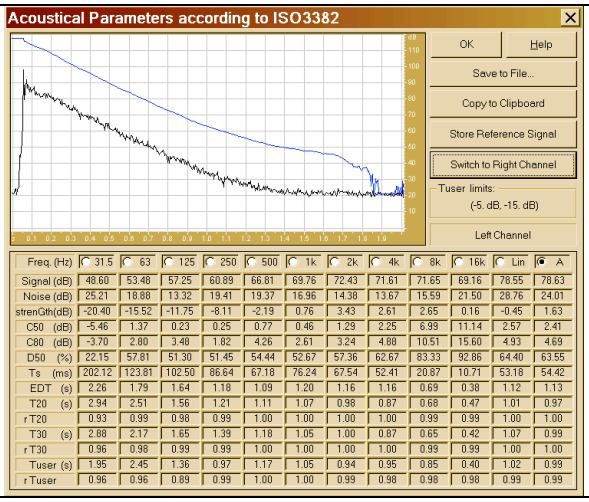
The playback system did include also a small subwoofer (Audio-Pro), set up with a cross-over frequency of approximately 60 Hz.

The "transparency" of the recording/reproduction chain is ensured by convolving the signal which feed the playback system with a pair of inverse-filters. They are designed based on a measurement of the transfer function of the playback system, which was performed placing the headphones over the same dummy head employed for the binaural impulse response measurements. This transfer function was numerically inverted making use of the Nelson-Kirkeby-Farina regularization method.

2.2. The theatres

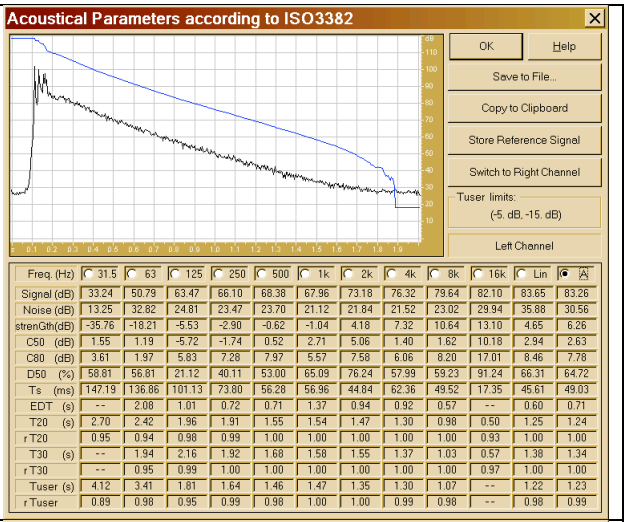
Here we provide a photo and a plot of the impulse response of the 5 theatres:

Teatro Regio, Parma

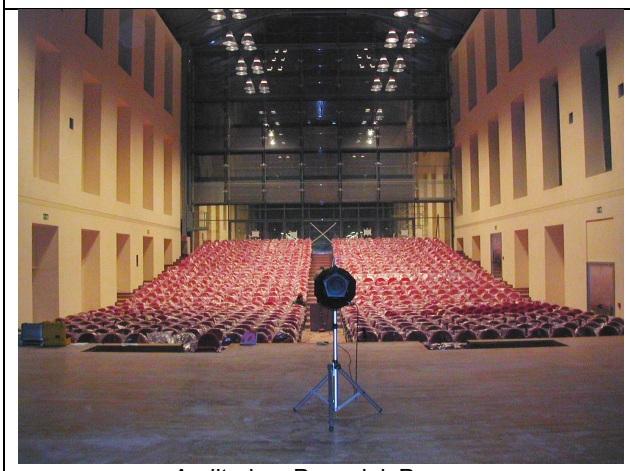


Freq. (Hz)	31.5	63	125	250	500	1k	2k	4k	8k	16k	Lin	A
Signal (dB)	48.60	53.48	57.25	60.89	66.81	69.76	72.43	71.61	71.65	69.16	78.55	78.63
Noise (dB)	25.21	18.88	13.32	19.41	19.37	16.96	14.38	13.67	15.59	21.50	28.76	24.01
strenGth(dB)	-20.40	-15.52	-11.75	-8.11	-2.19	0.76	3.43	2.61	2.65	0.16	-0.45	1.63
C50 (dB)	-5.46	1.37	0.23	0.25	0.77	0.46	1.29	2.25	6.99	11.14	2.57	2.41
C80 (dB)	-3.70	2.80	3.48	1.82	4.26	2.61	3.24	4.88	10.51	15.60	4.93	4.69
D50 (%)	22.15	57.81	51.30	51.45	54.44	52.67	57.36	62.67	83.33	92.86	64.40	63.55
Ts (ms)	202.12	123.81	102.50	86.64	67.18	76.24	67.54	52.41	20.87	10.71	53.18	54.42
EDT (s)	2.26	1.79	1.64	1.18	1.09	1.20	1.16	1.16	0.69	0.38	1.12	1.13
T20 (s)	2.94	2.51	1.56	1.21	1.11	1.07	0.98	0.87	0.68	0.47	1.01	0.97
r T20	0.93	0.99	0.98	0.99	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00
T30 (s)	2.88	2.17	1.65	1.39	1.18	1.05	1.00	0.87	0.65	0.42	1.07	0.99
r T30	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00
Tuser (s)	1.95	2.45	1.36	0.97	1.17	1.05	0.94	0.95	0.85	0.40	1.02	0.99
r Tuser	0.96	0.96	0.89	0.99	1.00	1.00	0.99	0.98	0.98	0.98	0.99	0.99

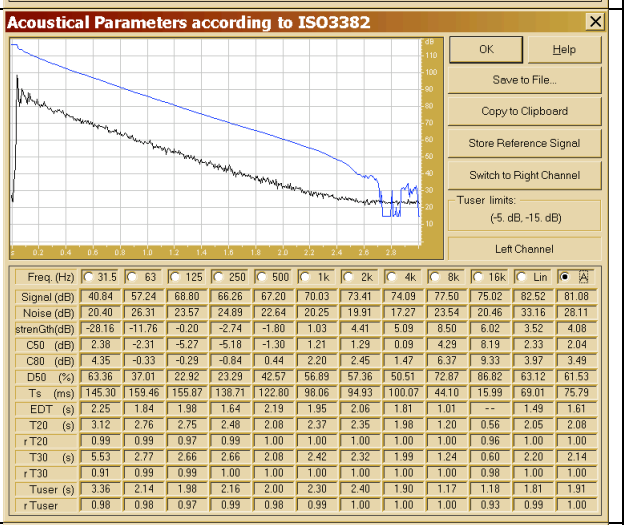
Teatro Valli, Reggio Emilia



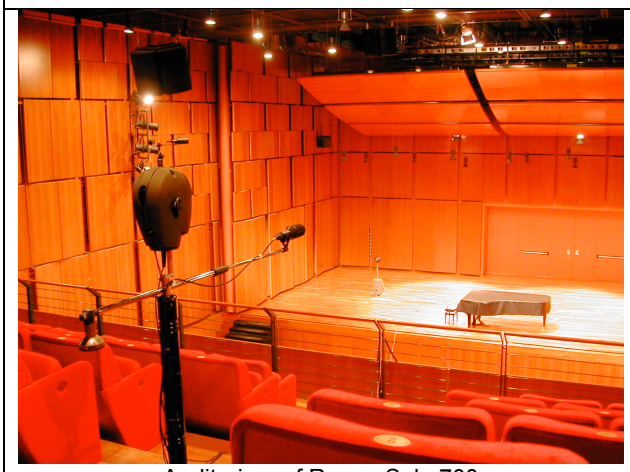
Freq. (Hz)	31.5	63	125	250	500	1k	2k	4k	8k	16k	Lin	A
Signal (dB)	33.24	50.79	63.47	66.10	68.38	67.96	73.18	76.32	79.64	82.10	83.65	83.26
Noise (dB)	13.25	32.82	24.81	23.47	23.70	21.12	21.84	21.52	23.02	29.94	35.88	30.56
strenGth(dB)	-35.76	-18.21	-5.53	-2.90	-0.62	-1.04	4.18	7.32	10.64	13.10	4.65	6.26
C50 (dB)	1.55	1.19	-5.72	-1.74	0.52	2.71	5.06	1.40	1.62	10.18	2.94	2.63
C80 (dB)	3.61	1.97	5.83	7.28	7.97	5.57	7.58	6.06	8.20	17.01	8.46	7.78
D50 (%)	58.81	56.81	21.12	40.11	53.00	65.09	76.24	57.99	59.23	91.24	66.31	64.72
Ts (ms)	147.19	136.86	101.13	73.80	56.28	56.96	44.84	62.36	49.52	17.35	45.61	49.03
EDT (s)	--	2.08	1.01	0.72	0.71	1.37	0.94	0.92	0.57	--	0.60	0.71
T20 (s)	2.70	2.42	1.96	1.91	1.55	1.54	1.47	1.30	0.98	0.50	1.25	1.24
rT20	0.95	0.94	0.98	0.99	1.00	1.00	1.00	1.00	1.00	0.93	1.00	1.00
T30 (s)	--	1.94	2.16	1.92	1.68	1.58	1.55	1.37	1.03	0.57	1.38	1.34
rT30	--	0.95	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	1.00
Tuser (s)	4.12	3.41	1.81	1.64	1.46	1.47	1.35	1.30	1.07	--	1.22	1.23
rTuser	0.89	0.98	0.95	0.99	0.98	1.00	1.00	0.99	0.98	--	0.98	0.99



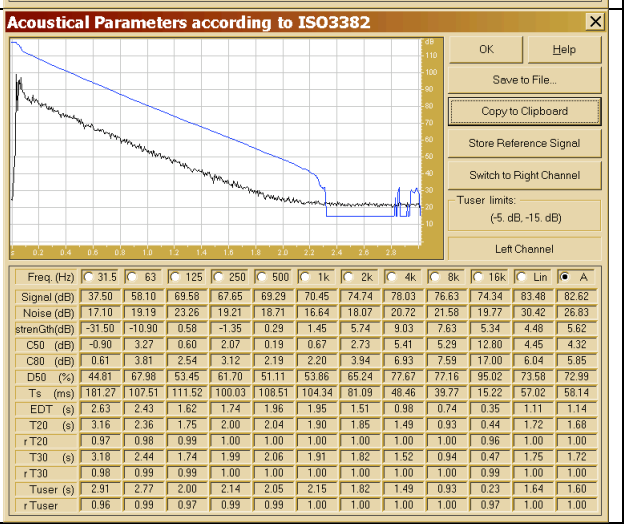
Auditorium Paganini, Parma



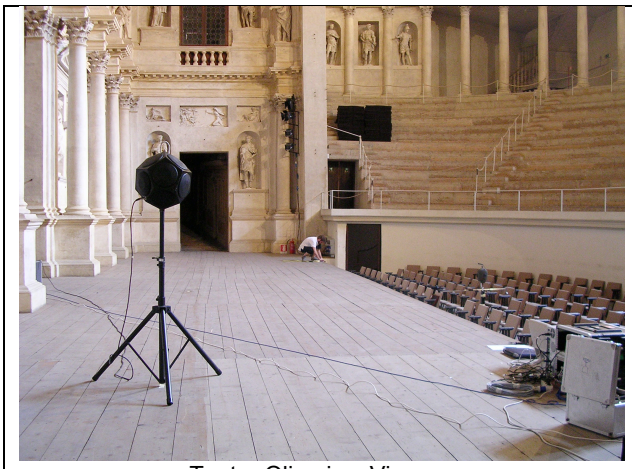
Freq. (Hz)	31.5	63	125	250	500	1k	2k	4k	8k	16k	Lin	A
Signal (dB)	40.84	57.24	68.80	66.26	67.20	70.03	73.41	74.09	77.50	75.02	82.52	81.08
Noise (dB)	20.40	26.31	23.57	24.89	22.64	20.25	19.91	17.27	23.54	20.46	33.16	28.11
strenGth(dB)	-28.16	-11.76	-0.20	-2.74	-1.80	1.03	4.41	5.09	8.50	6.02	3.52	4.08
C50 (dB)	2.38	-2.31	-5.27	-5.18	-1.30	1.21	1.29	0.09	4.29	8.19	2.33	2.04
C80 (dB)	4.35	-0.33	-0.29	-0.84	0.44	2.20	2.45	1.47	6.37	9.33	3.97	3.49
D50 (%)	63.36	37.01	22.92	23.29	42.57	56.89	57.36	50.51	72.87	86.82	63.12	61.53
Ts (ms)	145.30	159.46	155.87	138.71	122.80	98.06	94.93	100.07	44.10	15.99	69.01	75.79
EDT (s)	2.25	1.84	1.98	1.64	2.19	1.95	2.06	1.81	1.01	--	1.49	1.61
T20 (s)	3.12	2.76	2.75	2.48	2.08	2.37	2.35	1.98	1.20	0.56	2.05	2.08
rT20	0.99	0.99	0.97	0.99	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00
T30 (s)	5.53	2.77	2.66	2.66	2.08	2.42	2.32	1.99	1.24	0.60	2.20	2.14
rT30	0.91	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00
Tuser (s)	3.36	2.14	1.98	2.16	2.00	2.30	2.40	1.90	1.17	1.18	1.81	1.91
rTuser	0.98	0.98	0.97	0.99	0.98	0.99	1.00	1.00	1.00	0.93	0.99	1.00

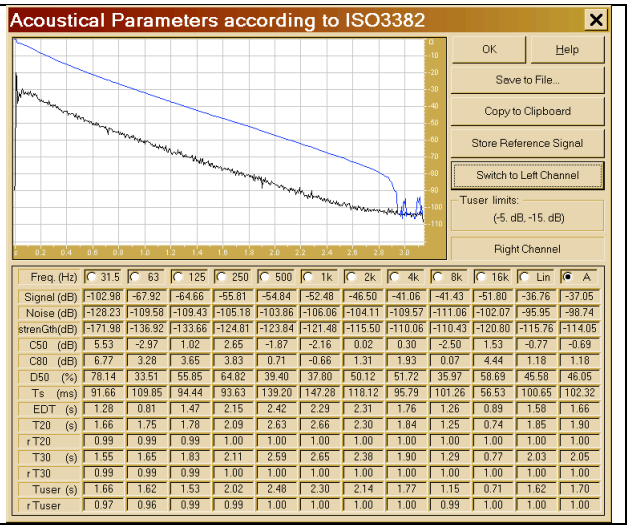


Auditorium of Rome, Sala 700



Freq. (Hz)	31.5	63	125	250	500	1k	2k	4k	8k	16k	Lin	A
Signal (dB)	37.50	58.10	69.58	67.65	69.29	70.45	74.74	78.03	76.63	74.34	83.48	82.62
Noise (dB)	17.10	19.19	23.26	19.21	18.71	16.64	18.07	20.72	21.58	19.77	30.42	26.83
strenGth(dB)	-31.50	-10.90	0.58	-1.35	0.29	1.45	5.74	9.03	7.63	5.34	4.48	5.62
C50 (dB)	-0.90	3.27	0.60	2.07	0.19	0.67	2.73	5.41	5.29	12.80	4.45	4.32
C80 (dB)	0.61	3.81	2.54	3.12	2.19	2.20	3.94	6.93	7.59	17.00	6.04	5.85
D50 (%)	44.81	67.98	53.45	61.70	51.11	53.86	65.24	77.67	77.16	95.02	73.58	72.99
Ts (ms)	181.27	107.51	111.52	100.03	108.51	104.34	81.09	48.46	39.77	15.22	57.02	58.14
EDT (s)	2.63	2.43	1.62	1.74	1.96	1.95	1.51	0.98	0.74	0.35	1.11	1.14
T20 (s)	3.16	2.36	1.75	2.00	2.04	1.90	1.85	1.49	0.93	0.44	1.72	1.68
rT20	0.97	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00
T30 (s)	3.18	2.44	1.74	1.99	2.06	1.91	1.82	1.52	0.94	0.47	1.75	1.72
rT30	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
Tuser (s)	2.91	2.77	2.00	2.14	2.05	2.15	1.82	1.49	0.93	0.23	1.64	1.60
rTuser	0.96	0.99	0.97	0.99	0.99	1.00	1.00	1.00	1.00	0.97	1.00	1.00


Teatro Olimpico, Vicenza



Freq. (Hz)	31.5	63	125	250	500	1k	2k	4k	8k	16k	Lin	A
Signal (dB)	-102.98	-67.92	-64.66	-55.81	-54.84	-52.48	-46.50	-41.06	-41.43	-51.80	-36.76	-37.05
Noise (dB)	-128.23	-109.58	-109.43	-105.18	-103.86	-106.06	-104.11	-109.57	-111.06	-102.07	-95.95	-98.74
strenGth(dB)	-171.98	-136.92	-133.66	-124.81	-123.84	-121.48	-115.50	-110.06	-110.43	-120.80	-115.76	-114.05
C50 (dB)	5.53	-2.97	1.02	2.65	-1.87	-2.16	0.02	0.30	-2.50	1.53	-0.77	-0.69
C80 (dB)	6.77	3.28	3.65	3.83	0.71	-0.66	1.31	1.93	0.07	4.44	1.18	1.18
D50 (%)	78.14	33.51	55.85	64.82	39.40	37.80	50.12	51.72	35.97	58.69	45.58	46.05
Ts (ms)	91.66	109.85	94.44	93.63	139.20	147.28	118.12	95.79	101.26	56.53	100.65	102.32
EDT (s)	1.28	0.81	1.47	2.15	2.42	2.29	2.31	1.76	1.26	0.89	1.58	1.66
T20 (s)	1.66	1.75	1.78	2.09	2.63	2.66	2.30	1.84	1.25	0.74	1.85	1.90
r T20	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
T30 (s)	1.55	1.65	1.83	2.11	2.59	2.65	2.38	1.90	1.29	0.77	2.03	2.05
r T30	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Tuser (s)	1.66	1.62	1.53	2.02	2.48	2.30	2.14	1.77	1.15	0.71	1.62	1.70
r Tuser	0.97	0.96	0.99	0.99	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00

Fig. 1 – The 5 Italian theatres

2.3. Objective parameters

Thanks to the "acoustical parameter" plugin of the Aurora software package [2] we did compute the most frequently used acoustical parameters, according to the ISO 3382/1997 standard. For each of them, a short description is given here.

2.3.1. T_{15}, T_{30}

Reverberation time calculated from the decay range between -5 and - 20 dB (T_{15}) and between -5 and - 35 dB (T_{30}) on the integrated Schroeder curve, in seconds.

Schroeder [3] found that the reverberant decay can be described by a backward integration of the impulse response:

$$\left\langle p^2(t)\right\rangle = N\int_t^\infty h^2(\tau)d\tau \qquad (1)$$

Where: $\left\langle p^2(t)\right\rangle$ = average of a infinite number of decay

$h^2(\tau)$ = impulse response.

Eq.(1) can be written as:

$$\left\langle p^2(t)\right\rangle = N\left(\int_0^\infty h^2(\tau)d\tau - \int_0^t h^2(\tau)d\tau\right) \qquad (2)$$

Eq. (2) can be represented in a (p^2,τ) diagram, as shown in fig. 2

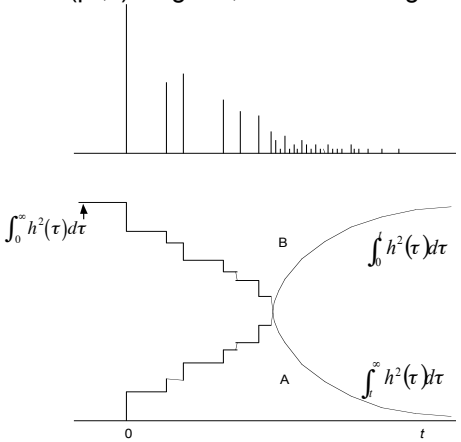


Fig. 2 - Schroeder plot as represented by Eq. (2)

2.3.2. Early Decay Time (EDT)

Since Jordan [4] demonstrated that the subjective perception of reverberation in correlated more strongly with the initial decay of the reverberant tail, he suggested to calculated the reverberation time from the decay range between 0 and - 10 dB on the integrated Schroeder curve, in seconds.

2.3.3. Center Time t_s

It was defined by Kürer [5], as *Schwerpunktzeit*, in the following equation:

$$\tau_s = \frac{\int_0^\infty \tau h^2(\tau)\, d\tau}{\int_0^\infty h^2(\tau)\, d\tau} \qquad (3)$$

It is the first-order momentum of the squared pressure impulse response, expressed in milliseconds.

2.3.4. Inter-Aural Cross Correlation (IACC-early)

As suggested by Ando [6], it is the normalized correlation coefficient between the first 50 ms of the pressure impulse responses measured at the two ears of the binaural microphone.

From the definition of the cross-correlation function, given by:

$$\rho(\tau) = \frac{\lim_{T\to\infty}\left(\frac{1}{2T}\int_{-T}^{T} h_d(\tau)\cdot h_s(\tau+t)d\tau\right)}{\lim_{T\to\infty}\left(\frac{1}{2T}\sqrt{\int_{-T}^{T} h_d^2(\tau)d\tau\cdot\int_{-T}^{T} h_s^2(\tau+t)d\tau}\right)} \qquad (4)$$

and taking into account that the limitations of the integral are set to 80 ms (Early), the IACC is defined as the maximum value of Eq. (4), that is,

$$IACC = \left|\rho(\tau)\right|_{MAX} \qquad \text{Where} \quad \tau \le 1\,ms. \qquad (5)$$

2.3.5. Strength (G)

It is the difference between the measured sound pressure level, and that produced by the same omnidirectional source in a free field, at 10-m distance from its center, and is expressed in decibels. It was defined in ISO 3382, and expressed in the following equation

$$G = 10\log\frac{\int_0^\infty h^2(\tau)d\tau}{\int_0^\infty h^2{}_{10}(\tau)d\tau} \qquad (6)$$

2.3.6 Klarheitsmass or Clarity C_{80} and C_{50}

It is defined [7] by the equation

$$C = 10\log\frac{\int_0^{80ms} h^2(\tau)d\tau}{\int_{80ms}^\infty h^2(\tau)d\tau} \qquad (7)$$

When the clarity is related to the musical perception, as in this research, the time interval is limited to 80 ms, whereas if the clarity is related to speech, the time interval is set to 50 ms. Reichardt, Abdel Alim and Schmidt defined such an acoustic parameter in order to relate the “transparence” of the music to an energetic parameter.

2.3.7 Lateral Fraction LF

It is defined by the equation

$$LF = 10\log\frac{\int_{5ms}^{80ms} h_{\infty}^{2}(\tau)d\tau}{\int_{0ms}^{80ms} h_{o}^{2}(\tau)d\tau} \qquad (8)$$

In which h_∞ is the impulse response measured with a "velocity" microphone, pointing outside the left ear of the listener, and h_o is the normal omnidirectional impulse response. If a Soundfield microphone is employed (as in this case), the h_∞ is obtained by the channel labeled Y, and h_o is obtained by the channel labeled W, provided that this is amplified by 3 dB (as the soundfield microphone outputs an omnidirectional signal which has a gain reduced by 3 dB in comparison with the other three velocity channels XYZ).

2.3.8 Tonal Balance TB and Bass Ratio BR

These two parameters were defined by Beranek [8], and are actually NOT considered in the ISO3382 standard (together with ITDG). They are defined as ratios between the reverberation times T20 averaged over different frequency ranges:

$$TB = \frac{T_{125} + T_{250}}{T_{2000} + T_{4000}} \qquad (9)$$

$$BR = \frac{T_{125} + T_{250}}{T_{500} + T_{1000}} \qquad (10)$$

2.3.9 Comparison between the 5 theatres

A first comparison is obtained looking at the more traditional parameter, the reverberation time T20. Fig. 3 shows a comparative plot of the spectra of T20 for the 5 theatres employed in the preliminary test.

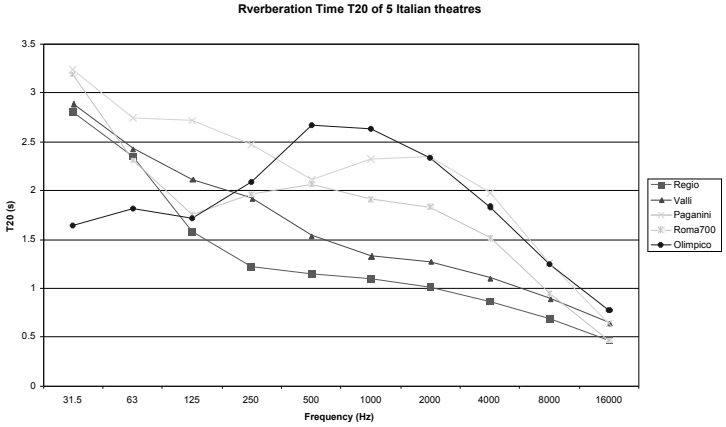


Fig. 3 - Reverberation time T20 of the theatres employed for the preliminary test

The following table reports the complete set of objective parameters, subsequently employed in the objective/subjective comparative analysis:

Param.	Regio	Valli	Paganini	Roma-700	Olimpico
C50 [dB]	1.82	7.48	-1.47	-2.45	0.03
C80 [dB]	4.93	9.28	0.81	0.83	1.20
D50 [%]	60	84	42	37	50
Ts [ms]	48	28	115	144	110
EDT [s]	1.08	1.26	2.09	1.98	2.43
T20 [s]	1.10	1.44	2.22	1.99	2.65
T30 [s]	1.11	1.55	2.24	1.99	2.64
LF	0.10	0.12	0.22	0.12	0.18
IACC (Early)	0.71	0.88	0.60	0.70	0.81
TB	1.47	1.70	1.20	1.11	0.91
BR	1.27	1.41	1.17	0.94	0.72

For the 9 ISO-3382 parameters, the average value between those in the 500 and 1000 Hz octave bands were taken. The last two parameters are already ratios between different frequency bands.

2.4 Listening tests

The questionnaire that we proposed is the fruit of a work made from Prof. A.Farina and Prof. L.Tronchin [1] in several years. From this work, a list of 9 couples of adjectives emerged, suitable for our tests: *"Pleasant-Unpleasant", "Round-Sharp", "Soft-Hard", "Diffuse-Localisable", Detached-Enveloping", "Dry-Reverberant", "Treble boosted-Treble reduced", "Bass boosted-Bass reduced", "Quiet-Loud"*.

In order to facilitate the test we used a software that permit the switch between the theatres in real time listening a track, as shown in *Figure 4*.

We used for the test 17 subjects: they were musicians, singers, professors of Music Accademy, audiophiles, music critics, musicologists.

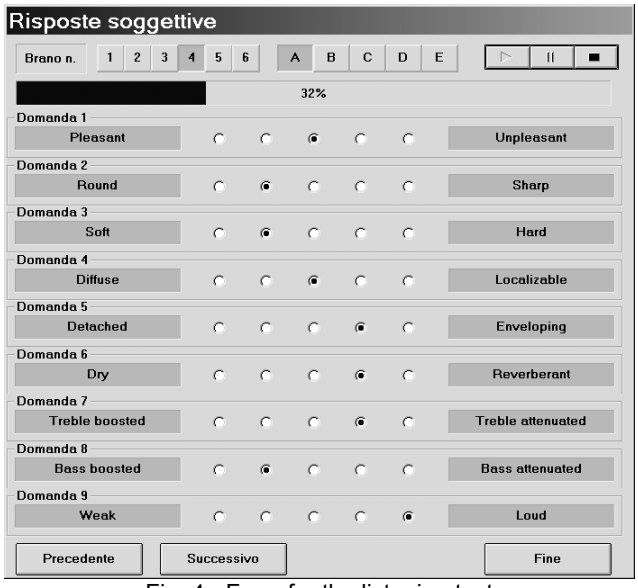


Fig. 4 - Form for the listening test

2.5 Statistical Results

For a first simple analysis we used the approach of a linear regression. For every crossing of the objective-subjective matrix we found out the coefficient of linear regression (*r*) and correlation. We assumed that for values of *r* higher than 0.3 the line starts to interpolate the points and consequently the objective parameter is at least partially correlated with the subjective one.

The following table contains the results of this linear correlation analysis for the orchestral music:

Coeff. Di Regressione Lineare ORCHESTRALI											
	C50	C80	D50	Ts	EDT	T20	T30	LF	IACC	TB	BR
Piacevole-Spiacevole	0,14	0,19	0,19	-0,24	-0,19	-0,20	-0,12	-0,21	0,16	-0,07	0,09
Rotondo-Spigoloso	**0,35**	**0,42**	**0,46**	**-0,51**	**-0,35**	**-0,37**	-0,22	**-0,51**	**0,32**	-0,22	0,15
Morbido-Duro	0,17	**0,33**	0,28	**-0,45**	**-0,40**	**-0,39**	-0,19	**-0,35**	0,28	-0,11	0,13
Diffuso-Localizzabile	0,17	**0,30**	0,22	**-0,36**	**-0,35**	**-0,37**	-0,26	-0,25	0,26	0,02	0,21
Distaccato-Avvolgente	-0,17	-0,24	-0,19	0,23	0,16	0,19	0,09	0,21	-0,21	-0,02	-0,17
Secco-Rimbombante	-0,24	**-0,42**	**-0,32**	**0,50**	**0,48**	**0,50**	**0,34**	**0,37**	**-0,36**	0,01	-0,26
Acuti Accentuati-Acuti Ridotti	-0,19	-0,28	**-0,35**	**0,45**	**0,31**	0,26	0,04	**0,44**	-0,22	**0,31**	0,02
Bassi Accentuati-Bassi Ridotti	0,22	0,29	**0,38**	**-0,48**	**-0,36**	**-0,32**	-0,18	**-0,46**	0,20	**-0,34**	-0,04
Sommesso-Sonoro	-0,08	-0,01	0,00	-0,11	-0,15	-0,09	-0,02	-0,05	-0,01	-0,12	-0,11

In this case the situation was not very bad, there are many values above ab absolute value of 0.30, and some even above 0.50.

But, going to the music with song, we found only very few positive results, like a correlation for vocal tracks between pleasantness and T30 as shown in the picture (*Fig. 5*), but they were far below our expectations. The same pleasantness, in according with orchestral and vocal tracks, didn't correlate with any other of objective parameters.

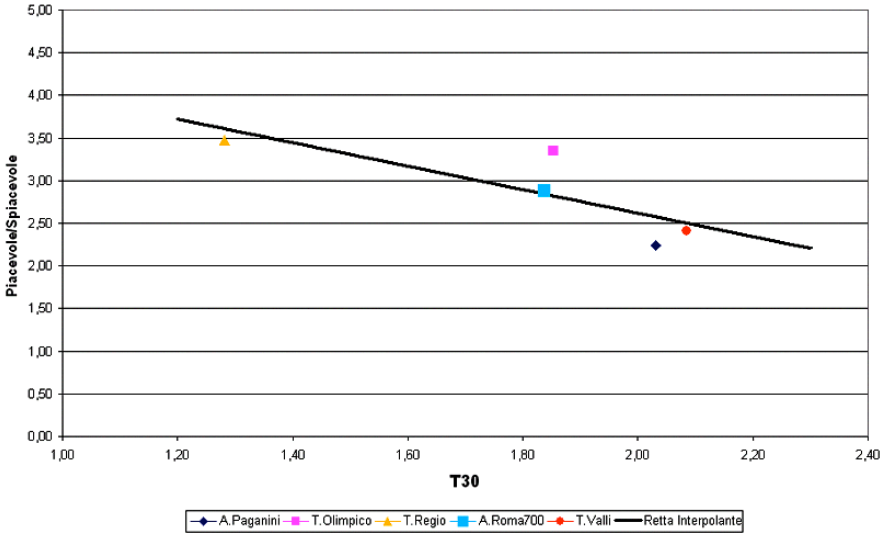


Fig. 5 - Correlation between Pleasant/Unpleasant and T30

For the orchestral tracks, the *Bass Ratio (BR)* didn't match with "*bass boosted – bass reduced*": the cause of this effect could be the use of the additive subwoofer.

A possible cause of this large uncorrelation could be found in the use of headphones for some problems explained in next paragraphs. Another problem is given from the "objectivity" of objective parameters: they are affected by different methods of measurement. For this reason it came out the need of a comparison between different reproduction systems in order to find out the best for a listening test.

Finally, this basic statistical analysis method revealed all its limits, and in the future more advanced analysis schemes will be employed (Principal Component Analysis, Factor Analysis, Multivariate Regression).

3. SETUP OF A NEW ROOM FOR MULTI-SYSTEM LISTENING TESTS

3.1 Acoustic treatment

The reproduction system was realized in a room inside the Casa della Musica in Parma. The room is parallelepiped shaped, the floor is 4.5 m x 3.2 m, the height 4.2 m.

The first step in definition of a hi-performance room is the acoustic treatment of the environment in order to obtain a suitable reverberation time and sufficiently flat frequency response. In theory any kind of virtual environment reproduction should take place in an anechoic room, for not adding any variation to the electronic signal, which should already describe the original environment: actually a completely anechoic ambience (beside being a very expansive solution) is not ideal, because of the odd sensation induced on the listener when music is off. So a little reverberation is suitable, but it should affect as little as possible the “timber” and “dynamics” of the reproduced sound. Reverberation (reflections) and frequency response are strongly bound together, since frequency response is given by different interference pattern between direct field and reflections. The nature of this phenomenon is different at high and low frequencies. For high frequencies the ear analysis window is quite short and space variability rate is very high (short wavelength): hence the perceived frequency response depend just on very early reflection and can change with little receiver movement, roughly maintaining a space averaged flat frequency response, or variations depending on not flat absorbing coefficients of the walls; the later part of the impulse response is instead perceived as reverb (with a timber typical of that room), which as already said is acceptable if sufficiently short.

For low frequencies, instead, the ear analysis time window is longer and spatial variability is very low; so the effect of reflections is mainly to make the perceived direct field stronger and longer at particular frequencies, depending on room shape and dimensions, strongly affecting timber and dynamics; this is like to say that here the dominant phenomenon is stationary waves, or resonance. The medium range is characterized by a transition behaviour between the two already analyzed.

The high frequency behaviour is not difficult to control, traditional absorbing panels can help to reach a suitable reverb time and inverse filtering of the loudspeakers signal can improve the flatness of frequency in a particular listening area. That’s what we did, as shown in figure 6, using glass wool end polyurethane foam.

Fig. 6 - Listening Room

For low frequency things get harder: is very difficult to damp the response of the room at low frequencies because surface roughness should be comparable with the wave length (some meters if we go below 200 Hz!). A typical partial solution is to put resonant open cavities in correspondence to point in which stationary waves have maximum amplitude, that is against the walls. As shown in figure 8, we adopted this technique using loudspeaker cases, cartoon boxes, and “home made” tube traps; also double side rigid and vertical absorbing panels put at a certain distance from the walls was used, to create a kind of cavity, and horizontal ones hung at one meter from the ceiling. Electronic control of low frequency resonance is also a difficult task, because low frequency narrow band filters must be very long, and moreover the deep lack of energy between two resonant frequencies leads the filter to stress very much the speakers. Also diffusing panel were used (like the wooden one visible in *Figure 6*), which are of great help inhibiting resonances in the medium frequency range and making reverb more diffuse.

3.2 Speakers positioning

As a further trick for lowering resonance, the axis of symmetry of the loudspeaker array was not aligned with the room, nor was the listener positioned in the room's centre. Loudspeakers are at a distance of 1.5 m from the listening position. Three couple of loudspeakers are arranged as shown in figure 7 for implementing different kind of reproduction. Dynaudio self-powered studio monitors are used for the conventional stereophonic pair, ±30º from the median plane. Genelec S30D self-powered loudspeakers are used for front stereo dipole, on their sides so that the tweeters were 22 cm apart, the mid-range drivers 43 cm apart, and the woofers 83 cm apart (measuring between driver centres). This corresponds to respective angles of 4º, 8º, and 16º from the median plane of symmetry (the angle seen by the subject between loudspeaker pairs is double these values). The rear stereo dipole pair are QSC AD-S82H passive loudspeakers, fed with a power amplifier, with driver centers separated by 45 cm, corresponding to a 9º angle from the midline. Last, Sennheiser HD580 headphones are available.

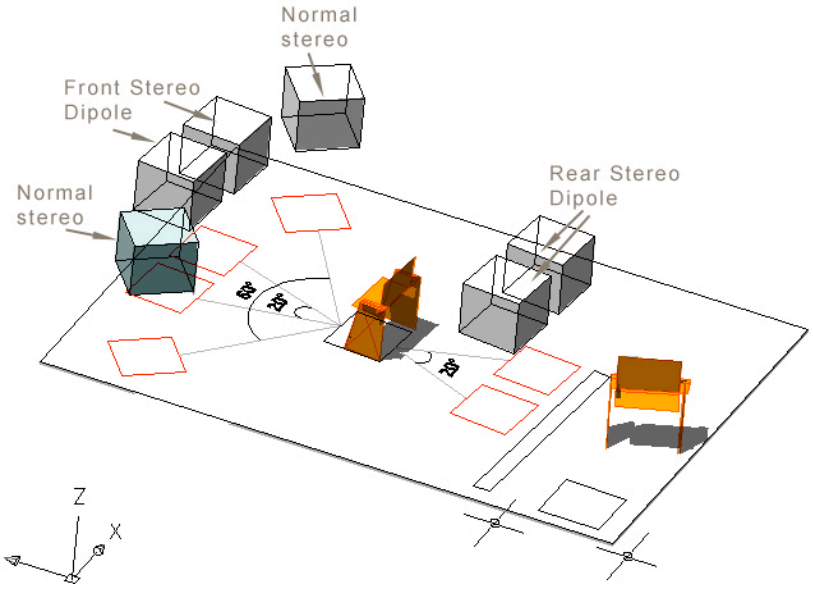


Fig. 7 - Disposition of the speakers inside the listening room

The final frequency response of the room is well described by the reverb time, shown in *Figure 9*. It is measured using the Genelec pair as test source and the dummy head as receiver: it's a matter of fact that the big increasing of energy at low frequency due to resonant modes can't be sufficiently damped.

Fig. 8 - Particular of the listening room

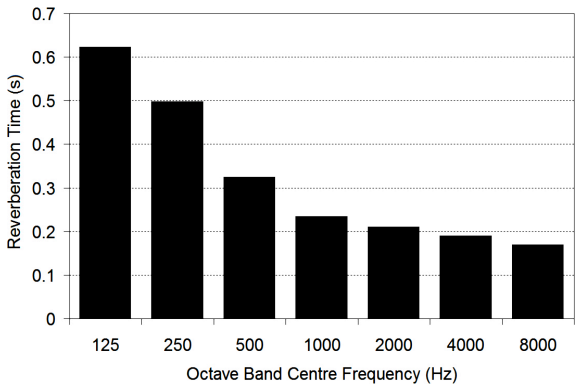


Fig. 9 - Reverberation Time of the listening room

4. Description of the three 2-channels reproduction systems

4.1. Recording

For describing the reproduction systems it's necessary to briefly introduce the corresponding recording methods. Headphones, stereo dipole and double stereo dipole are related to binaural recordings, which are made by means of two microphones capsules put inside the ears of a head, real or "dummy". The purpose of reproduction is in this case to reproduce the same sound pressure at the ears of the listener. The stereo pair is related with a more traditional kind of recording, the ORTF configuraton: two closely-spaced directional (cardioid) microphones with a distance of 170mm and an angle of 110°.

The main goal of the apparatus is to reproduce the impulse responses (or the live recordings) made with our recording system, shown in *Figure 10*. It is composed by three microphones configurations rigidly bound together: a Soundfield microphone, a Neumann dummy head for binaural recording and the ORTF Neumann pair fixed above it. The off axis Sounfield microphone is oriented to other kind of multi channel reproduction. The whole system can rotate on a turn table.

Fig. 10 - Microphone system

The system can be employed both for recording and for impulse response (IR) measurement (using the advanced method of logarithmic sweep described in [9]). In this second case several measurement can be made at different angles by means of the turn table. This may be used for sound tracking or further multi channel reproduction methods.

In general IRs are convolved with one or more anechoic recordings, to reproduce the virtual situation of the recorded source playing in the measured environment, as if it was in the same position of the test source used for IR measurement.

4.2 Playback

In general, a “transparent” sound reproduction system is one which surpasses the “photocopy of the photocopy” test. This test is performed placing the same microphones originally employed for recording or measuring inside the concert halls at the exact position where the head of the listener will be inside the playback room.

Playing through the reproduction system the signals originally recorded in the room, and re-recording them again, we should find that this “second copy” is identical to the “first copy” taken in the concert hall.

However, this does not ensure, by itself, that the listener placed in the listening position will perceive exactly the same sound as if he was in the original theatre (we are confronting two copies, we do not have the original to compare with…..).

If, after the “second copy” measurement is done, this differs from the “first copy”, then it is necessary to introduce in the reproduction chain a set of digital filters, designed with the goal to make the reproduction system fully transparent in terms explained above.

In the following subchapters we will see how these inverse filters are designed for reproduction systems with and without cross-talk.

4.2.1 Headphones

Stereo headphones are the more intuitive tool for reproducing binaural recording. They should put the right pressure directly where it was recorded, at the ears, maintaining the separation between the two (no “cross talk”) and not being affected by the room response. For achieving this target usually the transfer function from the headphones signal to the inner ear is measured, using the same dummy head (Neumann) used for recording; then two inverse filtering are calculated and applied at the binaural signal to make this path transparent, and reproduce exactly the signal recorded at the inner ear.

Reproduction would be very realistic if the head used for recording (or IR measurement) was the same of the listener head. For obvious reasons it is necessary to use a standard dummy head for recording, and this affects the reproduction in a non negligible way.

More over, the fact that the sound image reproduced is rigidly bound to eventual little movement of the listener head, plus the fact of wearing an object on the head which shield the listener from the natural extern background noise, represents psycoacoustic negative artifacts.

4.2.2 Stereo dipole

Stereo dipole aim is to recreate the correct sound signal at the two listener ears through a system of two loudspeakers, each fed with a processed version of the original binaural system, exploiting the technique of cross talk cancellation.

This technique (see *Figure 11*) uses a two by two matrix of (four) filters, calculated so that the system cancels the contribution of the left speakers to the right ear and viceversa. This matrix H is obtained inverting the original matrix of transfer function from speakers to ears C previously measured.

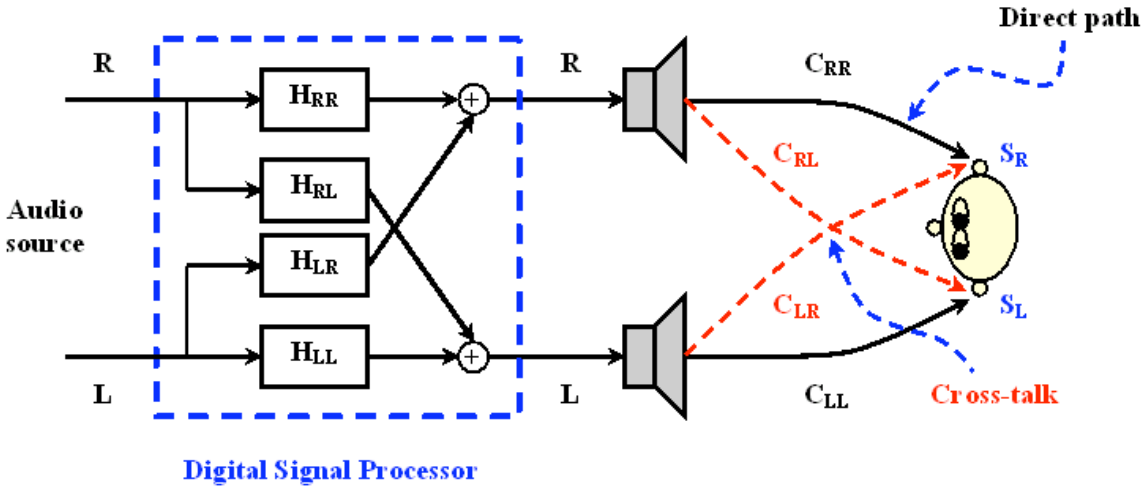


Fig. 11 – Scheme of Stereo Dipole

Kirkeby et al (1998) found that a configuration with a 10º interval between loudspeakers as seen by the listener minimises the ringing artefacts in the cross-talk cancellation filters, and expands the area in which the cross-talk cancellation is effective (allowing greater listener head movement).

This method gives of course a more natural sensation, not relying to a strange and close source like a headphones pair is. More over, the portion of sound coming from the front (in most recordings coincident with the direct field of a source) is spatially correctly reproduced, not only in the neighborhood of the ears, but on a wider area, inducing a natural spaciousness sensation when slightly moving the head.

Our speaker configuration, as described in the previous paragraph, shows an angle between tweeters which is a little narrower than 10°, and a wider one between woofers. According to us this should provide for a better synthesis of the front plane waves and a better separation between the ears of low frequencies.

4.2.3 Double stereo dipole

The double stereo dipole adds a rear pair to the normal stereo dipole configuration. The processing is made of two dipole matrix for H-front and H-rear, calculated inverting independently the two direct matrix C-front and C-rear.

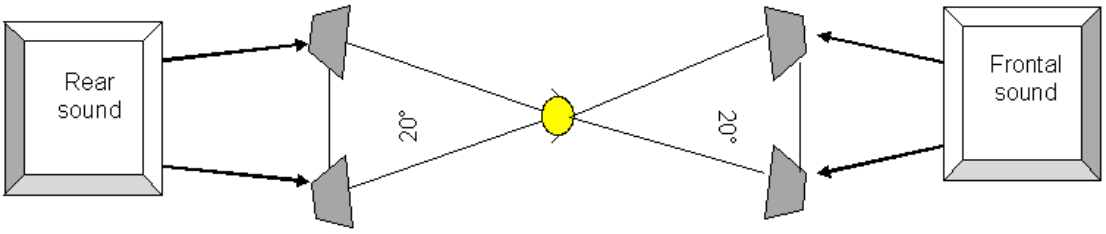


Fig. 12 – Dual Stereo Dipole system

With this approach, the ear pressure induced by the two stereo dipole, front and rear, should be exactly the same in ideal conditions (listener head coincident with measurement head, listener head perfectly positioned and still, ideal reproduction environment); in real conditions instead the double stereo dipole provides also for sound coming from behind the same advantages, previously described,

which the single dipole gives only for sound coming frontally: hence it is supposed to be more realistic in situations in which rear sound is particularly important (i.e. in a theatre, strong rear reflection or applauses).

4.2.4 Normal Stereo

In the normal stereo the sound picked up by the two cardioid microphones is fed to a pair of speakers forming an angle of ±30° (there are possible variation) with the symmetry axis with respect to the listener.

An inverse filtering may be done to flatten the response of the speakers, but, differently from the stereo dipole case, no filtering is performed for cancelling the cross-talk paths.

In this case the only spatial characteristic considered in the recording stage is sound direction, and with a very low angular resolution. The aim here is to discriminate in a very simple and reliable way the sound coming from the front in two principal contributions, left and right. The system is not supposed to give a realistic sensation about the specific spatial characteristic of the environment response, but just an idea of the frontal figure of direct field and first reflections and the temporal shape of reverberation tail, unless it has not particular characteristic of non isotropy.

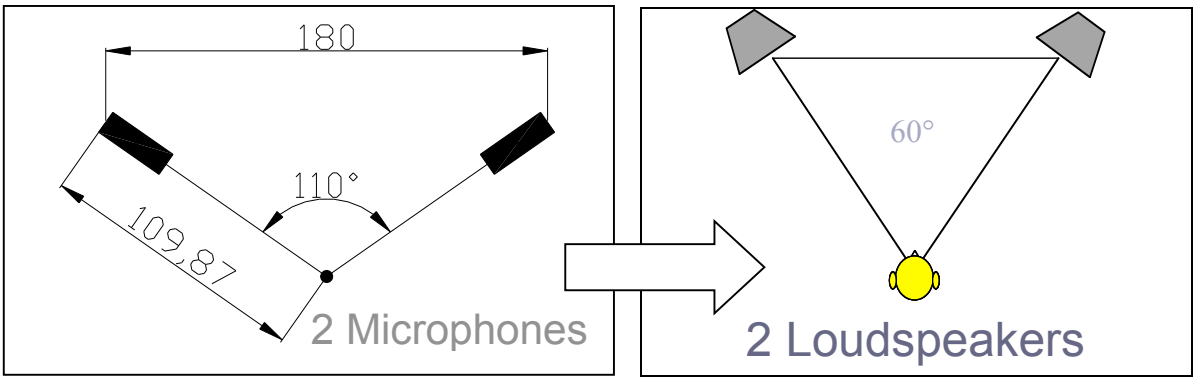


Fig. 13 – ORTF Stereo system

5. COMPUTATION OF THE INVERSE FILTERS

As shown in the previous chapters, any of the recording/reproduction system always requires the insertion of a set of numerical filters (we usually use FIR filters, due to the possibility to implement them very efficiently on modern hardware, and to the fact that designing their sets of coefficients is easier than with other architectures, such as IIR or Warped FIR).

In this chapter we explain the method employed for designing these inverse filters, both for single-input – single-output cases (such as for headphones or normal stereo), or for cross-talk cancelling cases (single and double stereo dipole).

5.1 Kirkeby inversion of a single-input, single-output system

We start with a measurement of one actuator-microphone system, for example obtained placing the headphones over the dummy head, and measuring separately the left-left transfer function (then everything will be repeated identical for the right-right). There are not cross-talk paths in this case.

Figure14 shows the headphones over the dummy head during this measurement.

Fig. 14 – dummy head and headphones

The result of the measurement is a single impulse response, h. We want to find another impulse response (typically longer than h, usually twice long), named f, so that the convolution of h with f is a

perfect Dirac's Delta function δ. This is expressed as follows, both in time-domain and in frequency-domain:

$$f \otimes h = \delta \qquad FFT(f) \cdot FFT(h) = FFT(\delta) \qquad (11)$$

As in frequency domain the convolution becomes a simple multiplication, to be performed independently for each spectral line, it appears to be advisable to get the searched filter f simply making the reciprocal of the complex spectrum of the measured transfer function h:

$$FFT(f) = \frac{1}{FFT(h)} \qquad (12)$$

Unfortunately, this simple approach does not work. In fact, in general h is "mixed phase", and does not admits a direct inversion. Only approximate inversion method can provide an inverse filter f which is stable, causal and of finite length, as required here.

Among various available approximate inversion methods, we did choose the Kirkeby-Nelson [10] inversion, and adapted it with further modification. In practice, the original method was based on taking the reciprocal in frequency domain but adding a small regularization quantity ε at the denominator:

$$FFT(f) = \frac{FFT'(h)}{FFT'(h) \cdot FFT(h) + \varepsilon} \qquad (13)$$

The value of ε has to be chosen with a trial-and-error approach, as it defines a compromise between the length of the inverse filter and the accurate inversion of the spectral peaks and dips.

In general, it is difficult to find a value of ε suitable for the complete wide-band inversion of a transducer-microphone pair, and the results, although stable and workable, are never optimal.

So we modified the original approach, making ε variable with frequency. The idea is to use a small value of ε in the central frequency range, where we want a very accurate inversion, and instead release the things at extremely low and high frequency, where there is no chance to control the transducers anymore, and where the human hearing is less sensitive to errors.

In practice, a suitable spectral variation of ε is as shown in figure15.

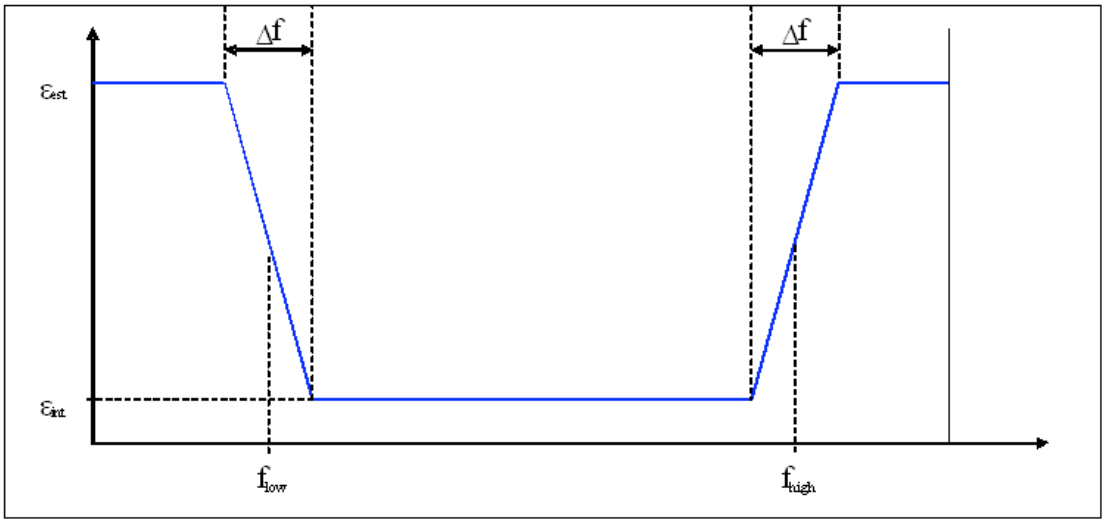


Fig. 15 – regularization parameter varying with frequency

In practice, usually the frequency limits f_{low} and f_{high} are chosen in correspondence of the declared frequency limits of the reproduction hardware. For the Sennheiser headphones shown in the previous figure 14, these limits were set to 40 Hz and 16000 Hz respectively.

Outside this frequency range, a value of ε typically 10 times greater the one used inside the range is used.

5.1 Kirkeby inversion of a cross-talk stereo system

The following fig. 16 shows the cross-talk phenomenon in the reproduction space:

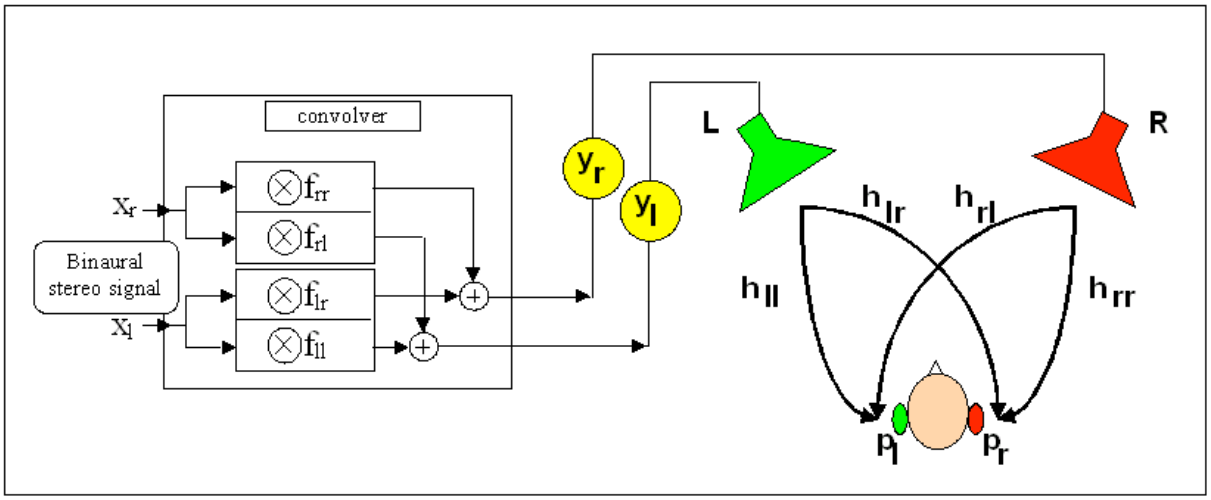


Fig. 16 – cross-talk cancelling scheme

The 4 cross-talk cancelling filters *f*, which are convolved with the original binaural material, have to be designed so that the signal collected at the ears of the listener are identical to the original signals. Imposing that $p_l=x_l$ and $p_r=x_r$, a 4x4 linear equation system is obtained. Its solution yields:

$$\begin{cases} f_{ll} = (h_{rr}) \otimes InvDen \\ f_{lr} = (-h_{lr}) \otimes InvDen \\ f_{rl} = (-h_{rl}) \otimes InvDen \\ f_{rr} = (h_{ll}) \otimes InvDen \\ InvDen = InvFilter(h_{ll} \otimes h_{rr} - h_{lr} \otimes h_{rl}) \end{cases} \quad (14)$$

The problem is the computation of the InvFilter (denominator), as its argument is generally a mixed-phase function. In the past, the authors attempted [11] to perform such an inversion employing the approximate methods suggested by Neely&Allen [12] and Mourjopoulos [13], but now the Kirkeby-Nelson frequency-domain regularization method is preferentially employed, due to its speed and robustness. A further adaptation over the previously published work [14] consists in the adoption of a frequency-dependent regularisation parameter. In practice, the denominator is directly computed in the frequency domain, where the convolutions are simply multiplications, with the following formula:

$$C(\omega) = FFT(h_{ll}) \cdot FFT(h_{rr}) - FFT(h_{lr}) \cdot FFT(h_{rl}) \quad (15)$$

Then, the complex inverse of it is taken, adding a small, frequency-dependent regularization parameter:

$$InvDen(\omega) = \frac{Conj[C(\omega)]}{Conj[C(\omega)] \cdot C(\omega) + \varepsilon(\omega)} \quad (16)$$

In practice, $\varepsilon(\omega)$ is chosen with a constant, small value in the useful frequency range of the loudspeakers employed for reproduction (80 – 16k Hz in this case), and a much larger value outside the useful range. A smooth, logarithmic transition between the two values is interpolated over a transition band of 1/3 octave.

Fig. 17 shows the user's interface of the software developed for computing the cross-talk canceling filters:

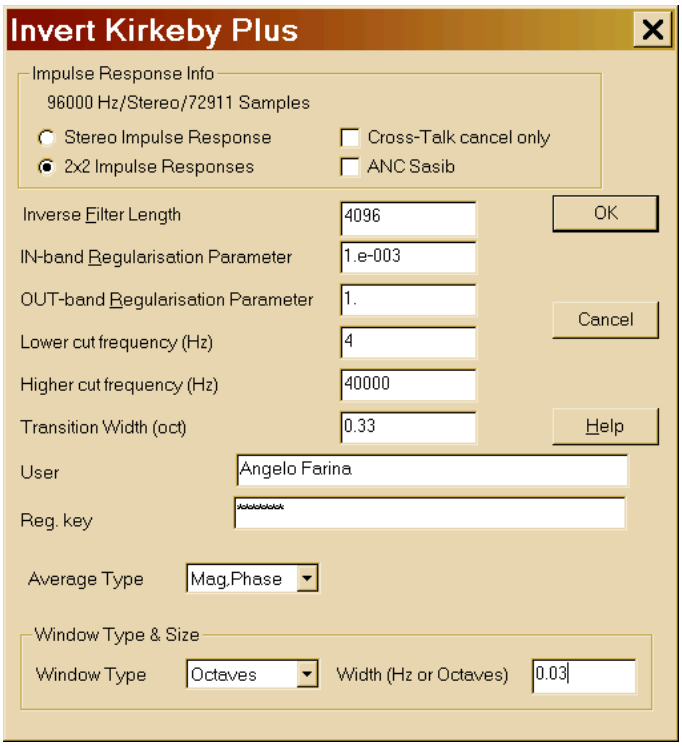


Fig. 17 – user's interface of the inverse filter module

This software tool was implemented as a plug-in for Adobe Audition (formerly known as CoolEdit), and it can process directly a stereo impulse response (assuming a symmetrical setup, so that $h_{ll}=h_{rr}$ and $h_{lr}=h_{rl}$), or a complete 2x2 impulse responses set, obtained placing first the binaural IR coming from the left loudspeaker, followed in time by the binaural IR coming from the right loudspeaker. In both cases, the outputted inverse filters are in the same format as the input IRs.

The computation is so fast (less than 100 ms) that it is easy to find the optimal values for the regularisation parameters by an error-and-trial method.

6. HARDWARE AND SOFTWARE

In order to compare all the four system at the same time, a special audio networking was developed. A notebook located in listening position was connected to a Soundcard, EDIROL 101 Firewire Capture Interface, and its four analogue output channel pairs are used to drive the four reproduction systems already described. So, using the special software described later, it is possible to select in real-time any of the 4 reproduction systems, and routing to it automatically the proper recording (O.R.T.F or dummy head).

The notebook contains the stereo recordings made with the ORTF microphones and with the dummy head, or the equivalent ORTF and binaural tracks obtained with a previous off-line convolution of an anechoic track with ORTF and binaural measured impulse responses. The stereo ORTF track is send to the stereo output which feeds the normal stereo loudspeakers, while the binaural track is sent on one of the three stereo outputs, connected respectively to headphones, stereo dipole and double stereo dipole systems. All the 8 output channel (4 stereo pairs) are connected to a separate PC, equipped with the Ardvaark Q10 audio interface, which provides for 4 stereo high quality input and output pairs.

The I/O is done at 24 bits, 96kHz, and all the processing, filtering, etc. is performed in floating-point with 32-bits precision.

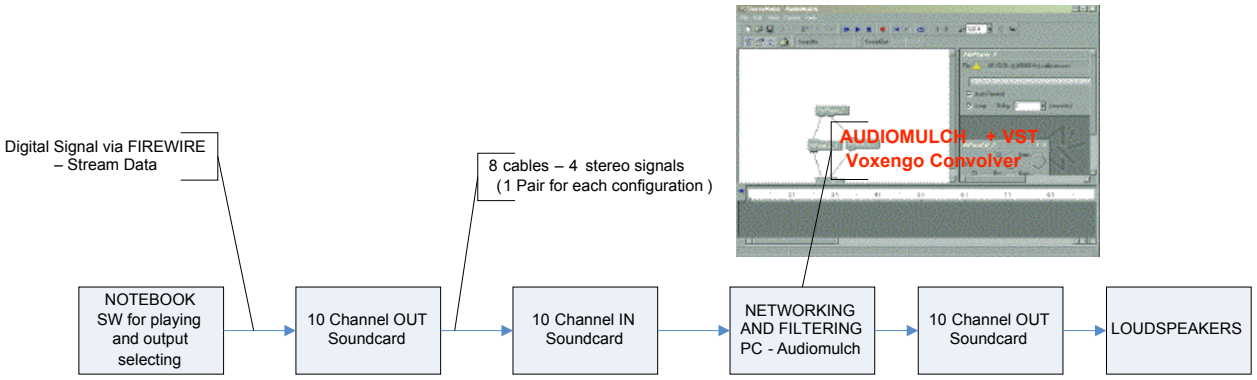


Fig. 18 - Chain of reproduction

The four stereo inputs of the Aadvark sound card receive the four stereo outputs coming from the notebook + EDIROL system. The four stereo channels pass through an application called Audiomulch: it is a multichannel VST host, inside which two instances of a plug-in called Voxengo Pristine Space are inserted. Pristine Space is a system able to handle multi-channel convolution of up to 8 incoming signals with up to 8 FIR filters.

Hence, setting in Audiomulch four stereo filtering configurations linked respectively to the four stereo pairs of the Q10, we drive separately the four reproduction systems. After being filtered each signal pair is send to correct Q10 outputs, connected to the corresponding loudspeakers. So this hardware-software tool made of PC, Q10, Audiomulch and Voxengo act as a four chanFor headphones, normal stereo, and stereo dipole, each input stereo pairs drives a single stereo output. In the case of the dual-stereo-dipole system, however, the input stereo pair is filtered with two independent sets of FIR filters, and drives two stereo outputs, one feeding the frontal stereo dipole, the second feeding the rear stereo dipole.

This long connection path is synthesized in the scheme in *Figure 18*.

It would be more clever to use just a single PC and a single multichannel sound card. But this would require a software solution capable of re-routing the outputs of an application as inputs to a second one. Although these software tools are existent nowadays, the system was not setup in this way yet, mainly for reasons of time, and for the fact that the computing power required would be too much for the low-end notebook employed as playback device.

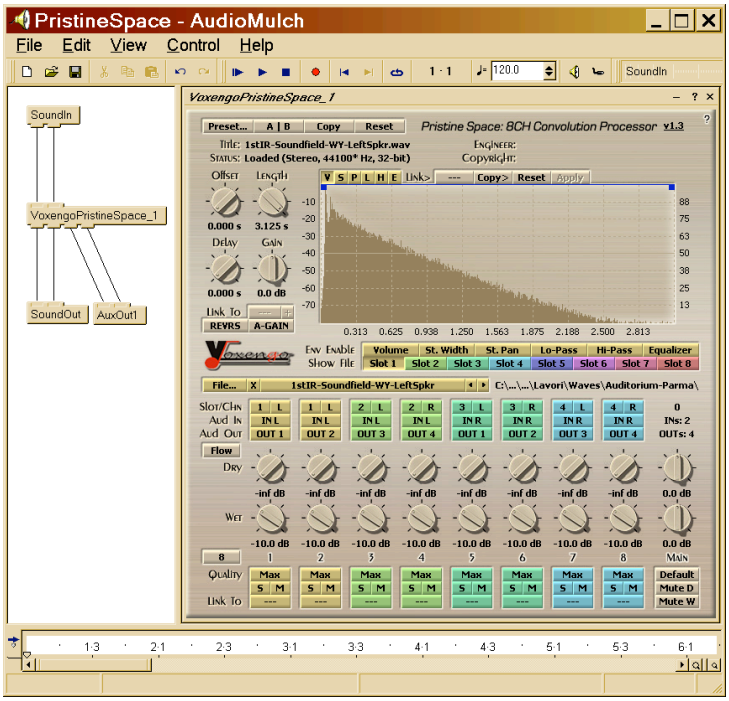


Fig. 19 – A simple dual cross-talk canceling network with AudioMulch and Pristine Space

The filters implemented in Audiomulch with the 2 instances of Voxengo Pristine Space are

1. the cross cancellation filters (2 by 2 matrix) for the frontal and rear stereo dipoles, which also automatically provide for flattening the response of the speakers employed (the first instance of Pristine Space has 2 inputs, 4 outputs and 8 FIR filters)
2. simple stereo inverse filters for headphones and normal stereo, which, as already explained, flatten the transfer function from headphones to ears, and for normal stereo flatten the loudspeaker's response (the second instance of Pristine Space has 4 inputs, 4 outputs and 4 FIR filters).

All these inverse filters were preliminarily designed, starting by measurement of the direct impulse responses. The details of the inversion were explained in chapter 6.

In all of these inversions the impulse responses are truncated just after the direct sound pulse: trying to invert also the residual contributions of the room response revealed to cause more problems

than those which are solved, and makes the system very unstable and not robust to listener movement.

Inverting only the direct sound has also other advantages:

1. The inverse filters are short (8192 coefficients at 96 kHz), so there is no problem for the PC to convolve 12 of these filters simultaneously employing just a fraction of the available computing power.
2. The latency of the convolver is short too, so that when the listener switches the sound sample, he hears immediately the new sound.
3. Modifications of the fittings of the room (which is continuously improved) do not require creating a new set of inverse filters.

Finally, it must be remarked here that the Voxengo Pristine Space convolver revealed to be very versatile, as it already contains the possibility to sum (mix) together the results of the convolutions before feeding the outputs. This revealed to be precious for implementing the cross-talk cancellation networks. This convolver revealed also very good performances in terms of CPU usage and latency, outperforming our Aurora convolution plugin, which simply could not stand up in comparison.

7. COMPARATIVE LISTENING TEST

Aim of this test session is to investigate comparatively the capabilities of the four systems to reproduce with realism the acoustic of a theatre. In order to make this comparison an anechoic track of an accordion was convolved with Binaural and ORTF Impulse Response recorded in five Auditoria and in different positions in the stalls. The software that we designed (*Figure 20*) permits to switch between ten acoustic situations: every situation is different from the other for position in the stalls, system of reproduction, theatre. This means that choosing one of the numbered buttons the software plays the binaural track or the ORTF one, selecting the corrispondent output of the Edirol soundcard as explained in the proevious paragraph.

Pressing another button, the track doesn't restart but continues playback giving the impression of a virtual "jump" between the theatres.

The subject has to answer to three questions. The first is about the perception of the room's dimension with an evaluation between *Small* and *Very Big*, the second is about the realism of the sound that the subject is earing and the third asks for the distance in meters of the accordion that is playing.

Till now we tested 24 subjects, all of them musicians or musically trained people.

The results of this test will be published in a nearly future, after proper statistical analysis of the results.

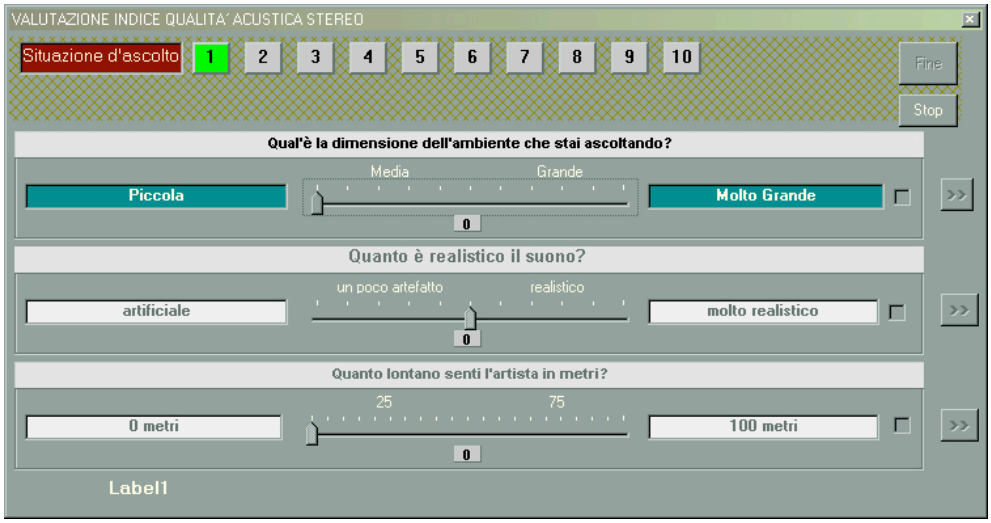


Fig. 20 - Software for the new listening test

8. CONCLUSIONS

A listening room equipped with three pairs of loudspeakers and one of headphones has been set up, and a lot of effort has been spent to design the filters for flattening the loudspeakers and to treat the room with passive and active means.

A complex hardware/software system has been developed, for allowing instantaneous switching between the 4 reproduction systems, and implementing the proper real-time digital filtering for each of them, so that the computer only needs to play the original soundtracks.

The system also features an easy user's interface, allowing for automatic collection of questionnaires, and giving to the subject the freedom to jump at will among the sound samples, re-listening to what he wants, and switching back and forth for A-B comparisons.

The goal is to rank the capabilities of the four systems to reproduce the spatial characteristics of the real acoustical spaces, by means of subjective tests which are currently going on. This should bring to the selection of the "optimal" playback system, which will be subsequently employed for other campaigns of listening tests.

The first future work is to repeat the tests for objective-subjective parameters correlation, using the system which will reveal to be the best one for reproducing frontal performances in theatre.

Then the study of reproduction quality versus system employed will be extended to more advanced multi channel systems, i.e. Ambisonics and Ambiophonics, and to sound samples including music, song and speech.

9. REFERENCES

[1] A. Farina, "Acoustic quality of theatres: correlation between experimental measures and subjective evaluations", Applied Acoustics, Volume 62, Issue 8, Pages 889-1023 (August 2001).

[2] The Aurora software plugins – HTTP://www.aurora-plugins.com

[3] M. Schroeder, "New method of measuring reverberation time", Journal of Acoustical Society of America, 37, 409-412 (1965)

[4] V.L. Jordan, "A group of objective acoustical criteria for concert halls", Applied Acoustics, 14 (1981)

[5] R. Kürer, "Zur gewinnung von eizahlkriterien bei impulsmessungeg in der raumakustik", Acustica, 21 (1969)

[6] Y. Ando, "Concert Hall Acoustics", Springer - Verlag, Berlin, 1985

[7] W. Reichardt, O. Abel Alim, W. Schmidt "Definition und Meßgrunglage eines objectiven Maßes zur Ermittlungder Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei musikdarbietung" Acustica, 32 126 (1975)

[8] L. Beranek, "Music, acoustics and architecture" John Wiley & Sons, New York (1962)

[9] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique", 108th AES Convention, Paris 18-22 February 2000.

[10] O. Kirkeby, P. A. Nelson, H. Hamada, "The "Stereo Dipole"-A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers" – *JAES* vol. 46, n. 5, 1998 May, pp. 387-395.

[11] A. Farina, F. Righini, 'Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering', Pre-prints of the 103rd AES Convention, New York, 26-29 September 1997.

[12] S.T. Neely, J.B. Allen, 'Invertibility of a room impulse response', *J.A.S.A., vol.66*, pp.165-169 (1979).

[13] J.N. Mourjopoulos, "Digital Equalization of Room Acoustics", *JAES vol. 42, n. 11*, 1994 November, pp. 884-900.

[14] A. Farina, E. Ugolotti - "Spatial Equalization of sound systems in cars" - Proc. of 15th AES Conference "Audio, Acoustics & Small Spaces", Copenhagen, Denmark, 31/10-2/11 1998.