SPATIAL EQUALIZATION OF SOUND SYSTEMS IN CARS

Angelo Farina (1), Emanuele Ugolotti (2)

(1)   Dipartimento di Ingegneria Industriale, Università di Parma, Via delle Scienze – 43100 PARMA
       tel. +39 0521 905854 - fax +39 0521 905705 –
       E-MAIL: farina@pcfarina.eng.unipr.it – HTTP://pcfarina.eng.unipr.it
(2)   ASK Automotive Industries, via Fratelli Cervi n. 79, 42100 Reggio Emilia
       tel. +39 0522 388311 - fax. +39 0522 388499 - E-MAIL: tec_ask@sirnet.it

ABSTRACT

This paper is about the equalization of the sound reproduction systems in cars. In particular, the whole equalization process can be seen as a cascade of multiple filters, each of them with different purposes: frequency response regularization, time domain realignment, spatial displacement of the sound sources, linearization of the loudspeaker's distortion, etc. .

A modular mathematical formulation is derived, making it possible to approach separately each stage of the equalization process: the greater effort is devoted here to the spatial equalization, which makes it possible to move the apparent position of the loudspeakers, and to modify the subjective spatial impression due to the volume in which the reproduction takes place. Nevertheless also the other equalization stages are discussed.

From the proper implementation of the above theory in a set of suitable software modules, it was possible to transport a measured sound field in a very different reproduction space. This was useful in performing two opposite tasks: reproducing the sound field of good listening rooms or theatres inside the compartment of a car, or reproducing the sound field of different cars in a listening room. The second task was useful for making blind subjective comparisons of the sound system of different cars, whilst the first task made it possible to recreate the preferred sound field inside another car.

In both cases it was first necessary to measure binaural impulse responses in the original space and in the reproduction space (making use of the same dummy head), after which a proper set of digital inverse filters was created, capable of transforming the second ones in the first ones. These filters are applied to the signal to be reproduced (music coming from CD tracks) by convolution, thanks to a real-time convolution software. The results were judged subjectively, thanks to an automated interactive questionnaire implemented as another software tool.

This paper describes in detail the theory developed for building the numerical filters for spatial equalization, the software developed for measuring the binaural impulse responses and for processing them, and the results of the subjective listening tests.

1   INTRODUCTION

The reproduction of sound in cars is not a trivial task. The proper disposition of loudspeakers, the right amount of them to be used, their features and the power trade-off require a deep knowledge of some variables not yet perfectly explained. Therefore, the evaluation and optimization of car sound systems is usually done both by means of objective measurements (frequency response, distortion, etc.) and by listening tests, in order to find out the optimal installation for a given car. The major efforts spent in these last years were devoted to modify the sound field in car's interior, basically by using, sometimes not properly, special electronic devices such as filters, delays and, recently, DSPs and FPGA.   All these devices can generally be regarded as equalizers: of course, what it's important is their goal and how they are implemented.

In a car compartment the sound field is heavily affected by the "strange" position of the sound sources, which is substantially different from the optimal "stereo triangle". Furthermore, the small volume of the compartment, and the fact that some surfaces (the windows) are highly reflecting, produces very evident resonances and reflections, which cause large alterations of the frequency response, and are perceived subjectively as a "small box" effect.

In general, the complete equalization of a reproduction system can be seen as the cascade of various separate effects, as shown in fig. 1.

It must be noted, anyway, that both frequency-domain and time-domain equalization are required for obtaining the modification of the perceived direction of provenience of the sound, because the human localization cues are due both to time-arrival differences (at low frequency) and to interaural level differences (at medium frequencies) and subtle spectral modifications (pinna colouring, at high frequencies).

For these reasons, a complete digital equalization of a car audio system has not to be intended simply for flattening the frequency response: also time-domain effects are included, for re-aligning temporally the sound coming from sources located at different distances from the listeners. In some cases, the digital filters are used to perform a "virtual displacement" of the sound sources, giving to the listener the subjective impression of listening at a pair of virtual loudspeakers, properly placed in symmetrical positions (Nelson [1,2]).
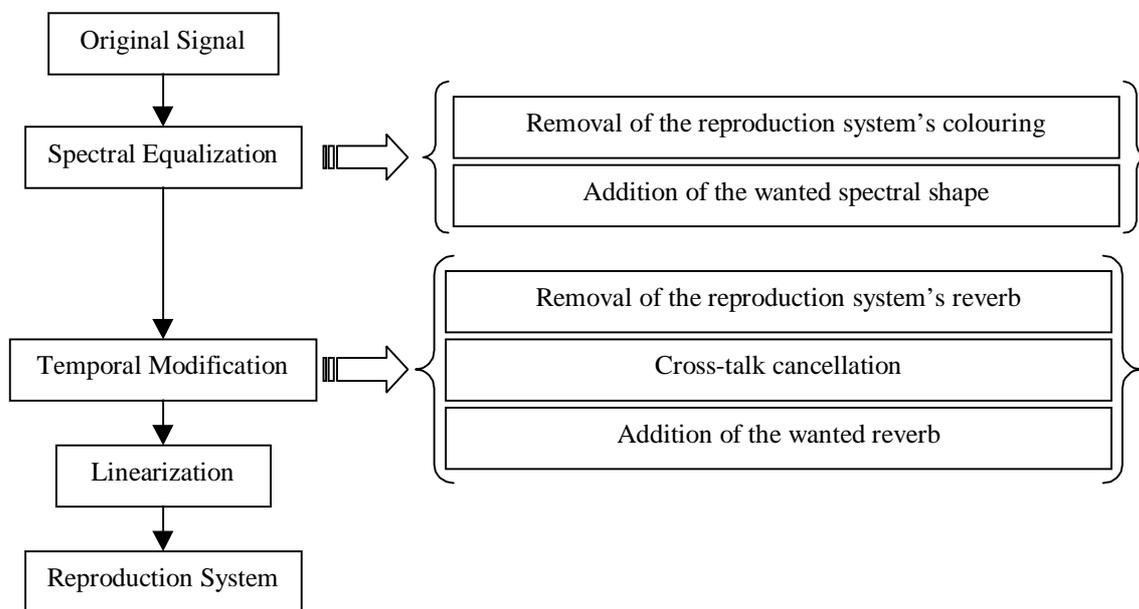
Fig. 1 – block diagram of a complete equalization system

In this work, the goal is even wider: the whole sound field is substituted with another one, sampled in a very different space, such as a theatre or a good listening room [3]. This means that also the reverberation and the discrete reflections have to be recreated successfully, with their spatial impression. This can be performed with the measurement of the Head Related Transfer Functions (HRTF) of the driver in two conditions: seated inside the car, and located in the "ideal space" to be reproduced. In principle, both these HRTF measurements have to be obtained using the driver's head: but in this case only dummy-head HRTFs were available, for the experimental measurements in theatres and for the numerical synthesis starting from computer simulations of non-existent spaces. So also inside the car the same dummy head was used.

After this, the construction of the digital filters is possible by means of a new processing procedure, which takes into account also the cross-talk effect, to ensure that at each ear of the driver only the signal originally received by the correspondent dummy-head's ear in the virtual space is perceived. Nelson et al. [1,2] suggested a global approach, in which the construction of the numerical filters is made solving simultaneously a set of equations, which minimize simultaneously the spectral and temporal error. In this work an alternative approach is developed, separating the construction of the single modules depicted in fig. 1: this makes it possible to minimize the length of the numerical filters, releasing some of the constraints in the modules of lower subjective importance, or even entirely eliminating some processing blocks.

The feasibility of the new approach was tested implementing the processing of two virtual sound spaces on a car equipped with 4 different reproduction systems, and performing a subjective comparison test.

But the same processing technique and apparatus can be used also in the other way: sampling the sound field inside a car and reproducing it in a suitable listening room. This makes it possible to conduct blind subjective tests of different cars, without the problems usually encountered in real tests: the subjects do not need to move between the cars, they can listen at each car how many times they want, and their responses are not biased by the knowledge of the car's maker or by non-acoustical effects due to the furniture of the car or to other comfort-related topics.

The employment of the numerical technique described here results in processed samples which are almost indistinguishable from original binaural recordings of the same sound sample directly recorded inside the car: but the new technique allows for the instantaneous processing of any other sound sample, and has the additional advantage of making use of loudspeakers for the reproduction instead of headphones.

The noise produced by the engine and by the tyres during the normal operation on-road is also reproduced: it is synthesized separately, starting from proper noise measurement conducted on each car during a road test, and mixed at the proper level over the numerically processed binaural signals. This means that the subjective tests conducted in the listening room are even more realistic than the tests conducted directly inside the car, as in the second case usually it is not possible to have the engine running, nor the tyre and aerodynamic noise.

Also for this second task, an experiment was conducted for the subjective comparison of 9 different cars. The collection of questionnaires was automated, thanks to a specially-built software which automates both the reproduction of the sound samples and the collection of the subjective responses for each of them.

Another very important point covered in this papers is the selection of the members of the judging panel, which was employed for both the above-described experiments: a preliminary listening test, conducted on a larger number of volunteers, revealed that there is a certain number of subjects who give inconsistent responses. If these subjects are included in any test, their "random" responses certainly degrade the overall coherence of the data, and often the spread can be so high that no evident correlation is found among the responses. The selection of the listeners revealed to be a very promising technique, for transforming

the subjective test technique from an "approximate" diagnostic tool to an exact measurement.

2 THEORY FOR THE COMPUTATION OF THE INVERSE FILTERS

Let we look first at the direct equalization problem, that is to reproduce the sound field of a good listening room inside a car compartment; it can be seen that the opposite task (reproducing the sound field of a car inside a listening room) can be obtained easily exchanging the impulse response of the two spaces.

Fig. 2 shows the paths of the sound from loudspeakers to ears in a theatre (or in a good listening room), and the same paths inside the car compartment. It is clear how in the second case the paths are not symmetrical, being present an unavoidable different delay along them. The goal of the equalization system is to create numerical filters, to be inserted in the reproduction path of the car, so that at the listener's ears arrives a sound field almost identical to the one perceived in the theatre. The schematic diagram of such a set of filters is shown in fig. 3.

First binaural measurements of the "ideal" impulse responses are performed in an high quality listening environment, which can be a concert hall or an esoteric hi-fi system. In a similar way, the "unwanted" impulse responses are measured inside the car, making use of the same dummy head already used in the ideal environment, or simply of the head of the particular listener, equipped with a wearable binaural microphone set (Sennheiser MKE 2002).
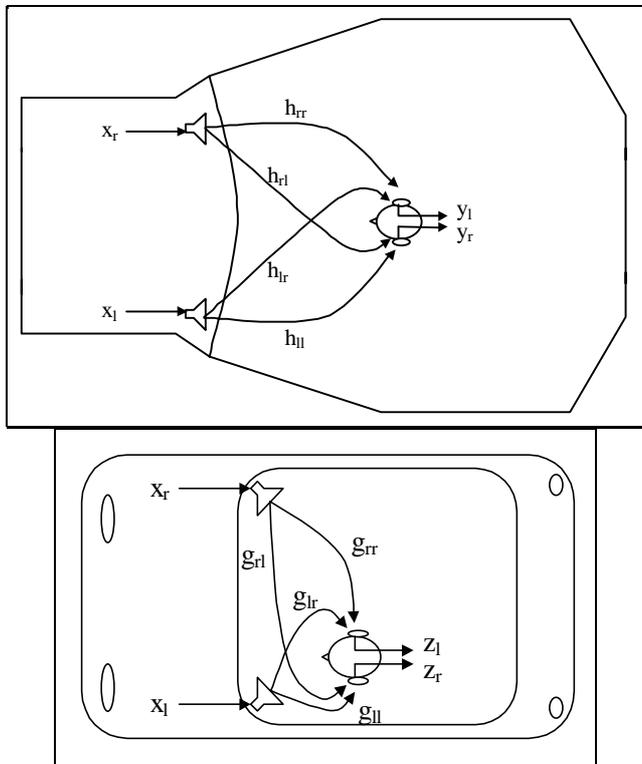

Fig. 2 – Ideal listening conditions in a theatre (above) and effective listening conditions inside a car (below)

Then, through the formulation described in this chapter, the proper inverse filters are numerically computed, and stored in the memory of a real-time convolver, shown in fig. 3. At this point it is possible to process (in real time) any kind of source signals, coming for example from a CD player or from a radio receiver, filtering them in such a way that they arrive at the listener's ears with the same characteristics as if they were played in the ideal listening environment, instead of in the car.

Obviously the whole process works only if each part of the system is perfectly linear, because the theory of linear systems is required for performing impulse-response based processing. Thus this digital equalizer cannot correct non-linear distortions, caused for example by the attempt to exceed the power handling limits of the system. As shown in fig. 1, anyway, the system can be made to behave linearly, within certain limits, by inserting another numerical filter before sending the signals to the loudspeakers: this task is not described here, although the authors already developed both the theory and some practical implementations of it [4].

If the original signals, coming from the CD player, are denoted as x_l and x_r, when they are played in the ideal listening environment, characterized by the four impulse responses denoted h_{ll}, h_{lr}, h_{rl} and h_{rr}, the ideal listening signals y_l and y_r are received at the listener's ears:

$$y_l = x_l \otimes h_{ll} + x_r \otimes h_{rl}$$
$$y_r = x_l \otimes h_{lr} + x_r \otimes h_{rr} \qquad (1)$$

Instead, if the same original signal is played inside the car, characterized by the four impulse responses denoted f_{ll}, f_{lr}, f_{rl} and f_{rr}, the perceived signals are:

$$z_l = x_l \otimes g_{ll} + x_r \otimes g_{rl}$$
$$z_r = x_l \otimes g_{lr} + x_r \otimes g_{rr} \qquad (2)$$

Now we introduce the digital equalizer, following the scheme shown in fig. 3, which also contains 4 impulse responses, denoted f_{ll}, f_{lr}, f_{rl} and f_{rr}. It processes the signal coming from the CD player, and sends to the loudspeakers modified signals, denoted w_l and w_r, which are given by:

$$w_l = x_l \otimes f_{ll} + x_r \otimes f_{rl}$$
$$w_r = x_l \otimes f_{lr} + x_r \otimes f_{rr} \qquad (3)$$

The goal of the digital equaliser is to make so that, filtering the signal coming from the CD player before sending them to the car's loudspeakers, at the listener's ears the ideal listening signals y_l and y_r are perceived; this means that it should be:

$$y_l = w_l \otimes g_{ll} + w_r \otimes g_{rl}$$
$$y_r = w_l \otimes g_{lr} + w_r \otimes g_{rr} \qquad (4)$$

Substituting in equations (4) the expression of y_l and y_r, coming from eqn. (1), and those of w_l and w_r, coming from eqn. (3), the following is found:
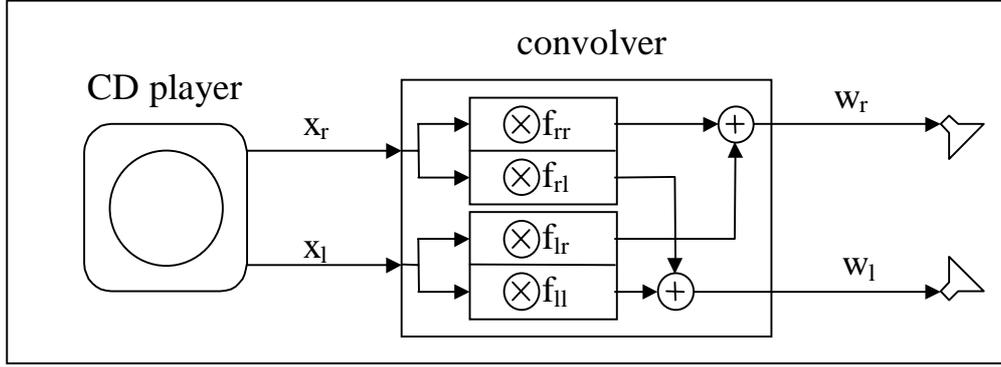
Fig. 3 – Block diagram of the convolver

$$x_l \otimes h_{ll} + x_r \otimes h_{rl} = (x_l \otimes f_{ll} + x_r \otimes f_{rl}) \otimes g_{ll} +$$
$$+ (x_l \otimes f_{lr} + x_r \otimes f_{rr}) \otimes g_{rl}$$
$$x_l \otimes h_{lr} + x_r \otimes h_{rr} = (x_l \otimes f_{ll} + x_r \otimes f_{rl}) \otimes g_{lr} + \quad (5)$$
$$+ (x_l \otimes f_{lr} + x_r \otimes f_{rr}) \otimes g_{rr}$$

For these equalities to be always true, it is needed that for any value of the input signals x_l and x_r, on the left and right term the factors multiplied (or, better, convolved) with them are the same. So it must be:

$$\begin{cases} h_{ll} = f_{ll} \otimes g_{ll} + f_{lr} \otimes g_{rl} \\ h_{rl} = f_{rl} \otimes g_{ll} + f_{rr} \otimes g_{rl} \\ h_{lr} = f_{ll} \otimes g_{lr} + f_{lr} \otimes g_{rr} \\ h_{rr} = f_{rl} \otimes g_{lr} + f_{rr} \otimes g_{rr} \end{cases} \quad (6)$$

After a few, easy mathematical passages, this linear system is solved, and we find the expressions for the wanted filters:

$$\begin{cases} f_{ll} = (g_{rr} \otimes h_{ll} - g_{rl} \otimes h_{lr}) \otimes \text{InvDen} \\ f_{lr} = (g_{ll} \otimes h_{lr} - g_{lr} \otimes h_{ll}) \otimes \text{InvDen} \\ f_{rl} = (g_{rr} \otimes h_{rl} - g_{rl} \otimes h_{rr}) \otimes \text{InvDen} \\ f_{rr} = (g_{ll} \otimes h_{rr} - g_{lr} \otimes h_{rl}) \otimes \text{InvDen} \\ \text{InvDen} = \text{InvFilter}(g_{ll} \otimes g_{rr} - g_{lr} \otimes g_{rl}) \end{cases} \quad (7)$$

In which the terms within brackets of the first four expressions are simply computed, being the sum of two convolutions: the problem is in computing the inverse of the denominator (InvDen), creating an inverse filter for the expression given by the fifth expression. The creation of the inverse filter for a mixed-phase impulse response is not an easy task, although it was addressed by many authors, and particularly by Mourjopoulos [5].
A software module which implements the Mourjopuolos least-squares approximate inversion has been already developed by one of the authors (Farina [6]), and can be used for this task. But in some cases, better results are obtained if a simple zero-phase (or equivalent minimum phase) inversion is done, following the well-known approach of Neely and Allen [7]; another software module is available for this task (Farina [6]). In this second way a simple frequency domain equalization is performed, and the all-pass component (which carries the reverberation) is

left unequalized, the complete filters f produced are more stable, and no audible artifact is introduced.
A third approach for the inversion of the denominator was suggested by Kirkeby [8]: the frequency spectrum is inverted taking into account also its phase, but a regularization value is added for avoiding that too high peaks appear at the frequencies where the modulus of the denominator is too great. This frequency domain approach gives a compromise between the previous two, being capable at a certain amount to remove echoes and reverberation, although not so effectively as the Mourjopoulos approach, and being fast to compute and robust as the Nelly and Allen approach.
If the reverberation, contained in the desired impulse responses h, is greater than the reverberation of the car compartment, as it happens when the first is, for example, a concert hall, then there is no need for removing the car reverberation, which is anyway masked under the room's reverberation. (The same holds in the case of the opposite application, when a reverberant car compartment has to be reproduced in an hemi-anehoic room).
Being the denominator the same for the 4 filters, the incomplete inversion of it does not affect the spatial perception of the sound field, and thus the "virtual displacement" of the sound sources is achieved anyway. If instead the "ideal" listening room is almost anechoic, as in the case of the audiophile stereo system, not removing the car reverberation is usually unacceptable, and causes a subjective perception significantly different from the ideal one, particularly for rapid transients or when the music suddenly stops.

3 SOFTWARE FOR DIGITAL SIGNAL PROCESSING

In recent years a lot of applications were developed making use of specialized hardware, such as DSP units or data acquisition boards equipped with DSPs. The programming of such units is usually slow and complex, and the sound quality is partially reduced by the fact that in most cases fixed-point math is used.
Although low-cost DSP "black boxes" are still an economic solution for series production, for research

purposes the modern approach is to use powerful, general purpose personal computers, which provide a favorable ratio between cost and development time. In fact it is easy to create small computing codes, which perform the required mathematical manipulation on the audio signals, and to run them from within sound editing programs, already equipped with a lot of standard filtering and processing capabilities.

In this case, the program Cool Edit Pro (by D.Johnston) was used as a starting point: it is a waveform editor, which can be easily expanded by writing little additional subprograms, in the form of custom DLLs (Dynamic Linked Libraries). These are automatically inserted in the main program menus, giving a smooth integration as if they were part of the original program.

Ten of these software modules were developed, covering many tasks encountered in the digital equalization of sound systems. In particular, the 6 modules relevant for the goals of this research are:

- generation of the excitation signal (Maximum Length Sequence) and deconvolution of the system's impulse response from the signal sampled at the output of it;
- creation of inverse filters of a given impulse response, with the Mourjopoulos complete inversion, with the Neely and Allen minimum-phase inversion and with the Kirkeby regularization;
- real-time convolution of arbitrary long (mono or stereo) signals with up to four impulse responses, producing a stereo output which is a filtered version of the input signal, passed through the FIR filters represented by the impulse responses.

In the following, a brief description of these modules is given.

3.1 MLS Measurement of the Impulse Responses

Usually the car-audio sound systems are typically subdivided in 4 separate sections, corresponding to the two sides (left and right) and to the front/back placement of the loudspeakers. A complete characterization can be obtained only measuring separately the acoustic response of each section; after this, the behaviour of the sound system when all the channels are being used together can be derived mathematically, both for the hypothesis of completely independent feeds, or for (partially) coherent signals.

The basis for the measurement is a software implementation of the Maximum Length Sequence technique [9], which allows for the determination of the Impulse Response between each input channel and each measurement microphone. Typically two dummy heads equipped with binaural microphones are used, placed one at the driver's seat, the other at the passenger seat behind the driver: so 16 impulse responses are measured.

From the measured impulse responses a lot of objective parameters can be obtained: steady state frequency response, reverberation time, inter-aural cross-correlation and many others. Furthermore, they can be used for the construction of numerical filters, as it was depicted in the previous paragraph.

The measurement with the MLS technique consists of two logical parts: generation of the excitation signal and deconvolution of the system's impulse response from the raw data sampled at its output.

The MLS excitation signal is well known since at least two decades [10, 11, 12]: it is a binary sequence, in which each value can be simply 0 or 1, obtained by a shift register, as shown in fig. 4.
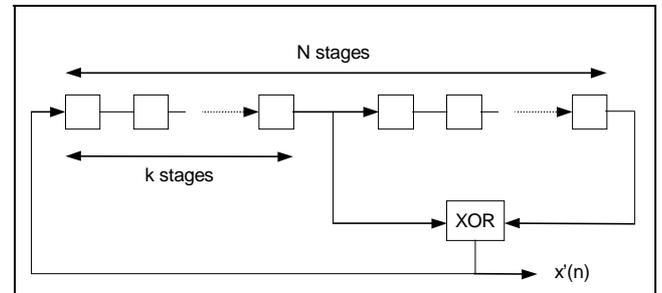

Fig. 4 – Shift register for the creation of the MLS signal

The obtained signal is periodic, with period of length L given by:

$$L = 2^{N-1} \qquad (8)$$

in which N is the number of slots in the shift register, also called the order of the MLS sequence. Thus an order N=16 means a sequence with a period of 65535 samples.

Another very important point is the position of the tap inside the shift register: it is possible to generate MLS sequences also with multiple taps, and the position of the taps influences the behavior of the sequence, particularly when it is used for the excitation of systems which are not perfectly linear. The work of Vanderkooy [12] suggests that some sequences are better than others, reporting a list of known taps positions and values to load in the shift register, which works reliably for various orders of the MLS sequence.

Older MLS systems were based on hardware shift registers for the generation of the excitation signal, but in this case a completely software implementation allows for the creation of MLS signals of any order and with arbitrary taps positions. Furthermore, the software implementation allows for the creation of pre-filtered MLS signals, following the original suggestion of Mommertz [13]. The creation of such pre-filtered signals is useful for compensating the uneven frequency responses of some elements of the measurement chain, such as the microphones, or for improving the S/N ratio in certain frequency bands: more details on how to use this technique along with the efficient creation of inverse filters are given in [6].

In any case, the response of the system is sampled simultaneously with the generation of the excitation signal, and the wanted impulse response is obtained by cross-correlation of the excitation signal with the system's response. This can be made very efficiently (and with very little computational noise) with the well known Fast Hadamard Transform (FHT), as originally suggested by Alrutz [10], and clearly explained by Ando [14] and Chu [11]. The Ando formulation was employed here, as

described in [15]. The process is very fast, because the transformation happens "in place", and requires only additions and subtractions. The computations are done in floating point math, and the computed IR is then converted in 16-bit integer format by scaling to a pre-defined scale factor, so that the absolute amplitude information is preserved.

This measurement procedure solves the problems already encountered with other software MLS implementations [6], where both the absolute time delay and amplitude of the measured impulse responses were lost: in fact, in car audio applications it is important to maintain the information of the absolute delay and amplitude, for making it possible to combine properly the responses of the various loudspeakers.

3.1.1 Hardware Implementation

The practical implementation of the measurement system was possible thanks to a multi-channel sound editing software (Cool Edit Pro), loaded on a PC with a 4-channels full-duplex sound board (Wave/4). In its multitrack mode, CoolEditPro is capable of synchronous playback and record, and thus maintains the absolute phase information between the input and output of the system under test. This makes it possible to generate a multichannel excitation signal which automatically feeds the four sections of the sound system in sequence, while the 4 microphone signals are simultaneously sampled. After this acquisition phase (typically shorter than 1 minute), the sampled signals are processed through a specially developed multiple MLS deconvolver, which will be described later.

The excitation signal is in this case made of two stereo sound tracks. The first is played through one stereo port of the sound board, connected to the frontal loudspeakers, and the second sound track is played on the rear loudspeakers through the second stereo port. This way, no relay switcher is required, but it is needed that the 4 power amplifiers are perfectly matched.

The two input stereo ports are connected through proper calibrated preamplifiers with two binaural microphones, placed on dummy heads at the front and rear seat of the car.

Fig. 5 shows a global layout of the measurement system. Actually all these units are placed in a single case, from which a single power cable connects to the DIN connector of the sound system, so that the connection of the measurement system with the car is fast and easy.
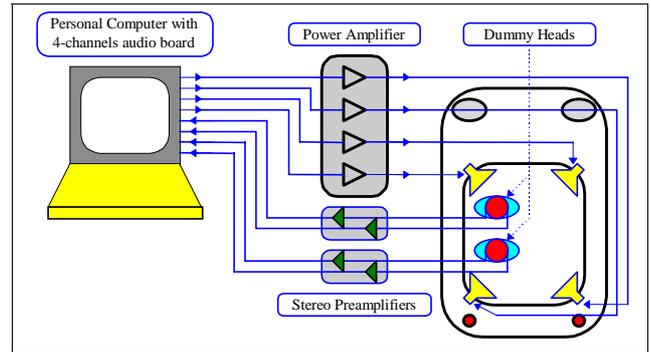

Fig. 5 - Layout of the measurement system

The power amplifier, the microphone preamplifiers and the power supplies were built in the ASK laboratory, whilst the dummy head microphones are actually two Sennheiser MKE 2002. In the future, it is planned to substitute them with other dummy heads, as Sennheiser does not produce them anymore.

After the deconvolution is performed, the results are 2 standard sound files (WAV format) containing, for each dummy head, the 4 binaural impulse responses of the 4 sections of the sound systems.

These results files can therefore be processed with an integrated analysis software, made specifically for this purpose, which also controls CoolEditPro during the acquisition. In this way, the 1/3 octave spectra and the other objective criteria can be displayed at will and stored to disk.

3.1.2 Software Implementation

The first module developed is the MLS generator. Fig. 6 shows the user's interface of this module, which asks for the MLS sequence to be generated and for the amplitude of it.
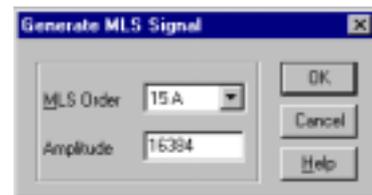

Fig. 6 – User's interface of the MLS generator

For making an automated measurement following the scheme of fig. 5, CoolEditPro is opened in its multi-track mode, with 4 stereo waveforms on the display. On the upper two waveforms the excitation signals are generated: 32 repetitions of the MLS sequence of order 15 are subsequently placed on the first left and right channels (front), and then on the second left and right channels (rear), so that only one channel is excited at each time. The whole excitation sequence is thus 32*32767*4 samples, which at 44.1 kHz gives a total length of 95.1 s.

Pressing the REC button, CoolEditPro automatically plays the first two tracks, which are not checked for recording. During the recording, a real-time display of the signals being recorded is given, as shown in fig. 7.
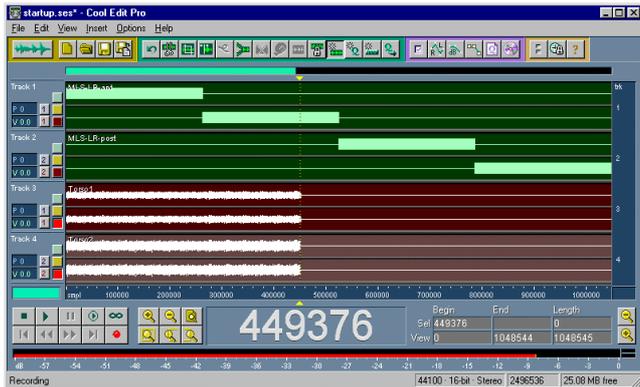
Fig. 7 – Cool Edit Pro during the simultaneous
Playback & Record

Each of the two dummy-head recordings, which are on the tracks 3 and 4, is processed through the Multiple Deconvolver module, as shown in fig. 8. The first of the 32 repetitions of the MLS sequence is discarded, as during its playback the system has not reached the steady state yet, and then the subsequent 31 repetitions are analyzed; this is repeated 4 times, corresponding to the excitation of the 4 loudspeaker sections. Only te first 8192 values of each impulse response are retained. The deconvolver places the results in the Windows Clipboard, from where CoolEditPro retrieves them into a new waveform file. Fig. 9 shows the results of a typical measurement inside a car: for each dummy head position, 8 impulse responses are obtained.
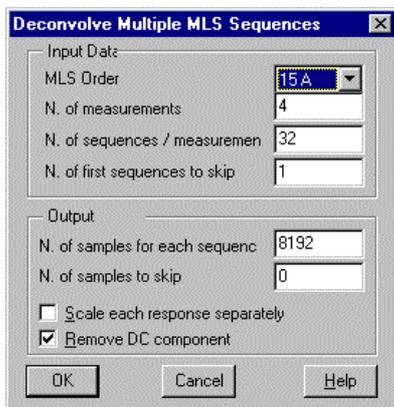

Fig. 8 – Interface of the MLS deconvolution module

Fig. 9 - A typical measurement result inside a car

After the measurements are completed, a Visual Basic program can be used for analyzing the results. The user can choose which loudspeaker section to analyze simply by clicking over it, or pressing the selector buttons above. Through them, also the combined responses can be computed, selecting two of the 4 loudspeakers, or all four. It must be noted that the summation between the responses of the corresponding front and rear loudspeakers is made as pressure summation (because usually the same excitation signal is applied to the front-left and to the rear-left loudspeakers). Instead, the left and right sections of the system are summed energetically, because it is assumed that the two channels are almost uncorrelated.

The spectra of the selected responses are displayed either in narrow-bands, as shown in fig. 10, or 1/3 octave bands. The vertical axis is calibrated in dB-SPL through the application of an offset based on a preliminary measurement, during which a 94-dB reference source is applied on the microphones.
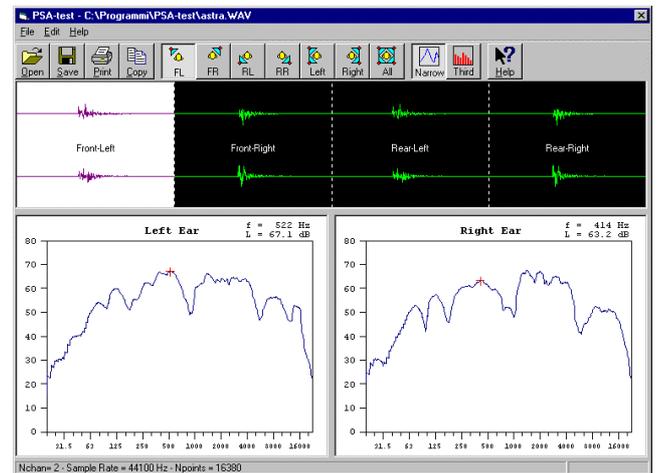

Fig. 10 – Post-processing window, narrow-band mode of a
single loudspeaker

The excitation MLS signal is inherently white, so no frequency equalization is required when looking at the narrow, constant bandwidth spectrum. Instead, with 1/3 octave spectra, the "whiteness" of the excitation signal causes a systematic increase of high-frequency levels, with a slope of 3 dB/octave. This effect is compensated for by the software, which leaves untouched the SPL value at 4 kHz, adjusting the others so that the global shape remains the same as for the narrow-band display. This means that the measured response can be compared both with an FFT spectrum obtained with white noise excitation, and with an 1/3 octave spectrum obtained with pink noise excitation having the same SPL at 4 kHz. The crossing-frequency of 4 kHz was chosen because it was verified that a white signal and a pink signal, which have the same total RMS amplitude in the frequency range 20Hz – 20 kHz, exhibit 1/3 octave band spectra which have the same value at 4 kHz.

The excitation signal is amplitude controlled, so that the total RMS voltage at the amplifier output is exactly 2V on a calibrated 4 Ω resistive load. This gives a nominal power of 1 W broadband white. This makes it possible to compare different sound systems not only in terms of spectral flatness, but also of absolute SPL produced. Thus, the measurement results carry an information not very dissimilar to the concept of Strength for concert halls, with the difference that in this case the measurement takes into account the efficiency of the loudspeakers and of their coupling with the car compartment.

3.2   Computation of the Inverse Filters

3.2.1   Minimum-Phase Spectral Equalization

This module creates an inverse filter, starting from a generic Impulse Response, and inverting only the Minimum Phase component of it. This means that the Magnitude of the Transfer Function is inverted, and the inverse filter, applied to a signal containing the effect of the original impulse response, perfectly equalizes it in terms of spectral flatness.

The algorithm employed is the one proposed by Neely and Allen [7]: the FFT of the original impulse response is taken, then the frequency response is properly smoothed, and the phase of the complex spectrum is thrown away. At this point an IFFT yields the required inverse filter.

It must be noted, anyway, that the non-minimum phase component is left not-inverted: as this component carries the reverberation, the equalized signal has a proper spectral frequency response   (depending on the degree of smoothing applied), but it still suffers of echoes or reverberation present in the original signal.

After having selected the impulse response to invert, activating the "Flatten Spectrum" module the Dialog Box shown in figure 11 appears. With this dialog box it is possible to select some parameters, which affect the computation of the Inverse Filter.
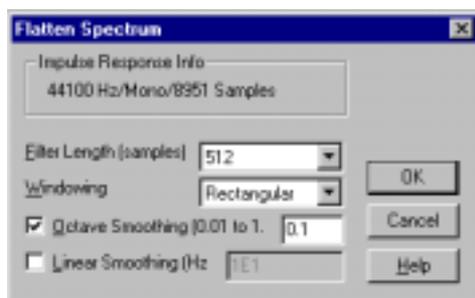

Fig. 11 – Input parameters for the "Flatten Spectrum" module

First of all there is the Filter Length: the larger this value, the most accurate is the frequency equalization. Then we have the Windowing function to be used before the FFT is applied to the original data.

The "minimum phase" inverse filter requires a certain amount of "frequency smoothing", to ensure that its length fits inside the required filter length: so we have to choose between Octave Smoothing or Linear Smoothing, specifying in either case the proper smoothing bandwidth.

3.2.2   Time-Domain Least-Squares Inversion

The creation of a causal, stable inverse filter for a given mixed-phase impulse response is generally an impossible task, and only approximate solution are available for this problem. This argument was deeply investigated by Mourjopoulos [5], particularly with the goal of obtaining a digital equalization of the room's response. Among the various techniques explored by him and others, the straightforward least-squares technique was chosen, as clearly explained in the Mourjopoulos paper.

A least-square problem is set up, requiring that the unknown inverse filter, g(t), when convoluted with the original impulse response h(t), produces as result a delayed Dirac's delta function. After proper formulation, a standard linear equation system is formed, in which a square matrix [R] is multiplied by the unknown vector {g}, producing the known terms vector {k}.

$$[R] \cdot \{g\} = \{k\} \qquad (9)$$

It turns out that the latter is simply a delayed, time reversed copy of the original impulse response h(t), while the [R] matrix is a Toepltiz matrix, in which the first row is simply the autocorrelation function of the original signal, and each of the subsequent rows is obtained by circular permutation (1 sample to the right) of the upper row.

Fortunately, the solution of a Toeplitz system is possible with standard algorithms, included in any library of scientific functions. This way only the first row of the matrix has to been generated and stored, and the computation is thus possible also for very long IRs on standard computers.

The successful creation of an inverse filter is anyway not always easy: if the frequency response of the original impulse response has very deep valleys, these turn out in strong peak of the inverse filters, which tends to become quite long. Furthermore, if the high frequency part of the spectrum of the original signal has low energy (this is common when an anti-aliasing filter is used), the inverse filter has maximum amplitude at these inaudible frequencies, and a very low gain at the lower frequencies, so that the filtered signal can fall below the numerical noise floor.

For these reasons a certain pre-processing of the original impulse response is usually needed before attempting its inversion, but in this field only the experience and skillness of the user can be useful. Anyway, the inverse filter can always be quickly checked, applying it to the original impulse response or to other raw signals, by means of the convolution module. Often our ears are the better instrument for checking if the inverse filter is acceptable, or if it produces intolerable artifacts.

Invoking the Transform / Inverse Filter module, a dialog box appears, as shown in fig. 12, asking for the length of the inverse filter (in samples) and for the wanted delay of the signal (this is necessary for ensuring that the

inverse filter is causal, and is usually chosen around half the length of the filter). Morjopoulos suggests a filter length approximately equal to the reverberation time, although in some case shorter filters works reasonably well.
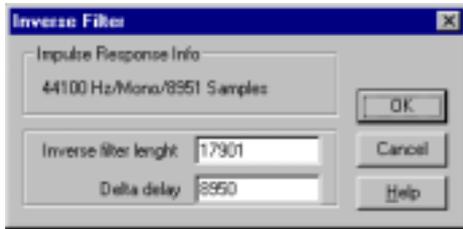

Fig. 12 – Interface of the Mourjopoulos inverse filter

Then the solution of the Toeplitz problem is started: the computation time increases with the square of the filter length, because it is an iterative algorithm, which loops for a number of times equal to the filter length, and the number of operations inside each cycle is also proportional to the number of unknowns. After some seconds, anyway, a dialog box is displayed, informing the user that the inverse filter has been successfully stored in the Windows clipboard.

3.2.3 Frequency-Domain Regularization

This algorithm for computing an inverse filter was developed by Kirkeby [8] specifically for creating binaural filters for the synthesis of virtual sound sources through a pair of close loudspeakers (stereo dipole). It is a classical regularization technique, in which a small, constant value is added to each datum prior of taking its reciprocal, to avoid that the result is too large when the original datum is very little.
The Kirkeby formulation, anyway, is very clever: in fact the regularization is applied in the frequency domain, after the original impulse response is transformed through an FFT, and preserving the phase information. Each (complex) spectral datum is inverted through the following formula:

$$G(f) = \frac{\text{Conjg}[H(f)]}{|H(f)|^2 + r} \qquad (10)$$

in which r is the regularization parameter, which has to be chosen by the user. After the frequency-domain inversion, an IFFT transforms back the inverse filter to time domain; an half-length rotational shift is applied at the end.
The processing is very fast, but the user has to make various attempts with different values of the regularization parameter r in order to obtain the better compromise between filter length and inversion accuracy. Fig. 13 shows the dialog box that allows the insertion of the computing parameters.
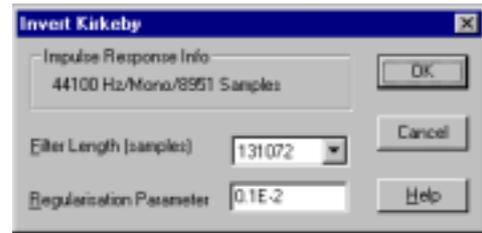

Fig. 13 – User's interface of the Kirkeby inverse filter

The inverse filters obtained with the regularization technique are usually as robust and reliable as the minimum-phase ones, although they retain a considerable capacity of removing also the reverberant tail. Furthermore, by applying a proper regularization parameter, it is avoided that the inverse filter has too high gain at some frequencies, improving the dynamic range and avoiding to push the reproduction system over its linearity limit. As this approach is a sort of compromise between the previous two, anyway, the frequency response is never equalized so well as with the minimum-phase approach, nor the time-domain dereverberation is so effective as with the least-squares formulation.

3.3 Real-Time Convolver

The convolution module developed for CoolEdit can be used in two different ways: it can make a batch processing on existing waveforms, which are filtered and saved again on the hard disk. Or it can do a real-time convolution, sending directly the result to the audio board for listening, but without saving it.
The second mode is useful for listening quickly at the results, after which the process can be started again in the first mode, saving the result.
The actual implementation of the algorithm makes use of the select-save algorithm, as described in [16]. The convolution of a continous input signal x(t) with a linear filter characterized by an impulse response h(t) yields an output signals y(t) by the well-known convolution integral:

$$y(\tau) = x(\tau) \otimes h(\tau) = \int_0^\infty x(\tau - t) \cdot h(t) \cdot dt \qquad (11)$$

When the input signal and the impulse response are digitally sampled (t = i·Dt) and the impulse response has finite length N, such an integral reduces to a sum of products:

$$y(i) = \sum_{j=0}^{N-1} x(i - j) \cdot h(j) \qquad (12)$$

The sum of N products must be carried out for each sampled datum, resulting into an enormous number of multiplications and sums! These computations are better made with float arithmetic, to avoid overflow and excessive numerical noise. For these reasons, the real time *direct true convolution* is actually restricted to impulse response lengths of a few hundredths points, while a satisfactory description of a typical concert hall transfer

function requires at least N=65536 points (at 44.1 kHz sampling rate).

However, the convolution task can be significantly simplified performing FFTs and IFFTs, because the time-domain convolution reduces to simple multiplication, in the frequency domain, between the complex Fourier spectra of the input signal and of the impulse response. As the FFT algorithm inherently suppose the analyzed segment of signal to be periodic, a straightforward implementation of the Frequency Domain processing produces unsatisfactory results: the periodicity caused by FFTs must be removed from the output sequence.

This can be done with two algorithms, called *overlap-add* and *select-save* [16]. In this work the second one has been implemented. The following flow chart explain the process:
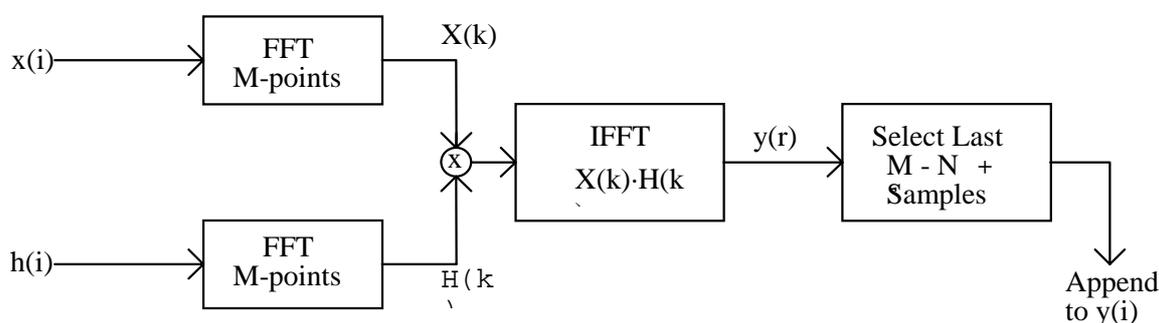
A zero-delay algorithm has been proposed by Gardner [18], and is actually implemented (and patented) in hardware DSP convolvers [19], but at the expense of a noticeable increase in the number of computations required. For these reasons, a new algorithm is under development, which will make it possible to listen immediately at the convoluted signal as soon as the "listen" button is pressed, but maintaining the computational needs at a level only slightly larger than the original Select-Save algorithm.

Here is how the user accesses the convolution module: first the original signal is loaded in CoolEdit, from a hard disk file or a CD. At this point the Transform-Convolver option is chosen, and the dialog box shown in Figure 15 appears. Here the user can select up to two stereo IRs, loading them from the clipboards or from files. How the input channels

Fig. 14 – flow-chart of the Select-Save algorithm

As the process outputs only M-N+1 convolved data, the input window of M points must be shifted to right over the input sequence of exactly M+N-1 points, before performing the convolution of the subsequent segment.

The tradeoff is that FFTs of lenght M>N are required. Typically, a factor of 4 (M=4·N) gives the better efficiency to the select-save algorithm: if N is 65536 (2^{16}), one need to perform FFTs over data segments of lenght M=65536·4=262144 points! This requires a very large memory allocation, typically 1 Mbyte, for storing the input sequence or the output signal. The overall memory requirement for the whole select-save algorithm is thus several Mbytes! This is the reason for which this algorithm cannot be easily implemented on widely diffused low cost DSP boards, which are not equipped with such a storage capability. Only dedicate DSP units can afford this task [17].

On the other hand, today's personal computers have plenty of RAM, and enough floating-point computing power for keeping-up in real time even with very long IRs: a Pentium Pro running at 200 Mhz can convolve a mono input signal (sampled at 44.1 kHz) with a mono IR of more than 800,000 points, and with a stereo (or binaural) IR of more than 200,000 points/channel. Further improvements will be possible with the new Pentium II, and the future is pink, because the computing power is doubling each year…

The only problem with the Select-Save algorithm is that it causes a substantial delay, typically 4 times the IR length.

are processed through the filters is chosen by pressing the push-buttons on the right: in the most complex case, a stereo input signal is passed through two stereo IRs, so that each input signal is convolved with two IRs, and the resulting "Left" and "Right" channels are mixed with the corresponding results of the other input channel. This is particularly useful when implementing a cross-talk cancellation system, for listening at binaural material through a pair of loudspeakers, as explained in chapter 2.

The lower-right section of the dialog box allows to start the processing, and the insertion of an overall scale factor, for adjusting the playback level during real-time listening (this control can be accessed also during the playback); the overall scale factor is disabled during batch processing, because in this case the result is always normalized to the maximum value before conversion to the integer format.

Being the entire process made with floating-point math, the computed result has a sonic quality and a dynamic range far better than those achievable with fixed-point DSP units.
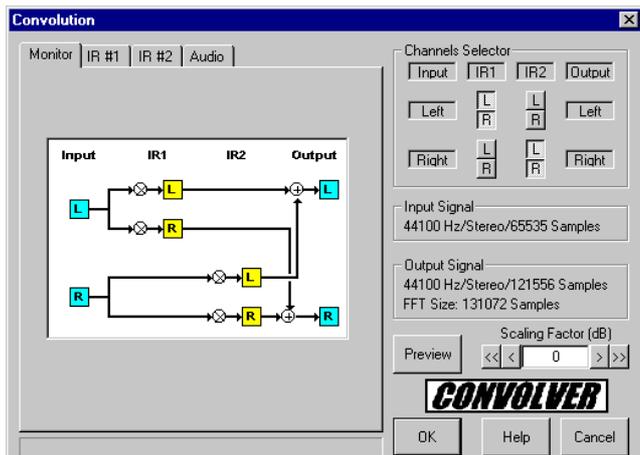
Fig. 15 – User's interface of the real-time convolver module.

4 DIGITAL EQUALIZATION EXPERIMENT

Making use of the software tools described in the previous section, the spatial equalization was attempted in a special test car, equipped with 4 independent sound system. These can be switched instantaneously through pushkeys. Two different "virtual" sound spaces were reproduced inside the car: a famous theatre and a high quality stereo listening room.

The performances of the equalizing filters were checked by subjective blind preference tests.

4.1 Measurements in the "Virtual" Sound Spaces

First of all, the "ideal" impulse responses of two different listening environments were measured: a famous theatre and an hi-fi listening room. The first was the Teatro "La Scala" in Milan, Italy, and the second the test room of ASK Industries, in Reggio Emilia, Italy. The measurements were performed with the MLS software described in sect. 3.

Fig. 16 reports the binaural response measured at "la Scala" in the middle of the stalls, while the loudspeaker was placed on the left of the stage (from the listener's point-of-view). The response with the loudspeaker on the other side was obtained simply by interchanging the left and right impulse responses, because the listening conditions are assumed to be perfectly symmetric.
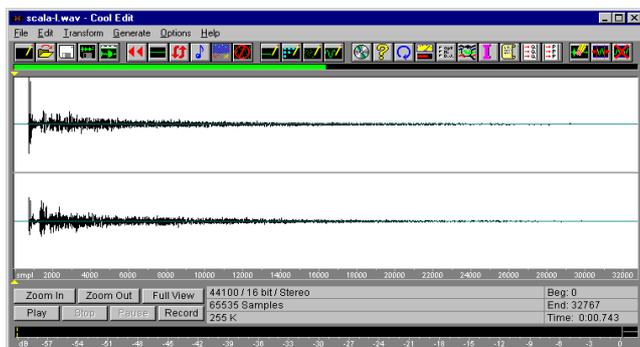

Fig. 16 – Binaural Impulse Response measured in the Teatro "La Scala" in Milan, Italy

The measurements in the ASK test room were performed employing the pair of high-quality, self-built loudspeakers already fitted in it and the Sennheiser dummy head.

Fig. 17 shows both the binaural impulse responses obtained by the two loudspeakers. In this case the loudspeaker response is considered part of the ideal listening system, and thus it was not removed from the measured IRs.
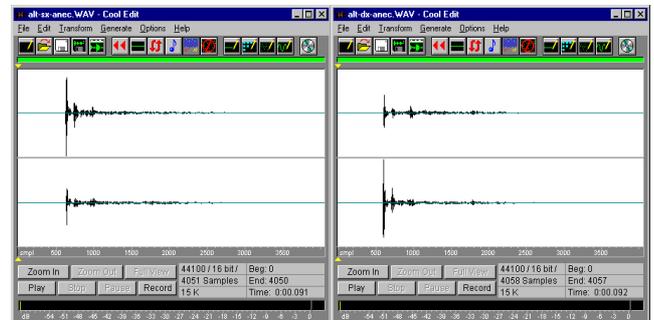

Fig. 17 – Binaural Impulse Responses of the ASK test room

4.2 Measurements in the Car

A test car, equipped with 4 different, switchable sound systems, was used for the listening tests. They were numbered from 1 to 4, following a ranking order based on the quality of the components: this means that the system #1 is the better one, and the system #4 is the one with the smallest, low quality loudspeakers. For each sound system, a set of 4 impulse responses was measured: fig. 18 reports the two binaural IRs for the system #1.
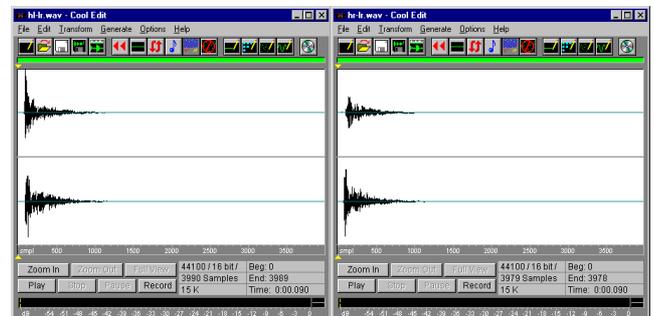

Fig. 18 – Binaural Impulse Responses of the test car – sound system #1

Also these measurements were performed with the same dummy head (Sennheiser MKE2002) as the previous ones, as shown in fig. 19.

Fig. 19 – The dummy head inside the car

4.3 Creation of the Digital Filters

At this point, the inverse filters were computed for the 4 sound systems and the two ideal listening environments, producing a set of 8 different cases. The creation of the numerators of eq. 7 was straightforward, making use of the software convolver, and saving the results in 2 new stereo files. Also the calculation of the denominator was made the same way.

The creation of the inverse of the denominator is a delicate point, as already explained in section 2: both the minimum phase and the complete inversion were attempted. In the first case, a 512-point inverse filter was created, with a frequency smoothing of 0.05 octaves and an Hanning time window. In the second case, a 2048-points inverse filter was built, with a delay of 1024 points.

After the computation of the inverse of the denominator, it was applied by convolution to the previously computed numerators, obtaining the required digital equalizing filters f. These filters were then employed again by means of the software convolver, yielding the equalized signals to be compared with the unequalized ones.

As an example, fig. 20 shows the inverse filters which, applied to the sound system #1, transform it in the ASK test room: these filters were computed with the Neely and Allen equivalent minimum phase inversion of the denominator: this way shorter filters are obtained, which correctly equalize the frequency response of the sound system, but which leave the natural reverberation of the car compartment not corrected.

Obviously, also the inverse filters for recreating the theatre sound field were obtained with the same approximation, and the same was repeated also for the other 3 sound systems installed on the test car.
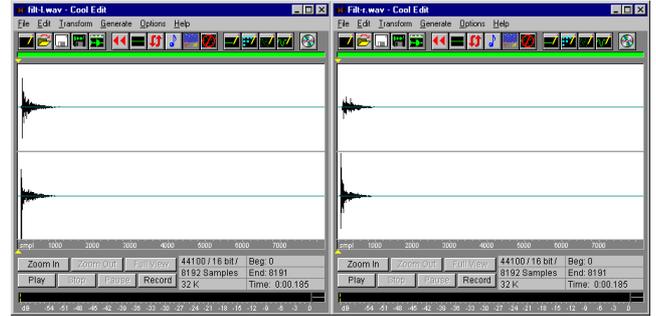

Fig. 20 – Inverse Filters for equalization of the sound system #1, recreating the ASK test room.

4.4 Subjective Evaluation of the Equalization Technique

The analysis of the performances of the digital equalizer was based exclusively on direct subjective tests, obtained with subjects seating inside the test car, listening at various sound samples. These were obtained by digitally transferring some music samples from commercial CDs to the hard disk of the PC.

Each sound sample, having a length of 90 s, was partially filtered through the digital equalizer, making use of the real-time software convolver already described; in more detail, half of the sample was filtered (sometimes the first half, sometimes the second), while the other part was left not filtered. The change between the two parts was very evident for any kind of music samples.

Each subject had then simply to express his preference for the first or the second part of each sample, without knowing which of the two was filtered. The following table summarizes the results of the preference test for the 4 sound systems and the two "virtual spaces":

N.	Virtual Space	Preference for the signals		
		Filtered	Unfiltered	Uncertain
1	"La Scala"	74 %	16 %	10%
2	"La Scala"	65 %	20 %	15%
3	"La Scala"	54 %	22 %	24 %
4	"La Scala"	33 %	45 %	22 %
1	Hi-Fi room	35 %	60 %	5 %
2	Hi-Fi room	28 %	65 %	7 %
3	Hi-Fi room	18 %	68 %	14 %
4	Hi-Fi room	0 %	85 %	15 %

It is clear how the digital equalizer was preferred only in the first three cases, with the theatre as "virtual space" and with the better loudspeakers. As the quality of the sound system decreases, the digital equalizer looses preference, and with the little, bad loudspeakers of the system #4 the unfiltered signals are always preferred, even with the theatre as "virtual space".

This fact was easily explained: the poorest sound system has a quite uneven frequency response, and the inverse filter tries to compensate for these deficiencies with strong peaks in its equalizing response. This causes a very little

overall gain, and the signal comes out at a very low level, as the filtered signal's magnitude is always scaled to fit within the 16-bit constraint of the D/A converter. The subsequent heavy amplification pushes the system over its linearity system.

It was concluded that the novel technique can be applied only to good quality systems, which already have an almost flat frequency response, and which can be driven by strong signals without causing distortion.

Furthermore, it is clear how the hi-fi listening room was not appreciated at all. In general, it resulted clear that the choice of the "virtual space" is critical: the limited number of spaces employed in this first step of the research does not allow for an analysis of the optimal one, but surely this aspect will be explored in the next future.

The subjects were asked to describe briefly the reasons for each preference choice: it turned out that the theatre gives a good effect of enveloping, while the hi-fi stereo system makes the sound to come only from the front (and in some cases from a virtual position even higher than the listener's ears), with a limited width of the sound scene, and globally a poor stereo effect. No one complained about too much reverberation with the theatre, but this is probably due to the particular acoustics of "La Scala", which is quite dry compared to its volume, although the spatial impression is wide.

The results with the hi-fi stereo system are surprising: it is a common believe that the ideal listening conditions are obtained when the sound sources are placed in front of the listener, at +/- 30° from the central direction, in an anechoic environment. Almost any sound purist tries to achieve these conditions in his listening room!

On the other hand, acoustics research about concert halls pointed out the importance of being enveloped by the sound (Ando [14]). This fact explains the great interest about movie-like surround systems, and the revival of the Ambisonics technology, which seemed dead only 5 years ago. The subjective results obtained here confirm the importance of the surround effect, which is naturally present in most car sound systems (due to the loudspeaker placement), and is completely removed by the digital equalization with the hi-fi listening room ideal responses.

Another fact must be kept in mind: as the listeners were seating inside a car, they were mentally prepared to listen at a car sound system. Hearing the sound coming from remote phantom sources in front of them seems unnatural, and the mind reject an auditive experience inconsistent with the information coming from the other senses. In this respect, also the theatre sounded a bit unnatural, because some listeners said that the sound seemed "larger" than expected.

5 COMPARATIVE SUBJECTIVE EVALUATION OF DIFFERENT CARS

Making use of the same processing techniques employed for the direct equalization, it is possible to set up a subjective comparison experiment, with the aim to find the most pleasant sound system. This task is accomplished by measuring the binaural impulse responses in many different cars, and reproducing their acoustic behavior in a proper listening room, where the subject can compare the different sounds and express their judgement. In order to recreate faithfully the listening conditions inside the cars, it is necessary to take in account also the background noise, generated by combination of aerodynamic, rolling and engine borne noises.

For simplifying the collection of data only the driver's position seat was considered. A collection of 9 different cars was tested, and the listening tests were performed in the ASK listening room, already described in the previous section.

5.1 Measurement of the Background Noise

In each of the 9 cars under tests, preliminary measurements of the interior noise, at various speeds, were conducted. The tests were made on an highway, at the three speeds of 90, 120 and 140 km/h.

A Bruel & Kjaer microphone type 4165 was mounted on a torso simulator, placed on the seat at the side of the driver. It was connected, through a B&K type 2131 Sound Level Meter, to a DAT recorder SONY DTC-790. A preliminary 94 dB, 1kHz calibration signal was recorded on each tape. A 20-minutes sound sample was recorded for each car, at each speed.

At the laboratory, the DAT recordings were played back and analyzed through a B&K type 2133 real-time analyzer, and the 1/3 octave spectra were stored to disk and converted into a spreadsheet.

Figure 21 shows some of the measured spectra at the speed of 90 km/h.

The background noise recordings were not used directly for mixing with the auralized signals: instead, they were used as shaping filters for the creation of an artificial background noise, having the required spectrum, as it will be explained at paragraph 5.3.2.
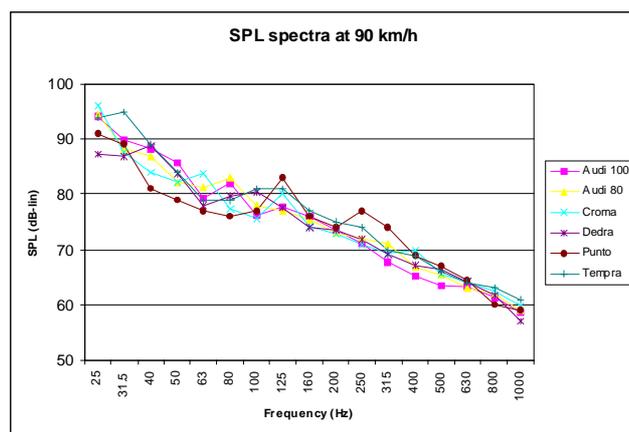

Fig. 21 – Background noise at 90 km/h

5.2 Signal Processing

The signal processing is made of three different steps: convolution, noise superposition and presentation to the subjects (playback).

The first step is accomplished making use of the software convolver. The original signals were samples of various kinds of music, digitally transferred from commercial CDs to the hard disk. Typically, the sound samples chosen for the tests were long between 30 s and 1 minute.

After the convolution was done, an equivalent background noise was generated, making use of the standard features of the sound editing program CoolEdit: we started with the generation of a "spatial-stereo" brown noise, and then we applied a proper frequency filtering, until the calculated 1/3 octave spectrum approached within +/- 1 dB the measured one, at the 90 km/h speed.

At this point the convolved signal was mixed with the background noise, taking into account the overall signal amplitude, in such a way that the absolute Equivalent Sound Pressure Level of the music at the ears of the subject was adjusted to 90 dB-lin, whilst the background noise was perceived with the same SPL as measured inside the car. This level adjustment revealed to be the most time consuming and delicate point of the whole signal processing.

This was also due to the fact that, listening to the reconstructed signals, it seemed that the background noise was too high compared with the music: the capabilities of our brain to concentrate only on the music when driving at a car, and neglecting the environmental noise, make so that anyone remembers his listening experiences on real cars with much lower noise than the reality. For discovering this point, and assuring the proper level calibration, some DAT recordings were made while playing a CD with the car running on the road: listening at such DAT recordings, the same anomalous high background noise is audible, while it was not perceived during the recording. This fact is one of the weak points of the new auralization technique, because it causes a systematic overestimate of the subjective effect of the background noise, compared to the subjective experience reported when driving a real car.

5.3 Subjective Testing System

For automating the process of playing the sound samples and expressing the subjective judgements for each of them, a new software tool was developed. It is a .WAV player, equipped with a graphical interface for collecting the responses to a set of predefined questions. Both the list of .WAV files and of subjective questions are stored in ASCII files, so that the same program can be used for different subjective tests. Each question is expressed as a couple of counter-posed terms (such as PLEASANT-UNPLEASANT), and the listener has to place a marker between them, on a 5-segment scale. So each response is represented by a numeric value, ranging between 1 (left term is more appropriated) to 5 (right term is more appropriated); 3 means that the response is in the middle.

The user can change at any time the sample being played, pause the playback or start it again, come back and change some responses after listening at other samples, and he is left completely free to listen again at the sound samples, or to change the responses, until he is completely satisfied.

Obviously he do not knows at what car each sound sample refers, nor he knows that the signals are artificially produced. Almost no one doubted that the signals were not naturally recorded inside running cars, but some complained about the noise "too loud". Fig. 22 reports the user interface of the subjective testing program.
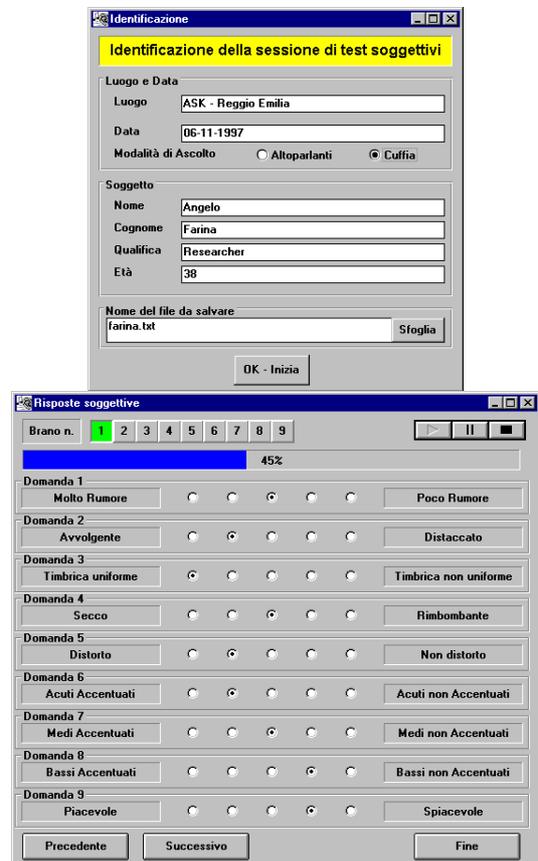

Fig. 22 – User's interface of the subjective test program

5.4 Subjective Results

9 cars were employed for this comparative test. The sound system was always the original one as delivered by the manufacturer. So the subjective test involved the comparison of 9 sound samples, one for each car: The music piece employed was a 40s excerpt from an Elton John's song, containing both voice and instrumental music. 9 questions were posed to the subjects, as reported here:

Q1	Much Noise	Little Noise
Q2	Enveloping	Detached
Q3	Uniform timber	Not uniform timber
Q4	Dry	Reverberating
Q5	Distorted	Not distorted
Q6	Treble enhanced	Treble not enhanced
Q7	Medium enhanced	Medium not enhanced
Q8	Bass enhanced	Bass not enhanced
Q9	Pleasant	Unpleasant

13 subjects were selected among 40 through a preliminary subjective test, as described in [20]: this ensured an high

quality of the subjective results, with low variance and good correlation. The following table reports the average score of each car to the 9 questions, ranked following the score obtained for the question #9; remember that the lower the score, the most pleasant is the judgement:

Car	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Audi 100	3.31	3.00	2.69	2.69	3.92	3.15	2.85	2.85	2.62
Audi 80	2.92	3.08	2.69	2.15	3.54	2.92	2.54	3.46	3.00
BMW 735	3.77	4.00	3.00	1.92	3.31	2.69	2.69	3.92	3.31
Cit. Evasion	2.77	2.54	2.69	3.69	3.15	3.23	3.23	2.31	3.38
Opel Astra	2.69	3.00	3.08	2.69	2.92	3.54	3.00	3.00	3.62
Fiat Croma	2.15	3.54	3.38	2.85	3.00	3.38	2.92	3.46	3.77
VW Passat	2.31	4.15	3.08	2.77	3.08	2.62	2.38	4.08	3.85
Dedra	2.15	3.92	3.31	1.92	2.85	3.08	2.85	4.23	4.15
Punto	2.08	3.69	3.85	2.62	2.31	3.92	2.69	4.31	4.46

Looking at the table, it can be observed how the overall pleasantness ranking is matched with the other subjective factors.

For making clearer these correlations, it is possible to make plots of the interdependence of pairs of subjective responses. Fig. 23 reports the diagram, which relates these two subjective parameters, and it is clear how the noise level is clearly negatively correlated with the acoustic quality. A linear trend is evident for 8 of the 9 cars, and the only exception is the BMW-735, which has very little noise, but nevertheless does not have a proportionally high quality judgement. Probably for this car other subjective parameters are more significant than background noise in explaining the overall quality score.
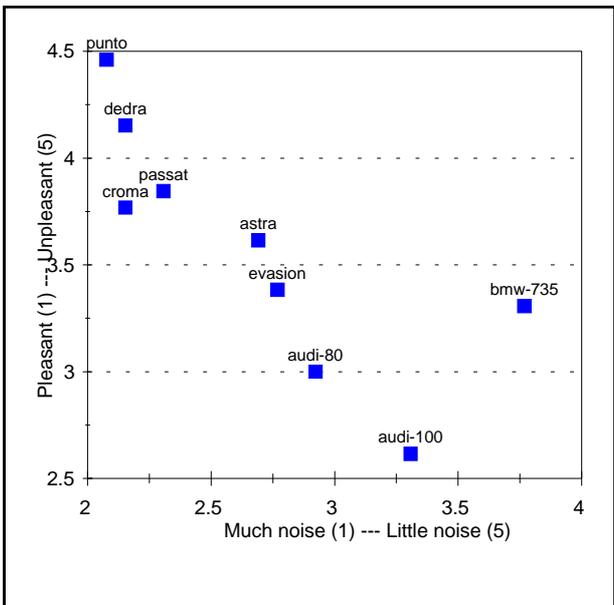

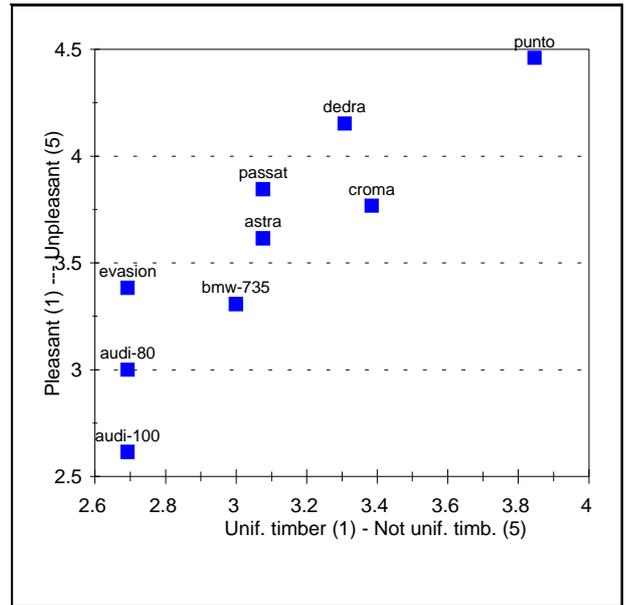Fig. 23 – Correlation between average responses at questions #1 and #9.

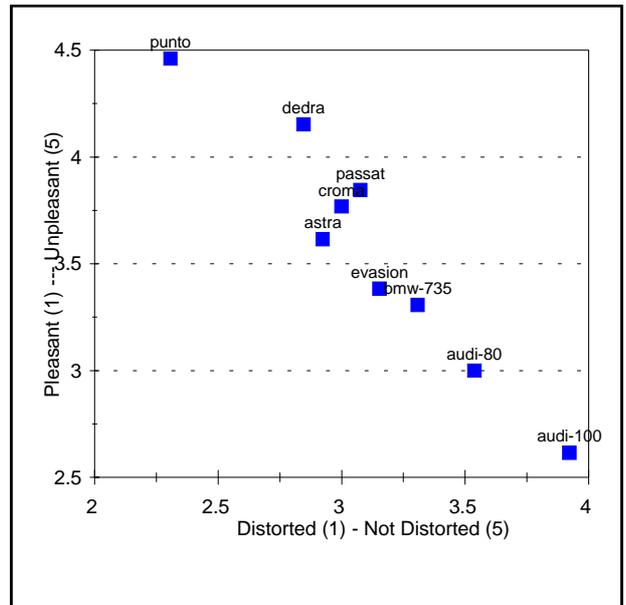Fig. 24 – Correlation between average responses at questions #3 and 9.

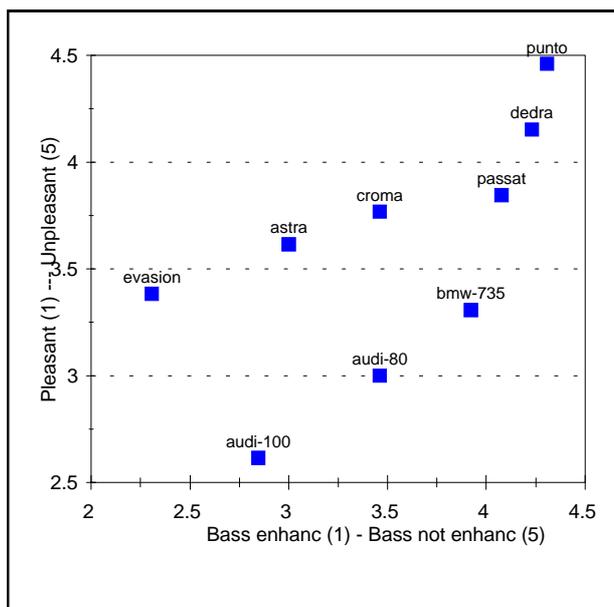Fig. 25 – Correlation between average responses at questions #5 and #9.

Fig. 26 – Correlation between average responses at
questions #8 and #9.

Other interdependencies are shown on figg. 24, 25, and 26: they reports the subjective responses which resulted well correlated with the global preference. The other subjective responses, instead, gave poor correlation with it.

Fig. 24 shows how the timbre uniformity is positively correlated with preference, although a certain spread is evident. Fig. 25 demonstrates how the distortion adversely affects the preference, with a quite good linearity of the plot, and fig. 26 relates the presence of bass frequencies with a general improvement of the preference, although in that case the spread is quite high. In this case a two-dimensional explanation can be attempted: in fact, the three cars with lower noise (Audi-100, Audi-80 and BMW-735) appear to lye on a lower line, while the other cars are grouped on an upper line: this means that a lot of bass is required particularly when the noise is high, for "covering" it.

Obviously at least these 4 subjective parameters interact each other in influencing the overall preference: the evaluation of the mutual effect of them is possible only with an advanced statistical technique, the factor analysis (Herman [21]), which in this case was not applied yet.

Nevertheless, the above information is enough for deducting some interesting conclusion, and for piloting the prosecution of the research.

6   CONCLUSIONS

The problem of inverting a mixed-phase impulse response has been solved with three different techniques, but from the results it came out that in this particular case the partial inversion of the equivalent minimum phase component gives better results than the complete least-squares inversion.

The choice of the "ideal environment", which has to be simulated inside the car, has revealed to be critical: on one hand the sound field has to seem natural for a car compartment, while on the other hand the possibility to give the impression of being in a large, reverberant space is usually appreciated by the listeners. Probably in the final commercial unit, a set of various sonic environments will be available, so that each listener can choose the sound field more suited to his tastes and to the musical piece being played.

Also artificial spaces can be simulated, making use of room acoustics computer programs equipped with auralization extensions (Farina, [22]). Anyway, for the success of the method, it is important that either experimental or computed impulse responses are including the HRTFs of the same dummy head used for the preliminary measurements inside the car.

It must be remarked that the new equalization technique is based on the linearity assumption: it cannot correct any kind of distortion, and furthermore the sound system needs to be already of good quality, for avoiding the risk of being driven out of his dynamic range by the equalized signals. The creation of not-linear numerical filters for the control of the distortion produced by the loudspeakers is currently being investigated in a parallel research [4].

The extension of the equalization technique to multi-channels systems is straightforward: it is quite easy to simulate more than two virtual loudspeakers, and this way an horizontal 5-speakers surround system can be emulated, or even an 8-speakers Ambisonics full-3D system. On the other hand, the use of more than two reproduction channels inside the car makes it possible in principle to obtain the virtual acoustics effect for more than one occupant. The software tools for processing B-format signals (instead of 2-channels stereo) have been already realized by the authors [23], but their application to the spatial equalization of the sound fields has yet to be tested.

Making an inverse use of the spatial equalization makes it possible to reproduce the sound field of various cars in a proper listening room: this exploits the acoustical difference between the cars, as the other factors potentially influent on the subjective results are maintained absolutely constant. Furthermore, the total time required for conducting the experiment is largely reduced in comparison with the traditional technique based on direct binaural recordings. The last advantage is the possibility to evaluate directly any modification to the sound system or to the car compartment, by digital filtering the measured impulse responses, for defining the most preferred characteristics of it.

The subjective results are consistent with already known design criteria, which demands for low noise inside the car compartments, for a reasonably flat frequency response, for absence of distortion and for wide extension of the frequency response toward the low frequencies. Although these requirements are substantially obvious, they confirm that the subjective method, applied to a selected panel of subjects, is a quantitative tool which gives sensitive results, also if the number of subjects is low.

From the results of above the equalization experiments (direct and inverse) it was concluded that the ideal field to

be created inside a car is that of another car: Thus the research will prosecute increasing the panel of subjects, and performing an exhaustive statistical analysis of the results. Furthermore, also the number of cars will be increased, with the goal of finding the "best car system", making extensively use of the new *virtual acoustics* technique.

The goal is to make it available a low-cost DSP unit, to be mounted in any kind of cars, after a proper in-situ programming of the FIR filters, based on MLS binaural measurements of the impulse responses on each car. For hi-end music lovers, even a personalized setup will be available, employing their own head both during the measurement of their preferred ideal sound spaces, and for the car compartment characterization.

The prosecution of the research will make use of numerical models of the sound field inside the compartment, for computing the Head Related Transfer Functions of cars at the first design stage [24, 25].

7   REFERENCES

[1] P.A. Nelson, F. Orduna-Bustamante et al., 'Multichannel signal processing techniques in the reproduction of sound', *JAES vol. 44, n. 11*, 1996 November, pp. 973-989.

[2] P.A. Nelson, F. Orduna-Bustamante et al., 'Experiments on a system for the synthesis of virtual sources', *JAES vol. 44, n. 11*, 1996 November, pp. 990-1009.

[3] A. Farina, E. Ugolotti, 'Spatial equalisation of sound systems in cars by digital inverse filtering', *Proc. of Reproduced Sound 13*, Windermere (GB) 23-26 October 1997.

[4] Bellini, G. Cibelli, E. Ugolotti, A. Farina, C. Morandi - "Non-linear digital audio processor for dedicated loudspeaker systems" - Proc. of International IEEE Conference on Consumer Electronics, Los Angeles, June 1998.

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

[6] 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.

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

[8] O. Kirkeby, P.A. Nelson et al. – "Fast Deconvolution of Multi-Channel Systems using Regularisation" – ISVR Technical Report no. 255, Aousthampton, April 1996.

[9] A. Farina, E. Ugolotti, "Automatic Measurement System For Car Audio Application", *Pre-prints of the 104rd AES Convention*, Amsterdam, 15 - 20 May, 1998.

[10] H. Alrutz, M.R. Schroeder, "A fast Hadamard transform method for evaluation of measurements using pseudorandoom test signals" Proc. of 11th International Congress on Acoustics, Paris, 6, pp. 235-238.

[11] W.T. Chu, "Impulse response and Reverberation Decay measurements made by using a periodic pseudorandom sequence", *Applied Acoustics*, vol. 29, pp. 193-205 (1990).

[12] J. Vanderkooy, "Aspects of MLS measuring systems", JAES vol. 42, n. 4, 1994 April, pp. 219-231.

[13] E. Mommertz, S. Muller, "Applying the inverse fast Hadamard transform to improve MLS-measurements", Proc. of ICA95 (International Conference on Acoustics), Trondheim (Norway) 26-30 June 1995, vol. IV, pp. 127-130.

[14] Y. Ando, 'Concert Hall Acoustics', *Springer-Verlag*, Berlin 1985.

[15] A. Farina – "MLS impulse response measurements for underwater bottom profiling" – Proc. of 4th European Conference on Underwater Acoustics, Roma, 21-25 September 1998.

[16] Oppheneim A.V., Schafer R., *Digital Signal Processing* - Prentice Hall, Englewood Cliffs, NJ 1975, p. 242

[17] Connoly B. - "A User Guide for the Lake FDP-1 plus" - Lake DSP Pty. Ltd, Maroubra (Australia) - Sept. 1992.

[18] W.G. Gardner – "Efficient convolution without input-output delay", JAES vol. 43, n. 3, 1995 March, pp. 127-136.

[19] A. Really, D.S. McGrath, "Convolution processing for realistic reverberation", Pre-Prints of the 98th AES Convention, 1995 February 25-28, Paris

[20] A. Farina, E. Ugolotti, "Subjective comparison of different car audio systems by the auralization technique", Pre-prints of the 103rd AES Convention, New York, 26-29 September 1997.

[21] Herman H.N., 'Modern Factor Analysis', *University of Chicago Press*, Chicago, 1968.

[22] A. Farina, 'Auralization software for the evaluation of a pyramid tracing code: results of subjective listening tests', *ICA95 (International Conf. on Acoustics)*, Trondheim (Norway) 26-30 June 1995.

[23] A. Farina, E. Ugolotti, "Software Implementation Of B-Format Encoding And Decoding", *Pre-prints of the 104rd AES Convention*, Amsterdam, 15 - 20 May, 1998.

[24] E. Granier, M. Kleiner et al., 'Experimental auralization of car audio installations' *Journal of the Audio Engineering Society*, vol. 44, n. 10, 1996 October, pp.835-849.

[25] A. Farina, E. Ugolotti - "Numerical model of the sound field inside cars for the creation of virtual audible reconstructions"- DAFX-98 Conference, November 19-21, 1998 Barcelona, Spain