

Spatial equalisation of sound systems in cars by digital inverse filtering

Angelo Farina (*), Emanuele Ugolotti (**)

- (*) Dipartimento di Ingegneria Industriale, Università di Parma,
Via delle Scienze - 43100 PARMA
tel. +39 521 905854 - fax +39 521 905705
E-MAIL: farina@pcfarina.eng.unipr.it
HTTP://pcfarina.eng.unipr.it
- (**) ASK Automotive Industries, via Fratelli Cervi n. 79,
42100 Reggio Emilia
tel. +39 0522 388311 - fax. 0522 388499
E-MAIL: tec_ask@xmail.ittc.it

Abstract

This paper describes the theory and an experimental application of digital inverse filtering applied to the correction of the response of sound systems in car compartments. The aim of this correction is not simply to equalise the frequency response of the system: by a proper implementation of the inverse filters, it is possible to create a transaural system, which makes it possible to move the “virtual position” of the sound sources. So the listening conditions for the driver are made similar to those originally intended for stereo reproduction, with the two sources at the same height of the ears and symmetrically located at +/- 30° from the frontal direction.

Furthermore, it is possible to create a “virtual acoustics” system, which substitutes the original sound field inside the car with a very different one, coming from measurements taken in brilliant concert halls, or even from numerical simulations of non-existing spaces.

This paper is divided in two sections: first a new set of measurement and signal processing software tools are presented, along with the theory which enables the creation of a virtual sound system through them. Then the possibility to implement in practice such a virtual sound system inside a car is evaluated, with a subjective listening test of the results.

1. Introduction

The possibility of using digital signal processing units for performing the equalisation of a sound system is widely diffused both for audio-pro applications and for consumer hi-fi home systems. Now some units are available also for car-audio systems: but this particular environment claim for different solutions and performances, making the design of these units quite different from the other applications.

In fact 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.

For these reasons, a digital equalisation of a car audio system is not 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 case, the digital filter is 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 [1,2].

In the present work a global approach to the problem is undertaken, making use of the most recent auralisation techniques. In this way, it is possible to “remove” the unwanted sound field characteristics, and substitute to them a completely new sound field, possibly coming from an appreciated concert hall, or from a high-end audiophile listening room. At the present stage, only two-way systems are considered both for the car audio system and for the “ideal room” system: this limits the virtual reconstruction to only one position in the car. This fact can anyway be acceptable, because it is well known that for nearly 80% of the time during which the sound system is used, only one person is in the car: the driver.

The system is based on 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.

The paper presents the details of the theory needed for the creation of the digital filters. Then an application example follows, in which the construction of the digital filter is made in a test car equipped with 4 different sound systems.

2. Theory

Fig. 1 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 unsymmetrical, being present an unavoidable different delay along them.

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).

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. 2. 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.

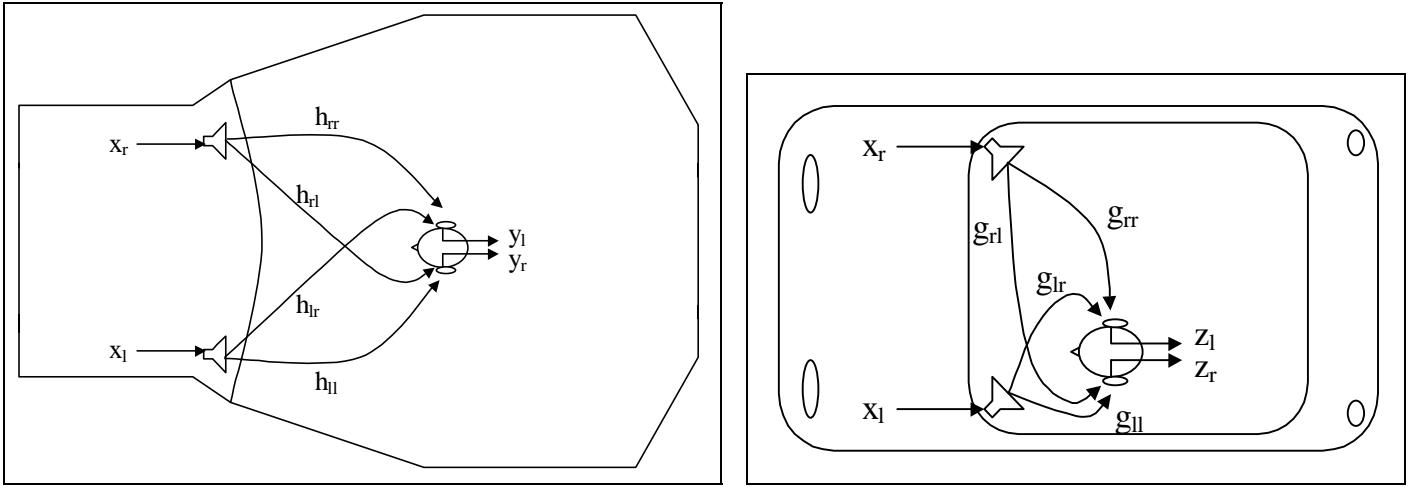


Fig. 1 – Ideal listening conditions in a theatre (left) and effective listening conditions inside a car (right)

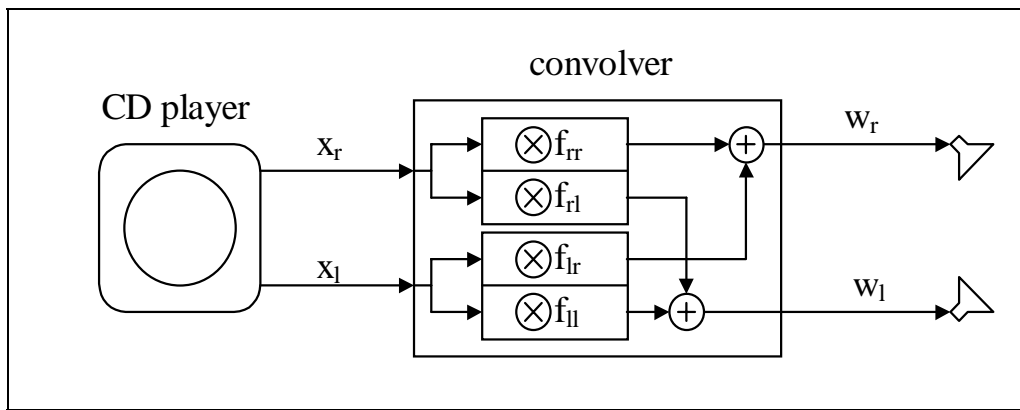


Fig. 2 – Block diagram of the convolver

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:

$$\begin{aligned} y_l &= x_l \otimes h_{ll} + x_r \otimes h_{rl} \\ y_r &= x_l \otimes h_{lr} + x_r \otimes h_{rr} \end{aligned} \quad (1)$$

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

$$\begin{aligned} z_l &= x_l \otimes g_{ll} + x_r \otimes g_{rl} \\ z_r &= x_l \otimes g_{lr} + x_r \otimes g_{rr} \end{aligned} \quad (2)$$

Now we introduce the digital equaliser, following the scheme shown in fig. 2, 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:

$$\begin{aligned} w_l &= x_l \otimes f_{ll} + x_r \otimes f_{rl} \\ w_r &= x_l \otimes f_{lr} + x_r \otimes f_{rr} \end{aligned} \quad (3)$$

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:

$$\begin{aligned}
y_l &= w_l \otimes g_{ll} + w_r \otimes g_{rl} \\
y_r &= w_l \otimes g_{lr} + w_r \otimes g_{rr}
\end{aligned} \tag{4}$$

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:

$$\begin{aligned}
x_l \otimes h_{ll} + x_r \otimes h_{rl} &= (x_l \otimes f_{ll} + x_r \otimes f_{rl}) \otimes g_{ll} + (x_l \otimes f_{lr} + x_r \otimes f_{rr}) \otimes g_{rl} \\
x_l \otimes h_{lr} + x_r \otimes h_{rr} &= (x_l \otimes f_{ll} + x_r \otimes f_{rl}) \otimes g_{lr} + (x_l \otimes f_{lr} + x_r \otimes f_{rr}) \otimes g_{rr}
\end{aligned} \tag{5}$$

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:

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

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

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

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 [3].

A software module which implements the Mourjopoulos least-squares approximate inversion has been already developed by one of the authors [4], 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 Neely and Allen approach [5]; another software module is available for this task [4]. In this second way a simple frequency domain equalisation is performed, and the all-pass component (which carries the reverberation) is left unequalised, the complete filters f produced are more stable, and no audible artefact is introduced. Both these modules, along with the fast convolution module, will be briefly described in the next chapter

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.

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 specialised 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 favourable 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 shareware program CoolEdit (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 5 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 an inverse filter of a given impulse response, both with the Mourjopoulos complete inversion and with the Neely and Allen minimum-phase inversion;
- 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.

Fig. 3 shows the user’s interface of the convolution module, which closely resembles the computational scheme shown in fig. 2.

The capabilities of these software tools substantially exceed those of hardware-based measure instruments and filtering units: in fact there is no limit in the order of the MLS signal, making it possible to measure incredibly long impulse responses, and also the subsequent processing of them have no limit about their length, apart of the memory available on the computer. Furthermore, the computations are always made with floating-point math, and only at the end the final results are converted to 16-bits integers, through a re-scaling process which maximises the dynamic range just prior of the D/A conversion.

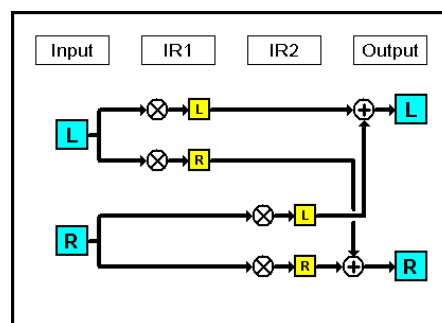
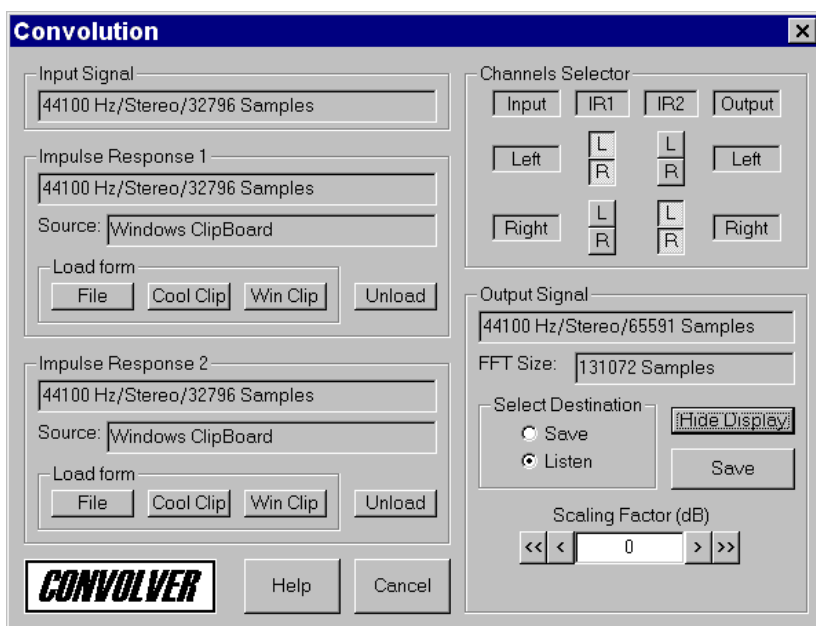


Fig. 3 – User’s interface of the real-time convolver module.

Also the real-time performances of the convolver are more than adequate: 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.

A low-cost, portable PC was used for this research, equipped with a Pentium 166 processor and 32 Mbytes of RAM: a Sound Blaster 32 audio board was inserted into it, and used both for the preliminary impulse response measurements and for playing back the processed signals during the listening tests.

For the impulse response measurements inside cars, an MLS sequence of order 15, at a sampling rate of 44.1 kHz was generated. As in this case the absolute delay and gain of each IR, relative to the others, is important, the measurement was made connecting through a switch just one of the two microphones to the PC right channel input, while the left channel input was directly wired to the signal output. In this way, each measured stereo IR contains always the same electric loopback signal on the left channel, with maximum amplitude and constant delay, and on the right channel the measured IR, with proper delay and relative amplitude. After stripping away the left channel information, the 4 measured IRs were packed into two stereo (binaural) IRs, and saved in .WAV format.

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 the previous chapter: 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 measured numerators, obtaining the required digital equalising filters f . These filters were then employed again by means of the software convolver running (in real time) on the PC [4], although the practical implementation for in-car installation will be based on a low-cost DSP unit.

1. Experiments

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.

Fig. 4 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.

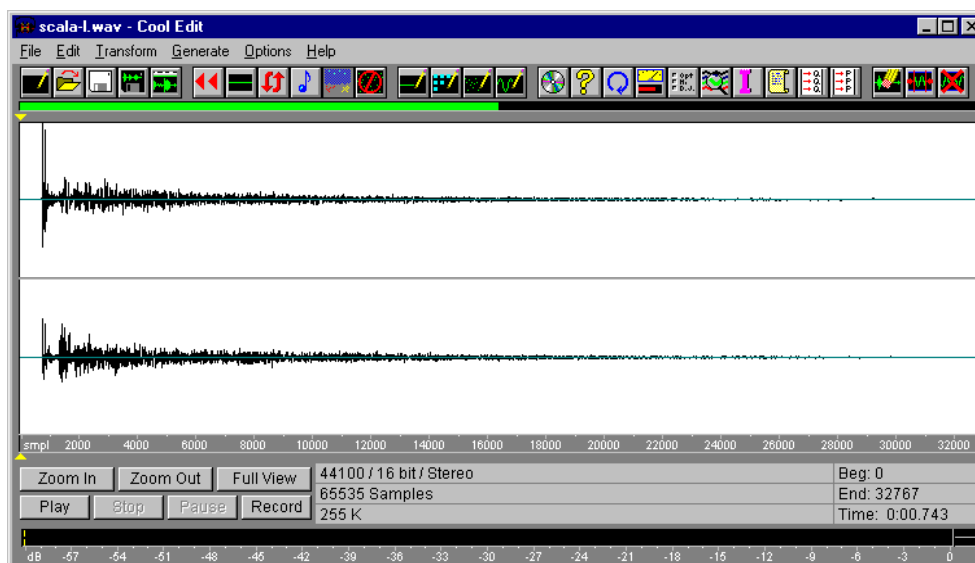


Fig. 4 – Binaural Impulse Response measured in the Teatro “La Scala” in Milan, Italy

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. 5 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.

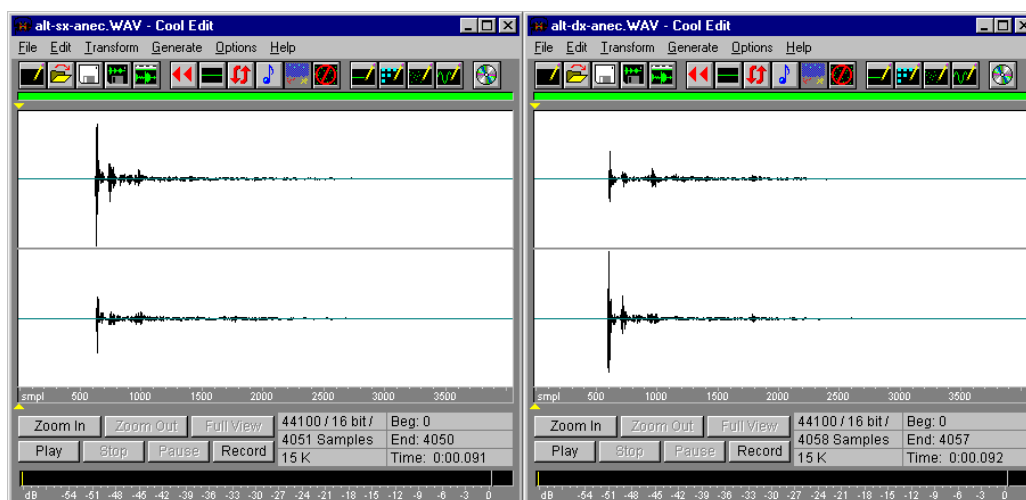


Fig. 5 – Binaural Impulse Responses of the ASK test room

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. 6 reports the two binaural IRs for the system #1.

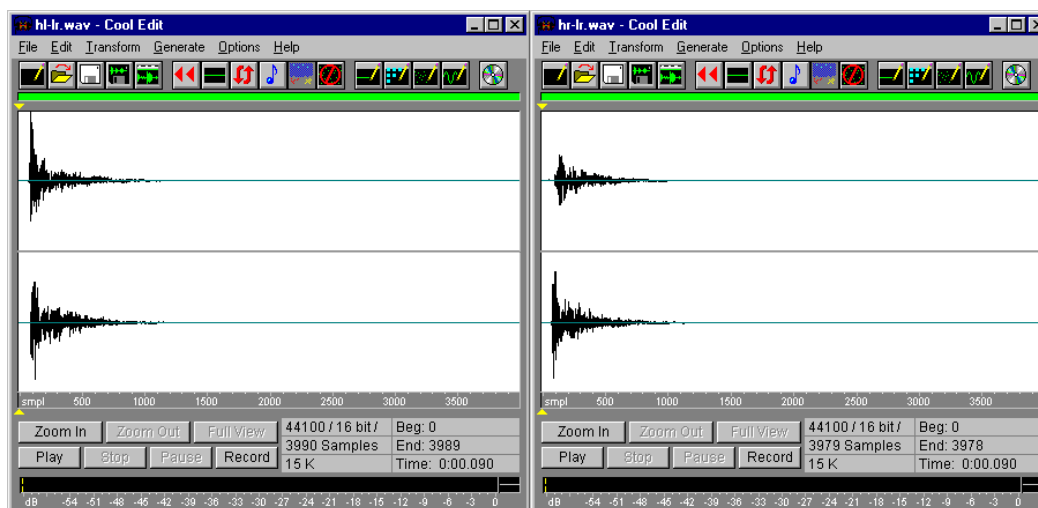


Fig. 6 – Binaural Impulse Responses of the test car – sound system #1

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.

As an example, fig. 7 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 equalise the frequency response of the sound system, but which leave the natural reverberation of the car compartment uncorrected. 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.

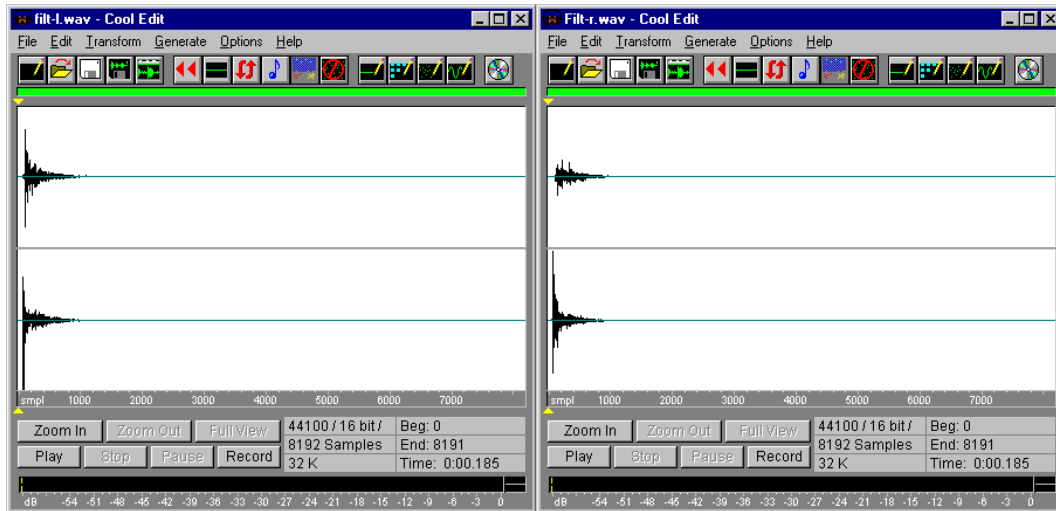


Fig. 7 – Inverse Filters for digital equalisation of the sound system #1, recreating the sound field of the ASK hi-fi listening room

2. Subjective tests

The analysis of the performances of the digital equaliser 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 equaliser, making use of the real-time software convolver already described in [4]; in more detail, half of the sample was filtered (sometimes the first half, sometimes the second), while the other part was left unfiltered. 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 summarises the results of the preference test for the 4 sound systems and the two “virtual spaces”:

Sound System N.	Virtual Space	Preference percentage for the signals		
		Filtered	Unfiltered	Uncertain
1	“La Scala” theatre	74 %	16 %	10%
2	“La Scala” theatre	65 %	20 %	15%
3	“La Scala” theatre	54 %	22 %	24 %
4	“La Scala” theatre	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 equaliser 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 equaliser loses 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 these deficiencies with strong peaks in its equalising 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.

Thus the signal has to be strongly amplified before being fed to the loudspeakers, which are driven well over their power handling capabilities. This causes evident distortion, the system is no more linear, and the

digital equalising filters (which are computed assuming a perfectly linear behaviour of the system) are no more effective in flattening the frequency response and recreating the proper spatial information.

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 $\pm 30^\circ$ 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 [6]. 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 equalisation 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.

1. Conclusion

From the results of the subjective listening tests, which were commented at the end of the previous chapter, it can be concluded that the first results of this research are encouraging. It must be noted also that the new scheme for computing the filter parameters is simple and straightforward, and it does not require advanced mathematical analysis, as it happens instead for the one presented in [1,2]. Thus the new equaliser can be easily implemented without the need of complex circuits or esoteric DSP algorithms.

The problem of inverting a mixed-phase impulse response has been solved with two 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 for example of room acoustics computer programs equipped with auralization extensions [7]. 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 equalisation 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 equalised signals.

The extension of the equalisation 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 research will prosecute along these lines, with the goal of making 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 personalised setup will be available, employing their own head both during the measurement of their preferred ideal sound spaces, and for the car compartment characterisation.

2. 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] J.N. Mourjopoulos, "Digital Equalization of Room Acoustics", JAES vol. 42, n. 11, 1994 November, pp. 884-900.
- [4] 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.
- [5] S.T. Neely, J.B. Allen, "Invertibility of a room impulse response", J.Ac.Soc.Am., vol. 66, pp. 165-169 (1979).
- [6] Y. Ando, *Concert Hall Acoustics*, Springer-Verlag, Berlin 1985.
- [7] A. Farina, "Auralization software for the evaluation of a pyramid tracing code: results of subjective listening tests" - ICA95 (International Conference on Acoustics), Trondheim (Norway) 26-30 June 1995.