Use of digital inverse filtering techniques for improving car audio systems

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 equalize the frequency response of the transducers: by a proper implementation of the inverse filters, it is possible to create a "virtual acoustics" system, which makes it possible to move the apparent position of the sound sources. This way, the original sound field inside the car is substituted with a very different one, coming from measurements taken in brilliant concert halls, or even from numerical simulations of non-existing spaces.

0. Introduction

The possibility of using digital signal processing units for performing the equalization of sound system is widely diffused both for audio-pro applications and for consumer hi-fi home systems. Now some units are available on the consumer market also for car-audio systems: but this particular environment claims 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 quite far from the optimal "stereo triangle". Furthermore, the small volume of the compartment, and the fact that some surfaces are highly reflecting (the windows), 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 equalization 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 auralization 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 an 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, anyway, can 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 has been implemented by first measuring the Head Related Transfer Functions (HRTF) of the driver seated inside the car, through each channel of the sound system. In principle, also the HRTF measurements taken in the “virtual space”, which has to be recreated inside the car, have to be obtained using the driver’s head: but in this case only dummy-head HRTFs were available, both for the case of experimental measurements 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 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.

The filters are actually implemented by means of a software convolver running (in real time) on a PC [3], although the practical implementation for in-car installation will be based on a low-cost DSP board.

Presenting to 5 subjects, seated inside the car at the driver’s position, the filtered signals in comparison with the non-filtered ones, a subjective evaluation of the performance of the system can be obtained. In this case a questionnaire was not employed, but each subject had simply to choose, for each sound system, if he preferred the first or the second (without knowing what of the two was the filtered one). The analysis of the results shows that the filtered signal gains a much higher preference than the unfiltered one, but also that the digital inverse filtering is easier and produces better results when applied to high-quality sound systems, as with low quality devices it can happen that the inverse filters push the system over its linearity limits, causing an unacceptable increase of distortion. In another series of test, instead, the improper choice of the virtual sound field caused an unnatural listening experience, causing the preference to shift toward the unprocessed signal.

1. Processing scheme

Fig. 1 shows the paths of the sound from loudspeakers to ears in a theatre (or in a good listening room), and fig. 2 shows 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 the 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.

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:

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

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

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

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:

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

The goal of the digital equalizer 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 eq. (1), and those of w_l and w_r, coming from eq. (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:

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

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 [4].

A software module which implements the Mourjopuolos least-squares approximate inversion has been already developed by one of the authors [3], 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 [3]. Although in this 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.

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.

2. 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.

In the theater, an omnidirectional dodechaedron loudspeaker was placed on the prescenium, while the fire curtain was closed. Binaural impulse responses were measured, with a sampling frequency of 44.1 kHz and an MLS order equal to 16. So two 65535-points impulse responses were measured for each loudspeaker position. A Sennheiser MKE 2002 dummy head was used. Fig. 4 reports one of these binaural responses, measured 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 effect of the loudspeaker and microphones was removed after the measurements, by convolving them with a proper inverse filter, obtained by the inversion of a free field measurement of the same loudspeaker and with the microphones dismounted from the dummy head. A 512-point least-squares inverse filter was employed in this case.

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.

As in this case the room is heavily damped, a 4096-points impulse response, sampled at 44.1 kHz, was enough for containing the complete reverberant tail.

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.

A test car, equipped with 4 different, switchable sound systems, was used for the listening tests. For each sound system, a set of 4 impulse responses was measured: fig. 6 reports the two binaural IRs for the 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. In this preliminary report about the research, only 2 cases were subjectively investigated: the two ideal listening environments for the sound system #1.

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 complete least-squares inversion of the denominator: as it can be seen, the filters are quite long, and they add some high-frequency artifacts and an audible pre-echo to the signals being filtered with them. Probably some manipulation of the denominator, prior of its inversion, was needed. Instead of trying to correct these problems, the inverse filters were recomputed making use of 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 uncorrected. These simplified filters are shown in fig. 8, and they were the ones employed for the subjective experiments.

Obviously, also the inverse filters for recreating the theatre sound field were obtained with the same approximation.

3. Subjective tests

The analysis of the performances of the digital equalizer is still in process. At the time of writing, only a little number of subjective results was available, but a more comprehensive statistical analysis will be presented elsewhere [6].

5 subjects were asked to seat inside the test car, and to listen at various sound samples. These were obtained by digitally transferring some music samples from commercial CDs, by the use of a Toshiba CD-ROM unit and the "Digital Domain" program by M.Overtoom, which makes it possible to get a perfect digital copy of the contents of an audio CD.

Each 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 [3]; 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 unfiltered and filtered part 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.

In the case of the virtual theatre, the average preference was 74% for the filtered signal, 16% for the unfiltered, and 10% of uncertain cases. In the case of the hi-fi system, instead, the preference was only 35% for the filtered signal, 60% for the unfiltered, and 5% of uncertain cases.

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 [7]. 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 little 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 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 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 unnatural, because some listeners said that the sound seemed "larger" than expected.

4. Conclusion

Apart the results of the subjective listening tests, conducted on a number of subjects too small, and with poor analysis of subjective aspects, the first results of this research are encouraging. The new scheme for computing the filter parameters is simple and straightforward, and does not require advanced mathematical analysis, as it happens for the one presented in [1,2].

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 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 [8]. 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.

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 makes in principle possible 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 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.

5. 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, 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.

[4] J.N. Mourjopoulos, “Digital Equalization of Room Acoustics”, JAES vol. 42, n. 11, 1994 November, pp. 884-900.

[5] S.T. Neely, J.B. Allen, “Invertibility of a room impulse response”, J.Ac.Soc.Am., vol. 66, pp. 165-169 (1979).

[6] A. Farina, E. Ugolotti, “Spatial equalisation of sound systems in cars by digital inverse filtering”, Proc. of 4th International Conference and Exhibition Comfort in the automotive industry - Bologna (Italy) October 2-3, 1997.

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

[8] 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.

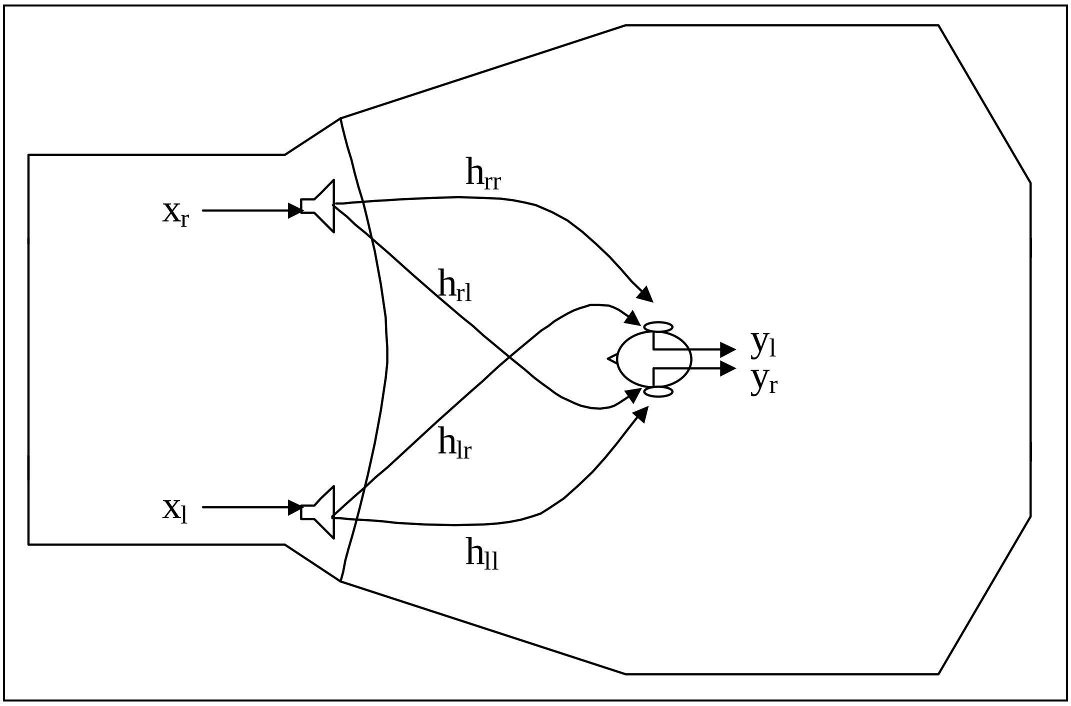

Fig. 1 – Ideal listening conditions in a theatre

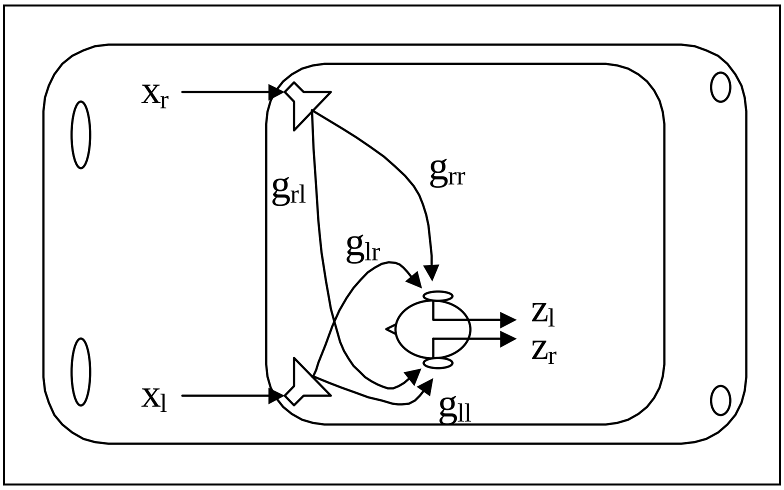

Fig. 2 – Effective listening conditions inside a car

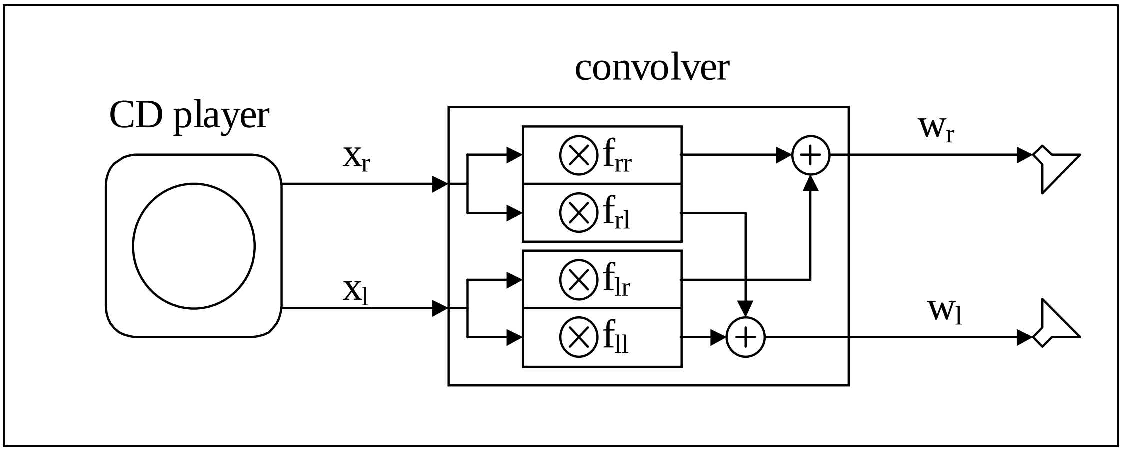

Fig. 3 – Block diagram of the convolver

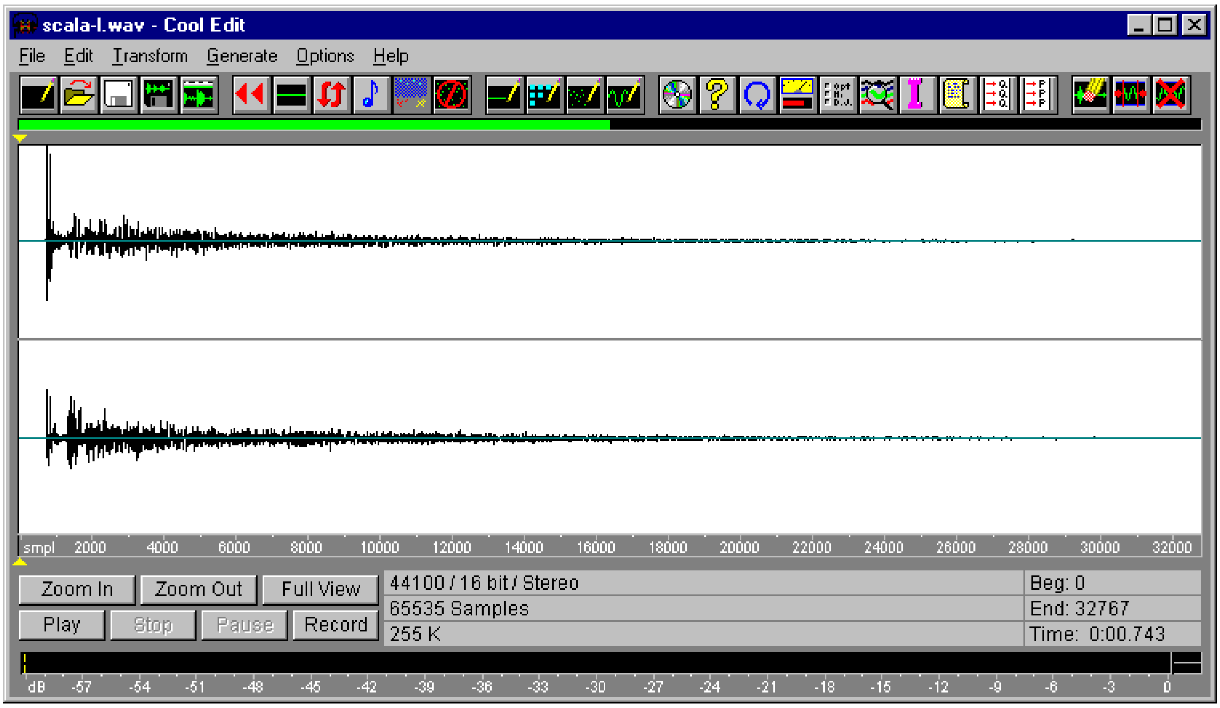

Fig. 4 – Binaural Impulse Response measured in the Teatro “La Scala” in Milan, Italy

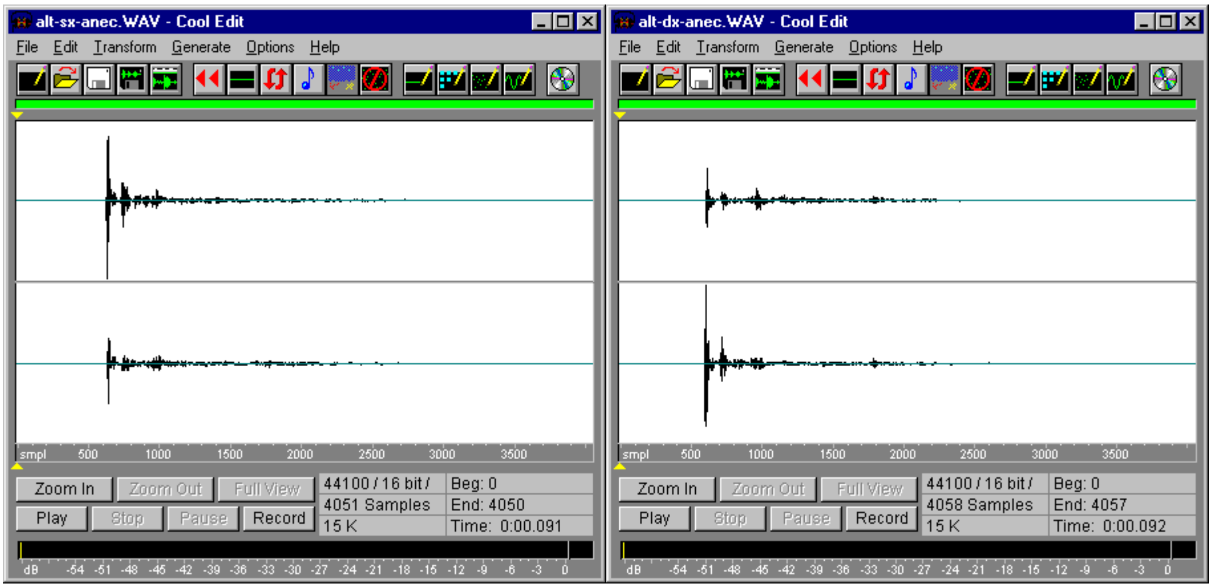

Fig. 5 – Binaural Impulse Responses of the ASK test room

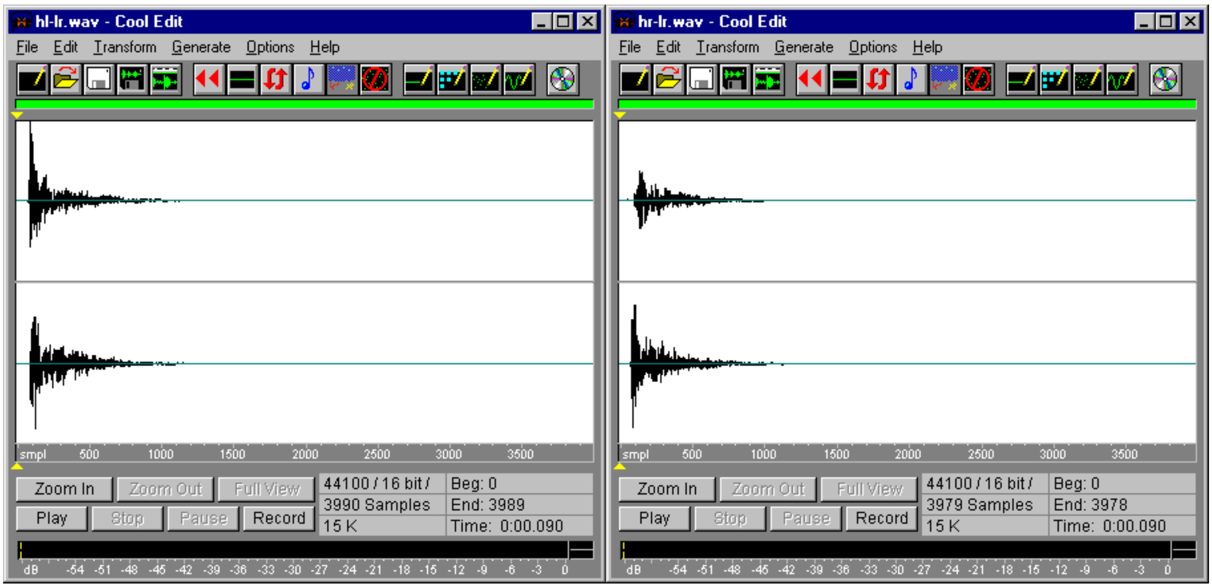

Fig. 6 – Binaural Impulse Responses of the test car – sound system #1

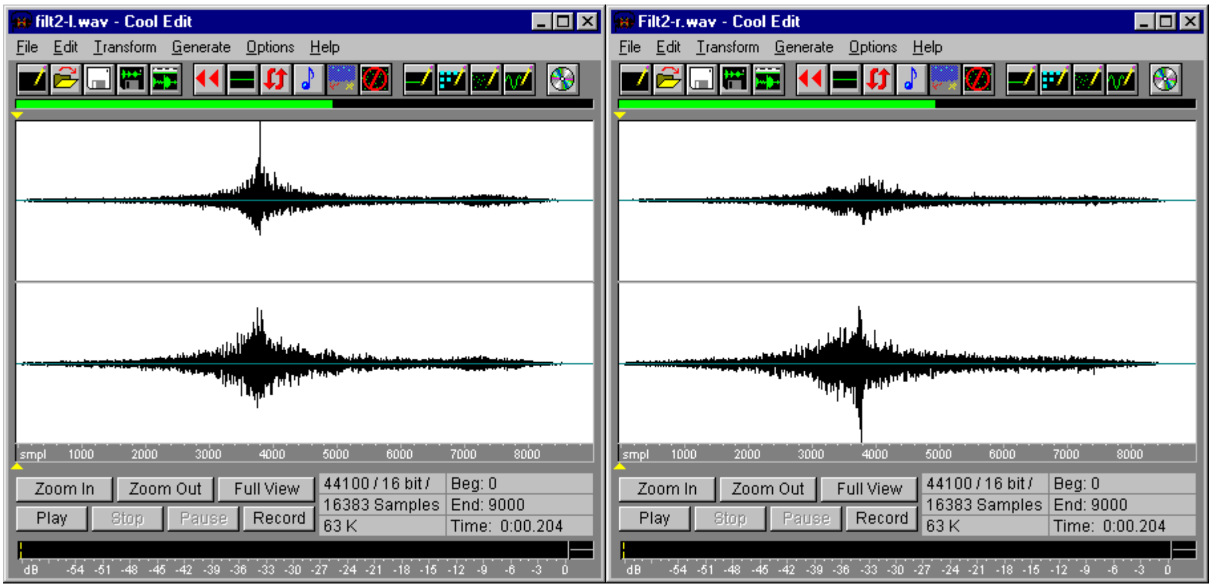

Fig. 7 – Inverse Filters for digital equalization of the sound system #1, recreating the sound field of the ASK test room, computed with complete inversion of the denominator through the Mourjopoulos least-squares technique.

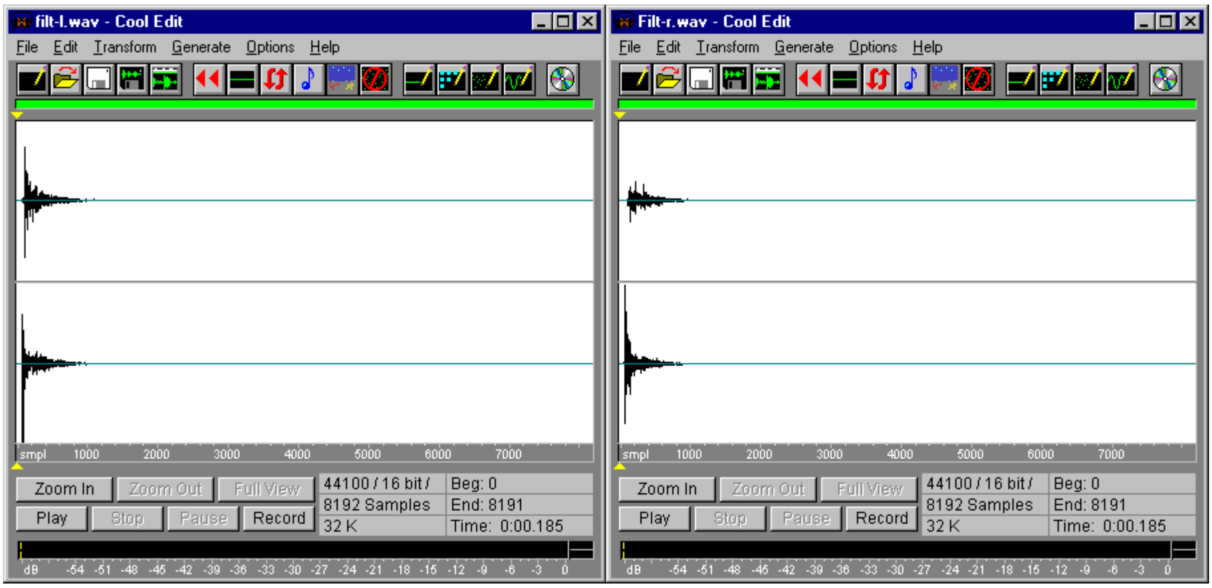

Fig. 8 – Inverse Filters for digital equalization of the sound system #1, recreating the sound field of the ASK test room, computed with minimum-phase inversion of the denominator through the Neely and Allen technique.