Audio Engineering Society
Convention Paper
Presented at the 112th Convention
2002 May 10–13 Munich, Germany

Effects of the Background Noise on the Perceived Quality of Car Audio Systems

Fabio Bozzoli[1], Enrico Armelloni[1], Emanuele Ugolotti[2], and Angelo Farina[1]

[1] *Industrial Eng. Dept., Università di Parma, Via delle Scienze - 43100 Parma – Italy – E-mail: bozzoli.fabio@hotmail.com*

[2] *ASK Automotive Industries, via Fratelli Cervi n. 79, 42100 Reggio Emilia – Italy – E-mail: ugolottie@askgroup.it*
Correspondence should be addressed to corresponding author (Email address)

ABSTRACT

The paper describes the recording/reproduction technique and the subjective listening experiment aimed to the assessment of the effect of the background noise on the perceived quality of the sound being reproduced inside a car. The noise inside 3 different cars was recorded at various speeds both with a binaural microphone and with a Soundfield microphone. These background noise recordings are reproduced inside a special listening room, by means of a very sophisticated reproduction chain, designed so that at the ears of the listeners the same sound pressure is presented as inside the original car. A computer-based system is finally employed for collecting subjective responses to sound stimula, constituted by the reproduction of music or speech on an automotive sound system in presence of the background noise.

INTRODUCTION

It is well known that the sound system installed in a car has to work properly in presence of the strong background noise produced by the car itself: this causes the fact that the optimal frequency response of these systems is not flat, but contains a substantial boost of lower frequencies, for being able to maintain a reasonable Signal-to-Noise ratio in presence of very loud noise at low frequency.

Although this fact is commonly accepted, and the sound systems are designed for providing the required low frequency boost, indeed the subjective evaluation of the quality of the reproduced sound is usually made with the car inside a laboratory, with the engine switched off, and without any realistic simulation of the background noise which instead is present during the normal usage of the system when the car is running on the road. Only in certain cases some listening test is done with the car running, but this results to be unpractical and gives very little reproducibility.

This paper is the first describing a huge research program, having the ambitious goal of investigating the effect of the background noise on the listening conditions inside a vehicle (including the passenger-to-passenger communications, the use of hands-free telephone systems, and the warning messages coming from GPS navigation systems or from proximity warning tools). In this paper the focus is mainly on

the effect of the background noise on the listening conditions for the sound being reproduced by the sound system of the car, with particular attention to the reproduction of music. Other papers will follow related to the investigation of other effects.
The methods scheduled for these research are two:

- reproduction of both the background noise and the sound coming from the car's sound system inside a specially-equipped listening room
- listening tests made inside a car in the laboratory, with a secondary sound system capable of realistic reproduction of the background noise.

Only the first method has been actually completed, although large part of the system employed for it will be useful also when the second one will be implemented.
In substance, the method starts from special multichannel recordings of the background noise made with the car running on the road under controlled conditions of speed, curvature of the path and conditions of the road pavement; these recordings are made with two complementary techniques: dummy-head binaural recordings and 4-channels, B-format recordings made employing a Soundfield microphone.
Fig. 1 shows a particular of the dummy head employed, retrofitted with the Soundfield microphone, which is placed just behind it.

Fig. 1 – Multichannel microphone setup

The recordings were made employing two separate DAT recorders: a Sony TCD-D100 is employed for the binaural recordings, and a TEAC RD-101T is employed for recording the 4 channels coming out from the Soundfield microphones. The two DAT machines were NOT synced together, and this caused some problems for getting a synchronous playback of the recordings.
Furthermore, the TEAC recorders halves the passband when working in 4-channels mode, although in reality the background noise is absolutely irrelevant above 12 kHz, and consequently this fact revealed to be of negligible importance.
The solution of these problems has been obtained by making use of a new, portable, multichannel sound board (Metric Halo Mobile I/O), which allows for the use of a notebook computer, equipped with Firewire interface, for performing the recordings on the car. Figure 2 shows this system.

Fig. 2 – Notebook equipped with a multichannel sound board

Unfortunately the new system was not available in time for collecting the experimental data discussed in this paper.
Once the recordings are done, these needs to be replayed inside the listening room, making use of advances digital signal processing techniques, for ensuring that the listener seating inside this special room is hearing exactly the same background noise which was present inside the car where the recordings were made.
In practice, the binaural recording is reproduced by means of two stereo dipoles (one in front, one behind the listener), as shown in figgs. 3 and 4.

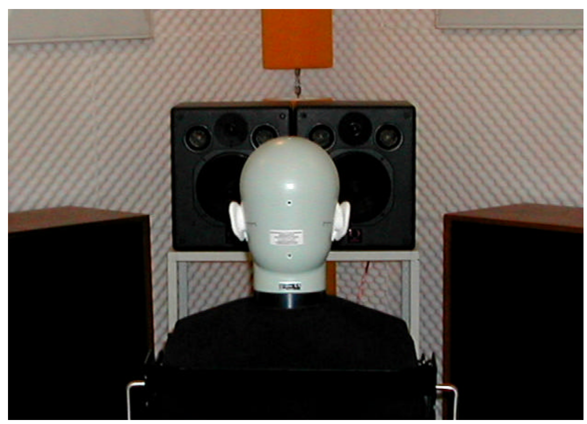

Fig. 3- Frontal Stereo Dipole

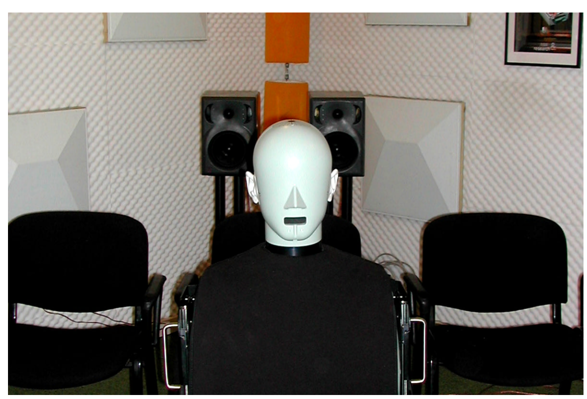

Fig. 4 – Rear Stereo Dipole

The frontal Stereo-Dipole is made with two Quested 2108 loudspeakers, the rear Stereo-Dipole is made with two Quested F11P monitors. Each loudspeaker pair is driven with a Crown K1 current-controlled power amplifier.

The signals driving the stereo dipole are processing with cross-talk canceling filters, implemented making use of the methods described in the next chapter. The processed signals are reproduced by means of the first 4 channels of an Echo Layla sound board.
Regarding the reproduction of the B-format soundtrack, a complete three-dimensional Ambisoncs system is employed. 8 identical loudspeakers (Turbosound Impact 50) are mounted in double-squatre configuration (4 are at the corners of an horizontal square, and the other four are at the crners of a vertical square, at the sides of the listener). Fig. 5 shows a panoramic view of the listening room, in which the position of the Ambisonics loudspeakers can be noticed on the walls and on the ceiling.

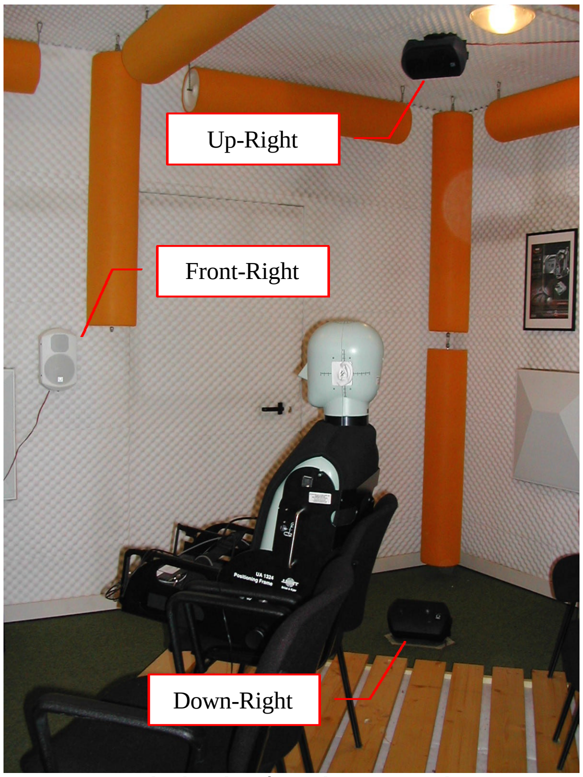

Fig. 5 – loudspeakers for Ambisonics reproduction

These 8 loudspeakers are powered by a single 8-channels amplifier (QSC CX168), which ensures that exactly the same gain is always maintained for the 8 channels.
The 4 channels of the B-format recording are played by means of the channels 5,6,7 and 8 of the Each Layla sound card. The conversion between the B.-format signal and the 8 loudspeaker feeds (usually called "Ambisonics decoding") is made be means of an external DSP processing unit (BSS Soundweb), which implements a suitable processing network, described in the following chapters.
The last two channels of the Echo Layla sound board (9, 10) are employed for driving a self-powered subwoofer (Audio Pro B1-20), which is located under the frontal Stereo Dipole, as shown in fig. 6.
Finally, the computer employed for driving the reproduction of the recordings, and for collecting the automated questionnaires described at the end of this paper, required to be completely noiseless: a liquid-cooled, fanless computer was employed for this (SID FutureClient). For the same reason, the cooling fans of the BSS Soundweb and of the QSC power amplifier were switched off.
The total sound pressure level produced by the whole stack of equipment is less than 26 dB(A) at 1m distance.
The computer is equipped with a large LCD display, and with wireless keyboard and mouse, so that it can be operated also from a listener seated at the center of the "sweet spot", 2.5m far from the computer itself.

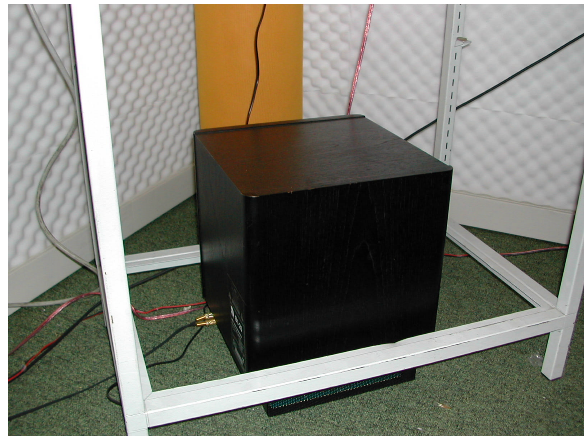
Fig. 6 – Active Subwoofer

The reproduction of the multichannel recordings is made employing a multichannel wave editor (CoolEditPro), which was also used, thanks to the availability of a set of specialized plugins (Aurora), for measuring the cross-talk functions and for computing the cross-talk canceling inverse filters.
Figure 7 shows the rack containing, from top to bottom, the Echo Layla sound board, the BSS Soundweb the two Crown K1 amplifiers, the QSC CX168 8-channels amplifier, and the liquid-cooled computer.


Fig. 7 – instrumentation rack.

The following chapters describe in detail the setup of the binaural reproduction over the stereo-dipoles, the Ambisonics reproduction of the B-.format soundtrack, the procedure for verifying that the reproduced background noise corresponds to the noise really existent in the original cars, and the subjective listening test system, employed for collecting the questionnaires being compiled by the subjects.

STEREO DIPOLE

The theory of the Stereo-Dipole reproduction system is widely known [1,2,3], and consequently it needs to be recalled here only very shortly.

As with other "transaural" reproduction systems, the concept is to cancel the cross-talk signal which is going from the left loudspeaker to the right ear, and vice versa.

The approach employed here is derived from the Research about the Stereo Dipole implementation originally developed by Kirkeby and Nelson [4,5].

The 4 cross-talk cancelling filters *f*, which are convolved with the original binaural material, have to be designed so that the signal collected at the ears of the listener are identical to the original signals. The following fig. 8 shows the cross-talk phenomenon in the reproduction space:

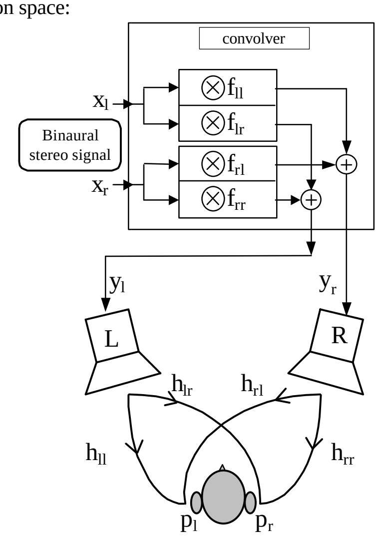

Fig. 8 – cross-talk cancelling scheme

Imposing that $p_l=x_l$ and $p_r=x_r$, a 4x4 linear equation system is obtained. Its solution yields:

$$\begin{cases} f_{ll} = (h_{rr}) \otimes InvDen \\ f_{lr} = (-h_{lr}) \otimes InvDen \\ f_{rl} = (-h_{rl}) \otimes InvDen \\ f_{rr} = (h_{ll}) \otimes InvDen \\ InvDen = InvFilter(h_{ll} \otimes h_{rr} - h_{lr} \otimes h_{rl}) \end{cases} \tag{1}$$

The problem is the computation of the InvFilter (denominator), as its argument is generally a mixed-phase function. For this task the Kirkeby-Nelson frequency-domain regularization method is preferentially employed, due to its speed and robustness. A further adaptation over the previously published work [6] consists in the adoption of a frequency-dependent regularisation parameter. In practice, the denominator is directly computed in the frequency domain, where the convolutions are simply multiplications, with the following formula:

$$C(w) = FFT(h_{ll}) \cdot FFT(h_{rr}) - FFT(h_{lr}) \cdot FFT(h_{rl}) \tag{2}$$

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

$$InvDen(w) = \frac{Conj[C(w)]}{Conj[C(w)] \cdot C(w) + e(w)} \tag{3}$$

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

The advantage of this approach is that it takes into account the real performance of the loudspeakers and of the binaural dummy head, because the computation of the inverse filters is based on the measured transfer functions.

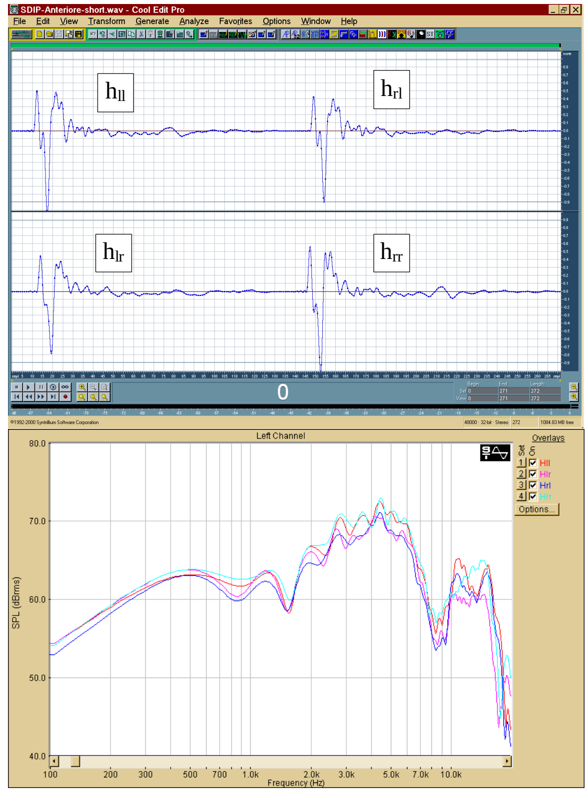

Fig. 9 – measured transfer functions of the frontal Stereo Dipole

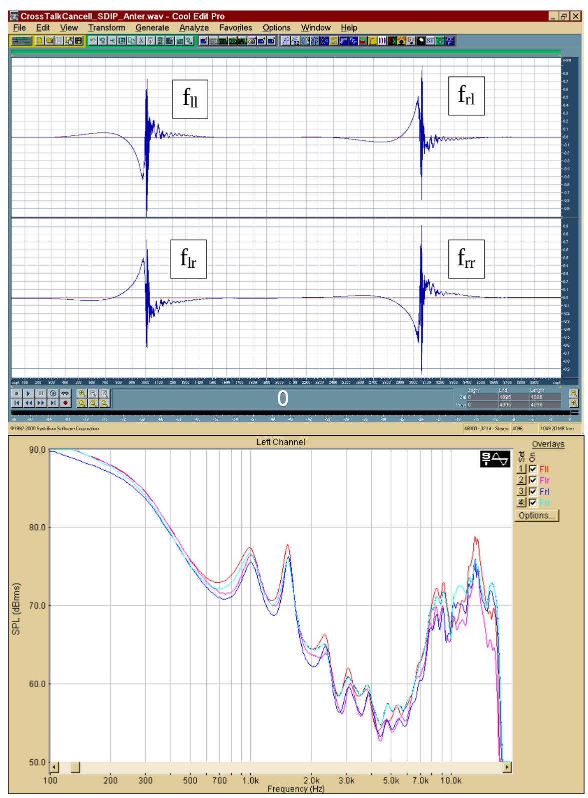

Fig. 10 – inverse filters for the frontal Stereo Dipole

Figure 9 and 10 show respectively the measured transfer functions of the frontal stereo dipole and its inverse filters (both in terms of impulse responses and of frequency responses), obtained by a multiple MLS measured method, which was already presented in previous papers [7,8].
In practice, the inverse filters shown in fig. 10 do not simply provide effective cross-talk cancellation; they completely equalize both the magnitude and the phase response of each loudspeaker.
Fig. 11 shows the verification of the effectiveness of the cross-talk cancellation and of the frequency response linearization obtained by means of the inverse filters on the frontal Stereo Dipole.

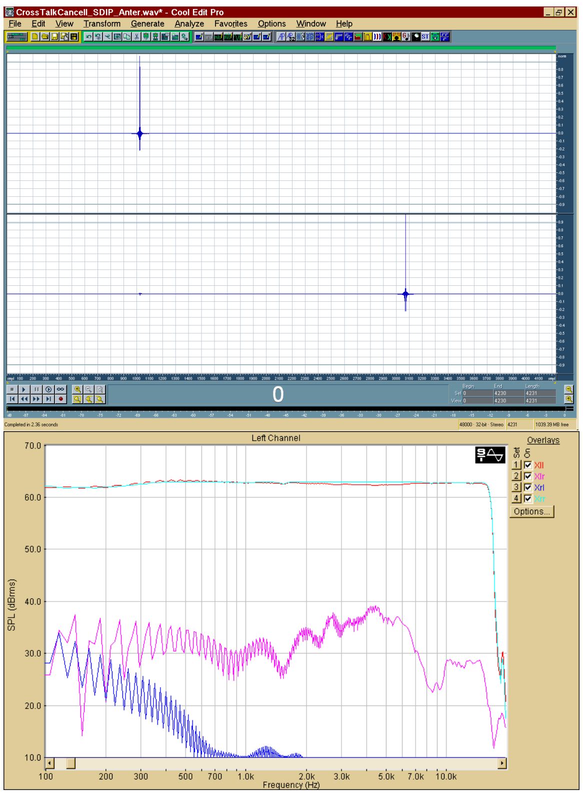

Fig. 11 – verification of the frontal Stereo Dipole

It can be seen how the ipselateral paths are now perfectly equalized, and the cross-talk paths are attenuated by more than 40 dB over the whole frequency range.
As both stereo-dipole systems are being operated simultaneously, this reproduction system is substantially more robust to head movements than traditional systems which make use of just two loudspeakers. In practice the system revealed to be almost immune from problems such as front-back confusion and apparent rotation of the sound field together with the rotation of the listener's head: this means that the listening is much more natural than what can be obtained by means of headphones.

AMBISONICS REPRODUCTION SYSTEM

The Ambisonics array of loudspeakers employed here is a derivation of the system already presented in [9]. The principal difference is the different position of the 8 loudspeakers, but many other improvements were made, thanks to the better hardware now available, and to the fact that the Soundweb can be easily programmed for performing a very complex real-time digital signal processing.
The location of the 8 loudspeaker chosen was the so-called "dual square" array, This is shown in fig. 12. In substance, a first quadrilateral array is created in the horizontal plane XY (resembling the old "quad" setup), and a second vertical array is created in the YZ plane.
If the two quadrilateral are perfect squares, the array is NOT isotropic: a third square lying on the XZ plane would be required for completing the symmetry. An isotropic array is obtained if the sides along X and Z are longer than the sides along Y, (with a ratio equal to $\sqrt{2}$). This not being the case, as the available room does not have enough height, it is necessary to compensate for the non-isotropic array inside the Ambisonics decoder.

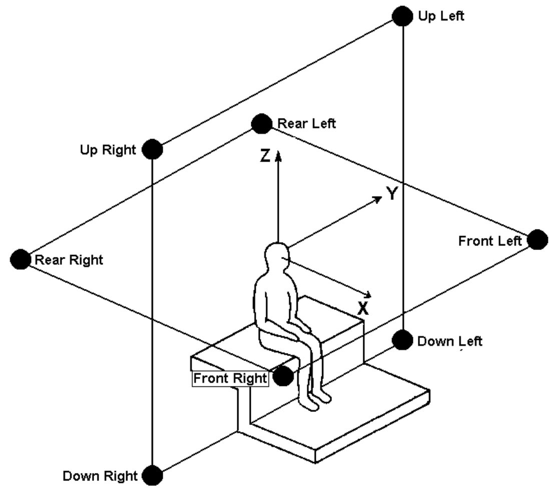

Fig. 12 – geometry of the dual-square loudspeaker array

In practice, the use of the BSS Soundweb makes it possible to introduce easily variable gains, so that the Y signal is slightly reduced in comparison with X and Z (which instead are boosted): this corrects for the fact that the not isotropic array tended to give more emphasis to the Y component of the B-format signal.
Fig. 13 shows the processing network implemented on the BSS Soundweb. This network incorporates:

- Correction for the -3 dB offset of the W channel caused by the Soundfield microphone (standard B-format)
- Shelf-filters with a crossover frequency of 500 Hz, which implement max-Rv decoding at low frequency and max-Re decoding at high frequency [10].
- Reduction of the gain of the Y channel and increase of the gain of X and Z channels.
- Matrixing of the WXYZ modified signals for deriving the theoretical speaker feeds
- FIR equalization of each theoretical speaker feed for driving the corresponding loudspeaker, so that any minor difference between individual transducer is completely compensated.

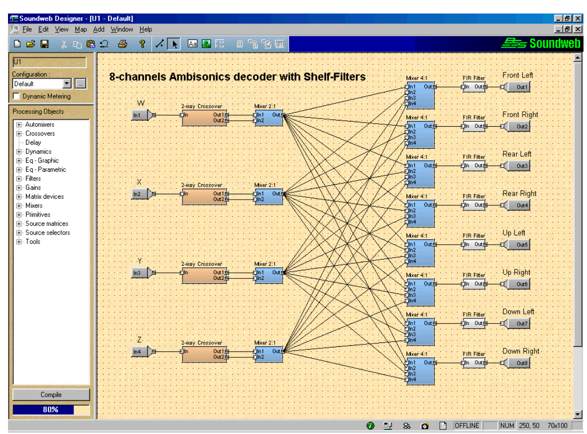

Fig. 13 – Soundweb processing network for Ambisonics decoding

The above features make this decoder implementation unique, and corrects for most of the problems previously encountered with simpler implementations (which did not include shelf filtering, nor FIR equalization of the individual reproduction channels).

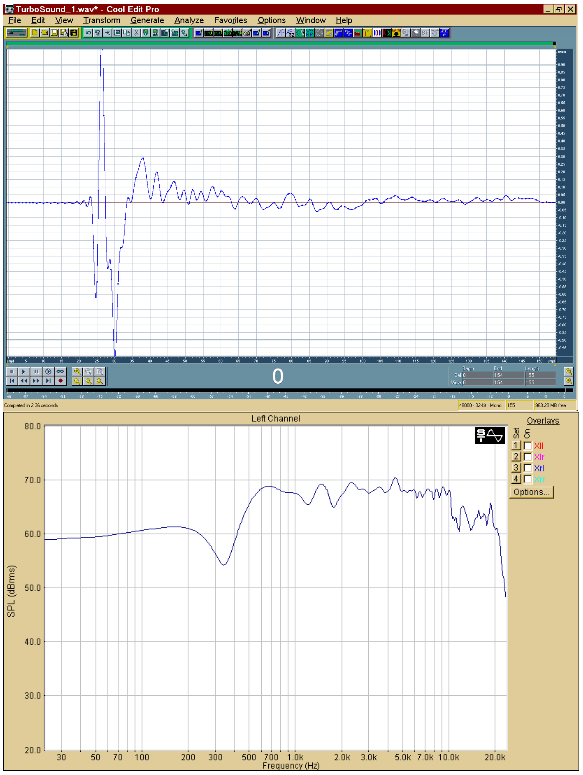

Fig. 14 – response of the Turbosound loudspeaker Front-Left

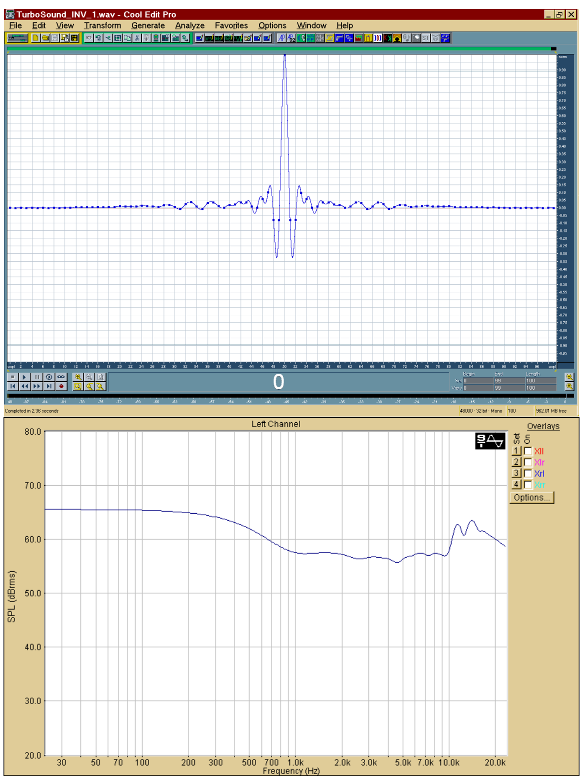

Fig. 15 – 100-taps, minimum-phase inverse filter

Of course, this approach required that the impulse response of each loudspeaker was separately measured and inverted. The processing for the loudspeaker Front-Left is shown here: fig. 14 is the measured impulse response, fig. 15 the corresponding inverse filter, and fig. 16 the verification of the equalization obtained.

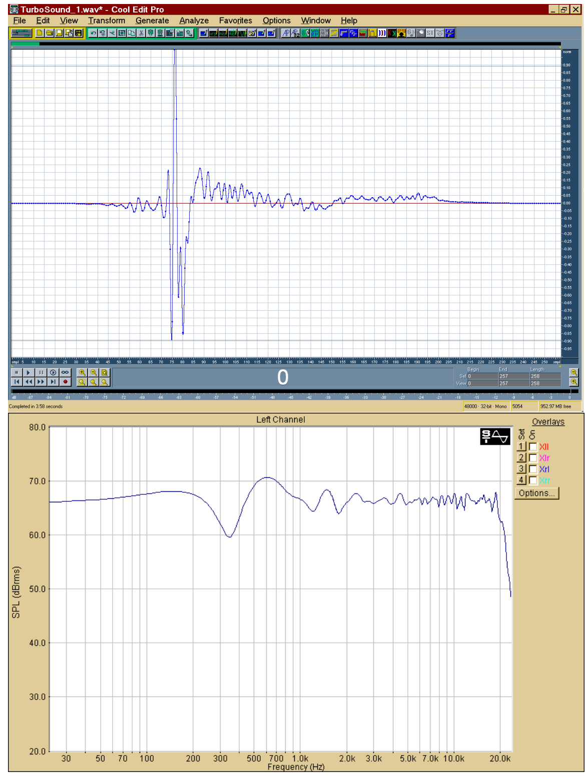

Fig. 16 – Verification of the equalization

Although the equalized curve is not perfectly flat, it can be seen how also employing a FIR filter of very limited length (only 100 taps) it is possible to improve significantly the reproduction fidelity.

RECORDINGS OF NOISE INSIDE CARS

As this research is just at the beginning, only a very limited number of multichannel recordings were made till now. In practice, for assessing the capability of the background noise reproduction system, it was chosen to attempt to reproduce the noise of three cars, which have very similar characteristics. The three cars are of the same brand (Opel) and mount similar engines (2.0 Diesel). Consequently, also the noise was quite similar: this is a good “transparence” test, because only a very good reproduction system can allow for detecting such minor difference between the background noise of the three cars. The recordings were made at various speeds, on flat and straight road (the highway between Parma and Reggio Emilia). The following three pictures show the 1/3 octave spectra obtained from the binaural recordings (the dummy head was on the front passenger seat), at a speed of 120 or 130 km/h (the same RPM was kept, but one of the three cars had longer transmission ratio, and consequently required to run at 130 km/h instead of 120 km/h for giving the same RPM as the other two cars).

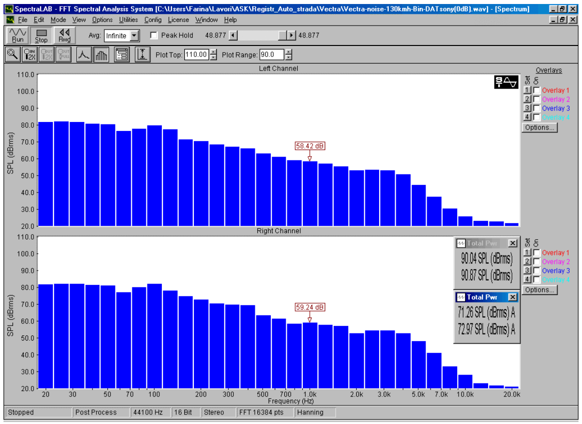

Fig. 17 – 1/3 octave spectra, VECTRA 2.0 DTI, 130 km/h

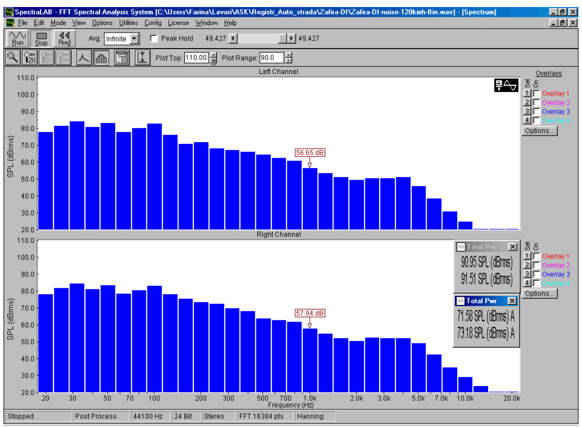

Fig. 18 – 1/3 octave spectra, ZAFIRA 2.0 DI, 120 km/h

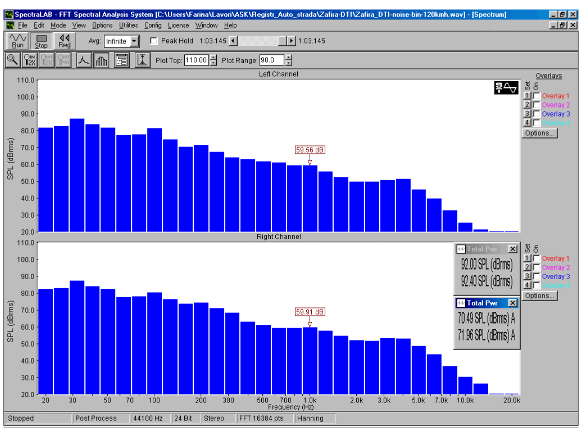

Fig. 19 – 1/3 octave spectra, ZAFIRA 2.0 DTI, 120 km/h

The similitude of the three recordings is self-evident. Nevertheless, carefully listening by headphones to the binaural recordings, some minor difference appears (particularly between the first car and the other two, which instead are more difficult to separate).

Consequently, it is expected that also with the loudspeaker reproduction system it will be possible to distinguish between the different "character" of the noise inside these cars.

For checking if the reproduction system is capable of maintaining the detection of such very subtle differences a subjective listening test was performed, as described in the next chapter.

The final remarks is related to the procedure implemented for ensuring that the same absolute SPL and spectrum is presented at the ears of the listeners.

After setting up the reproduction chain for the proper playback of each binaural recording and of each B-format recording, a "transparency" test was performed (also called the "photocopy of the photocopy test").

The same microphone employed for the original recording (binaural dummy head or Soundfield ST-250) was placed on the listener's seat inside the reproduction system. Then a second re-recording was done, while the original recording was being played over the reproduction system.

Finally, the re-recorded signal was compared with the original recording, and the system was adjusted until these differences were reduced to less than 1 dB.

Finally, the gain of each part of the sound reproduction system (Stereo Dipole, Ambisonics array) was reduced by 3 dB, so that when both the systems are running simultaneously, the energetic sum of the two background noises, being reproduced by the two subsystems, combine together, yielding exactly the same SPL value which was present inside the car during the original recording.

COLLECTION OF SUBJECTIVE RESPONSES

The collection of subjective responses is also based on the use if the same computer, which is being employed for reproducing the sound samples. In practice, these two functions were integrated, by means of the creation of a special software tool. This program makes it possible for the listener to play at will several sound samples, and to fill up on the screen the questionnaire pertinent to each group of stimula.

This way the listener can easily compare the samples in pairs or in triplets, switching back and forth among them, and this makes it possible to detect also very subtle differences.

Fig. 20 shows the user's interface of this Audio Testing software.

Fig. 20 – software for questionnaire collection

In practice, the software can be easily reconfigured for collecting questionnaires employing a different number of questions, with different attributes as left and right terms of each question. Furthermore, the number of sound samples (1,2,3,4 in the above example) for each car (A,B,C in the above example) can also be varied without the need to recompile the program.

The user has to fill up a different questionnaire for each of the cars being presented (3, in this cases), making use of 4 different speech or music samples recorded in each car. Switching the car causes automatically the appearance of the corresponding questionnaire on the screen. Furthermore, switching the car also switches the WAV file being reproduced, so that the listener can easily check if the differences in the responses to the questions fit with the listening perception. The sound switching is smooth and silent, and the exact position inside the sound file is kept, so that the switching causes no intrusion in the listening experience.

The first results of these subjective tests simply shown that some differences between the background noise of the three cars described at the previous chapter can be perceived. As these background noises were really very similar, this was considered a good result, because it demonstrates that the listening system is capable of keeping detectable also very subtle differences. On the other side, there is the risk that this very selective comparative listening method overestimates these difference, making it detectable effects, which are of no practical importance in the real life.

As it is not easy to conduct a similar listening experiment on real cars, it is not possible to establish now if this artificial evaluation method is corresponding to the perceptions in the real life. On the other hand, the availability of this method will be in any case a powerful tool for the design of better sound systems, and for the assessment of their performance under conditions certainly closer to the real life than in traditional listening tests performed in absence of any background noise.

CONCLUSIONS AND FUTURE WORK

This preliminary paper reported on the steps taken for setting up a reproduction system capable of recreating a realistic reconstruction of the background noise existent inside a car compartment.

The method is based on the use of true multichannel recordings of the background noise, which are replaced inside a special listening room, equipped with two integrated reproduction chains: a dual Stereo Dipole for transaural presentation of binaural recordings, and an advanced Ambisonics decoder for periphonic (3D) presentation of B-format recordings. The two systems can be operated separately or simultaneously, provided that, in the latter case, a proper correction for the gain of the two systems is applied.

The paper described the theoretical principles and the actual implementation of the methods, alongside with the software employed. The reproduction of the sound samples employed for the listening tests is driven by a specially written software, which also enables for the automatic collection of questionnaires.

The system resulted to place the listener in condition to discriminate very subtle differences in the reproduced background noise, and consequently this opens the possibility to assess the effect of the noise on the perceived quality of the sound reproduced by the car's sound system when the car is running on the road.

This allows for the prosecution of the research, which will move to the execution of several listening tests, aimed principally to defining the optimal listening conditions in terms of the system's frequency response, of the spatial attributes of the sound field, and of the intelligibility of the speech inside the car.

ACKNOWLEDGEMENTS

This work was possible thanks to a research cooperation program between the University of Parma and ASK Automotive Industries of Reggio Emilia, Italy. The authors want also to express their gratitude to Rieter Automotive Management AG, Winterthur, Switzerland, for the financial support to the Ph.D. program of the University of Parma, under which the main author is working, and to the Ambiophonics Institute, USA, for the substantial funding, support and encouragement.

REFERENCES

[1] H. Moller – "Fundamentals of Binaural Technology" – Applied Acoustics vol. 36 (1992) pp. 171-218.

[2] 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.

[3] O. Kirkeby, P. A. Nelson, and H. Hamada – "Virtual Source Imaging Using the Stereo Dipole"", Pre-prints of the 103rd AES Convention, New York, 26-29 September 1997.

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

[5] O. Kirkeby and P. A. Nelson – "Digital Filter Design for Virtual Source Imaging Systems", Pre-prints of the 104th AES Convention, Amsterdam, 15 - 20 May, 1998.

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

[7] 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.

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

[9] A. Farina, E. Ugolotti, "Subjective comparison between Stereo Dipole and 3D Ambisonics surround systems for automotive applications", 16th AES Conference, Rovaniemi (Finland) 12-14 April 1999.

[10] J. Daniel, J.B. Rault, J.D. Polack - "Ambisonics encoding of other audio formats for multiple listening conditions" - Pre-prints of the 105th AES Convention, S.Francisco, 26 - 29 September, 1998.

INTERNET REFERENCES

The Aurora software employed for performing the measurements and computing the inverse filters can be downloaded for free from HTTP://www.ramsete.com/aurora.

The measured impulse responses, the inverse filters, and the Ambisonics Decoder setup for the Soundweb processor can be downloaded from HTTP://pcangelo.eng.unipr.it/public/AES112.