Proceedings of the 1st Human-Centered Transportation Simulation Conference, The University of Iowa, Iowa City, Iowa, November 4-7, 2001 (ISSN 1538-3288).

A cost effective architecture for realistic sound rendering in the SCANeR II driving simulator

Byline

A. Heidet[1], O. Warusfel[2], G. Vandernoot[2], B. Saint-Loubry[1], A. Kemeny[1],

1- Renault, Research Department, 1 av. du Golf, 78288 GUYANCOURT, FRANCE
2- IRCAM, Room Acoustic Team, 1 place Igor Stravinsky, 75004 PARIS, FRANCE

For more than ten years, the Driving Simulation and Virtual Reality Research Group within Renault's Research Department has been developing interactive and real-time driving simulators. These developments have resulted in the SCANeR II comprehensive driving simulation package. The simulators are used as a tool for product development and driver and vehicle studies, in such areas as ergonomics, electronic driver aid systems (ACC for example), driver behaviour and accident studies.
In the past years, much attention has been given to improving vehicle dynamics and visual simulation fidelity, as well as motion rendering, raising the requirements on the realism of the overall simulation, for the sake of coherence. This has brought Renault and IRCAM (research institute on acoustics and music) to collaborate more specifically on the issue of sound rendering.

The approach to sound rendering has been to research a cost-effective and tuneable frame-work providing realistic and sufficient cues, relevant to general driving simulation. Based on standard PC hardware and Windows-based software APIs, the sound rendering engine provides the cues to perceive the configuration of the global driving environment (position of obstacles and other vehicles), and to the behaviour of the vehicle. The realism of the simulation is ensured by using techniques developed in the field of room acoustics, and transposing them to the driver's cockpit. Samples are recorded on a real vehicle taking into account the characteristics of the recording method and the transfer function of the vehicle. Then they are fed-back during the simulation, compensating for the acoustic transfer functions of the cockpit used for sound restitution. Extension of these techniques to simulators using Head Mounted Display technology will also be considered.

1 INTRODUCTION

The range of applications that can be run on a given driving simulator is directly linked to its level of realism. A given level of realism, even low, can be useful for a specific purpose, but when it comes to evaluating vehicle systems in a development process by an engineering team, the requirements on realism become quite high in order to immerse the test driver into a familiar environment and therefore facilitate the use of the simulator as a prototyping tool.

In that respect, an important amount of work has been continuously put into increasing the realism of vehicle models [1], of motion rendering [2,3], image generation [4], or traffic generation [5,6]. The results have been integrated into simulators that are used in the framework of various studies in fundamental research, engineering or training [7,8]. In the recent works, sound generation has often been under-stressed, being seen as an easy to implement item.

However, there are justifications to a higher level of importance for sound rendering. First, the perceived over-all realism of a simulation is much based on the homogeneity of its elements: the perceived realism of the dynamics model is improved if the engine noise sounds realistic and reacts correctly to driver input. Second, visual immersion is strongly relayed by sound immersion, as sound helps to feel the volumes in and out of the vehicle.

Based on those reflections, we have implemented a new sound rendering process in the dynamic driving simulator of Renault aiming at simulating the various cues used in driving. The stress has especially been put on defining vehicle-independent sound samples and vehicle transfer function for possible simulation of the sound of different vehicles, and on analysing the specific response of the sound rendering system, including the physical driving station, to take it into account in the sound generation process.

Using the framework provided by SCANeR II and the cues it makes available about the status of the simulation, which are described first, the new architecture of the sound rendering module has been built upon the process described in the second part of the article, using a series of pre-processed sound samples. The recording, processing and final validation of those samples is described last.

2 SOUND RENDERING FRAMEWORK FOR THE RENAULT DRIVING SIMULATOR

2.1 Context of the Driving Simulators at Renault

The SCANeR II driving simulation package created and developed by Renault for more than ten years includes the functionality needed to run training applications as well as vehicle development applications. Since 1995, SCANeR is commercialised through Oktal, and more than ten copies are installed world-wide. Two simulators running at the Renault's Technocentre are equipped with SCANeR II and are involved in tests of new on-board systems (such as AICC (Automatic Intelligent Cruise Control) or ESP (Electronic Stability Program), or new Human-Machine Interfaces) [9] and new research paradigms (in accidents study, comfort, attention monitoring). The simulator, as a virtual prototyping tool, offers a wide range of tuneable parameters for each of its modules, and especially in the dynamics model. The dynamics model used in real-time is also the one used for off-line development and prototyping by vehicle development teams, which allows for a faster integration of the simulator in the vehicle systems development process.

As sound perception participates in the interaction between the driver and its vehicle, if the sound inside the cockpit is not right, the comfort of the driver is impacted and its judgement of the tested systems might be biased. Therefore, even if the goal of the simulator is not to test sound ambiences inside the vehicle, it is important to get the sound right.

2.2 Architecture Highlights

SCANeR II is composed of specialized software modules that exchange data on a local network through a proprietary UDP-IP-based protocol. This architecture choice is allowing to run each function of the simulator (traffic generation, dynamics model, image generation, sound generation, ...), on the hardware platform that is best adapted to the level of performance that is expected: the network is abstracting each platform from the others. Figure 1 is describing a typical implementation.

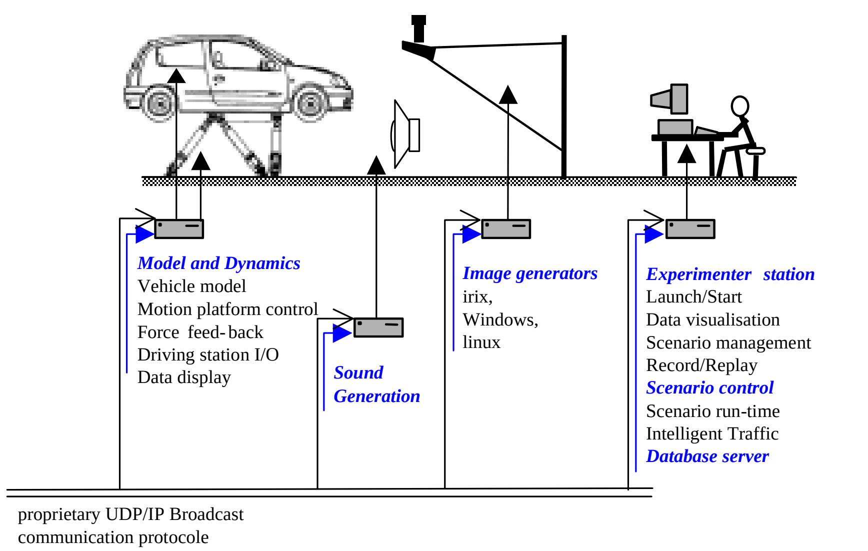


figure 1: A description of SCANeR II architecture

This architecture is allowing to upgrade functionality in a given area of the simulation without impacting other modules, therefore favouring stability. As an example, in the near past, have been used in the SCANeR II architecture sampler-based sound generation, consumer-level PC boards and real-time sound synthesis systems, independently or in parallel.

2.3 Data available for sound rendering

In the SCANeR II architecture, the sound generation process can get its information from the databases that describe the static environment and are interrogated in real-time during the simulation, or from the data made available on the communication network by the other modules. The use of those sources and the combination of the information they provide is offering a wide range of possible effects, continuous, such as engine noise, or based on conditions local to the scenario, such as man holes.

2.3.1 Databases

The static world is described through separate databases that handle the description of the world through its visual, road network, road surface and collision objects descriptions.

A world database in Open Flight format is used for visualisation purposes and is not providing information for sound rendering.

The road network is described as a list of roads. Each road has got characteristics applying to the whole road, such as its width, road markings, nature (urban, country-side, highway,...). An example of a scene and its interpretation is given in figure 2.

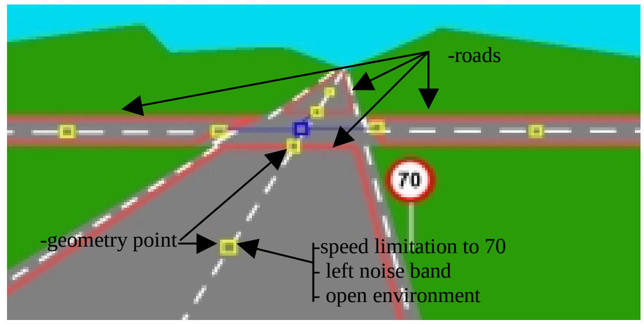


figure 2 : Example of a scene and modelled information

The roads are built of linear lists of points describing the geometry, and each point is a potential attachment for “special sound effects” tags. The tags available today allow to change the type of reverberation of the sounds that is provided by the environment (e.g. <OUT_ENV> for open spaces, <WALL> for semi-open spaces, <CITY> for spaces where high objects are very close on both sides, and <TUNNELS>, for closed spaces), and to play a sample at specific points down the road (e.g. <DRAIN> allows for the simulation of pot-holes and man-holes, <CART_TRACK> is linked to rail track samples, <SPEED_RAMP> for noisy road markings perpendicular to the road, <EMERGENCY_BAND> for noisy road markings on the sides of the road). As the loader of the database has been made flexible, it is easy to extend the format of the file to include new effects when they are made available in hardware.

The road surface is described through a network of Bézier patches that covers all drive-able areas, mainly roads and hard shoulders, but also road sides, side-walks, or any area that is supposed to be driven onto during a simulation. The network of patches is allowing to describe a continuous surface where needed for the vehicle dynamics calculations and is also providing road surface state data.

This data is obtained through the “road query” algorithm, which is giving altitude under each wheel as well as the nature of the road, that can be asphalt, concrete, grass, snow, cobblestones, and a coefficient of “spatial sound” that quantifies the roughness of the surface.

The use of Bézier patches is giving a flexible way to model local changes in the road nature that have impact on sound generation.

The collision objects database is a list of selected objects of the world that are queried for physical interaction with the driven vehicle. Results of the query are available for starting sound samples, such as shock noises or vocal warnings.

2.3.2 Network shared information

As shown in figure 1, each module in SCANeR is linked to a local network. Each module that is producing data is making it available to the simulation and is collecting data needed for its purpose on the network. In the case of the sound rendering process, relevant data is generated in the dynamics model, the traffic generation modules and the scenario.

The traffic generation modules are providing longitudinal speed, engine rotation speed and vehicle position and orientation in the world for each vehicle. The vehicle can be of various types (TRUCK, CAR, VAN, CYCLE) allowing to generate appropriate sound signatures. When attached to a given vehicle, the sound is positioned in the world and the Doppler effect is managed.

The dynamics model is producing the corresponding values for the interactively driven vehicle as well as slip, skid and road status independently for each wheel. The interface to the simulation of the driven vehicle has the same format as the one of any other vehicle, meaning that the sound generation process can be attached to any participant of the simulation, whether interactive or not.

The scenario run-time module is providing event-based information, configurable by the user. This includes collision information and vocal messages for directions to the driver and warnings.

2.4 Sound rendering specifications

2.4.1 Sound rendering architecture choice

The base sound rendering architecture that is used in the Renault driving simulators is shown in figure 3. This architecture is cost efficient and is well adapted to static simulators. The relation between the sound rendering system and the listener is not changed by motion of the cockpit in respect to the room. The sounds played by the speakers are naturally filtered by the vehicle cockpit before reaching the observer. It is therefore possible to use samples directly recorded from a real vehicle and use them in the simulator with no specific treatment other than assuring seamless looping.

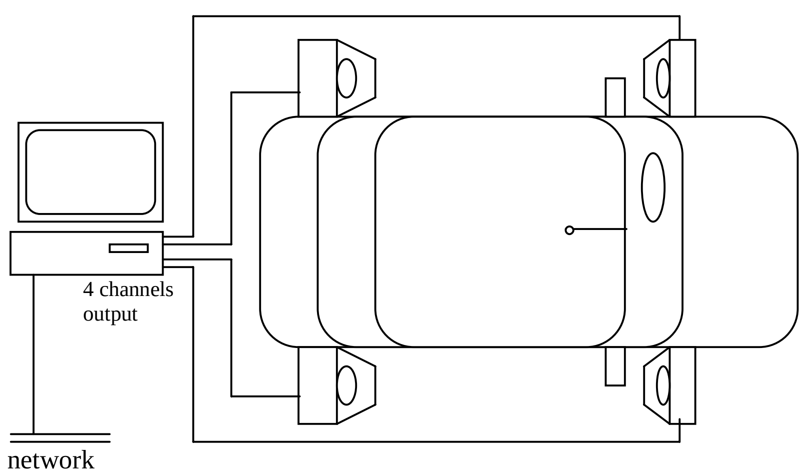


Figure 3 : Standard sound generation configuration

However, this configuration suffers from losses in power transmission, as the sounds are generated in an open space, often forcing the driver to open the windows of the cockpit to have a better perception of the sound environment, losing part of the filtering effect of the cockpit and the balance between the channels that generate the 3D effect.

The advantage of using a real cockpit is furthermore balanced by the drawback of not being able to simulate sound ambiences of vehicles other than the one that is used for the cockpit of the simulator, that one not necessarily being a real vehicle and therefore not having the same filtering characteristics. The coherence of the sound rendering with the vehicle model is therefore easy to loose, as it is a main feature of the simulator that the vehicle model is easy to modify. This incoherence can produce biases in the simulation sessions, especially when it comes to space perception.

Finally, the ideal position for the speakers in such a configuration would be in the same plane as the observer, in order to generate realistic cues for the aerodynamic sound and the surrounding vehicles. This position interferes with the display system, forcing the system designers to place the rendering system closer to the ground, in a position that makes the observer feel that all the sound is produced close to the ground. All the sound therefore feels like it is produced uniquely by the contact of the vehicle with the road.

When the simulator is taken onboard a motion platform, the standard configuration is further handicapped: the speakers cannot stand close to the cockpit without adding embarked weight, because of the range of motion. In order to avoid their interfering with the display system, they have to stand even lower than in a static configuration, increasing the degradation in the perception of the elevation of the sound. As the cockpit is put in motion by the platform (+/- 20° in all angles) a static source of noise will be felt as moving.

The option of taking the sound rendering system inside of the cockpit had been rejected in the past because available equipment did not allow it. High quality professional or semi-professional systems are voluminous and one cannot afford to insert them into a cockpit, especially on the driver's side. Standard vehicle sound systems, that are perfectly adapted to vehicle volumes, are on the other hand limited in performance and do not allow to feed-back to the driver all the dynamic of a vehicle noise ambience. However, Renault developed in co-operation with a provider a high bandwidth compact sound restitution system for use in the Clio2 "MTV" special edition in year 2000. This system provides higher sound dynamics and opens the way for embarking the sound restitution inside the cockpit.

This solution allows for controlling the sound generation direction in respect to the driver, to control the specialisation of the restitution depending on the vehicle that has to be simulated and the characteristics of its noise response, through adapted pre-processing filtering.

2.4.2 Real-time processing hardware and software choice

The architecture of the sound rendering is based on widely available PC hardware and software. The consumer market based on PC architecture is very dynamic and research results that previously took years to be translated into hardware are made available to users in a matter of months. The PC hardware being available on operating system such as Microsoft's Windows or open source Linux, with generic Application Programming Interfaces (APIs) have therefore a good functionality to global-implementation-cost ratio compared to professional systems.

The APIs that are chosen for the application, Microsoft's DirectSound3D and Creative Labs' EAX, have become *de facto* standards of the industry, and are therefore supported by most of the manufacturers, making them a good development platform, with frequent updates. Even if today there are limitations in the modularity of the API and the possibility to insert one's own processing steps, this modularity is coming in future releases. It is therefore possible to invest development time on an application that will be directly portable on new hardware when made available. The range of available hardware products is also giving flexibility in the choice of a price-point when building up a new simulator for customer use, with a unified software equipment.

Those APIs are providing access to hardware acceleration to a range of sound effects such as sound positioning, sound pitch shifting, automatic management of the Doppler effect, automatic attenuation with distance, management of 3D positioning and reverberation that are of use in our sound system implementation.

3 GENERAL REQUIREMENTS FOR 3DSOUND REPRODUCTION / SYNTHESIS

Figure 4 shows the different steps of the general architecture proposed for processing the 3D simulation. The successive functions are written into the different blocks with their associated DSP operations. The blocks are ordered according to the successive physical transformations of sounds external to the cockpit and that shall be reproduced on the vehicle loudspeaker system.

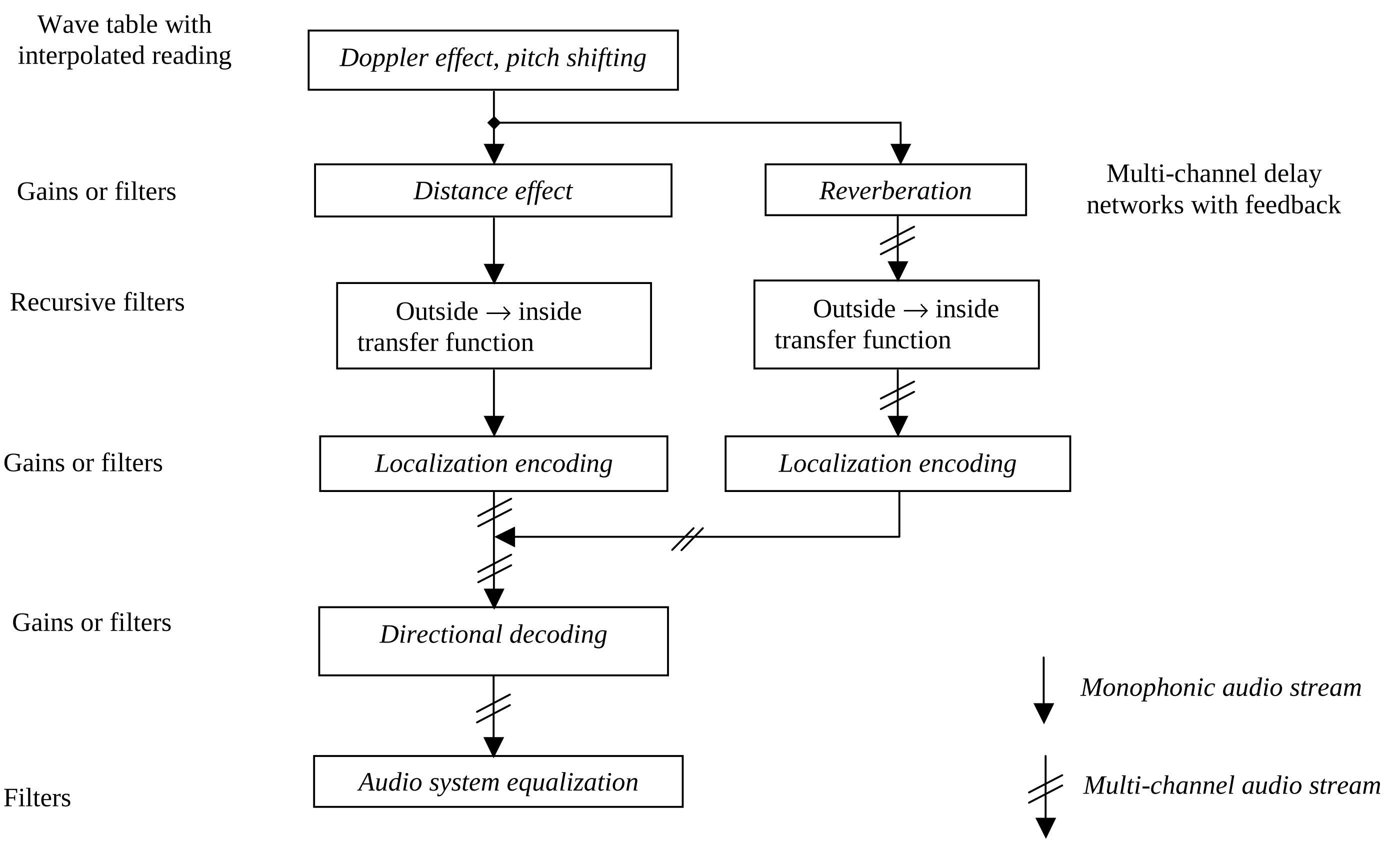


Figure 4: General architecture of sound processing

Doppler Effect, pitch shifting

In order to reproduce the Doppler effect linked to moving sources (related to the vehicle position) the recorded samples are read from a looped buffer with varying speed of the table lookup. As the table lookup is not driven with the same sampling rate as the original one, this process requires an interpolation filter. The same algorithm allows simulating the RPM of the motor from a limited number of samples recorded with constant RPM.

Cabin related transfer function

This step corresponds to the transformation of external sounds heard inside the cockpit. This transfer function will depend on each specific vehicle and also on the recording conditions. In the case of reverberation effects (tunnel, urban environments, effect of road-side objects such as fences...), such transfer function will also be required. However, it will probably be different from the one used for the filtering of the engine, because acoustical conditions of coupling are not the same.

Directional localization encoding

This feature is mainly required for reproducing opposing and overtaking vehicles, or aerodynamic noise. However, because of the conventional layout of loudspeakers in a car cabin, conventional pan pot laws may not be very convincing. For example, as the driver is located near the side front speaker, the directional reproduction may be strongly distorted, and could lead to a monophonic sensation. Different signal processing operations can be envisioned to overcome this drawback. A de-correlation process can be proposed. In order to avoid a mismatched sensation of elevation, due to the layout of the speaker below the head, a binaural encoding could be applied.

Directional decoding
Some of the localization techniques are based on a multi-channel format that must be decoded first before feeding the reproduction setup (ambisonic, transaural).

Reproduction equalization
This step is required for compensating the loudspeaker spectral response and time misalignment, together with the spectral correction of the simulator cockpit.

Owing to the specific application of vehicle simulation, some simplification can be proposed in the previous block diagram.

The main modifications concern grouping the different filtering steps into single module. Figure 5 shows the new block diagram where this global equalization module has been moved before the car cabin transfer function. This modification is only possible if the cockpit compensation is not made independent from the loudspeakers. In case different sound sources are synthesized simultaneously this module should be shifted, on the contrary, to the last step of the process.

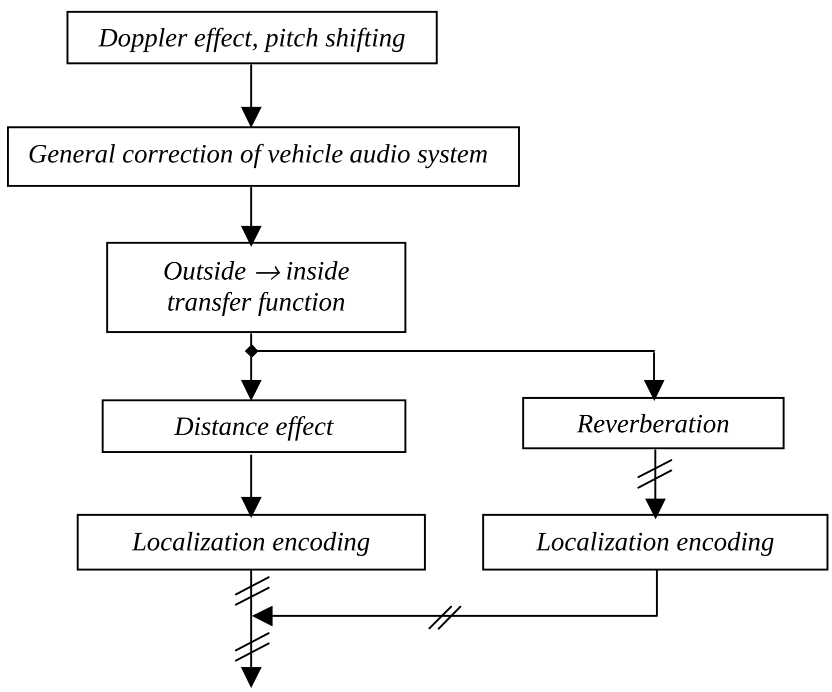


Figure 5: General architecture of sound sources processing, with vehicle equalization module moved upstream of the 3D sound processing

The same modification can be applied to the cabin related transfer function, provided that it is supposed to be common to all sound samples. Hence, all spectral corrections can be grouped into a single one to save calculation power.

The second modification concerns the decoding step which has been removed in the case of a conventional pair-wise panning technique for localization synthesis.

3.1 Software implementation

The software processing is first envisioned within the Windows environment using DirectSound 3D API from Microsoft and EAX from Creative Labs. This facilitates the programming of the application for it provides automatic mixing of the different sound streams and pre-programmed acoustic functions like pan-pot laws, Doppler effect that depends on the sound scene control parameters (source position and movements). However, the current state of the chosen APIs (at time of simulator development) showed some limitations with regard to the planned processing. Figure 6 shows the general DS3D architecture. Sound files, associated to the different sources of the auditory scene are loaded into secondary buffers (3D sound objects).

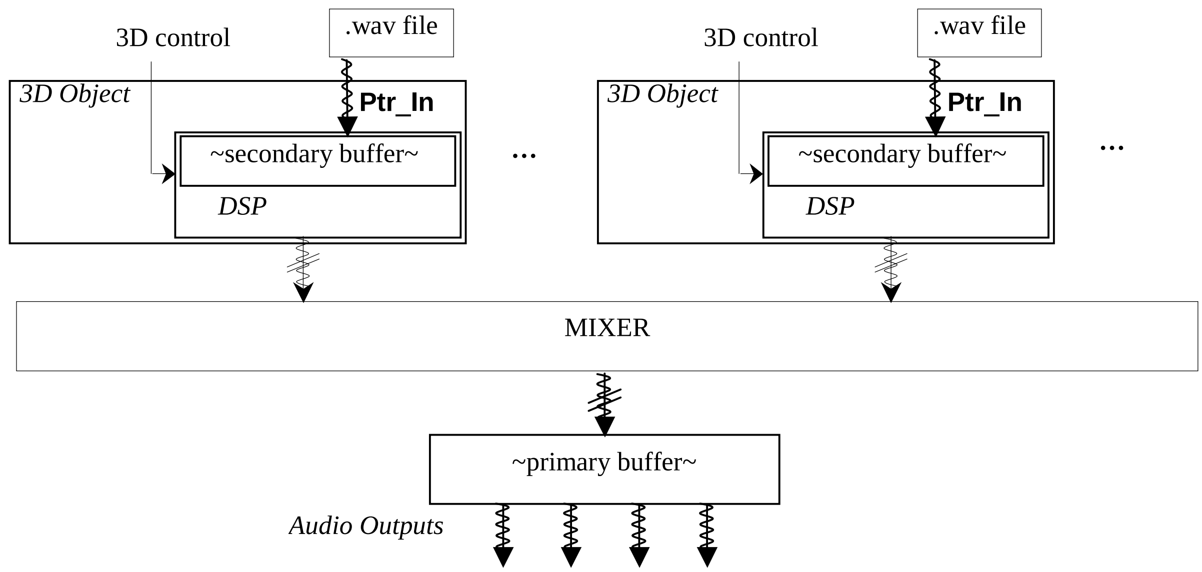


figure 6: General architecture of the 3DSound API

According to the control parameters they are processed in the DSP board (pitch shifting, filtering, pan-pot). The multi-channel outputs are then mixed in the primary buffer before being sent to the audio output of the board. The only way to insert specific processing is between the sound file reading and the input of the secondary buffer. This constrains the specific processing to functions that can be calculated before any operation taken in charge by the API. Another possibility would be to insert one's operation before the input of the primary buffer, but this would mean re-programming all the provided functions.

These constraints require preprocessing the sound samples with the different spectral corrections linked to the cockpit-related transfer function and to the cockpit compensation. Unfortunately, as was shown on figure 4 and 5, Doppler effect and pitch shifting should be executed before any other process. The main drawback of processing first the spectral correction is that this one will be pitch shifted during the simulation of the RPM variations instead of being kept constant. This behavior is not physically correct and can lead to non-relevant perceptual effects. In order to work around this drawback a solution is to increase the data base of RPM samples so that each one is used within a narrow speed range (thus avoiding high table lookup speed ratio).

Another limitation of the API is the constrain on the loudspeaker positions which can not take into account the actual lay out of cockpit transducers. Moreover this does not allow the use of localization format that need a decoding step that should be inserted at the very end of the processing, just before the audio output.

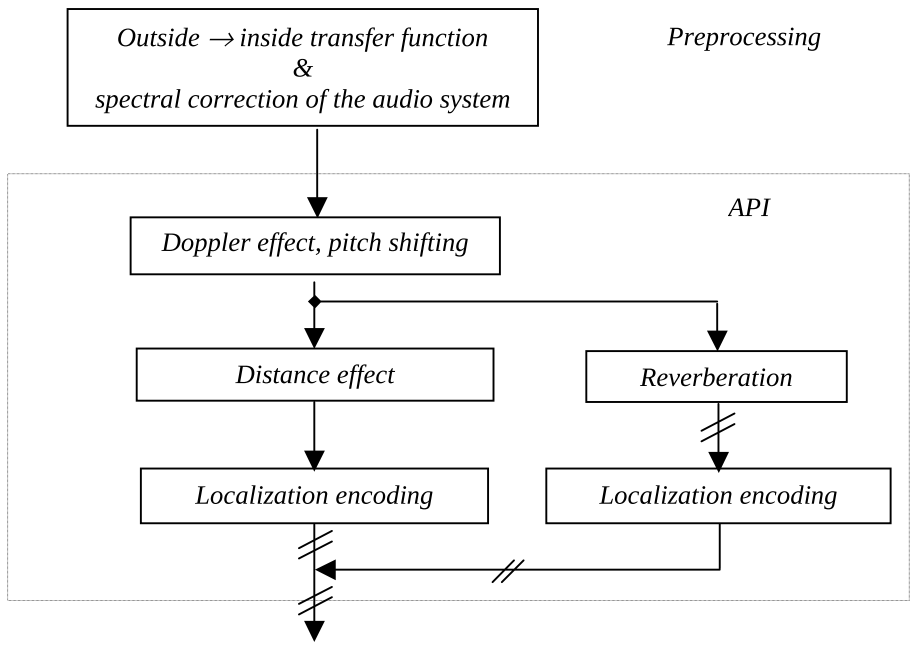


Figure 7: General architecture of sound processing currently available with the chosen APIs

Control and Signal processing proposed by the API

The main control parameters provided by the API are :

- position, orientation, and speed of the source (absolute or relative)
- position, orientation, and speed of the receiver
- simple directivity model
- roll off factor for the simulation of sound attenuation with distance

The main signal processing operations required for the spatialization process are :

- gain control (pan-pot laws, attenuation, ...)
- encoding techniques for directional laws
- filtering (air absorption simulation, binaural techniques,...)
- fractional delay (Doppler effect, some directional encoding techniques,..)
- delay lines (distance, room effect,...)
- reverberation

Most of these functions are provided by the API. However, a possibility to insert user plug-ins would allow more flexibility into applying the treatments that our application requests.

4 RECORDING AND PROCESSING THE SOUND SAMPLES

In this paragraph we present how the different sound samples are recorded and processed in order to apply the necessary spectral corrections. Examples are focused on the processing of RPM samples, but the method is also applied to the other samples.

4.1 Sound samples

Sound samples of the engine have been recorded on a car category similar to the simulator cockpit (Renault Clio 2). The vehicle was put on a roll bench in order to reproduce the conditions of a loaded engine. The following recording setup was used to record the sound samples and to get the necessary information in order to estimate the car cabin related transfer functions.

Figure 8 shows the different recording transducers used for the recording of the different samples. The signals that will be used for the simulation are recorded with a pressure microphone positioned in the bonnet of the vehicle and with an accelerometer. The signals dedicated to the estimation of the cabin related transfer function are recorded with a Bruel&Kjaer microphone and a Head Acoustics dummy head, both set at the driver position. These transducers will be used respectively for a reproduction on the cockpit loudspeakers or on headphones.

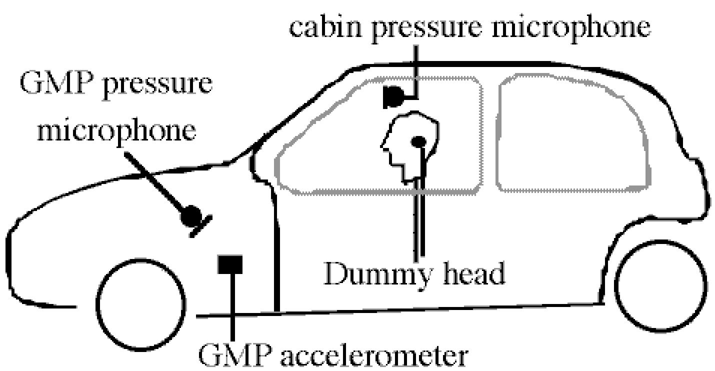


figure 8: Localization of the sensors used for engine noise measurement : cabin microphone (B&K 4165 ½ inches), dummy head (HEAD ACOUSTICS HMS II), GMP microphone (B&K 4165), and GMP accelerometer

4.2 Cockpit characterization

In order to compensate for the restitution setup of the cockpit, measurements have been made using the same microphones (Bruel&Kjaer and Head Acoustics Dummy head). The characterization is undertaken with impulse response measured between each of the loudspeakers and the driver position. This setup is presented in figure 9, and the processing is described in the following paragraph.

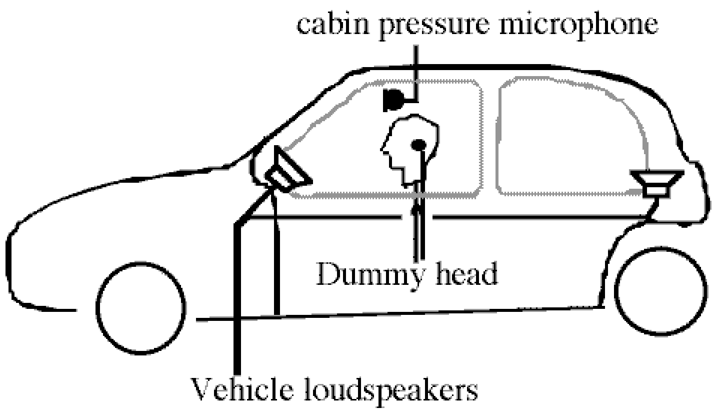


figure 9: Localization of the sensors used for audio system characterization.

4.3 Filtering process

According to block diagram 4, the spectral correction is made of two successive steps. The fist one is related to the transfer function between outside and inside the car cabin, while the second is dedicated to compensation of the reproduction system setup.

Equations 4.1 to 4.6 describe the different transfer functions corresponding to the recording vs. reproduction process. The reference is the signal captured by an ideal microphone located at the driver position. This signal is noted X(t).

Recording this signal with a microphone introduces a first filtering step corresponding to the microphone response. This signal is written :

$$S(t) = X(t) * BK(t) \quad (4.1)$$

where BK(t) is the response of the microphone.

When reproducing the signal S(t) on the cockpit loudspeakers it will introduce different transformations linked to the loudspeakers and to the cockpit effect. This transfer function is described in equation 4.2.a

$$Y(t) = S(t) * [\Sigma_i HP_i(t) * V_i(t)] \quad (4.2a)$$

where HP_i(t) describes the open air response of the loudspeakers mounted in their support (doors, dash-board), and V_i(t) corresponds to the transfer function between these loudspeakers and the listening point. Equation 4.2.a may be written in a more synthetic form:

$$Y(t) = S(t) * R(t) \quad (4.2b)$$

where $R(t) = [\Sigma_i HP_i(t) * V_i(t)]$ is the audio system "meta-response".

Looking for a transparent reproduction system requires to inverse the function R(t) that will be noted R´(t). On a practical stand point this response R(t) is measured with the same microphone as the one used for recording the sound samples. Hence, we get a set of responses (one for each loudspeaker) noted :

Q(t) = R(t) * BK(t), which inverse is Q' (t)= R´(t)*BK-1(t)

Inserting Q´(t) in the reproduction process leads to :

Y(t) = S(t) * R(t) * Q´(t) = X(t)*BK(t)*R(t)*R´(t)*BK-1(t) which is equivalent to

$$Y(t) = S(t) * R(t) * R'(t) \tag{4.3}$$

Equation 4.3 shows that using the same microphone for recording the sound samples and for cockpit characterization allows to compensate for its possible effects. If R(t) may be inverted, then the reproduction set up can be made transparent. However, before studying the inversion of R(t), the validity of the sound samples S(t) must be evaluated. On a theoretical stand point they seem valid, however, the actual sound samples were disturbed by periodic perturbations linked to the bench rolls.

<u>Outside cabin transfer function vs. Inside cabin transfer function</u>

An alternative solution consists of substituting these sound samples with the ones captured in the bonnet, provided that they would be less disturbed by the rolling noise. This solution, which have been preferred to the previous one, needs to evaluate a transfer function between outside recordings E(t) and inside signals S(t). This allows to introduce the car cabin related transfer function noted H(t) .

$$S(t) = E(t) * H(t)$$

Theoretically, this function H(t) cannot be interpreted as the real bonnet vs. cabin transfer function. As a matter of fact this one is complex because it is related to different coupling functions (air propagation, solid transmission), that could be non linear and independent from the particular RPM. Hence, E(t) represents a specific "view" of the engine excitation. In order to try different situations the samples were simultaneously recorded with a pressure microphone and an accelerometer that would allow to introduce a combination of these transducer in order to obtain a more reliable characterization as described in equation 4.5

$$S(t) = [\Sigma_i E_i(t) * H_i(t)] \tag{4.5}$$

where $E_i(t)$ represents different sensor types or positions, and $H_i(t)$ the corresponding transfer functions.

In conclusion, the signal reproduced at the driver location is written as :

$$Y(t) = E(t) * H(t) * R(t) * Q'(t) \tag{4.6}$$

where H(t) and Q´(t) have to be estimated in order to proceed to the simulation of the engine noise.

4.4 Transfer function estimation

4.4.1 Outside vs. Inside cabin transfer function

The transfer function allowing to synthesize the engine noise heard inside the cabin from external recordings leans on the computation of the ratio between their respective spectrograms. However, this spectral division may lead to artifacts because the spectrogram of an engine RPM shows sharp resonance and other low level areas that can produce a bad signal to noise ratio.

A solution could be to compute this transfer function from RPM sweeps that would cover a large spectrum, hence avoiding singular division. In our case, an intermediate solution was used, averaging the different RPM samples that were recorded at 1000, 2000, 3000, 4000, 5000 and 5500 rpm.

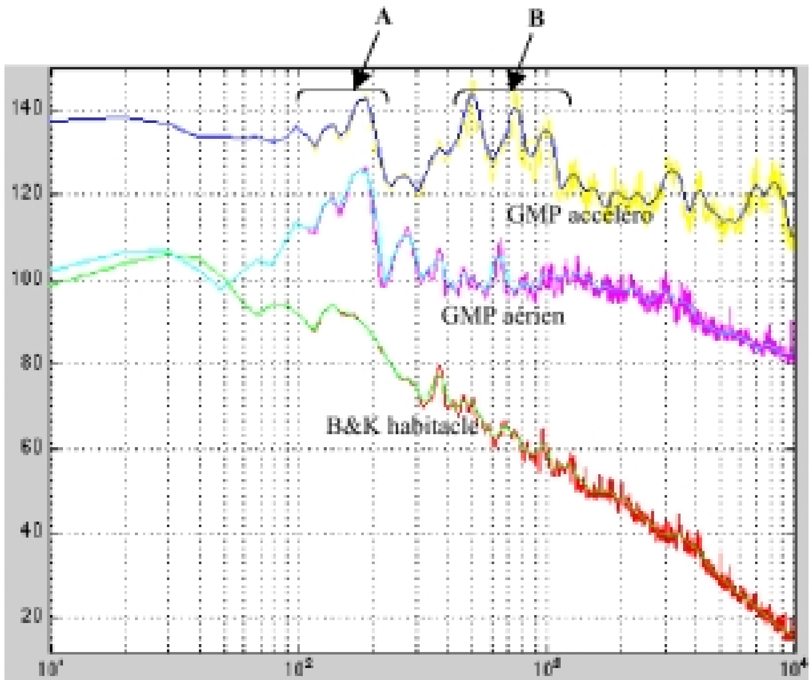


Figure 10 : cumulated and averaged spectra on the various engine RPMs, as measured by the accelerometer (blue), aerial microphone (cyan) and the B&K microphone inside the cabin (green).

Figure 10 shows these cumulated and averaged spectrograms captured with the pressure microphone (GPM aerien), the accelerometer (GPM accelero) located in the bonnet and the pressure microphone (B&K) located in the cabin. For each microphone results are shown with raw and smoothed data.

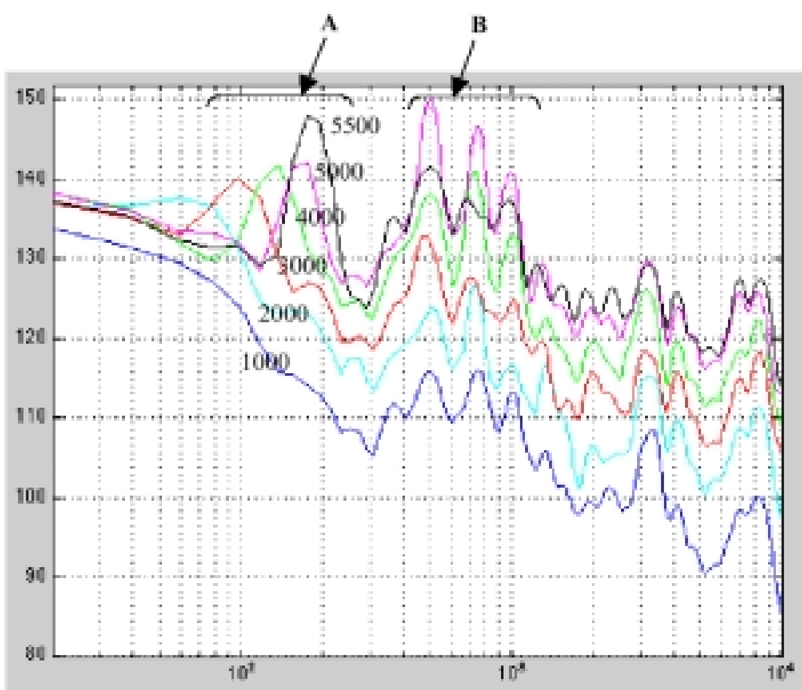
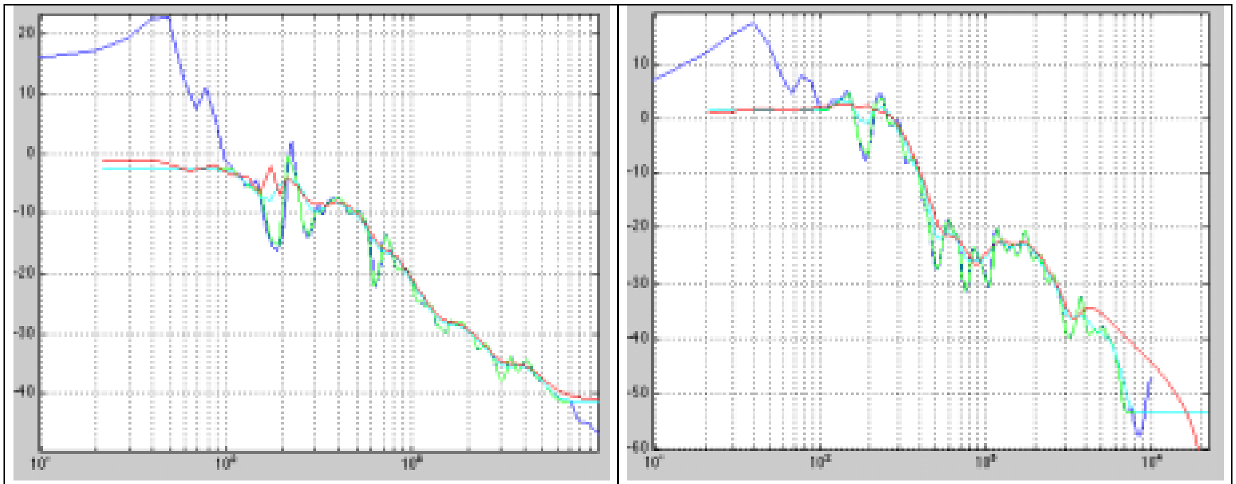


Figure 11 : cumulated spectra for the various RPMs as recorder by the accelerometer.

As cumulated spectrograms have been estimated using a limited number of RPM samples and not a continuous sweep one can wonder if the resonance still observed in regions A and B are revealing a general behavior or are only related to the specific recorded samples. A detailed analysis shown on figure 10 allows observing that in region A each RPM shows a specific resonance linked to the rotation speed, while in region B the resonance stay constant and are less speed dependant. Hence, the resonance between 100Hz and 300Hz are RPM dependant and should be smoothed. On the contrary the resonance in the frequency range [400-1000Hz] should be preserved in the transfer function estimation.

Therefore, the first solution to estimate the transfer function H(t) of equation 4.6 is to compute the ratio between the spectrogram of the reference microphone inside the cabin and the external microphone spectrogram after averaging all the RPM samples. Figures 12.a and 12.b show the results when considering the pressure and the accelerometer transducers. For each transducer, the figure shows the raw magnitude of the transfer function and different derived magnitude curves when applying smoothing, frequency band limitation and parametric estimation with a 12th order ARMA filter.



12.a **12.b**

Figure 12: Corrective filters, cabin/exterior. 12.a aerial microphone. 12.b Accelerometer microphone. Direct transfer function (blue),limited in frequency (green), smoothed (cyan), estimated with a 12th order ARMA filter (red).

This first solution presents the advantage of computing a common filter for all RPM samples and to avoid spectral division artifacts. On the other hand, this solution leans on the hypothesis that the transfer function is independent from the RPM, which can be wrong when considering the complex coupling vibration modes of the vehicle. An alternative solution has also been studied, considering a specific transfer function for each RPM, applying smoothing and thresholds functions on each RPM spectrogram in order to avoid singularities during the division. Both solutions are currently under validation tests.

4.4.2 Equalization of the reproduction setup

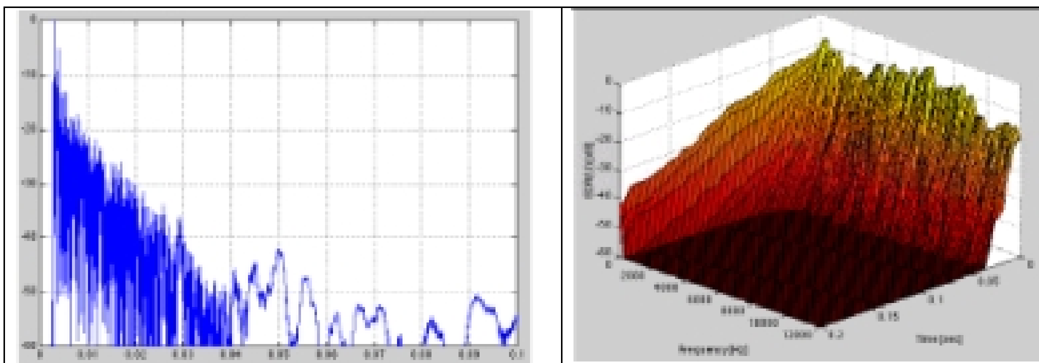
As mentioned in the previous section, the reproduction setup must be equalized in order to ensure its transparency. Unfortunately, this equalization is difficult for it includes the response of the car cabin. As in any closed space a direct de-convolution of the enclosure response is difficult because of the presence of zeros in the spectrum. Moreover, such a direct de-convolution would only be valid for a very limited area in the cockpit and would be wrong as soon as acoustical conditions vary (temperature, presence of the driver...).

We therefore compensate for the global magnitude spectrum independently from the time response of the vehicle which is anyway very short (< 50ms). However the choice for a magnitude characterization is not obvious for the cabin response, like any enclosure, is characterized by a time vs. frequency distribution of energy. If we consider the response of the enclosure to be roughly composed of a direct sound and a reverberation field we can envision different approaches depending how we consider our perception of timbre to be mainly driven by :

- the direct sound,
- the global energy
- the reverberated energy

On an objective stand point this choice may depend on the relative importance of the energy associated to the direct or reverberated energy. In a car cabin, in spite of a short time response both energy are similar and it has been shown that our perception is able to segregate the spectral response linked to the diffuse field and to the direct one, even in car cabin conditions [10].

The frequency contributions of the direct sound and the diffuse field have been estimated, using a time frequency analysis described in [11]. From a short time Fourier transform applied to the impulse response the time vs. frequency distribution of energy may be computed, and various spectral quantities may be derived such as the direct sound spectrum or the diffuse field spectrum. Such analysis is illustrated in figure 13 which shows for a given loudspeaker the echogram and the time vs. frequency distribution associated to the corresponding impulse response.



13.a 13.b

figure 13: 13.a loudspeaker echogram. 13.b time vs. frequency distribution associated to impulse response

Figure 14 shows the spectral densities associated to the direct sound, to the global energy and to the diffused energy after averaging the response of each loudspeakers of the cockpit. The main differences are located in the low frequency band [100-800Hz].

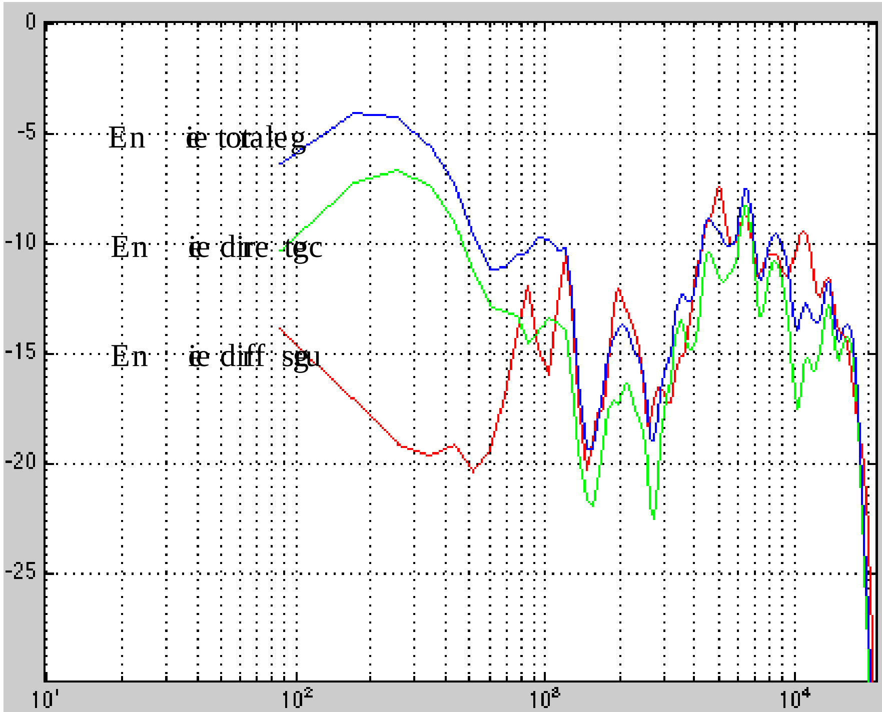
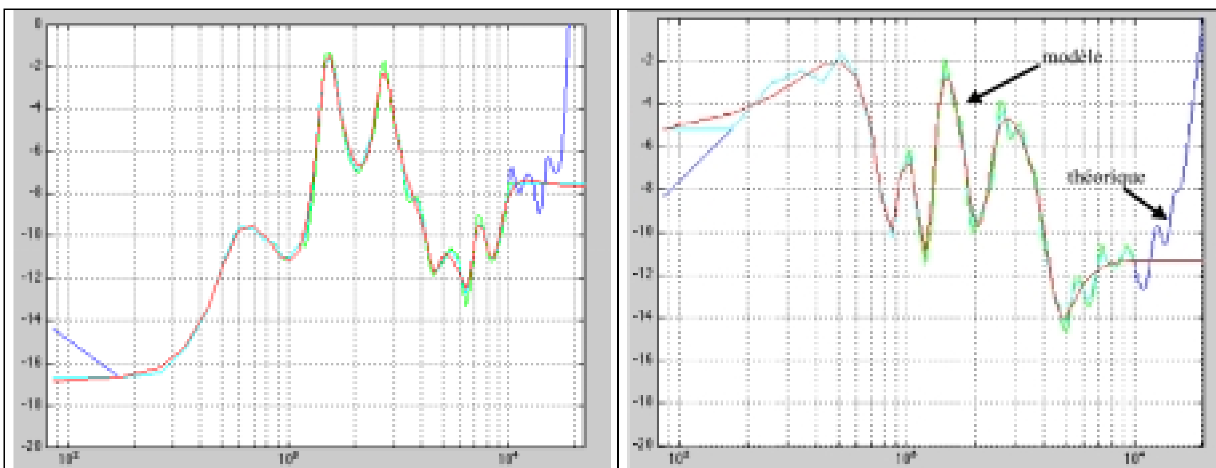


Figure 14 : spectral densities associated to the direct sound (green), to the global energy (blue) and to the diffused energy (red) after averaging the response of each loudspeakers of the cockpit

The choice for the most relevant spectrum is not obvious. On the one hand, engine noise being stationary, it would lead to choose the global energy spectrum. However, the sensation of engine noise in a car is quite diffuse, hence it could also be relevant to select the diffuse field spectral characterization. Perceptual validation tests are being performed to decide which is the most relevant solution.



15.a

15.b

figure 15: correction filter magnitude derived from the global (15.a) or diffuse (15.b) field spectral characterization. Direct transfer function (blue),limited in frequency (green), smoothed (cyan), estimated with a 12^{th} order ARMA filter *(*red

Figure 15.a and 15.b show the correction filter magnitude derived from the global or diffuse field spectral characterization. Once again, the raw transfer function is presented together with derived

magnitude curves corresponding to smoothing and band limiting functions and to an estimated parametric filter of order 12.

4.5 Validation

As explained in section 3, the current implementation of the correction filter is made by pre-processing the different RPM samples. These samples are then controlled by the SCANeR II real-time software according to the driving conditions.

Objective and subjective tests are currently undertaken in order to investigate the performance and validity of the different correction solutions. Objective tests consist in recording the results with the same cabin microphone and with the same RPM as during the characterization session. Non-formal perceptual tests are undertaken as well.

The first conclusion is that the reproduction equalization driven with the global spectral correction is the most relevant. It can be shown that considering a common cabin transfer function for all RPM was not totally convincing on both objective and perceptual standpoints. Further investigation are being made in order to test the solution where a specific transfer function is applied for each RPM sample.

5 CONCLUSION AND FURTHER DEVELOPMENTS

The processing of samples in real-time, with PC hardware proved successful to give the cues the pilot expects in the driving task. The pre-processing applied to the samples allows to abstract the characteristics of the recording and rendering devices, and therefore to spatialize the sound, immersing the driver more completely into an environment with positioned sources. Finally, a procedure was defined that can be transposed to any simulated vehicle – physical driving station couple.

To some extent the response of the vehicle to world and vehicle sounds can be taken into account. Some information lost to the recording method that was used for establishing the transfer function of the cockpit could not be regenerated through post-processing. A different method will be used in the next steps of development allowing to use more completely air and solid borne energy.

In applications using a virtual reality helmet, the physical driving station can be reduced to the few elements needed to give tactile feed-back. Those applications are pushing virtual prototyping and the associated flexibility a step further, allowing to change the configuration of the cockpit very quickly, or even in real-time. In that environment, the method that is presented here would give full control over the sound perception for the driver.

Improvements to that technique for use in a virtual reality environment, would include generating the sound environment into a pair of earphones, allowing to use effectively binaural sound generation techniques. Those techniques, based on Head Related Transfer Functions (or HRTF) allow for an even more accurate generation of spatial cues based on the knowledge of the geometry of the head and of the auditory channels. Binaural techniques generate elevated and more precisely positioned sources. However, earphones require tracking of the head movements in order to have the sound sources stable in the world when the observer is moving his head. The influence of transport delay in the re-positioning of the sources would therefore become a major axis of study.

REFERENCES

1. R.W. Allen, T. Rosenthal, B. Aponso (Systems Technology, USA), A. Harmsen (A&A, The Netherlands), S. Markham (Valentine Technologies, UK), Low cost, PC-based techniques for driving simulation implementation, DSC 1999 Proceedings.
2. Erwin R. Boer, Ahna R. Girshick, Tomohiro Yamamura, Nobuyuki Kuge (Nissan Cambridge Basic Research, United States), Experiencing the same road twice: A Driver Centered Comparison between Simulation and Reality, DSC 2000 Proceedings.
3. Gilles Reymond, Andras Kemeny, Motion Cueing in Renault driving Simulator, Vehicle System Dynamics, In Press, 2000.
4. P. Lecocq, J.-M. Kelada, A. Kemeny, Interactive Headlight simulation, DSC 1999 Proceedings.
5. Omar Ahmad, Yannis E. Papelis, Matt Schikore, Ginger Watson (The university of Iowa, United States), An autonomous Driver Model and Graphical Authoring of Scenarios : Advanced Traffic Simulation Tools for the NADS Simulator, DSC 2000 Proceedings.
6. M. Fernandez, I. Coma, G. Martin, S. Bayarri (University of Valencia, Spain), A hierarchical object oriented architecture for the management of complex driving simulation scenarios, DSC 1999 Proceedings.
7. Alain Flipo (TT&S, France), TRUST: The Truck Simulator for Training, DSC 2000 Proceedings.
8. F. Panerai, J-M. Kelada, A. Kemeny, Séminaire Renault juin 2001.
9. Gilles Reymond, A. Heidet, M. Canry, A. Kemeny (Renault, France), Validation of Renault's Dynamics Simulator for Adaptive Cruise Control Experiments, DSC 2000 Proceedings.
10. J.-M. Jot, L. Cerveau and O. Warusfel, "Analysis and synthesis of room reverberation based on a time-frequency model", AES 103rd convention preprint, #4629, September 1997.
11. G. Vandernoot and O. Warusfel, "Spectral optimization of listening conditions in car cabins", proceedings of ICA2001, September 2001.
12. A. Bellini, G. Cibelli, E. Ugolotti, F. Bruschi, A. Farina, "Experimental validation of equalizing filters for car cockpits designed with warping techniques", AES 109th convention, #5278, September 2000