Towards the timbre modeling of interior car sound

Jean-François Sciabica,, Marie-Céline Bezat, Vincent Roussarie, Richard Kronland-Martinet, Sølvi Ystad

TOWARDS THE TIMBRE MODELING OF INTERIOR CAR SOUND

J-F. Sciabica°, M-C. Bezat*, V. Roussarie*, R. Kronland-Martinet°, S. Ystad°*

*PSA Peugeot-Citroën, 2 route de Gisy, 78943 Vélizy-Villacoublay, France
{jean-francois.sciabica, marieceline.bezat, vincent.roussarie}@mpsa.com
° LMA, Centre National de la Recherche Scientifique, 31 chemin Joseph-Aiguier, 13402 Marseille cedex 20
{kronland, ystad}@lma.cnrs-mrs.fr

ABSTRACT

Quality investigations and design of interior car sounds constitute an important challenge for the car industry. Such sounds are complex and time-varying, inducing considerable timbre variations depending on the driving conditions. An interior car sound is indeed a mixture between several sound sources, with two main contributions, *i.e.* the engine noise, the aerodynamic and tire-road noise. That's why masking phenomena between these two components should be considered when studying perceptive attributes of interior car sounds. Additive synthesis is used to simulate the harmonic engine noise. Nevertheless, this synthesis is controlled by a large number of parameters and no relation between these parameters and their perceptive relevance has been clearly identified. By combining sensory analysis and signal analysis associated with an auditory model, we can find a relation between a reduced number of signal parameters and perceptive attributes. This study develops a method to simplify the timbre description of interior card sounds and presents the first results of auditory model application on such sounds.

1. INTRODUCTION

Since global noise level in and outside cars has considerably decreased in the past decades, interior car sounds are no longer considered as a discomfort, but rather as a sound design question. Interior car sounds indeed contribute to the identity and the perceived quality of the car.
Sounds perceived in car passenger compartments are the result of three well known acoustic sources: the engine source – called the harmonic part-, the tire-road source and aerodynamic source –called the noise part.
In order to study perceptive attributes of interior car sounds, we would like to develop a perceptively based synthesis model that can be used as a tool for the investigation of the perceptive relevance of the signal parameters. The identification of these parameters will make it possible to directly control the timbre of such sounds.
The first step towards the construction of our synthesis model consists in determining the components of the harmonic part which are actually audible in the noise part. A considerable amount of information contained in the harmonic part can be reduced by considering masking phenomena. In addition to the masking between the noise part and the harmonic part, masking phenomena between harmonics should also be considered due to the small frequency range between engine harmonics

In the literature, the investigation of car sound quality has most often been studied through investigations of perceptive aspects and evocations. Bisping [1] described two perceptive factors forming a four quadrant scheme of sound quality with one axis defined by pleasant-unpleasant" ant the other axis defined by "powerful-weak" Kubo [2] proposed a car sound map with two axes "not sporty…sporty" and "simple...luxurious". A semantic diffential method and a factor analysis method were used to create their sound map. With semantic differential techniques, the meaning of engine sounds has been widely investigated and a quantity of adjectives was used to describe this kind of sounds. Chouard [3] proposed a list of 232 pairs of adjectives to illustrate the sound of an engine noise.
However, some studies aimed at finding the link between perceptive attributes and signal parameters. Hansen [4] analysed the impact of tonal components on the sound quality. With an algorithm based on DIN 45681, he identified and extracted engine harmonics that are considered as prominent in interior car noise and showed their relation to the perception of tonal phenomenon like whistling.

In a previous study [5], we proposed a procedure to predict audibility threshold of harmonics produced by the engine noise. Our algorithm is a linear model based on the knowledge of noise level in the critical bands centered on the harmonic frequency. We reproduced here an audibility threshold experiment [6] with stimuli and conditions adapted to automotive context.
Indeed, the psychoacoustic approach seems to be the right way to get a robust description of interior car sounds. Bezat [7] also proposed an original analysis/synthesis approach to investigate the link between the quality of car door closure noise and signal parameters. The perceptive aspects of such sounds were first evaluated with a sensory panel. This method consists in finding acoustic descriptors to qualify the sound. Then, the panel can rate the sound for each descriptor and create a sensory profile. Secondly, a time/ERB description was chosen to analyze the signal. Combining these two ways to analyze signals, she proposed a synthesis model controlled by perceptive criteria.

Hence, in order to apply this approach to interior car sounds, we analyzed our signals with an auditory model in order to visualize harmonic components which are audible with the noise part. Future investigations on the output of the auditory model should lead to a good description of masking phenomena between the harmonic part and the noise part.
In this paper we describe the specificities of interior car sounds and the different source contributions. We then describe the signal using a time-frequency analysis technique and a synthesis

tool. Further, the perceptive analysis of the signal effectuated with a sensory panel is presented, and finally the auditory model is applied to the signal to point out perceptually relevant signal components.

2. THE INTERIOR CAR SOUND

2.1. Interior car sound spectrum

2.1.1. From the piston rotation to the spectrum of the harmonic part

Good knowledge of motor mechanics improves the comprehension of automotive sound. The harmonic part of an interior car sound is the result of the piston rotation. These harmonics are called engine harmonics (with the prefix H) and the number of an engine harmonic is called engine order.

If we consider the case of a 4 cylinders motor, pistons fire successively every half motor rotation so a given piston fires every two motor rotations. The periodicity between two firing pistons is not ideal, that's why we have to consider these periodic phenomena (successive firing pistons and a given firing pistons)

Nevertheless, we consider that the spectrum fundamental called H1 corresponds to a motor rotation even if this phenomenon doesn't physically generate the harmonics. It's our way to combine the two phenomena which generate their own fundamental and their own partials. If we call *Nrpm* the number of motor rotations per minute, H1 is given by the following relation:

$$H1 = \frac{Nrpm}{60} \qquad (1)$$

The first phenomenon, the most energetic, corresponds to the periodicity between two firing pistons (a half motor rotation) and generates harmonics called "odd harmonics" with it own fundamental frequency called H2 which is twice H1 (H2 = 2xH1) and with its own partials (H4 = 2xH2 = 4xH1, H6 = 3xH2=6xH1, H8 = 4xH2=8xH1). These harmonics are very often the most energetic in the spectrum. If the periodicity between firing pistons were ideal, engine sounds should only contain these odd harmonics.

So we have to consider a second phenomenon, the periodicity of one given firing piston which correspond to 2 motor rotations and generates harmonics called "semi-harmonics ". It own fundamental frequency is called H0.5 which is half H1 (H0.5 = 0.5xH1).
These harmonics are less energetic, but their contribution to the sound phenomenon cannot be neglected. They also respect the same labeling and numbering (H1=2xH0.5, H1.5=3xH0.5, etc.)

2.1.2. Aerodynamic noise and tire road noise

The noise part of an interior car sound is composed of aerodynamic noise and tire road noise. Aerodynamic noise is a broadband noise whose global sound increases with speed. It mainly has a low frequency energy distribution (below 400 Hz), but its contribution to the perception is also important in the high frequency domain. Indeed, aerodynamic noise mainly masks high engine orders (~>12), but its impact can also be observed at low engine order. Tire-road noise depends on three main parameters: car speed, tire texture and road texture. The contact between tire and road generates low frequency noise.

2.2. Impact of driving situation on the interior car sound

The spectrum of an interior car sound also depends on the driving situation. According to the speed of the car, the balance between the noise part and the harmonic part will change. At slow motion, the harmonic part is predominant, while at fast motion, the noise part becomes more and more audible.
Moreover, the distribution of the harmonic amplitudes can vary rapidly in time. Indeed, engine sound transfer from the motor to the car interior depends on frequency. Some harmonics which were clearly audible at a given rotation speed can be less energetic at an other one. The balance between harmonics will also be modified in time, changing the timbre perception. In this case, masking phenomena between harmonics must be considered.

During an acceleration, the engine speed varies from 800 rpm to 6 000 rpm, which means that H2, the fundamental frequency of odd harmonics spans from 27 Hz to 200 Hz. Similarly, the frequency gap between semi-harmonics varies form 6 to 50 Hz creating good conditions for roughness phenomena. The density of semi harmonics in an ERB band is indeed important. For example, engine harmonics H3.5, H4, H4.5 and 5 are separated by 25 Hz at 3000 rpm. They belong to the same ERB band and interact to create sound modulations. The resolution of harmonics by the auditory system will therefore depend on the engine speed and the corresponding frequency spans of the harmonics.

3. ANALYSIS TOOLS

3.1. Time/ Frequency Analysis

In order to study the frequency evolution of the signal obtained inside an accelerating car, we performed a time-frequency analysis (short Time Fourier with a window length of 0.3 seconds and a window step of 10 ms)
The signal was recorded in an accelerating car with 100% accelerator load (*i.e.* full acceleration during the whole acceleration period) and a dummy head in the passenger position. The chosen frequency range allows visualizing the harmonics until the 12th engine order.
The spectrogram on figure 2 shows that a lot of harmonics are physically present in the signal. We can notice that many semi-harmonics contribute to the signal (H4.5, H5, and H5.5). We also notice that some harmonics emerge during the acceleration. However, this kind of representation does not tell us which harmonic is responsible for eventual timbre variations of the interior car sound

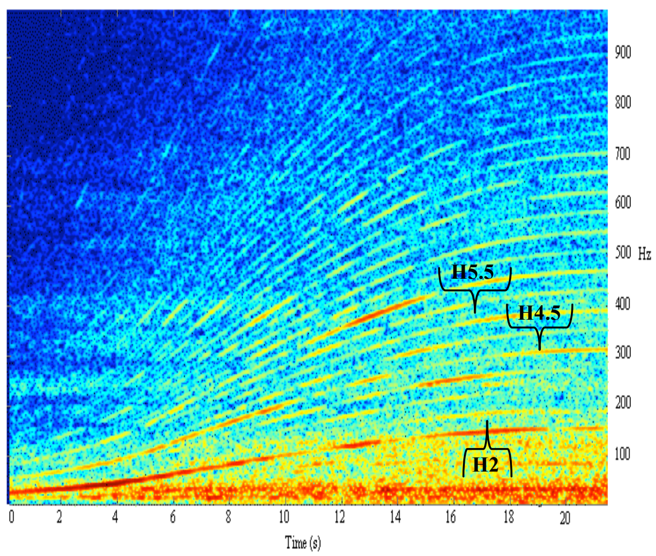

Figure 2*: Time-frequency representation of a signal recorded in an accelerating car*

3.2 Synthesis

In order to test some perceptive signal properties, a synthesis system call Hartis (Harmonic Real Time Synthesis) was developed at the PSA Research Center [8]. This tool is used to create signals for perceptive tests and to study the perceptual impact of specific signal parameters.
The first step of this sound design process is the separation of the original sound in a harmonic part and a noise part. For this purpose, we use the Additive software [9] to extract acoustic characteristics of the engine partials (frequency, amplitude and phase) according to the number of motor rotations per minute (RPM). Additive also extracts a residual noise which corresponds to the original sound minus the detected harmonic components. This residual noise contains aerodynamic noise and tire-road noise.

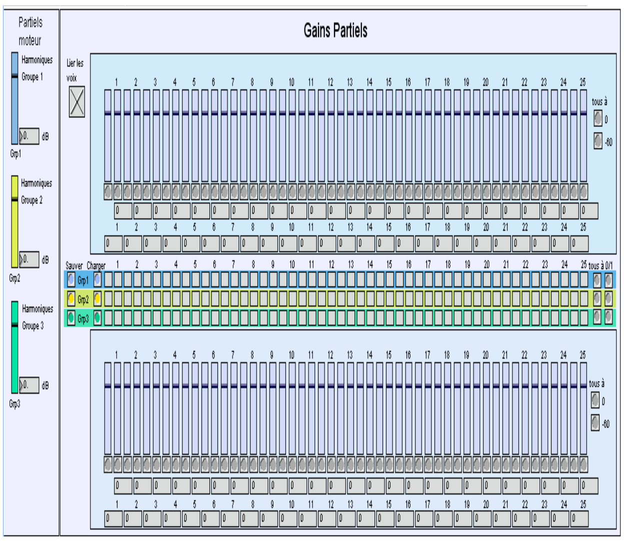

Figure 3*: Hartis Interface: Software for the real-time synthesis of interior car sound.*

The synthesis software Hartis was then used. This software can replay the original sound and allows real time control by acting on basic parameters of the signal like engine partials and noise level. To resynthetize the harmonic parts, an additive synthesis with the inverse FFT algorithm was used [10]. The noise part was generated by a smooth overlap granular synthesis [11].
This software allows the simulation of different harmonic profiles and balances between different sources. Four grains of 160 ms are played together to create the sound.

We have the possibility to modify the amplitude of 50 harmonics, but we have no control on the perceptive relevance of each parameter. We therefore need to reduce the number of parameters by using auditory models.

4. PERCEPTIVE ANALYSIS

4.1. Sensory analysis

We used a sensory panel as analytic noise evaluation tool. The panel was made of 9 naïve and normal hearing subjects, called judges, who were paid for their participation. After some training sessions, the panel was asked to define a minimal number of descriptors to qualify the sound they hear (2 x 2 hours). Then the panel proposed a scaling for each descriptor (4 x 2 hours). The panel gave an evaluation (2 x 2 hours) of 12 stimuli of 2 seconds recorded in standard and sporty cars with 100% accelerator load in 3rd gear. Sound starts with motor at 3500 rpm). Sounds were recorded in cars from different constructors. The total procedure was divided in 8 sessions of 2 hours. Through this method, we obtain a sensory profile for each sound. The success condition of a sensory panel is sound discriminability, the judges' repetitivity in time and their consensus. This is the reason why such a long procedure is necessary to transform naive subjects into efficient judges in order to obtain their consensus on all the descriptors.

The sensory panel found two main descriptors [8]. First, an "ON" which characterizes engine noise booming. It mainly depends on the audibility of odd harmonics (H2 –H4 – H6). The second descriptor "REU" is used to describe the roughness of the sound. This roughness is strongly related to the interaction between semi-harmonics (for example the interaction of high engine order harmonics H9-H12).

4.2. Application of auditory model

We applied an auditory model to understand how the human ear reacts to an interior car sound. In this manner, signal parameters that have a real impact on perception can be identified. This is also a global approach to the masking phenomenon which can give us information about the audibility threshold in the noise part and information about masking between engine harmonics.
The auditory model simulates different stages of our auditory system. The transfer of sounds through the outer and middle ear can be modeled using a single FIR filter [12]. It is also constructed to simulate the effects of head-phone presentation. The output of the filter can be considered as symbolizing the sound reaching the cochlea.
The cochlea can be described as a bank of bandpassfilters called auditory filters which center frequencies span from 50 to 15000 Hz and which bandwidth increases with increasing center frequency. We chose a 4th order linear gammatone filter, which

is well adapted to our automotive preoccupations, which parameters are the center frequency and the damping coefficients [13]. The derivation of those parameters leads to the design of a non-linearly spaced auditory filterbank. The frequencies of the auditory filterbank are linearly spaced on the so called ERB frequency scale. The step size determines the density of the filters on the ERB Scale. In this study we chose an important number of filters by ERB in order to improve the visualization at the output of our model. The center frequencies of each filter were distributed according to the harmonic frequencies. Hereby we had the possibility to follow the perceptive impact of each harmonic. Frequency centers are calculated with the knowledge of the RPM in time. We used a 20 ms interval between two updates of the center frequency filters. We draw the output of our auditory model by indexing the output of each gammatone filter by the engine number. That's why we have all the information of a given harmonic on the same line. This representation is helpful to describe the interaction between harmonics.

The next stage is the modeling of the inner cells: This stage was well described by Meddis [14]. The unidirectional excitation of inner hair cells is modeled by a half-wave rectification (HWR) followed by a low pass filter (LPF). We add compression with a power law in order to accentuate the variation [15].
The output of our model can be considered as the cochleagram defined by Slaney et al. [16] and can be considered as the cochlea movement.

4.3. Discussion

The representation on Figure 4 confirms the important role played by the harmonics around H2-H4 and H6 to the perception of the "ON" phenomenon. This description first shows that high order harmonics do not have a great impact on the perception of the global timbre of an interior car sound. They are perceptively mixed with aerodynamic noise. Secondly, we believe that the booming zone we can hear in the sound can be identified thanks to such a representation. This zone corresponds to the important shift in the cochleogram (For example between time t=16 s and t = 20 s). The representation, figure 4, facilitates the understanding of the transitions between these different events.

Moreover, we can identify perceptive formants on our representation. Actually, a combination of harmonics that seems to contribute to the perception can be observed. For example, we can see on Figure 4 that at the instant t=12 s, we have two formants: a first formant resulting from a combination between the harmonics H1.5, H2 and H2.5 and a second formant linked to the harmonics H5, H5.5, H6 and H6.5. The presence of this second formant clearly changes the sound. The next question we have to study is how these formants interact to quantify timbre variations. Formant trajectories can also inform us about timbre evolution and give a partition of timbre change. We can suppose that there is a relationship between these formants distribution and the descriptors used by the sensory panel .In the sound we present in the figure 4, we can hear a timbre change when we have two formants on our representation (for example at 12 seconds). The sound is clearly different when we only have one formant.

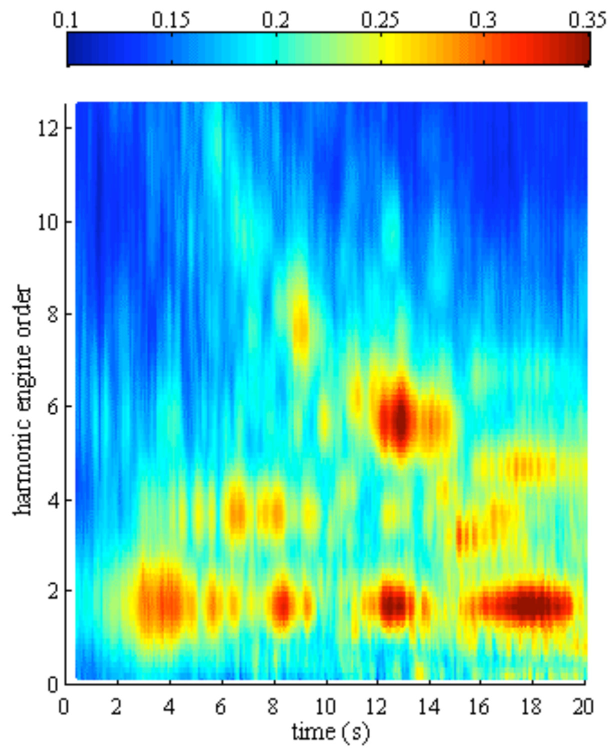

Figure 4: *Output of Auditory Model on the same accelerating car*

By considering the harmonic distribution through a few formants, we can reduce the information and hereby concentrate our investigation on fewer and perceptively relevant parameters to characterize perceptive aspect of the interior car sound. Further investigations are needed to interpret the cochleogram. In particular, masking phenomena that occur between the harmonic part and the noise part should be identified through the analysis of such representations.

5. CONCLUSION AND PERSPECIVES

Time-frequency representations are inadequate to describe the perceived interior car sounds because of the presence of multiple masking phenomena between the harmonic part and the noise part. Auditory models give better representations to catch the perceived impact of signal parameters. We show here our first auditory representation of interior car sound and we here described our first attempt to build a complete model of interior car sound timbre. The first step is the reduction of signal parameters by determining the audibility of engine harmonics in the interior car noise. That's why, in the future, we will model the engine harmonic audibility by comparing the auditory representation of the engine part on one hand and the auditory representation of the noise part on the other hand. Then, we will study the link between signal parameters and these perceptive formants.

In addition, sensory analysis is a good way to describe interior sounds. By combining the knowledge of the sensory profile based on typical descriptors of interior car sounds (ON and REU) with the description of signal impact on perception by perceived formants, timbre variation could be identified. These findings will make it possible to construct a synthesis model of interior car sounds that efficiently controls perceptive aspects of such sounds.

6. REFERENCES

[1] R. Bisping, "Car Interior Sound Quality: Experimental Analysis by Synthesis," *Acta Acustica united with Acustica,* Vol. 83, 1997, pp. 813-818, 1997.

[2] N. Chouard and T. Hempel, "A semantic differential design especially developed for the evaluation of interior car sounds," *J. Acoust. Soc. Am.* Volume 105, Issue 2, 1999.

[3] N. Kubo, V. Mellert, R. Weber and J. Meschke, "Categorization of engine sound," *Proceedings of the Internoise,* Prague 2004.

[4] H. Hansen, R. Weber and U. Letens, "Quantifying tonal phenomena in interior car sound," *Proceedings of the forum acusticum*, Budapest 2005.

[5] F. Richard, F. Costes, J-F. Sciabica, V. Roussarie, **"**Vehicle Acoustic specifications using masking models," *Proceedings of the Internoise*, Istanbul 2007

[6] J. H. Hawkins and S. S. Stevens ''The masking of pure tones and of speech by white noise,'' *J. Acoust. Soc. Am.* 22, 6–13. 1950

[7] M.C. Bezat, *Perception des bruits d'impact: Application au bruit de fermeture de porte automobile.* Thèse de Doctorat en acoustique, Université de Provence, 2007.

[8] V. Roussarie and F. Richard, "Sound design in car passenger compartment: Process and tool for the control of engine sound character," Journées du Design Sonore, 2004.

[9] X. Rodet, Musical Sound Signal Analysis/Synthesis: Sinusoidal Residual and Elementary Waveform Models, TFTS'97 *(IEEE Time-Frequency and Time-Scale Workshop 97),* Coventry, Grande Bretagne, 1997

[10] X. Rodet and P. Depalle, "Spectral envelopes and inverse FFT synthesis," *Proceedings of the 93rd Audio Engineering Society Convention,* 1992.

[11] L. Zadorin, *Granular synthesis: an introduction*. Queensland University of technology, 1997.

[12] B.R. Glasberg and B.C.J. Moore, "Development and Evaluation of a Model for Predicting the Audibility of Time-Varying Sounds in the Presence of Background Sounds," *J. Audio Eng. Soc*, vol. 53, pp. 906-918, 2005.

[13] V. Hohmann, "Frequency analysis and synthesis using a Gammatone filterbank". *Acta Acustica United with Acustica.* , vol. 88, pp. 433–442, 2002.

[14] M.J. Hewitt M.J and R. Meddis: Implementation details of a computation model of the inner hair-cell/auditory-nerve synapse, *J. Acoustic. Soc. Am.*, vol.87, no.4, p. 1813-1816, April 1990.

[15] D. Gnansia, "*Modèle auditif en temps reel*," Mémoire de stage, Rapport de stage du Master ATIAM, Université Pierre et Marie Curie, 2005.

[16]]M. Slaney, D. Naar and R.F. Lyon, "Auditory Model Inversion for Sound Separation," *Proc. IEEE-ICASSP*, Adelaide, II.77-80, 1994.