PERCEPTIVELY BASED DESIGN OF NEW CAR HORN SOUNDS

Guillaume Lemaitre, Patrick Susini, Suzanne Winsberg and Stephen McAdams*

Ircam-CNRS, Music Perception and Cognition Team
1, place Igor Stravinsky, F-75004 Paris, France.
*also Laboratoire d'Acoustique de l'Université du Maine
Avenue Olivier Messiaen, F-72085 Le Mans cedex 9, France.
lemaitre@ircam.fr

ABSTRACT

Due to the technologies used, there exist only a few different kinds of car horn sounds. We propose a method for the design of new sounds, based on a perceptual study of the actual sounds of car horns.
Firstly we deal with recordings of existing car horns. We show that the different kinds of horn sounds can be divided into nine main families. Within these families we demonstrate secondly that the perception of timbre results from the integration of three elementary sensations. Thirdly, another experiment reveals that some sounds are better identified as car horns than others. A relationship between the perceived timbre of the sounds and their ability to be identified as car horns is established. Finally, we generalize our results to synthesized sounds. The synthesis method was designed to explore and enlarge the perceptual space. Studying these sounds confirms and generalizes the previous results. A model is proposed, which is able to synthesize sounds and predict their ability to be identified as car horns.

1. INTRODUCTION

The design of sounds is an important issue nowadays for the automotive industry. A lot of attention is paid to all the sounds produced by cars. Car horns are no exception. Designing car horn sounds allows car builders to fit the sound of the horns to vehicule categories. However, and before this marketing aspect, the function of a car horn is to alert people to a potential danger related to a car. Designing the sound of the car horn thus involves a compromise between the need to customize the sounds and the necessity of providing efficient warning signals. This paper is a part of a broader project, the aim of which is to help car horn builders to invent new sounds, keeping the main function of a car horn: to alert people.
There exist nowadays only a few different kinds of car horn sounds, but they are well known to people and are identified without ambiguity. People do have a sort of "prototypical" representation of car horn sounds ([1]). The goal of the work presented here is to define how to design sounds which would be perceived as car horn sounds without any ambiguity. Our assumption is that if a new sound is immediately identified as a car horn, it will automatically alert people to the potential danger of a car.
This paper is divided into five parts.
In the first part we describe the usual sounds of car horns.
Secondly we define the main families of sounds perceived.
Thirdly we describe the perception of the timbre of these families, using perceptual spaces and psychophysical descriptors.
Fourthly we show that some sounds are better identified than others. Relationships between psychophysical descriptors and identification rates are then established.
The fifth part aims to generalize the first results. A method for synthesizing new sounds is described. The new sounds are created to explore the perceptual space and to enlarge it by creating sounds, the perceptual dimensions of which have values outside the range occupied by the usual sounds. Studying the identification of the sounds allows us to extend the relationships found in the first part. It is then possible to synthesize new sounds and to predict how well theses sounds would match the prototypical representation of car horn sounds. This will help car horn builders to design efficient new sounds: new sounds that are, however, identified as a car warning signal.

All the tests were performed in the sound booths at Ircam. The experiment was run on a Personal Computer under Linux, and the graphical interface was implemented under Matlab. The sounds were amplified through a Yamaha amplifier and sent to Sennheiser HD 520 II headphones.

2. THE CAR HORNS SOUNDS

Before going further, let us take a close look at car horns. There exist three kinds of devices. The first kind, that we will call the *horn-like* device is based on an electrodynamical driver loaded by a horn. The second kind is also made of an electrodynamical driver, but there is a metal plate attached to the membrane. We will call this one the *plate-like* device. The third kind of car horns are pneumatic driver loaded by a horn. We will call it the *pneumatic-driven* horn.
The horn sounds were recorded in the anechoic chamber at Ircam. They were all 550 ms in duration, and were equalized in loudness in a preliminary adjustment experiment. The corpus included sounds of horn-like, plate-like, and pneumatic-driven devices. They were recorded individually (monophonic case), and in some cases in twos or threes to obtain chords (multiphonic case). In the experiments that we describe, the sounds were calibrated in level compared to that of a 1-kHz pure tone at 83 dB SPL. A total of 43 sounds were recorded.

3. THE MAIN FAMILIES OF SOUNDS

The aim of this part was to determine families of sounds with similar timbre.

Class	Label	Number
Class 1	Special new monophonic horn-like	1
Class 2	Standard monophonic horns	11
Class 3	Ship horn	1
Class 4	Special very high-pitched horns, both mono- and multiphonic	6
Class 5	Multiphonic standard plates	5
Class 6	Multiphonic standard horns	6
Class 7	Special plates, mono- and multiphonic	6
Class 8	Standard monophonic plates	6
Class 9	Special very low horn	1

Table 1: The perceived families of car horn sounds

Experiment: 28 subjects (15 men and 13 women) were given written instructions explaining the task. Emphasis was placed on what timbre is (neither pitch, nor duration nor loudness). The subjects saw a white screen on which stars labeled from 1 to 43 were drawn, each star corresponding to the sound of a different car horn. They could hear the sound by double-clicking on a star. They were asked to move the stars in order to group together the sounds they heard as having the same timbre.
Analysis: We first derived a hierarchical tree representation of the data using an unweighted arithmetic average clustering (UPGMA) analysis procedure. Then in order to evaluate the most stable level of clustering across listeners responses we used a bootstrap algorithm [4]. This procedure was used to find an optimal partition of the sounds which corresponded to perceptual categories. As a result we found a stable tree with nine classes. Table 1 lists the type of horns corresponding to each class. Note that due to confidentiality restrictions, the references of the car horns are not detailed. According to this classification, sounds of car horns of the same technology are classed together (horns with horns, multiphonic with multiphonic, etc.: a typological classification).
Discussion: The relationship between this timbre-based classification and the typological classification of the car horns should be considered. Three large families appear. The first one (classes 1 to 4) groups together all the monophonic devices, including the pneumatic horn. The second one (classes 5 and 6) is composed of multiphonic sounds: plate-like or horn-like devices. The last one (Classes 7 to 9) groups together the monophonic plate-like devices. The distinction between monophonic and multiphonic sounds is quite clear. At this step, we can nonetheless conclude that there are two main families of timbre for the car horn sounds: horn-like and plate-like devices. Within these families, monophonic and multiphonic devices are grouped together. Somes specific sounds are set apart. Further discussions on this experiment can be found in [3]
We used this classification in order to choose a subset of sounds, both representative of the variety of the car horns and not too large in number. We sampled 22 sounds spread over the nine families

4. THE PERCEPTION OF TIMBRE

The strategy followed here is based on [2]. The main idea is first to represent the proximity between the sounds heard by the subjects using a spatial (Euclidian) representation. This spatial model is considered to represent the perceptual dimensions that underlie the dissimilarity ratings. The problem is thus to find the psychoacoustic correlates that match the perceptual dimensions. We used descriptors discovered by previous studies on timbre (see [5], [6], or [7] for example)
Experiment: 41 subjects (20 men and 21 women) were told that they were to make judgments on the timbre (not pitch, duration or loudness). All 241 different pairs among the 22 sounds were presented. On each trial the subject had to assess the dissimilarity between two sounds using a horizontal slider on the computer screen. The scale was labelled *Very Similar* at the left end and *Very Dissimilar* at the right end.
MDS analysis: A multidimensional scaling analysis was performed using CLASCAL [8]. We found a spatial model with 3 dimensions and 6 latent classes of subjects. The distances computed in the spatial model were correlated with the raw proximity data ($df = 239$, $r = 0.93^{**}$, p¡0.01). The latent classes are only interpreted as an image of the variety of response. It should be noted that the output of the CLASCAL algorithm doesn't need to be rotated, if more than one latent class is included in the model.
Pyschoacoustic correlates: Using the previous studies on timbre ([5], [6], [7]), we compute the psychoacoustic descriptors that those studies revealed to be relevant for timbre perception. These descriptors are based on physiological models of the auditory system and computed directly on the signal. Three descriptors were found to match the perceptual dimensions.
The first descriptor is *spectral spectroid* (df=20, r=0.96, p$<$0.01), as computed in [7]. It corresponds to the semantic attribute of "brightness".
The second descriptor is *roughness* (df=20, r=0.86, p$<$0.01) as described in [5] and [9].
The third descriptor is *spectral deviation* ([7] and [6]). It is related to fine structure of the spectral envelope (df=20, r=0.82, p$<$0.01) .
Discussion: We found psychoacoustic descriptors that explain a significant portion of the variance in the perception of the sounds of car horns.
The descriptor matching the second dimension (roughness) characterizes the classification between monophonic and multiphonic sounds. It must be noted that subjects performed a continuous rating from pure periodic sounds (one single harmonic series) to sounds composed of two periodic sounds. The spectral analysis of the intermediate sounds reveals that they are made of harmonic series based on the fundamental frequency, added to a second attenuated sub-harmonic series, which progressively increases the perceived roughness.
The descriptors matching the first dimension (brightness) and the third dimension (spectral deviation) distinguish between the sounds of horn-like and plate-like devices.
We are now able to describe the timbre of car horns by computing these three descriptors. Further discussion on these results can be found in [3].

5. IDENTIFICATION OF THE SOUNDS

This part sought to describe the ability of the sounds to be identified as car horn sounds.
Experiment: 20 subjects listened to each of the 22 sounds. They had to answer the question: "Do you recognize a car horn?". Their responses were counted, and for each sound, an identification rate was computed.
Analysis The identification rates range from 30% to 95%. We used an analysis method developed by Suzanne Winsberg (LCREG) to find a relationship between the descriptors of the sound and the

identification rate. It computes the best model (both explicative and made of a small number of parameters) of additive spline functions of the descriptors that account for the variance of the identification rate. We found a model made of a second order spline function (1 interior knot) of the spectral deviation and of a third order spline function of the fundamental frequency (no interior knot). The functions are represented in Figure 1. The prediction of this model was correlated with the identification rate measured (df=20, r=0.88, p<0.01).

Discussion: The range of the identification rates shows that some sounds are better identified than others, even though all the sounds were recordings of car horns. There are relationships between one descriptor of timbre (spectral deviation) and fundamental frequency[1] on the one hand, and idenfication rates on the other. These relationships define values for which the sounds are the best identified: $spectral\ deviation < 8.5$ and $F_0 > 450\ Hz$. However the power of these relationships is limited by the small number of sounds used for this test. The curves are not well sampled.

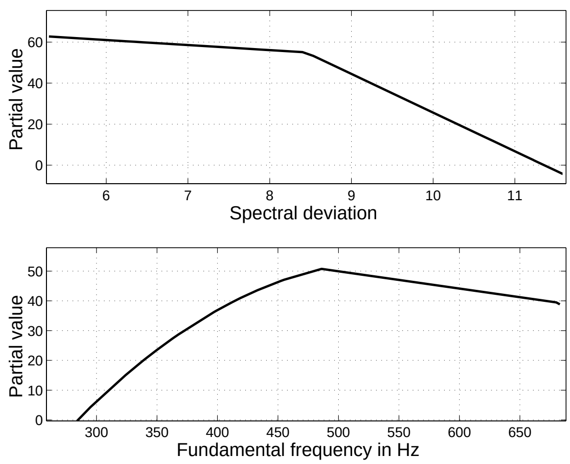

Figure 1: *Identification prediction model. Only real horn sounds are tested.*

6. PERCEPTION AND IDENTIFICATION OF SYNTHESIZED SOUNDS

6.1. The synthesis of the sounds

The results above have been found using only real car horn sounds. To extend these results, we decided to create new sounds. But, as suggested in [10], completely different sounds may have a different "semantic content", and may be set apart from the others. To counter this potential problem, we decided to create sounds that were perceptually close to the sounds of real car horns. The timbre of the new sounds had to share the same perceptual dimensions as the real sounds. But the values of the descriptors were chosen in such a way as to fill the gaps within the perceptual space, and to enlarge it with values outside the range occupied by the original sounds. Hence, we hoped to build a new perceptual space that was larger and denser than the former one.

The three descriptors defining the sounds of the car horns only deal with the spectral properties of the sounds (spectral centroid and spectral deviation) and the short-term temporal properties (roughness). To create new sounds that would be perceptually close to the real sounds, we decided to keep the temporal envelope and the non-harmonic noise of the sounds, and only tune the spectral and short-term temporal properties.

To synthesize a new sound, we used the temporal envelope of each partial and the non-harmonic noise of a real sound extracted using the ADDITIVE algorithm [11]. Then we defined an arbitrary spectral envelope (to explore spectral centroid and spectral devation), and we set the value of roughness by fine-tuning the frequencies of the partials. The resynthesis produces sounds which differ from the real sounds only by their spectral content and temporal fine-structure. We made 46 new sounds. They were all 550 ms in duration and were equalized in loudness in a preliminary experiment. The range of the values of the descriptors was larger than for the real sounds and the dimensions were better sampled. We hypothesized that the timbre of these sounds will be defined by the same descriptors as the real sounds.

6.2. The perception of the sounds

In order to check the validity of our synthesis method, we performed a dissimilarity judgment experiment. 19 sounds were chosen among the 46 synthesized sounds and mixed with 4 sounds that had already been tested in the previous experiments. The procedure was the same as in paragraph 4, but with only 20 subjects.

Analysis: We performed an MDS analysis using CLASCAL [8]. We found a spatial model with 2 dimensions correlated with the raw dissimilarity data (df=249, r=0.93, p<0.01). The first dimension was correlated with spectral centroid (df=21, r=0.79, p<0.01), and the second dimension was correlated with roughness (df=21, r=0.77, p<0.01). The dissimilarities for the sounds tested in the previous experiment were consistent with the previous data.

Discussion: We found that the timbre of this corpus of sounds was defined by the same first two descriptors as the real sounds. The correlation coefficients are weaker than for the experiment described in section 4, and the third dimension (spectral centroid) does not appear here. However, since we had only 20 subjects, we conclude that the perception of the timbre of the new sounds is close to the timbre of the real sounds.

6.3. The identification of the sounds

19 listeners took part in an identification task based on a corpus made of the 19 new sounds and the 22 real sounds tested beforehand. The procedure was the same as in section 2.4.

Analysis: We used the LCREG analysis method to find a relationship between the descriptors of the sound and the identification rate. We found a model composed of a second-order spline function (1 interior knot) of spectral deviation, and of a third-order spline function of fundamental frequency (no interior knot). The functions are represented in Figure 2. The prediction of this model was correlated with the identification rate measured (df=39, r=0.69, p<0.01).

Discussion: In the experiment in which we tested synthesized sounds, we again found that spectral deviation and fundamental frequency are able to predict the results of the identification task. The range of these two descriptors was larger here, and the functions are better defined. We conclude that the best identified sounds have a spectral deviation between 6 and 9, and a fundamental frequency between 300 Hz and 500 Hz.

[1] Fundamental frequency does not match the definition of timbre and is related to pitch

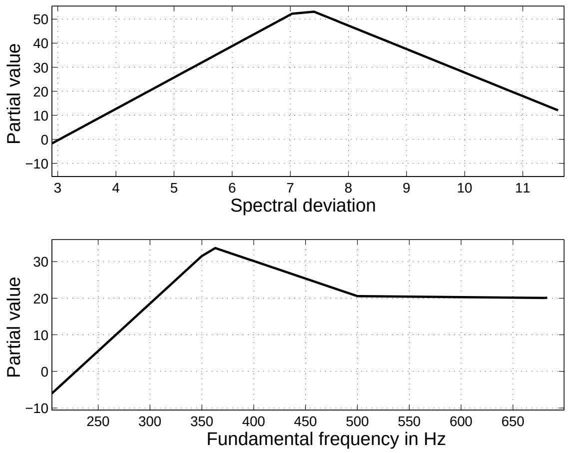

Figure 2: *Identification prediction model for both real and synthesized sounds.*

7. CONCLUSION

This paper briefly describes results concerning timbre and pitch perception for car horns and proposes a method to design new sounds that will be identified as car horn sounds. We assume that if a sound is identified as a car horn, it will warn listeners of the potential danger of a car.

We have shown that the sounds of the car horns can be divided into nine families. Within these families, the timbre of the sounds results from the integration of three elementary sensations. These sensations are correlated with three psychophysical descriptors. Two of them are classical descriptors (spectral centroid and roughness). The third descriptor (spectral deviation) is less well-known. It is related to the spectral fine-structure of sounds. Then we asked listeners to identify the sounds. Even if all the sounds were recordings of real sounds, we found that some sounds were not identified as coming from cars by most of the listeners. We showed that there exists a relationship between one of the timbral descriptors (spectral deviation) and fundamental frequency on the one hand, and identification rate on the other.

To generalize the results, we created new sounds, using an original method. Using an additive analysis-synthesis algorithm, we designed sounds which differ from real sounds only by their spectral content and their temporal fine-structure. These new sounds have the same global temporal envelope as real sounds. They were designed to be close enough perceptually to the real sounds to be compared to them, but to extend the perceptual space. The values of the descriptors were spread over a much larger range than for the recordings. Studying the identification of these sounds allows us to confirm and generalize our results.

Emphasis should now be placed on the definition of identification rate. All the experiments indeed took place in a laboratory context, and we only tested sounds that were not too different from one another. The results should not be interpreted as an absolute definition of what the sound of a car horn is among other sounds. Our results should rather be interpreted within the context of a homogeneous corpus of harmonic sounds with a defined temporal envelope.

In spite of this restriction, we can predict, for a given synthesized sound, how it will be perceived and how it will be identified. This prediction is based on a descriptor of timbre and on the fundamental frequency. The fundamental frequency, because the sensation of pitch is perceptually distinct from that of timbre [7], did not appear in the timbre study. But this is an important factor for the identification of car horn sounds. Listeners better identify sounds that have the same fundamental frequency as normally encountered car horns (400 Hz).

Spectral deviation is a descriptor that has only been studied a little, but it is shown here that it is the most important factor for the identification of car horn sounds. People do identify car horns using their fine-grained spectral properties.

8. ACKNOWLEDGEMENTS

This work forms part of the doctoral research of the first author and is funded by SCE Klaxon, part of the FIAMM group, in collaboration with Ircam-CNRS and the University of Le Mans.

9. REFERENCES

[1] J. A. Ballas, “Common Factors in the Identification of an Assortment of Brief Everyday Sounds” in *Journal of Experimental Psychology: Human Perception and Performance*, vol.19(2), pp. 250-267, 1993,

[2] P. Susini, S. McAdams and S. Winsberg, “A Multidimensional Technique for Sound Quality Assessment”, *Acta Acustica united with Acustica*, vol. 85, pp. 650-656, 1999,

[3] G. Lemaitre, P. Susini, S. Winsberg and S. McAdams, “Perception of the timbre of car horns”, *Proc. of the Forum Acusticum Sevilla, Spain 2002* (S. Hirzel Verlag [CD-ROM], Sevilla, Spain),

[4] O. Houix, “Catgorisation auditive des sources sonores”, *PhD Thesis*, Université du Maine, Le Mans, France, 2003. To be published,

[5] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, Springer-Verlag, Heidelberg, Germany, 1990.

[6] J. Krimphoff, “Analyse acoustique et perception du timbre”, *Master’s Thesis*, Université du Maine, Le Mans, France, 1994,

[7] J. Marozeau, A. de Cheveigné, S. McAdams and S. Winsberg, “The dependency of timbre on fundamental frequency”, *Submitted*,

[8] S. Winsberg and G. de Soete, “A latent class approach to fitting the weighted euclidian model, CLASCAL”, *Psychometrika*, vol. 58(2), pp. 315-330, 1993,

[9] D. Pressnitzer, “Perception de rugosité psychoacoustique: d’un attribut élémentaire de l’audition l’écoute musicale”, *PhD Thesis*, Université Paris 6, France, 1998,

[10] B. Gygi, G. R. Kidd and C. S. Watson, “Environmental Sounds: Detection, Identification and Similarity”,*Proc. of the Neural Information Processing System Conference (NIPS99)* , Workshop on Neural Mechanisms of Music Processing,

[11] X. Rodet, “The additive analysis-synthesis package”, *http://www.ircam.fr/equipes/analyse-synthese/DOCUMENTATIONS/additive*, Ircam