



HAL
open science

Spatial Sound Design in a Car Cockpit: Challenges and Perspectives

Théophile Dupré, Sébastien Denjean, Mitsuko Aramaki, Richard Kronland-Martinet

► **To cite this version:**

Théophile Dupré, Sébastien Denjean, Mitsuko Aramaki, Richard Kronland-Martinet. Spatial Sound Design in a Car Cockpit: Challenges and Perspectives. 2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA), Sep 2021, Bologna, France. pp.1-5, 10.1109/I3DA48870.2021.9610910 . hal-03456711

HAL Id: hal-03456711

<https://hal.science/hal-03456711v1>

Submitted on 30 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spatial Sound Design in a Car Cockpit : Challenges and Perspectives

Théophile Dupré^{*†}, Sébastien Denjean[†], Mitsuko Aramaki^{*} and Richard Kronland-Martinet^{*}

^{*}Aix-Marseille Univ, CNRS, PRISM, Marseille, France

Email: {dupre, aramaki, kronland}@prism.cnrs.fr

[†]Scientific department, Stellantis, Vélizy, France

Email: sebastien.denjean@stellantis.com

Abstract—With the development of electric motor vehicles and autonomous driving systems, the domain of automotive sound design addresses new issues, and is now concerned by creating suitable and pleasant soundscapes inside the vehicle to ensure comfort and safety for the users. In particular, the integration of informative, alerting and intuitive auditory feedback becomes essential. For instance, previous studies proposed relevant sonification strategies to convey the information about the vehicle dynamics in electric cars. However, while the sound feedback is informative per se, users reported its lack of blending with the surrounding natural sound field of the vehicle and the artificial character of the sound itself. The challenge is therefore to integrate such sound feedbacks in the car interior soundscape in natural way by adjusting both timbre and spatial parameters. In this research, we propose an experimental setup to investigate perceptual integration of synthesized and environmental sound streams spatialized in car cockpit. By combining synchronized 3D acoustic field measurements and 3D video recordings at seat viewpoint, we constructed immersive audiovisual environments in the laboratory to get close to ecologically valid driving situations while avoiding time consuming in-situ experiments. The proposed setup also allowed to virtually control sound sources (both their timbre and location) in the car interior and to study the spatial perception of the overall designed soundscape before vehicle integration. While we here address the case of car cockpit, this approach may be used for different contexts in which sounds need to be spatially integrated in natural soundscapes.

Index Terms—Sound Design, 3D Sound, Virtual Reality, Immersion

I. INTRODUCTION

Car industry is facing a major shift from Internal Combustion Engine (ICE) vehicles to Electric Vehicles (EV). Major developments are now focused on EV. Apart from different dynamic behavior of electric motor, user experience is mainly affected by a different acoustic environment or soundscape.

The term soundscape has been introduced by the canadian composer R. M. Schafer in [21]. He stated that the auditory environment must be considered as a composition of contextually related elements rather than an addition of isolated and independent acoustic events. This concept addresses the limitations of traditional noise management based on energetic indicators which have limited perceptual impact on user experience. In this paper, we intend to adopt soundscape paradigm to the changing auditory experience of an interior car.

Traditional ICE vehicle acoustic environment is composed of three main acoustic sources. At low and medium speed, harmonic engine sound with frequencies proportional to engine speed and low frequency tire-road contacts are predominant. At higher speed, wide-band aerodynamic noise tends to mask engine sound [18]. Engine sound is very useful to the driver because it conveys information about the dynamic of the vehicle [7] and vehicle quality [22], [19]. Tire-road contact noises and aerodynamic noise compose the noisy part in the sense of unwanted noise.

On the other hand, silence and quiet environment describes our expectation of EV acoustic environment. While the overall energy of the sound field is lower due to the absence of predominant engine sound, EV are composed of noisy elements only. Engine brings less acoustic feedback on the vehicle dynamic and add high frequency content [9]. This contribution is unexpected and reported as annoying by users. Other noises that were previously masked are audible in EV and new noise sources appears (i.e. the battery cooling system).

The soundscape of an interior car is drastically changing. This shift is an opportunity to design a controlled acoustic environment and integrate virtual sources to bring auditory feedbacks to the driver and passengers while masking unwanted noises. For example, strategies have been investigated to add auditory feedback about the vehicle dynamic in the cockpit through sonification of speed and acceleration [6]. Perceptual evaluation showed improvement of speed dynamic perception in the presence of sonification, increasing safety and driving comfort. Nonetheless, some users reported a lack of coherence and belonging within the existing soundscape and perceived the proposed sounds as artificial. The next challenge is to naturally integrate virtual sound sources such as electric motor sonification by adjusting timbral and spatial parameters.

We previously investigated the impact of timbral features, vehicle structure resonances and random frequency and amplitude modulations observed on ICE noise, on naturalness of electric motor sonification sounds. Results showed no impact of modulations and a positive impact of structure resonances but with high variance between subjects. Results have to be balanced with a lack of contextualization reported by the subjects during the experiment. The previous study seems to indicate that the problem of integration is more linked

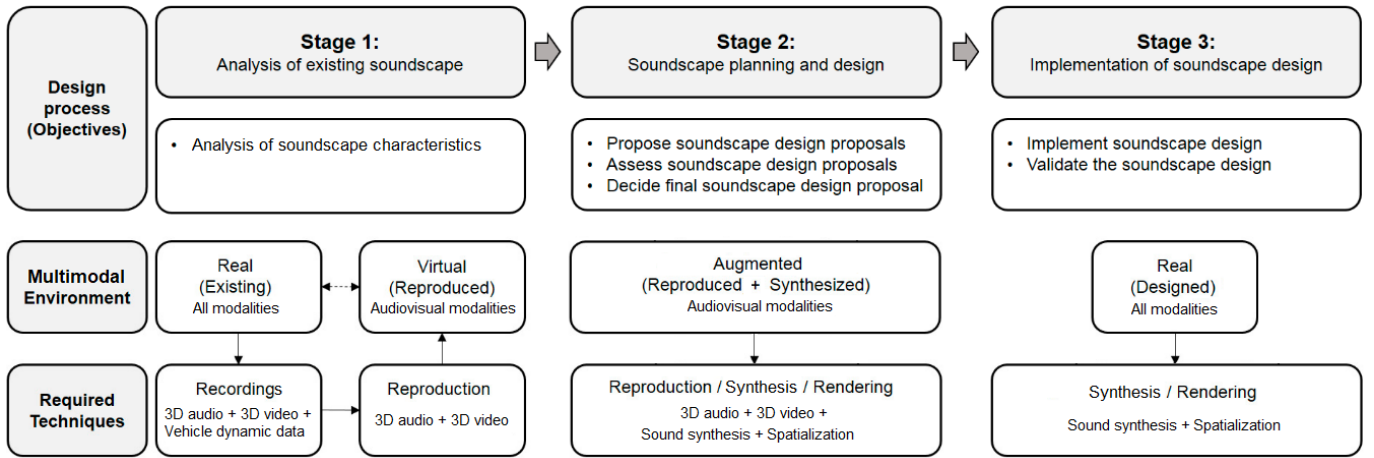


Fig. 1. Proposed approach for soundscape design development in car cockpit adapted from [12]

to the environment (i.e. resonances) and how the sound is integrated that the sound itself. Therefore, we focus here on the investigation of spatial features of the environment and the virtual sound sources while paying particular attention to contextualization.

Context plays a critical role in the perception of soundscape. Especially in a car cockpit, our expectation of specific multimodal stimulation is closely linked to the context. Our perception of an audio recorded driving situation listened through headphones is completely different from the actual auditory perception we experience while driving. The study of car soundscape in a laboratory requires an immersive environment to be as close as possible to real multimodal experiences of car users. This paper proposes a methodology developed for urban soundscape planning [12] to study spatial auditory perception of an interior car and to design and validate solutions to integrate virtual sound sources ecologically in an EV cockpit.

First, the proposed methodology is presented. Then, the experimental setup to record and render cockpit environment is explained and soundscape analysis procedure is detailed. Finally, future works and perspectives are developed.

II. PROPOSED METHODOLOGY

A. Methodology in urban soundscape planning

Urban soundscape planning faces the same contextualization issue in laboratory experiments. Laboratory-based experiments minimize the influence of uncontrolled factors encountered in-situ and allow fast solution testing. The challenge is to construct immersive environments in laboratory to maximize the generalization of results to in-situ environments while keeping the ease of use of controlled environments. A three step methodology has been introduced in [12].

B. Adapted methodology and hypothesis

In this paper, we propose to apply this methodology to test and evaluate spatial integration solutions of virtual sound sources in the cockpit. A car cockpit is a very specific acoustic environment, our perception of this soundscape is highly

influenced by contextual aspects. High ecological validity of laboratory experiments is crucial to generalize results in real driving contexts. An illustration of the adapted methodology is shown on Figure 1. In step 1, the cockpit soundscape is recorded and reproduced in an audiovisual virtual environment. As we are mostly concerned by spatial auditory sensation, recording and restitution of the soundfield must be carefully chosen. Based on the measurements, analysis of the soundscape is done to identify wanted and unwanted sounds. This analysis can be based on acoustical and perceptual considerations (c.f. IV). In step 2, sound design proposals based on the analysis are tested and evaluated in the virtual environment. Preferred solutions are then implemented in-situ in step 3 and evaluated in the targeted soundscape.

Here, we focus on a specific virtual sound source: sonification sound of electric motor. The sonification sound is based on the Shepard-Risset illusion [6]. A modular and spatial synthesizer has been developed and allows controlling timbral and spatial characteristics of the sound. We assume that if the spatial impression of this virtual source corresponds to the spatial impression of an engine noise in ICE vehicles, the sound will match the expected presence of vehicle dynamic feedback and blend naturally in the soundscape.

III. COCKPIT ENVIRONMENT RENDERING

In order to evaluate the integration of the virtual engine sound source, the cockpit environment has been recorded in-situ and then rendered through virtual environment in the laboratory.

A. Recording and restitution of 3D acoustic field

Traditional monophonic recordings are unable to capture faithfully 3D auditory environments apart from spatial impression due to reverberation. Multichannel solutions have been developed over the years that can be categorized in three approaches: direct, non-parametric and parametric. Direct methods are the most commonly used by sound engineers



Fig. 2. Experimental setup to record 3D video and 3D audio from passenger seat viewpoint in driving situations

and consist in manually arranging microphones and loudspeakers based on heuristics. Those methods will not be further detailed here because microphone and loudspeaker layouts are interdependent. Please refer to [5] for a design methodology of such systems.

Non-parametric and parametric methods allows flexibility between microphone array and loudspeaker layout. An intermediate representation of the sound field is constructed, mixing and unmixing matrix are computed depending on the microphone array and the loudspeaker layout respectively. Non-parametric methods usually use linear transformation such as ambisonic. This recording format also known as B-format was developed by Gerzon in the 70's for four channel microphones [10]. Input channels are encoded in the spherical harmonic domain and can further be decoded on any loudspeaker configuration. Later, Higher Order Ambisonic (HOA) extended the original 1st order ambisonic format to more than 4 channels. The relation between the order N and the number of channels M is $N = (M + 1)^2$, the more the better in terms of spatial resolution [4]. Ambisonic framework supposed that the recorded sound field is composed of plane waves. In a car cockpit, sources are located in the near field. This condition is not met especially at low frequency. However, according to [3], spherical waves are correctly reconstructed at the center of the reconstruction area. This area increased with the ambisonic order and at low frequency. Car passengers are not moving. Therefore, the auditory environment is correctly rendered for one person by ambisonic recordings and decodings.

Parametric methods process microphone array signals in the time-frequency domain and estimate spatial parameters based on sound field models. As an example, Directional Audio Coding (DirAC) [16] algorithm estimates in each frequency band the direction of the energy and the amount of diffuseness by computing the correlation between B-format channels. Omnidirectional channel of the B-format recordings is used to



Fig. 3. Multichannel rendering system

reconstruct the sound field by weighting between directional and diffuse field in each reproduced region. An extended version of DirAC have been developed for HOA recordings [15]. Parametric techniques are mainly used in spatial audio coding because they allow to downmix the recordings and achieve good perceptual reconstruction.

Perceptual recording and reconstruction known as binaural techniques are also possible with stereophonic setup. It consists in recording the sound field at the ear entrance with intra-auricular microphone or dummy head. The recorded signals contain the binaural cues that we use to perceive the spatial auditory environment. Direction-dependent filtering called Head Related Transfer Function (HRTF) processes the signals for headphone playback. Interindividual differences influence the quality of the reconstruction, personalized HRTFs are required to avoid externalization problem (feeling that the sound is coming from inside the head) and front/back confusions [23]. Also, headphones move with listener's movement. Head-tracking system is required to decode the audio streams accordingly and keep the auditory image independent of head rotation. This process adds latency which decreases the sensation of presence and immersion.

For the technical difficulties mentioned above and the immersion quality of loudspeaker-based rendering, loudspeaker restitution have been preferred to binaural rendering in the virtual environment.

B. In-situ recordings

In-situ recordings involve measurements of the acoustic field, the visual field and dynamic parameters of the vehicle. The acoustic field is recorded using a spherical array composed of 32 microphones distributed on a sphere (*Eigenmike32*). This type of microphone allows encoding the acoustic field in 4th order ambisonic to ensure a good reproduction in all directions. The visual field is recorded by a 3D camera (*Insta360 Pro*) composed of 6 optic lenses on the azimuthal plane, each having a 6k resolution. Reconstruction processing of the visual field is done by Insta360 professional stitching software allowing 6k 3D stereoscopic videos (*Insta360 Stitcher*). Stereoscopic rendering simulates parallax effect which is the

main perceptual cue of vision depth. Vehicle speed, engine speed for ICE vehicle measurements and throttle opening are recorded simultaneously. The camera is placed on the front passenger seat at head height with the microphone just over it as in Figure 2. Time and orientation alignments of the recordings are done by hand claps. An exponential sine sweep test signal is played through car loudspeakers to equalize the virtual environment in loudness and frequency.

C. Virtual environment

We constructed a virtual environment to render driving situations previously recorded in-situ. A Virtual Reality (VR) headset is used to render the visual field (*Oculus Quest*). Ambisonic channels are decoded on 42 loudspeakers (*Genelec 8020C*) distributed on a sphere and 2 subwoofers. The setup is placed in an anechoic chamber to avoid any influence of the room on the reconstruction of the sound field. An illustration of the simulation environment is shown on Figure 3. A Unity application has been developed to play back 3D video recordings in the VR headset. Virtual sound engine synthesizer is based on a modular synthesizer developed by Stellantis [8]. A Max/MSP patch controls the recorded scene and the synthesizer. Spatialization of virtual sources is controlled in real-time by IRCAM SPAT library [2]. Video and audio synchronization is done via the OSC protocol.

Perceptual distortion of the environment is induced by the reproduction system especially localization blur and distance distortion known in ambisonic rendering [11]. We don't intend to characterize quantitatively these distortions but rather concentrate on the immersion quality to ensure that the subjects are properly immersed and contextualized in a moving vehicle. In order to validate the latter, questionnaires have been developed to measure presence (i.e. the sense of being present in a environment) [1]. A pilot study will be conducted prior to any investigation of sound integration to measure the sense of presence in a car cockpit created by this virtual environment.

IV. COCKPIT SOUNDSCAPE ANALYSIS

In this study, we are interested by spatial features of the sound field. Ultimately, we need a perceptual understanding of the cockpit sound field but acoustic analysis might be of interest. Non-parametric acoustic analysis techniques are usually based on beamforming and can be adapted to spherical harmonic signals. Beam patterns of selected directivity estimate the energy arriving from a specific direction. By probing in every directions, a powermap can be constructed to show the energy distribution in space. Several algorithms based on different principles have been developed, please refer to [17] for a review. Such approaches have been used for automotive applications, especially acoustic leak detection and noise management in general [14]. Parametric models such as DirAC have been designed in a perceptually motivated way. Model parameters can give relevant perceptual information on the sound field.

Characterization of interior car sound field can also be done through perceptual tests. Rapid sensory analysis protocol has

been developed to assess the quality of automotive audio [13]. The subjects are asked to freely describe different acoustic environments during a first phase to construct a common subjective description framework. In a second session, subjects evaluate the same acoustic environments on the constructed scales. The benefit of this protocol is to be subjective and avoid biases induced by instructions or descriptors imposed by experimenters.

Analysis of the measured soundscapes are based on the methods mentioned above. We recorded controlled driving situations on closed roads on ICE vehicles and EV such as acceleration, deceleration and constant speed. We are analysing spatial energy distribution of engine sound in ICE vehicles with beamforming techniques. This approach is well suited for ICE noise because the engine noise is predominant at low speed. For EV soundscapes, no sound predominates, perceptual approach has to be used. Therefore, we are designing a protocol based on rapid sensory analysis to better understand EV soundscape perception and expectation and compare with ICE vehicle soundscapes. Open road measurements have been recorded for this analysis to be closer to real driving situations. This procedure corresponds to the first stage in Figure 1.

V. FUTURE WORKS AND PERSPECTIVES

In this paper, we proposed a methodological framework to study spatial integration of virtual sound sources in a car cockpit. The interest of this methodology is to create a simulation environment with good ecological validity and contextualization where testing, tuning and validation is faster than in situ. First, we are testing the methodology on sonification sound of the electric motor. We collected in-situ data and constructed an audio-visual immersive environment. Data analysis is still in progress.

Driven by the analysis of ICE sound and the expectation of EV soundscape, spatial integration solutions will be proposed and evaluated in the immersive environment. Specific perceptual spatial features of the source such as location, apparent width and envelopment will be studied [20]. Besides the validation of our assumption, other integration solutions with moving sources for example might be investigated. This part constitutes the second stage in Figure 1. The final stage is to implement the validated solutions in a car cockpit and evaluate in real driving contexts (c.f. third stage in Figure 1).

If the results are promising, it will be extended to the increasing number of informative and alerting sound sources in vehicles in order to have a global control of the soundscape. In terms of technical improvement of the virtual environment, the ecological validity could be improved by adding vibration feedback in the seat. Furthermore, this methodology could be employed in any context in which virtual sound sources needs to be spatially integrated in soundscapes.

REFERENCES

- [1] Sarvesh Agrawal, Adèle Simon, Søren Bech, Klaus Bæntsen, and Søren Forchhammer. Defining immersion: Literature review and implications for research on audiovisual experiences. *Journal of the Audio Engineering Society*, 68(6):404–417, 2020.
- [2] Thibaut Carpentier. A new implementation of spat in max. In *15th Sound and Music Computing Conference (SMC2018)*, pages 184–191, 2018.
- [3] Jérôme Daniel. Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia. 2000.
- [4] Jérôme Daniel and Sébastien Moreau. Further study of sound field coding with higher order ambisonics. In *Audio Engineering Society Convention 116*. Audio Engineering Society, 2004.
- [5] Enzo De Sena, Hüseyin Hacıhabiboğlu, and Zoran Cvetković. Analysis and design of multichannel systems for perceptual sound field reconstruction. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(8):1653–1665, 2013.
- [6] Sebastien Denjean, Richard Kronland-Martinet, Vincent Roussarie, and Sølvi Ystad. Zero-emission vehicles sonification strategy based on shepard-risset glissando. In *14th International Symposium on Computer Music Multidisciplinary Research*, 2019.
- [7] Sebastien Denjean, Vincent Roussarie, Richard Kronland-Martinet, Solvi Ystad, and Jean-Luc Velay. How does interior car noise alter driver’s perception of motion? multisensory integration in speed perception. In *Acoustics 2012*, 2012.
- [8] Gaylord Desoeuvre, Florent Richard, Vincent Roussarie, and Marie Céline Bezat. Hartis, a re-synthesis tool for vehicles sound design. *The Journal of the Acoustical Society of America*, 123(5):3247, 2008.
- [9] Maud Geoffriault, Emmanuel Godoy, Dominique Beauvois, and Gwennaél Favennec. Active reduction of vibrations in synchronous motors. In *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society*, pages 3293–3298. IEEE, 2013.
- [10] Michael A Gerzon. Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, 33(11):859–871, 1985.
- [11] Catherine Guastavino, Véronique Larcher, Guillaume Catusseau, and Patrick Boussard. Spatial audio quality evaluation: comparing transaural, ambisonics and stereo. In *International Conference on Auditory Display*, 2007.
- [12] Joo Young Hong, Jianjun He, Bhan Lam, Rishabh Gupta, and Woon-Seng Gan. Spatial audio for soundscape design: Recording and reproduction. *Applied sciences*, 7(6):627, 2017.
- [13] Neofytos Kaplanis, Søren Bech, Sakari Tervo, Jukka Pätynen, Tapio Lokki, Toon Waterschoot, and Søren Holdt Jensen. A rapid sensory analysis method for perceptual assessment of automotive audio. *Journal of the Audio Engineering Society*, 65(1/2):130–146, 2017.
- [14] Daniel Pinardi, Lorenzo Ebrri, Costante Belicchi, Angelo Farina, and Marco Binelli. Direction specific analysis of psychoacoustics parameters inside car cockpit: A novel tool for nvh and sound quality, 2020.
- [15] Archontis Politis, Juha Vilkkamo, and Ville Pulkki. Sector-based parametric sound field reproduction in the spherical harmonic domain. *IEEE Journal of Selected Topics in Signal Processing*, 9(5):852–866, 2015.
- [16] Ville Pulkki. Spatial sound reproduction with directional audio coding. *Journal of the Audio Engineering Society*, 55(6):503–516, 2007.
- [17] Boaz Rafaely, Yotam Peled, Morag Agmon, Dima Khaykin, and Etan Fisher. Spherical microphone array beamforming. *Speech Processing in Modern Communication*, pages 281–305, 2010.
- [18] Florent Richard, Francois Costes, Jean-Francois Sciabica, and Vincent Roussarie. Vehicle acoustic specifications using masking models. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 2007, pages 3153–3162. Institute of Noise Control Engineering, 2007.
- [19] V Roussarie, F Richard, and MC Bezat. Perceptive qualification of engine sound character; validation of auditory attributes using analysis-synthesis method. *Proceedings of the CFA/DAGA*, 2004.
- [20] Francis Rumsey. *Spatial audio*. Taylor & Francis, 2001.
- [21] R Murray Schafer. *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster, 1993.
- [22] Jean-François Sciabica, Marie-Céline Bezat, Vincent Roussarie, Richard Kronland-Martinet, and Solvi Ystad. Towards the timbre modeling of interior car sound. In *15th International Conference on Auditory Display*, 2009.
- [23] Elizabeth M Wenzel, Marianne Arruda, Doris J Kistler, and Frederic L Wightman. Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, 94(1):111–123, 1993.