Urban environment audio simulation for contextual evaluation of Quiet Vehicles' sound design

Nicolas MISDARIIS[1]; Julien GERBER; Julien ALEONARD

STMS Ircam-CNRS-UPMC, France

ABSTRACT

The silence of Quiet Vehicles – especially, Hybrid or Electric Vehicles – is a crucial issue in terms of safety, ergonomics or even ecology which becomes to be largely studied either in a research or development point of view. Most of the works done in that domain are based on experiments which try to measure presence and emergence of such vehicles in real, or simulated, environment and to evaluate their detectabilty, in the same contexts, while being equipped with extra sounds specifically designed for fulfilling basic functional conditions with regards to nearby pedestrians or other road users. In this necessary experimental workflow, we assume that simulation approach – or, more generally, virtual reality – is a potentially powerful and valid paradigm to test the contextual impact of a large number of propositions and to conduct controled laboratory experiments for evaluation. This paper will then present the first step of the development of an audio simulation based on Ambisonics field recordings in various urban locations, multi-channel audio rendering and mock-up of different audio scenarios for insertion of one or multiple QV sound signatures in a given sound scene. Next steps – together with general stakes and perspectives – of this approach will also be presented and discussed.

Keywords: Electric Vehicle, Sound design, Simulation
I-INCE Classification of Subjects Number(s): 13.2; 52.3; 63.7; 68.2.

1. INTRODUCTION

The issue addressed by Quiet Vehicle relative quietness becomes to be progressively well known. Moreover, it seems to be integrated in broader and regular scientific studies or actions like, for instance, the recent European project eVADER [1]. In fact, after having dealt with this fundamental question: is silence of electric vehicles a blessing or a curse ? [2, 3] and its corollary: are vehicles driven in electric mode so quiet that they need acoustic warning ? [4], most of the people agree with the QRTV United Nations working group conclusions: vehicles propelled in whole or in part by electric means present a danger to pedestrians [5] and the consequence that Quiet Vehicles need an artifical extra sound – or sound signature – mainly for two reasons: signalling its presence for nearby users (pedestrians, cyclists, etc.) and giving a sonic feedback for the driver [6]. At this step, the chalenge is to define which sound will be the best candidate to play this role, or more precisely, what should be the nature of the sound (spectro-temporal characteristics, interactive properties, etc.) to fulfill these functional requirements and, in the same time, answer to aesthetic criteria of each entity of the automotive industry.

To this end, and quite conventionally, most of the scientific studies use an experimental approach to test their hypothesis. Considering uppermost the external question (sound for signalling presence outside of the vehicle), paradigms used in the different studies are often very similar: a stimuli implemented in a usage scenario (traditionnal pass-by or more complex ones, like lateral straight or turn paths [7]) and a listening test involving participants asked to achieve a detection task (raise an hand or push a button as soon as the coming vehicle is heard). This paradigm is developed in two different experimental conditions: i/ in situ when people stand at the side of the road where the vehicles are passing and realise the task in a real environment [8] – providing an optimal realism but dealing with some difficulties to control the experimental factors; ii/ in laboratory when people listen to an electroacoustic rendering of the sound scene (background + artificial sounds) and realise the

1 nicolas.misdariis@ircam.fr

task in a sort of virtual environment – allowing an optimal control of the experimental factors but needing some strong approximations in terms of realism. In that latter case, the technical set-up often stays quite simple and is rather based on stereo devices [9, 10, 11], even if multichannel solutions are sometimes proposed [7]. Nevertheless, it is worth noticing that, in most cases, few details on the rendering protocol is given in the respective studies (recording, decoding, calibration, etc. …).

Between these two extremes experimental conditions, a third approach consisting in developing an efficient simulation tool – based on consistent virtual reality paradigms – appears to be the right compromise. At a first level of complexity, this tool could only use the audio modality, but it could also moves towards an audio/visual combination – as the ESS by B&K [12] – in order to enhance the two main criteria targeted by this kind of application: realism and immersion.

In this paper, we will present the preliminary work done in the way to develop a simulation tool for Quiet Vehicles in urban soundscape. The paper is an experiment report type and, then, is structured following the workflow: first, definition of a urban soundscape typology in order to state on relevant recording locations with regards to a representativeness criterion; then, investigations on QV's sound signature to be inserted in the simulated scene and, finally, development of the simulation tool itself by mixing these two elements together. A synthetic discussion on several aspects of the current tool is proposed afterwards, in order to draw perspectives given by this work.

2. A SIMULATION TOOL FOR QUIET VEHICLES IN URBAN SOUNDSCAPES

2.1 The *Ground*: Urban Soundscape

The purpose of this first part was to adress the issue of urban soundscape typology in a pragmatic perspective: to define several recording locations that should be representative of a city – the elicited city being Paris, for sake of convenience. For that, we extracted from the literature several representativeness criteria of different natures (structural, acoustical, perceptual) and resulted in defining an *ad hoc* 3-group classification associated with parisian locations where we have been able to record some specimens of similar sound scenes.

2.1.1 Elements of Typology

On a **structural/functional** point of view, urban soundscapes – or more generally, landscapes – can be described either in terms of their components (number of lanes, pavement typology, etc.) – which merely defines town-planning configurations – or in terms of their contents (objective type of existing activities or sources, number/speed/type of vehicles, etc.) – which defines effective usage configurations. In that focus, and based on previously collected theoretical factors influencing urban soundscapes (running conditions, number of lanes, road width, etc.), Polack et al. have determined a morpho-typology of urban spaces, i.e. a typology based only on morphological – or structural – variables. This typology contains 4 main classes: A/ double-way, 3 or more lanes, very heavy traffic, open fabric; B/ one-way, 3 or 4 lanes, very heavy traffic, U-shaped streets, width below 40 m. ; C/ double-way, 2 lanes, medium traffic; D/ one-way, 1 or 2 lanes, weak traffic, U-shaped streets, width below 20m. [13].

On an **acoustical** point of view, urban soundscapes can be defined by their acoustic properties: sound pressure level (SPL) – or equivalent level L_{eq}, as defined by Kang [14] and used, for instance, in the European Directive [15] –, reverberation conditions, mainly due to architectural components (buildings, open space, etc.) or pavements.

On a **perceptual** point of view, urban soundscapes can be assessed as pleasant/unpleasant, annoying, ideal, comfortable/uncomfortable, variable, quietness/lively, etc. All these attributes have something to see with the 3 main components of a soundscape, as theorized by Schafer: "tonality", related to background noise having the amorphous property [16], "signals", related to source events having a property of emergence and "sound markers" related to a specific sound with regards to a given location or population [17]. In this focus, Kang et al. have experimentally determined key factors characterizing soundscapes in urban open public spaces: i/ relaxation, including comfort, quietness or enjoyment notions; ii/ communication, including social, meanings or notion of calm; iii/ spatiality, including echo or distance notions; iv/ dynamics, including amplitude levels notions [18]. Another study by Guastavino has determined other representative criteria for sound quality assessment of soundscapes: i/ variety, related to the type and structure of neighborhood activities; ii/ tranquility, related to how quiet and relaxing is a given environment; iii/ animation, referring to human activities; iv/ non-aggressiveness, i.e. warmhearted and peacefull [19].

Relying on all of these description systems aforementionned, we finally proposed our own typology based on 4 attributes: **dynamics**, **dimensions**, **sound level** of the location and **nature of the source** which are composing it; the latter attribute being able to be discriminated, at a first level, between articificial or natural types.

2.1.2 Gross Categories and Recording Sites

On the basis of these last elements of typology, 3 gross categories have been arbitrarily defined, namely: calm, noisy but bearable, unbearable; and described upon the four elicited criteria as depicted in Table 1 below:

Table 1 – description of the 3-class typology of urban soundscapes

	Calm	**Noisy but bearable**	**Unbearable**
DYNAMICS	low	high	low
DIMENSIONS	narrow lanes	varying	wide lanes
SOUND LEVEL	< 50 dB(A)	~70 dB(A)	> 80 dB(A)
NATURE OF SOURCES	mostly natural, minimum of saliency	Mostly artificial, with few human contributions	Mostly artificial, of all types

Starting from that description, we were able to select representative urban sites. We also made the selection by having in mind the type of passing scenarios we would like to test, so as to be able to embed a QV in the simulated scene as coherently as possible with regards to the existing vehicle flux. We finally elicited three different sites for each category (operating at one, two or three different local positions, according to each site) as follows:

- **Calm**: rue Maspero (XVIth district), rue Norvins street (XVIIIth), rue Clerc (VIIth)

Figure 1 – rue Maspero (XVIth district) topographical configuration

- **Noisy but bearable**: rue de Rivoli (IVth district), rue Beaubourg (IIIth), rue Caulaincourt (XVIIIth)

Figure 2 – rue Beaubourg (IIIth district) topographical configuration

- **Unbearable**: boulevard Ney (XIXth district), boulevard Barbes (Xth), Place de l'Etoile (VIIIth)

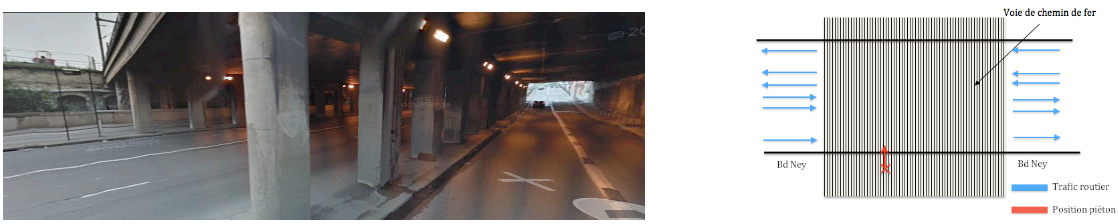

Figure 3 – boulevard Ney (XIXth district) topographical configuration

The technical apparatus was identical for each *in-situ* session and included:

- a Soundfield (ST250) microphone – attached to a Zoom H6 recorder – allowing a 4-channel Ambisonics recording, directly formatted in B-format (W, X, Y, Z);
- an ORTF stereo couple (Schoeps MSTC 64 U) – attached to a Tascam HDP2 recorder – allowing a standard stereophonic recording, useful for having ear references either in terms of spectral or spatial responses;
- a PHONIC PAA3 sonometer, allowing the monitoring of an average sound level in dB(A) at each recording spot, useful for an *a posteriori* calibration of the rendering system.

2.2 The *Figure*: Quiet Vehicles

The purpose of this second part was to process the question of QV's sound signature in order to define which type of **sound signature** will have to be put in the first developed version of the application. For that, after having summarized the general concept on sonification of QV, we chose one successfully completed solution from the whole set of guidelines, recommendations and implementations that existed at that time. Again, for sake of simplicity and coherence, this solution came from a work previously done in collaboration with the French car manufacturer Renault and dealing with one commercialized model of their electric brand (Zoe).

2.2.1 Basic concepts

On a conceptual point of view, the Quiet Vehicle can be described as a silent object evolving in an heterogeneous and noisy environment and thus can be considered as a sort of Unidentified Moving Object with respect to people in its direct contact or vicinity. Then, this new type of vehicle obviously adresses new issues in term of usage or user experience, and especially:

- coexistence with actual noisy combustion engine vehicles that is likely to generate a potential dangerous situations for persons evolving nearby (pedestrians, cyclists);
- ergonomics of driving as the lack od audio feedback in the inner of the car significantly modifies driver's perception with regards to the vehicle's movement;
- global (urban) soundscape modification involved by the introduction of this new object in its environment, rather similar to the arrival of a new species in a given ecosystem.

Considering the sound design itself, an overview of the literature led to define three main formal criteria that seems to be of a major importance for fulfilling the basic requirements in terms of functionality, aesthetic and ecology: i/ sound level which is the first cue for detectability and strongly depends on the type of backgrond noise [10]; ii/ nature of the signal that can maximize – or minimize – perceptual effects like masking or can lead to consider causal sounds linked to a certain semantic value; iii/ interactivity that can deal with the necessity for QV to be inserted in a continuously time-evolving environment, for instance by adjusting its sound signature with the surrounding sound level, the day period or some specific weather conditions.

2.2.2 One particular solution from the whole

Since the end of the 90's and the publication of the first patents detailing "apparatus for generating noise sound which is equipped in industrial electric cars with little or no driving sounds" [20, 21], a lot of actions have been undertaken at different levels.

At a national level, several countries have worked on standards and recommendations to address the question of safety for Quiet Vehicles: the Japan and the work of the Japan Automobile Standards Internationalization Center (JASIC), the United States with the publication of the Pedestrian Safety Enhancement Act of 2010 edited after works done by the National Higway Traffic Safety Administration (NHTSA) and mainly supported by American blind associations or the European Commission which have recently made temporary demands – on the basis of conclusions of the World Forum for Harmonization of Vehicule Regulations (WP29) of the United Nations Economic Commission for Europe (UNECE) – that should be included in the European rules from 2019.

At an international level, the informal working group named Quiet Road Transport Vehicle (QRTV) works since 1999 inside the WP29 for elaborating a Global Technical Regulation (GTR) on the topic of Quiet Vehicles sound signature. The aim of this work is to determine – under scientific studies and objective data – if QVs need an extra sound for signalling itself, and in that case, what kind of properties should have this sound. The main work-in-progress criteria concern sound level, spectral content, temporal modulation and speed-driven evolution [22].

At an industrial level, some automotive manufacturers have transposed the first specifications for Audible Vehicle Alert Systems (AVAS) into their own workflows and products. For instance, Nissan has developped the concept of Vehicule Sound for Pedestrian (VSP) for its Leaf model. Acoustically, the VSP is based on two spectral peaks – approx. 600 and 2500 Hz – and one dip – around 1 kHz –; moreover, the first peak is slightly modulated and this modulation is driven by the car's speed. Other indutrial entities (Ford, GM, Hyundai, etc.) has also developed their own proprietary solutions – see the wikipedia "Electric vehicle warning sounds" webpage for more details on that point.

For the present study, we chose to work with one particular solution yet developed within the team in collaboration with the composer Andrea Cera, taking advantage of a long term project with the french car manufacturer Renault [23]. On a technical point of view, the sound signature is synthesized by a wavetable technique using 4 buffers, each of them storing one component of the sonification signal. Each component plays a precise role by being attached to a certain part of the spectral bandwidth (100-400 Hz / 1-2 kHZ, 5 kHz), a temporal morphology (stationary layer, modulation, impulsion) and a perceptual or ergonomic function (for instance, localisation, notification of presence, etc.). This sound synthesis engine is driven by output data from the vehicle (mainly, speed from CAN bus) that are transformed into gain and pitch variables in the synthetizer. The interactivity is assumed by the fact that these variations occur either in local (in each buffer) and global step, ensuring that the audio content is always changing during the movement of the car. The prototype version of this application is developed in a Max/MSP patch giving access to the four buffers and the control parameters (see Figure 4). Three different solutions corresponding to three different moods are finlally implemented on a microchip embedded in the car: "pure", "sport" and "glam" (see https://soundcloud.com/renaultze/ for audio demonstrations).

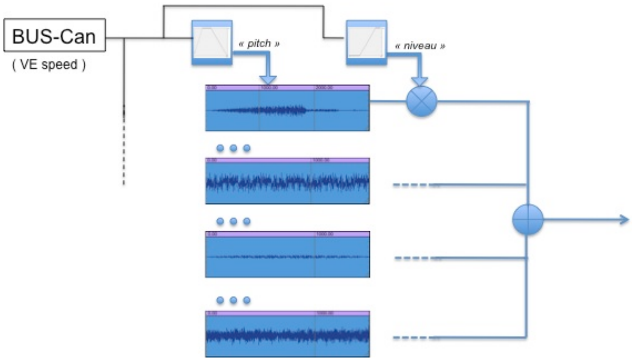

Figure 4 – Sound synthesis engine architecture of the Zoe's sound signature

Moreover, the output signal is tuned in terms of sound pressure level according to specifications measured as required by current standards. In other words, we know exactly the SPL value that the vehicle emits at different speeds; these data will be used for calibrating the simulation tool.

2.3 The Simulation Tool

The purpose of this third part is to develop a technological framework able to simulate the presence of a QV in a urban environment in order to **design and evaluate** sonic solutions dedicated to this issue. This development involved several operational steps including post-production of the Ambisonics field recordings, authoring presence and evolution of Quiet Vehicle signature and mixing theses two elements in a coherent virtual sound scene.

2.3.1 Decoding Ambisonics

The concern of this first step was to convert the recorded raw data – Ambisonics encoded in B-format – into an optimized multi-channel audio reproduction format in order to propose a convincing and ecologically valid auditory scene for the simulation tool.

We mainly based our reflexion on the work of Guastavino et al. who carried out comparative evaluation of different rendering configurations (1D, 2D or 3D) together with verbalizations on the criteria used to assess the auditory scene reproduction [24, 25]. As in our case, their experiments were using Ambisonics recordings of urban (parisian) soundscapes and, then, the main findings on which we relied were:

- first, that stereophonic reproduction is ecologically valid for source identification but not for processing complex auditory scene; and furthermore, 2D configurations provide better performances in terms of presence, spatial definition and image distance than 3D [25].
- second, that the relevant perceptual criteria used by people to judge the validty of the scene were mainly: readability, presence, distance, localization, coloration and stability of the

image [24]. These terms have been forming a common baseline for our subjective listening during all the development and especially for the next step, when there has been a need to elicit the better Ambisonics decoding paradigm.

This precise work – and all the next audio development works – was achieved in a post-production studio at Ircam (Studio 8) equipped with 5.1 system based on 5 loudspeakers and 1 subwoofer and placed in a standard configuration: C at 0°, L and R at +30° and -30° respectively, Ls and Rs at 110° and -110° respectively. Because of the limitation of total available channels and the complex routing we needed (different audio interfaces, audio softwares and main mixer of the studio), we chose to exclude the subwoofer. We thus chose to limit our work to a 5.0 setup in terms of mixing and recording. In our conditions, this was acceptable because of the really good frequency response of the satellite speakers we used (+/-1 dB between 40Hz to 20kHz).

Moreover, we calibrated the monitoring system in order to allow the recordings to be reproduced with the exact sound pressure measured in each outside location. This enabled us to avoid audition artifacts due to listening level when listening to the recorded ambiances, comparing decoders and building audio scenarios.

Then, from this technical set-up, we conducted a comparative informal listening between a series of Ambisonics codecs (available in the form of plug-ins), in order to choose the better one with regards to the perceptual criteria mentioned above. These different algorithms were:

- SurroundZone 2, a proprietary algorithm supplied with the Soundfield hardware;
- SpatDecoder, an algorithm developed by Ircam/EAC team and used in the Ircam Spat© [26];
- Harpex-B, a commercial plug-in newly released and based on quite original model for decoding [27].

Our (informal) listening protocol was based on two tangible elements: i/ a test file involving a recording scan at regular angle (22.5°) of the Ambisonics microphone with calibration signals (pink noise and 1000 Hz): fixed loudspeaker, controllably rotating microphone (turntable) in nearfield conditions of a reasonably dry recording studio; ii/ an outside singular recording containing a clear passing-by (right-left) of an ambulance in a lateral street with medium traffic (rue Beaubourg) and few structural artefacts (opening perpendicular streets), so that this recording could constitute a very good calibration file either for immersion and source localization.

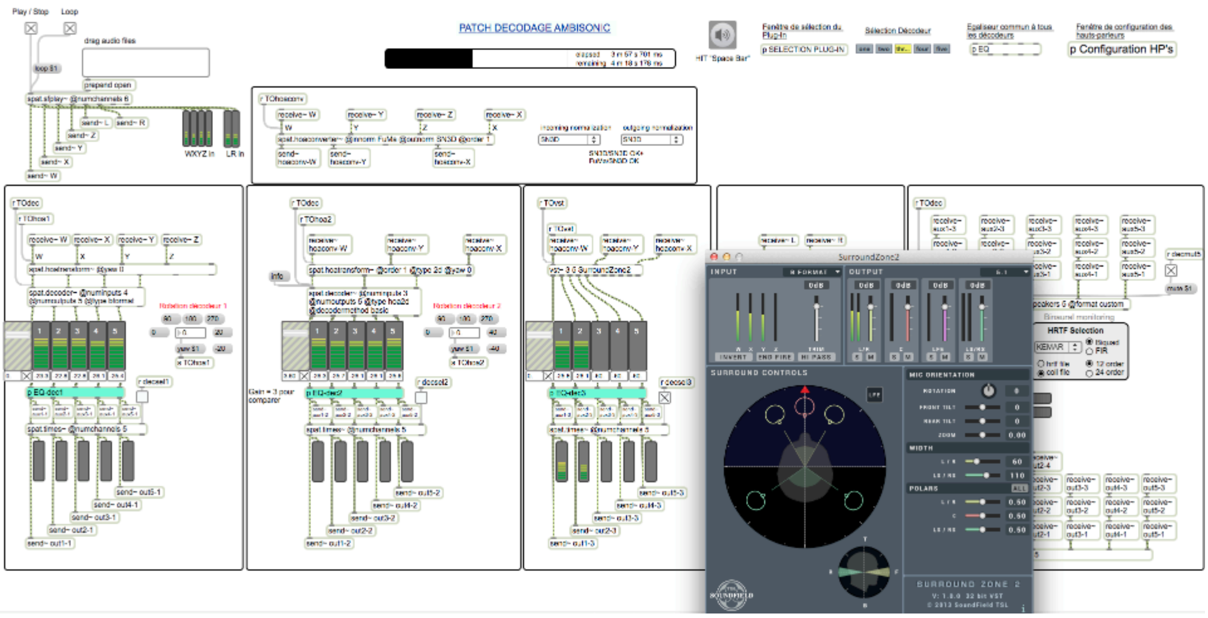

Figure 5 – Overview of the control patch for Ambisonics codec selection (Max/MSP)

After having implemented a Max/MSP patch specially dedicated to this choice task (containing several functionalities like source selection, channel switch, master level tuning, etc (see Figure 5 for overview), we reached the conclusion that, in our specific case, the Harpex-B codec was the best as it provided the best compromise between source events localization and background noise rendering in terms of coherence and immersion.

We then bounced all the Ambisonics recorded materials thru Harpex to make "ready-to-play" 5.1

files. For some pragmatic reasons, this work has been done by making communicate the Max/MSP patch with the Digital Performer sequencer by means of the OSC (OpenSoundControl[2]) communication protocol.

2.3.2 Embedding additional sources

This being, a second step of development has consisted in embedding a module dedicated to the additional source used to sonify the Quiet Vehicle. As mentioned above (sect. 2.2.2), the retained solution was the prototype developed for Zoe (Renault), available in the form of a Max/MSP implementation, in 3 different versions; these 3 instances of the patch have then become the second main audio source to be inserted in the whole environment of the simulation tool (Figure 6).

For that, a key element was needed to be fixed: definition and implementation of the evolution scenario of the source, either in spatial terms, considering the topography of a given site (distribution/orientation of lanes, etc.) but also temporally, i.e. according to other occuring events, in order to be fully coherent with the recorded scene. The software environment used for this task was the Ircam Spat© and more precisely, the SpatOper module that allows manipulation/visualization of sources into a virtual auditory scene reproduced by the Spat© audio rendering engine (Figure 6).

Note that, in terms of audio spatialisation, the Spat was only used in terms of localisation properties: because of the use case configurations (outdoor, nearly free field, etc.) – and to a first approximation –, the reverberation functionalities were completely disabled.

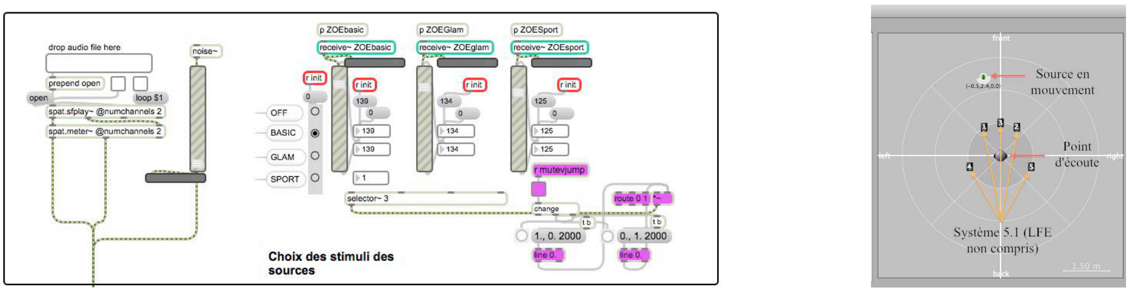

Figure 6 – control panel of the source in the simulation tool

sound signature selection (left); SpatOper interface (right)

The trajectories were defined by means of two modes: i/ manually, by playing directly with the source icon inside the SpatOper interface (Figure 6-right); ii/ automatically, by fixing the departure and arrival point coordinates, plus the traveling time (the trajectory being able to be split into several segments in order to define a non-linear displacement – as in a BPF paradigm); from these data, an average speed value was computed and added to the dataset attached to each scenario.

Again for pragmatic reasons, the Max/MSP module was coupled with the Digital Performer sequencer (DP) by means of the same OSC protocol (see sect. above); this architecture especially allowed us to record the scenario trajectories as MIDI automation variables in DP and then to get a symbolic representation of the trajectories that can be edited, fine-tuned directly in the sequencer and finally independently applied to several kind of sources.

Then, from these symbolic data, we were able to dynamically generate one of the three sound signatures of Zoe, according to a pre-defined passing scenario compatible with the background behind which the source aims at being embedded.

Furthermore, since we knew all the properties of the Zoe sounds in terms of level, we calibrated the Spat in order to reach the correct sound level when the vehicle was positioned 1m forward the listening position. From this point, the Spat© audio engine rendered all the modifications due to distance (attenuation) and speed (Doppler effect), with respect to that initial calibration.

2.3.3 Building a virtual scene

The last step of the simulation tool development simply consisted in adding up the two previous modules in order to generate the mix of a virtual auditory scene containing a real background urban noise and a coherently added source event that corresponds to the passing of a Quiet Vehicle.

On a technical point of view, the architecture is as follows: the 5.1-encoded background urban soundscape is played by Digital Performer together with the trajectory automation variables (coordinates+speed) that are sent to a Max/MSP module attached to the dynamic event source – Zoe

[2] http://opensoundcontrol.org

signature – that delivers a modulated signal, encoded in the same 5.1 format and sent back to DP for recording. By mixing a traditional sequencer (Digital Peformer) with an advanced real-time environment (Max/MSP), this architecture allows to take advantage of the interactive and time-evolving nature of the additional sources. All theses tracks are finally bounced into a 5.1 mix containing the two sources: background noise and sound signature.

3. DISCUSSION AND PERSPECTIVES

At the present step of development of the simulation tool, this section will try to sum up benefits and drawbacks of the options that have been selected to reach this point and will try to draw the short and mid term perspectives of such a tool within the field of virtual reality for sound design and perceptual evaluation. Different points will be tackled: the audio technics for recording/encoding/decoding, the effective power of spatial audition – considering also potential interactions with other sensory modalities (sight) –, and the role of virtual reality in sound design applications.

3.1 Audio Technics

On the recording side, the choice of Ambisonics technics to capture urban soundscapes has proven to be relevant for many reasons:

- comparable to a lot of other scientific studies that have used this same approach;
- acceptable in terms of audio performances (frequency response, dynamics) that led to reasonably colored recordings without any significant audio distorsion;
- portable and useful solution, an important point when dealing with field recordings that needs, for instance, autonomous power supply or acceptable weight to move easily from one place to another.

On the rendering side, the choice of the standardized 5.1 loudspeaker arrangement has matched a compromise between advanced spatial audio technology and available devices for working (Ircam / studio 8). Nevertheless, even if decoding Ambisonics materials into 5.1 format appears to be rather standard, a large variety of coding/decoding algorithms appeared to exist and surprisingly lead to quite different perceptual results in terms ecological validity, according to the goal that is targeted (essentially, immersion vs. localisation). Then, even if the Harpex-B plugin seemed to fulfill our application requirements, this part certainly needs to be deeply investigated and formalized.

Besides, as the B format is, by nature, output-independent, it could be easily conceived to consider other rendering configurations. From this point of view, the binaural technique would be worth considering either for recording or reproduction. In fact, at the very end of the study, we have made some preliminary trials to decode the raw materials taken with the Soundfield microphone into a binaural representation but they were not fully successful. In spite of this, binaural could offer the benefit to be more portable, even if some issues like individualized HRTFs still exist.

3.2 Spatial Audition

Despite all the technical questions addressed in the previous section, another important one is to evaluate the efficiency of the simulation tool in terms of representation of the auditory scene reality. In other words, the question should be, at least: does the apparatus is enough convincing that a sonified QV is passing in a urban context ? (the question of “feeling exactly as being in the recording place” being utopic as soon as you listen to an outdoor content inside a room !).

Objectively, at the current precision level of the application, it is important to say that, whereas the background rendering is quite adequate (*immersion* criterion), the recorded-source rendering (*localisation* criterion) may be still better improved. In fact, in the only auditory modality, the pre-defined trajectories are mostly very hard to understand and to be mentally represented, except for the simplest ones like lateral passing-by.

One way of improvement could be either to further investigate the question of localization in the Ambisonics approach or to work on a better physical/acoustical description of the source itself inside the Spat©. But, another direction of work could also be to consider a multimodal representation in order to mix, for instance, auditory and visual sensory modality and take advantage of the constructive interactions that can result from this association. Without initially considering complex auditory stimuli (images or video), a first step of perceptual improvement could be just to give a graphical representation of the mobile source inside a simplified representation of the context (lanes, junctions, etc.). For instance, the source may be symbolized by a colored spot evolving inside schematic elements of urban context, in synchronicity with the audio trajectories defined in the simulation tool.

3.3 Virtual Reality

The aim of the present work was not exactly to validate the solutions that have been implemented in Zoe, but rather to develop an open framework able to help sound designer to achieve their design and evaluate their ideas. At this point, the simulation tool addresses the issue concerning the use of virtual reality in a sound design process and, in the specific case of the QV's sound signature, opens questions like: will the virtual reality be enough realistic to be efficient ? To what extent will it help to conceive sounds / validate intermediate solutions / communicate on sounds ?

These questions may be investigated by means of experimental protocols – like, for instance, hands-on session – involving sound designers in order to observe their respective practice and get their direct reactions about the application.

4. CONCLUSION

The present work has proposed a first step of investigation on the role of virtual reality for designing and evaluating sounds in an interactive and applied context. The application framework was actually sound signature of Quiet Vehicle.

All the necessary modules of an operational tool for rendering the simulation of a passing-by quiet vehicle into a coherent urban soundscape have been realised and implemented: recording and rendering the soundscape on a audio 3D format (Ambisonics for recording, 5.1 for rendering), defining a scenario for the virtual source (QV sound signature) and mixing it in a coherent manner with the pre-defined background scene.

The first release of the simulation tool is yet rather operational and can be used to evaluate and improve the sound design process in that particular use case.

ACKNOWLEDGEMENTS

Authors would like to thank the Ircam Production staff – especially Cyril Beros, head of the Department – for their concern in the present matter, and jury members at ENS Louis Lumière – namely, Jean-Pierre Halbwachs, Claude Gazeau, Laurent Millot and Gerard Pele – for their judgements and helpful comments.

REFERENCES

1. eVADER – Electric Vehicle Alert for Detection and Emergency Response. http://www.evader-project.eu
2. Otto N., Simpson R., Wiederhold J. Electric vehicle sound quality, Journ. of SAE 1999-01- 1694, 1999.
3. Cocron P., Bühler F., Franke T., Neumann I., Krems J. F. The silence of electric vehicles – blessing or curse, in Proceedings of the 90th Annual Meeting of the Transportation Research Board, Washington, DC. January, 2011.
4. Sandberg U., Goubert L., Mioduszewski P. Are vehicles driven in electric mode so quiet that they need acoustic warning signals, Proceed. of 20th Internat. Congress on Acoustics, Sydney, Australia, 2010.
5. QRTV-1. Terms of Reference and Rules of Procedure for the GRB informal group on Quiet Road Transport Vehicles (QRTV), Document QRTV-01-02-e, ECE/WP29/GRB, Geneva, Switzerland, 2010.
6. Misdariis N., Cera A. Sound signature of Quiet Vehicles: state of the art and experience feedbacks. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 247, No. 5, pp. 3333-3342). Institute of Noise Control Engineering, 2013.
7. Ashmead D. H., Grantham D. W., Maloff E. S., Hornsby B., Nakamura T., Davis T. J., Pampel F., Rushing E. G. Auditory perception of motor vehicle travel paths, Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(3), 437-453 (2012).
8. Emerson R. W., Naghshineh K., Hapeman J., Wiener W. A pilot study of pedestrians with visual impairments detecting traffic gaps and surges containing hybrid vehicles, Transportation research part F: traffic psychology and behaviour, 14(2), 117-127 (2011).
9. Nyeste P., Wogalter M. On adding sound to quiet vehicles, Proceedings of the Human Factors and Ergonomics Society, 52nd annual meeting (2008).

10. Menzel D., Yamauchi K., Völk F., Fastl H. Psychoacoustic experiments on feasible sound levels of possible warning signals for quiet vehicles, Proceed. of DAGA'11 conf., Düsseldorf, Germany 2011.
11. Misdariis N., Gruson A., Susini P. Detectability study of warning signals in urban background noises: a first step for designing the sound of electric vehicles, Proc. of Intern. Congress on Acoustics (ICA), Montreal, Canada. Juin 2013.
12. Gillibrand A., Suffield I., Vinamata X., Williams R., Brückmann A. An Initial Study to Develop Appropriate Warning Sound for a Luxury Vehicle Using an Exterior Sound Simulator (No. 2011-01-1727). SAE Technical Paper, 2011.
13. Polack J. D., Beaumont J., Arras C., Zekri M., Robin B. Perceptive relevance of soundscape descriptors: a morpho-typological approach. Journal of the Acoustical Society of America, 123(5), 3810, 2008.
14. Kang J. Urban sound environment. CRC Press, 2006.
15. European Directive 2002/49/CE on evalutation and management of noise in the environement http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000337482&dateTexte=20130123
16. Maffiolo V., Castellengo M., Dubois D. Qualitative judgments of urban soundscapes. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings (Vol. 1999, No. 2, pp. 1251-1254). Institute of Noise Control Engineering, 1999.
17. Schafer R. M. The tuning of the world, 1977
18. Kang J., Zhang M. Semantic differential analysis of the soundscape in urban open public spaces. Building and environment, 45(1), 150-157, 2010.
19. Guastavino C. The ideal urban soundscape: Investigating the sound quality of French cities. Acta Acustica United with Acustica, 92(6), 945-951, 2006.
20. Cha Y. W., Ahan H. S. U.S. Patent No. 5,517,173. Washington, DC: U.S. Patent and Trademark Office. 1996.
21. Koike M., Kitagawa M., Ishiguro K. U.S. Patent No. 5,635,903. Washington, DC: U.S. Patent and Trademark Office. 1997.
22. Pardo L. F., Misdariis N. Warning system for electric vehicles – Security and sound design. Techniques de l'ingenieur, 2014. (under submission)
23. Misdariis N., Cera A., Levallois E., Locqueteau C. Do electric cars have to make noise ? An emblematic opportunity for designing sounds and soundscapes. Congrès Français d'Acoustique, Nantes, 2012.
24. Guastavino C., Katz B. F. Perceptual evaluation of multi-dimensional spatial audio reproduction. The Journal of the Acoustical Society of America, 116(2), 1105-1115, 2004.
25. Guastavino C., Katz B. F., Polack J. D., Levitin D. J., Dubois D. Ecological validity of soundscape reproduction. Acta Acustica united with Acustica, 91(2), 333-341, 2005.
26. Vaananen R., Warusfel O., Emerit M. Encoding and Rendering of Perceptual Sound Scenes in the Carrouso Project. In Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio. Audio Engineering Society, 2002.
27. Berge S., Barrett N. High angular resolution planewave expansion. In Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics May (pp. 6-7), 2010.