PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

Traffic Flow Auralisation based on Single Vehicle Pass-by Noise Synthesis

Yang FU[1], Damian MURPHY[1], Alex SOUTHERN[2]

[1]AudioLab, Department of Electronic Engineering, University of York, UK, yf852@york.ac.uk
[2]Acoust. Dept., AECOM Ltd., 7th Fl., Aurora, 120 Bothwell St., Glasgow G2 7JS, UK, alex.southern@aecom.com

Abstract
Research into the auralisation of road traffic noise has drawn more attention in the recent years, as it can provide a more direct and immersive experience with which to communicate the implication of major urban design decisions with city planners and the public. Some synthesis models of single vehicle pass-by noise, and methods for outdoor sound propagation modelling have been proposed, which may be integrated for auralisation of traffic flow scenes. In this paper, a framework for traffic flow auralisation is presented. The synthesis model is composed of multiple single vehicle pass-by examples, based on tyre noise and propulsion sound synthesis. A first-order image source method is implemented, as a first approximation to represent sound propagation across a local region of a wider urban environment. The resulting plausibility of the proposed traffic flow auralisation model under different traffic flow scenes are under assessment using subjective listening tests.

Keywords: Auralisation, Traffic Flow, Procedural Audio

1 INTRODUCTION

Road traffic is a widespread source of environmental noise in our daily life. The World Health Organization (WHO) recognizes that traffic noise can cause serious public health problems, especially under the circumstances of higher traffic volumes and speeds. In order to control or reduce road traffic noise more effectively, some published guidelines for traffic noise evaluation (1) and noise prediction methods (2) are usually used for reference. Most of these guidelines and methods are set based on noise levels in dB(A), which can be inconvenient when communicating with non-acousticians. In some situations, noise levels might be misleading as studies on traffic noise annoyance show that there can be a significant difference in annoyance judgment for different traffic flow conditions with the same L_{Aeq} (the equivalent A-weighted noise level within a period) (3). Auralisation can be potentially used as an alternative to predict and evaluate road traffic noise in addition to noise level metrics, as it provides subjective interpretation of the varying time-frequency dependent characteristics of the traffic noise. With an optionally detailed road traffic auralisation system available, people will be able to have an audible experience of the planned road network, traffic management, or traffic control developments.

According to the definition of auralisation in (4), there are three fundamental elements to be considered for road traffic scenes: 1) Moving sound source models for different vehicle types, speeds, and traffic volumes; 2) sound propagation models for an outdoor environment; and 3) sound reproduction system for high plausible auditory rendering. As there are no standardized frameworks to implement such auralisations for road traffic, the perceived plausibility may vary considerably, depending on what models and algorithms are used, and how complex they are in each of the elements above.

Based on the literature review presented in (5), two primary sound sources of pass-by vehicles should be considered - tyre rolling noise, and propulsion noise generated by engine, intake system, and exhaust system. These sources can be synthesized by different methods, e.g. granular synthesis (6), spectral modelling synthesis (7), or physical modelling synthesis (8), with pros and cons for each method, leading to differences in plausibility.

Considering traffic noise auralisation for a ‘microscale’ urban area (e.g. a street, a square), two different types of sound propagation models can be considered, defined as engineering methods (based on geometrical acoustics), and wave-based methods (solving the governing physical equations). Engineering methods work well for scenarios where only a limited number of reflections from buildings for simulation, and when meteorological conditions can be simplified (9). Wave-based methods should be used if high levels of detail and physically accurate results are required. However, wave-based methods are more computationally demanding, and are sometimes limited through a lack of robust physical input data regarding materials, surface impedance, and meteorological conditions, etc.

Regarding the sound reproduction system, the techniques used for traffic noise auralisation resemble those used for indoor environments in most literature, without additional considerations for the outdoor conditions. Head Related Transfer Function (HRTF) techniques are mostly used in recent traffic noise auralisation studies, if spatial audio rendering is required.

The structure of this paper is as follows: in Section 2, some of the previous models for vehicle pass-by noise auralisation are reviewed. In Section 3, the development of an auralisation model for traffic flow is introduced. Section 4 discusses results and plausibility assessment methdos of the proposed auralisation model and is followed by the conclusions in Section 5.

2 REVIEW OF SOME PREVIOUS MODELS

2.1 Auralisation Frameworks

Nilsson & Forssén et al. (10) developed a demonstrator, called LISTEN-Demonstrator, to illustrate the potential of auralisation as a tool for evaluating future sound environments or soundscapes. The propulsion sound sources are first modelled via additive synthesis, and then granular synthesis as an improvement. Spectral modelling synthesis (SMS) techniques (20) are used for heavier road vehicles because of the more prominent tonal characters of the diesel engine sound. Tyre noise is modelled according to engineering methods Harmonoise and Nord2000. Sound propagation effects including distance attenuation, air attenuation, ground effect, and diffraction, are modelled by Harmonoise and Nord2000 methods across one-third octave bands from 25Hz to 20kHz. The CIPIC HRTF database (21) is used in post processing for spatial audio rendering. The demonstrator is implemented in Pure Data (PD) for real-time performance. They conducted listening tests to investigate the perceived realism, annoyance, velocity, and similarity to real recordings of the demonstrator, and used subjective evaluation results to optimize the auralisation parameters.

Piernen et al. (7) proposed an auralisation model for accelerating passenger cars. Tyre noise is synthesized by a modified version of Harmonoise model, with an extra consideration of road surface correction, and horizontal directivity pattern for the sound source. The engine sound signal is assumed to consist of a deterministic signal representing the most important engine orders, plus with a quasi-stochastic signal in this work, and so SMS techniques are used. The time-varying characteristics of the tonal signal and the stochastic signal are assigned with control parameters including engine speed, engine load, and emission angle, etc. The propagation effects involved in this model include ground reflection and air absorption. Ground reflection is implemented by a secondary signal path with an additional delay and a complex reflection factor which is calculated from an empirical model according to the surface impedance of ground. Air absorption is implemented by introducing linear-phase FIR filters, for which the coefficients are designed according to the standard ISO 9613-1. This study provided some samples of synthetic sounds using this model, but did not conduct subjective evaluation for the plausibility of the auralisation method.

Southern & Murphy (11) presented a framework for road traffic noise auralisation for capture, analysis and synthesis of moving road vehicles within a virtual auditory environment. Only tyre/road noise was considered. The analysis and synthesis process are performed in the time-frequency domain and incorporate changes at the listening position based on an approximation of the vehicle's directional noise radiation pattern.

2.2 Source synthesis models

Maillard & Jagla (12) proposed a method of synthesising engine sound and tyre noise using a granular synthesis technique. Grains are extracted from the analysis of a recorded engine sound signal corresponding to continuously varying engine speeds, and tyre noise recordings via the Close-Proximity Method (CPX), a standard measurement method for noise levels emitted from a passenger car tyre when rolling over a road surface. Each grain is assigned with a vehicle speed or engine speed as the first control parameter, in addition to engine load as a second control parameter. This method is computationally effective for real-time usage of engine sound and tyre/road noise synthesis, but its flexibility in terms of synthesising different types of vehicle sound is relatively limited as it is difficult or impossible to take recordings of all the vehicles required.

2.3 Propagation models

Georgiou & Hornikx et al. (13) proposed a method to auralise a car pass-by using binaural impulse responses (BIRs) computed with a 2-D wave-based pseudospectral time-domain method (PSTD), which can provide physically accurate predictions for urban sound propagation at low frequecies. They implemented the proposed auralisation under simplified scenarios where buildings are absent, and a long flat wall is located behind the car, opposite to the receiver. They investigated the effect of the spacing between the discrete source positions in the perception of the auralised car pass-by noise using subjective listening tests, and found that it is difficult to distinguish between the auralised sounds when different angular spaces (2° - 10°) corresponding to the discrete source positions are used. However, the Doppler effect was not simulated within this work, which can be crucial for the perceived plausibility of vehicle pass-by auralisation, and might also be relevant with the audibility of switching between discrete source positions.

3 DEVELOPMENT OF TRAFFIC FLOW AURALISATION BY PROCEDURAL AUDIO

3.1 Considerations on procedural audio

The authors previously presented a framework for traffic flow auralisation (14), in which only road/tyre noise was considered, and the directivity pattern of the source was not implemented at that time. The framework is further developed in this paper, by adding propulsion sound synthesised via a physically-inspired model, an image source model simulating sound propagation effects in an outdoor environment, and horizontal tyre/road source directivity patterns, in order to improve the plausibility of traffic flow auralisation. The improved model has been implemented using *procedural audio* methods, which are defined as "non-linear, often synthetic sound, created in real-time according to a set of programmatic rules and live input" (15). Although procedural audio is originally used for sound design in a computer game, it has also been extended into a series of tools for rendering virtual acoustics for the auditory effects of virtual reality (VR) (16). Compared with sample-based methods, procedural audio has higher flexibility, with continuous real-time parameters being applied, which makes it more suitable for interactive applications, e.g. allowing movement around a virtual environment, or auditory changes due to the modification of some elements in the scene. As there are no pre-existing audio samples involved, procedural audio has variable cost of computational power – the more complex the sound model is, the more work it requires. This can be seen as an advantage under the limited conditions of computational resource for real-time usage, as the cost can be tailored by adjusting the complexity levels of the models. Plausibility is one of the most important factors for consideration when using procedural audio. Some sound models are complicated, and may cause severe timbral distortion due to the lack of sufficient details at a low complexity level, or sometimes it is impossible for real-time usage because of the high computational power required.

3.2 Rendering traffic flow scenario

Figure 1(a) shows an example of a traffic flow scenario in an urban environment, which is simulated in this paper. Each vehicle is allocated a constant vehicle speed, from left to right (V_n), or vice versa (V'_n). A distribution function is used which describes the enter/leaving time of each vehicle. The length of the road is L(m), and the listening position S facing the road is set D(m) away from the middle of the road, with free access to movement along the road. There are buildings facades set separately on each side of the road, of finite length H_{a1}(m), H_{a2}(m), H_{b1}(m), and H_{b2}(m), respectively. The visual simulation is implemented within the game engine Unity3D, utilizing an open-source simulator AirSim (17) developed by Microsoft AI & Research, and a screenshot is shown in Figure 1(b). Although visual aspects are not the main concern of this paper, these scenes provide opportunities for interactive VR applications in terms of vehicles pass-by in an urban environment.

(a) (b)

Figure 1. (a) Example of a traffic flow scenario in an urban environment; (b Interactive simulation of traffic flow scenario in an urban environment in Unity3D

3.3 Framework for traffic flow auralisation

Figure 2 demonstrates the auralisation framework used in this work. The source synthesis models are implemented in MAX/MSP, following procedural audio concepts to render the vehicle pass-by sound in real-time. An open-source tool called 'μ Max-Unity Interoperability Toolkit' (18) is used for the two-way parameter communication between MAX/MSP and Unity3D in real-time. Compared with the previous implemented model in (14), the main development includes: propulsion sound synthesised via a physically-inspired model, an image source model simulating sound propagation effects in an outdoor environment, and horizontal tyre/road source directivity patterns. These development are further discussed in the rest of this section.

The propulsion sound synthesis follows the idea in (19), which is a modified version of Farnell's engine sound model (8) by replacing the signal-based phasor with a series of functions that depict the actual acoustic contribution of exhaust valve operation, fuel ignition, intake valve operation, and piston motion. The relevant components of the engine are treated as resonant cavities and represented by mutually interacting digital waveguides. The inputs, delays and feedback coefficients for each waveguide are then derived from the functions simulating the moving components of a four-stroke engine, including the exhaust valves, fuel ignition, intake valve operation, and piston motion. The harmonic part of the engine sound is mainly determined by simulating the motion of the piston attached to a rotating crank via a sine wave, together with the fuel ignition effect via the positive half of a sine wave scaled by time varying parameter which represents the time required for fuel explosion. Intake and exhaust valves are used to modulate the feedback coefficients: more energy is fed back to the delay lines during the closing period of valves, while more energy is sent to the output side of the waveguides during the opening period of valves. For passenger car engines, cylinders are connected to the intake and exhaust systems. These systems are modelled as waveguides with fixed length, fixed feedback on the free end, and variable feedback on the valve end. A separate white noise signal is fed into the intake system, which is then modulated by the intake valve operation to simulate the turbulence of the fuel-air mixture.

The image source model is implemented within the Unity3D game engine, utilizing the plug-in Wwise Reflect developed for Wwise, an audio middleware, which allowing non-programmers to have better control of how, when, and where sounds are triggered corresponding to inputs from the game engine. The image source method is implemented as a multi-tap time-varying delay line with filters to simulate early reflections in Wwise Reflect. This effect is assigned to each sound emitting game object called an 'emitter'. The positions of the image sources can be calculated using the Wwise Spatial Audio API with the parameters of the listener, the emitter,

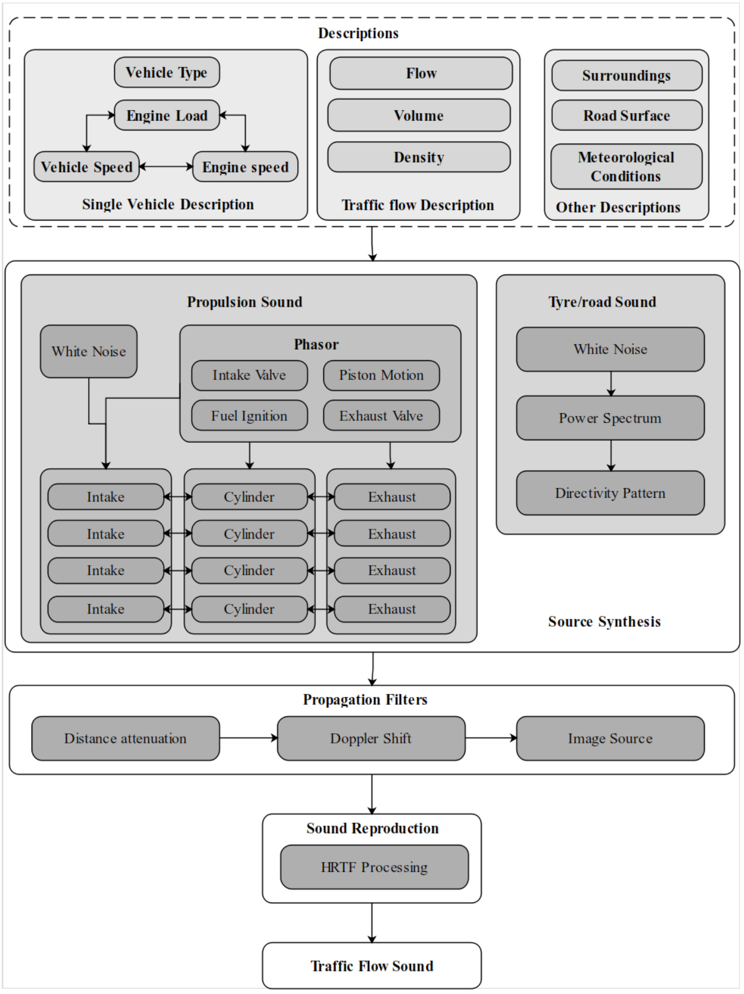


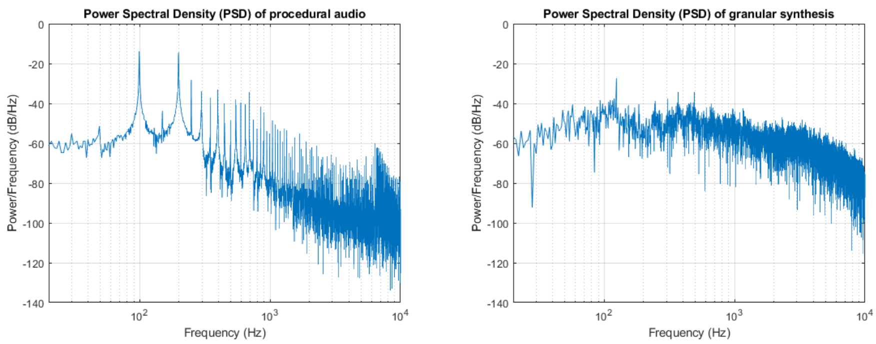
Figure 2. Framework for traffic flow auralisation using procedural audio

and the reflective surfaces defined in the Unity3D game engine. For the current version of Wwise Reflect plug-in, the acoustic property of the reflective surfaces can be manipulated by absorption coefficients over four frequency bands (250–4000Hz). Although up to the 4th-order of reflections is supported in Wwise Reflect, the computational power will increase exponentially with each extra order of reflection. Therefore, only 1st reflection is implemented in this work in order for the real-time performance. Some other propagation effects, such as diffraction, scattering, and meteorological effects, are not considered in the current work.

The horizontal tyre/road source directivity patterns are also implemented in the Unity3D game engine using the audio middleware, Wwise. The directivity pattern of the sound emitter can be implemented by depicting an attenuation cone-shaped boundaries to an omni-directional source. Energy within the attenuation cone will be scaled as a function of direction relative to the source's coordinate system. Although this method is not physically correct as the source directivity patterns can vary significantly at different frequencies, the perceptual accuracy of such methods used for sound design and interactive audio, specifically for the perceived realism of vehicle pass-by auralisation, has not yet been fully studied.

4 RESULTS

Figure 3 shows the power spectral density (PSD) of a propulsion sound signal (4 cylinders, with constant engine speed 4000RPM and 60% engine load) synthesised by different methods. Figure 3(a) is the PSD of the propulsion sound synthesised by procedural audio within the proposed framework. For comparison, the PSD of propulsion sound synthesised by a granular synthesis model (sample-based audio) is presented in Figure 3(b). As constant engine speeds and engine loads result in steady state sound, we took 0.5s of the signal with zero-padding for analysis using a 32768 points FFT. As can be seen from the figures, there is significant difference in spectral energy distribution when using different methods to synthesis the propulsion sound. In the procedural audio sound, there are sharp peaks at about 100Hz and 200Hz, while only one sharp peak in the granular synthesis sound appears around 125Hz. There is steady broadband sound component in both models, but the energy above 1kHz in the granular synthesis sound is higher than that in the procedural audio sound. For the frequency range about 200Hz–1000Hz, the energy distribution in the granular synthesis sound is more broadband, while in the procedural audio sound the energy distributes as a pattern with multiple harmonics.



(a) (b)

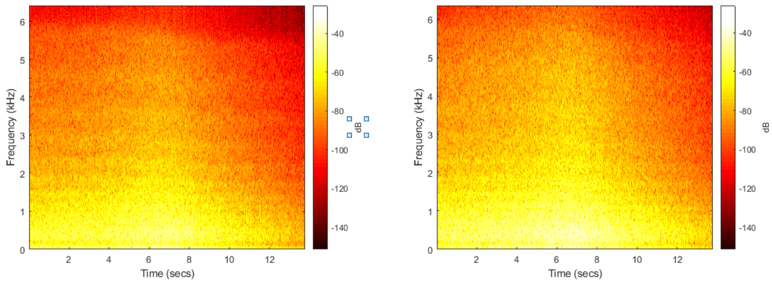
Figure 3. PSD of propulsion sound synthesis using: (a) procedural audio; (b) sample-based audio

Figure 4 shows spectrograms of a single vehicle pass-by sound signal, consisting of propulsion sound together with tyre/road sound (4 cylinders, with constant engine speed 4000RPM and 60% engine load, at constant driving speed 40km/h) synthesised by different methods. Figure 4(a) is the spectrogram synthesised by procedural audio within the proposed framework, with propagation filters implemented. For comparison, a spectrogram of propulsion sound synthesised by a granular synthesis model is presented in Figure 4(b), with the same propagation filters implemented. As can be seen from the figures, the energy distribution shares a similar pattern although different synthesis methods for propulsion sound are used. We are doing subjective listening tests to investigate how this kind of similarity in spectrograms correspond to the perceived difference for a single vehicle pass-by auralisation. The assessment results will be published in a future paper.

Figure 5 shows the spectrogram of a traffic flow scenario synthesised by the proposed auralisation framework, using procedural audio (Figure 5(a)), and granular synthesis (Figure 5(b)), respectively. The traffic flow consists of 6 vehicles on a two-lane road. The descriptive parameters of each single vehicle, such as speed, enter/leaving time, etc. are dereived from a microscopic multi-modal traffic flow simulation software named *PTV VISSIM*. We are doing listening tests to evaluate the plausibility of such traffic flow auralisations and the results. The assessment results will be published in a future paper.

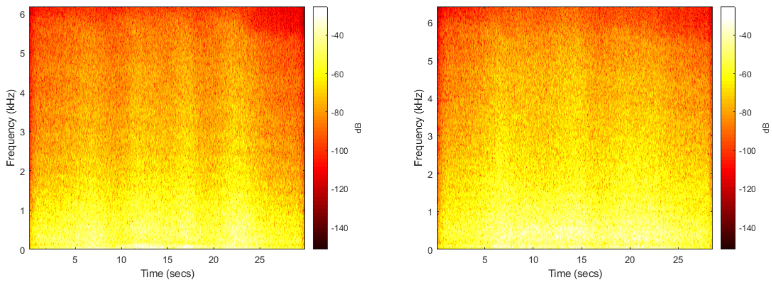
5 CONCLUSIONS

In this paper, a framework for traffic flow auralisation is presented. The model is a further development of a previous approach proposed by the authors, by adding the propulsion sound synthesised via a physically-



(a) (b)

Figure 4. Spectrograms of single vehicle pass-by synthesis using: (a) procedural audio; (b) sample-based audio



(a) (b)

Figure 5. Spectrograms of traffic flow auralisation using: (a) procedural audio; (b) sample-based audio

inspired model, an image source model simulating sound propagation effects in an urban environment, and horizontal tyre/road source directivity patterns, so as to improve the plausibility of the original model. The model implementation follows the procedural audio concepts, creating synthetic sound in real-time according to a set of programmatic rules, and live input. The simulation is conducted within the Unity3D game engine for interactive applications, with the support of AirSim for the visual part, and the support of Wwise for the auditory part. The plausibility of the proposed traffic flow auralisation model is under assessment by subjective listening tests. The assessment results will be published in a future paper.

ACKNOWLEDGEMENTS

This study is supported and funded by University of York - China Scholarships Council (CSC) joint research scholarships. We thank our colleagues in Audiolab who reviewed this work and assisted the research.

REFERENCES

[1] Boer, Eelco, and Arno Schroten. Traffic noise reduction in Europe: Health effects, social costs and technical and policy options to reduce road and rail traffic noise. CE Delft, Solutions for environment, economy and technology, 2007.

[2] Kephalopoulos, Stylianos, Marco Paviotti, and Fabienne Anfosso-Lédée. "Common noise assessment methods in Europe (CNOSSOS-EU)." Common noise assessment methods in Europe (CNOSSOS-EU), 2012.

[3] Kaczmarek, Tomasz, and Anna Preis. "Annoyance of time-varying road-traffic noise." Archives of acoustics 35, no. 3 (2010): 383-393.

[4] Vorländer M. Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. Springer Science & Business Media; 2007 Oct 27.

[5] Braun, Michael E., Stephen J. Walsh, Jane L. Horner, and R. Chuter. "Noise source characteristics in the ISO 362 vehicle pass-by noise test: Literature review." Applied Acoustics 74, no. 11 (2013): 1241-1265.

[6] Jagla, Jan, Julien Maillard, and Nadine Martin. "Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method." J Acoust Soc Am, no. 5 (2012): 3098-3108.

[7] Pieren, Reto, Thomas Bütler, and Kurt Heutschi. "Auralization of accelerating passenger cars using spectral modeling synthesis." Applied Sciences 6, no. 1 (2016): 5.

[8] Farnell, Andy. Designing sound. MIT Press, 2010.

[9] Hornikx, Maarten. "Ten questions concerning computational urban acoustics." Building and Environment 106 (2016): 409-421.

[10] Nilsson, M., Jens Forssén, Peter Lundén, Andrew Peplow, and Björn Hellström. "LISTEN Auralization of Urban Soundscapes." Stockholm University, Chalmers University of Technology, Sonic Studio, KTH Royal Institute of Technology, University College of Arts, Crafts and Design (2011).

[11] Southern, A.P. & Murphy, D.T. "A Framework for Road Traffic Noise Auralisation." Proc of Euronoise, Maastricht, the Netherlands, 2015.

[12] Maillard, Julien, and Jan, Jagla. "Effect of load on engine noise for the auralization of road traffic." Proc of Euronoise, Maastricht, the Netherlands, 2015.

[13] Georgiou, Fotis, Maarten Hornikx, and Armin Kohlrausch. "Auralization of a Car Pass-By Using Impulse Responses Computed with a Wave-Based Method." Acta Acust united Ac 105, no. 2 (2019): 381-391.

[14] Fu, Y., & Murphy, D. T. "A Comparative Overview of Traffic Flow Modelling Approaches for Auralistion." Proc of Euronoise, Crete, Greece, 2018.

[15] Farnell, Andy. "An introduction to procedural audio and its application in computer games." 2007.

[16] Masullo, Massimiliano & Fırat, Hasan & Maffei, Luigi. (2018). "Virtual acoustics with game engines.“ Proc of 25th International Congress on Sound and Vibration 2018 (ICSV 25), Hiroshima, Japan, 2018.

[17] Shah S., Dey D., Lovett C., Kapoor A. (2018) AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles. In: Hutter M., Siegwart R. (eds) Field and Service Robotics. Springer Proceedings in Advanced Robotics, vol 5. Springer

[18] Bukvic, I. I., and J. S. Kim. "myu Max-Unity3D interoperability toolkit." In Proceedings of the 35th International Computer Music Conference. Huddersfield, UK, 2011.

[19] Baldan, Stefano, Helene Lachambre, Stefano Delle Monache, and Patrick Boussard. "Physically informed car engine sound synthesis for virtual and augmented environments." In 2015 IEEE 2nd VR Workshop on Sonic Interactions for Virtual Environments (SIVE), pp. 1-6. IEEE, 2015.

[20] Serra, Xavier, and Julius Smith. "Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition." Computer Music Journal 14, no. 4 (1990): 12-24.

[21] Algazi, V. R., Duda, R. O., Thompson, D. M., & Avendano, C. (2001). The cipic hrtf database. In Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575) (pp. 99-102). IEEE.