Comparison of road tyre noise auralisation methods

Alex SOUTHERN[1&2]; Damian MURPHY[2]

[1] AECOM, Aurora, 120 Bothwell St, Glasgow, G2 7JS, UK

[2] AudioLab, Department of Electronics, University of York, Heslington, York, YO10 5DD, UK

ABSTRACT

Since the Environmental Noise Directive was introduced there has been considerable increase in research into the impact and abatement of road traffic noise. The World Health Organisation has recognised road traffic noise as a serious problem for public health, and annoyance with some aspect of our daily soundscape is well recognised as a common complaint. Auralisation tools can allow designers, planners and relevant stakeholders to listen to the acoustic consequences of a planned development and any associated noise mitigation for those most directly affected by it. An auralisation generally consists of three key components: sound sources, acoustic transmission paths and a calibrated soundscape listening system. The overarching goal of this work is to achieve a detailed road traffic noise auralisation system where the acoustic emission of every vehicle on the road network is accounted for at the desired listening position. This work extends a previously presented method for synthesising road tyre noise based on a small dataset of roadside recordings and validates the plausibility of this method in comparison to a recently published approach.

Keywords: Auralization, rolling contact noise perception, tires and road-tire interactions. I-INCE 11.7.1

1. INTRODUCTION

Research institutions, regulatory agencies and government bodies have recognised the importance of sustainably managing noise pollution, and, in particular, that arising from road traffic [1][2]. However, given the ephemeral qualities of sound it can be challenging for planners, consultants and the public to predict and understand the day-to-day consequences and impact of early stage design decisions for a planned development or intervention on the health and well-being of those most affected by it. Auralisation technology is more typically associated with room acoustic design, allowing the user to hear the effect of changes to the design, or acoustic treatments applied to, for example, a concert hall, lecture theatre or an open plan office [3]. In this way, auralisation can offer a more intuitive way to understand the perceptual impact of acoustic design considerations before the work is started. More recently, auralisation has been applied to outdoor environments, and for example, given that a larger proportion of people are exposed to road traffic noise in most European countries, see [4], it is unsurprising that some researchers have focused on methods to auralise arbitrary road traffic scenes, starting first with the synthesis of the noise caused by a single road vehicle [5][6].

More specifically, the method presented in by Piernen controls the time varying spectrum of tyre noise based on empirically derived noise level correction parameters, where linear regression coefficients are calculated based on repeated noise level measurements at e.g. different rolling tyre speeds and on different road surfaces. However, the underlying motive behind this method was to predict sound pressure levels due to radiated noise rather than for informing an auralisation system. As a consequence there is the potential for some audible sound characteristics to be lost in the synthesis of the road tyre noise that could lead to a less plausible auralisation experience for the listener in some situations, particularly when in close proximity to the moving vehicle as such specific nuances are less likely to be masked by the surrounding environmental soundscape, or when the road surface has deteriorated over time introducing surface discontinuities.

[1] alex.southern@aecom.com

[3] damian.murphy@york.ac.uk

In this paper, an alternative method to synthesising road tyre noise is discussed, as originally presented in [7]. It is proposed that this method inherently captures and reproduces features of a recorded vehicle pass-by such as *tyre thudding,* which may be considered to be an important feature of rolling tyre noise auralisation when incorporating the age of the road. This implementation is further perceptually validated in comparison to an implementation similar to that presented in [6].

The structure of this paper is as follows; in Section 2 the previously presented methods for road tyre noise synthesis are introduced. In Section 3 the listening test methodology is described. Section 4 presents the results of the listening test and is followed by the conclusions in Section 5.

2. PREVIOUS ROAD TYRE NOISE SYNTHESIS METHODS

Maillard proposed a method of synthesising tyre noise using a granular synthesis technique [5]. The method requires that the tyre to be synthesised is first recorded whilst mounted to a specially designed trailer. The microphone is positioned close to the tyre and the recording captures the tyre noise as the wheel speed is gradually reduced to a halt. The resulting recording is post-processed and divided into short sound samples called grains. Generally, each grain is assigned a speed and a new rolling tyre sound can be synthesised by concatenating the grains together in an appropriate manner. This method can be effective and is very computationally light, however it is predicated on the acquisition and application of a specialist trailer which is both costly and timely to use, and so is not considered further in this work.

More recently Piernen proposed the synthesis of road tyre noise based on a pink noise model [6]. The pink noise is passed through a one-third octave bandpass filter network and the relative gain of each band may be controlled by a well-established logarithmic relationship between vehicle speed and noise level. Empirically derived level correction terms also account for spectral differences due to the road surface type and the instantaneous horizontal angle with respect to the listening position. However, it is worth noting that the angular level correction was originally intended for predicting sound pressure levels rather than informing an auralisation system. As such the presented method, while simple and elegant, can potentially omit certain nuances of the sound source that would otherwise make the resulting auralisation more natural and plausible.

In 2015 the authors presented a new method for synthesising road tyre noise based on roadside pass-by recordings [7]. The approach here is to record single vehicle pass-bys in free-field conditions. The effects of amplitude changes due to the distance of the vehicle from the microphone (wave propagation divergence) and to Doppler shift are compensated as part of a post-processing phase. The resulting signal can therefore be considered as a representation of the angle and frequency dependent noise emission characteristics of the vehicle. The original work in [7] investigated the optimal encoding strategy with which to store the directivity dataset for each pass-by recording, although this aspect is not considered any further in this present paper. During the resynthesis of a vehicle's movement, the directivity characteristic dataset is used to filter a white noise source dependent on the angle of the vehicle to the listening position. Doppler shift and wave propagation divergence effects are then also applied as appropriate.

3. METHODOLOGY

The method presented in [7] uses a specific vehicle characteristic dataset on a per vehicle basis to synthesise road-tyre noise that is representative of the original recorded tyre noise. Similarly, Maillard's approach also results in a tyre specific dataset and therefore for brevity these approaches will be referred to as employing a *specific characteristic noise model*. A consequence of these methods is that they inherently have the capacity to capture and resynthesize the road's textural qualities that lead to occasionally prominent tyre vibrations due to local surface undulations or discontinuities in the road. These tyre vibrations can be described more commonly as *tyre thudding*.

In contrast, the *generic characteristic noise model* of road-tyre noise, refers to methods that synthesise the road-tyre interaction of a single vehicle based on filtering noise using empirically derived parameters encapsulating the averaged radiated noise from multiple vehicles/tyres. These approaches therefore do not inherently capture and resynthesize the road's textural qualities because the synthesis is based on a more generalised model of road/tyre interaction rather than that from a specific vehicle/tyre/road surface combination. Furthermore, while such a simplified model is appropriate for the prediction of road tyre-noise sound pressure levels over a duration i.e. using CNOSSOS-EU [8], it does not follow that they alone describe the complete perceived sound characteristic of rolling tyre noise in the context of a more complete auralisation.

This paper investigates the hypothesis that more generic synthesis methods, that do not consider tyre thudding, result in less plausible rolling tyre noise auralisations due to the perception of the smoothness of the resulting sounds. In the remainder of this section, a more detailed description of both implementations to be compared is given. It should be noted that both methods were used to synthesise the same recorded vehicle pass-bys and therefore a description of the recording methodology is first given in this section. Finally, a description of the test methodology used in this paper to compare the synthesis methods is provided.

3.1 Vehicle Recording

It was assumed that free-field conditions could be met by recording single vehicle pass-by events in open country-side free of any substantial acoustic reflections. The surrounding ground was soft grassy vegetation and there was generally no notable wind, but there was an occasional breeze that appears to have been mitigated by the use of a microphone windshield. A public road was used and none of the vehicle pass-bys were controlled - this was deemed important as it makes the resulting synthesis methods more robust and less costly. The recording setup comprised of a Tascam DR-680 portable recorder and an Earthworks M30 measurement microphone with environmental windshield. The microphone was positioned at 2.6m from the nearside edge of a dry 60 mph (~97 km/h) single carriageway road and at a height of 1.5m. The two methods discussed in sections 3.2 and 3.3 both use the same vehicle pass-by recordings described here to synthesise a vehicle pass-by event.

3.2 Specific Characteristic Noise Model

The synthesis approach presented in [7] is used here to represent specific characteristic noise models. The aim of this previous work was to explore the use of different encoding schemes to reduce storage memory requirements while aiming to retain a perceptually plausible representation of the recorded vehicle. In this study, to remove any unwanted bias in the results, we assume that memory limitations are not a concern and employ the perceptually optimal *Non-Compressed* (NC) scheme from [7] to synthesise the vehicle pass-by. Briefly, the NC scheme uses a sliding window FFT to generate a time-frequency representation of the directional magnitude response which must be smoothed to avoid tonal artefacts after synthesis.

3.3 Generic Characteristic Noise Model

The proposed generic characteristic noise model is represented by two different implementations. Each implementation synthesises the original recording with varying levels of detail in order to represent different types of generic model. For brevity, the most detailed method is labelled Type 1 (T1) and the less detailed method is labelled Type 2 (T2). Both T1 and T2 require that pink noise is first filtered by a one-third octave bandpass filterbank. The T1 method is defined as follows:

$$y_{T1}(t) = \sum_{i=1}^{N} 10^{\left(\frac{L_{T1,i}(v(t),\varphi(t))}{20}\right)} \cdot \varepsilon_i(t) \tag{1}$$

where $y_{T1}(t)$ is the synthesised rolling tyre noise according to the T1 approach, The term $\varepsilon_i(t)$ is the pink noise signal in the i^{th} one-third octave frequency band, of which there are N bands. Note that (1) is in the form presented by Piernen [6], however in this paper the noise level term $L_{T1,i}$ is derived by direct analysis of the recorded pass-by as opposed to using an empirically derived directivity function and linear regression parameters that determine methods for relating noise levels to speed and road surface type. In this way, T1 has a more favourable chance of generating an auralisation of rolling tyre noise that is a plausible representation of the original recording. This is because the effects of speed, road surface and source directivity are present in the recording and are therefore directly transferred into the synthesised audio sample used in the listening tests described in Section 3.4. The time-frequency dependent noise level function $L_{T1,i}$ is obtained here by fitting two straight lines to the logarithm of the time domain waveform, with an increasing and decreasing gradient as the vehicle approaches and departs the microphone position. The two lines generally do not meet at the vehicles passing-point sample, and so sudden discontinuities in the resulting noise level are avoided by a clipping procedure. Figure 1 shows a typical set of $L_{T1,i}$ noise level functions used in (1) to shape the pink noise, the plateaued top is a result of the clipping procedure.

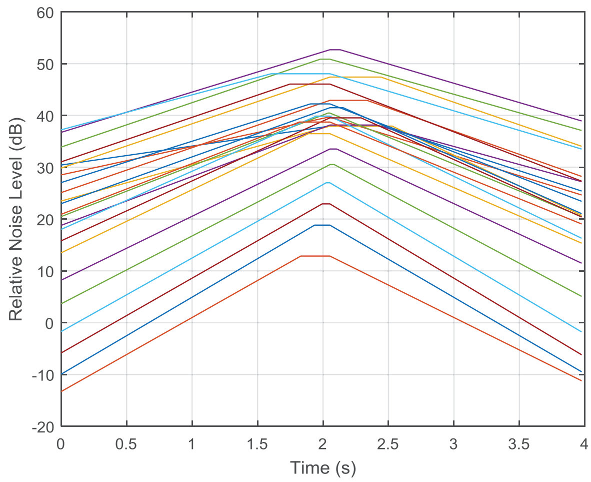

Figure 1 – Typical dataset of $L_{T1,i}$ noise level functions. Each function corresponds to the noise level evolution in a one-third octave frequency band and has been calculated from a vehicle recording to be synthesised.

The T2 method is defined according to (2). Note that this approach will result in an intentionally inferior auralisation to the other methods considered here, so that it may be used as a reference in the listening test described in the remainder of this paper.

$$y_{T2}(t) = 10^{\left(\frac{L_s(t)}{20}\right)} \cdot \sum_{i}^{N} 10^{\left(\frac{L_{T2,i}}{20}\right)} \cdot \varepsilon_i\,(t) \tag{2}$$

where $y_{T2}(t)$ is the synthesised rolling tyre noise according to the T2 approach. $L_{T2,i}$ is the i^{th} one-third octave band RMS noise level in dB, calculated over an approximate duration of 3 seconds as the recorded vehicle passed the microphone. The term $L_s(t)$ is the smoothed temporal envelope of the recorded pass-by in dB. The T2 approach therefore neglects to model any time dependent variations in the frequency domain and should therefore be perceived as being perceptually less plausible compared to the T1 and NC approaches.

3.4 Testing

As with [7], the aim is to synthesize pass-by events such that they are considered a similar and plausible representation of the actual recorded pass-by, rather than to synthesise a pass-by that is indistinguishable from the original recording. In the context of this specific study it is required to determine if the presence of tyre thudding in the synthesised signal leads listeners to perceive a more plausible auralisation over examples where this is omitted – does this particular aspect of the overall auralisation influence perceptual accuracy?

This listening test is therefore designed to find out if either the NC or T1 methods are considered to produce a more plausible auralisation of a recorded pass-by in the presence of a real pass-by recording as a reference and when participants are asked to listen critically for reproduced road-tyre interaction. To this end, the testing approach is close to MUSHRA although this listening test focuses on *plausibility* rather than *sound quality* [9].

The test data consisted of 9 recorded passenger car pass-bys, collected according to Section 3.1, that were presented as the reference sounds for each of 9 questions. These were selected ensuring that the vehicle was moving at speed but not accelerating, so as to avoid the presence of strong engine noise. The participants were then asked to rate the plausibility of 5 other audio samples, referred to as REC, NC, T1, T2 and T2R in the results that follow, on a continuous scale from 0 to 100.

Given that T1 in this work has been derived directly from the recording to be synthesised, it was thought worthwhile to also investigate how a more generalised representation of road-tyre noise i.e. using only spectral data as in CNOSSOS-EU [8], might compare in terms of plausibility. Therefore to favourably simulate this general case, two other synthesis approaches are used, namely T2 as described in Section 3.3 and a randomly chosen T1 rendering (T1R) corresponding to a different recording from the chosen dataset of 9. One of the samples presented (REC) was the hidden reference recording, with NC being the method applied from [7] as noted above.

4. RESULTS

The results of the listening test are presented in Figure 2 and summarise the mean scores and 95% confidence intervals for all 10 participants over the 9 sample questions. The mean scores show that as a group the participants ranked the NC method as the most plausible representation of the reference in each of the questions. Unsurprisingly, the results also suggest that the participants can distinguish between even more generalised methods (T2 and T1R) which further confirms that listeners are indeed sensitive to the temporal and spectral nuances that contribute to rolling tyre noise.

It is also important to comment on the extent of the confidence intervals which suggest that there were occasions where participants failed to identify the hidden reference and also rated T1 higher than NC. There were in fact two occasions where participants rated NC higher than the reference and an additional two occasions were they rated T1similarly. Given that the test was relatively short (10 to 20 minutes) it is perhaps not possible to attribute these as errors due to fatigue. In [7] NC was sometimes confused with the hidden reference and so this is of course possible again. However, this would also suggest that participants sometimes thought T1 was indistinguishable from the reference as well. This is also a reasonable assumption, supported by the fact that participants did sometimes rank T1 higher than NC. It is thought that this may be due to the original recording being considered as sounding *smooth* and therefore free of any prominent tyre thudding. This indicates that T1 is clearly not always going to produce a less plausible auralisation. However, given that it is unrealistic for rolling tyre noise to always be considered free of tyre thudding sounds, particularly on used and worn roads, T1 is therefore only appropriate for a limited and more ideal, set of road-tyre interactions.

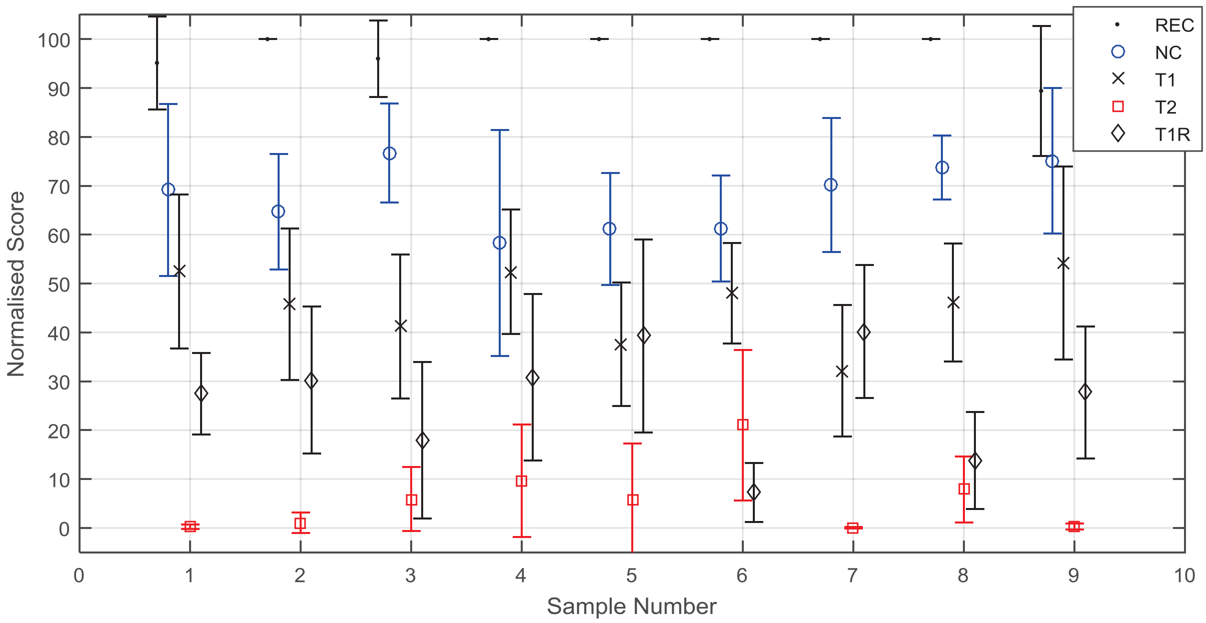

Figure 2 – The results the listening test. The mean score and 95% confidence interval are provided for each synthesis method for each of the 9 sample questions. Note the mean values and associated confidence intervals have been offset on the x-axis for readability.

5. CONCLUSIONS

This paper has discussed some preliminary results comparing the plausibility of synthesised vehicle pass-by noise to real recordings. Two different types of synthesis methodologies were tested, namely *specific characteristic noise models* and *generic characteristic noise models*. The specific models can capture and resynthesize the nuances of road/tyre interactions, such as those caused by local road surface undulations and discontinuities associated particularly with worn roads. The generic models reduce the synthesis complexity by parameterisation of a dataset of rolling tyre recordings. In particular, Piernen recently used a one-third octave band pink noise model controlled by empirically derived linear regression coefficients to auralise rolling tyre noise. The consequence of using such methods is that they can sound too smooth due to neglecting to reproduce the noise resulting from these road surface undulations and discontinuities.

The purpose of this paper was therefore to present the results of a listening test designed to determine whether synthesised pass-by events were in fact considered plausible representations of real rolling tyre noise if any tyre thudding evident in the original recording was reproduced as part of this synthesis. It should be noted that the overall aim was not to attempt to synthesise a pass-by indistinguishable from an original recording under critical listening conditions. A MUSHRA inspired listening test was used with one hidden reference recording and four synthesised rolling tyre representations. The specific characteristic noise model, NC, was represented using the synthesis approach previously presented in [7]. The generic characteristic noise model was represented by synthesis approaches of varying quality; the best of which, T1, being similar to that presented by Piernen in [6], but modified to give the approach a greater chance of being regarded as a more plausible representation of the recorded rolling tyre noise.

The results generally suggested that NC was the more plausible representation of real rolling tyre noise. This result is not surprising considering that specific models as in [5] and [7] were designed with auralisation in mind, whereas the method discussed by Piernen has adapted data originally intended for noise level prediction. That said, Piernen's approach is a simpler process to implement and it should be noted that there were occasions where some participants rated T1 to be better than NC. This suggests that there are some auralisation scenarios where this simpler method could be appropriate, such as for newly laid roads, distant roads where surface details cannot be perceived, roads masked by closer sound sources and ambient noise. However for auralisation of a specific vehicle, particularly in close proximity to the listening position, such specific methods are more appropriate.

Consequently this suggests that a complete auralisation system that is optimised for computational efficiency and perceptual accuracy would incorporate both approaches. With this in mind, it should be noted that the specific characteristic models of rolling tyre noise discussed in this paper can also be trivially adapted for synthesis based on the approached defined in (1) and as adopted by Piernen, making them suitable for a wide variety of auralisation scenarios depending on the required implementation aim.

ACKNOWLEDGEMENTS

Alex Southern was seconded to the Audio Lab, Department of Electronics at the University of York from AECOM supported by a Royal Society Industry Fellowship.

REFERENCES

1. European Commision Evironment. The Environmental Noise Directive (2002/49/EC). 2015. url: http://ec.europa.eu/environment/noise/directive_en.htm
2. World Health Organisation Europe. Night Noise Guidelines for Europe. 2009. url: http://www.euro.who.int/__data/assets/pdf_file/0017/43316/E92845.pdf
3. Fritschi, L, Brown A.L, Kim R, Schwela D.H, Kephalopoulos S. (Eds). Burden of Disease from Environmental Noise: Quantification of healthy life years lost in Europe. World Health Organisation, Regional Office for Europe, Bonn, and European Commission Joint Research Centre. 2011.

4. Southern A, Siltanen S, Murphy D, Savioja S. Room Impulse Response Synthesis and Validation Using a Hybrid Acoustic Model. IEEE Transactions on Audio, Speech, and Language Processing 21.9 (2013), pp. 1940–1952. issn: 1558-7916. doi: 10.1109/TASL.2013.2263139
5. Maillard J, Maillard J. Real time auralization of non-stationary traffic noise-quantitative and perceptual validation in an urban street. Proceedings of AIA-DAGA, 2013
6. Pieren R, Bütler T, Heutschi K. Auralization of Accelerating Passenger Cars Using Spectral Modeling Synthesis. Applied Sciences. 2015 Dec 24;6(1):5.
7. Southern A, Murphy D. A Framework for Road Traffic Noise Auralisation. 2015. Paper presented at EuroNoise 2015, Maastricht, Netherlands.
8. European Commission. Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council. Off. J. Eur. Union 2015, 58, 1–823.
9. Sporer T, Liebetrau J, Schneider S. Statistics of MUSHRA Revisited. Audio Engineering Society Convention 127. 2009. url: http://www. aes.org/e-lib/browse.cfm?elib=15021.