Spectral Modelling Synthesis of Vehicle Pass-by Noise

Yang FU[1]; Damian MURPHY[2]

[1,2] Audiolab, Department of Electronic Engineering, University of York, Heslington, York, YO10 5DD, UK

ABSTRACT

Spectral Modelling Synthesis (SMS) is a sound synthesis technique that models time-varying spectra of given sounds as a collection of sinusoids plus a filtered noise component. Although originally utilized to produce musical sounds, this technique can also be extended for analysis, transformation and synthesis of a wide range of environmental sounds, such as traffic noise. Simplifications based on psychoacoustic analysis can be conducted during the modelling process to avoid redundant data, which leads to perceptual similarity between synthesized sounds and the original recordings of vehicle pass-by noise. In this paper, we investigate if this perceptual similarity can be described by objective metrics, and how to improve the synthesis by tuning the parameters in the SMS algorithm. The results showed that vehicle pass-by sounds characterized by tyre and engine noise can be well synthesized with different parameter sets in the SMS algorithm. Furthermore, it is found that Zwicker Roughness is a sensitive metric for measuring the perceptual similarity between original recordings and synthesized sounds as it varies significantly when tuning SMS parameters.

Keywords: Spectral Modelling Synthesis, vehicle pass-by noise, Roughness

1. INTRODUCTION

Road traffic noise is an important aspect of our acoustic environment and can lead to environmental distress including complaints, dissatisfaction, and also health problems (1). When dealing with traffic noise issues, it is worthwhile to pay more attention to prediction and planning at an early stage before infrastructure work is started. Nosie mapping software based on numerical models are widely used to predict noise level indicators in specific areas (2, 3). Although many indicators have already been standardized for the comparison of different scenarios and lead to useful guidance, they provide only limited information of sound quality in terms of hearing perception. It is still difficult for planners and the public to clearly understand what the scenarios sound like with only specific noise level indicators given in a numerical form.

With the development of auralisation technology applied for outdoor environments, an intuitive impression of road traffic scenarios is becoming feasible. A virtual sound environment based on a 3D audio rendering system can be implemented to simulate real sound field for road traffic (4). The virtual sound environment should be authentic enough for perceptual evaluation, and also be interactive for assessment of different design and noise attenuation solutions in the planning stage (5). Three main tasks are involved for this kind of auralisation (4): the synthesis of the source signals (such as engine and tyre noise); the sound propagation model for filtering the source signals; and the spatial rendering of the sound field. Many of previous studies focused mainly on the latter two aspects and used vehicle pass-by recordings (6, 7). This limits the interaction property of the auralisation system as the source signals lack flexibility for a wide range of scenarios – the greater the modification to the recordings, the more unnatural the sound will be. This dilemma can be treated by using synthesized sounds as source signals because it is much easier to modify the synthesis by tuning some parameters in the algorithms. The synthesized sound can be still perceived as natural if the algorithm and the parameters are well defined.

Nevertheless, when using synthesized source signals, it is required to evaluate whether the synthesized sounds are plausible enough. This process can often be time consuming, especially when

1 yf852@york.ac.uk

2 damian.murphy@york.ac.uk

listening tests are required. A general idea to simplify this process is to develop some objective metrics that directly relate to the subjective evaluation. Although there is no standard metric to evaluate the plausibility of traffic noise synthesis so far, the objective sound quality metrics applied in the automobile industry could be taken for reference as the correlation between subjective ratings and these objective metrics has been widely studied. For example, Zwicker metrics (8) consisting of Loudness, Sharpness, and Roughness have been widely used to evaluate the perception of sound quality of in-cabin noise (9), door closing sound (10), engine roaring sound (11), etc., which provides useful cues for sound quality research and product design.

This paper presents a novel approach for vehicle pass-by noise synthesis which could be utilized as source signals for outdoor auralisation work. This approach is based on the Spectral Modelling Synthesis (SMS) technique which models time-varying spectra of given sounds as a collection of sinusoids plus a filtered noise component. In order to simplify the evaluation process of the plausibility of the synthesis, Zwicker metrics including Loudness, Sharpness, and Roughness have been calculated when tuning the SMS algorithm to explore how these metrics change when different parameter sets are used.

The structure of this paper is as follows: in Section 2, the framework and the implementation of SMS for vehicle pass-by noise are introduced. Section 3 describes the concept and the calculation methods of the objective sound quality metric including Loudness, Sharpness, and Roughness. The results and discussion of the synthesis and the objective metric calculation are presented in Section 4. Finally, Section 5 summarizes concluding remarks and perspectives on future work.

2. SPECTRAL MODELLING SYNTHESIS

2.1 The Framework of Spectral Modelling Synthesis

Spectral modelling synthesis (12) is a technique that models time-varying spectra as: 1) a collection of sinusoids controlled through time by piecewise linear amplitude and frequency envelopes (the deterministic part), and 2) a time-varying filtered noise component (the stochastic part). Although originally developed in computer music industries for creating musical sounds, SMS has also been successfully implemented for analysis, transformation and synthesis of traffic noise. For example, Pendharkar (13) proposed an SMS model to generate perceptually similar sounds to recordings. A listening test was conducted to verify the perceptual fidelity of the synthesized sounds.

Mathematically, in an SMS model, a sound signal $s(t)$ is seen as a sum of a series of sinusoids plus a stochastic part $e(t)$, defined as (12, 14):

$$s(t) = \sum_{r=1}^{R} A_r(t)\cos[\theta_r(t)] + e(t) \tag{1}$$

Where R is the number of sinusoids, $A_r(t)$ and $\theta_r(t)$ are the instantaneous amplitude and phase of the r^{th} sinusoid at time t, respectively, and $e(t)$ is the noise component at time t (in seconds).

The model assumes that the sinusoids are stable partials of the sound and that each one has a slowly changing amplitude and frequency. The instantaneous phase is then taken to be the integral of the instantaneous frequency $\omega_r(t)$, and therefore satisfies:

$$\theta_r(t) = \int_0^t \omega_r(\tau)d\tau \tag{2}$$

Where $\omega_r(t)$ is the frequency in radians, and r is the sinusoidal number.

By assuming that $e(t)$ is a stochastic signal, it can be described as filtered white noise:

$$e(t) = \int_0^t h(t,\tau)u(\tau)d\tau \tag{3}$$

Where $u(\tau)$ is white noise and $h(t,\tau)$ is the impulse response of a time-varying filter at time t.

Figure 1 shows a block diagram of SMS model. A series of spectra based on the sliding window and Short-Time Fourier Transform (STFT) technique is calculated and detected as the deterministic components of the original sound. Then the residual part can be obtained by subtracting the deterministic components from the original sound. After that, the shapes of time-varying spectral envelopes of the residual part can be derived by filtering white noise with the spectral envelop shape of each frame. Both the deterministic and the stochastic components are analysed and calculated in the frequency domain. These two parts are synthesized separately by Inverse Fast Fourier Transform (IFFT), and then summed together using the overlap-add method in the time domain, for each frame, to create the synthesized sound.

Peak detection and peak tracking are critical steps for capturing the sinusoidal components in SMS. A peak is defined as a local maximum in the magnitude of the spectrum. However, a simple maximum detection searching for the gradient of the magnitude spectrum from positive to negative is not sufficient. Since the frequency information in the spectrum is discretized in STFT, it is usually impossible to find an exact value of a peak. A combination of zero padding and quadratic interpolation technique with a suitable choice of window function has been utilized to improve the accuracy of peak tracking for this work (14). Once all the peaks in each frame have been detected, peak tracking technique is utilized to find the gradually changing sinusoids from the current time instant to the next time instant. In this process, the phase information of the original sound signal is disregarded except for the starting phases of the sinusoids based on which the phase are reconstructed during the synthesis process.

Once all the peaks are detected and tracked, the deterministic components can be obtained by additive synthesis. The residual part can be calculated by subtracting the deterministic components from the original sound signal. As noise perception is based on energy levels in bands rather than individual spectral peaks (12, 15), it is feasible to simplify the noise synthesis. The spectral envelope modelling approach can be conducted to detect the general shape with line-segment approximations (14).

During the synthesis process, deterministic components and stochastic components should be treated separately as they have different characteristics in spectra. For the deterministic component synthesis, the phase information should be reconstructed based on the initial phase in such a way that transitions between each of the frames are smooth in order to generate continuous sinusoids. Interpolation of the amplitude and phase information is required for each time instant. For stochastic component synthesis, the line segments should be interpolated to make each frequency index have a specific magnitude. A random phase can be assigned to the magnitude spectra as noise perception is mainly associated with power spectral density (12, 15).

2.2 SMS implementation for vehicle pass-by noise

Two types of vehicle pass-by noise are synthesized using SMS to explore the plausibility and sound quality produced by this method. The first is characterized by engine noise generated from a motorcycle pass by (at a speed approximately 50km/h), and the second is characterized by tyre noise produced by a single commercial car (at a speed approximately 30km/h). Both of these two types of sounds are common road traffic noise in daily life but they sound quite different. The engine noise sounds more ‘tonal’ while tyre noise sounds more similar to white noise. The specific information of the vehicle such as brand and model is disregarded as the aim of this study is to evaluate the suitability of SMS for traffic noise synthesis in terms of plausibility and to explore an objective metric to describe the plausibility.

2.2.1 Source Data

The source data was extracted from an Ambisonic B-format recording along a section of Scarcraft Road in the city of York, U.K. This road section is located in an urban area and traffic is usually not so busy, but with several single vehicles passing in the daytime, making it suitable for capturing the source data. The spatial information is discarded and only the W-channel recordings are used as mono audio files. The single motorcycle and single commercial car pass-by with as little interference signal (e.g. birds chirping, wind blowing, etc.) as possible are then selected. The instantaneous A-weighted sound pressure level (L_A) of the vehicle pass-by noise is also measured using a sound level meter for calibration and reproduction of the sound field in the following steps.

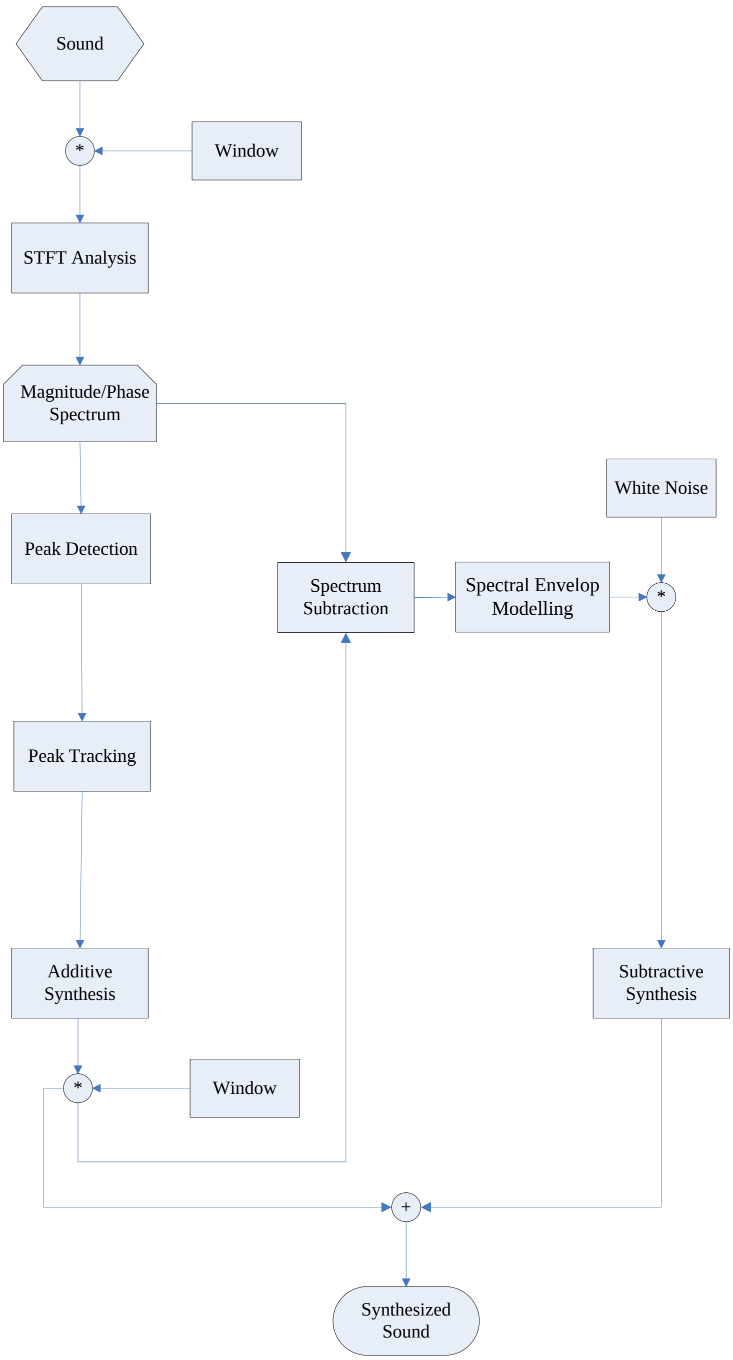

Figure 1 – Block diagram of SMS model

2.2.2 Parameters Tuning

When implementing SMS for a specific sound, some parameters in the algorithm should be well tuned to ensure the synthesis is effective and efficient. The parameter values may vary for different types of sounds according to their spectral analysis results.

Window size and FFT size are fundamental parameters for SMS. A good resolution of the spectrum is needed since the process that tracks the partials has to be able to identify the peaks which correspond to the deterministic component. Hamming window can be considered a suitable type of window because it has a relatively narrow main lobe (4 bins) with low side lobe (-43dB). The window size should be large enough to resolve the most closely spaced sinusoidal frequencies. However, the window should not be too long as this leads to poor temporal resolution. Therefore, there should be a suitable window size in terms of this time-frequency trade-off. The FFT size is set as the first power of two that is at least twice the window size with zero-padding for the difference of

FFT size and the window size. According to the guidelines above, several window size and FFT size values have been tested and evaluated in this work.

Noise threshold is a parameter that any signal having an amplitude below this threshold should be treated as noise and not be reconstructed using sinusoids. In this work, the background noise without any traffic noise in the near field has been measured as a reference value of the threshold (corresponding to approximately -80dB). In theory, any value lower than the reference can be used as a threshold as will not lose the sinusoidal information in the deterministic part (if the number of sinusoids is large enough. Several noise threshold values below the reference value have been tested and evaluated in this work.

Maximum number of sinusoids is a parameter that limits the maximum number of sinusoids per frame. The number of sinusoids should be set large enough to capture all the audible tonal components and leave the residual similar as much like noise as possible in order to get effective synthesis of the stochastic part. Several values of the maximum number of sinusoids have been tested and evaluated in this work.

Stochastic approximation factor is the smoothing factor of the down sampling for the stochastic synthesis. This value ranges from 0 to 1. The larger the stochastic approximation factor is, the better the frequency characteristics of the stochastic part will be, as a trade-off of computation power and time. Several values of stochastic approximation factors have been tested and evaluated in this work.

There are also other parameters that can be tuned in SMS, such as minimum duration of sinusoidal tracks, maximum frequency deviation in sinusoidal detection, etc. These parameters are just set fixed and have not been explored in this paper as these fixed values fit well with the different scenarios presented here. After obtaining a series of synthesis using different parameter sets, a preliminary listening test has been conducted based on a subjective evaluation of plausibility of different sounds and this is summarized as Table 1 and Table 2. As this is not a formal listening test, only the authors of this paper participated in this preliminary evaluation and the subjective evaluation ratings are not standardized or quantified. However, in this way the plausibility and quality of the synthesized sounds can be generally distinguished. For instance, in the case T1, the noise threshold is deliberately set too large and the synthesized sound lacks plausibility as expected. As another example, in the case T11 all the parameters are set for ‘an ideal synthesis’ at the cost of computation time and power so the outcome sounds highly plausible**.**

Table 1 – Different parameter sets for SMS implementation

Case No.	Scenario (Characterized sounds)	Window size	FFT size	Noise threshold (dB)	Maximum number of sinusoids	Stochastic approximation factor	Subjective evaluation of plausibility
T1	Tyre noise	883	1024	-60	100	1	Low
T2	Tyre noise	883	1024	-80	100	1	Medium
T3	Tyre noise	883	1024	-100	100	1	Medium
T4	Tyre noise	883	1024	-100	10	1	Low
T5	Tyre noise	883	1024	-120	100	1	Medium
T6	Tyre noise	1325	2048	-80	100	1	High
T7	Tyre noise	1325	2048	-100	100	0.8	Medium
T8	Tyre noise	1325	2048	-100	100	1	High
T9	Tyre noise	1325	2048	-100	120	1	High
T10	Tyre noise	2647	4096	-70	100	1	Medium
T11	Tyre noise	2647	4096	-120	160	1	High
T12	Tyre noise	2647	4096	-80	60	1	Low
T13	Tyre noise	2647	4096	-80	160	0.8	High

Table 2 – Different parameter sets for SMS implementation

Case No.	Scenario (Characterized sounds)	Window size	FFT size	Noise threshold (dB)	Maximum number of sinusoids	Stochastic approximation factor	Subjective evaluation of plausibility
E1	Eng noise	2647	4096	-120	120	1	High
E2	Eng noise	2647	4096	-60	120	1	Low
E3	Eng noise	2647	4096	-80	120	1	High
E4	Eng noise	2647	4096	-90	10	1	Low
E5	Eng noise	2647	4096	-90	120	0.1	Low
E6	Eng noise	2647	4096	-90	120	0.4	Medium
E7	Eng noise	2647	4096	-90	120	0.7	High
E8	Eng noise	2647	4096	-90	120	1	High
E9	Eng noise	2647	4096	-90	160	1	High
E10	Eng noise	2647	4096	-90	40	1	Low

3. SOUND QUALITY METRIC

During the process of tuning parameters, we have to listen to the synthesized sound directly in order to assess its plausibility and draw up some clues to improve the synthesis and then repeat these steps in several cycles. However, this subjective tuning approach may fail due to listener fatigue, and when we cannot distinguish the difference between different versions of the synthesis. When the parameters are finally tuned well and the sound is synthesized, a listening test is often required to evaluate its plausibility and sound quality. However it is also worthwhile to explore some objective metrics to evaluate the plausibility of the synthesized sounds. So far, there is no standard metric for this kind of evaluation. In this paper, Zwicker sound quality metrics as applied in the automobile industry including Loudness, Sharpness and Roughness are used to explore whether these metrics are sensitive for different parameter sets used in the synthesis. If a metric is sensitive to the variation of parameters, it can provide useful information for the objective evaluation of plausibility.

3.1 Loudness

Loudness is a metric representing the effect of the energy content of sound which is related to the decibel (dB) scale. In addition to energy consideration, Loudness is also dependent on the frequency content of a sound corresponding to the characteristics of human hearing. Mathematically, Loudness can be calculated as follows (8):

$$N = \int_{0}^{24\,\mathrm{Barks}} N'dz \qquad (4)$$

Where N is the overall Loudness (in Sone), N' is the specific Loudness for the critical band (in Sone), and z is the critical band rate (in Bark). The calculation is based on the programme in ISO 532/R (16) to obtain a series of time-varying Loudness values.

3.2 Sharpness

Sharpness is a measure of the high frequency content of a sound. The greater the proportion of high frequencies the ‘sharper’ the sound. Although it is a metric that is still not standardized, the Sharpness metric proposed by Zwicker has been widely used as it provides a useful measure of the high frequency content which is considered important to the quality of the product. The unit of Zwicker Sharpness is ‘acum’. A sound with Sharpness of 1 acum is defined as ‘a narrow band noise one critical band wide at a centre frequency of 1kHz having a level of 60dB’. Mathematically, Sharpness can be calculated as follows (8):

$$S = 0.11 \cdot \frac{\int_0^{24 Barks} N' g'(z) \cdot z \cdot dz}{\int_0^{24 Barks} N' dz} \quad (5)$$

Where S is Sharpness (in acum), N' is the specific Loudness for the critical band, z is the critical band rate (in Bark), $g'(z)$ is a weighting function with different values for different critical band rate (an example of weighting function can be found in (8)).

3.3 Roughness

Roughness is a complex effect which quantifies the subjective perception of rapid (15-300 Hz) amplitude modulation of a sound. The unit of Roughness is 'asper'. A sound with Roughness of 1 asper is defined as 'a 60dB, 1kHz tone that is 100% modulated in amplitude at a modulation frequency of 70Hz'. Although it is a metric that is still not standardized, Roughness has also been widely used in the automobile industry to describe the sound quality of the product, such as the engine roaring and the tyre noise. Mathematically, Roughness can be calculated as follows:

$$R = cal \cdot \int_0^{24 Barks} f_{mod} \cdot \Delta L \cdot dz \quad (6)$$

Where R is Rharpness (in asper), cal is a calibration factor, z is the critical band rate (in Bark), f_{mod} is the frequency of modulation (in Hz), and ΔL is the perceived masking depth (in dB, representing the differences between the maximum and the minimum of the temporal masking pattern).

4. RESULTS AND DISCUSSION

For the original recordings and each version of the synthesized sounds, the sound quality metrics of Zwicker Loudness, Sharpness and Roughness are calculated. The instantaneous values in dB(A) read from sound level meter during the measurement process are used as calibration values to calculate these sound quality metrics. For tyre noise scenario the calibration value is set 65.0 dB(A), while for engine noise scenario the calibration value is set 72.0 dB(A).

The hypothesis tested is that if the synthesis sounds more plausible, the metric values should be more similar to that of the original recording. In order to verify this hypothesis, we chose the synthesis cases T4 (Low plausibility), T7 (Medium plausibility), and T11 (High plausibility) in Table 1 for tyre noise scenario, and E4 (Low plausibility), E6 (Medium plausibility), and E1 (High plausibility) in Table 2 for engine noise scenario. For each case, we calculated the metrics and compare the corresponding metric with that of original recordings. The selected sounds are considered 'typical' to represent different subjective evaluation of plausibility (low, medium, high). As all the metric calculation results are time-varying, the percentile statistics of each metric are presented in order to represent their characteristics.

The results of percentile statistics of Loudness (N), Sharpness (S), and Roughness (R) for different versions of tyre noise synthesis and engine noise synthesis in comparison with original recordings are shown as Figure 2, Figure 3, and Figure 4, respectively. Here percentile is defined as a number where a certain percentage of values are larger than this number. For example, the value of N80 of the original recording of tyre noise is 7.6 Sone. This means over 80% of the loudness values of the original recording of tyre noise are larger than 7.6 Sone.

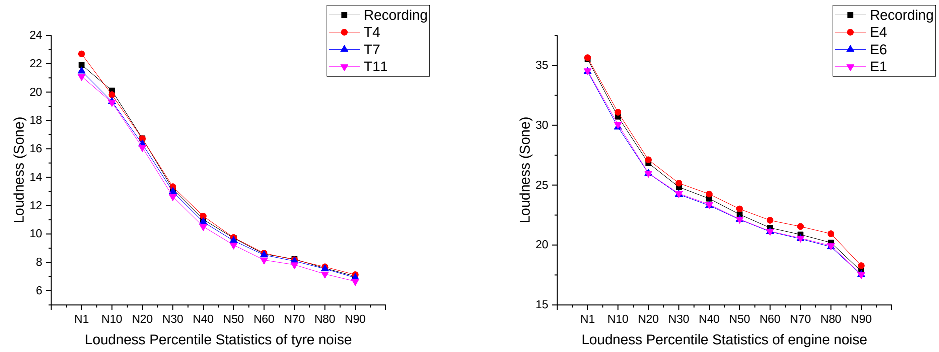

Figure 2 – Percentile statistics of Loudness of tyre noise synthesis (left) and engine noise synthesis (right)

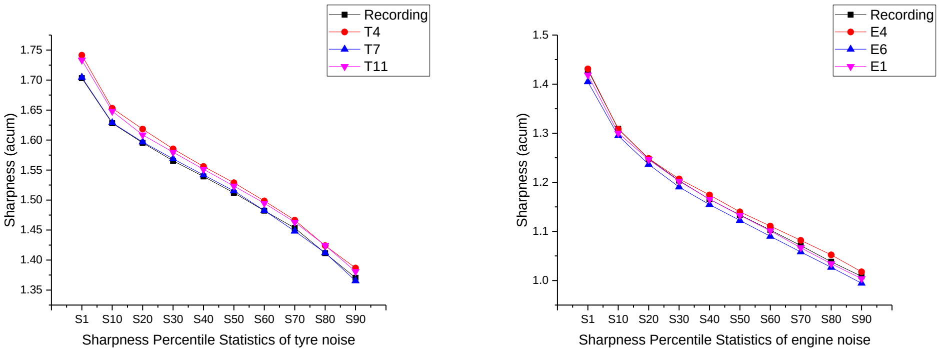

Figure 3 – Percentile statistics of Sharpness of tyre noise synthesis (left) and engine noise synthesis (right)

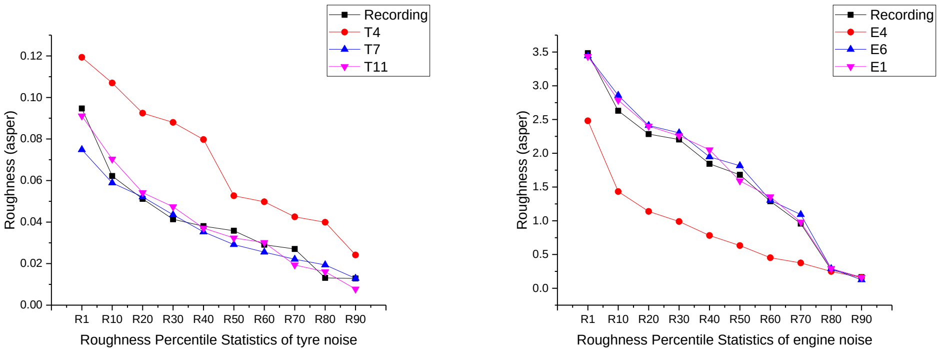

Figure 4 – Percentile statistics of Roughness of tyre noise synthesis (left) and engine noise synthesis (right)

As can be seen from the Figure 2, Figure 3, and Figure 4 above, Roughness is a relatively sensitive metric as the distribution values vary obviously in terms of sounds with different plausibility versions. The Roughness distributions vary drastically when the synthesized sounds with different plausibility versions are imported to calculate. The more plausible the synthesis is, the less the difference between the synthesis and the original recordings. This pattern can be seen in both of the scenarios characterized by tyre noise and characterized by engine noise but cannot be found in the calculation results of Loudness and Sharpness. The recordings and all the synthesized sounds share similar Loudness and Sharpness percentile statistics distributions regardless of how plausible they are.

This pattern can be also verified by Mann-Whitney U Test which is a non-parametric statistical hypothesis test suitable for small sample sizes. The null hypothesis (H0) is defined as there is no significant difference in terms of Roughness distribution in percentile statitics between the synthesized sounds and original recordings. In the 'Decision' column, H0 indicates the null hypothesis cannot be rejected, while H1 indicates the null hypothesis should be rejected with a 5% probability of error, which is equivalent to a 95% confidence level. The results of this significance test are shown as Table 3.

Table 3 – Mann-Whitney U Test Results in terms of Roughness

Case No.	Scenario	Element 1	Element 2	p-value	Decision
T4	Tyre noise	Recording	Low plausibility synthesis	0.019	H1
T7	Tyre noise	Recording	Medium plausibility synthesis	0.748	H0
T11	Tyre noise	Recording	High plausibility synthesis	1.000	H0
E4	Eng noise	Recording	Low plausibility synthesis	0.089	H0
E6	Eng noise	Recording	Medium plausibility synthesis	0.796	H0
E1	Eng noise	Recording	High plausibility synthesis	0.912	H0

As can be seen from Table 3, for each scenario the significance value (p-value) increases when the plausibility improves. Although the null hypothesis can be only rejected for Case T4 with a 5% probability of error, the significance values are small for synthesized sounds with low plausibility (0.019 for T4, and 0.089 for E4).The p-value increases when the plausibility improves. This provides strong evidence that Roughness is sensitive to the variation of plausibility.

5. CONCLUSIONS

An approach to vehicle pass-by noise synthesis using SMS technique is presented in this paper. The theory and framework of SMS are first introduced, followed by the specific implementation in terms of vehicle pass-by noise. Two different types of scenario have been synthesized, including one that is characterized by tyre noise and the other characterized by engine noise. As these two scenarios have different spectral characteristics, the parameter sets in SMS including window size, FFT size, noise threshold, maximum number of sinusoids, etc. should be well tuned for each scenario in order for a plausible synthesis. A series of parameter sets are presented concerning these two scenarios and each set results in a different version of plausibility.

In order to simplify the evaluation process of the plausibility of the synthesis, it is worthwhile to explore objective metrics rather than subjective listening tests for each case. In this paper, Zwicker metrics including Loudness, Sharpness, and Roughness have been explored to find if these parameters are sensitive in distinguishing different versions of plausibility of synthesis. It is found that Roughness can be considered a sensitive metric as its distribution of percentile statistics vary significantly for sounds with different subjective plausibility evaluation. It is also found that the difference of Roughness between original recordings and synthesized sounds decreases when the subjective plausibility evaluation improves. This pattern has also been verified by a null hypothesis test in terms of the variation trend of the significance value.

For future work, a formal listening test should be designed and conducted to explore how the variation in Roughness correlates with the subjective listening test results. It is also important to find out more objective metrics that can be used to measure the plausibility of synthesis. Another interesting direction is to explore the autocorrelation approach for the different parameters in SMS. This may be realised more easily if robust objective metrics have been proposed and taken for reference.

ACKNOWLEDGEMENTS

This research was supported and funded by University of York – China Scholarships Council (CSC) joint research scholarships. We thank our colleagues from Audiolab who provided insight and expertise that assisted the research.

REFERENCES

1. Dratva J, Zemp E, Dietrich DF, Bridevaux P-O, Rochat T, Schindler C, et al. Impact of road traffic noise annoyance on health-related quality of life: Results from a population-based study. Quality of Life Research. 2010;19(1):37-46.
2. Salomons EM, Zhou H, Lohman WJ. Efficient numerical modeling of traffic noise. The Journal of the Acoustical Society of America. 2010;127(2):796-803.
3. Maisonneuve N, Stevens M, Niessen ME, Steels L. NoiseTube: Measuring and mapping noise pollution with mobile phones. Information technologies in environmental engineering. 2009:215-28.
4. Maillard J, Jagla J, editors. Auralization of non-stationary traffic noise using sample based synthesis - Comparison with pass-by recordings. Internoise 2012, International Congress and Exposition on Noise Control Engineering; 2012.
5. Southern AP, Murphy DT, editors. Comparison of road tyre noise auralisation methods. 2016. Paper presented at Internoise, Hamburg, Germany.
6. Nilsson ME, Rådsten-Ekman M, Alvarsson J, Lundén P, Forssén J, editors. Perceptual validation of auralized road traffic noise. 40th International Congress and Exposition on Noise Control Engineering, Osaka, Japan, September 4-7, 2011; 2011: Institute of Noise Control Engineering/Japan & Acoustical Society of Japan.
7. Maillard J, Martin J. A simulation and restitution technique for the perceptive evaluation of road traffic noise. Proc Paris, France, Euronoise 2008.
8. Zwicker E, Fastl H. Psychoacoustics: facts and models. Springer series in information sciences (1999.
9. Noumura K, Yoshida J. Perception modeling and quantification of sound quality in cabin. SAE Technical Paper; 2003. Report No.: 0148-7191.
10. Parizet E, Guyader E, Nosulenko V. Analysis of car door closing sound quality. Applied acoustics. 2008;69(1):12-22.
11. Sellerbeck P, Nettelbeck C, Heinrichs R, Abels T. Improving diesel sound quality on engine level and vehicle level-a holistic approach. SAE Technical Paper; 2007. Report No.: 0148-7191.
12. Serra X, Smith J. Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. Computer Music Journal. 1990;14(4):12-24.
13. Pendharkar C. Auralization of road vehicles using spectral modeling synthesis. 2012.
14. Serra X. Musical sound modeling with sinusoids plus noise. Musical signal processing. 1997:91-122.
15. Goodwin M, editor Residual modeling in music analysis-synthesis. Acoustics, Speech, and Signal Processing, 1996 ICASSP-96 Conference Proceedings, 1996 IEEE International Conference on; 1996: IEEE.
16. ISO R. 532 B: Acoustics–Method for Calculating Loudness Level. International Organization for Standardization, Geneva. 1975.