Roughness calculation for randomly modulated sounds

Arne Oetjen[1], Uwe Letens[2], Steven van de Par[3], Jesko Verhey[4] and Reinhard Weber[5]
[1] *C.v.O. University Oldenburg - Acoustics Group, 26129 Oldenburg, Germany, Email: arne.oetjen@uni-oldenburg.de*
[2] *Daimler AG, 71059 Sindelfingen, Germany, Email: uwe.letens@daimler.com*
[3] *C.v.O. University Oldenburg - Acoustics Group, 26129 Oldenburg, Germany, Email: steven.van.de.par@uni-oldenburg.de*
[4] *O.v.G. University Magdeburg - Experimental Audiology, 39120 Magdeburg, Germany, Email: jesko.verhey@med.ovgu.de*
[5] *C.v.O. University Oldenburg - Acoustics Group, 26129 Oldenburg, Germany, Email: reinhard.weber@uni-oldenburg.de*

Introduction

The psychoacoustic sensation of roughness is an important feature in the overall characteristics of environmental sounds such as machinery noise. Especially the perception of the engine noise in passenger cars is strongly influenced by the magnitude of the roughness sensation, for example how sportive a car sounds.
In [1] we introduced an approach for roughness calculations based on subjective roughness ratings for different synthetic sounds. In contrast to the established algorithms this approach not only accounts for roughness of sinusoidally amplitude modulated tones [2] but also for a various set of tones modulated with more complex periodic waveforms like differently shaped triangles [3] and smoothed triangles [4].
In addition, there is a high correlation between experimental data and prediction of our algorithm for vehicle sounds in different operating conditions, at least within each operating condition (Figure 1). Correlation is still high when the operating conditions one and two are analyzed together, while the operating condition 3 (constant speed of 60 km/h) is overestimated by the algorithm.
Another limitation of the algorithm is that the absolute magnitude in asper seems to be too high since up to one asper was predicted for an idle operating condition (operating condition 1). These limitations motivated a modification of the algorithm based on a statistical approach that is described below.

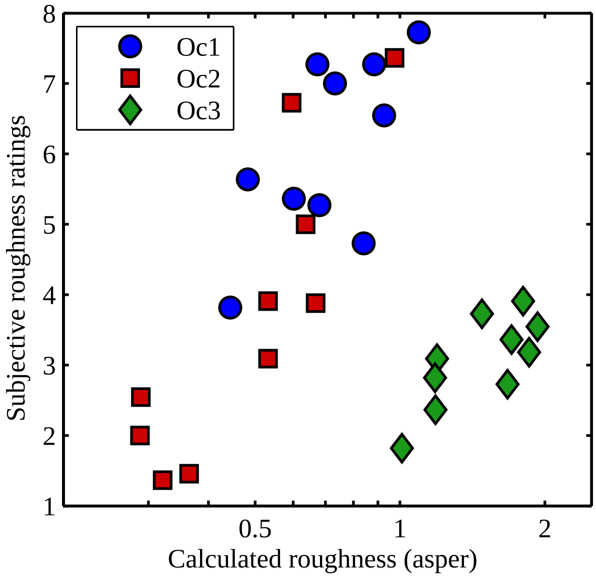

Figure 1: Roughness calculations (X-axis) and categorical subjective ratings (Y-axis) for vehicle sounds in three different operating conditions: engine idling recorded outside (Oc1, blue circles) and inside (Oc2, red squares) the vehicles and driving at a constant speed of 60 km/h (Oc3, green diamonds). The correlation coefficient of the calculations and the ratings is 0.10 for all sounds and 0.86 if the sounds in Oc3 (green diamonds) are taken out. The calculations were made with the approach suggested by Oetjen et al. [1].

Roughness predictions

Envelope analysis

When calculating the roughness from the envelope of a sound it is important to consider the sound sources generating this envelope. For example, vehicle interior sounds recorded at a sufficiently high constant traveling speed contain audible wind noise. Analyzing these kinds of sounds with the roughness calculation algorithm [1] by short, overlapping time windows ($\approx$ 0.2 s) results in envelope signals with high amplitudes and distinct modulation frequency spectra. Therefore high roughness values are calculated although the roughness of these sounds is not perceived that high.
A comparison of the modulation spectra of these sounds with the ones calculated from rough sounds such as certain engine idle sounds shows no qualitative differences in the single modulation spectra of one single time frame. The main difference can only be observed by inspecting modulation spectra calculated in consecutive time frames. Observing the absolute values of the modulation frequencies with the highest amplitudes shows low variations for sounds with distinct roughness perception whereas for vehicle interior sounds containing large amounts of wind noise these frequencies are distributed much less regularly. On the basis of this observation a weighting factor for calculated roughness values based on statistics on modulation frequency time series was introduced in the algorithm, as described in the next sections.

Statistics on modulation frequencies

A measure of the randomness of the distribution of a discrete set of samples from information theory is provided by the Shannon entropy [5] (Equation 1), where p_i denotes the probability that a sample value is observed

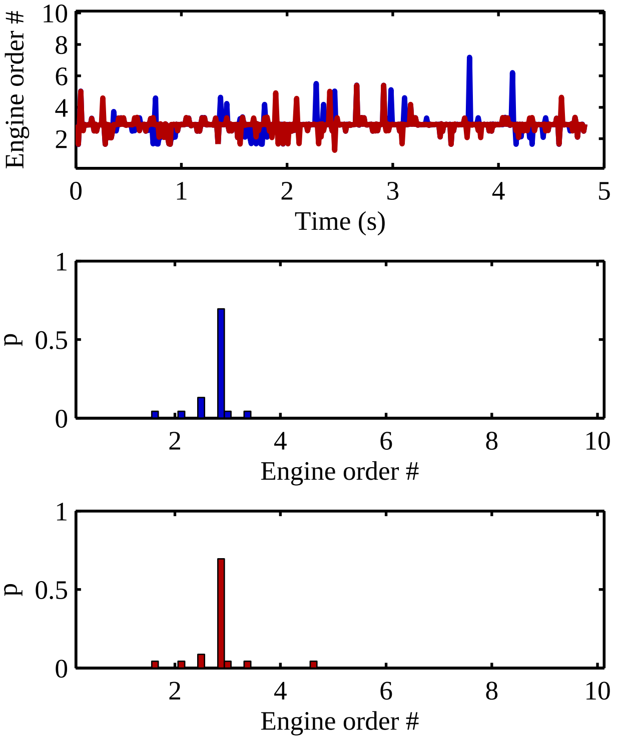

Figure 2: Modulation frequency statistics showing the chronological sequence (top) and histograms for the last 0.5 s (middle and bottom) of the quotient of the dominant modulation frequency and half the frequency of the first engine order provided from the instantaneous rotational engine speed for an engine idle recorded on outside the vehicle. The Shannon entropies are 1.53 for the blue and 1.65 for the red channel.

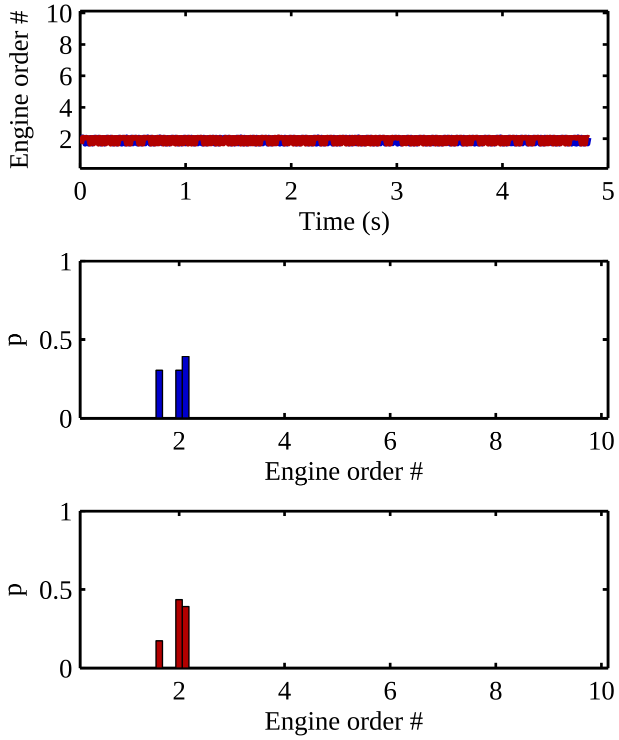

Figure 3: Modulation frequency statistics for an engine idle recorded inside the vehicle. The configuration of this figure is the same as in Figure 2. The Shannon entropies are 1.57 for the blue and 1.49 for the red channel.

within the i-th observation interval.

$$H = -\sum_i p_i \cdot \log_2(p_i) \qquad (1)$$

To apply this measure to the distribution of modulation frequencies over time, the frequencies are analyzed in observation intervals of a certain frequency width, i.e the occurrence of modulation frequencies in different frequency bins during a certain time period is represented in a histogram. For this data, the entropy H is calculated and provides information about the distribution of the modulation frequencies. The more randomly they are distributed, the higher the entropy.

In order to make this approach also suitable for vehicle sounds in transient operating conditions, such as acceleration processes, the detected dominant modulation frequencies are divided by the current rotational engine speed being available in an additional channel of the sound recording. This quotient gives the modulation frequencies in terms of the engine order as it is assumed that modulation frequency changes with a sufficiently small gradient will not reduce the magnitude of the overall roughness sensation. Examples of these statistics for different stationary operating conditions are shown in the Figures 2 to 4, the time interval in which the modulation frequencies were observed for the histogram statistics is the last 0.5 s of the modulation frequency time series shown in the upper panels of the Figures 2 to 4.

For the sound examples from engine idle processes (Figure 2 and 3) the histograms show a non-random distribution of modulation frequencies and therefore entropy ranges between $H \approx 1.5$ and $H \approx 1.7$. In contrast, the histograms for the example recorded at a constant speed of 60 km/h (Figure 4) show a more or less uniform distribution over a certain range of engine orders and an entropy of approximately 2.6.

Weighting factor for roughness calculations

One may conclude that the sensation of roughness also depends on the regularity of the frequencies in the envelope fluctuations over a certain time period as modulated sounds showing rapidly altering changes in modulation frequency do not seem to be perceived as rough. A time frame of 0.5 s seems to be an appropriate observation range for this consideration.

Hence, a weighting factor based on the distribution statistics of modulation frequencies was developed. The corrected roughness value in the time step τ, $\mathcal{R}_{corr}(\tau)$, is calculated as the quotient of the non-corrected roughness value $\mathcal{R}(\tau)$ obtained from the previously introduced calculation algorithm [1] and the Shannon entropy obtained from the last 0.5 s of the signal.

As it was assumed that the statistical measure of the Shannon entropy is not linearly related to the roughness sensation an adding constant $b > 0$ and an exponent $a > 0$ are included in the weighting factor (Equation 2).

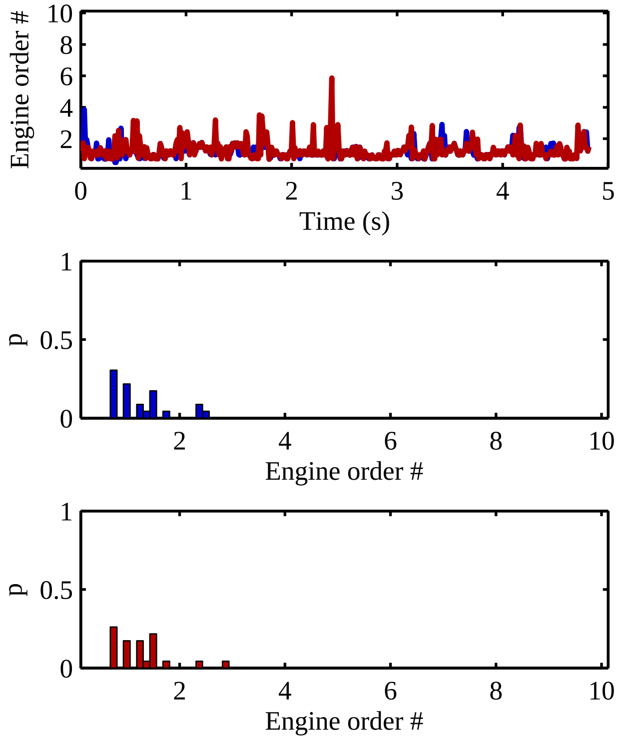

Figure 4: Modulation frequency statistics for a vehicle driving at a constant speed of 60 km/h. The configuration of this figure is the same as in Figure 2. The Shannon entropies are 2.64 for the blue and 2.65 for the red channel.

The values of the constants a and b were numerically fitted to a large set of subjective roughness ratings, in a first approach they were set to $b = 1$ and $a = 3$.

$$\mathcal{R}_{corr}(\tau) = \frac{\mathcal{R}(\tau)}{(b + H(\tau))^a} = \frac{\mathcal{R}(\tau)}{(b + \sum_i p_i \cdot \log_2(p_i))^a} \quad (2)$$

Introducing this weighting factor to the roughness calculation algorithm [1] results in an improved correlation between the calculated values and subjective roughness ratings for different vehicle sounds (Figure 5). The presence of the Shannon entropy weighting resulted in lower roughness values for sounds with randomly distributed modulation frequencies but did not affect the good correlation between calculations and ratings for sounds with clearly rough components that has already been established by the algorithm without weighting (Figure 1).

This also holds for a set of synthetic sounds that is used for tuning parameters in the algorithm which consists of amplitude modulated pure tones with different modulator waveforms such as sinusoids [2] or differently shaped triangles [3] and smoothed triangles [4]. As in this case the analyzed sounds all have a clearly defined stationary modulation frequency the usage of this weighting factor does not influence the roughness calculation.

Also the absolute values in asper seem to be much more realistic. For transient operating conditions, such as acceleration processes, and mixed (transient and stationary) processes the time series of calculated roughness values mostly correspond to the roughness development observed by expert listeners.

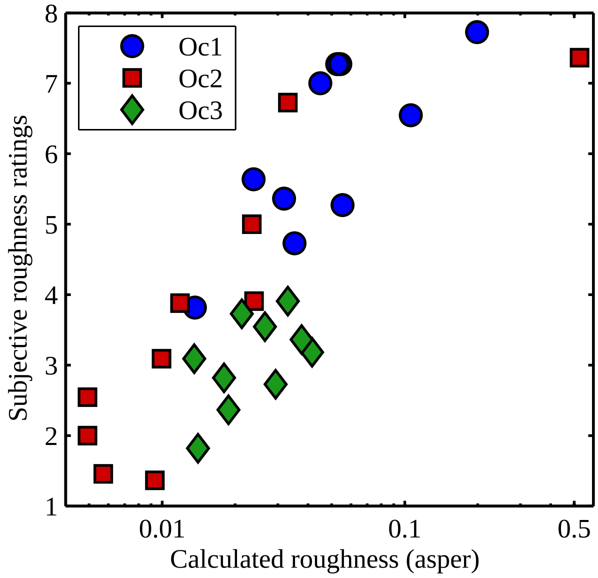

Figure 5: Rougness calculations (X-axis) and categorical subjective ratings (Y-axis) for vehicle sounds. The configuration of this figure is the same as in Figure 1. The correlation coefficient of the calculations and the ratings is 0.79. The calculations were made with the approach suggested by Oetjen et al. [1] and additional weighting with a measure derived from the detected modulation frequency sequence.

Summary and Conclusion

Based on observations made in the modulation frequency domain of vehicle sounds analyzed in short time windows a weighting factor for roughness calculations based on the Shannon entropy of dominant modulation frequencies found in previous time steps was introduced. Applying this weighting factor to an existent roughness calculation algorithm highly improved the correlation between roughness predictions and subjective ratings for a large set of different natural and synthetic sounds known from literature and own experiments.

References

[1] Oetjen, A., Weber, R. and Verhey, J.: Rauigkeitsberechnung unter Berücksichtigung der Einhüllendenform. Fortschritte der Akustik - DAGA 2012

[2] Fastl, H. and Zwicker, E.: Psychoacoustics - Facts andModels. Springer-Verlag, Berlin Heidelberg, 2007

[3] Yasui, N. und Miura, M.: Perception of Roughness on sounds amplitudemodulated with triangular wave. Forum Acusticum 2011

[4] Weber, R., Oetjen, A. and Oetjen, H.: Shape of the envelope and roughness perception. Fortschritte der Akustik - DAGA 2012

[5] Shannon, C. E. (1948), A mathematical theory of communication, Bell System Technical Journal 27, 379-423 and 623-656.