

Predicted Roughness Perception for Simulated Vehicle Interior Noise

Alois Sontacchi and Robert Holdrich
Inst. of Electronic Music and Acoustics

Josef Girstmair and Hannes Allmaier
Virtual Vehicle Competence Center

Stefan Bikker
Daimler AG

Alfred Rust
AVL LIST GmbH

ABSTRACT

In the past the exterior and interior noise level of vehicles has been largely reduced to follow stricter legislation and due to the demand of the customers. As a consequence, the noise quality and no longer the noise level inside the vehicle plays a crucial role. For an economic development of new powertrains it is important to assess noise quality already in early development stages by the use of simulation. Recent progress in NVH simulation methods of powertrain and vehicle in time and frequency domain provides the basis to pre-calculated sound pressure signals at arbitrary positions in the car interior. Advanced simulation tools for elastic multi-body simulation and novel strategies to measure acoustical transfer paths are combined to achieve this goal.

In order to evaluate the obtained sound impression a roughness prediction model has been developed. The proposed roughness model is a continuation of the model published by Hoeldrich and Pflueger [1]. Within the model simultaneous as well as temporal masking effects are considered. In addition, specific model parameters have been adjusted to predict subjective ratings of 18 experienced subjects, including mechanical engineers and audio engineers. The adapted roughness model has been developed by the usage of real sound stimuli measured in the car interior for different pre-defined engine types. Regression analysis shows that in most cases the subjectively perceived roughness can be predicted with good accuracy. Finally, the development model is tested with new stimuli not used in the development of the model; also for these new stimuli a good agreement of $R^2 \geq 88\%$ could be achieved.

After the discussion of the roughness prediction model, parameter variations for an automotive internal combustion engine (ICE) are discussed and compared with the aid of the new roughness model. From the results it is shown that the developed model is well suited to assess design changes and their consequences on the perceived roughness. Therefore, it can be used to develop roughness optimized solutions already in early design stages.

CITATION: Sontacchi, A., Holdrich, R., Girstmair, J., Allmaier, H. et al., "Predicted Roughness Perception for Simulated Vehicle Interior Noise," *SAE Int. J. Engines* 5(3):2012, doi:10.4271/2012-01-1561.

INTRODUCTION

As modern powertrains becoming more and more silent, their noise quality becomes one of the key parameters in the determination of the vehicle interior sound field. Typically cost intensive subjective assessments by experts are used to evaluate the noise quality. One important sound characteristic

of this noise quality is the engine roughness. Roughness is a well known psycho-acoustic measure, which describes the perception of competing frequencies that are close together so they cannot be separated by the hearing process and too far away to be in tune [2]. In order to reduce this time consuming evaluation procedure and to enable the determination of roughness already in early stages of development, a

roughness model is developed. While there exist several approaches to determine subjectively perceived roughness of synthetic sounds, their application to real engine sounds shows typically only a weak correlation to the subjective evaluation [1,8,11] and a new model is applied in the following.

To obtain a suitable data basis for the roughness model a number of listening tests were conducted. Deliberate planning of the listening tests was considered to be a crucial requirement as otherwise listening tests might be affected by a number of unintended factors. Even if the objective appears obvious several aspects have to be considered. In [3] several related perspectives concerning the term sound quality are discussed. If the requested definition is given the related properties have to be determined and evaluated. Considering our research aim, we do not head for sound quality benchmarking in general we seek to describe singular accessible perceived sound characters related to the term roughness. In common the assessment can be obtained via objective or subjective evaluation among defined framework conditions. Objective evaluation methods are based on modeling the human sound perception; they can deliver qualitative and quantitative acoustical descriptions of the investigated products.

However, only a limited number of application areas can be treated reliably [4]. The predominant crucial aspect due to the complex interaction of acoustical stimulus and auditory perception is neither linear and additive nor time-invariant. Overlapping and competitive acoustical stimuli can be instructively assessed by a group of so called expert listeners [5]. However, specific demands on the listening panel are claimed to extrapolate the results of the panel to the general population involved with the evaluated set of stimuli. Moreover, a well trained and proper selected panel will show consistent vocabulary usage and reliable rating behavior. Reduced variability within evaluation rating will shrink the confidence interval. Reliable and satisfying statistical conditions are obtained even with a reduced number of subjects and this will save money and time.

The main aim of this contribution is to establish an objective model that provides the possibility to predict the perceived roughness sensation. In the following section the development procedure of the objective model is briefly described. In the second section the roughness prediction based on pre-calculated sound pressure signals at arbitrary positions in the car interior using elastic multi-body simulation and considering transfer functions is presented.

THE DEVELOPMENT OF AN OBJECTIVE MODEL

As mentioned in the introduction, an expert listener panel was used for the listening tests as their familiarity with engine roughness gives more consistent and reliable results. To this task, a number of NVH engineers from Daimler, AVL and the Virtual Vehicle, as well as a number of audio engineers

formed an expert listener panel of 22 people with 18 people remaining after the post-selection.

To obtain reliable data from listening tests that can be used for the development of a predictive model, some preparations have to be taken. Firstly, the available subjects have to be examined concerning their hearing expertise. Therefore, hearing abilities of each subject have to be accessed by audiometric measurements. In addition, several hearing experiments with predefined correct answers have to be examined to obtain the specific hearing profiles of each subject. Secondly, to focus on particular sound attributes rather than on global quantities these attributes have to be identified. The identification process implicates that these attributes can be assessed by all subjects, are interpreted consistently among the subjects and further guarantee that these attributes are most widely complementary.

Within a preliminary listening test (no 1) the individual hearing thresholds of the subjects, envisaged members of the expert listener panel, were determined using the adaptive stimuli approach. In a two alternative forced choice (2AFC) paradigm subjects have to identify the correct temporal appearance of the hearable stimuli by using the three-down-one-up method. Starting from an ensured general hearable level threshold, the loudness level is stepwise decreased after three correct answers. If the answer is wrong, the level is increased immediately and the predefined level step size is reduced. As mentioned before, three additional experiments should highlight the subjects hearing profile in regard to repeatability and hearing acuity, as well as discrimination ability. To this task, following sound samples were prepared: pink noise to test the loudness perception, filtered pink noise for timbre differentiation and an amplitude modulated sine waves to examine analytical roughness perception. For each of these three experiments four sortable stimuli realizations with different characteristics are predetermined. Whereby, the largest and the smallest stimuli realization can be easily differentiated. However, the two stimuli in mid-position are at the threshold of noticeable difference. Based on these four stimuli six different pair combinations can be realized to investigate the discrimination abilities of the subjects. Dependent on the discrimination threshold the resulting stimuli pairs are repeated in the experiments several times.

From the preliminary listening test we obtain the following results: the hearing threshold loss and the intra-rater reliability. The later was calculated using the Kendall-coefficient for every listener. The Kendall-coefficient is a measure of the agreement between decisions of listeners. Figure 1 shows the score obtained from the Kendall-coefficients for all listeners and confirms the general suitability of the participants. The obtained results are profile characteristics of each panel member and this data will be used in the subsequent listening tests to interpret individual responses.

After identifying the listeners' ability profile the determination of relevant assessable attributes has been conducted in listening test no 2. These attributes have been

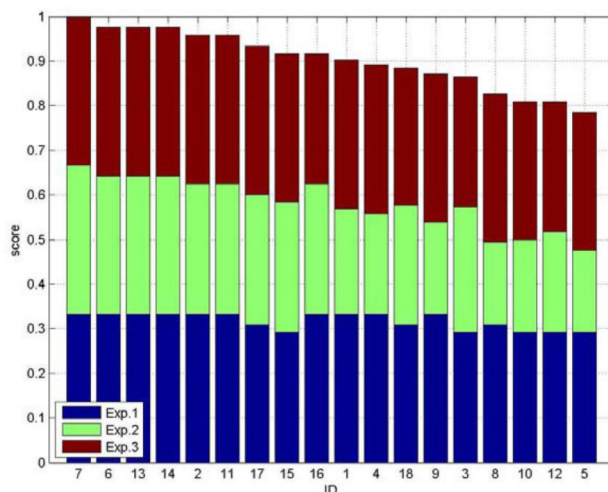


Figure 1. Intra-rater reliability calculated for every member of the expert listener panel

gained from individual attribute word lists after statistically examining their corresponding evaluation and discussing resulting candidates in an expert panel. The individual word lists were found by applying the Repertory Grid Technique (RGT) [12, 13] to pre-selected real engine sounds. Within this test procedure the following attributes were identified: roughness, r-roughness, buzzing, booming and humming.

In the subsequent listening test (no 3) the resulting attributes were evaluated by assessing real engine sounds and modified versions thereof in order to reveal the complementarities of these attributes as well as their differentiation against existing psycho acoustic measures (loudness, sharpness, tonality and impulsiveness). To this task, nine different sounds of real engine run ups were used as basis for the sound samples. For the listening tests, four equidistant samples of 5 second length were taken from each engine run up. These 36 samples were further extended by including 8 additional stimuli that were obtained by changing loudness (+/- 5 Phon) and/or sharpness (+/- 10%) of three different engine run-up samples. Results of this experiment indicate that the listeners on average assessing roughness will vote slightly higher or lower depending on increased or decreased loudness levels. However, assessing loudness they will vote in such cases significantly higher or lower. Therefore, reliable properties of our listeners under real conditions can be deduced. Examining the correlation of the different attributes - new assigned attributes as well as long term established psycho acoustic measures - we have to face with highly correlating items considering the new proposed attributes (see Figure 2). Therefore, roughness and r-roughness can be still treaded as one item, booming is related to the complex of loudness, sharpness and tonality, buzzing and humming can be also subsumed.

The actual listening test (no 3) consisted of two parts, where the intra-rater reliability and the inter-rater agreement were investigated. In the first part, nine different attributes had to be distinguished (roughness, r-roughness, humming,

booming, buzzing, loudness, sharpness, tonality and impulsiveness) and to test the intra-rater reliability some stimuli were presented more often to the listener than others. The intra-rater reliability (IRR) is then calculated as ratio of the variance of the repeated ratings σ^2_R to the total variance σ^2_{tot} :

$$IRR = 1 - \frac{\sigma^2_R}{\sigma^2_{tot}} \quad (1)$$

In the second part only the four attributes roughness, humming, booming, buzzing had to be distinguished and the consistency of the evaluations was determined. From the results one finds that the strongest correlations between the listeners exist for the roughness attribute and the weakest for humming, which provides a good basis for the objective roughness model.

Considering the determined IRR value for the roughness attribute in the second part in average about 0.67 between all members of the listener panel were achieved (see Figure 3, and compare IRR~0.74 value for loudness within the first part of this test series). However, based on the calculated IRR values and previous collected subject profiles, subjects with the IDs 3, 13, 18 and 22 are removed from the data to establish the objective model.

The results obtained from the listening test no. 3 were used as a basis for the generalized psycho acoustical model developed by Hoeldrich and Pflueger [1]. This model is developed for interior car sounds. It mimics the peripheral auditory system by taking into account the outer-middle ear filtering, non-linear masking properties and the absolute hearing threshold. By definition, roughness is a modulation of the envelope of a signal with a modulation frequency between 15 and 300 Hz [7]. Therefore, the sensitivity to roughness depends on a number of factors, like the center