Applied Acoustics 113 (2016) 149–161

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Sound quality prediction of vehicle interior noise using deep belief networks

Hai B. Huang [a,b], Xiao R. Huang [a], Ren X. Li [a], Teik C. Lim [b], Wei P. Ding [a,*]

[a] Institute of Automotive Engineering Research, Southwest Jiaotong University, 610031 Cheng Du, Si Chuan, China
[b] Vibro-Acoustics and Sound Quality Research Laboratory, College of Engineering and Applied Science, University of Cincinnati, 45221 Cincinnati, OH, USA

ARTICLE INFO

Article history:
Received 11 December 2015
Accepted 21 June 2016
Available online 30 June 2016

Keywords:
Sound quality
Vehicle interior noise
Deep belief networks
Machine learning
Feature fusion

ABSTRACT

The sound quality of vehicle interior noise strongly influences passengers' psychological and physiological perceptions. To predict the sound quality of interior noise, a vehicle road test with four compact cars has been conducted. All recorded interior noise signals have been denoised via a discrete wavelet transform (DWT) denoising procedure and subsequently evaluated subjectively through the anchor semantic differential (ASD) test by a jury. In addition, a novel prediction method, namely, regression-based deep belief networks (DBNs), which substitute the support vector regression (SVR) layer for the linear softmax classification layer at the top of the general DBN's structure, has been proposed to predict the interior sound quality. The parameter selection of the DBN model has been compared and studied using a grid search. In addition, four conventional machine-learning-based methods have been introduced to enable a comparison of the performance with the newly developed DBNs. Furthermore, the feature fusion ability of DBNs has been studied by varying the amount of information that the dataset offers. The results show the following: (1) The accuracy and robustness of the proposed DBN-based sound quality prediction approach are better than those of the 4 other referenced methods. (2) The multiple-feature fusing process can strongly affect the prediction performance. (3) Finally, the unsupervised pre-training process of the DBNs can enhance the information fusing ability. Finally, the newly proposed regression-based DBN approach may be extended to address other vehicle noises in the future.

1. Introduction

Vehicles improve efficiency and have become one of the most important non-human entities in individuals' daily lives. Many people spend substantial amounts of time in vehicles; thus, the comfort and sound quality, i.e., NVH (Noise, Vibration and Harshness), of a vehicle have become increasingly important to passengers. In addition, vehicle noises, which constitute more than 70% of urban noise [1], have been studied for many years. Vehicle interior sound quality evaluation (SQE), or sound quality prediction (SQP), has accordingly become increasingly important in vehicle design [2]. A bad vehicle interior sound quality would not cause hearing damage; however, it would negatively affect passenger psychology and physiology [3]. In addition, the sound quality of a vehicle is a key factor in customers' purchasing decisions. Therefore, it is necessary and reasonable to apply sound quality prediction to passenger cars, as has been performed by many researchers and engineers in recent decades [2,4–10].

Different types of cars have different types of interior sound quality, including sporty, conservative, powerful, and refined. To improve the sound quality in vehicle designs, researchers and engineers should first understand how to evaluate interior noise [11]. Consequently, a substantial number of research studies related to sound quality evaluation have been conducted and have produced a relatively satisfactory subjective evaluation system [7]. The characteristics of interior noise perceived by passengers have been found to be slightly dissimilar to the characteristics of the sound being emitted [12]; to some extent, the characteristics are the result of a process related to the physical structural effects of the human hearing system. Therefore, many psychoacoustics-based metrics, such as loudness, sharpness, fluctuation strength, and roughness, have been developed and applied to predict interior sound quality in, for example, vehicles, trains [13] and airplanes [14], due to their good quantitative explanation of noise characteristics. However, further studies on the prediction of interior sound quality have shown that the relationship between human sensations and acoustical motivation presents strong nonlinearity, and

* Corresponding author.
E-mail addresses: huanghaibo214@my.swjtu.edu.cn (H.B. Huang), xiaoronghuang@my.swjtu.edu.cn (X.R. Huang), rxli@home.swjtu.edu.cn (R.X. Li), limt@ucmail.uc.edu (T.C. Lim), dwp@home.swjtu.edu.cn (W.P. Ding).

http://dx.doi.org/10.1016/j.apacoust.2016.06.021
0003-682X/

as such, no dominant psychoacoustics-based metric representation for sound quality prediction has emerged.

The strong nonlinear relationship between vehicle interior noise and subjective perception makes it an especially active research problem for the application of signal processing and feature learning. Thus, various non-stationary, or transient, analysis methods and machine learning techniques have been introduced to address this issue. Correspondingly, a wide range of practical applications for vehicle interior sound quality prediction have been reported in the literature, where successful methods include the wavelet transform (WT) [6], empirical mode decomposition (EMD) [15–17], artificial neural networks (ANNs) [18] and statistical learning models (SLMs) [19,20]. In addition to general interior noise, there are various other types of vehicle noise, such as heating, ventilation and air conditioning noise (HVAC) [9]; suspension shock absorber rattling noise [3]; and axle gear noise [21,22], which have been objectively predicted based on the above advanced methods. Despite their success, the more effective prediction of vehicle interior sound quality remains an intricate problem, and substantial work remains for the following three reasons. First, the representational ability of a sound quality feature is limited. Although there are many sound quality metrics for estimating human hearing perception, features seldom perform overwhelmingly better than other features. Second, most generally used models, such as ANNs, support vector machines (SVMs) and logical regression (LR), are shallow models, namely, these models only have one or even no hidden layer (Multi-hidden-layer ANNs will not be discussed in this paper due to their defect of gradient diffusion) and thus cannot fully consider the space distributions of features or learn the deeper feature representations alone. Third, the information fusion performance obtained based on the advantages of using multiple sound quality features is not fully utilized. Machine-learning-based methods have been found to be superior to non-machine-learning-based methods in certain regards when applied to the feature fusion task; however, fusion methods continue to lack adequate study.

Despite the challenges in predicting vehicle interior sound quality, a new deep machine learning method, called deep belief networks (DBNs), brings new hope to addressing the above problems. DBNs were first proposed by Hinton in 2006 [23] and currently receive substantial attention in both the signal processing field and machine learning field while being successfully applied to pattern recognition problems, especially image classification [24], speech recognition [25], and natural language feature learning [26]. DBNs represent a probability generation model with a hierarchical structure and are stacked by many restricted Boltzmann machines (RBMs) [23]. In addition, DBNs provide various advantages such as the ability to encode deeper and higher order network topological structures and the ability to prevent overfitting and falling into local minimum through a special unsupervised pre-training phase [27]. Analysis of the similarity between sound quality prediction and pattern recognition in the above successful applications has motivated collaboration with their analysis process. Therein lies the great potential in utilizing state-of-the-art deep machine learning techniques to address the challenges faced in the prediction of vehicle interior noise. However, generally, almost all DBN research is focused on classification problems; the application of regression issues has received very little attention. Thus, in this paper, we propose a novel regression-based DBN method that replaces the top classification layer of DBNs by a regression layer. Moreover, this method could be easily extended to any regression-based machine learning problem.

This paper is organized as follows. In Section 2, we briefly review related work on sound quality prediction methods. In Section 3, the basic techniques of DBNs are introduced, and a regression-based DBN model is proposed. In Section 4, we describe the experiments and methods used in sound quality prediction. Section 5 reports the experimental results and further analyzes the superior performance of DBNs. Finally, Section 6 gives the conclusion and presents some future research topics.

2. Related works

The prediction of vehicle interior sound quality has been studied for many years. Roughly speaking, the reported methods for sound quality prediction in the literature can be briefly categorized into two classes, namely, psychoacoustics-based methods and machine-learning-based methods, both of which will be introduced below.

2.1. Psychoacoustics-based interior sound quality prediction

To numerically quantify acoustical stimuli, the ISO (International Standardization Organization) has generalized objective sound quality evaluation methods into two types [28]. The first type is the octave-band-based analysis for linear or weighted sound pressure levels. The most generally used analysis using this method is the one-third octave-band analysis with A-weighting due to its high calculation speed and simplicity. The second type is the critical-band-based analysis for the human sensation of loudness. The masking effect considered in the latter method makes it more effective and accurate than the former method. Typically, the standard ISO 532B [29] has adopted the second method to calculate loudness levels and has been proven to be a very powerful method for sound quality prediction in automotive engineering [7,30]. In addition to loudness, there have been many other psychoacoustic metrics developed by scientists and engineers, such as sharpness, roughness, fluctuation strength, articulation index, and tonality, although they have not been completely recognized. These sound quality metrics have been widely utilized in research on interior noise.

Wang et al. [8] used a roughness index to evaluate vehicle interior noise and proposed a modified roughness metric, HAP-RM, based on human auditory perception; their method provided good results in analyzing stationary and non-stationary sound signals. Parizet et al. [31] analyzed the car door closing sound quality based on some psychoacoustic metrics and found that two timbre parameters, namely, frequency balance of the sound and its cleanness, were the most effective indices for evaluating this special type of noise. Lee et al. [32] found that it is difficult to predict the gear whine sound quality when only using the A-weighted sound pressure level but that the loudness, articulation index and tonality can be used to objectively and efficiently evaluate the gear whine noise. Huang et al. [33] adopted four psychoacoustic metrics to evaluate the vehicle suspension shock absorber ratting noise and found that all generally used sound quality metrics cannot accurately reflect this specific noise. Other studies have been conducted on psychoacoustic metrics and sound quality problems, as noted in Refs. [8,34].

In addition, many researchers have integrated the characteristics of practical noises with human hearing procedures as well as advanced signal processing methods, therein developing new sound quality metrics. Lee et al. [6] introduced a special sound quality index, namely, HFEC (High Frequency Energy Contribution), using WT to evaluate the impact and shock noise of a car. To predict the squeak and rattle (S&R) noise of vehicle suspension shock absorber, Huang et al. [3] proposed a quantifiable metric, namely, SMWVD (Sound Metric based on the Wigner-Ville distribution), that can extract the S&R parts of suspension shock absorber noise. Zou et al. [35] presented an IEI (Impact Energy Index) metric that uses order analysis and spectral analysis to evaluate and improve