Mechanical Systems and Signal Processing 84 (2017) 245–267

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Evaluation of vehicle interior sound quality using a continuous restricted Boltzmann machine-based DBN

Hai B. Huang [a,b], Ren X. Li [a], Ming L. Yang [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 10 April 2016
Received in revised form
6 July 2016
Accepted 8 July 2016

Keywords:
Sound quality evaluation
Vehicle interior noise
Feature extraction
Continuous restricted Boltzmann machines
Deep belief networks

ABSTRACT

The perception of vehicle interior sound quality is important for passengers. In this paper, a feature fusion process for extracting the characteristics of vehicle interior noise is studied, and an improved deep belief network (DBN) that uses continuous restricted Boltzmann machines (CRBMs) to model continuous data is proposed. Six types of vehicles are used for recording interior noise under different working conditions, and a corresponding subjective evaluation is implemented. Psychoacoustic metrics and energy-based criteria using the wavelet transform (WT), wavelet packet transform (WPT), empirical mode decomposition (EMD), critical-band-based pass filter, and Mel-scale-based triangular filer approaches have been applied to extract interior noise features and then develop a fusing feature set combining psychoacoustic metrics and critical band energy based on comparisons. Using the obtained fusion feature set, a CRBM–based DBN (CRBM-DBN) model is developed through experiments. The newly developed model is verified by comparing its performance relative to multiple linear regression (MLR), backpropagation neural network (BPNN), and support vector machine (SVM) models. The results show that the proposed CRBM-DBN model has a lower prediction error and higher correlation coefficient with human perception compared to the other considered methods. In addition, CRBM-DBN outperforms BPNN and SVM in terms of stability and reliability. The presented approach may be regarded as a promising method for evaluating vehicle noise.

1. Introduction

Currently, the demands for vehicle quality have increased. Specifically, vehicle sound quality has become one of the most important characteristics to vehicle consumers and manufacturers. Passengers can be considered as part of a vibro-acoustical system. Vehicle interior noise is the most intuitive parameter felt by passengers and could affect their decision to purchase a vehicle. Technical competition with competitors urges vehicle companies to improve the interior sound quality of their vehicles. Furthermore, increasingly strict government standards for vehicle noise have been implemented because vehicle noise accounts for more than 60% of urban noise [1]. Moreover, vehicle interior noise, which consists of airborne noise and structure-borne noise, negatively influences the physical and psychological perception of passengers. The acoustic

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

http://dx.doi.org/10.1016/j.ymssp.2016.07.014
0888-3270/

characteristics of a vehicle are an integral part of product identity and could also be regarded as the 'sound DNA' of the vehicle [2]. Therefore, the methods used for evaluating the interior sound quality of vehicles should be investigated further.

The sound quality of vehicle interior noise is influenced by three variables: sound field, auditory perception, and auditory evaluation [3], which cause the evaluation of sound quality to be a multidimensional task. The A-weighted sound pressure level is widely used to objectively evaluate the quality of vehicle interior noise because it is simple to apply and can easily be understood. Evaluations using the A-weighted technique have demonstrated that human hearing has highly complex level-dependent evaluation mechanisms [3]. In general, low-frequency sound waves can be transmitted over a longer distance than high-frequency sound waves on the basilar membrane. Therefore, high-frequency noise can be masked more easily by low-frequency noise [4]. The perception of a sound event is affected by the sound pressure level and by psychoacoustic characteristics, such as loudness, sharpness, roughness, fluctuation strength and other extended metrics (Some of the psychoacoustic metrics and their attributes are summarized in Table 1). Thus, the perceived properties of a sound are not identical to those of a sound that is being emitted. Perception involves the physics and psychology of the human hearing process. Previous research has found that loudness and sharpness constitute at least 50% of the subjective assessment of a sound [5–7]. Particularly, loudness and sharpness could constitute more than 70% of the subjective assessment for stationary vehicle noise [8]. However, psychoacoustic metrics are largely designed to represent the specific perceived characteristics of a noise, and it is difficult to find a psychoacoustic model to extract noise features that can precisely describe the perception responses for all individuals. Therefore, the individual psychoacoustic criterion is not entirely adequate for the overall objective quantification of vehicle sound quality [2].

People will experience different hearing perceptions when they receive different sound waves with different frequencies. Vehicle interior noise, including stationary and non-stationary vehicle noise, can be represented in the time and frequency domains. Thus, it is important to select a signal processing approach for extracting sound perception features based on human auditory properties. Recently, many acoustic scholars and engineers have extracted noise features by using advanced signal processing methods, including wavelet transform (WT), wavelet packet transform (WPT), empirical mode decomposition (EMD), Wigner–Ville distribution (WVD), and other methods frequently mentioned in the literature [8,16–18]. WPT has been applied to extract sound features for evaluating vehicle interior noise [18] and can be used as a specific filter bank that is similar to the critical bands in the human hearing system. The WVD-based method was developed and used to assess noise resulting from vehicle suspension shock absorbers [16]. The proposed sound quality criterion, the Sound Metric based on the Wigner–Ville distribution (SMWVD), was highly correlated with the corresponding subjective evaluation of rattling noise. The EMD in the Hilbert–Huang transform (HHT) method was adopted to estimate the noise resulting from slamming a car door [19], which means that EMD can extract the main impact characteristics of the noise resulting from door slamming. In addition, the critical-band-based method [18] and Mel-scale-based approach [20] have been applied to extract noise characteristics and have potential advantages for acoustic modeling. Human auditory perceptions are related to sound characteristics; hence, the extraction of noise features is an important sound quality evaluation issue, and extraction methods depend on the original sound signals and practical applications.

Various intelligent pattern recognition methods have been introduced because of the complexity and nonlinearity of human hearing perceptions and noise characteristics. Multiple linear regression (MLR) is the most commonly used method for predicting vehicle interior sound quality and has two major advantages: the small amount of required experiment data and the ease of interpreting the results. Lee et al. [17] and Kim et al. [21] used the MLR method to evaluate vehicle impact noise and modified the vehicle suspension components to improve the sound quality based on the developed MLR model. However, MLR has limitations when fitting the highly nonlinear characteristics of the human auditory perception process because of its linear property. The support vector machine (SVM) is a statistical learning method that can be used to map the input data from a low-dimensional feature space to a high-dimensional feature space by using nonlinear mapping and then performing linear pattern recognition in the high-dimensional space. Liu et al. [22] used the SVM model to predict engine-radiated sound quality and found that the SVM was successful for establishing a nonlinear relationship between the subjective and objective evaluations. However, because the characteristics of sound are distributed widely, the predefined kernel function in the SVM model might not be sufficiently powerful to express the characteristics of all features simultaneously. Another well-known pattern recognition method is neural networks (NNs), among which the backpropagation neural network (BPNN) is the most

Table 1
Sound quality metrics and their attributes.

Sound metrics	Attributes
Loudness	The auditory characteristic related to the intensity of sensations [9].
Sharpness	The auditory characteristic involving the high-frequency portion of a sound [10].
Roughness	The auditory perception property related to a frequency modulation of approximately 70 Hz and the amplitude modulation of a sound [11].
Fluctuation strength	The auditory perception characteristic involving a frequency modulation of approximately 4 Hz and the amplitude modulation of a sound [10].
Articulation index	The quantitative measure of the intelligibility of speech [12].
Tonality	The auditory perception property related to the pitch strength of sounds [13].
Tone-to-noise ratio	Used to record the prominent discrete tone in a noise [14].
Noise criterion	The specific measurement of indoor background noise [15].