Expert Systems with Applications 40 (2013) 1477–1483

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Improvement of accuracy in a sound synthesis method using Evolutionary Product Unit Networks

M. Dolores Redel-Macías [a,*], Francisco Fernández-Navarro [b], Pedro A. Gutiérrez [b], A. José Cubero-Atienza [a], César Hervás-Martínez [b]

[a] Engineering Project Area, Department of Rural Engineering, University of Córdoba, Campus de Rabanales, Edificio Leonardo da Vinci, 14014 Córdoba, Spain
[b] Department of Computer Science and Numerical Analysis, University of Córdoba, Campus de Rabanales, Edificio Albert Einstein, 14014 Córdoba, Spain

ARTICLE INFO

Keywords:
Airborne source quantification
Auralization
Sound synthesis
Sound quality
Evolutionary computation
Product Unit Neural Networks

ABSTRACT

Auralization through binaural transfer path analysis and synthesis is a useful tool to analyze how contributions from different sources affect the perception of sound. This paper presents a novel model based on the auralization of sound sources through the study of the behavior of the system with respect to frequency. The proposed approach is a combined model using the airborne source quantification (ASQ) technique for low-mid frequencies (⩽2.5 kHz) and Evolutionary Product-Unit Neural Networks (EPUNNs) for high frequencies (>2.5 kHz), which improve overall accuracy. The accuracy of all models has been evaluated in terms of the Mean Squared Error (MSE) and the Standard Error of Prediction (SEP), the combined model obtaining the smallest value for high frequencies. Moreover, the best prediction model was established based on sound quality metrics, the proposed method showing better accuracy than the ASQ technique at high frequencies in terms of loudness, sharpness and 1/3rd octave bands.

1. Introduction

Optimizing the acoustic package of a vehicle requires the identification of its radiating sources and an analysis of how they propagate. This can be done by means of auralization, i.e., the process of rendering audible the sound field of a physical sound source in a space, in such a way as to simulate the listening experience at a given position in the space modeled (Berckmans, Kindt, Sas, & Desmet, 2010; Redel-Macias, Berckmans, & Cubero-Atienza, 2010). For auralization, sound synthesis models have been developed that make possible both real listening experiences and the determination of sound quality with metrics. In this sense, the recent change in the ISO 362 procedure for vehicle Pass-by Noise testing can be cited as an example of the importance of the auralization model. Generally, the first step taken in these models involves the modeling of all contributing noise sources and the determination of all relevant transfer paths by a source-transmission path-receiver model. To determine these transfer paths, direct measurement techniques are either not applicable or very cumbersome in practice, as well as being time consuming and expensive; therefore, indirect techniques are preferred, such as sound intensity (Prasad & Crocker, 1983), nearfield acoustic holography (NAH) (Leclere & Laulagnet, 2009; Martin, Le Bourdon, & Pasqual, 2011; Maynard and Lee, 1985; Pasqual & Martin, 2012; Pézerat, Leclère, Totaro, & Pachebat, 2009) or airborne source quantification (ASQ) (Berckmans, Kindt, Sas, & Desmet, 2010). The main problem of indirect techniques is, generally, that their limited resolution is due to an ill-posed problem and the need for regularization strategies in the vast majority of cases.

An Artificial Neural Network (ANN) is a powerful learning technique used to perform complex tasks in highly nonlinear dynamic environments (Castaño, Fernandez-Navarro, Gutierrez, & Hervas-Martinez, 2012; Tasdemir, Saritas, Ciniviz, & Allahverdi, 2011; Torres, Hervas, & Garcia, 2009; Yilmaz & Kaynar, 2011). Several distinguishing features of ANN, such as their ability to learn based on the optimization of an approximation of nonlinear functions, make them valuable and attractive for regression purposes or classification tasks, among others (Canakci, Ozsezen, Arcaklioglu, & Erdil, 2009; Feng & Chou, 2011; Oğuz, Saritas, & Baydan, 2010; Tallon-Ballesteros & Hervas-Martinez, 2011). Different types of ANN have been proposed for regression purposes (Haykin, 2009), including: MultiLayer Perceptron (MLP) neural networks (Yuste & Dorado, 2006), where the transfer functions are logistic, or hyperbolic tangent functions, Radial Basis Function Neural Networks (RBFNN) (Bishop, 2006), General Regression Neural Networks (GRNN) proposed by Specht (1991), Product Unit Neural Networks (PUNN) (Martinez-Estudillo, Hervas-Martinez, Gutierrez, &

* Corresponding author. Address: Escuela Politécnica Superior, Edificio Leonardo da Vinci-Campus de Rabanales, Ctra. Madrid km 396, Spain. Tel.: +34 957 21 22 22; fax: +34 957 21 85 50.

E-mail addresses: mdredel@uco.es (M. Dolores Redel-Macías), i22fenaf@uco.es (F. Fernández-Navarro), pagutierrez@uco.es (P.A. Gutiérrez), ir1cuata@uco.es (A.J. Cubero-Atienza), chervas@uco.es (C. Hervás-Martínez).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.08.041

Martinez-Estudillo, 2008), etc. An additional problem related to the application of ANN models is the selection of the most appropriate network architecture to be used. Classical neural network training algorithms assume a fixed architecture (number of layers, number of hidden layer nodes and number of connections) but it is very difficult to establish beforehand the best structure of the network for a given problem. In the last few years, Evolutionary Algorithms (EAs) (Fernández-Navarro, Gutierrez, & Carbonero-Ruz, 2011) have demonstrated their great accuracy in designing near optimal architectures, with different kinds of ANNs (Martinez-Estudillo, Hervás-Martínez, Martinez-Estudillo, & García-Pedrajas, 2005; Yao, 1999).

In this paper, a PUNN is trained to model and identify transfer functions in the auralization of the sound process. The main advantage of PUNNs is their capacity to reflect high order relationships among the input variables and the output. But PUNNs have a major drawback: their training is more difficult than the training of MLP or the RBFNN. As mentioned above, in the standard sound synthesis model, transfer functions between the sound source and the receptor are difficult to establish and the main drawback is their ill-conditioned problem which makes it necessary to apply regularization strategies involving a change in accuracy depending on the regularization method used (Berckmans, Pluymers, Sas, & Desmet, 2008; Kim & Nelson, 2004; Leclere, 2009). After studying the behavior of the system with respect to the frequency, a novel strategy based on a combined model is proposed to determine the transfer functions between the sound source and the receptor of the sound synthesis. ASQ techniques for sound synthesis at low and mid frequencies ($f \leqslant 2.5$ kHz) and an Evolutionary PUNN (EPUNN) model at high frequencies ($f > 2.5$ kHz) were combined because the system shows bad conditioning at high frequencies as will be proven. To evaluate the accuracy of the model proposed, different sound quality metrics were used to assess the real experience in a receptor position.

The remainder of the paper is organized as follows: the first section offers an introduction and the motivation behind the work. Section 2 summarizes the different modeling techniques used in this work which are extended in the corresponding subsection to include the description of the method proposed based on EPUNNs. Section 3 presents the configuration set-up of the experimental design to measure and characterize the sound sources. Section 4 presents the results comparing other traditional techniques and the novel method proposed. Section 5 discusses the main conclusions of the paper.

2. Modeling techniques

The main problem associated with the inverse ASQ method is that the transfer functions present an ill-conditioned problem, regularization strategies being needed. The conditioning is represented by the condition number of the measured transfer

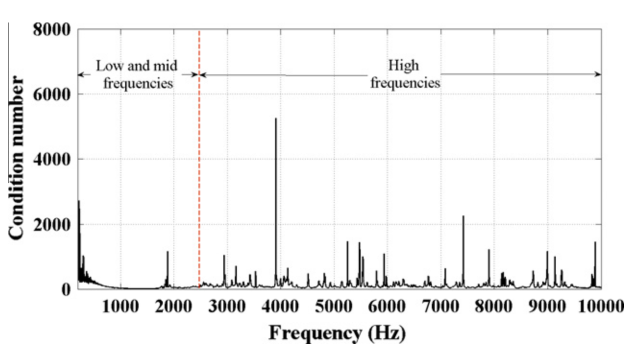

Fig. 1. Condition number of the system.

function matrix **H** in function of the frequency, which is shown in Fig. 1. Kim and Nelson (2004) established that the condition numbers can be said to be small when they are below 10^3. Therefore, it can be observed that for frequencies lower than 2.5 kHz, the condition number remains low, indicating a well-conditioned quantification problem. For high frequencies ($f > 2.5$ kHz), however, the condition number increases considerably and thus regularization strategies will probably be useful to come up with a physically relevant solution.

This paper considers two different models following the behavior of the system: the substitution monopole model based on the ASQ technique for the whole range of frequencies, and a combined model with an ASQ technique for low and mid frequencies, and an EPUNN model for high frequencies, comparing their accuracies based on the sound quality metrics of the best models.

2.1. Airborne source quantification model

The first step is to obtain the model of the sound sources using reciprocal measurements for which monopole sound sources are required. Then, in case of an experimental inverse quantification, several indicator field points are needed around the sound source; the operating sound pressures are captured at these points. After that, the transfer paths between each monopole location and all target field points are determined by means of reciprocity (Fahy, 2003; Leclere, 2009; Leclere & Laulagnet, 2009; Verheij, VanTol, & Hopmans, 1995). This is due to the fact that there are very different space requirements for sound sources and sensors, the measurements of acoustic transfer functions often being much more easily carried out by reciprocity, when the source and sensor are interchanged. Next, all substitute sources are quantified. In an experimental study, the source strengths $q_1 \ldots q_M$ are usually calculated through an inverse procedure. Generally, regularization has to be applied then to stabilize this inversion. Finally, a forward problem is solved for the calculation of the total sound field in the target field point, as a sum of the contributions of all monopole sources. The flow of the complete procedure is as follows:

I. $\mathbf{p}_{N\times 1} = \mathbf{H}_{N\times M}\mathbf{q}_{M\times 1}$
The relation between the complex volume velocities **q** of the *M* monopoles and the pressures **p** in *N* field points is modeled by means of a frequency response function matrix **H**.
II. To measure by reciprocity the transfer functions **H** between each of the *M* monopoles and the *N* field points.
III. $\mathbf{q}_{M\times 1} = \mathbf{H}^{-1}_{M\times N}\mathbf{p}_{N\times 1}$
The sound sources are determined by the inverse procedure.
IV. To measure by reciprocity the transfer functions **H** between each of the *M* monopoles and the *K* target microphone.
V. $\mathbf{p}_{K\times 1} = \mathbf{H}_{K\times M}\mathbf{q}_{M\times 1}$
Once the sound sources are known, the sound pressure at target position is obtained as a forward problem.

2.2. Combined ASQ-Evolutionary Product-Unit Neural Networks Model

A combined linear ASQ and EPUNN model is proposed in this paper, according to the behavior of the system regarding the frequency:

$$f(x,f,\theta) = \begin{cases} \text{ASQ model}, & \text{if } \mathbf{f} \leqslant 2.5\text{kHz} \\ f_{PUNN}(\mathbf{x},\boldsymbol{\theta}), & \text{if } \mathbf{f} > 2.5\text{kHz} \end{cases}$$

where **f** is the vector of frequencies, the ASQ model was presented in Section 2.1 and $f_{PUNN}(\mathbf{x},\boldsymbol{\theta})$ is the EPUNN model which is described in Section 2.2.1. Only low-mid frequencies are used to fit the ASQ model, and high frequencies are used to adjust the $f_{PUNN}(\mathbf{x},\boldsymbol{\theta})$ model, where $\boldsymbol{\theta} = (\boldsymbol{w},\boldsymbol{\beta})$ with $\mathbf{w} = (w_1,\ldots,w_m)$ and $\boldsymbol{\beta} = (\beta_0,\beta_1,\ldots,\beta_m)$, respectively.