

three diminsional sound reproduction in Vehicles based on data mining techniques

Maosheng Zhang, Ruimin Hu, Lin JIang

computer school, wuhan university
Bayi road, Wuhan, China

Abstract. The abstract should summarize the contents of the paper and should contain at least 70 and at most 150 words. It should be written using the *abstract* environment.

Keywords: Vehicles, sound, data mining, reproduction

1 Introduction

Sound systems for vehicle have been well researched by scientists and engineers. Akitoshi Yamada developed a sound reproduction system for vehicle using only a pair of loudspeakers in 1982[1]. The system comprised a transfer function, a delay circuit, and a reverberation circuit. With the help of these components, a surrounding sound system was implemented. Honda Motor designed a sound reproducing apparatus for vehicle in 1990[2]. The apparatus takes advantage of a acoustic duct and a loudspeaker placing in the duct. In 2003, Takeshi reproduced a required sound image for the specified seat with a sound system consisting of two loudspeakers for Vehicles[3]. In addition, sound systems using more than two loudspeakers are developed to generate surrounding ambience acoustic effects[4][5]. A sound entertainment system for determined positions in a vehicle is proposed by David in 2007. This system provides ultrasonic waves and cancels the unwanted noise[6]. FORD Motor Company invented a multichannel sound reproduction system for vehicles and applied for a patent in 2017[5]. The embodiments mentioned in this patent composed of several loudspeakers, including a low-frequency loudspeaker or sub-woofer, placing in pillars, door frames and vehicle roof. Obviously, the sound reproduction system for entertainment in vehicle are well researched. Lots of patents about sound reproduction in vehicle are applied by vehicle companies to recreate acoustic environment[6][7][8][9]. However, sound spatial perception is far from satisfactory and acoustic virtual reality has not been implemented in vehicle. Three-dimensional (3D) sound reproduction system provides immersive perception about sound sources and thus enhances the sensation of reality[10][11][12]. It is necessary to reproduce 3D sound to enjoy realistic acoustic environment and sound events in vehicle.

The most there popular 3-D sound reproduction algorithms include Wave Field Synthesis(WFS), Ambisonics and Amplitude panning. WFS is based on Huygens principle and it is able to reproduce the whole sound field and thus the

real sound immersion was recreated[13]. However, WFS method is not practical since there are too many loudspeakers required in WFS system. Ambisonics system reproduces the sound pressure at listening point in the center of spherical loudspeaker array. While, it is impossible to configure a spherical loudspeaker array in vehicle. Amplitude panning is a widely used sound reproduction technique due to its computational efficiency. Vector base amplitude panning (VBAP) is a popular sound reproduction technique to render the sound direction and distance. And thus VBAP is considered as a promising technique to recreate sound events[14]. Unfortunately, VBAP system requires a spherical loudspeaker array, which is not satisfied in vehicle.

Though there are several methods to reproduce sound field in physics or mathematics method, the intrinsic goal of sound field reproduction is to reproduce sound pressure at every point in the listening field. Obviously, this is a typical data mining problem. In this paper, a regression algorithm is proposed to reproduce 3D sound in a specific location, the driving seat for example. A sound pressure data set is set up at the specific listening point based on physical sound theory. The sound pressure is related to the number of loudspeakers, the location of loudspeakers, frequency of sound, the distances between listening point and loudspeakers. Since both inputs and outputs are known, a supervised learning model is built to demonstrate the relationship between received sound pressure and the all factors. The mean square error (MSE) shows the proposed regression model predicts sound pressure accurately.

The rest of this paper is organized as follows: the theory foundation, including 3D sound filed reproduction method and regression model, is introduced in the next section. The proposed method to reproduce the 3D sound in vehicle is developed in section (3) experiment is conducted in section (4). And section (5) concludes this paper.

2 Fundamental Theory

2.1 3D sound reproduction

A realistic 3D sound reproduction system needs to reproduce sound pressure at the listening point or every point in a listening area. The Fourier transform of sound pressure generated by a sound source is shown in equation (1) [12][15][16].

$$p(r, \xi) = G \frac{e^{-ik|r-\xi|}}{|r-\xi|} s(\omega) \quad (1)$$

where the parameters are explained as follows:

ξ : $\xi = (\xi_x \ \xi_y \ \xi_z)$ is sound source location;

e : a constant irrational number which is the base of natural logarithm;

i : imaginary unit;

k : wave number, which is equal to $\frac{2\pi f}{c}$;

f : sound frequency;

c : sound velocity;

G : a constant number which represents the proportionality coefficient between sound pressure at a unit distance from a loudspeaker and the input to the loudspeaker;

$s(t)$: sound signals in time domain;

$s(\omega)$: sound signals in frequency domain, i.e. Fourier transformation of sound signal $s(t)$;

$r = (x, y, z)$: listening point.

It is concluded sound pressure is related to several factors after analyzing the sound pressure formula. In a multichannel reproduction system, the computational complexity grows with an exponential trend as the number of loudspeakers increases. However, whatever the 3D sound reproduction methods are, the ultimate goal of the reproduction is to recreate the sound pressure at listening point. As long as the volume of sound pressure data is large enough, data mining technology is an appropriate method to predict sound pressure.

2.2 Regression algorithm

Regression methods, popular techniques to predict system outputs, are well researched by scientists and engineers. Linear regression is one of the most well-known regression algorithms [17][18]. A linear equation is constituted by the product of a constant and a single variable with the power of either one or zero. A typical example of linear equation is shown in equation (2).

$$a_1x_1 + a_2x_2 + \dots + a_nx_n + b = 0 \quad (2)$$

where $a_i, i = 1, 2, \dots, n$ is a coefficient and b is a constant.

if $a_2 = a_3 = \dots = a_n = 0$, in other words, if there is only one variable, equation (2) is the simplest linear equation, which is written as shown in equation (3).

$$ax + b = 0 \quad (3)$$

where a, b are constants.

Linear regression is a linear approach to demonstrate the relationship between inputs $X = (x_1, x_2, \dots, x_n)$ and output Y . In mathematics, the relationship is shown in equation (2.2).

$$y = a_1x_1 + a_2x_2 + \dots + a_nx_n + b \quad (4)$$

In data mining, given the known dataset X and Y , the linear regression is established as shown in equation (2.2).

$$y_i = a_1x_{i1} + a_2x_{i2} + \dots + a_nx_{in} + b_i \quad (5)$$

In practice, matrix form is used to show the relationship, which is shown in equations (2.2) and (2.2).

$$y_i = x_i^T A + b_i \quad (6)$$

$$Y = X^T A + B \quad (7)$$

where $Y = (y_1, y_2, \dots, y_n)$, $B = (b_1, b_2, \dots, b_n)$

gebraic equation in which each term is either a constant or the product of a constant and (the first power of) a single variable with one variable

[19]

[20] abstract = Examining articulatory compensation has been important in understanding how the speech production system is organized, and how it relates to the acoustic and ultimately phonological levels. This paper offers a method that detects articulatory compensation in the acoustic signal, which is based on linear regression modeling of co-variation patterns between acoustic cues. We demonstrate the method on selected acoustic cues for spontaneously produced American English stop consonants. Compensatory patterns of cue variation were observed for voiced stops in some cue pairs, while uniform patterns of cue variation were found for stops as a function of place of articulation or position in the word. Overall, the results suggest that this method can be useful for observing articulatory strategies indirectly from acoustic data and testing hypotheses about the conditions under which articulatory compensation is most likely. [21] abstract=To develop speaker adaptation algorithms for deep neural network (DNN) that are suitable for large-scale online deployment, it is desirable that the adaptation model be represented in a compact form and learned in an unsupervised fashion. In this paper, we propose a novel low-footprint adaptation technique for DNN that adapts the DNN model through node activation functions. The approach introduces slope and bias parameters in the sigmoid activation functions for each speaker, allowing the adaptation model to be stored in a small-sized storage space. We show that this adaptation technique can be formulated in a linear regression fashion, analogous to other speaker adaptation algorithms that apply additional linear transformations to the DNN layers. We further investigate semi-supervised online adaptation by making use of the user click-through data as a supervision signal. The proposed method is evaluated on short message dictation and voice search tasks in supervised, unsupervised, and semi-supervised setups. Compared with the singular value decomposition (SVD) bottleneck adaptation, the proposed adaptation method achieves comparable accuracy improvements with much smaller footprint. [22] abstract=Personalization techniques can be applied to address the subjectivity issue of music emotion recognition, which is important for music information retrieval. However, achieving satisfactory accuracy in personalized music emotion recognition for a user is difficult because it requires an impractically huge amount of annotations from the user. In this paper, we adopt a probabilistic

framework for valence-arousal music emotion modeling and propose an adaptation method based on linear regression to personalize a background model in an online learning fashion. We also incorporate a component-tying strategy to enhance the model flexibility. Comprehensive experiments are conducted to test the performance of the proposed method on three datasets, including a new one created specifically in this work for personalized music emotion recognition. Our results demonstrate the effectiveness of the proposed method.,

3 three dimensional sound reproduction algorithm in Vehicle

4 experiment

5 conclusion

References

1. A. Yamada, "Sound reproduction system for motor vehicle," *Journal of the Acoustical Society of America*, vol. 72, no. 3, pp. 1101–1101, 1982.
2. K. Terai, S. Saiki, K. Murata, K. Satoh, Y. Kumura, Y. Nakama, M. Ogawa, and S. Obata, "Sound reproducing apparatus for use in vehicle," *Journal of the Acoustical Society of America*, vol. 88, no. 5, pp. 2518–2518, 1990.
3. T. Enya, Y. Sato, and I. Aichi, "Sound output apparatus for an automotive vehicle," 2003.
4. D. L. Clark and J. W. Steuber, "Vehicle audio system," *Journal of the Acoustical Society of America*, vol. 104, no. 6, p. 3155, 1998.
5. M. C. A. S. N. M. Orellana, Fernando Mar (Benito Juarez, "Loudspeaker arrangement in a vehicle," Patent 9725047, August, 2017. [Online]. Available: <http://www.freepatentsonline.com/9725047.html>
6. D. S., B. E., D. VallWendell, and C. Johnson, "Audio reception control arrangement and method for a vehicle," *Journal of the Acoustical Society of America*, vol. 155, no. 6, pp. 3151–3156, 2007.
7. J. G. G Simon, "Method and apparatus for control of personal digital media devices using a vehicle audio system," Patent 10/870,424, August, 2005. [Online]. Available: <https://academic.microsoft.com/#/detail/1600054936>
8. M. Vu, B. Boblett, N. Penke, K. Hsieh, and J. Nuxoll, "Vehicle audio system interface," Patent 13/671660, August, 2014. [Online]. Available: <http://www.freepatentsonline.com/y2014/0096003.html>
9. J. S. Gibson, "Vehicle human machine interface with auto-customization," Patent 14/672698, August, 2015. [Online]. Available: <http://www.freepatentsonline.com/y2015/0277735.html>
10. A. Gupta and T. D. Abhayapala, "Three-dimensional sound field reproduction using multiple circular loudspeaker arrays," *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 19, NO. 5, JULY 2011*, vol. 19, no. 5, pp. 1149–1160, July 2011.
11. D. Comminiello, S. Cecchi, M. Scarpiniti, M. Gasparini, L. Romoli, F. Piazza, and A. Uncini, "Intelligent acoustic interfaces with multisensor acquisition for immersive reproduction," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 591–598, AUGUST 2015.

12. M. Zhang, R. Hu, S. Chen, X. Wang, D. Li, and L. Jiang, "Spatial perception reproduction of sound events based on sound property coincidences," in *International conference on Multimedia and Expro*. IEEE, 2015.
13. G. Firtha, P. Fiala, F. Schultz, and S. Spors, "Improved referencing schemes for 2.5d wave field synthesis driving functions," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 5, pp. 1117–1127, MAY 2017.
14. V. Pulkki, "Spatial sound generation and perception by amplitude panning techniques," Ph.D. dissertation, Helsinki University of Technology, 2001.
15. A. Ando, "Conversion of multichannel sound signal maintaining physical properties of sound in reproduced sound field," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 6, pp. 1467–1475, 2011.
16. S. Wang, R. Hu, B. Peng, Y. Yang, and H. Wang, "Sound intensity and particle velocity based three-dimensional panning methods by five loudspeakers," in *IEEE International Conference on Multimedia and Expo (ICME 2013)*, 2013.
17. J. M. Hahne, F. B. Jiang, H. R. Farina, F. C. M.-R. Muller, and L. C. Parra, "Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 269–27, 2014.
18. P. C. Austin and E. W. Steyerberg, "The number of subjects per variable required in linear regression analyses," *Journal of Clinical Epidemiology*, vol. 68, no. 6, pp. 627–636, 2015.
19. C. Lewis-Beck and M. S. Lewis-Beck, *Applied Regression: An Introduction*. SAGE Publications, 2015.
20. A. Khasanova, J. Cole, and M. Hasegawa-Johnson, "Detecting articulatory compensation in acoustic data through linear regression modeling," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 925–929, 1 2014.
21. Y. Zhao, J. Li, J. Xue, and Y. Gong, "Investigating online low-footprint speaker adaptation using generalized linear regression and click-through data," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015, pp. 4310–4314.
22. Y. A. Chen, J. C. Wang, Y. H. Yang, and H. Chen, "Linear regression-based adaptation of music emotion recognition models for personalization," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 2149–2153.