Research on Virtual Reality Sound Effects for High-Speed Train Simulation System

Ruidan Su[1]
1. College of Information Science and Engineering, Northeastern University, Shenyang, China
Email: suruidan@hotmail.com

Kunlin Zhang[3], Weiwei Yan[3], Zhandong Zhao[2], Huaiyu Xu[3*], Tao Wen[1*]
2. Integrated Circuit Applied Software Lab, Software College, Northeastern University, Shenyang, China
3 Service Science Research Center, Shanghai Advanced Research Institute, Chinese Academy of Sciences
* Corresponding author's Email: huaiyu@cs.ucla.edu, wentao@ise.neu.edu.cn

***Abstract*—In this paper, virtual reality sound effects for high-speed train simulation system using wavelet denoising algorithm is studied. As one important part of the high speed train operation system, virtual reality system visualizes most parameter in train operation process. Sound effect system provides sound information to the whole simulation system. It is greatly significance to the whole effect and reality of the simulation. Here we will introduce high-speed train virtual reality sound analysis and acquisition method. Based on wavelet denoising algorithm, we analyze the filtering mechanism and realization of high-speed train virtual reality sound effect. Combining audio files and the type of virtual reality scene, we present a new virtual reality sound model and realization mode; eventually offer a new ideal and flexible solution to high-speed train virtual reality sound effect system.**

***Index Terms*—High-Speed Train; Virtual Reality; Simulation; Wavelet Denoising; Sound Effects; Wavelet Denoising Algorithm**

I. INTRODUCTION

High speed operation system contains virtual reality system, train traction simulation, braking system simulation and etc. High speed train virtual reality system vividly shows train running process, traction and braking control system working process and their effects on the train running. It also reflects the whole high speed train simulation effect and provides a direct platform for efficiency evaluation of the system and for driver's training. The sound effect system is an important branch of the simulation system. [1] In the system, audio files are collected and processed. Loading sound effects according to different scenes is researched, which includes loading different audio files alone and together, controlling the volume, audio frequency modulation and so on. The collecting and processing of audio files is the key part of the system. Finally, this paper provides a solution for virtual reality sound effect system. It consists of the following parts: audio acquisition, wavelet de-noising, virtual reality sound effect modeling and loading sound effect based on DirectX.

II. AUDIO ACQUISITION

Acquisition of sound material is the foundation of sound effect simulation. It is related to the success of sound simulation.[2] The types of sound material in different conditions are first analyzed. Based on this, audio acquisition is performed. Audio analysis is to identify the direct sound material of high-speed train virtual simulation sound and confirm the reason, while audio extraction is the description of sound material requirement.

A. Sound Analysis

In different conditions, train running situations are different. To simulate train running sound effect, all conditions should be considered. Different conditions are divided into two kinds: mixed sounds and single sound. According to the requirement of sound virtual environment such as the position, velocity and other train parameters, single sound is a kind of voice material that can be dynamically dispatched to simulate the operation condition of high-speed train. Mixed sounds are formed by overlapping variant sound according to different conditions. In the sound effect simulation, any actually existing sound should be simulated. However, realization of this could be limited by hardware, technique, reverberation, obstruction and other factors.[3] Simplified model should be used. Considering different train running conditions and convenience of sound collection, only single sound sources are collected. Single sound sources include the followings: sound in carriage, sound of motors when train is accelerating or decelerating, sound outside the train, sound of raining, sound when train pass tunnels, sound when two train encounter and sound when train whistles. The other conditions can be composed by the above seven sound sources. The framework is shown below figure1.

doi:10.4304/jsw.7.8.1919-1922

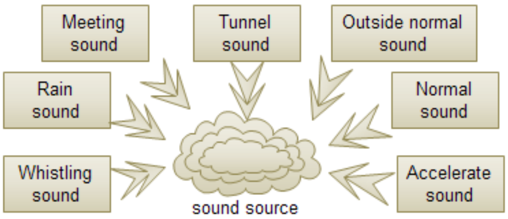

Figure 1. The framework of audio source.

B. Audio Extraction

In order to obtain good sound materials, in the process of collecting sounds in different conditions, the distortion caused by recording equipment should be minimized. To ensure sound quality, it is necessary to adopt high sampling frequency. Sampling frequency should be higher than twice of the maximum frequency of the recording sound to eliminate the distortion. And also earshot of human beings is 20Hz～20KHz. Therefore, sampling frequency should be 20k×2 = 40 KHz to ensure that there is no low frequency distortion.[4] According to this, in the sound effect simulation system, sampling frequency is 44.1HZ, sampling digit is 16Bit and channel is 2. After interception and de-noising artificially, the above seven sound files which should be as short as possible are obtained.

III. Wavelet De-Noising Theory

Only after reliable sound materials are obtained and analyzed effectively, can actual sound effect be simulated. In this part, wavelet de-noising theory and processing of sound files based on this theory are analyzed.

Due to the limit of audio source collection environment and equipment, each single audio inevitably contains noise which is man-made noise or other single audio source's effect. For a specific single audio source, its spectral characters can relatively keep stable for a short time. Because of this property, we can handle its noise and eliminate the waveform whose frequency or amplitude changes greatly, and keep audio stability as much as possible. Therefore, we will use wavelet denoising algorithm for the processing of single audio.

A. Introduction of Wavelet Denoising Theory

Based on the Mallat algorithm of multi-resolution analysis of wavelet transformation, signal S(t) can be decomposed by the following formula[5]:

$$f(t)=\sum_k c_{j,k} 2^{-j/2}\phi(2^{-j}t-k)+\sum_k d_{j,k} 2^{-j/2}\psi(2^{-j}t-k) \quad j,k \in Z \tag{1}$$

Then, decompose the signal according to below formula.

$$s(t)=\sum_k c_{j,k}\phi_{j,k}(t)+\sum_k d_{j,k}\psi_{j,k}(t) \quad j,k \in Z \tag{2}$$

After that, signal S(t) is decomposed into L(t) part whose frequency is below 2j and H(t) part whose frequency is between 2j and 2j+1. Usually, useful signals have low frequency and their low frequency coefficient is Aj, noises have high frequency and their high frequency coefficient is Dj. Useful signals and noises have opposite properties in different scales of wavelet transformation, namely, with the increasing of scale, the wavelet transformation coefficient of useful signals increases while that of noises decreases.[6] Thus, after several wavelet transformations, the wavelet transformation coefficient of noises has been eliminated or very small. Then the coefficient of useful signals remains. Wavelet de-noising schematic is shown in the following figure2.

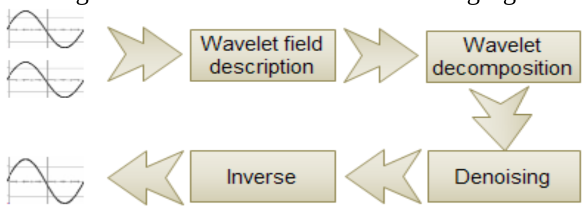

Figure 2. The schematic structure of walevet denosing.

B. Wavelet De-noising Processing

- Obtain wavelet signals.

Since different wavelet bases on time domain and frequency domain have different local performances, the capability for wavelet transformation to characterize local signals is different.[7] Based on wavelet transformation's effectiveness and frequency resolution requirements, db wavelet is adopted to analyze in this paper. To some specific wavelet base, different orders have different abilities to characterize signals' local features. In general, higher orders have stronger abilities, but relatively computational complexity is larger. Considering comprehensively, we set wavelet base to 3. When there are more noises in the audio signals, more times of wavelet transformation are needed, but computational amount are larger.[8] In this paper, the time of wavelet transformation is determined by the amount of noises. Then according to the settled parameters, wavelets of different frequencies can be obtained by decomposing wavelets using approximate coefficients and detail coefficients. After the above process, through carrying out wavelet transformation to single sound fragments with noise, wavelet signals of different frequency bands are obtained. Then characteristic audio and white noises can be separated from the sound fragments.

- Then determine the threshold value of wavelet threshold de-noising method based on Neyman—Pearson principle.
- Filter each wavelet signal according to the threshold value.
- Do the inverse wavelet transformation to each processed wavelet signal to obtain de-noised audio signal.
- Listen to the de-noised audio signals. If the signals have good qualities, then stop de-noising. Otherwise, reselect the threshold value and do the process above again until get a satisfactory result.

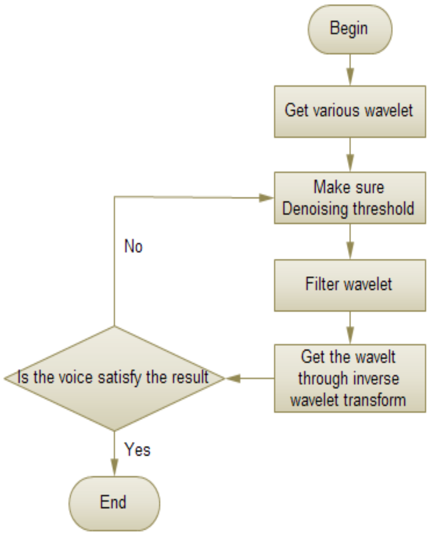

Figure 3. The schematic structure of walevet denosing.

Take the single audio source of normal sound for example, we can decompose noise from it by wavelet donoising processing, and calculate the noise intensity distribution in different frequencies. The breakdown drawing is shown in figure 4.

Similarly, for the single audio source of normal sound, we use "bd" wave function wavelet analysis and take 3 as the wavelet base number. Based on the Fujian value denoising method, we get the result as shown in the figure 5 after 4 times' transformation.

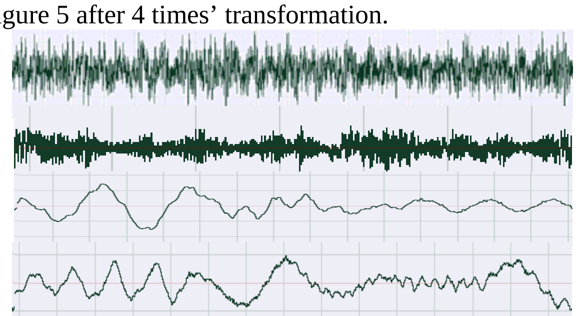

Figure 4. The noise of walevet denosing.

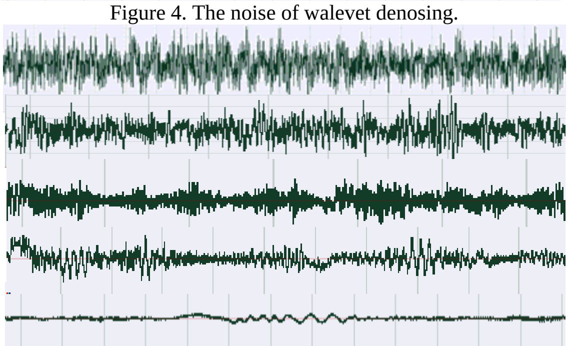

Figure 5. The result of walevet denosing.

IV. Virtual Reality Sound Effect Modeling

Source modeling refers to the analysis of voice signal spectrum and audio files characteristics after the original single audio source collection process. Then, get a different single audio source after denoising and according to the characteristics of train running different position, velocity, etc, superpose and compose the signal audio files. Its work flowchart is as figure6.

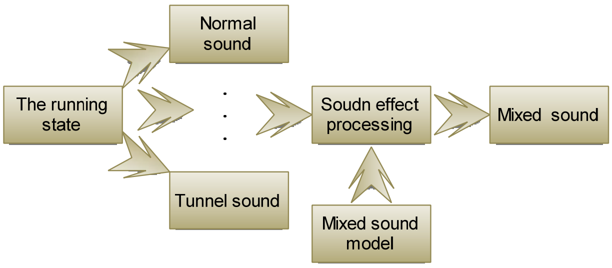

Figure 6. The flowchart of sound effect modeling.

In the train running process, train interacts with airflow, track and other external factors all the time. These factors are closely linked to the train velocity. Thus, it is important to model the relationship of train sound effect and its velocity. Through investigations of information of train running sound effect and its velocity, a basic model between them is preliminarily built:

$$f(v) = \begin{cases} A + \int_0^C \frac{a}{v-b} dv (0 \le V \le C) \\ 100 (C < V) \end{cases}$$

f(v) represents the sound effect volume linked to certain velocity; A represents the initial volume of every condition; C represents the maximum velocity above which the volume no longer increases; a and b are related coefficients in the sound effect model. In certain velocity range, the train running effect volume increases with larger velocity, but the increasing amplitude is decreasing. Meanwhile, in different conditions, related coefficients are not all the same. In the conditions of passing tunnel and train encountering, initial volume, namely initial value of A, is relatively large. We relate audio data in different conditions to the above model and finally discover the relationship between volume and velocity.

There are various kinds of sound in the background sound and the simulation methods of different types of sound are varied in detail. A good sound simulation system should simulate different kinds of train running sound effect in order to enhance people's sense of immersion in simulation.[9] The system should have good ability of audio mixing and can mix different kinds of sound. The sounds of train running are not in order and some are overlapped. The system also can process some special sound effect, such as echoes, reflections and Doppler effects.

As a result, this model takes little resource and has better sound quality, especially while simulate high-speed train virtual reality system mixing effect.

V. Realization Of Sound Loading Based On DirectX

DirectSound is a component of DirectX provided by Microsoft. It provides two specialties for sound programmers: fast development and good controlling. DirectSound provides functions of audio mixing, hardware acceleration and direct access to sound devices. By using it, programmers can realize sound files' loading quickly, composing and playing. In this paper, with DirectX, audio can be loaded dynamically in real time.

A. Create a device object and set its degree of collaboration

B. Create an auxiliary Buffer which can also be called background buffer

Through the created background buffer, object processed sound data can be obtained. DirectSound can mix sounds from several background buffers into the main buffer and then export the sounds to the sound output devices to achieve the effect of audio mixing. There are altogether seven types of single sound sources in the system. Therefore, seven auxiliary Buffers are created so as to acquire the effect of single sounds and mixing sounds.

C. Obtain PCM data

WAV files or other resource data are loaded to the buffers.

D. Load other information to the buffers

Lock method and Unlock method should be used when writing operation is performed to a buffer.

E. Play the sound data in the buffer

Play and control audio data in the buffers through play method and stop method. If several buffers are created, then the audio data in these buffers can be played at the same time.[9] These sounds will be mixed automatically. According to various train running properties and velocity, sounds' volume and frequencies can be loaded and regulated based on the sound effect model, the processing is shown as figure7.

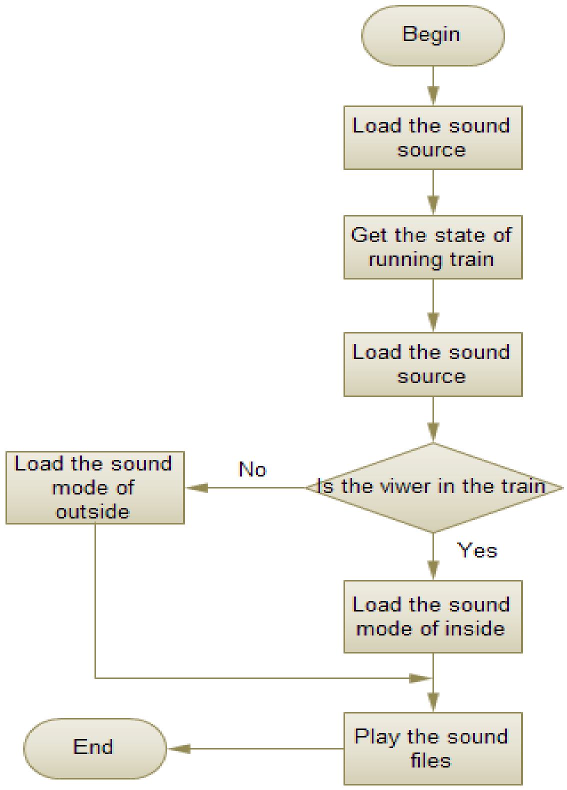

Figure 7. The process of loading sound source.

VI. CONCLUSIONS

In this paper, wavelet de-noising algorithm is researched to realize high speed train virtual reality sound effect system and a new solution to virtual reality sound simulation is provided. At the respect of sound source de-noising, wavelet de-noising algorithm is adopted. This algorithm makes up the weaknesses of Fourier transform de-noising such as time-frequency signal's limitation and non-stationary signal's buffering. This algorithm leads to very good de-noising results. Besides, the modeling of high speed train sound effect in this paper also provides important basis for sound effect system.

Moreover, this paper does not provide a scheduling scheme optimizing solution in the sound loading process and lacks application of sound buffer pool. System resources will be wasted if the quantity of sound files increases.

REFERENCES

[1] e Zhengwei. The Design and Implementation of Real—time 3D Sound Simulation System [J].APPLICATION RESEARCH OF COMPUTERS, 2002, 23-25. (2002).

[2] Meng Shu.The Design and Development of Sound Simulation System of Flight Simulator[D].Beijing:Civil Aviation University of China. (2007).

[3] Luo Fuyuan. Sound Rendering Technology and its Application in Virtual Environment [J].JOURNAL OF SYSTEM SIMULATION, 1999,11(5),364—368. (1999).

[4] Luo fu. Study on sound generation and display technology for virtual environment [D].Beijing:Beijing University of Aeronautics and Astronautics, (1999).

[5] S. Mallat. A wavelet tour of signal processing[M]. Academic Press, USA, (1998).

[6] Tang Yuanyan,Wang Ling.Wavelet Analysis and Text Character Recognition[M], Science Press , (2004).

[7] Wu Shuzhen , Feng Chenglin.methods for speaker-dependent noisy speech recognition [J].ACTA SCIENTIARUM NATURALIUM UNIVERSITATIS PEKINENSIS, 2001, 37(3),365-370 (2001).

[8] S. Mallat.A theory for multiresolution signal decomposition: The wavelet representation[J], IEEE Trans. Pattern Anal. Machine Intell., vol.11, 1989, 674–693. (1989)

[9] Donoho D L.Nolinear Wavelat Methods for Recovery of signals,Densities and Spectraform Indirect and Noisy Data[C].Proceeding of Symposia in Applied Mathematics and Standford Repon (1989)