24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

METHODOLOGY FOR VIRTUAL SOUND SYNTHESIS FOR GRAPHICAL OBJECT INTERACTIONS

Hyunin Jo, Muhammad Imran, Jin Yong Jeon

Department of Architectural Engineering, Hanyang University, Seoul Korea
Correspondence should be addressed to Author Jin Yong Jeon (jyjeon@hanyang.ac.kr)

Modal synthesis procedures are extensively used in virtual reality applications to estimate the vibrations of rigid bodies and to identify the shapes of modes when subjected to an external force. Modal analysis involves the statistics of material parameters and damping factors. The damping functions are significant in virtual sound generation because the duration of oscillations depends on damping, i.e., how long the sound will be perceptible before fading out completely. The computed damping coefficients to be applied for synthesizing the subsequent virtual sounds when rigid bodies come in contact with each other in a virtual environment. In this study, we present a pipeline to generate virtual sounds for 3D object interactions in dynamic virtual environments.

Keywords: Modal sound synthesis, damping models, virtual reality, graphical object interactions.

1. Introduction

In a virtual domain, the coupling of realistic sounds with visuals accompanied by high-resolution graphics can significantly improve the sense of immersion and enhance the users' involvement in interaction with virtual objects [1]. Sound synthesis approaches are mainly classified into two types: a granular synthesis technique and physically based synthesis. The granular synthesis technique is broadly used for synthesizing sounds by the researchers, whereas physically based sound synthesis is relatively new for producing realistic sounds accompanied by visual rendering. Numerous sounds of fluid dynamics with stimulating features have been proposed by Chadwick, Zheng, and James [2]. A rigid body interaction methodology is presented by O'Brien [3]. This methodology simulates rigid body vibrations that lead to variations in perceptual sound pressure. However, this approach is ineffective for handling interactive real-time applications [1].

Recent up-to-date methods to synthesize virtual sounds for graphical object interactions are based on physically based sound synthesis methods. A recently introduced approach that achieves convincing results is based on modal analysis methods. In this approach, material parameters are extracted and subsequently used in a modal synthesis process for the synthesis of sounds. Another approach, known as a Rayleigh approximation, is based on the same methodology and uses linear approximation in the computation of damping functions. However, such techniques are often unable to handle the nonlinear modes of vibrations that create naturalness in the perception of sound.

2. Modal Analysis

It is chiefly focused by modal synthesis method to deliver an accurate simulation of particular realistic entities for their acoustic properties [4]. This method does so by considering each object as a set of many small vibrating substructures. The material properties of the object, along with its shape, are evaluated in order to generate the modes of vibration by decomposing the object vibrations. These vibrational modes are provoked when the object has an encounter with some external force. A unique manner is defined by every normal mode of vibration through which the object can be restructured sinusoidally along periods. These vibrations can be approximately indicated as a linear combination of normal modes with different amplitudes, frequencies, and phases [5]. By using the linear deformation equation, displacement vectors can be considered, as presented in Eq. (1).

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{C}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} = \boldsymbol{f} \qquad (1)$$

Here, M, C, and K are the mass, damping, and stiffness matrices, respectively. Because the surface vibrations during the contact of virtual bodies produce small-scale level damping, an approximation of the damping matrix is considered by a Rayleigh approximation. In other words, a damping matrix can be represented as a linear combination of the mass matrix (M) and stiffness matrix (K), and is given by Equation 2.

$$\boldsymbol{C} = \boldsymbol{\alpha}\boldsymbol{M} + \boldsymbol{\beta}\boldsymbol{K} \qquad (2)$$

The generalized eigenvalue problem given in Equation 3 is solved to decouple the system into the form described in Equation 4.

$$\boldsymbol{K}\boldsymbol{U} = \Lambda\boldsymbol{M}\boldsymbol{U} \qquad (3)$$

$$\ddot{\boldsymbol{q}} + (\boldsymbol{\alpha}\boldsymbol{I} + \boldsymbol{\beta}\Lambda)\dot{\boldsymbol{q}} + \Lambda\boldsymbol{q} = \boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{f} \qquad (4)$$

Here, Λ is a matrix with the eigenvalues of Equation 3. $\boldsymbol{U}$ is the eigenvector matrix that converts $\boldsymbol{x}$ to the decoupled deformation bases q and is represented as $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{q}$. The solutions to Equation 4 are damped sinusoidal waves. The ith mode is given in Equation 5.

$$\boldsymbol{q_i} = \boldsymbol{a_i}\boldsymbol{e}^{-\boldsymbol{d_i}\boldsymbol{t}}\mathbf{sin}(2\boldsymbol{\pi}\boldsymbol{f_i}\boldsymbol{t} + \boldsymbol{\theta_i}) \quad (5)$$

Here, f_i is the frequency, and d_i is the damping coefficient of the ith mode. a_i is the amplitude, and θ_i is the initial phase. In Equation 5, the frequencies, damping, and amplitudes are referred to as the feature φ_i of that mode, which is represented as

$$\boldsymbol{\varphi_i} = (\boldsymbol{f_i}, \boldsymbol{d_i}, \boldsymbol{a_i}) \qquad (6)$$

These features, as expressed in Equation 6, depend on the material properties, geometries, and exciting forces during real-time interactions. The general formulation is given below in Equation 7 and Equation 8.

$$\boldsymbol{d_i} = \frac{(\boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{\lambda_i})}{2} \qquad (7)$$

$$\boldsymbol{\varphi_i} = \frac{1}{2\boldsymbol{\pi}}\sqrt{\boldsymbol{\lambda_i} - \left(\frac{\boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{\lambda_i}}{2}\right)^2} \qquad (8)$$

3. Proposed Methodology

In this study, we designed a virtual sound synthesis pipeline for object collisions in virtual environments, as shown in Figure 1. This section describes the computational framework of the proposed methodology.

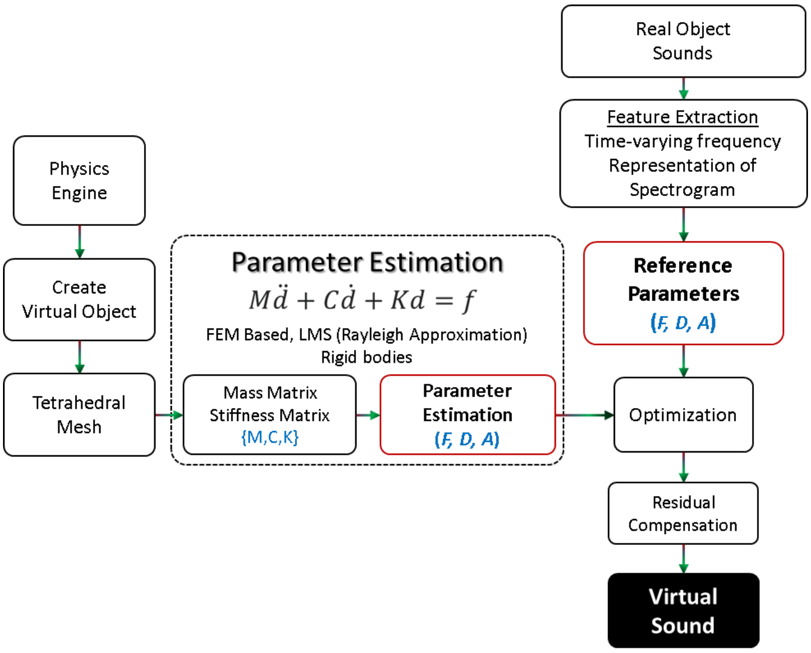

Figure 1. Computational Process and Methodology: Conceptual flow

As a first step, audio clips for different rigid-body impact sounds are recorded in an anechoic chamber hall and subsequently subjected to a feature extraction processing chain. High-level features are extracted with a spectrogram analysis of these sound clips by splitting them into short intervals of overlapping window frames in the time domain. A power spectrogram is obtained by applying a short time Fourier transform (STFT) using a Hanning window function to each frame of the sound clips.

We applied a search algorithm to determine potential peaks in the time-frequency spectrogram. These peaks correspond to the potential modes of the recorded sound in the form of damped sinusoids. The potential modes are filtered out based on the criterion of selecting the strongest peaks for a specified time window. These features are detected in the form of frequencies, damping, and amplitude envelopments $\varphi_i = (f_i, d_i, a_i)$, which are known as extracted features.

Second, a material parameter estimation algorithm is applied based on an optimization process. These parameters are represented as $\varphi_i = (f_i, d_i, a_i)$. To calculate the parameters, we created a virtual object of the same size and geometry as that of a real object from which the features were extracted. This object is tetrahedralized to compute the mass $\boldsymbol{M}$ matrices and stiffness $\boldsymbol{K}$ matrices by providing initial values of Young's modulus $\boldsymbol{E_0}$, Poisson's ratio $\boldsymbol{\nu_0}$, and mass density $\boldsymbol{\rho_0}$. The eigenvalues $\boldsymbol{\lambda_i}$ are computed as follows:

$$\boldsymbol{\lambda_i} = \frac{\boldsymbol{\gamma}}{\boldsymbol{\gamma}_0} \boldsymbol{\lambda}_i^0 \qquad (12)$$

Here, $\gamma = E/\rho$ is Young's Modulus ratio, and $\gamma_0 = E_0/\rho_0$ is the ratio of the initial assumed values of mass and density. A unit impulse is applied to the object that generates excitations of the eigenvalues as given in Equation 5, and is denoted by a_o^j. Finally, the parameters are computed by combining Equations 7, 8, and 9. The sound is synthesized using the estimated parameters in Equation 10, given below:

$$s[n] = \sum_{j} \left(a_i e^{-d_i \left(\frac{n}{F_s}\right)} sin \left(2\pi f_i \left(\frac{n}{F_s} \right) \right) \right) \quad (13)$$

4. Discussions and Future Work

The advantage of this approach is that the properties of recorded audio can be transferred to different objects of different sizes and of various geometries. In the final step, residual compensation is applied in order to compensate the non-harmonic part of the real sound example. A linear modal approach is unable to detect the non-harmonic part during the sound synthesis process; therefore, the quality of the sound remains unnatural. In our method, we followed a technique for residual compensation based on subtracting the representative sound, synthesized from material parameters, from the original sound.

Acknowledgments

This work was supported by the Global Frontier R&D Program on <Human-Centered Interaction for Coexistence> funded by a National Research Foundation of Korea grant funded by the Korean government (MSIP) (2013M3A6A3079356).

REFERENCES

[1] M. Imran and J. Y. Jeon, “Feature Based Impact Sound Synthesis of Rigid Bodies Using Linear Modal Analysis for Virtual Reality Applications,” in *Audio Engineering Society Conference: 61st International Conference: Audio for Games*, 2016.
[2] C. Zheng and D. L. James, “Harmonic fluids,” *ACM Transactions on Graphics (TOG),* vol. 28, p. 37, 2009.
[3] J. F. O’Brien, “Synthesizing sounds from physically based motion,” in *ACM SIGGRAPH 2001*, 2001, p. 59.
[4] Z. Stefano, “Accurate sound synthesis of 3D object collisions in interactive virtual scenarios,” 2012.
[5] R. Corbett, K. van den Doel, J. E. Lloyd, and W. Heidrich, “TimbreFields: 3D interactive sound models for real-time audio,” *Presence Teleoperators Virtual Environ.*, vol. 16, pp. 643–654, 2007.