

## MODELING AND RENDERING FOR VIRTUAL DROPPING SOUND BASED ON PHYSICAL MODEL OF RIGID BODY

*Sota Nishiguchi*

Graduate School of Computer and Information Sciences  
Hosei University  
Tokyo, Japan  
sota.nishiguchi.4b@stu.hosei.ac.jp

*Katunobu Itou*

Faculty of Computer and Information Sciences  
Hosei University  
Tokyo, Japan  
katunobu.itou@hosei.ac.jp

### ABSTRACT

Sound production by means of a physical model for falling objects, which is intended for audio synthesis of immersive contents, is described here. Our approach is a mathematical model to synthesize sound and audio for animation with rigid body simulation. To consider various conditions, a collision model of an object was introduced for vibration and propagation simulation. The generated sound was evaluated by comparing the model output with real sound using numerical criteria and psychoacoustic analysis. Experiments were performed for a variety of objects and floor surfaces, approximately 90% of which were similar to real scenarios. The usefulness of the physical model for audio synthesis in virtual reality was represented in terms of breadth and quality of sound.

### 1. INTRODUCTION

The sound in computer generated (CG) animation is created manually by the sound designer. Some sounds cannot be always realized, for example, the sound of a giant robot or a fictional weapon. Such sounds are prepared by processing or by using synthetic sounds that match the image in every single scene. Experience and the creative sense of the creator are important in the processing and synthesis of good sound. It is inconvenient to create a large number of sounds manually. In immersive content, it is necessary to create sounds to match the user's movement and the situation. However, it is difficult to always synchronize the sound with the video timing and impression. In this paper, the term "impression" indicates whether the sound matches the object and the phenomenon seen in the video. With a physical model, it is possible to generate a good sound from the physical information of the CG generation. In addition, irregular phenomenon, in which it is difficult to match the sound with the image, can be managed using synthesized sound. It can also be used for a fictional phenomenon. This study addresses the physical model for sound synthesis of the sound of a falling object, namely a huge sword, and the virtual dropping sound of the weapon with the aim of creating an automatic sound generation system corresponding to them.

There are two approaches to generate sounds of irregular movements such as dropping of objects. The first method is a statistical model that creates large quantities of transition models and attenuation models based on dropping sounds[1]-[4]. With this method, it is possible to generate dropping sounds that are very close to the real sounds and match the related images. In addition, by creating a model in advance, the generation process is completed in real time. Therefore, real-time sound generation is possible. However, it is not suitable for generating a virtual dropping sound because it is impossible to prepare a real thing. Another approach, which is a physical model, is a method for reproducing object vibration using

physical simulation[5]-[11]. It is possible to generate dynamically natural sound based on rigid body simulation of CG animation and three-dimensional data in this method; it is also possible to generate virtual dropping sounds by appropriately setting the physical model and parameters.

The problem with the physical model is that the simulation cost is enormous and it takes time to generate the required sound. In the previous study[9], it was necessary to limit the frequency of vibration due to the cost of computation. As a result, high-frequency sound could not be generated and the sound had a boxy impression. Because it cannot handle huge objects due to high calculation cost, it can be said that it is insufficient as a synthesizer for virtual dropping sounds. In the above-mentioned research, the natural vibration mode is calculated by the finite element method (FEM) based on the three-dimensional data of the object, and the vibration of the object is precisely reproduced. The vibrations of all shapes such as complicated shapes and objects composed of a plurality of parts can be reproduced, whereas there are modes that do not need to be considered depending on the shape of the object; therefore, extra calculations are assumed to occur.

In this paper, we restrict the shape of the object to a bar, and generate a falling sound using a simple model that considers only sounds that can be heard. When limited to bars, we assume that the sonic vibration is only bending vibration. We simulate the vibration with a bending vibration model with reduced lattice points and dimensions when compared with the conventional method. Because the torsional vibration of the rod is smaller in amplitude than the bending vibration and the stretching vibration is an ultrasonic wave in the audible range or higher, it can be anticipated that the sound does not change considerably even when using a simple model with only bending vibration. We will construct a system that reduces the computational complexity while maintaining the quality of sound, and which can be used to generate the dropping sounds of huge objects.

### 2. PHYSICAL MODEL FOR DROPPING SOUND SYNTHESIS

#### 2.1. Various dropping sounds

Various situations can be considered for the dropping phenomenon. The phenomenon also changes if the shapes and materials of the objects are different. In addition, the movement undergoes complex changes depending on the manner of dropping. Due to these factors, various dropping sounds having different tone pitch, volume, tone color, attenuation, and timing are generated. Particularly, it is difficult to control the movement at the time of fall; therefore, it is very difficult to retrofit or prepare in advance a

falling sound that matches the image. Figure 1 shows the spectrograms of sounds when the same object is dropped on different floor surfaces. The main physical phenomena that emit sounds are the collision between the floor surface and the object and the vibration of the object due to the collision.

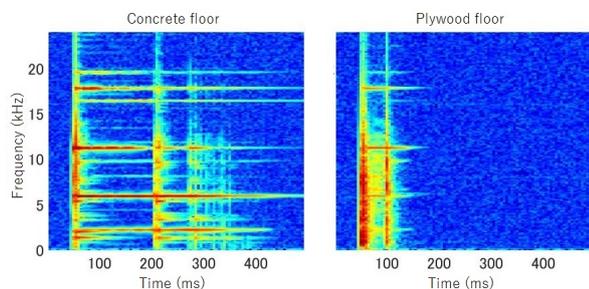


Figure 1: Spectrograms of sounds of a falling aluminum rod. The rod is a round bar having a length of 15 cm and a diameter of 1 cm.

Collision sounds are represented as pulse shock waves that vary depending on the physical properties of object, the floor surface, and the collision speed. A vibration sound is represented by a periodic function, which is the sum of a number of eigenmodes. There are three types of eigenmodes: bending vibration (transverse wave), stretching vibration (longitudinal wave), and torsional vibration. The intensity and attenuation of each eigenmode change depending on the position and intensity of the external force. These phenomena are repeated by the rebounding of the dropping object. Because speed and attitude of the object vary depending on each bounce, the collision sound and vibration sound also change accordingly. With recorded samples and general sound effects, it is difficult to reproduce dropping sounds that variously change depending on conditions. In a virtual space, objects are grasped and moved with high degree of freedom. Therefore, an acoustic generation engine without restrictions on movements is necessary.

## 2.2. Related researches

Sound source generation using physical equations and physical parameters has been studied for phenomena accompanying irregular movements such as the sound of fluids like water, the rubbing sound of clothes, and the sound of a flame [12]-[14]. There are also studies on the dropping sound of objects. Sound sources suitable for individual objects are generated by calculating the eigenmodes of falling objects by eigenvalue analysis[5]-[11].

As these studies use a three-dimensional vibration model, the amount of calculation is huge. Therefore, it is necessary to set the upper limit of the frequency, whereupon the generated sound becomes a muffled impression. Furthermore, it cannot handle huge objects that make the number of nodes extremely large. Three-dimensional vibration analysis can reproduce all eigenmodes bending, stretching, and torsional vibrations. However, depending on the shape of the object, there are eigenmodes whose frequencies are outside the audible range; therefore, it is necessary to select the vibration model that is most suitable for sound generation.

In addition, the initial condition of vibration is focused only on simple impulse excitation. Depending on the difference in the

floor surface and the collision speed, the force applied to the object at the time of collision changes differently, and it is expected that it will also affect the subsequent vibration. To practically apply the physical model, it is necessary to consider a dropping sound generation model that can cope with a greater variety of phenomena and can ensure good sound quality.

## 3. PROPOSED METHOD

### 3.1. Overview of dropping sound generation system

The sound generation system consists of two processes. The first process is sound modeling, which generates sound sources using physical information and physical models. Another process is sound rendering, which generates an acoustic field based on the generated sound source and the spatial information. A desired sound is generated by the arrival sounds based on the sound field obtained through these processes and the observation point.

Physical information and spatial information are input to the system. In this system, the numerical values related to physical properties, shape, speed of the object, and the observation point are set. The sound generation is performed according to the phenomena in the image by quoting the shape and behavior data from the physical engine of the CG animation. Various parameters are assigned to the physical model of collision and vibration to generate the collision waveform and the vibration waveform. The vibration model is a bending model in which an object is regarded as a rod. The sound that reaches the ears is generated through sound rendering of these waveforms. In this paper, we synthesize the sound of a dropped object by solving the wave equation with the collision waveform and vibration waveform as the boundary conditions.

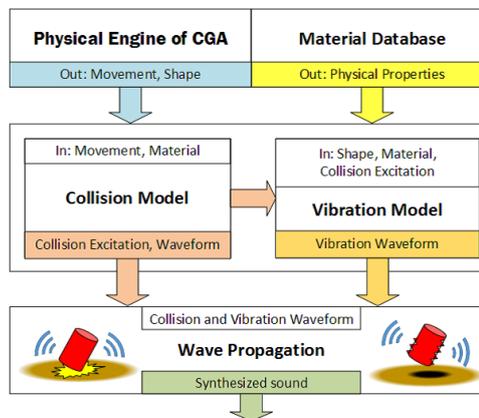


Figure 2: Overview of a dropping sound generation system

### 3.2. Bending vibration model

The Euler–Bernoulli beam is a model expressing the vibration of a bar only by the deflection deformation. Because of the bending deformation, the inside of the curve shrinks and the outside elongates in the axial direction, and the restoring force is created. For each material, the ratio of the restoring force to the deformation is given as the elastic modulus or Young’s modulus  $E$ . The fundamental equation of the Euler–Bernoulli beam is obtained from

the elastic curve equation and the equation of motion. The elastic curve equation, which expresses the displacement when a bar receives an external force, is given as follows.

$$M = -EI \frac{\partial^2 w}{\partial x^2} \quad (1)$$

$M$  is the bending moment. Let  $w(x, t)$  be the displacement at position  $x$  and time  $t$ . Further,  $I$  is the moment of inertia of the area,  $\rho$  is the mass density of the object, and  $A$  is the cross-sectional area. The relationship between the bending moment  $M$  and the stress  $V$  applied to the object can be expressed by the following equation.

$$\frac{dM(x)}{dx} = V(x) \quad (2)$$

By applying this to the equation of motion, we obtain the fundamental equation of the Euler–Bernoulli beam. Considering the small section  $\Delta x$  of the bar, the external force is  $F = \Delta x \frac{dV(x)}{dx} = \Delta x \frac{d^2 M(x)}{dx^2}$  and weight is  $\mu \Delta x$  ( $\mu$  is the linear density of the bar); the acceleration can be expressed as  $d^2 w/dt^2$ . Therefore, the equation of motion related to the bending of the bar is as follows.

$$\Delta x \frac{\partial^2 M(x)}{\partial x^2} = \mu \Delta x \frac{\partial^2 w}{\partial t^2} \quad (3)$$

By deleting  $\Delta x$  on both sides and substituting the expression (1) for  $M$  on the left side, the following equation is obtained.

$$\frac{\partial^2 w}{\partial t^2} + \frac{EI}{\mu} \frac{\partial^4 w}{\partial x^4} = 0 \quad (4)$$

The Euler–Bernoulli beam is a model that considers only the bending deformation of the object, so it is strictly different from the actual vibration. Therefore, when handling a short bar with a small slenderness ratio of the object, the frequency of vibration calculated by the above equation becomes higher than the actual value. This is an error, which occurs because the rotational inertia and the shear deformation of the entire bar are not considered. It is easier to rotate the entire object with the shorter bars, which increases the rotational inertia. Moreover, when the ratio of the cross section to the length increases in the short bar, the ratio of shear deformation to the bending deformation also increases and cannot be ignored. The Timoshenko beam is a model that considers these effects, and is therefore used in this paper.

The Timoshenko beam expresses the vibration of a bar by flexural deformation and shear deformation. Shear deformation is a displacement in the cross-sectional direction, and a restoring force against the displacement occurs. The elastic modulus is expressed by the rigidity rate  $G$ . Rotational inertia is applied to the Timoshenko beam to derive the basic equation.

$$\begin{aligned} \frac{\partial}{\partial x} \left[ AG\kappa \left( \frac{\partial w}{\partial x} - \phi \right) \right] &= \rho A \frac{\partial^2 w}{\partial t^2} \\ AG\kappa \left( \frac{\partial w}{\partial x} - \phi \right) + EI \frac{\partial^2 \phi}{\partial x^2} &= \rho I \frac{\partial^2 \phi}{\partial t^2} \end{aligned} \quad (5)$$

Let  $\phi(x, t)$  be the rotation angle of the section at position  $x$  and time  $t$ , and let the Timoshenko coefficient be  $\kappa$ .

Next, these partial differential equations are solved numerically by calculus of finite differences. The differential terms are substituted with the central difference of the second-order precision with respect to both temporal and spatial derivatives of the fundamental equation.

$$\begin{aligned} \frac{G\kappa}{\rho} \left( \frac{w_j^{i+1} - 2w_j^i + w_j^{i-1}}{\Delta x^2} - \frac{\phi_j^{i+1} - \phi_j^{i-1}}{2\Delta x} \right) &= \frac{w_{j+1}^i - 2w_j^i + w_{j-1}^i}{\Delta t^2} \\ \frac{AG\kappa}{\rho I} \left( \frac{w_j^{i+1} - w_j^{i-1}}{2\Delta x} - \phi_j^i \right) + \frac{E}{\rho} \frac{\phi_j^{i+1} - 2\phi_j^i + \phi_j^{i-1}}{\Delta x^2} &= \frac{\phi_{j+1}^i - 2\phi_j^i + \phi_{j-1}^i}{\Delta t^2} \end{aligned} \quad (6)$$

$i$  is an index on space, and let  $w_j^i$  be the displacement at time step  $j$ .  $\Delta x, \Delta t$  are discrete widths of space and time. In order to ensure the stability of the calculation, the set value of the discrete width needs to satisfy the following Courant–Friedrichs–Lewy(CFL) condition.

$$\begin{aligned} |\alpha| < 1, \quad \alpha &= \frac{G\kappa}{\rho} \left( \frac{\Delta t^2}{\Delta x^2} - \frac{\Delta t^2}{2\Delta x} \right) \\ |\beta| < 1, \quad \beta &= \frac{AG\kappa}{\rho I} \left( \frac{\Delta t^2}{2\Delta x} - \Delta t^2 \right) + \frac{E}{\rho} \frac{\Delta t^2}{\Delta x^2} \end{aligned} \quad (7)$$

Because the propagation speed of the elastic vibration is different depending on the physical property data, the discrete width is set so that  $\alpha, \beta$  become sufficiently small for any physical property.

The oscillation of the object is simulated by time evolution of the differentiated basic equation. The free edge boundary condition is used, as the object is free at the time of vibration after the collision. The free edge boundary condition in the Timoshenko beam model is given by the following equation.

$$\left. \frac{\partial \phi}{\partial x} \right|_{x=0,l} = 0, \quad \left[ \frac{\partial w}{\partial x} - \phi \right]_{x=0,l} = 0 \quad (8)$$

The derivative terms in the boundary condition are also substituted with the difference formulas.

$$\begin{aligned} \frac{\phi_j^2 - \phi_j^0}{2\Delta x} = 0, \quad \frac{w_j^2 - w_j^0}{2\Delta x} - \phi_j^1 &= 0 \\ \frac{\phi_j^N - \phi_j^{N-2}}{2\Delta x} = 0, \quad \frac{w_j^N - w_j^{N-2}}{2\Delta x} - \phi_j^{N-1} &= 0 \end{aligned} \quad (9)$$

$N$  is the end of the spatial index, and the elements of the indices 0 and  $N$  are dummy elements for the free boundary condition. For each calculation process of the explicit method, we update the displacement and rotation angle at both ends of the object using these boundary conditions.

Next, the vibration attenuation model will be explained. Rayleigh damping is used in this system. Rayleigh attenuation is a model that takes into consideration two types of attenuation: external damping (viscous damping) and internal damping (viscoelastic damping of the object). The following equation is obtained by adding the attenuation term to the Euler–Bernoulli beam in the expression

$$\frac{\partial^2 w}{\partial t^2} + \frac{EI}{\mu} \frac{\partial^4 w}{\partial x^4} + \eta \frac{\partial}{\partial t} \left( \frac{EI}{\mu} \frac{\partial^4 w}{\partial x^4} \right) - \gamma \rho A \frac{\partial w}{\partial t} = 0 \quad (10)$$

$\gamma$  and  $\eta$  are coefficients for external attenuation and internal attenuation. As approximate values of these attenuation coefficients can be defined for each material, a database of attenuation

coefficients is prepared along with physical property values such as density and Young’s modulus.

Similarly, considering the attenuation in the Timoshenko beam, the fundamental equation becomes as follows.

$$\begin{aligned} \frac{\partial}{\partial x} \left[ AG\kappa \left( 1 + \eta \frac{\partial}{\partial t} \right) \left( \frac{\partial w}{\partial x} - \phi \right) \right] - \gamma \rho A \frac{\partial w}{\partial t} &= \rho A \frac{\partial^2 w}{\partial t^2} \\ AG\kappa \left( 1 + \eta \frac{\partial}{\partial t} \right) \left( \frac{\partial w}{\partial x} - \phi \right) + EI \left( 1 + \eta \frac{\partial}{\partial t} \right) \frac{\partial^2 \phi}{\partial x^2} &= \rho I \frac{\partial^2 \phi}{\partial t^2} \end{aligned} \quad (11)$$

The relationship between  $\gamma$  and  $\eta$  and the loss factor  $\xi$  is expressed by the following equation.

$$\xi = \omega \gamma + \frac{\eta}{\omega} \quad (12)$$

$\omega$  is the angular frequency of vibration and the loss factor has frequency characteristics. The external attenuation shows a characteristic that is inversely proportional to the frequency, and the internal attenuation has a characteristic proportional to the frequency. In this study, these attenuation coefficients are cited from [10].

### 3.3. Collision model

Impact noise is the change in air pressure exerted by the deformation of the object occurring during the action of the impact force. The literature [16] is a study targeting collision sounds of steel balls. A steel ball with a diameter of 5 cm is used in the experiment, but in this case, the fundamental mode of the vibration sound is beyond the audible range and analysis is performed by considering only the impact sound. According to the literature [16], the impulsive sound is a pulse-like waveform, and the peak sound pressure of the pulse is proportional to the acceleration and volume of the object and is inversely proportional to the distance.

$$p(x, y, z; t) = \frac{\rho a^2}{4R} \frac{\partial}{\partial t} \left\{ U \left( x', y', z'; t - \frac{R}{c} \right) \right\} \quad (13)$$

$p$  is the sound pressure at time  $t$  at the observation point  $(x, y, z)$ , and  $U$  represents the velocity at the point  $(x', y', z')$  of the sound source. In addition, let  $a$  be the radius of the sphere,  $R$  be the distance between the sound source and the observation point,  $\rho$  be the mean density of air, and  $c$  be the sound velocity.

The object vibrates due to the collision. Therefore, the force applied to the object by the impact is calculated and used as the boundary condition of the vibration model. The force applied to the object at the time of collision can be predicted from Hertz’s solid contact theory. The deformation  $d(t)$  due to collision when the collision surface is spherical is obtained by the following formula.

$$d(t) = F(t)^{2/3} \left( \frac{C^2}{R} \right)^{1/3} \quad (14)$$

$F(t)$  is the force working at time  $t$  and  $R$  is the radius of the object.  $C$  is defined as follows.

$$C = \frac{3}{4} \left( \frac{1 - \nu_0^2}{E_0} + \frac{1 - \nu_1^2}{E_1} \right) \quad (15)$$

Let  $E_0$  be the Young’s modulus of the floor and  $\nu_0$  be Poisson’s ratio of the floor. The constant with subscript 1 is the corresponding value of the object. The contact time  $\tau$  can also be calculated from these parameters.

$$\begin{aligned} \tau &= \frac{4\sqrt{\pi}\Gamma(2/5)}{5\Gamma(9/10)} \left( \frac{m_r^2}{g^2 v_i} \right)^{1/5} \\ g &= \frac{4}{5C} \sqrt{R}, \quad m_r = \frac{m_0 m_1}{m_0 + m_1} \end{aligned} \quad (16)$$

$m_r$  is the relative mass of the object and the floor surface,  $m_0$  is the mass of the floor, and  $m_1$  is the mass of the object. The following equation is obtained from the relationship between the momentum of the falling object  $p = m_1 v_i$  and the excitation force  $F(t)$ .

$$F_{\text{ave}} = \frac{2m_1 v_i}{\tau} \quad (17)$$

Using this  $F_{\text{ave}}$ , the time waveform of the collision excitation force is modeled as a sine wave.

$$F(t) = F_{\text{ave}} \left( 1 - \cos \left( \frac{2\pi t}{\tau} \right) \right) \quad (18)$$

The vibration corresponding to the change of collision can be generated by giving the time waveform of the external force as the boundary condition for vibration simulation. Even when the object has a shape other than a sphere, the external force waveforms can be calculated using appropriate models [17]. The pulse sound of the collision itself is reproduced by giving the time history  $d(t)$  of deformation due to collision as the boundary condition of the sound propagation simulation.

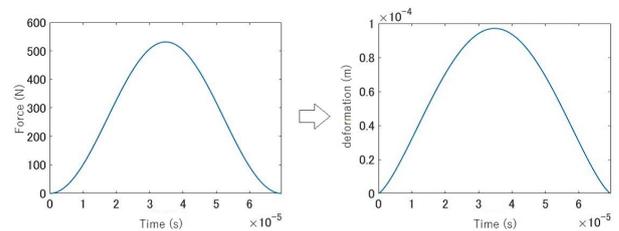


Figure 3: Force history  $F(t)$  and coordinate history  $d(t)$

### 3.4. Dropping rigid body simulation

The vibration model and the collision excitation model are applied to the dropping of the object. The collision speed can be obtained with  $\sqrt{2gh}$  using the object’s height  $h$  and the gravitational acceleration  $g$ . Next, using the coefficient of restitution of the floor and the object, the height after reflection is obtained and the speed at the time of re-impact is calculated. By repeating this, it is possible to obtain the vibration condition for the object bouncing from the floor surface. However, the coefficient of restitution is a numerical value involving complicated physical properties and shapes, and it is difficult to prepare a physical property such as density or Young’s modulus for each material. Moreover, at the time of falling, the object rotates about its center of gravity; therefore, repulsion breaking processing is required.

In the generation of CG animation, the movements of a rigid body are reproduced by performing arithmetic based on the physical law so as to express falling and collision with more realistic behaviors. Bullet is a free physics engine, which is used in many 3DCG creation software (Maya, Blender, etc.). In order to obtain a sound consistent with the image, it is appropriate to use rigid body simulation information contained in the image. In this paper, we use rigid body simulation with Bullet. In Bullet, it is possible to set the coefficient of restitution and the friction coefficient for each object. Using these values as parameters for sound generation, we obtain the collision timing, velocity, and shape of the colliding surface from the rigid body simulation in Bullet.

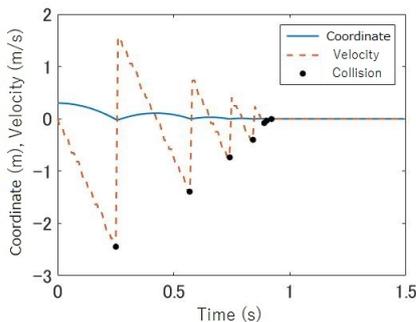


Figure 4: *Coordinate, velocity, and collision timing of an object obtained from Bullet*

### 3.5. Sound rendering

Acoustic processing based on the positional relationship between the sound source and the observation point is required to obtain the result of the vibration simulation as a sound. In the previous study [9], the amplitude of each mode was calculated from the positional relationship between the sound source and the observation point by FFAT(Far-Field Acoustic Transfer) map. An FFAT map is a model of the sound field around the object based on the Helmholtz equation, which is obtained by calculating the phase and amplitude for each mode by eigenmode analysis.

$$\nabla^2 p(\mathbf{x}) + k^2 p(\mathbf{x}) = 0 \quad (19)$$

$p(\mathbf{x})$  is the sound pressure at position  $\mathbf{x}$ , and the wavenumber is  $k = \omega/c$  (sound speed  $c$ , each frequency  $\omega$ ).

With this differential equation, the sound field is estimated by assigning the displacement of the object surface obtained from sound modeling as the Neumann boundary condition. A sound field is created for each vibration mode, and the sound fields of all modes are synthesized to generate the sound field of the vibration sound.

In this research, sound wave propagation is reproduced by directly solving the wave equation, which is the basis of the Helmholtz equation in the FDTD method. The wave equation is given by the following equation.

$$\nabla^2 p(\mathbf{x}, t) + \frac{1}{c^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} = 0 \quad (20)$$

The boundary condition based on the vibration of the object is as follows.

$$\frac{\partial p(\mathbf{x}, t)}{\partial n} = -\rho_m a_n(\mathbf{x}, t), \quad \mathbf{x} \in \Omega \quad (21)$$

Let  $\partial/\partial n$  be the normal derivative of the object surface  $\Omega$ .  $\rho_m$  is the density of the medium. Air is usually the medium, and air vibration on the object surface can be obtained by using  $\rho_{air} = 1.2041 \text{ kg/m}^3$ .  $a_n$  is the acceleration of the surface of the object and can be obtained by differentiating the second-order displacement of the object surface with respect to time. The boundary conditions given as differential equations are differentiated, and conditional expressions for the sound pressure at the boundary between the object and the medium are obtained. We simulate sound propagation using this condition and the wave equation (Fig. 5).

The recording environment is reproduced for the evaluation of the generated sound. Because the recording environment is an anechoic room, a perfectly matched absorption boundary layer is set around the calculation area so that reflection of waves from the wall do not occur [18].

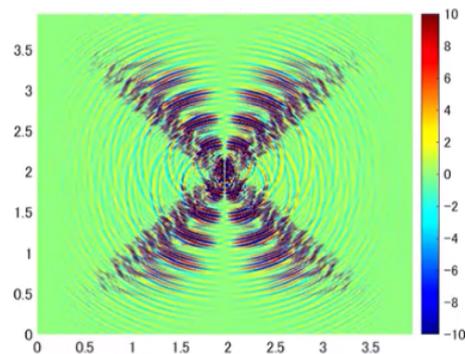


Figure 5: *The generated spatial acoustic field*

## 4. EVALUATION

In order to evaluate the bending vibration model, we simulated the vibration of the object and compared it with the recorded vibration sound. Next, dropping sounds were generated under various conditions, and spectrograms were compared with the actual sound sources. The shape and physical property parameters of the targeted object were set as shown in Tables 1 and 2.

### 4.1. Evaluation of vibration sound

For object A in Table 2, vibration was simulated with shock applied to the end of the rod as the initial condition. The air vibration was simulated at a position 5 cm away from the end of the rod in the direction perpendicular to the axis. It was found that although there was a numerical error from the actual sound, a sound with no incongruity in terms of the sound impression and height was generated. Figure 6 compares the spectrums of the actual sound and the generated sound. It was understood that the natural frequencies were roughly coincident, and the amplitude of the mode component was almost faithfully reproduced. For eigenmodes above 15,000 Hz, the actual sound and intensity were different and could not be reproduced satisfactorily, but because the frequency was

Table 1: Physical property parameters of the material

Material	Density (kg/m <sup>2</sup> )	Young's modulus (GPa)	Rigidity ratio (GPa)	Poisson ratio	Internal damping	External damping
Aluminum	2698.9	70.3	26.1	0.345	3E-8	5
Brass	8411	100.6	37.3	0.35	3E-8	5
Iron	7874	211.4	81.6	0.293	4E-8	0.1
Wood	800	11	4.23	0.3	2E-6	60

Table 2: Shape parameters

Object name	Material	Length(m)	Width(m)	Thickness(m)	Cross section shape
Object A	aluminum	0.15	0.01	0.01	circle
Object B	aluminum	0.15	0.01	0.01	rectangle
Object C	brass	0.15	0.01	0.01	rectangle
Object D	iron	0.30	0.008	0.008	circle
Object E	wood	0.15	0.005	0.005	rectangle
Object F	iron	10	0.3	0.07	rectangle

close to the upper limit of the audible range, it was considered that there was no great influence on the sound impression; therefore, it was not evaluated this time.

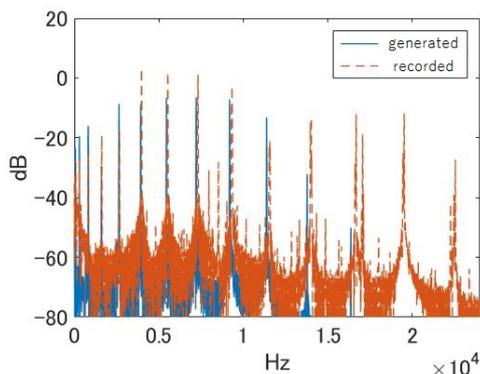


Figure 6: Vibration sound spectrum of object A. Comparison of synthesized and recorded sounds.

#### 4.2. Evaluation of dropping sound

The sound of the falling object was evaluated for the objects B to E. We set various parameters and generated the dropping sounds. The object B had a metallic lightweight sound impression. The object C was heavier and softer than B, and the object D had a hard metallic sound. Natural dropping sounds were generated for aluminum, brass, and iron bars. Compared with the actual dropping sound, the height of the vibration sound of each object could almost be reproduced, and the vibration corresponding to the difference in material and shape could be generated. Furthermore, the object E which was a wooden rod, had a fast decaying dry sound, which was close to the impression of the actual sound. By defining the damping coefficient for each material, a natural falling

Table 3: Comparison of natural frequencies.

Mode No.	Recorded	Generated	Relative error
1	303 (Hz)	304 (Hz)	0.33%
2	831	823	0.96%
3	1627	1606	1.29%
4	2675	2643	1.20%
5	3979	3928	1.28%
6	5532	5454	1.41%
7	7319	7211	1.48%
8	9330	9191	1.49%
9	11,568	11384	1.59%
10	14,006	13,779	1.62%
Average	—	—	1.26%

sound was obtained even for an object made of material with great difference in hardness and weight.

The objective evaluation of each generated sound is as follows. The recall ratio of the eigenmode is obtained by dividing the matching modes of the recorded sound and generated sound by the number of all modes. The relative error of the frequency is used as the condition for the match. A mode in which the relative error was less than 6% was regarded as the matching mode. The relative error of 6% was approximately the same as the chromatic scale, which was the minimum unit of the pitch.

We also generated the virtual falling sound of a huge sword using the system. The parameters correspond to object F in Table 2. The shape (the length, width and thickness) of a general Japanese sword was measured as a square bar and the value was magnified by 10. An image was created using the same physical parameters. A heavy metal sound was generated according to the movement of the drop, and a virtual falling sound with an impression suitable for the image was obtained.

Table 4: Evaluation data

Object	Generated mode	Recall ratio	Frequency relative error	Average power error
Object B	4/6	66.7% (4/6)	3.43%	5.65 dB
Object C	3/3	100% (3/3)	1.52%	17.12 dB
Object D	11/13	69.2% (9/13)	2.66%	12.15 dB
Object E	5/8	62.5% (5/8)	4.11%	8.16 dB

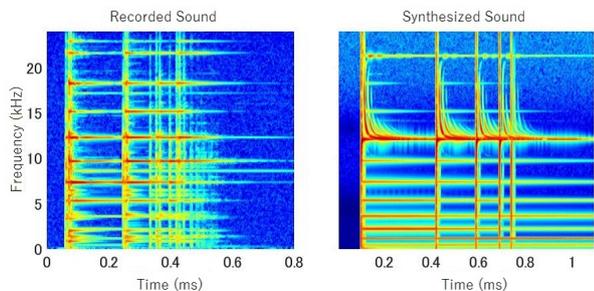


Figure 7: Dropping sound spectrogram of object D. Comparison of recorded and synthesized sounds.

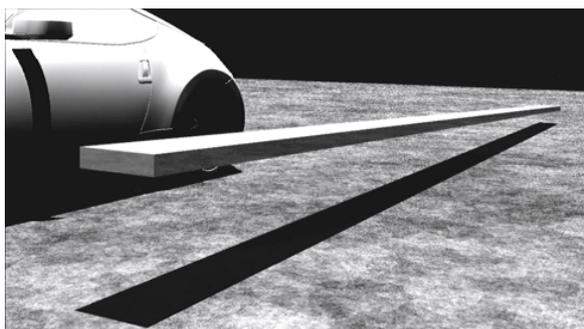


Figure 8: A huge falling sword simulated by Ballet

### 4.3. Discussion

For bars with a circular cross section, sufficient vibration sound reproduction is possible even with models with only flexural vibration. This is because the torsional vibration does not generate sound waves in a circular cross section. However, for bars with rectangular cross section, a strong torsional vibration mode occurred depending on the collision position. In the model with only bending vibration, a monotonous impression sound was generated rather than the actual vibration sound due to the lack of torsional vibration mode.

It is presumed that the power error is due to the mismatch between the attenuation and error of the excitation condition. Vibration suitable for the sound made by bars of various metals and wood was obtained by setting the damping coefficient for each material. However, because we did not consider energy absorption from the object to the floor, we could not sufficiently reproduce the drop to the soft floor surface. The floor, which was thin and easily vibrated, was not reproduced because of the same reason. We be-

lieve that the exchange of energy can be applied to the generation of the sound that causes the floor surface to vibrate, reproducing the rapid attenuation by the contact with the floor after the end of the bouncing phase.

By using the collision excitation force waveform for the vibration sound, the higher order mode of the vibration sound gradually weakened each time it bounced back the characteristics of the dropping sound. Moreover, by considering the collision sound caused by the collision deformation, strong feeling of attack on the falling sound was born. However, the spectrum did not change much in any of the generated collision sounds. In this system, the point-to-point collision model was applied to all collisions. The presumed reason is that it was not possible to reproduce the collision from line to point, line to line, and more. In addition, a strong collision generates a shock wave. There may also be a phenomenon wherein the falling object could not fully cope with only by atmospheric pressure change due to deformation and sound propagation.

The excessive attenuation of the component of 15,000 Hz or more of the generated sound is caused by the numerical dispersion. Numerical dispersion is the dispersion occurring due to change in the phase velocity depending on the wave number in the numerical solution. Actually, the phase velocity is constant irrespective of the wave number. As the wave number increases, the numerical dispersion increases. We consider that the components above 15,000 Hz could not be properly simulated with the mesh width used in this simulation. To suppress the numerical dispersion, it is necessary to set a discrete scheme where the CFL condition is sufficiently satisfied.

The simulation of sound wave propagation in a two-dimensional space with nothing around the object was performed for simplicity, but faithful sound generation is possible by further propagating sound waves in three dimensions considering shields and other objects. However, the amount of computation required is proportional to the power of the number of dimensions. A simulation can require extensive computations in three dimensions. In this paper, we adopted FDTD for solving the original partial differential equation for both vibration and propagation directly to clarify the relationship between the models. With regard to the amount of computation, it is considered that it is effective to use a radiation model [19] to simulate the propagation.

It is expected that sounds can be improved by implementing the attenuation due to absorption at the contact points and defining the attenuation rate of the vibration model by physical property parameters. Although we have implemented only the flexural vibration model this time, it is necessary to consider a framework to apply the optimal vibration model as compared with the model considering torsional vibration and stretching vibration.

## 5. CONCLUSION

We have proposed a simplified model of vibration in this study. As a result, it was found that the round bars were almost reproducible only by the bending vibration model. We could reduce the computation for round bars while maintaining the quality of the generated sound. A virtual falling sound of a huge object was generated from the physical model, and an appropriate sound was obtained. By considering sound waves caused by collision deformation, an impact effect was imparted to the dropping sound, leading to a more natural dropping sound generation.

It is necessary to conduct experiments with various parameters to confirm the versatility of the created system. During the evaluation, the generated sound was played along with the image, and the subjective evaluation of the degree of coincidence with the image and the sound quality were important criteria in the evaluation. To realize realistic sound generation, it is important to solve the problems inside the system such as expansion of the vibration model and collision model, use of shock wave propagation simulation, study of the discrete width of simulation, and the like.

## 6. REFERENCES

- [1] Langlois, Timothy R., and Doug L. James. "Inverse-foley animation: Synchronizing rigid-body motions to sound." *ACM Transactions on Graphics (TOG)* 33.4 (2014): 41.
- [2] Zheng, Changxi, and Doug L. James. "Rigid-body fracture sound with precomputed soundbanks." *ACM Transactions on Graphics (TOG)*. Vol. 29. No. 4. ACM, 2010.
- [3] Chadwick, Jeffrey N., Changxi Zheng, and Doug L. James. "Faster acceleration noise for multibody animations using precomputed soundbanks." *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Eurographics Association, 2012.
- [4] Ren, Zhimin, Hengchin Yeh, and Ming C. Lin. "Synthesizing contact sounds between textured models." *Virtual Reality Conference (VR)*, 2010 IEEE. IEEE, 2010.
- [5] James F. O'Brien, Perry R. Cook, Georg Essl, "Synthesizing Sounds from Physically Based Motion," *ACM SIGGRAPH*, 2001.
- [6] Van Den Doel, Kees, Paul G. Kry, and Dinesh K. Pai. "FoleyAutomatic: physically-based sound effects for interactive simulation and animation." *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, 2001.
- [7] O'Brien, James F., Chen Shen, and Christine M. Gatchalian. "Synthesizing sounds from rigid-body simulations." *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*. ACM, 2002.
- [8] Bonneel, Nicolas, et al. "Fast modal sounds with scalable frequency-domain synthesis." *ACM Transactions on Graphics (TOG)* 27.3 (2008): 24.
- [9] Chadwick, Jeffrey N., Steven S. An, and Doug L. James. "Harmonic shells: a practical nonlinear sound model for near-rigid thin shells." *ACM Trans. Graph.* 28.5 (2009): 119-1.
- [10] Zheng, Changxi, and Doug L. James. "Toward high-quality modal contact sound." *ACM Transactions on Graphics (TOG)*. Vol. 30. No. 4. ACM, 2011.
- [11] Langlois, Timothy R., et al. "Eigenmode compression for modal sound models." *ACM Transactions on Graphics (TOG)* 33.4 (2014): 40.
- [12] Doel, Kees van den. "Physically based models for liquid sounds." *ACM Transactions on Applied Perception (TAP)* 2.4 (2005): 534-546.
- [13] Langlois, Timothy R., Changxi Zheng, and Doug L. James. "Toward animating water with complex acoustic bubbles." *ACM Transactions on Graphics (TOG)* 35.4 (2016): 95.
- [14] An, Steven S., Doug L. James, and Steve Marschner. "Motion-driven concatenative synthesis of cloth sounds." *ACM Transactions on Graphics (TOG)* 31.4 (2012): 102.
- [15] Hideo Tsuru, "Numerical analysis of vibration of xylophone by Finite Difference Method," *Acoustical Society of Japan*, Vol. 67, no. 7, pp. 296-301, 2011.
- [16] Genrokuro Nishimura, Koichi Takahashi, "Impact Sound of Steel Ball Collision," *Journal of the Society for Precision Mechanics of Japan*, Vol. 28, no. 4, pp. 220-230, 1962.
- [17] S.P.Timoshenko and J.N.Goodier, "Theory of Elasticity THIRD EDITION," McGraw-Hill Book Company Inc., pp. 423-436, 1970.
- [18] Yuan, Xiaojun, et al. "Simulation of acoustic wave propagation in dispersive media with relaxation losses by using FDTD method with PML absorbing boundary condition." *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 46.1 (1999): 14-23.
- [19] Eston Schweickart, Doug L. James and Steve Marschner, "Animating elastic rods with sound," *ACM Transactions on Graphics* 36(4):1-10, 2017.