



HAL
open science

Sound Synthesis by Simulation of Percussive Gesture of a Virtual Animated Agent

Sylvie Gibet, Alexandre Bouënard

► **To cite this version:**

Sylvie Gibet, Alexandre Bouënard. Sound Synthesis by Simulation of Percussive Gesture of a Virtual Animated Agent. GIMS, May 2006, Leeds, United Kingdom. pp.3:1-3:6. hal-00497291

HAL Id: hal-00497291

<https://hal.science/hal-00497291>

Submitted on 3 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sound synthesis by simulation of percussive gesture of a virtual animated agent

Sylvie Gibet, Alexandre Bouenard

VALORIA, University of Bretagne Sud, Vannes, France

sylvie.gibet@univ-ubs.fr

Abstract

This paper presents a system which simulates a physically-based sound synthesis algorithm, controlled by a virtual animated agent. This system provides a global framework for studying the interaction between virtual autonomous and realistic humanoids and audionumeric processes. We are more specifically concerned by physically based simulation, both for the virtual character and the virtual instruments, since we aim at producing natural gestures thus leading to realistic musical interaction and credible synthesized sounds.

1. Introduction

Most interactive sound synthesis applications controlled by gestures try to express some relationship between the gestural actions and the effects on the produced sounds. Different methods are employed. For the approaches which directly manipulate acoustical parameters of sound, the main difficulty concerns the proper definition of mapping strategies between the gesture and the sound parameters. With such approaches, we generally observe a lack of realism in the produced sounds and in the interaction process. Other approaches use recorded sound clips for providing sounds corresponding to object interaction in a scene. Although this approach has the advantage that the sounds are realistic, there are many physical effects which cannot be taken into account with this technique. In particular, when interacting with percussive instruments, the location and

the magnitude of the force impulse play a major role in the produced sound timbre.

In recent years, there has been more and more emphasis placed on physical modeling in order to restore the causal link as the natural basis for sound synthesis. A basic hypothesis which supports this causal approach relies on the motor theory of perception which states that the objects of the perception emerge out of the image of physical processes that produce it. Following this idea, we believe that the organization of acoustical structures implicitly contains gesture patterns. Moreover, this approach makes possible the fine modeling of the interaction process between sound and gesture.

In this paper we propose a general physical framework, both for gesture modeling and sound synthesis. The sound-producing instrument (mechanical or acoustic) is represented as a collection of vibrating structures responding to external demands. The model consists in simulating the propagation of the vibrations along these structures. This approach requires the development of a numerical algorithm which simulates this wave equation. Furthermore, producing natural sounds necessitates the modeling of interaction mechanisms between the excitatory device and the vibrating structure, as well as the high-level control of such mechanisms.

This physical modeling approach is placed in a broader simulation and animation framework, in which a virtual humanoid dynamically interacts with a physical simulated instrument. The idea here is to show the influence of the gesture control in the context of sound simulation. This influence can be evaluated as the subtle effects of the interaction processes when objects are

colliding but also can be measured by the quality of the preparatory gesture. This simulation framework makes also possible the analysis of the spatio-temporal coordination of motion responsible for musical expressivity. In this paper we present the foundation of this system and the first simulation modules that are being developed. The paper is organized as follows. Related works are reviewed in section 2. In section 3 we present a general overview of the system. The environment for the dynamical modeling of virtual characters is described in section 4. The physical model of a specific percussion instrument is proposed in section 5. The interaction methodology is presented in section 6. Finally, we conclude with further perspectives.

2. Related works

Previous work concerns both the sound synthesis using physical models and dynamic animation of virtual characters.

Modeling the wave equation for vibrating objects was first proposed by Hiller and Ruiz in 1971 [1]. Different methods have been proposed in the last two decades. Finite element methods can be used for boundary value differential equations. These methods were proposed for numerical integration of vibrating systems [2-3]. Some approaches use an assembling of masses and damped springs to model physically-based sound synthesis [4]. The main drawback with this approach is the loss of accuracy in the simulation results due to the propagation of numerical errors. Another physical modeling technique, called modal synthesis computes analytically the different modes of vibration instead of numerical integration [5]. A software for modal synthesis called *Modalys* has been developed by IRCAM (France). More recently, several approaches proposed to link the vibration modes to physical models and to fit the sound synthesis with real time graphical synthesis. O'Brien and al [6] found a correspondence between arbitrary shapes and tetrahedral volume elements modeled by finite element equations. A similar approach is proposed in [7], where the surface vibrations are modeled by a system of point-masses and damped springs, which fit an analytical solution. A simpler method, using

finite difference technique has been proposed in [8] to simulate a vibrating circular membrane with forces applied on various impact points.

Concerning human simulation, several physics environments are now available to design complex graphical scenes including physical entities. A specific physics engine, called PhysX (AGEIA) helps to construct rigid articulated bodies, to define collision shapes of the bodies, materials that define the surface properties and constraints between objects and their environment [9]. Numerous methods have been proposed to control dynamical characters. Physical models of humanoids were first designed by Hodgins et al. [10]. More recently, controllers based on automatics theory have been proposed to control dynamical characters [13-14]. In [11-12], methods based on learning algorithms have been developed, and are used to control cinematic or dynamical hand-arm systems.

In this study, we use the AGEIA PhysX framework, which aims at modeling and animating in real time several entities in a scene [9]. The humanoid is constructed by the attachment of joints between segments.

3. Overview of the system

An experimental platform has been built to test our humanoid animation and sound synthesis, using the physical engine PhysX. As shown in Figure 1, the functional architecture is divided into three main parts. The first one deals with the motion database. A task specification, directly extracted from captured motion controls the second module, called *Animation Engine*. This one is composed of a set of specific controllers which drive the mechanical model of the humanoid through the specification of forces and torques applied on specific joints. The third module, called *Instrument Simulation* is dedicated to the simulation of numerical vibrating objects. Both the humanoid and the instrument are displayed in real time, thanks to a rendering engine.

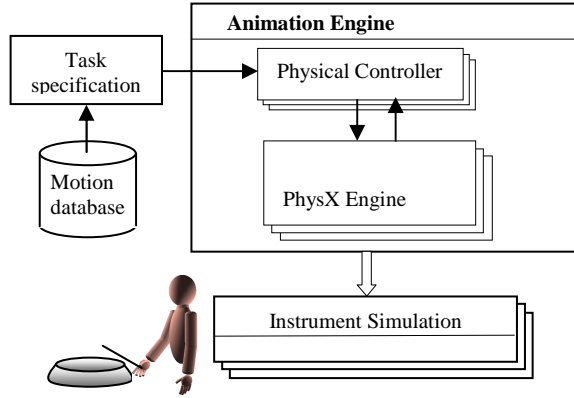


Fig. 1: Architecture of the system

4. Simulation of a simplified drum

We model in this application a 2D elastic circular membrane. The projection of the membrane on the horizontal plane defines an open surface Σ whose boundary is Γ . The material properties of the drum are represented by the tension τ , the uniform distribution of mass ρ and the damping parameter ξ of the membrane. Let define a vertical force $\tau \cdot f(p, t) \cdot dp$ by unit of surface, applied on a specific point p of this membrane, with $p = (x, y)$ a point of Σ . The vertical displacement of the points located on the membrane satisfies the two-dimensional wave equation, with boundary conditions:

$$\frac{1}{c^2} \frac{\partial^2 \phi(x, t)}{\partial t^2} + \xi \frac{\partial \phi(x, t)}{\partial t} - \nabla \phi(x, t) = f(x, t)$$

with $c^2 = \frac{\tau}{\rho}$, the initial conditions:

$$\phi(x, 0) = \phi_0(x), \quad \frac{\partial \phi}{\partial t}(x, 0) = \phi_1(x)$$

and the boundary condition:

$$\phi(x, t) = 0 \quad \forall x \in \Gamma$$

For numerical simulation, we discretize the surface with a set of uniform squares of equal size and keep only those contained in Γ . We use the finite difference method. This method approximates the Laplacian and the second derivative with respect to time with Taylor formula. This method does not make any assumption concerning the form of the solution. It has also the advantage to be very easily implemented, and to behave well numerically.

The only substantial difficulty which arises with such an approach is the real-time requirement necessary if we want to interact with the instrument using forces provided by the virtual agent.

5. Gesture interaction with the drum

Different modes of interaction are considered, depending on whether both systems – the stick manipulated by the virtual agent and the drum – are or not physically coupled. When these systems are separated, each mechanical system independently evolves. Their respective evolution can be described by the wave equation for the drum (the forces being null) and the motion equation for the virtual agent. When they are in contact, these equations are coupled through interaction forces equal in magnitude and acting in opposite sense (Action-Reaction mechanical principle). The condition of the separation of the two systems depends on the nature of the motion. In the analysis of simple percussive gestures we consider that the rupture occurs when the interaction forces are nearly zero.

In our approach, we make the assumption that the quality of the produced sounds is directly linked to the control of the percussionist gesture, taking into account both the planning of the gesture and the fine interaction process. The proposed methodology used to study these different levels of influence between gesture parameters and sound parameters can be stated as follows. First, we analyze the influence of the physical properties of the drum, in particular its elasticity and its shape. Second, we analyze two categories of gestures and the effect of their variation on the synthesized sounds: hitting gestures with a stick and sliding gestures with the finger. For both classes, the influence of the gesture parameters are qualitatively analyzed through the observation of impact forces and velocity profiles. The corresponding acoustical spectrum around the instant of the impact is also studied.

5.1. Hitting gestures

For hitting gestures we want to emphasize the importance of the force magnitude and the force profile, and also the effect of the point

impact of the stick on the membrane. In a first step we only consider vertical forces applied on the horizontal membrane. With our gesture control model, the planning gesture is defined by two target points: the initial target represents the beginning of the gesture; the final target the planned objective. This objective can be located directly on the membrane, but it can also be located under the surface, leading to a gesture with a greater velocity at the impact point. Therefore, by programming a fictive target more or less close to the membrane, we can act on the force of the strike. The total duration to reach the target is not a priori known. It depends on the task and on the motion controller.

Furthermore, we want to analyze how the specification of different set of mechanical parameters for the arm-stick system (stiffness of the joints and inertia of the segments) as well as the physical properties of the membrane (mostly the elasticity) may influence the dynamics of the interaction. If for instance the end of the arm is flexible during the preparatory motion, a narrow peak appears in the force profile.

5.2. Sliding gestures

In order to produce sliding movements on the membrane, we specify a slightly moving target.

Here again, we can modify the trajectory of the gesture, spatially and temporally, and analyze the influence on the produced sounds.

6. Conclusion and perspectives

We have presented in this paper a general physically-based framework for analyzing gesture interaction between a virtual percussionist and an artificial drum. This approach requires both the design of a virtual actor and the computing of a virtual 2D membrane. The actor is modeled using the AGEIA PhysX dynamic engine, and dedicated controllers are designed and applied on the actors' joints. A methodology for studying the influence of the physical parameters of both the gesture and the instrument, as well as the influence of the preparatory gesture is presented. This is a work in progress. First results have already been obtained for the

humanoid control and a simple 2D membrane. The synthesis of sound in real-time according to various hand-instrument interaction, and with little loss in perceived sound quality will require the development of acceleration techniques.

7. References

- [1] Hiller, L., Ruiz, P. Synthesizing Musical sounds by solving the wave equation for vibrating objects. *J.A.E.S.*, Vol. 19, n°6, pp. 462-470, 1971.
- [2] Chaigne, A. and Doutaut, V. Numerical simulations of xylophones, *Journal of Acoust. Soc. Am.* 101, 1, 539-557, 1997.
- [3] O'Brien, J.F., Cook, P.R., and Essl, G. Synthesizing sounds from physically based motion. In *SIGGRAPH'01: Proceedings of the 28th annual conference on Computer Graphics and interactive techniques*, ACM Press, NY, USA, 529-536, 2001.
- [4] Florens, J.L. and Cadoz, C. The physical model: modeling and simulating the instrumental universe. In *Representations of Musical signals*, G.D. Poli, A. Picciali, and C. Roads, Eds. MIT Press, Cambridge, MA, USA, 227-268, 1991.
- [5] J.M. Adrien, The missing link : modal synthesis synthesis, *Representations of musical signals*, pp. 269-297, 1991.
- [6] O'Brien, J.F., Shen, C., and Gatchalian, C.M. Synthesizing sounds from rigid-body simulations. In the *ACM SIGGRAPH 2002 Symposium on Computer Animation*, ACM Press, 175-181, 2002.
- [7] Raghuvanshi, N. And Lin, M.C. *Interactive Sound Synthesis for Large Scale Environments*, I3DG 2006, 2006.
- [8] Gibet, S., and Marteau, P.F., *Gestural control of sound synthesis*, ICMC'90, 1990.
- [9] <http://www.ageia.com/>

- [10] J.K. Hodgins, W.L. Wooten, D.C. Brogan, J.F. O' Brien, Animating human athletics, Proceedings in SIGGRAPH pp. 71-78, 1995.
- [11] S., Marteau P.F., Julliard F. Models with Biological Relevance to Control Anthropomorphic Limbs : A Survey. Lecture Notes in Artificial Intelligence, LNAI 2298, Ipke Wachsmuth and Martin Fröhlich Eds., Springer Verlag, London, 2002.
- [12] Gibet S., Marteau P.F. Expressive Gesture Animation Based on Non Parametric Learning of Sensory-Motor Models, CASA 2003, Computer Animation and Social Agents, 2003.
- [13] P. Faloutsos, M. van de Panne, D. Terzopoulos, Composable controllers for physics-based character animation, SIGGRAPH Computer Graphics Proceedings pp. 251-260, ACM Press / ACM SIGGRAPH, 2001.
- [14] A. Shapiro, P. Faloutsos, F.H. Pighin, Hybrid control for interactive character animation, 11th Pacific Conference on Computer Graphics and Applications, pp. 455-461, 2003.