Enhancing the Visualization of Percussion Gestures by Virtual Character Animation

Alexandre Bouënard, Sylvie Gibet, Marcelo M. Wanderley

Enhancing the visualization of percussion gestures by virtual character animation

Alexandre Bouënard
Samsara / VALORIA
Univ. Européenne Bretagne
Vannes, France
alexandre.bouenard@univ-ubs.fr

Sylvie Gibet*
Bunraku / IRISA
Univ. Européenne Bretagne
Rennes, France
sylvie.gibet@irisa.fr

Marcelo M. Wanderley*
IDMIL / CIRMMT
McGill University
Montreal, Qc, Canada
marcelo.wanderley@mcgill.ca

ABSTRACT

A new interface for visualizing and analyzing percussion gestures is presented, proposing enhancements of existing motion capture analysis tools. This is achieved by offering a percussion gesture analysis protocol using motion capture. A virtual character dynamic model is then designed in order to take advantage of gesture characteristics, yielding to improve gesture analysis with visualization and interaction cues of different types.

Keywords

Gesture and sound, interface, percussion gesture, virtual character, interaction.

1. INTRODUCTION

Designing new musical interfaces is one of the most important trends of the past decades. Efforts have constantly been made to elaborate more and more efficient devices in order to capture instrumental gestures. These technical advances have given rise to novel interaction opportunities between digital instruments and performers, and the creation of new sound, image or tactile synthesis processes. Our main guideline aims at providing a set of pedagogical tools for helping the study of percussion gestures. Among these, rendering real instrumental situations (interaction between performers and instruments) and exploring the gestural space (and its corresponding visual, gestural and sounding effects) are of great interest. Eventually, our final goal is to build new virtual instrumental situations, especially with gesture-sound interactions controlled by virtual characters. This paper offers a new tool for visualizing percussion gestures, which exploits both the analysis and synthesis of percussion gestures. The analysis process is achieved by capturing the movements of performers, while a physical model of virtual character is designed for the synthesis. The visualization is composed of different views of both the virtual character and the instrument. It is finally enhanced with interactions between graphics modeling, physics synthesis of gesture and sound replay.

The paper is organized as follows. In section 2, previous work and motivations are discussed. The analysis process of percussion (timpani) gestures is detailed in section 3. Visualization and interaction concerns are discussed in section 4. Finally, we conclude with further perspectives.

2. RELATED WORK

Previous works concern both percussion-related models and interfaces, and works combining virtual character animation and music.

Most of the work about percussion gesture and sound deals with the design of new electronic percussion devices, thus creating either new interfaces (controllers) and/or new sound synthesis models and algorithms.

On the one hand, new interfaces are based on increasingly efficient devices that are able to track gestures. Electronic percussions such as Radio Baton [1], Buchla Lightning [3], Korg Wavedrum [20] and ETabla [14] are digital musical instruments that are improving or emulating acoustic phenomena, by taking into account gesture cues such as position, touch and pressure. More recent work take advantage of various techniques, such as magnetic gesture tracking [17], computer vision [16] or the physical modeling of the drum skin [13].

On the other hand, it is also achieved by designing sound synthesis models and algorithms, ranging from purely signal-based to physically-based methods [9]. These works rarely include the study of the instrumental gesture as a whole, especially regarding to its dynamical aspects or its playing techniques, even if some take into account real measurements [2] and physical parameters mapping with percussion gesture [5]. Playing techniques can be qualitatively observed and used ([14] [12] [8]) for a better understanding of percussive gestures.

They can also be quantified thanks to capturing techniques [24], among which the most used is motion capture by camera tracking. But whichever method is used to reproduce the quality of the instrumental gesture, it generally fails to convey its dynamic aspect. That is why we explore in this paper the possibility to physically animate a virtual character performing percussive gestures so that its intrinsic features are available to our interface.

As for previous work combining virtual character animation and music, very few studies are available, especially

*Also with Samsara / VALORIA, Université Européenne de Bretagne (UEB), Vannes, France

NIME08 Genova, Italy

in a mean of taking advantage of virtual character animation for helping the visualization of gestures. The DIVA project[1] used virtual character animation for audiovisual performances output driven by MIDI events [11]. Hints about motion capture characteristics towards the quality of re-synthesis of the movement [15] have been proposed. The influence of music performance on virtual character's behavior [23] has also been emphasized. Some work aims at extracting expressive parameters from video data [4] for enhancing video analysis. Eventually, a solution consists in directly animating virtual models from the design of sound[2]. These studies are nevertheless out of the scope of virtual character animation as a gestural controller for enhancing the visualization and the analysis of instrumental situations.

3. TIMPANI PERFORMANCE

There are many classifications of percussion instruments, one of the most established typologies is based on physical characteristics of instruments and the way by which they produce sound. According to this classification, timpani are considered as *membranophones*, "producing sound when the membrane or head is put into motion" [6].

3.1 Timpani Basics

Timpani related equipment is mainly composed of a bowl, a head and drumsticks (Figure 1). In general, timpanists have to cope with several timpani (usually four) with bowls varying in size [19]. As for timpani drumsticks, they consist of a shaft and a head. They are designed in a wide range of lengths, weights, thicknesses and materials [6] and their choice is of great importance [18].

Figure 1: Timpani player's toolbox: bowl, head and drumsticks.

Timpani playing is characterized by a wide range of playing techniques. First, there are two main strategies for holding drumsticks (Figure 2, left side): the "French" grip (also called "thumbs-up") and the "German" grip (or "matched" grip).

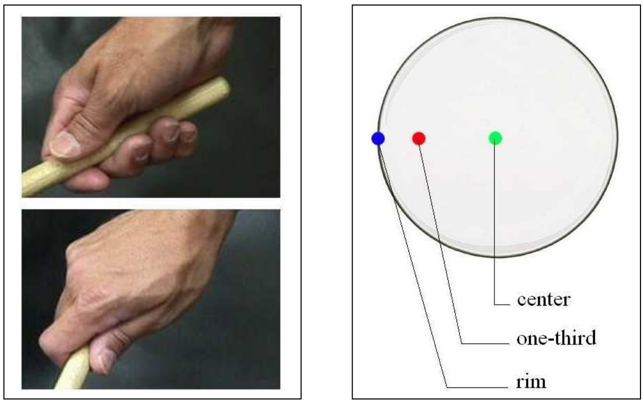

Figure 2: *Left:* French (top) and German (bottom) grips; *Right:* Impact locations on the drumhead.

Players commonly use three distinct locations of impacts (Figure 2, right side). The most used is definitely the one-third location, while the rim appears rather rarely.

A database of timpani gestures has been created and is composed of five gestures: legato, tenuto, accent, vertical accent and staccato. Each gesture is presented on Figure 3, showing the space occupation (Y-Z projection) of each drumstick's trajectory, and highlighting the richness of timpani playing pattern variations.

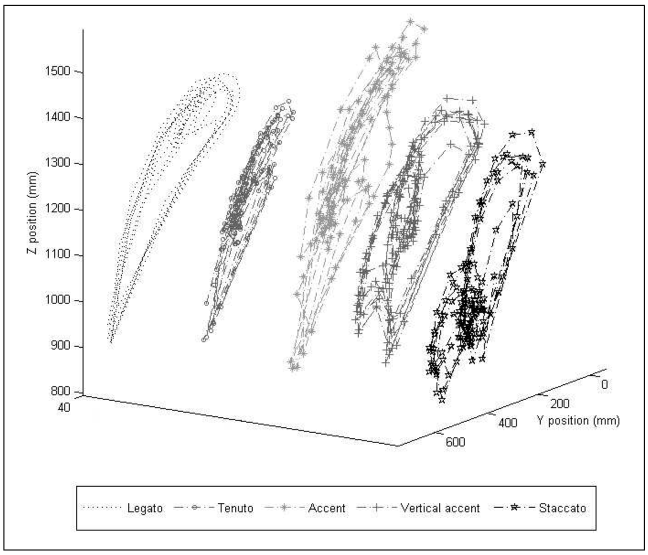

Figure 3: Timpani playing variations - Tip of the drumstick trajectories (Y-Z projection). Legato is the standard up-and-down timpani gesture. Tenuto and accent timpani variations show an increase in velocity and a decrease in space occupation (in the Y direction). Vertical accent and staccato timpani variations also show an increase in velocity, and are characterized by an increase of space occupation (in the Y direction) for a more powerful attack and loudness.

Taking into account these various features, timpani gestures are thus characterized by a wide variability. Next session will concern the quantitative capture of these variations.

[1]DIVA project : *www.tml.tkk.fi/Research/DIVA*
[2]Animusic : *www.animusic.com*

3.2 Motion capture protocol and database

We propose to quantitatively characterize timpani gestures by capturing the motion of several timpani performers. We use a camera tracking Vicon 460 system[3] and a standard DV camera that allow both the retrieval of gesture and sound.

The main difficulty using such hardware solutions is then the choice of the sampling frequency for the analysis of percussive gestures (because of the short duration of the beat impact [7]). For our experiments, cameras were set at 250 Hz. With a higher sampling frequency (500 Hz and 1000 Hz), we could expect to more accurately retrieve beat attacks, but the spatial capture range is significantly reduced so that it is impossible to capture the whole body.

In order to retrieve beat impacts, markers have also been placed on the drumsticks. The smaller timpani (23") has been used to emphasize sticks rebounds.

Figure 4: A subject performing the capturing protocol. The number of markers and their positions follow Vicon's *plug-in Gait* indications.

Three performers (c.f. Figure 4) were asked to perform our timpani-dedicated capturing protocol, yielding our timpani gestures database. Table 1 proposes a summary of the playing characteristics for each subject that has performed our capturing protocol. The differences between performers namely lie in their degree of expertise (Professor or Master student), the grip strategy that is used (French or German), their dominant (Left or Right) hand, and their gender.

Table 1: Timpani gestures data.

Subject	Expertise	Grip	Handedness	Gender
S1	Professor	F	Right	M
S2	Master stud.	G	Left	M
S3	Master stud.	G	Right	F

Each performer has been asked to perform a single stroke roll of each gesture variation (legato, tenuto, accent, vertical accent and staccato) presented in section 3.1. And for each of these gestures, the performer has been asked to change the location of the beat impact according to Figure 2 (right side). Finally, our database is composed of fifteen examples of timpani playing variations for each subject, and to each example corresponds five beats per hand. This database will be used when studying in detail the variations of the timpani gesture.

The use of widespread analysis tools integrated in Vicon software allow for the representation of temporal sequences as cartesian or angular trajectories (position, velocity, acceleration), but one can easily observe that such a representation isn't sufficient to finely represent the subtility of gesture dynamics, and cannot be easily interpreted by performers.

In the instrumental gesture context, we are mainly interested in also displaying characteristics such as contact forces, vibration patterns, and a higher-level interpretation of captured data (space occupation, 3D trajectories, orientation of segments).

4. VISUALIZATION

Our visualization framework proposes the design of a virtual instrumental scene, involving the physical modeling and animation of both virtual characters and instruments. Timpani gestures are taken from the database and physically synthetized, making available both kinematic and dynamic cues about the original motion.

4.1 Virtual instrumental scene

A virtual instrumental scene is designed using both graphics and physics layers. The OpenGL graphics API is used for rendering the virtual character, the timpani model, and rendering motion cues of these entities. It also allows users to explore the virtual instrumental space and to visualize the scene from different points of view.

The ODE physics API [22] is used for the physical simulation the virtual character and collisions.

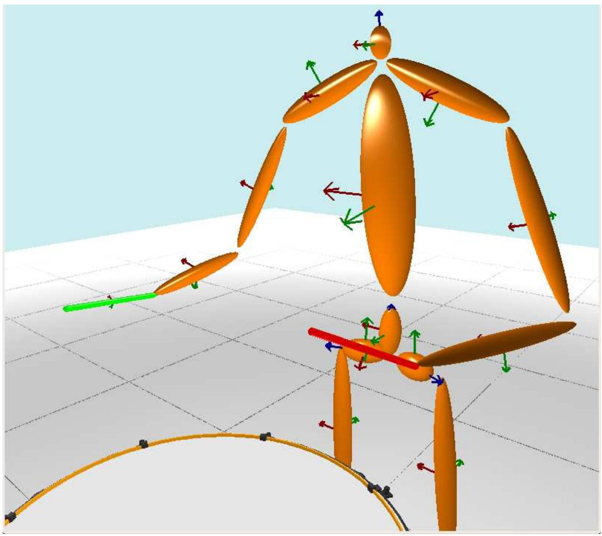

Figure 5: Real-time visualization of segments' orientations.

These graphics and physics layers build the primary visualization framework. It is possible to enrich this visualiza-

[3]Vicon : *www.vicon.com*

tion with both meaningful kinematic and dynamic motion cues since the overall gesture is available.

4.2 Kinematic cues

Kinematic motion cues can be of different types. Firstly, positions and orientations of any joint and segment composing the virtual character can be visualized (Figure 5) in real-time by the rendering of coordinate references.

Temporal trajectories describing the motion can be traced (Figure 6). These include position, velocity, acceleration, curvature trajectories, as well as plane projections, position/velocity and velocity/acceleration phase plots of segments and joints.

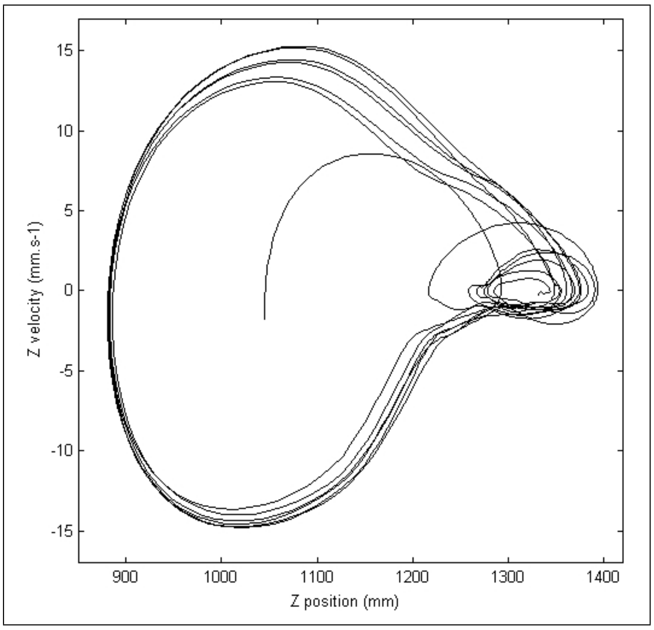

Figure 6: Example of kinematic trajectory plot. Tip of the drumstick : position/velocity phase along the Z axis.

Figure 6 shows an example of such plots, the trajectory represents the position/velocity phase (projected on the Z axis) of the drumstick.

Although temporal trajectories (Figure 6) convey helpful information about the motion, they cannot be visualized for the moment at the same time as our virtual instrumental scene rendering. We propose the real-time visualization of 3D trajectories and their corresponding bounding boxes (Figure 7). This helps in identifying the gestural space actually used during the performance.

In addition of these kinematic cues, we offer the visualization of dynamic characteristics of percussion gestures by physically modeling, simulating and controlling a virtual character.

4.3 Dynamic cues

The aim of the visualization of gesture's dynamic profiles is to facilitate the visualization of the interaction between the virtual character and the percussion model. Interaction information is available, thanks to physical modeling and simulation of instrumental gestures.

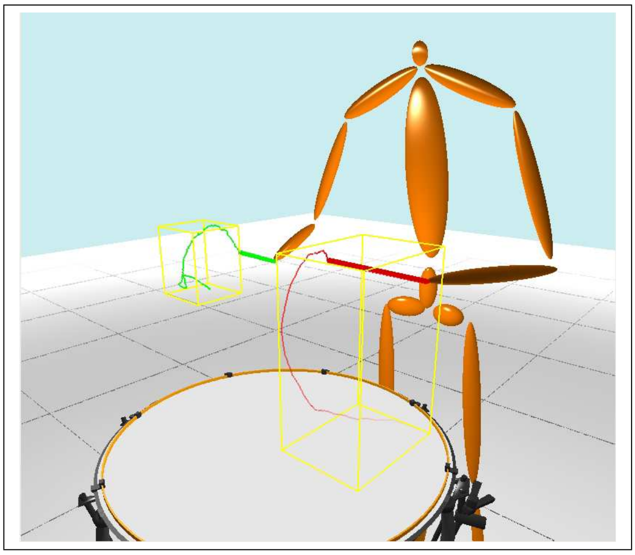

Figure 7: Real-time rendering of 3D trajectory and bounding box - drumstick tip trajectories helps in identifying the gesture space that is actually used.

4.3.1 Virtual character modeling and simulation

The dynamic simulation of instrumental gestures has been achieved by firstly proposing a dynamic model of a virtual character, and secondly by putting this physical model into motion through a simulation framework.

The virtual character is both modeled by its anthropometry and its physical representation. As for the anthropometry, it directly comes from motion capture. The physical representation of the virtual character is composed of segments (members) articulated by joints, each represented by its physical parameters (mass, volume, degrees of freedom).

The simulation framework is composed of two modules. The first one is the simulation of motion equations. Equations 1 and 2 describe the evolution of a solid S of mass m. The acceleration of a point M of the solid S is a_M and F_M is the resulting force applied on S at point M. The inertia matrix of S expressed at the point M is I_M , while Ω_S represents the angular velocity of S. Finally τ_M is the resulting torque applied on S at the point M.

$$m.a_M = F_M \tag{1}$$

$$I_M.\dot{\Omega}_S + \Omega_S.I_M.\Omega_S = \tau_M \tag{2}$$

Once the joints and members of the virtual character can be simulated by the emulation of motion equations, we offer a way to physically control the virtual character with motion capture data thanks to a Proportionnal - Integral - Derivative (PID) process (Figure 8).

The PID process translates the motion capture trajectories into forces and torques. Knowing angular targets from motion capture Θ_T and $\dot{\Theta}_T$, and knowing the angular state of the virtual character Θ_S and $\dot{\Theta}_S$, the PID computes the torque τ to be applied. K_p, K_i and K_d are coefficients to be tuned. This process ends the simulation framework and makes the virtual character able to dynamically replay instrumental timpani sessions.

The interactions between the virtual character, percussion model and the sound are then discussed. It is achieved by taking advantage of the dynamic characteristics that are available thanks to our virtual character dynamic model.

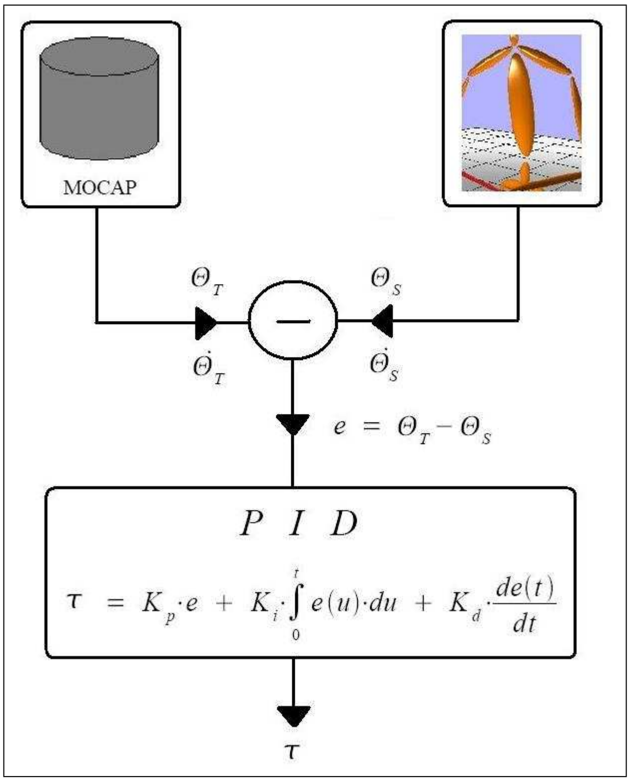

Figure 8: PID process. From motion capture data targets (angles Θ_T and angular velocities $\dot{\Theta}_T$), joints' current state (angles Θ_S and angular velocities $\dot{\Theta}_S$) and coefficients (K_p, K_i and K_d) to be tuned, torques τ are processed to physically control the virtual character.

4.3.2 *Interaction*

In order to account for the interaction between the virtual character's sticks and the timpani model, we suggest to render a propagating wave on the membrane of the timpani when a beat impact occurs. Although the rendering of such a wave isn't the theoretical solution of the wave equation, this model can take into account the biomechanical properties of the limbs and the properties of the sticks. Once the collision system detects an impact, kinematic and dynamic features - such as the velocity and the impact force - can be extracted. These features instantiate the attributes of the propagation of the wave making it possible the visualization of the position and the intensity of the impact (Figure 9).

Once kinematic and dynamic features of motion and physical interactions are obtained, we can set up strategies of sound production. In this paper, we limit ourselves to the triggering of pre-recorded sounds available from motion capture sessions. These sounds are played when the impacts of the virtual character sticks are detected on the membrane of the timpani model.

One can notice that the time when the sound is played doesn't depend on motion capture data, but depends on the physical simulation and interaction between the virtual performer and the percussion model. This provides an extensive way of designing new gesture-sound interactions based on both kinematic and dynamic gesture features.

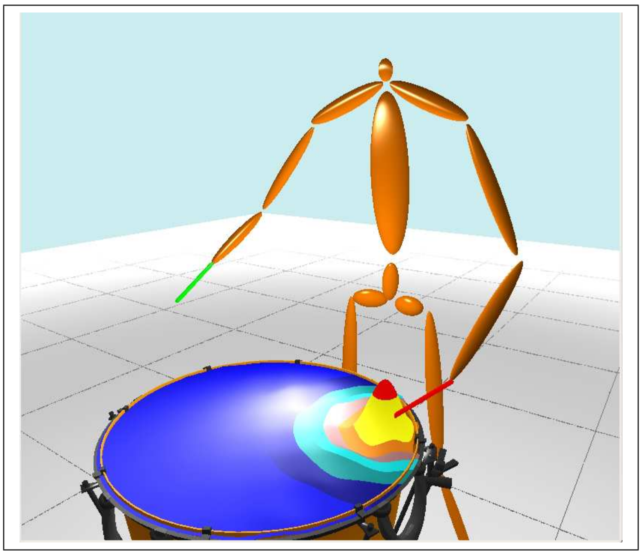

Figure 9: Dynamic cues about beat impact: visualization of the location and magnitude of the attack by the propagation of a wave.

5. CONCLUSION

We have presented in this paper a new interface for visualizing instrumental gestures, based on the animation of a virtual expressive humanoid. This interface facilitates the 3D rendering of virtual instrumental scenes, composed of a virtual character interacting with instruments, as well as the visualization of both kinematic and dynamic cues of the gesture. Our approach is based on the use of motion capture data to control a dynamic character, thus making possible a detailed analysis of the gesture, and the control of the dynamic interaction between the entities of the scene. It becomes therefore possible to enhance the visualization of the hitting gesture by showing the effects of the attack force on the membrane. Furthermore, the simulation of movement, including preparatory and interaction movement, provides a mean of creating new instrumental gestures, associated with an adapted sound-production process.

In the near future, we expect to enrich the analysis of gesture, by extracting relevant features from the captured motion, such as invariant patterns. We will also introduce an expressive control of the virtual character from a reduced specification of the percussion gestures. Finally, we are currently implementing the connection of our simulation framework to well-known physical modeling sound-synthesis tools such as IRCAM's Modalys [10] to enrich interaction possibilities of this framework. A similar strategy to existing frameworks, such as DIMPLE [21], using Open Sound Control [25] messages generated by the simulation engine, is being considered.

6. ACKNOWLEDGMENTS

The authors would like to thank the people who have contributed to this work, including Prof. Fabrice Marandola (McGill), Nicolas Courty (VALORIA), Erwin Schoonderwaldt (KTH), Steve Sinclair (IDMIL), as well as the timpani performers. This work is partially funded by the *Natural Sciences and Engineering Research Council of Canada* (Discovery and Special Research Opportunity grants), and the *Pole de Competitivite Bretagne Images & Réseaux*.

7. REFERENCES

[1] R. Boie, M. Mathews, and A. Schloss. The Radio Drum as a Synthesizer Controller. In *Proc. of the 1989 International Computer Music Conference (ICMC89)*, pages 42–45, 1989.

[2] R. Bresin and S. Dahl. Experiments on gesture : walking, running and hitting. In Rochesso & Fontana (Eds.): *The Sounding Object*, pages 111–136, 2003.

[3] D. Buchla. Lightning II MIDI Controller. http://www.buchla.com/. Buchla and Associates' Homepage.

[4] A. Camurri, B. Mazzarino, M. Ricchetti, R. Timmers, and G. Volpe. Multimodal analysis of expressive gesture in music and dance performances. In A. Camurri, G. Volpe (Eds.): *Gesture-Based Communication in Human-Computer Interaction*, LNAI 2915, Springer Verlag, pages 20-39, 2004.

[5] K. Chuchacz, S. O'Modhrain, and R. Woods. Physical Models and Musical Controllers: Designing a Novel Electronic Percussion Instrument. In *Proc. of the 2007 International Conference on New Interfaces for Musical Expression (NIME07)*, pages 37–40, 2007.

[6] G. Cook. *Teaching Percussion*. Schirmer Books, 1997. Second edition.

[7] S. Dahl. Spectral Changes in the Tom-Tom Related to the Striking Force. *Spech, Music and Hearing Quarterly Progress and Status Report*, KTH, Dept. of Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden, 1997.

[8] S. Dahl. Playing the Accent: Comparing Striking Velocity and Timing in Ostinato Rhythm Performed by Four Drummers. *Acta Acoustica with Acoustica*, 90(4):762–776, 2004.

[9] C. Dodge and T. A. Jerse. *Computer Music: Synthesis, Composition and Performance*. Schirmer - Thomson Learning, 1997. Second edition.

[10] N. Ellis, J. Bensoam, and R. Causse. Modalys Demonstration. In *Proc. of the 2005 International Computer Music Conference (ICMC05)*, pages 101–102, 2005.

[11] R. Hanninen, L. Savioja, and T. Takala. Virtual concert performance - synthetic animated musicians playing in an acoustically simulated room. In *Proc. of the 1996 International Computer Music Conference (ICMC96)*, pages 402–404, 1996.

[12] K. Havel and M. Desainte-Catherine. Modeling and Air Percussion for Composition and Performance. In *Proc. of the 2004 International Conference on New Interfaces for Musical Expression (NIME04)*, pages 31–34, 2004.

[13] R. Jones and A. Schloss. Controlling a physical model with a 2D Force Matrix. In *Proc. of the 2007 International Conference on New Interfaces for Musical Expression (NIME07)*, pages 27–30, 2007.

[14] A. Kapur, G. Essl, P. Davidson, and P. Cook. The Electronic Tabla Controller. *Journal of New Music Research*, 32(4):351–360, 2003.

[15] M. Peinado, B. Heberlin, M. M. Wanderley, B. Le Callennec, R. Boulic, and D. Thalmann. Towards Configurable Motion Capture with Prioritized Inverse Kinematics. In *Proc. of the Third International Workshop on Virtual Rehabilitation*, pages 85–96, 2004.

[16] T. Mäki-Patola, P. Hämäläinen, and A. Kanerva. The Augmented Djembe Drum - Sculpting Rhythms. In *Proc. of the 2006 International Conference on New Interfaces for Musical Expression (NIME06)*, pages 364–369, 2006.

[17] M. Marshall, M. Rath, and B. Moynihan. The Virtual Bodhran - The Vodhran. In *Proc. of the 2002 International Conference on New Interfaces for Musical Expression (NIME02)*, pages 153–159, 2002.

[18] F. W. Noak. *Timpani Sticks*. Percussion Anthology. The Instrumentalist, 1984. Third edition.

[19] G. B. Peters. *Un-contestable Advice for Timpani and Marimba Players*. Percussion Anthology. The Instrumentalist, 1984. Third edition.

[20] G. Rule. Keyboard Reports: Korg Wavedrum. *Keyboard*, 21(3):72–78, 1995.

[21] S. Sinclair and M. M. Wanderley. Extending DIMPLE: A Rigid Body Simulator for Interactive Control of Sound. In *Proc. of the ENACTIVE'07 Conference*, pages 263–266, 2007.

[22] R. Smith. Open Dynamics Engine. www.ode.org.

[23] R. Taylor, D. Torres, and P. Boulanger. Using Music to Interact with a Virtual Character. In *Proc. of the 2005 International Conference on New Interfaces for Musical Expression (NIME05)*, pages 220–223, 2005.

[24] A. Tindale, A. Kapur, G. Tzanetakis, P. Driessen, and A. Schloss. A Comparison of Sensor Strategies for Capturing Percussive Gestures. In *Proc. of the 2005 International Conference on New Interfaces for Musical Expression (NIME05)*, pages 200–203, 2005.

[25] M. Wright, A. Freed, and A. Momeni. Open Sound Control: The State of the Art. In *Proc. of the 2003 International Conference on New Interfaces for Musical Expression (NIME03)*, pages 153–159, 2003.