



HAL
open science

From action to sound : a challenging perspective for haptics

Annie Luciani, Jean-Loup Florens, Nicolas Castagné

► **To cite this version:**

Annie Luciani, Jean-Loup Florens, Nicolas Castagné. From action to sound : a challenging perspective for haptics. World Haptics Conference 2005, 2005, Pise, Italy. pp.[5]. hal-00910648

HAL Id: hal-00910648

<https://hal.science/hal-00910648>

Submitted on 26 Jun 2014

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.

From Action to Sound: a Challenging Perspective for Haptics

Annie Luciani¹, Jean-Loup Florens², Nicolas Castagné²

¹ ICA laboratory INPG, ² ACROE

INPG, 46 av. Félix Viallet, 38000 Grenoble, France

E-mail: Annie.Luciani@imag.fr, Florens@imag.fr, Nicolas.Castagne@imag.fr

Abstract

The link from action to sound is under-explored in Virtual Realities. The article categorizes the action-sound relations, proposing the term “ergotic” to refer to the case when there is a mechanical interaction between the subject and the sound source. Force-feedback devices and physically-based models are indeed required for implementing virtual ergotic action-sound systems, able to engrave the energetic consistency of the physical action in the sound.

1. Introduction

Nowadays, in virtual realities (VR), haptic devices are mainly used for interacting with 3D geometrical objects. As for it, sounds implementations put the emphasis on the 3D spatialization for 3D sound localization. However, as another very basic function, sound conveys the identification of the mechanical properties of the sounding object and of its interaction with the physical environment. This function that has been a major concern in computer music, in which force feedback interaction leads to a promising shift within the digital musical instruments, from the concept of *parameter control* of signal based models, to the concept of an energetically coherent *gesture interaction* with a physically-based digital artifact. Inspired by this potential shift in computer music, we assume that ensuring such energetic consistency throughout the general chain from gesture to sound is a major issue for VR and haptics.

2. Typology of Action-Sound Relations

In non-electrified environment, there are two types of action – sound relations:

(1) When action produces directly sounds by direct and physical manipulation of a physical object. Most objects are silent objects until we act on them. The main property of these action-sound relation is that there is a physical energetic consistency between the physical action and the produced sounds.

(2) When the sound is not produced by a physical human action on a physical body, as for sounds which exist autonomously (the sound of the wind, of a river, etc...) or by non-physical human actions.

Among the known epistemic and semiotic functions, Claude Cadoz [1] has introduced a complementary function - the ergotic function - to operationally categorize the human-environment relations in three functions.

The epistemic function is the function through which the environment is known. One can speak about the epistemic function of touch [2] and the epistemic function of seeing [3]. Similarly, we can speak about “epistemic hearing”, as when we listen a sound source to identify it. The semiotic function refers to a symbolic activity. It is considered here in a restrictive use to qualify actions that are used at a symbolic level during the man-environment interaction. A typical example is the action to show a target with the finger: people look the pointed target and not the finger. All the situations in which the dynamical properties of the sound are not correlated to the dynamic of the actions can be described by relations between pure semiotic action and pure epistemic perception, (as loops over two conversely oriented grey arrows of figure 1). They are pure semiotic-epistemic or *non-ergotic* relations, characterized by the fact that there is no need of energetic consistency between the actions and the produced sensorial event, i.e. here, sounds.

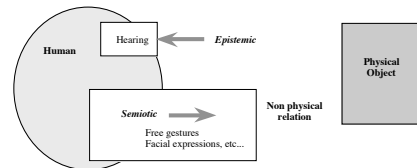


Figure 1. Pure semiotic-epistemic, i.e. non-ergotic, action-sound relations

The term “ergotic” identifies a property of human-environment interaction that cannot be supported by any association of the semiotic and the ergotic functions, that is the energetic consistency characterizing the physical interactions. The term

« Haptic », that is often used to state this function, covers unfortunately several meanings in the perception domain and in the devices (force and/or tactile). « Ergotic », from « ergos » which means « physical work, energy », represents clearly the core property of such function. All the situations in which there is a physical energetic interaction between humans and a physical object are ergotic situations (figure 2).

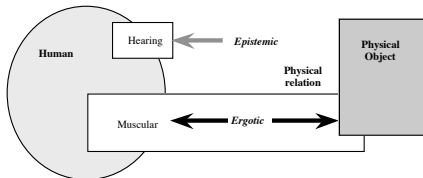


Figure 2. The ergotic action-sound relation

The operational Cadoz' typology allows to answer clearly to the question "Where and Why Haptic transducers are – or are not – a necessity?".

The non-ergotic relation between actions and sounds in computer concerns all the cases where the sound is triggered by a digital action and/or where the sounds parameters are controlled by an action. Since there is no energetic consistency between action and sounds, such relations do not require necessarily neither force feedback devices nor physically-based simulation.

Ergotic action-sound relations were the last to be considered in computer environments, mainly in Computer Music [4,5,6], though this type of situation is the most frequent in our daily live. Ergotic action-sound situations suppose an energetic consistency throughout the chain from the action to the sound. Their implementation in a computer context requires the energetic consistency not to be broken at any stage from the haptic and the acoustical transducers. For example, the rendering of hitting or rubbing a sound object, requires not only force feedback devices but also adequate real-time simulation of adequate physically-based models and adequate links between them. This means that the designing of all the elements is a modeling process as a whole.

3. Quantitative Requirements

3.1. Non-ergotic Computer Chain

In non-ergotic situation, the computer chain is generally composed of pure sensors and of a sound synthesis process, linked through a control process, as a *mapping* of the gesture on the sound process [7].

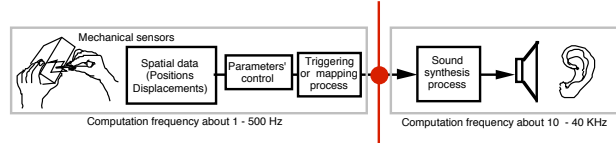


Figure 3. The electrified « non-ergotic » action-sound relation

Such chain can be clearly cut in two non-retroactive parts (figure 3). The sensors and the mapping process can run at a computation rate of about 1 - 500 Hz. The sound synthesis part has to be computed at a rate about 10 - 40 KHz. The link between the two parts is unidirectional, from the left to the right, gesture part *controlling* the sound part. Examples of corresponding implementations, eventually featuring haptic transducers, can be found in VR and in Computer Music [7,8,9,10,11]. The major remark that can be pointed out here is that, in such case, when an force feedback is used, it does not impact the sound quality.

3.2. Ergotic Computer Chain

In comparison with the non-ergotic case, the basic ergotic computer chain (Figure 4) is necessarily:

- augmented by force feedback actuators
- transformed by replacing the computational processes by computation of physically-based models.

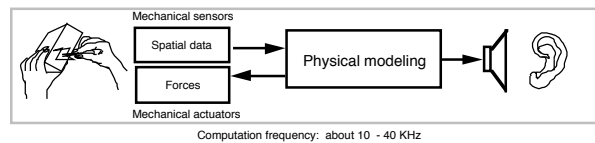


Figure 4. Basic computer chain for an ergotic action-sound relation

This chain presents noticeable differences with the previous non-ergotic computerized chain, in bandwidth, dynamic ranges in the representation of physical variables and temporal latencies.

3.2.1. Bandwidth

The cut-off frequency of the phenomena along the natural chain is close to 10Hz for the large displacements of the gesture and to few 10 KHz for the small acoustical deformations. Computer implementations should *a priori* be fully computed at the acoustical frequency (10 to 50 KHz). Any simplification leading to damage the energetic coherence of the chain belongs to modelling choices.

3.2.2. Dynamic range for the physical variables

We performed measurements of the forces circulating during a hit between a finger and a rigid surface. For the softest hit possible, i.e. when the hit

produces both the lowest audible sound and a force feedback at the limit of the sensitivity of the tactilo-kinesthetic sensation, the forces are about 10 mN. During hard acoustical shocks, the forces are impulsive peaks of about 1000 N, lasting less than 50 microseconds. The dynamic range of forces is thus at least 10^5 (100dB). Physically speaking, the 100 dB dynamic range for the forces corresponds necessarily with a similar ratio in the displacements and the velocities within the acoustical vibrating parts. Consequently, to obtain similar acoustical quality in a computer implementation, the overall dynamic range that must be considered for the variables throughout the chain is close to 100 dB. This concerns as well the variables within the virtual manipulated object, the input and output variables on the force feedback device, and the output variables on the loudspeaker.

3.2.3. Temporal latencies

In the non-ergotic situation, the only latency to be considered is the delay between the action and the sounds, that is a soft constraint of about 10-100 ms. In ergotic electrified situation one plays a real vibrating object within the hands, the object is around 40 cm far from the body. This corresponds to a delay of about 1 ms from gesture to heard sounds. More important is that these acoustical frequencies exist in the interaction between the body and the sound object. Thus, their interaction has, *in principle*, to be considered at the acoustical sampling frequency (44Khz), leading to a latency between action and force feedback of no more of the sampling period, that is up to 25 microseconds.

4. Optimized Ergotic Action-Sound Chain

All these quantitative issues are very hard to implement in a computer context. A first optimisation can be performed by considering that, taken separately, the gestural manipulation (without sound production) and the sound object, have different dynamic properties, leading to represent the physical object by two physical interacting parts (figure 5). One which deforms at the acoustical scales and one which interact with the hands. The first needs to implement: (1) a computation frequency of 40Khz, (2) displacements about 10 microns, and (3) dynamic ranges for the forces and displacements of $10^5 - 100$ dB. The second needs to implement properties of non-sounding mechanical objects: (1) a computation frequency of about 500-4Khz, (2) displacements about a few centimeters, and (3) dynamic ranges for the forces and displacements of $10^5 - 100$ dB.

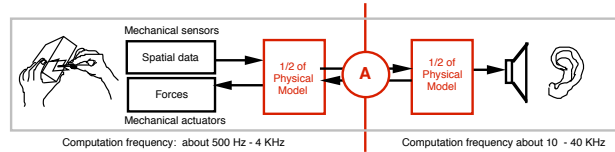


Figure 5. Optimization of the electrified ergotic chain

Thus, the the two parts can be optimize separately. The decorrelation between the range of the forces to be returned to the hand and within the physical sound object allows to choose a returned force of 100 N over 500 microseconds, corresponding to the same motion quantity of 1000 N over less than 50 microseconds measured in an acoustical hard shock, that is impossible to obtain with mechanical actuators.

But, this decomposition in two parts running at two dynamic scales rises the question of the place and the nature of the frontier (the red line and circle), to maintain the energetic consistency between the two parts. Such a problem cannot be solved in a general way. We present in the next paragraph a solution for an exemplary case of ergotic action-sound situation.

5. A typical case of ergotic action-sound situation

Friction is a typical example of the ergotic action – sound relation. During friction interactions, actions and sounds are permanently closely physically linked during the rubbing, producing typical non predictable properties of the sound that cannot emerge from a sound parameters' control process. One very relevant is the non predictable timbre changes and modulations, showing that in such ergotic action-sound situation the properties of the sounds are closely correlated to the properties of the closed loop physical action.

The two real-time computer implementations of the action-sound rubbing effect, “the finger on a glass” and the “bowed string” (Figure 6), have been implemented to render such properties. A physically-based model of physical string coupled with a non-linear interaction [6] to the ERGOS versatile force feedback device [12] produces in real time believable glass-finger friction and bowed string sounds, with the whole ranges of action and auditory expressions one can find in the corresponding natural situations.

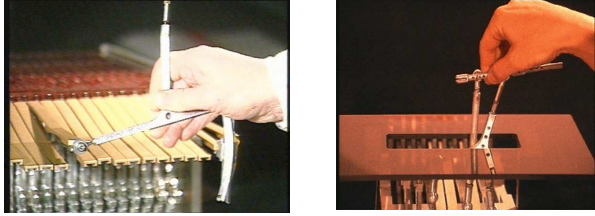


Figure 6. The finger on the glass and the bowed string

The rubbed objects are represented by two distinct 1D physically-based models (Figure 7): one (upper line) for the lateral deformations under the friction interaction and one (lower line) for the transversal collision and pressure deformations. The bow is represented by two spatially independent components: 1D sliding bow and 1D pressure bow, coupled with the 2 actuators-sensors of the force feedback devices by two independent 1D visco-elastic buffers. The spatial concordance of the two chains is obtained only by the coupling of the two bows with a stick or stylet morphology and by means of the modulation of the friction non-linearity law by the pressure forces. Thus, the only spatial features within the whole model are the threshold of the collision buffer and the morphology of the haptic device.

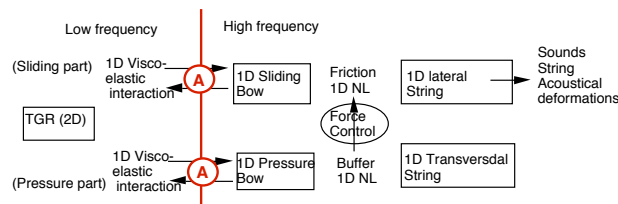


Figure 7. A functional representation of the two models

The whole simulation process puts the emphasis on the dynamic properties by implementing a very low temporal latencies (less than 0,3ms), a very small minimal returned friction forces (0,5mN at 3 KHz) for the feeling of friction as a texture, a wide dynamic range on the friction axis (10N for the attack forces) and for collision-pressure forces on the other axis (50N). Despite the morphological dissemblance with a natural surface rubbing, users manipulate with a high level of accuracy such “virtual violin” or “virtual rubbed glass”, taking pleasure to play with them as “instruments”. No noticeable discomfort was reported, though one could think given the lack of complex geometrical and spatial properties within the system (3D interaction, 3D control of the string, 3D wide displacements, etc), and the major relevant effects that characterize such situations are obtained: pizzicati, timbre or pitch modulations, very accurate modulation

of the friction and the pressure and of their correlation, way of attacks, creaking, etc.

6. Conclusion

Many important action-sound relations, that we call *ergotic*, are necessarily supported by a mechanical interaction between a player and a sound source. They feature an energetic continuity throughout the chain from the hand to the ear. Implementing such relations requires a drastic quality in the dynamics. Bandwidth, latencies and dynamic ranges, as evaluated in this article, are noticeably higher than usually considered.

Although if they seem too much “extreme”, taking them into account leads to shift the modeling process from a predominance of 3D shapes and geometric coherence, to a predominance of dynamics and energetic consistency. A subsequent property is then to reduce significantly the geometric and spatial complexity of the scene, that are computational demanding, and to open new prospects in better balanced specification of the tasks, from spatially-oriented tasks to dynamically-oriented tasks.

Acknowledgements

This work has been supported by the French Ministry of Culture and by the FP6 Network of Excellence IST-2002-002114 - Enactive Interfaces

References

- [1] Cadoz C., Wanderley M. (2000). “Gesture-Music” - in “Trends in Gestural Control of Music”, M. Wanderley and M. Battier, eds, pp. 71-94 -©2000
- [2] Hatwell Y., Streri A., Gentaz E.. “Touching for knowing : Cognitive psychology of haptic manual perception”. John Benjamins Ed.. 2004.
- [3] Dretske, F.I: “Seeing and Knowing” - The University of Chicago Press, 1969.
- [4] Cadoz C., Luciani A., Florens J.L.. “Responsive Input Devices and Sound Synthesis by Simulation of Instrumental Mechanisms: The Cordis System”. Computer Music Journal, 8, N°3, pp. 60-73. M.I.T. Press, Cambridge Mass. 1984.
- [5] Gillespie B. “The Virtual Piano Action: Design and Implementation“ Proceedings of the International Computer Music Conference, Aarhus, Denmark, Sept 12-17, 1994. pp. 167-170.
- [6] Florens JL. “Real time Bowed String Synthesis with Force Feedback Gesture”. Invited paper. Forum Acousticum. Sevilla. November 2002.
- [7] Hunt A. D., Paradis M., and Wanderley M., “The importance of parameter mapping in electronic

- instrument design,” *Journal of New Music Research*, vol. 32, no. 4, pp. 429–440, December 2003.
- [8] O’Modhrain S., Chafe C.: “Incorporating Haptic Feedback into Interfaces for Music Applications” - in proceedings of ISORA, World Automation Conference, 2000.
- [9] Wood J., Magennis M., Arias E. F. C., Gutierrez T., Graupp H., Bergamasco M. “The Design and Evaluation of a Computer Game for the Blind in the GRAB Haptic Audio Virtual Environment”, *Proc. of Eurohaptics*, 2003.
- [10] Nichols C: “the vBow: a Virtual Violin Bow Controller for Mapping Gesture to Synthesis with Haptic Feedback” – in *Organised Sounds* – Leigh Landy ed. - Leicester, United Kingdom, 2002.
- [11] Rimell S, Howard D M., Tyrrell A M., Kirk R, Hunt A – “Cymatic: Restoring the Physical Manifestation of Digital Sound using Haptic Interfaces to Control a New Computer Based Musical Instrument” – *Intern. Computer Music Conference*, Sweden, 2002.
- [12] Florens JL, Luciani A, Castagne N, Cadoz C. “ERGOS : a Multi-degrees of Freedom and Versatile Force-feedback panoply”. *Proceedings of Eurohaptics*, pp356-360, Germany, 2004.