MaD: Mapping by Demonstration for Continuous Sonification

Jules Françoise, Norbert Schnell, Frédéric Bevilacqua*
STMS Lab, IRCAM–CNRS–UPMC, Paris, France.

1 Introduction

Gesture has become ubiquitous as an input modality in computer system, using multi-touch interfaces, cameras, or inertial measurement units. Whereas gesture control often necessitates the visual modality as a feedback mechanism, sound feedback has been significantly less explored. Nevertheless, sound can provide both discrete and continuous feedback on the gestures and the results of their action. We present here a system that significantly simplify and speedup the design of continuous sonification. Our approach is based on the concept of *mapping by demonstration*. This methodology enables designers to create motion-sound relationships intuitively, avoiding complex programming that is usually required in common approaches. Our system relies on multimodal machine learning techniques combined with hybrid sound synthesis. We believe that the system we describe opens new directions in the design of continuous motion-sound interactions. We developed prototypes related to different fields: performing arts, gaming, and sound-guided rehabilitation.

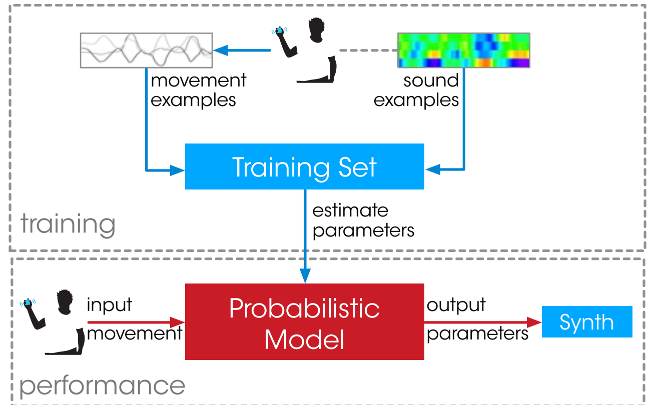

Figure 1: *Mapping by Demonstration Workflow*

2 Technology

Our system is innovative on two different levels. The first level concerns the user and/or designer: the system allows for rapid prototyping of motion-to-sound mapping by demonstration. The mapping is automatically learned by the system when movement and sound examples are jointly recorded (see Figure 1). In particular, our applications focus on using vocal sounds - recorded while performing action - as primary material for interaction design. The second level is technical and concerns the integration of specific probabilistic models of the mapping (Multimodal and Hierarchical Hidden Markov Models [Françoise et al. 2013; Françoise et al. 2012]) with hybrid sound synthesis models. In particular, the system supports the simultaneous use of granular, concatenative and physical modeling synthesis, which offer a diverse set of sound qualities. Importantly, the system is independent of the type of motion/gesture sensing devices, and can be used with different sensors such as cameras, contact microphones, and inertial measurement units.

3 Applications

Our system is currently applied in different fields, which demonstrates its versatility. First, it is used in applications that require - learning specific movements, such as performing arts. As shown in the videos[1], the use of vocalization, is particularly interesting since it is inherently used in these activities to support action or describe movement. Second, the system is currently evaluated for rehabilitation that could benefit from auditory feedback. In this case, the use of either vocal sounds or physically-inspired models allows for the design of motion-sound interaction metaphors that should facilitate movement training. Third, we believe that our system can provide vocally impaired persons with alternative ways to express themselves sonically through movement.

4 Presentation at Siggraph

We propose two different gaming scenarios that will allow attendees to try the system in a playful manner. In the first game (*Sonic Stories*), players first record vocal sketches that can then be used to communicate messages through their gestures. A set of cards that symbolize a variety of characters, actions, and objects suggest players to invent stories together. The goal of the second game (*Sonic Choreographies*), is to challenge the players' ability to learn complex movements with the help of auditory feedback. The players are sonically guided through movement sequences that become more and more challenging from one level to another! How fast can you learn?

Acknowledgements

We acknowledge support from the ANR project Legos 11-BS02-012 and the Labex SMART (ANR-11-LABX-65).

References

FRANÇOISE, J., CARAMIAUX, B., AND BEVILACQUA, F. 2012. A Hierarchical Approach for the Design of Gesture–to–Sound Mappings. In *Proceedings of the 9th Sound and Music Computing Conference*, 233–240.

FRANÇOISE, J., SCHNELL, N., AND BEVILACQUA, F. 2013. A Multimodal Probabilistic Model for Gesture–based Control of Sound Synthesis. In *Proceedings of the 21st ACM international conference on Multimedia (MM'13)*, 705–708.

*{jules.francoise, norbert.schnell, frederic.bevilacqua}@ircam.fr

[1] Additional material and examples can be found on: http://ismm.ircam.fr/siggraph2014-mad/

SIGGRAPH 2014, August 10 – 14, 2014, Vancouver, British Columbia, Canada.

ACM 978-1-4503-2961-3/14/08