Sound ”Gesturefication”, sound/gesture-mapping for the interaction with recorded sounds

Baptiste Caramiaux and Norbert Schnell

IRCAM, CNRS-UMR STMS, 1 Place Igor Stravinsky, 75004 PARIS, France
{baptiste.caramiaux,norbert.schnell}@ircam.fr

1 Introduction

With ”Gesturefication” we designate the creation of a mapping between gestural control and sound synthesis parameters based on the recording and analysis of a sound and related gestures. These gestures may have actually produced the sound or simply describe the sound from a particular point of view such as instrument performance, music conducting or active and embodied music listening.

Our working hypothesis is that mappings can be derived from the coherence of data sets describing the gestures on one hand and the sound on the other. In the ideal case, these mappings have to take into account their temporal evolution as well as the recurrence of patterns in both processes. This paper describes a first approach using an adaptive model for the mapping between instantaneous motion and audio features.

2 Related Works

Mapping strategies can be divided into two sub-problems [3]. First, *Explicit Mapping* where input to output is explicitly defined. Second, *Implicit Mapping* where a generative method provides the mapping. In this last issue, Lee and Wessel in [4] used a multi-layer neural network to control sound synthesis in real-time. Another well-known model is Hidden Markov Model (HMM) which can implement a gesture follower [1]. Whereas neural networks and HMM need training and are used for temporal features, this paper provides a new mapping method based on data coherence called Canonical Correlation Analysis (CCA) [2].

3 Method

CCA is a way of measuring the linear relationship between two sets of variables. It finds two bases maximizing the cross-correlation between two sets of random variables, whatever their dimension. The first set of variables is N_1 motion descriptors from a controller and the second set is N_2 sound descriptors with $N_2 > N_1$.

The algorithm gives us two projection matrices involving two sets of projected variables : the canonical components. The i-th component of the first set corresponds to the i-th component of the second according to a correlation coefficient c_i. Therefore, the projection process through these previous matrices defines the mapping. Since we based the method on observations taken at different times, temporal coherence is ensured by an alignment before the CCA process as shown in figure [1].

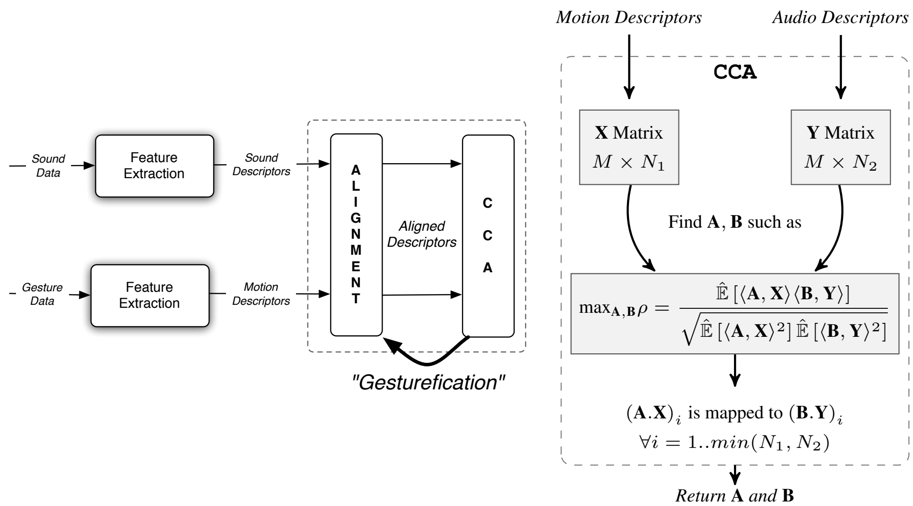

Fig. 1. The "gesturefication" iterative process : a temporal alignment following by the canonical correlation analysis. CCA process returns the projection matrices defining the mapping

4 Results and conclusion

The algorithm gives us audio features explained by motion descriptors. Thus, sound descriptors are projected into a motion descriptors space. The synthesis process uses a corpus of audio grains which are described by a set of values. In this work, such grains are characterized by a set of audio canonical component values. Gesture canonical components ensure the audio grain selection by minimizing a distance between its values and those describing a grain.

A first validation of the described algorithm has been based on artificial mappings in MatLAB. A developed Max/MSP real-time synthesis application permits further qualitative validation and experimentation with the algorithm. In the first experiments with this application the gestures are performed on a graphics tablet. The user can first analyze a recorded gesture corresponding to a sound file and then re-perform the sound material in real-time. In this application gesture sonification is implemented by inverting the calculated sound gesturefication.

References

1. Frédéric Bévilacqua. A gesture follower for performing arts. Gesture Workshop, 2005.
2. Hotelling, Harold. Relations between two sets of variates. *Biometrika*, 28(3/4):321–377, December 1936.
3. Hunt, Andy and Wandereley, Marcelo M. Mapping performer parameters to synthesis engines. *Organised Sound*, 7:97–108, 2002.
4. Lee, Matthew and Wessel, David. Connectionist models for real-time control of synthesis and compositionnal algorithms. page 277, San Francisco, USA, 1992. ICMC92, Proceedings of the 1992 International Computer Music Conference.