Mapping sound to human gesture: demos from video-based motion capture systems

Frédéric Bevilacqua
Beckman Laser Institute,
University of California, Irvine
+1-949-824 4104
fbevilac@uci.edu

Christopher Dobrian
Music Department
University of California, Irvine
+1949-824-7288
dobrian@uci.edu

Jeff Ridenour
Information and Computer Science
University of California, Irvine
+1-949-854-7473
jridenou@ics.uci.edu

ABSTRACT

We report research on gestural control of digital music, currently performed at the University of California Irvine. Experiments with various motion capture systems are performed, including single video camera systems and a 3D optical motion capture system (Vicon 8). We present a series of demos of various approaches for motion analysis and gesture-to-sound mapping.

Keywords

gestural control, mapping, motion capture, gesture recognition

1. INTRODUCTION

The use of unfettered gesture to control electronic music is an old idea, as evidenced by the Theremin invented in the mid 1920's. Since then, several devices have been designed and used in performances and art installations[1]-[4]. Nonetheless, the control of sound/music based on movements is still considered as an outstanding challenge.

Such an approach has gained attention in the last decade, partially due to enabling new technologies such as new motion capture systems and cost-effective powerful personal computers. Concepts for gestural control of music have also benefited from the cross-fertilization with other expanding research fields and applications, such as human-computer interaction, human gesture recognition, computer animation, and biomechanics, to name a few.

At the University of California, Irvine (UCI), a small group of artists and scientists is experimenting with various motion capture systems and software for the development of interactive music/dance performances and/or art installations [6]-[8]. We will show demos and videos, showing motion analysis algorithms and gesture-to-sound mapping that we are currently developing.

2. MOTION CAPTURE SYSTEMS

Our experiments are based on two types of motion capture systems: 3D optical motion capture (Vicon 8) and single video camera. These two systems described below are complementary, in terms of technical specifications and price.

In both systems, the transformation of the movement to sound follows several steps, as summarized in Fig.1:

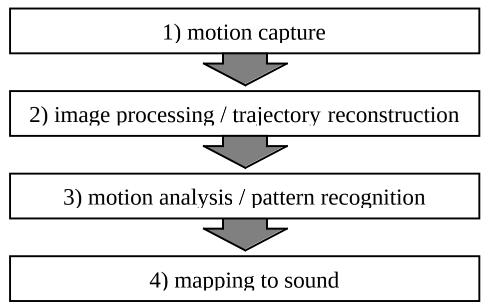

Fig. 1 Basic steps for gestural control of music

3D motion capture

We are using a commercial 3D motion capture system, Vicon 8, primary designed for animation purposes or biomechanics studies. Comprehensive information about the system can be found at the Vicon website.[9]

The Vicon system at the University of California, Irvine (UCI) School of the Arts is based on the simultaneous recording by eight video cameras of small reflective balls, known as “markers”, attached to a dancer. Usually, a set of 30 markers is used, allowing for the reconstruction of the 3D movements with standard animation software. Note that the steps 1) and 2) shown in Fig. 1 are directly operated by the Vicon proprietary hardware and software.

The system at UCI does not operate in real-time, but an update to the new Vicon Real-time system is expected. For this project, Vicon Motion Systems provided their in-house RTEmulator software, which allows one to emulate the behavior of the Real-time Engine.

The accuracy of the Vicon system and its 3D capability make it well suited to assess various algorithms for movement analysis and recognition, and to test gesture-to-sound mappings.

Single video capture

The Vicon system is expensive and not easily transportable, which make it impractical for many performances. In contrast, the use of a standard video camera is simple and cost-effective.

We are programming image processing necessary to extract motion parameters in the Max/MSP environment [10], using various extensions for video processing such as Jitter, Cyclops or SoftVNS. This corresponds to the step 2 described in Fig. 1

Gesture Segmentation and Recognition

Some of the research work in the UCI Motion Capture Studio has focused on the detection of individual gestures within the stream of motion capture data, and on recognizing particular kinds of gestures. Such gesture detection has been demonstrated to be feasible and musically useful. This corresponds to step 3 in Fig 1. Two approaches have been pursued [7]:

1) Gesture segmentation, using laplacian edge detection on acceleration curves. Parameters can then be deduced from the gesture segments to produce discrete, parameterized events, which can be used in conjunction with more traditional continuous position and velocity data generated by motion capture.

2) Principle component analysis (PCA) or independent component analysis (ICA) is used to analyze classes of short motions (training set). Subsequent movement sequences can be then compared to the training set. In particular, "similarity" between the examined movement and the training set can be quantified.

Mappings

Step 4 in Fig. 1 corresponds to the mapping of the movement parameters—obtained after the movement analysis process—to sound. This step is generally intrinsically linked to a particular musical composition.

We are developing software in Max/MSP for the mapping to MIDI messages and digital signal processing, as well as digital video (in Jitter).

Complex mappings will be demonstrated, including using recognized gestures to control discrete events, and using continuous parameters to control algorithmic processes. Different paradigms for the relationship between the dancer and music will also be discussed (controller/instrument, improviser/collaborator, etc.)

3. DEMONSTRATIONS

The whole motion capture systems cannot be demonstrated, however software for mapping and videos illustrating the whole process will be shown.

1) We have recorded a series of short dance sequences and generic movements (walking, running, etc) with the 3D motion capture system. We are developing a program, called MCMmax, for the simultaneous sound generation and the display of the animations created, from the mocap data.

The mapping can be altered in real-time. Several different examples of motion analysis and mappings will be demonstrated, and compared on the same set of movements.

2) A series of videos demonstrating the setup and software we are developing with a single-camera video system. In contrast to the examples from the 3D motion capture system, the performer benefits from the real-time feedback of the sound produced by his/her movements.

The comparison of these different approaches for motion capture, movement analysis, and sound mapping will serve as a basis for discussion. Both technical and artistic considerations will be addressed ([7]-[8]).

Equipment needed:

The material will be in the form of short videos and Max/MSP programs, presented from a Macintosh PowerBook. A standard digital projector with good audio capability is required.

4. ACKNOWLEDGMENTS

This project has been supported by the facilities of the Realtime Experimental Audio Laboratory and the Motion Capture Studio at the University of California, Irvine. Vicon Motion Systems has provided us with their RTEmulator software to assist us in this software development. We acknowledge the contribution of Lisa Naugle, John Crawford, Alexia BonvinBevilacqua, Isabel Valverde for motion capure, animation and performance, Cayci Suitt, Gene Wie, Maybelle Tan, Mark Magpayo, Andre van der Hoek for programming, and David Cuccia for his contribution with the Principal Component Analysis.

5. REFERENCES

[1] "Trends in Gestural Control of Music", edited by: Marcelo Wanderley and Marc Battier, Ircam - Centre Pompidou - 2000.

[2] J. A. Paradiso, "The Brain Opera Technology: New Instruments and Gestural Sensors for Musical Interaction and Performance", Journal of New Music Research Vol. 28 No. 2, pp.30-149, 1999.

[3] A. Mulder, S. Fels and K. Mase. “Design of Virtual 3D Instruments for Musical Interaction”, Proceedings Graphics Interface '99, S. Mackenzie and J. Stewart Eds., 76-83, 1999.

[4] A. Camurri, S. Hashimoto, M. Ricchetti, A. Ricci, K. Suzuki, R. Trocca, and G. Volpe "EyesWeb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems", Computer Music Journal, Vol. 24 No. 1, pp. 57-69, 2000, or see http://musart.dist.unige.it/

[5] F. Paracino, G. Davenport, and A. Pentland "Media in performance: Interactive spaces for dance, theater, circus, and museum exhibits", IBM Systems Journal, Vol. 39, No. 3&4, 2000.

[6] F. Bevilacqua, L. Naugle and I. Valverde, “Virtual dance and music environment using motion capture” Proc. of the IEEE - Multimedia Technology And Applications Conference, Irvine CA, 2001.

[7] Bevilacqua, F., J. Ridenour, and D. Cuccia, “Mapping Music to Gesture: A study using 3D motion capture data”, Proceedings of the Workshop/Symposium on Sensing and Input for Media-centric Systems, Santa Barbara CA, 2002.

[8] Dobrian, C., “Aesthetic Considerations in the Use of ‘Virtual’ Music Instruments” Proceedings of the Workshop on Current Research Directions in Computer Music, Institut Universitari de l'Audiovisual, Universitat Pompeu Fabra, Barcelona, Spain, 2001.

[9] Vicon, www.vicon.com,

[10] Max/MSP, www.cycling74.com