Music control from 3D motion capture of dance

Frédéric Bevilacqua, Lisa Naugle, Christopher Dobrian

University of California, Irvine

fbevilac@uci.edu, lnaugle@uci.edu, dobrian@uci.edu

ABSTRACT

Research is currently being conducted at the University of California Irvine to develop novel approaches in music performance and composition generated from dancer's gestures. We report here specifically the modification of the "Vicon 8" motion capture system to control digital music. This system allows for the capture of a dancer's movement in 3D. Software is being developed, in the "Max/MSP" environment, to use the motion capture data for sound generation and/or alteration, through MIDI parameters or by controlling signal processing algorithms.

This approach is promising for the extensive study of various possible relationships between gesture and music. This paper describes the method currently developed, illustrated by a short movie example. We also briefly discuss future directions of this work in progress. Movie examples can be found at www.zotnet.net/~fbevilacqua/mocap/

INTRODUCTION

Most of the electronic music controllers have been created based on existing acoustic instruments (for example the piano keyboard). Such electronic controllers have the obvious advantage to be used relatively easily by "traditionally" trained musicians. However, there is no fundamental reason to be restricted to such types of controller. Indeed, the relation between sound generation and touch or gesture in electronic music is always artificially defined, on the contrary to acoustic instruments. In fact, the possibility of designing the interface between the gesture (or the touch) and the sound is a fascinating feature of digital music. Various unconventional types of controller have already been developed. See for example the review by M. Cutler, G. Robair and Bean in Electronic Musician[1]. Further detailed examples and discussions can be found in references 2 and 3. As generally noted, there is an emergence of whole new types of controllers, promoting fresh works, but also raising many questions about the new artistic language(s) they require.

Different groups[4,5] have recently been involved in electronic music triggered by dancer gestures, which might be considered as the "ultimate" sound controller, merging musical and visual arts.

In order to transform motion into sound one must generate a data stream that represents the movement of interest. This is usually accomplished through the direct placement of sensors sensitive to flexion or acceleration directly on the body or through the video recording(s) of the motion. Available software and hardware such as VNS (Very Nervous System)[6] or Big Eye[7] allow for the processing of such video images into MIDI parameters. Nevertheless, note that systems such as VNS or Big Eye can handle only a two dimensional (or planar) projection of a given movement.

At the University of California Irvine, we have the opportunity to use a commercial 3D motion capture system (Vicon 8), primary designed for animation purposes or biomechanics studies. Interestingly, this system can also be potentially used as a powerful music controller. Indeed, it offers the possibility to track the trajectory of multiple points on the dancer's body simultaneously in three dimensions. The 3D data (position, velocity or acceleration) of each of these points can be used to control music "parameters", following various algorithms which need to be defined.

We started research to implement software that can transform the movement data provided by this system into various music data, either in a post-processing mode or in real-time. The scope of this paper is to report the state of this ongoing multidisciplinary research.

GOALS

The general goal of this research is to explore various possible relationships between gesture and music. Two challenges are faced: 1) the technical aspect of designing or modifying systems fulfilling this goal, and 2) the artistic aspect of "mapping" music to gesture.
We report here work in progress that we are performing with the Vicon motion capture system. We describe the basics of software currently developed to transform the motion capture data in music parameters. The results of this research will be concretized soon in several artistic works, video or live performances of dance/music.

PRINCIPLE

Vicon motion capture

Comprehensive information about the system can be found at the Vicon website[8]. We summarize here only the basic principles of this system.
The Vicon system at UCI is based on the simultaneous recording by 8 video cameras of small reflective balls attached to a subject. The balls are light and do not interfere with the movement. The 8 video cameras are placed around the subject. Because each ball is simultaneously imaged by several video cameras, the 3D coordinates of each ball can be computed (a calibration being performed first). After the recording, the Vicon software reconstructs the trajectory of each ball. The standard Vicon system does not process the data in real-time. However, this system can be upgraded for real-time processing.

Motion Capture Data

Once the Vicon system has processed the data, it outputs them as a list of 3D coordinates of each reflective ball, frame after frame (typically 60 frames per second). If the balls are placed at chosen points of the "skeleton", the data can then also be expressed in term of angles between various human joints. Each data structure corresponds to different aspect of the movement. The basic data, i.e. the 3D coordinates, allow one to follow the position of the dancer relative to the room. Additionally, the data referring to the angle of bones have a direct biomechanic or anatomical interpretation of a particular gesture (for example the rotation of the wrist).

Method

For animation purposes, a minimal set of 33 balls is normally used, which correspond to a total of 99 parameters recorded each 16 ms!
Clearly, one of the challenges is to fully exploit such an important set of data. Nevertheless, the first approach is naturally to start with only few parameters, as will be illustrated in this paper.
We developed a program called "MoCap" in the Max/MSP environment[9]. Max/MSP is a high level graphical programming language (not to be confused with the animation software 3D Studio Max). It is designed for controlling synthesizers through a MIDI interface, and/or digital audio recording.
The purpose of the program "MoCap" is to access the motion capture data and transform them into either MIDI parameters or parameters controlling signal processing. At this stage of the research, the system is not used in real-time, but as a post processing software. It can be therefore seen as a composition tool, in which parameters are derived from gesture. At the end of the process, the music track can be added to either a standard video recording of the dancer, or an animation based on the motion capture.
A basic example of the program is shown in Figure 1, and the result is presented in movie 1. The goal of these examples is only to illustrate the method. As it will be discussed, the actual possibilities of the program are much broader.

Program basics

When activating the start button, the program outputs the data, i.e. the XYZ coordinates of each recorded point, frame by frame.

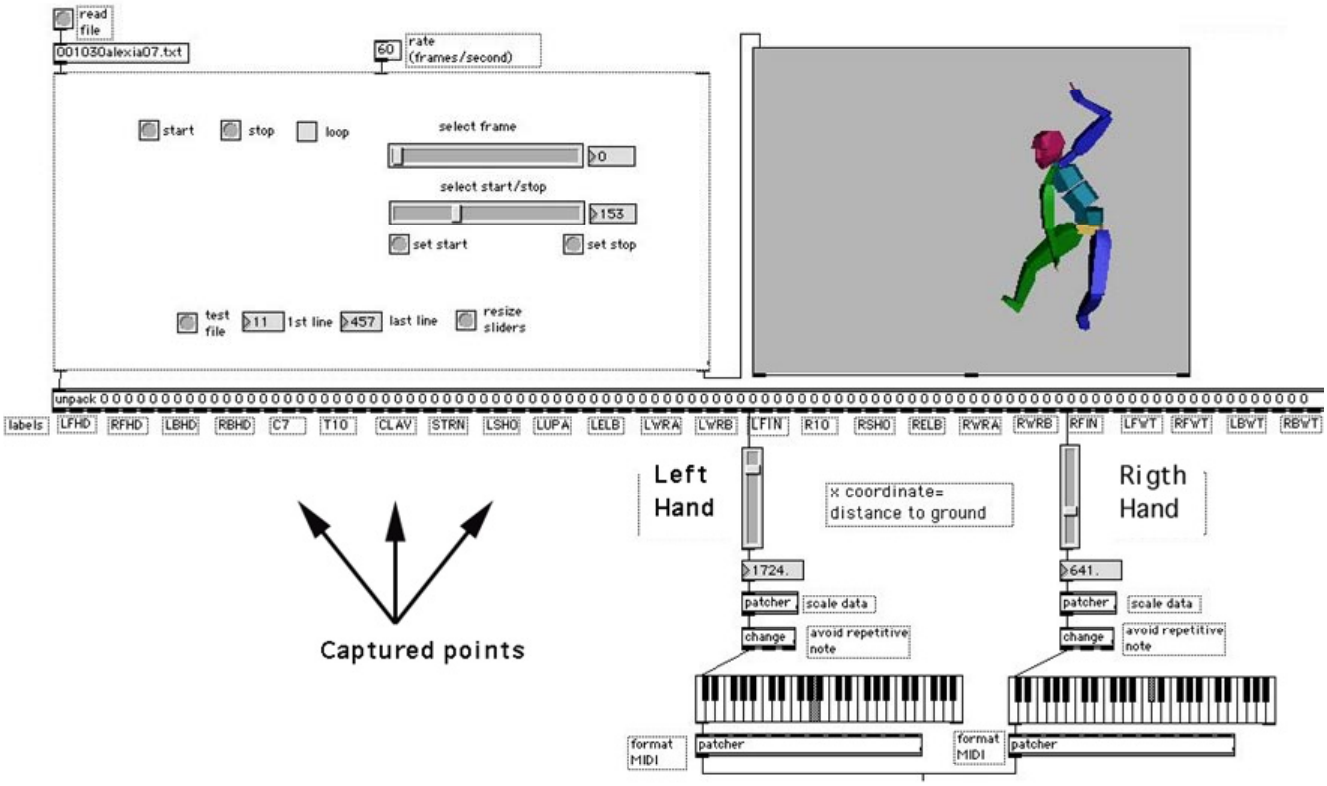
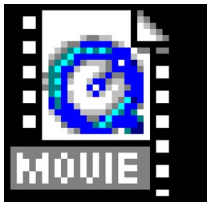


Figure 1. Basic implementation of the Max "MoCap" program. The data, the distances between the ground and the left and right hands, respectively, are scaled to MIDI parameters, such as a note.

The output of the data is synchronized with a QuickTime movie representing the motion capture. The movie is used as a monitor of the subject movement.
The Max function "unpack" is used to access all XYZ coordinates of each point separately. All coordinates appear therefore as possible connections form the "unpack" function.



Al1.mov

EXAMPLE

A very short example is presented. As already mentioned, the goal of this example only is to illustrate the basics of the program (this example is not supposed to have any artistic value at this point).
For the sake of simplicity, motion of only left and right hands are used to generate MIDI notes. Each hand plays a different sound.

Movie 1 : Note heights are generated from the distance of the hands to the ground (the higher the hand, the higher the note) . *Warning:* If the link to the movie is lost, access it through Internet at www.zotnet.net/~fbevilacqua/mocap/AL1.MOV

Other examples can be found at the following website: www.zotnet.net/~fbevilacqua/mocap/.

DISCUSSIONS

The linear relationship between the hand position and the note height, as shown in the previous example, is certainly one of the simplest possible. More complex relationships are currently being tried. First, not only the position of a particular location on the body can be considered, but also the velocity and the acceleration, each of which expresses different characters of the movement. Second, relationships other than linear can be introduced, such as probabilistic, adaptive, or based on pattern recognition. "Negative" relationships can be equally simple and useful. Also, introducing delay between the gesture and the resulting sound can potentially create pleasing effects. Nevertheless, in order to maintain an interest to both the performer and the audience to this approach, any relationship should be somehow felt. This sets limits to the complexity of the relationship between gesture and sounds that can be defined. As already pointed out, one challenge with working with a 3D motion capture system is to fully take advantage of its capabilities. For example, simple up and down movements of the arm can be well tracked by a single video camera and therefore do not require complex motion capture. On the contrary, rotation of the body usually requires sophisticated motion tracking. Since subtle movements can be accurately measured by the Vicon system, the music should somehow be specifically sensitive such movements. Also, this system enables us to work with both absolute distance measurements (relative to a fixed reference in the room) and relative distance measurements between various points of the body. Such features could also be reflected in the music.

CONCLUSION

The use of the Vicon motion capture system to control digital music has been demonstrated. This system is potentially a powerful tool for music control, even if it was not designed for such an application. The challenge for such an approach is to create music that takes advantage of the system. Each type of instrument set limits in terms of what type of music can be achieved technically. In the case of "free movements" as it is the case here, the limits are clearly different from standard musical instruments. A vast number of mappings between gesture and music can be implemented and explored. Based on these experiments, we will present soon music/dance works.

ACKNOWLEDGEMENTS

School of the Arts, UCI.

Alexia Bonvin, Lara James, Irma Castillo, and H.

REFERENCES

1. M.Cutler, G.Robair and Bean, “The outer limits: a survey of unconventional musical input devices”, Electronic Musician, pp 48-72, August 2000.

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

3. 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.

4. See the DIEM digital dance project, www.daimi.au.dk/~diem/dance.html

5. 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.

6. VNS, see www.interlog.com/~drokeby/vnsII.htm

7. Big Eye, see www.steim.nl/bigeye.html

8. Vicon 8, see www.vicon.com.

9. Max/MSP, see www.cycling74.com.