Proceeding of the IEEE Multimedia Technology And Applications Conference (MTAC), Irvine, 2001.

# Virtual dance and music environment using motion capture

**Frédéric Bevilacqua**
Beckman Laser Institute, University of California, Irvine
fbevi@bli.uci.edu

**Lisa Naugle**
School of the Arts, Department of Dance, University of California, Irvine
lnaugle@uci.edu

**Isabel Valverde**
Department of Dance History and Theory, University of California, Riverside
idvalverde@ucdavis.edu

## ABSTRACT

We present a multimedia project incorporating music and dance. We used a 3D motion capture system to produce animations from dance and generate the soundtrack from the dancer's movements. Movement analysis is performed to extract the important features of a particular gesture. Based on the parameters chosen from this analysis, various mappings between gesture and music are applied. In particular, the motion capture data is used to trigger and modify the timbre of sounds. This paper describes the method and the interactive environment that are under development.

## INTRODUCTION

The use of new media and technology challenges the traditional frontiers between artistic disciplines such as music, dance, studio art and drama. It also implies stronger collaboration between artists and scientists. We present an ongoing multi-disciplinary project using motion capture that fits within such a new paradigm. Other aspects of this collaboration are described in a companion proceeding by I. Valverde et al.[1]

In this project, motion capture of dance concurrently generates video animations and controls digital sounds. The video and music tracks are therefore linked together, producing new types of interplay between gesture and sound. In electronic music, the relationship between sound generation and touch/gesture is artificially defined, contrary to the way in which sound is produced in 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. Most electronic music controllers that have been created are 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. Nevertheless, there is an emergence of whole new types of controllers (see for example refs 2-7), promoting fresh works, and raising many interesting questions about the new artistic language(s) they require.

Several groups have developed systems that use dancer gestures to trigger music (Refs 8-14). In order to transform motion into sound, one must generate a data stream that represents the movement. This can be accomplished through the direct placement of sensors, sensitive to flexion or acceleration, on strategic parts of the body, or through the video recording(s) of the dancer.

The approach presented here is based on 3D optical motion capture. The motion capture system is not used as a real-time music controller, but as a basis to generate a 3D "virtual" environment where the visual and sonic components are connected together. Based on this primary material, either traditional animations or interactive multimedia work can be produced.

## PROJECT DESCRIPTION

Fig.1 gives an overview of the production process. Each step is described in the following sections.

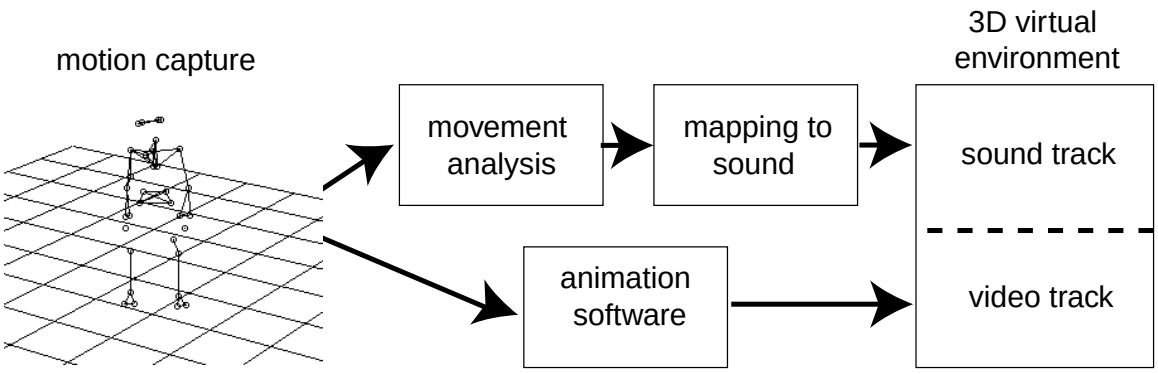


Figure 1. Diagram of the animation and music production.

### Motion capture and animation

We are using a commercial 3D motion capture system (Vicon 8), primary designed for animation purposes or biomechanics studies. We summarize here only the basic principles of this system. Comprehensive information about the system can be found at the Vicon website.[15]

The Vicon system at the University of California Irvine, School of the Arts, is based on the simultaneous recording by 8 video cameras of small reflective balls attached to a dancer, at 120 Hz. The balls are lightweight and interfere minimally with the dancer's movements. The 8 video cameras are placed around the subject to create a capture volume. Because each ball is simultaneously imaged by several video cameras, the 3D coordinates of each ball can be computed using triangulation (a calibration being performed first). The standard Vicon system does not process the data in real-time. However, this system can be upgraded for real-time processing.

Once the Vicon system has processed the data, it outputs them as a list of 3D coordinates of each reflective ball, frame by frame. 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.

For animation purposes, a minimal set of 33 balls is normally used. The data can be imported in animation software to be transformed in 3D animation (3DStudioMax, Maya). The final product can be either rendered as a 2D movie or kept in 3D format using, for example, VRML.

### Movement analysis

The use of motion capture to build animation is standard practice. The novelty of our approach is to also use the motion capture data to "drive" the soundtrack. One of the challenges is to fully exploit the set of data representing the dancer's movements (99 parameters). In order to simplify the task, the data processing is divided in two steps: movement analysis and sound mapping. The goal of the movement analysis is to reduce the raw data in few parameters that describe relevant features of the movements. These parameters then translate the dancer’s movement language. This process involved subjective choice and is an integral part of the artistic choice in the work. The complexity of such movement analysis can vary from choosing a limited set of recorded points to performing pattern recognition. We present below a few guidelines that we established:

The parameters can be classified in two categories:

1) Position/Distance/Angle

The 3D coordinates of the captured points allow for the calculation of various distances and angles. In particular:

- Position of a given point of the dancer relative to the stage.
- Distance/Angle between various points of the body.

Using such parameters, the dance can be deconstructed as a series of footsteps and arm movements, and other characteristics significant to the choreography.

2) Velocity/Acceleration.

The velocity and acceleration derived from the distance/angle parameters naturally expressed the dynamic of the movement. The velocity as a scalar quantity obviously corresponds to the speed of a particular movement. As a vector, it can be used to indicate direction changes. It enables, for example, one to derive “pulsation” from a dance figure. Other movement qualities can be expressed by the calculation of acceleration (velocity variation): “sustained” corresponding to low acceleration, “percussive” corresponding to high acceleration.

**Mapping to sound**

We developed a program to map aspects of motion to qualities of sound in the Max/MSP environment[16]. Max/MSP is a high level graphical programming language (not to be confused with the animation software 3D Studio Max). It is designed to perform real-time signal processing and to control synthesizers through MIDI interfaces.

The program reads the motion capture data and transforms it into either MIDI parameters or parameters controlling signal processing. The program in itself can be seen as part of the "score" of a musical composition that uses the gesture as input.

A video rendering of the animation can be played simultaneously to the music generation. The mapping gesture-sound can be modified in real-time.

Various mappings, from triggering discrete musical events to continuous control of sonic parameters, are currently the subject of experiments:

1) Sampled sounds, including text, are triggered based on criteria such as:

- Location of the dancer relative to the stage. The stage can be divided in “virtual” parts, each of one reacting specifically to the presence of the dancer.
- Change of direction of a particular body part. Such a mapping works well to produce rhythmic musical structure with “vibratory” or “pendular” movements.

2) Alteration of preexisting recorded musical soundtrack. Various filters can be continuously driven via the dancer’s movements. Promising results were obtained by using filters shaped by frequency modulation. The modulation index, the carrier and modulation frequencies are controlled by the movement, which produce a complex filtering function.

3) Sound synthesis. The pitch and/or timbre can be controlled by the movement. In particular, granular synthesis has been implemented.

One-to-one mapping between gesture and pitch usually sounds mechanical. More expressive music is achieved by mapping gesture to the timbre of the sound. Generally, dramatic and "organic" compositions are produced when the gesture/sounds are not directly predictable. Also, introduction of delay between the gesture and the resulting sound can create potentially interesting effects. Nevertheless, in order to maintain the interest of the audience to this approach, a relationship between visual and sonic components should still be discerned. This sets limits on the complexity of the relationship between gesture and sounds that can be defined.

## CURRENT AND FUTURE WORK

Using the scheme describing above, several projects are in progress simultaneously. First, we are developing a motion capture library with specific movement qualities (rhythmic, slow, etc.) or embodied states. This library is the basis for experimentation with movement analysis and sound mapping. Thus, two other libraries are emerging from this experimentation: a "sound" library and a "mapping" library.

An interesting feature of the method is the possibility of interactively selecting "motions", "sounds" and "mapping". We are exploring various forms of interactive multimedia presentation in which the viewer can make such selections. In particular, we anticipate building an interactive website and installations. Finally, this work will serve as a basis for live performances that would incorporate the animation environment, and/or real-time motion capture systems.

## ACKNOWLEDGEMENTS

Christopher Dobrian, Alexia Bonvin, Lara James, Irma Castillo, and H. Shahani.

## REFERENCES

1. I. Valverde, L.Naugle and F.Bevilacqua, “Embodied Shifts: Towards an experiential approach to dance-technology interfaces”, MTAC Proceeding 2001.
2. M.Cutler, G.Robair and Bean, “The outer limits: a survey of unconventional musical input devices”, Electronic Musician, pp 48-72, August 2000.
3. "Trends in Gestural Control of Music", edited by: Marcelo Wanderley and Marc Battier, Ircam - Centre Pompidou - 2000.
4. Interactive Systems and Instrument, Design in Music Working Group, http://www.notam.uio.no/icma/interactivesystems/wg.html
5. 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.
6. 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.
7. J. Paradiso, "Electronic Music Interfaces: New Ways to Play," IEEE Spectrum, 34(12), 18-30, 1997
8. DIEM digital dance project, www.daimi.au.dk/~diem/dance.html
9. Palindrome Intermedia Performance Group, http://www.palindrome.de
10. Troika Ranch MidiDancer http://www.troikaranch.org/mididancer.html
11. DIST - Laboratorio di Informatica Musicale, EyesWeb, http://musart.dist.unige.it/
12. Active Space Project, L.Naugle and J. Crawford, http://www.arts.uci.edu/lnaugle/
13. 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.
14. 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/
15. Vicon, http://www.vicon.com.
16. Max/MSP, http://www.cycling74.com.