

# The Virtual Mandala

J. Cecilia Wu, François Conti

Media Arts and Technology, University of California, Santa Barbara  
Robotics Laboratory, Stanford University  
U.S.A.

jcw@umail.ucsb.edu  
conti@cs.stanford.edu

## Abstract

This paper presents the *Mandala*, a multimedia live performance piece that explores the use of body movement and virtual imaging to enrich musical expression. This project combines an interactive 3D bimanual interface with a computer imaging application designed to graphically render and physically simulate the interactive construction and deconstruction of the ancient Tibetan Buddhist sand arts, called *Mandala*. *Mandala*, an agent of ancient arts and philosophy, uses technology to lure the next generation into traditional cultural practices, disrupting the boundaries between ancient arts and leading-edge interactive imaging technologies.

## Keywords

Tibetan mandala sand arts, multimedia live performance, CHAI3D, body motion, human computer interaction, augmented-reality, dynamic simulation.

## Introduction

Human beings are deeply expressive. We use body language, and speech to communicate our thoughts and feelings to other individuals. During musical performances, audiences often connect to the performer and mentally model thoughts and feelings in the performers mind, in what is called emotional contagion (Hatfield et al., 1993). In addition, humans are highly skilled at reading bodies, and voices; this ability begins at infancy and is refined to an art by adulthood (Darwin, 1872). Body motion and gestures are natural artistic expression forms worthy of further study and implementation in live performance. The aspiration of the Mandala project is to capture the natural forms of human expression and embody them artistically in real-time by bringing an ancient art form of Tibetan Buddhism to the digital world.

As new technologies have appeared, inventors and artists have been driven to apply new concepts and ideas to improving multimedia live performance or creating entirely new means of controlling and generating dynamic sound and graphics. In the past two decades, with the explosion of digital technology, the computer has become an essential tool for creating, manipulating, and analyzing audio and visuals. Its precision, high computational performances and capacity to generate or simulate audio-visual signal make it a compelling platform for artistic expression and

experimentation. Moreover, the availability of many new forms of human interfaces, ranging from camera trackers to tactile displays, offers an infinite spectrum of possibilities for developing new techniques that map human gestures on sophisticated audio-visual events. Although many approaches have been proposed to use body movement to drive interactive arts and musical instruments (Hunt et al., 2000; Knapp and Cook, 2005; Machover and Chung, 1989; Overholt, 2001; Wanderley et al., 2001; Winkler, 1998), to the best of our knowledge, few studies have explored the use of vocalists' digitalized body motion to simultaneously control real-time dynamic simulation and to enhance musical expression during multimedia live performance.

## Related work

Previous work in movement tracking can be categorized in two groups: movement sensor-based and position tracking based. Movement sensor-based recognition is most suitable for mobile applications in which the user may traverse large workspaces. These systems typically rely on small accelerometers or joint sensors that capture the movements of the human body. Such systems have been mounted on instruments and used for musical performance conducting systems (Rekimoto, 2001). Furthermore, wearable devices such as glove-based interfaces designed for capturing sign language (Hideyuki and Hashimoto, 1998) have demonstrated the capacity to sense even subtle finger motions. Tsukuda and Yasumura (2002) introduced the Ubi-finger, a wearable device that uses acceleration, touch and bend sensors to capture a fixed set of hand gestures.

In contrast, position tracking based devices are used for stationary applications and generally require a tracker or camera to be set up, and often calibrated, at a fixed location. Iason et al. (2011) present different strategies for recovering and tracking the 3D position, orientation and full articulation of a human hand from marker-less visual observations obtained by 3D Kinect sensor. More advanced multi camera systems such as the Vicon 8 have also been used to capture human motion during large stage performances where optical occlusions may occur (Dobrian and Bevilacqua, 2003).

In the area of musical expression, systems for interactive control of sound and visuals in virtual reality can also be traced back to 1990s. Meskers et al. (1999) used the electromagnetic to track the "Flock of Birds" in shoulder motion research. Robyn et al. (2006) further explore the use of motion tracking techniques to extract musical data to gen-

erate responsive imagery. Through a series of his own work, Pritchard (2010) theoretically analyzed “the body as object”, and how body, body motion and gesture can be used in both musical and visual expressions. Our research attempts to expand upon the musical applications of these technologies as well as the theoretical implications of gesture control in live, multimedia performance.

## The Mandala

In Tibet, as part of a spiritual practice, monks create mandalas with colored sand (Bryant, 2003). The creation of a sand mandala may require days or weeks to complete. When finished, the monks gather in a colorful ceremony, chanting in deep tones (Tibetan throat singing) as they sweep their Mandala sand spirally back into nature. This symbolizes the impermanence of life.

In this project we strive to create an intriguing work that translates this ancient philosophy to an arts and cultural experience as spiritual expression through interactive multimedia art and music composition. As the engineer Richard Hamming pointed out (1997, cited by Roads 2015), it is insignificant if the novelty and innovation make no cultural impact on human history. Through this piece, we innovatively provide a quick, yet rich presentation to simulate the Tibetan Mandala sand arts with musical expression; we aesthetically explore the ancient Buddhist philosophy of impermanence; and we increase the visibility of traditional cultural practices to the “twitter generation”, who has less patience to rediscover those transcendent human treasures.



Figure 1: Construction and destruction of a sand Mandala.

## The Piece and the Performances

As part of a multimedia live performance created at Stanford University and an invited performance at the Third International Buddhist Youth Conference in the Hong Kong Institute of Education, an interactive display of sand Mandala was dynamically simulated and choreographed with the vocal and instrumental compositional-improvisation. The musical and visual expressions through body motion were further explored and documented. The first performance at Stanford University took place before an audience of engineers and young Silicon Valley entrepreneurs, while the second performance occurred in Hong Kong, before an eastern, Buddhist audience. Both audiences showed enthusiasm for the culture content, and the sci-

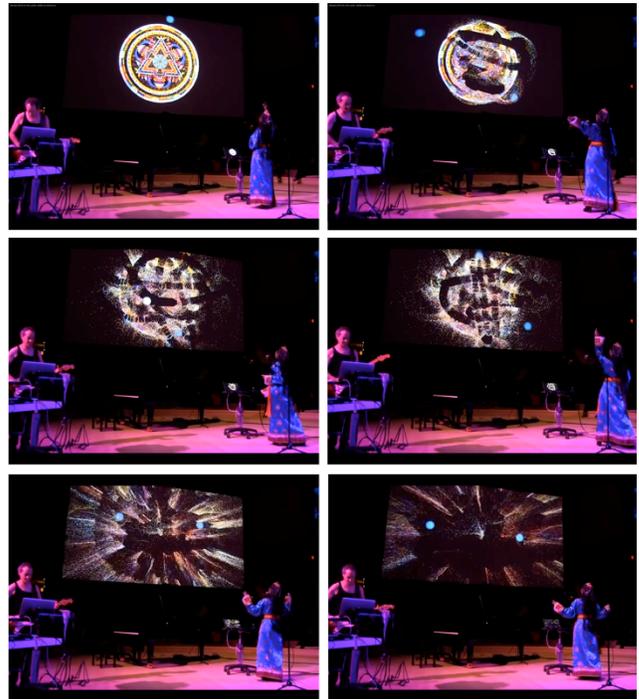


Figure 2: Live performance of the *Mandala*.

ence and technologies behind the scene that creates the magic, both during and after the performances.

The proposed system is composed of two applications that operate on a single laptop computer and handle both the motion tracking and audio-visual processing events. The motion tracking application connects to the Hydra Razer interface and acquires the position, orientation, velocity and acceleration values of each input device at 100 times per second. Additionally, the states of the input switch and configuration of the miniature joysticks mounted at the extremities of both end-effectors are also retrieved. These measurements are then converted into events that are used to interact with the dynamic model of the virtual Mandala and real-time audio processing. The different parts of the systems are further developed in the following paragraphs.

The piece *Mandala* includes three movements: *Construction*, *Climax*, and *Destruction*. At the beginning of the piece, hand motions from the vocalist were directed to initiate the *Construction* of the virtual Mandala. Music composition for this part includes implementing additive synthesis and FM synthesis techniques (Chowning, 1977) to create an ocean-wave-like musical landscape with dynamic high pitch notes moving through the body movement. The backbone of the piece is made of rhythmic loops in Logic Pro with outputs and lives guitar processor inputs, routed through a Mark of the Unicorn 828mkII interface for 10 channels of spatialized output. Live guitar sounds are generated by a Fractal Audio AxeFX Ultra processor.

Mixed sounds of electronic guitar, rhythmic drumming, Buddhist chanting, synthesized sound waves and cello start to create and build up an imaginary world from tran-

quility and emptiness to intense passion and complex glamour. During the *Climax*, musicians were focusing on musical improvisation while the virtual Mandala was self-directed and presenting animation, such as rotation, twisting and glitter effects. Granular synthesis, along with audio panning and spatialization techniques, are implemented to create bursts of variously-pitched pieces of sound fractions which, when put into interaction with experimental low-tone throat singing and explosions of violent soprano interjections, comes to symbolize the extreme conflicts and ecstasy in life, the extraordinary highs and lows, the sharp rises and falls of life as it is experienced, ending with a sharp fade out into the deconstruction of the mandala. The *Destruction*, which underscores the eternal nature of impermanence, occurs near the coda of the audio-visual piece through the hand motions from the singer.

A video of this multimedia live performance piece's premier can be viewed at: (<http://tinyurl.com/ku5cg4f>)

### 3D Human Interface

The vast majority of motion-controllers available today rely on small accelerometers that detect hand motions generated by the operator. Although such inertial sensors are used in almost any game controller or smart phone available today, their precision remains limited, and they lack the ability to track absolute positions in 3D space. For this project we selected a new type of interface called the Razer Hydra. As compared to accelerometer based, ultrasonic or camera driven motion sensors, the Razer Hydra uses magnetic motion sensing. Magnetic motion sensing enables this new interface to track in true 3D space for absolute position and orientation with its magnetic field. This technology also allows for full six degree-of-freedom movements tracking with sub millimeter accuracy, while eliminating the need for a line of sight to operate.

The user interface adopts a paired system, with a base station emitting a low energy magnetic field that both hand controllers record to estimate their precise position and orientation. The controllers and the base station each contain three magnetic coils. These coils work in tandem with the amplification circuitry, digital signal processor and positioning algorithm to translate field data into position and orientation data. The raw data are then processed using a first order filter and converted to a MIDI protocol before being broadcast to the audio stage.

The audio processing application, designed using the ChucK language (Wang and Cook, 2003, 2008), interprets MIDI signals from the motion-tracking module and activates a selection of audio filters and generators according to the desired musical expression. The application communicates all input and output audio signals by using the Unicorn 828mkII interface, which supports multiple channels that include both voice and electronic instruments.

Because of the real-time nature of the technology, the interface delivers near instant response between the hand motions generated by the operator and the actions commanded by the computer. An illustration of the device is presented in Figure 3.



Figure 3: The bimanual Razer Hydra interface

### Dynamic Simulation

The virtual Mandala is composed of 4.5 million digitized sand particles that are physically modeled and updated 100 times a second. Force interactions are modeled in real-time by capturing the position of the performer's hands and by computing collision events with and between the individual sand particles. A graphic and dynamic modeling framework CHAI3D (Conti, 2003) is used to simulate in real-time the physical interaction between the captured hand motions of the performer and the small mass particles composing the virtual Mandala. The mapping between the hands of the performer and their models in the simulation is performed using a new mapping technique (Conti and Khatib, 2005) called *workspace expansion* that relies on progressively relocating the physical workspace of the devices mapped inside of the virtual environment towards the operator's area of activity without disturbing his or her perception of the environment. This approach allows the operator to interact with high accuracy in a very large virtual environment while moving his or her hands within a reasonably small workspace.

### Conclusion

In this paper, we presented a human interactive simulation framework to explore the imaginative relationships between singing, dancing and natural emotional expression by using gestural and motion control to modulate and shape the texture and contour of sound and visuals. By demonstrating the motivation, compositional structure and methodology to achieve our aesthetics and technical goals of the multimedia live performance piece *Mandala*, we combined state-of-the-art technology in simulation and human interaction with ancient eastern philosophy and Tibetan Buddhist arts. Repeated performances of this piece with multiple audiences have confirmed not only the successful application of this technology, but also the ability of this technology to work with music to bring to life ancient art forms and philosophies for worldwide modern audiences. These preliminary results validated the intuitive design of our application and its capacity to respond with real-time audio-visual performances and confirmed our original idea toward exploring natural human motion and gesture as an effective way to instinctively enrich musical and artistic expression.

## Acknowledgements

We thank Chris Chafe, Jay Kadis, Curtis Roads, JoAnn Kuchera-Morin, Matt Wright, Marcos Novak, and George Legrady for their collaboration and support.

## References

1. Bryant, Barry. *The wheel of time sand mandala: Visual scripture of Tibetan Buddhism*. Snow Lion Publications, 2003.
2. Chowing, John M. "The synthesis of complex audio spectra by means of frequency modulation." *Computer Music Journal* (1977): 46-54
3. Conti, François. *The CHAI libraries*. No. LSRO2-CONF-2003-003. 2003.
4. Conti, François, and Oussama Khatib. "Spanning large workspaces using small haptic devices." In *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005*. World Haptics 2005. First Joint, pp. 183-188. IEEE, 2005.
5. Darwin, Charles. *The expression of the emotions in man and animals*. Oxford University Press, 1998.
6. Dobrian, Christopher, and Frédéric Bevilacqua. "Gestural control of music: using the vicon 8 motion capture system." In *Proceedings of the 2003 conference on New interfaces for musical expression*, pp. 161-163. National University of Singapore, 2003.
7. Gabriëlsson, Alf, and Patrik N. Juslin. "Emotional expression in music performance: Between the performer's intention and the listener's experience." *Psychology of music* 24, no. 1 (1996): 68-91.
8. Hamming, Richard R. *Art of Doing Science and Engineering: Learning to Learn*. CRC Press, 1997.
9. Hatfield, Elaine, and John T. Cacioppo. *Emotional contagion*. Cambridge university press, 1994.
10. Sawada, Hideyuki, and Shuji Hashimoto. "Gesture recognition using an acceleration sensor and its application to musical performance control." *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)* 80, no. 5 (1997): 9-17.
11. Hunt, Andy, Marcelo Wanderley, and Ross Kirk. "Towards a model for instrumental mapping in expert musical interaction." In *Proceedings of the 2000 International Computer Music Conference*, pp. 209-212. 2000.
12. Knapp, R. Benjamin, and Perry R. Cook. "The integral music controller: introducing a direct emotional interface to gestural control of sound synthesis." In *Proceedings of the International Computer Music Conference (ICMC)*, pp. 4-9. 2005.
13. Machover, T., and J. Chung. "HyperInstruments: Musically Intelligent and Interactive performance and Creativity Systems. Intl." In *Computer Music Conference (Columbus, Ohio, 1989)*.
14. Meskers C G M, Fraterman H, Van der Helm F C T, et al. Calibration of the "Flock of Birds" electromagnetic tracking device and its application in shoulder motion studies[J]. *Journal of biomechanics*, 1999, 32(6): 629-633.
15. Oikonomidis, I. Kyriazis, N. Argyros, and A. Efficient Model-based. "3D Tracking of Hand Articulations using Kinect." In *Proceedings of the British Machine Vision Conference 2011, Dundee UK*, pp. 101-1.
16. Overholt, Dan. "The MATRIX: a novel controller for musical expression." In *Proceedings of the 2001 conference on New interfaces for musical expression*, pp. 1-4. National University of Singapore, 2001.
17. Pritchard, Robert. "The Body as Object: From Body Image to Meta-Body." *Critical Studies* 33, no. 1 (2010): 203-215.
18. Rekimoto, Jun. "Gesturewrist and gesturepad: Unobtrusive wearable interaction devices." In *Wearable Computers, 2001. Proceedings. Fifth International Symposium on*, pp. 21-27. IEEE, 2001.
19. Roads, Curits. *Composing Electronic Music: A New Aesthetic*. New York: Oxford Press. 2015
20. Taylor, Robyn, Pierre Boulanger, and Daniel Torres. "Real-Time Music Visualization Using Responsive Imagery." In *8th International Conference on Virtual Reality*, pp. 26-30. 2006.
21. Tsukadaa, Koji, and Michiaki Yasumurab. "Ubi-finger: Gesture input device for mobile use." In *UbiComp 2001 Informal Companion Proceedings*, p. 11. 2001.
22. Wanderley, Marcelo M. "Gestural control of music." In *International Workshop Human Supervision and Control in Engineering and Music*, pp. 632-644. 2001.
23. Wang, Ge, and Perry R. Adviser-Cook. *The chuck audio programming language. a strongly-timed and on-the-fly environment/mentality*. Princeton University, 2008.
24. Winkler, Todd. "Motion-sensing music: Artistic and technical challenges in two works for dance." In *Proceedings of the International Computer Music Conference*. 1998.

## Authors Biographies

**Jiayue Cecilia Wu:** Originally from Beijing, Jiayue Cecilia Wu is a scholar, composer, multimedia performer, and audio engineer. Cecilia earned her Bachelors of Science degree in Fashion Design and Engineering in 2000. Upon winning the MTV Asian Beat contest, Universal Music identified her talent and signed her as a music producer in Hong Kong. In 2010, Cecilia produced her original album of spiritual electronic music, *Clean Your Heart*. In 2013, Cecilia obtained her Master's degree in Music, Science and Technology at Center of Computer Research in Music and Acoustics at Stanford University, where she focused on computer-assisted composition and audio engineering. As an audio engineer, she received a graduate grant award from Audio Engineering Society. As a musician, she received an award from the California State Assembly for being a positive role model in sharing Chinese culture. Currently, she is pursuing her PhD in Media Arts and Technology at UC Santa Barbara.

**François Conti:** Originally from Switzerland, François Conti is a co-founder of Force Dimension, a company that has earned international recognition for designing and manufacturing high precision haptic interfaces that are used today in research, entertainment, and medical fields. François also leads the development of the CHAI3D open-source haptic visualization libraries and is a lecturer of the Experimental Haptics course at Stanford University. François holds a B.S., M.S. from the Swiss Federal Institute of Technology in Lausanne, and a PhD from Stanford University.