MELODY OF MOTION: INTERACTIVE SPATIAL COMPOSITION WITH 3D SENSOR AND AMBISONICS TECHNIQUE

HSIEH, Ching-Ying[1]

Graduate Institute of Architecture,
National Chiao Tung University,
1001 University Road, Hsinchu, Taiwan 300,
ROC
olifahsieh@arch.nctu.edu.tw

Prof. HOU, June-Hao[2]

Graduate Institute of Architecture,
National Chiao Tung University,
1001 University Road, Hsinchu, Taiwan 300,
ROC
jhou@arch.nctu.edu.tw

ABSTRACT

The relationship between body kinetics and the music creation is essential. For examples: the conductor gives a series of gestures when he conducts a symphony orchestra; every musician creates a diversity of marvellous sounds through differently gestures in act and gesticulation with the instrument. Nowadays, the application of sensors can more precisely capture body motions and their resulting data. By using the data to construct hyper-connections, the interactive relationship between the body and space will be further established. This essay is based on the author's sound installation *Melody of Motion (2019)*, created from a chain of interactive sound installations. It will illustrate how the spatial audio technique Ambisonic transmits the sounds and how 3D Sensor applies the body motion data to connect synchronously with immersive sound. In this 3D construction of Body/Sound, it does not only show how these two elements interplay intimately, but also unfolds an immersive soundscape experience for the auditor.

1. INTRODUCTION

*"*If by using sound, I can create a different audio perception through an individual's every movement within a finite space, this space can become infinite. I imagine that sound can be tactile like water, and it has the full potential to fluid and change without limits in space. I would like to simulate the sense of touch by using sounds..*"* - HSIEH, Ching-Ying

"聲洄 Melody of Motion" is a participatory interactive soundscape installation that combines the spatial audio technology of Ambisonic and 3D depth sensor technology. When audiences enter the exhibition space, they not only feel surrounded by sounds projected at them from all around, the sensors will generate real-time data according to their motions to further modify the dynamic soundscape of the space and acutely capture their movements to immediately create electronic sounds, forming real-time interaction between body and sound with a vivid sense of space. All the invisible sounds form an absolute closeness with the involvement of the body. Sounds, like water in this case, change and flow through the finite space in a soft and delicate way, composing an immersive sensory experience that is both audible and tactile in the seemingly infinite space.

Melody of Motion takes the image of the Mandarin Chinese character "洄" to represent the concept of water swirling, and the character as a pictograph as well displays the symbol of a transmitting dynamic of energy in motions of surrounding, whirling, flowing from or around an invisible center of an energy source. After the sound installation project, A Sound Journey at National Taichung Theater, sound artist Hsieh, Ching-Ying again applies the full-sphere surround sound format, Ambisonic, and combines it with the technology of 3D motion detection to set up Melody of Motion, the immersive sound installation. Stepping in the installation, the audience can feel they are surrounded by sounds coming from all directions. The changing coordination of their physical movements causes a series of data that promptly adjusts and changes the contents of sounds. The dynamic of movements creates immediate interactions between the participants' bodies and the sounds, so that all the invisible sounds forms a kind of intimacy with the engaging bodies. Therefore, sound becomes palpable and visible, moves like water softly and subtly, and creates an immersive experience of hearing and touching in the unlimited, fluid space.

2. INSPIRATIONS BY LÉON THEREMIN

In 1919, Russian scientist Léon Theremin invented Theremin, a musical instrument that produces sounds and performs music by using different levels of pitch and volume from interactions between weak electromagnetic fields around its antenna and human static electricity. It is the world's first electronic musical instrument that applies spatial sensor data from body movements to produce sounds. In other words, the Theremin is a "virtual musical instrument"[1] that can produce sounds without requiring any "physical contacts" between the body and other objects.

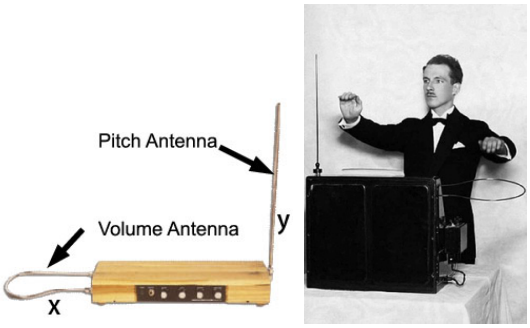

Figure 1: Léon Theremin demonstrates playing the Theremin as its inventor.

Under the influences of his dancer wife, Theremin further transformed the Theremin from a horizontal configuration into a vertical one in 1932. He named it Terpsitone. In addition to producing music without requiring any physical contacts, Terpsitone controls the sound pitch and volume through whole body movements. This breaks through the

limitation of Theremin to only produce music by using the upper body (arms and hands). The invention of Terpsitone not only inspired electronic musical instruments to generate sounds by dance moves, but it also created the music performance style of "dance vocalization."

Figure 2: The transformation from Theremin to whole-body Terpsitone performance by Léon Theremin

Terpsitone was designed to control pitch by body height, and control volume by the distance between the body and instrument. The body must be "within a specified range" to perform music, and the dancer is unable to walk or move freely. Therefore, the design remains a 2D (x-axis + y-axis) sensor interface mechanism.

Since the control mechanism is "invisible" and "untouchable," playing the Theremin and Terpsitone are highly difficult. In comparison with typical musical instruments, the Theremin is lacking the reference to "body touch" and "visual positioning." Players must experiment with specific pitches and volume in mid-air and they must possess highly sensitive motor control abilities to achieve the expected performance effects. This is why only a handful of musicians know how to play even to this day[2]. Furthermore, realizing the artistic expression of dance vocalization requires the body to move freely within the space. To the author, performing vocalized dance in empty space is a very interesting concept. It led the author to wonder whether more sensitive spatial sensors can be applied to extract kinetic data—not only from the upper body but from the entire body—to design a method of producing electroacoustic sounds in 3D space[3] by expanding the sensing range and evolving from 2D to 3D (x-axis + y-axis à x-axis + y-axis + z-axis). Can we develop a set of virtual musical instrument installation that combines immersive audio experiences with spatial interactions? In close connections with body movements by enabling the dancer to freely move his/her body within the space, and combining it with the surround audio playback of Ambisonic audio technique? Moreover, with respect to the original lack of visual guidance, can symbolic instructions that facilitate bodily connections be designed for body movements to produce sounds more precisely? The experimental and research objective of this creative project is to combine 3D spatial sensing with Ambisonic technique to construct a special, immersive interactive space experience.

3. MELODY OF MOTION - AN IMMSIVE SOUNDSCAPE INSTALLATION

In the Chinese title of this sound installation work, the word "Hui" has a meaning of swirling, flowing water. The Chinese character also represents an imagery of energy flowing outward from the center to form dynamic motions of enclosure, scattering, spiraling, and flowing. *Melody of Motion* mainly applies spatial audio and 3D sensing techniques to construct an invisible sensory experiential space. By using 3D sensors above the center of the space to extract bodily data, and setting up the content of electroacoustic sounds activated by moving body positions and styles of motion, an immersive sensory experience of interactive sounds[4] can be constructed, whereby sounds are affected by body movements just like water.

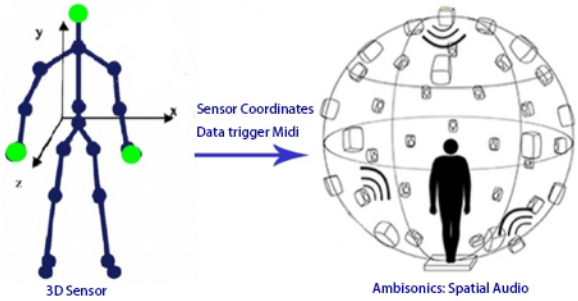

Figure 3: A combination of dynamic 3D sensing and Ambisonic immersive audio techniques.

4. AN IMMERSIVE CONSTRUCTION OF INTERACTIVE SPACE

The 3D Depth Sensor in this creative project (*Melody of Motion)* was sponsored by Taiwan's FocusVision. The advanced sensing technology allowed this project to achieve a level of unprecedented expressiveness. The chart shows the difference of data transmission between Kinect and FocusVision production.

Kinect V2		FVT ToF sensor	
Items	**Specifications**	**Items**	**Specifications**
Sensor type	TOF + RGB	Sensor type	ToF + RGB
Measure Range	Up to 8m	Measure Range	Up to 14m
Measurement accuracy	1%	Measurement accuracy	1%-5%
Depth resolutions & frame rate	512 x 424@30fps	Depth resolutions & frame rate	320 x 240 @ 120 fps
Dynamic standard deviation	HD	Dynamic standard deviation	HD
EEL (Laser) #/LED#	3EEL	EEL (Laser) #/LED#	4 EEL
Power consumption	16W	Power consumption	Max. 5W

Figure 4: A Data Comparison between the Two Sensor Technologies: KINECT and FOCUSVISION

These sensors have the following advantages in comparison with Kinect 3D sensors:

1. **FOCUSVISION** has a minimum sensing distance of 15CM and a maximum sensing distance of 12M. **KINECT** has a range between 50CM~8M.
2. **FOCUSVISION** has a reading rate of 120 FPS, and **KINECT** has a reading rate of 30 FPS.

3. Open source codes are not available for **KINECT** SDK, while they are available for **FOCUSVISION**.
4. **FOCUSVISION** does not require any parts of the body to be scanned in advance.
5. **FOCUSVISION** allows multiple spatial blending.

In order to create different sound designs in a limited space with multi-channels audio, the project started by suspending the sensors in the center of the space. From the top to bottom, the sensors produced a rectangular sensing area. The 3D depth data and flat address of any objects within this rectangular sensing area were extracted at the frequency of 1/30sec.

Furthermore, the most basic requirements for the sound-producing "musical instrument" include: 1) whether the mechanical design of each part meets ergonomic requirements; 2) whether the instrument can precisely produce quality tones and pitches under synchronous operation with the body. Therefore, the body interaction data in *Melody of Motion* were gathered by testing a real body in the sensing area; the data were reviewed and filtered after connections to electroacoustic sounds were made. After multiple experiments, spatial data for the body position, body height, and respective hand positions were selected to be the activation criteria for specific electroacoustic sounds. Then, a spectrum of corresponding tones and pitches were further designed within this range of motion. This was the basic structural method for spatial composition that was used to create unique sound content.

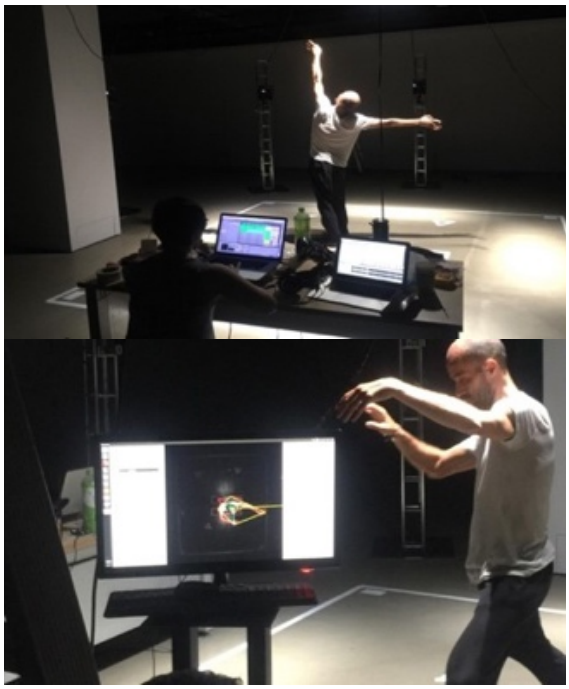

Figure 5. Experimenting with connections between body data and sound composition.

5. GRAPHICAL NOTATION DESIGN FOR 3D SPATIAL COMPOSITION

Since Theremin/Terpsitone playing lacks a visual benchmark and sense of touch to determine specific high pitches and sound volume, different geometric shapes were set up on the floor to signal the activation of specific sounds within the sensing area (Figure 6) during the construction of *Melody of Motion's* interactive 3D space. This does not only provide a reference for the audience when they enter into the sound installation, but it also creates a unique acoustic structure for each personalized path. Taking water as an example, it is analogous to moving in a pool to feel different water temperatures and ripples. In the center-right area marked by a square within a square is where body data synchronously produce electroacoustic sounds. Any acute motion can produce sounds that form a variety of rippling effects as they connect with the body just like water. This creates an immersive spatial audio experience that is covered by interactive sounds. Below please find a real scene of the sound installation and its graphical notation design on the floor:

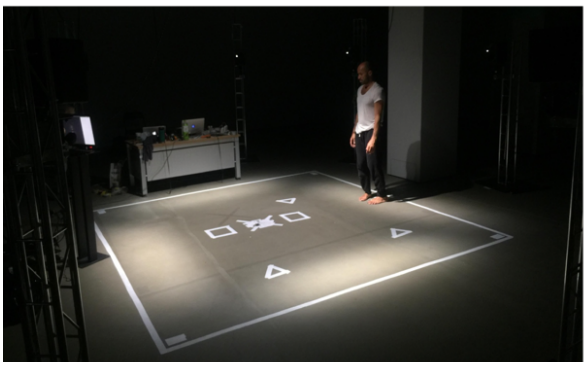

Figure 6: A real scene of the sound installation.

With the above design, every player will be closely tracked by sensors when he or she enters the space and makes body movements that activate present electroacoustic sounds within the space. In addition to being surrounded by immersive sounds, data from the player's body movements and positions immediately change the content of the sounds. This is further incorporated into the participant's dynamic body data, producing immediate interactions between the body and the sounds. The setup described above allows sounds to be affected by body positions and motions in real time, which breaks through typical designs of using timeline-based composition structures. In this way, all the invisible sounds form an absolutely intimate relationship with the body. Sounds become touchable, visible, and versatile, soft, delicate like water. This creates an audible and touchable immersive sensory experience under an infinite space of the moment.

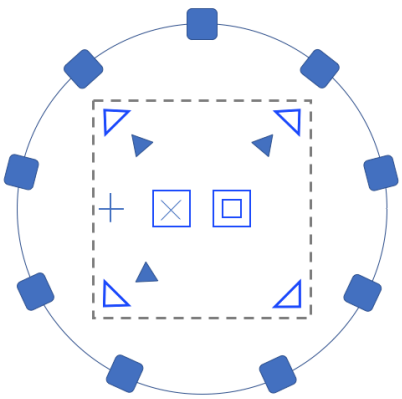

Figure 7: The graphical notation design on the floor.

Figure 8: The explanation of sound symbols.

6. BODY-SOUND INTEGRATION: A SOUND PERFORMANCE THROUGH DANCE

As dancers usually exercise their bodies based on the requirements of visual expressions, most of their movements are not linked to the sound content when they first enter into the sound installation's sensing area. Dancers would produce chaotic sounds in the space if they perform a free dance in their original styles. After several trainings, Mauro Sacchi was constantly reminded that specific sounds would be activated by any movements and body positions. The body must get rid of its original dancing method, and the dancer should explore the space by starting with the simplest movements while "listening to" and "remembering" the movements and their consequent sounds.

Figure 9: Dancer Mauro Sacchi experiencing the expression of body-sound integration.

Just like in learning any musical instruments, the first step is to understand and observe the instrument's capability. After trying out different kinds of movements to produce the best musical sounds, the body is ready to move by acting as a medium between the brain and musical instrument. All body movements are based on the goal of producing sounds instead of mere visual expressions. After much practice, the dancer gradually eliminated unnecessary movements and developed a sound-based way of dancing after feeling the sounds from different movements in various surface positions and longitudinal motions. Not only did the final outcome express the artistic aesthetic of a visual (body performance) and audio (electroacoustic sounds) integration, but it also expanded the sound installation's limit of only allowing one player at a time. In other words, the audience can enter the sensing area to experience the sound field that is tailor-made for them, as well as watch other players' body expressions and interact with the sound installation from outside the sensing area. This adds a level of "participating the player's experience" to this art work.

7. CONCLUSION

Under the ever-evolving and rapid developments in technological interactive applications, connections between the human body and space have transitioned from a basis on tangible materials to intangible links of "digital virtualization." Through accelerated real-time transmission of the wireless Internet, different methods of communication in empty space are produced to break previous "spatial" limits defined by visual materials, and thereby achieving far-reaching innovative applications.

This creation of *Melody of Motion* attempted to combine interactive technology with spatial audio technique Ambisonic. It also further explored the possibilities between space and body. In addition to sound content, other versatile, abstract materials such as light, color, and temperature can also be processed in the future by connecting body movements to symbols. This creates an opportunity to introduce a greater variety of abstract sensory elements into the work and elaborate on creative imaginations about the performance-based architecture. Recently, the concept of this work was further applied to advanced creative experiments in virtual reality space, with the hope of creating more diverse content applications in the future and expanding the possibilities between spatial perception and interaction.

8. ACKNOWLEDGMENT

This research was mainly supported by Digital Art Center, Taipei(DAC), which offered us a great exhibition space to perform our art work. We also like to show our gratitude to our sponsors MidiMall. Inc. and FocusVision Tech., both of them provided hi-tech equipment and technical informations to help us made this imagination come true. And also thank to Chiyou Dean that greatly assisted the research and Prof. Jun-Hao Hou for comments that greatly improved the manuscript from the beginning.

9. REFERENCES

[1]. S. M. Martin, *Theremin - An Electronic Odyssey*. 1995 (documentary movie)

[2]. O. Mattis, *An Interview with Leon Theremin.* In *Thereminvox*. 1989, Retrieved August 10, 2019 from http://thereminvox.com/story/495/

[3]. L. Wang, T. Meydan, and P. Williams, "A Two-Axis Goniometric Sensor for Tracking Finger Motion," *Sensors*., vol. 17, no. 4, pii: E770, April 2017.

[4]. C. Y. Hsieh, "Melody of Motion," In *Olifa Hsieh Works.* Retrieved August 10, 2019 from http://olifahsieh.net/melodyofmotion

[5]. W. B. Kleijn, "Directional Emphasis In Ambisonics," *IEEE Signal Processing Letters.*, vol. 25, no. 7, pp. 1079-1083, March 2018.

[6]. D. Michel, C. Panagiotakis, and A. A. Argyros, "Tracking the articulated motion of the human body with two RGBD cameras," *Machine Vision and Applications*., vol. 26, no. 1, pp. 41-54, December 2014.