Appeared In Proceedings of Audio Mostly 2008 – a Conference on Interaction with Sound, pp. 123-127.

Audio Interface as a Device for Physical Computing

Kazuhiro Jo, RCAST, University of Tokyo / Culture Lab, Newcastle University, jo@jp.org

Abstract. In this paper, we would like to describe the employment of audio interface as a device for physical computing. We compare the audio interface with other devices and describe its characteristics. We also present examples of the employment with three different art works, Monalisa "shadow of the sound", The SINE WAVE ORCHESTRA stay amplified, and AEO. We explain the implementation of each work with different physical components. Finally, we discuss some of the potential of the employment of audio interface for future implementation.

1. Introduction

An audio interface is a device that allows a computer to receive/send audio signal from/to the outside. It converts a stream of bits (digital) into a time-varying voltage (analog) and vice versa. Most of computers currently have a built-in audio interface. The interface generally provides two (stereo) built-in input/output channels with 44.1KHz sampling rate and 16-bit sampling resolution. People employ such audio interface for listening or recording audio in the field of music, film, and/or gaming. Some of external audio interface serves more input/output channels, higher sampling rate, and higher sampling resolution (e.g. 56-ch, 192KHz, 24-bit).

Physical computing is an approach to sensing and controlling the physical world with computers [8]. People create a conversation between the physical world and the virtual world of the computer by employing not only standard computer interface (e.g. keyboard, mouse, display, speaker) but also several kinds of physical components (e.g. sensor, switch, actuator). The conversation comes into sight as a form of musical controller, gaming console, interactive installation etc.

To realize such conversation, people have been developed several input/output devices, such as Infusion I-CubeX (http://infusionsystems.com/), Arduino (http://www.arduino.cc/), and MAKE Controller Kit (http://www.makezine.com/controller/). These devices allow a computer to communicate with the physical world by employing physical components. The devices convert the signals from/to the components with several protocols. I-CubeX employs MIDI, arduino employs serial, and MAKE Controller Kit employs OSC (Open Sound Control). The programming environment such as Cycling74 MaxMSP (http://www.cycling74.com/), Adobe Flash (http://www.adobe.com/), and Processing (http://www.processing.com/) manages multiple input/output signals from/to the physical world and maps them into other types of representation (e.g. graphic, text, sound).

In this paper, we would like to propose the employment of audio interface as a device for physical computing. It allows a computer to communicate with the physical world by audio. The signals from/to the components in the physical world are treated as audio and processed on the programming environments in the computer.

Audio interface is unique to compare with the other devices, I-CubeX, Arduino, and MAKE Controller Kit, at least in three points: attachment, sampling resolution, and sampling rate (Table 1).

Table 1. Comparison of devices
(specifications are taken from the website)

	I-CubeX	Arduino	MAKE Controller Kit	Audio Interface
Protocol	MIDI	Serial	OSC	Audio
Attachment	External	External	External	Internal/ External
Sampling Resolution	12-bit	10-bit	10-bit	16bit to 24bit
Sampling Rate	244Hz	14.4KHz (112500bps)	1KHz	44.1KHz to 192KHz

Attachment:
Most of computers currently have a built-in audio interface with two (stereo) input/output channels. It enables people to connect with several components (e.g. headphone, microphone, speaker) without additional input/output equipments. External audio interface offers extra input/output channels.

Sampling resolution:
Sampling resolution is a measure of how preciously the device treats the quantity of object. The audio interface basically offers 16-bit sampling resolution, while the other devices provide 10 or 12-bit resolution. Some external audio interface has higher resolution such as 20-bit or 24-bit.

Sampling rate:
Sampling rate is a measure of how frequently the device looks at the input from the outside. Most of the audio interfaces support 44.1KHz or more sampling rates because of the sense of hearing (about 20Hz to 20KHz). Owing to the limitation of the protocol or the hardware, other devices offer lower sampling rate than the audio interface.

2. Implementation

We have employed the audio interface as a device for physical computing with three different art works, Monalisa "shadow of the sound", The SINE WAVE ORCHESTRA stay amplified, and AEO. In each work, we employed different physical components with the audio interface: a push button for Monalisa "shadow of the sound", a fader controller and a foot switch for The SINE WAVE ORCHESTRA stay amplified, and two accelerometers, a set of distance sensor, and two custom built converter boxes for AEO. In following sections, we explain the implementation of each work in both hardware and software settings.

2.1. Monalisa "shadow of the sound"

Monalisa "shadow of the sound" is an installation created with Norihisa Nagano. It is a part of software platform Monalisa which enables to "see the sound, hear the image" by treating all the image and the sound as the sequence of numbers represented by binary codes [7]. The work was premiered at Open Space at NTT InterCommunication Center from 9th June 2006 to 11th March 2007. The work consists of a set of computers with custom version of Monalisa applications, projector, camera, microphone, push button, and speaker situated in a room (Figure 1). We briefly describe the procedure of the work below.

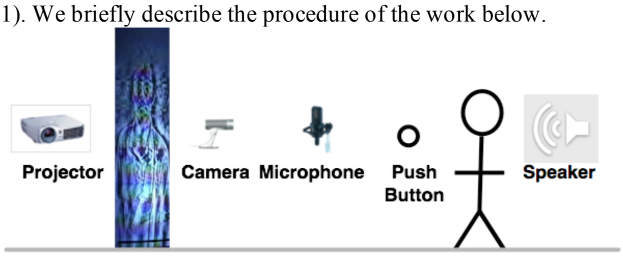

Figure 1: Monalisa "shadow of the sound"

When entering the room, each participant saw his/her image, caught with the camera, projected on the screen. When he/she pushed the push button, the image was captured as a static bitmap image data. The image data was transformed to the application and automatically played as a stream of sound through the speaker. The microphone simultaneously captured the stream of sound and transformed it to the application. The application sequentially re-projected the incoming sound data as an image on the screen from top left to down right of the screen. Because of the acoustics of the room, the re-projected image was blurred like a shadow with reverberations of the sound.

2.1.1. Push button

We employed a push button and a set of input/output channel of a built-in audio interface to receive the action from participants. The push button has two terminals. The terminals conduct with the push of the button. To detect the conduction with the audio interface, we cut a standard audio into two fragments. Each fragment of the cable is attached to each terminal of the button. Each other side of the fragments connected to the input/output channel of the audio interface (Figure 2).

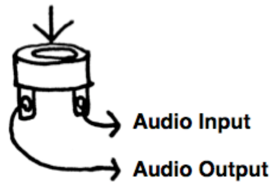

Figure 2: Push button

We detected the conduction of the button as audio signal on MaxMSP. We assigned a sine wave for the output channel. We set up a gate for the input channel. When the terminals conduct with a push of a participant, the gate identifies an incoming audio signal. With the identification, the application captures the image of the participant.

2.2. The SINE WAVE ORCHESTRA stay amplified

The SINE WAVE ORCHESTRA stay amplified is a work which was premiered at WAVES exhibition at the Exhibition Hall ARSENALS of the Latvian National Museum of Art from 25th August to 17th September 2006 [3]. It is a work of a participatory sound performance project The SINE WAVE ORCHESTRA (http://swo.jp). The author serves as the one of four core organizers of the project. In the project, under the basic concept that each participant plays a sine wave, people are invited to create a sea of sine waves as a collective sound representation. Each of the work in the project employed the different instrument with different temporal, physical, environmental, and procedural settings [6]. This work consists of a set of control computer, sound synthesis computer, light, foot switch, fader controller, rotational controller, and 16 speakers circularly situated in a room (Figure 3). We briefly describe the procedure of the work below.

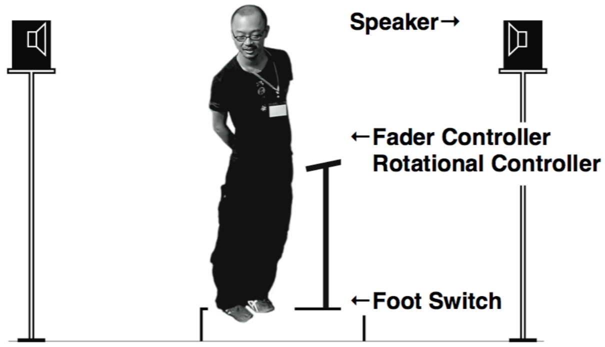

Figure 3: The SINE WAVE ORCHESTRA stay amplified

When entering the room, each participant was exposed to a collective sound representation. The sound representation consisted of sine waves that previous participants had produced. 16 speakers that horizontally encircled the participant produced the sounds. When a participant stood on the foot switch in front of the controllers, a new sine wave started to play. As he/she moved the fader controller, the frequency of the sine wave was changed. As he/she rotated the rotational controller, the sound source position (i.e. speaker) of the sine wave was changed. He/she selected the frequency and the sound source position of the sine wave. By leaving from the food switch, he/she left his/her own sine wave as a part of collective sound representation. The volume of the sine gradually attenuated and disappeared after the exhibition period. As more participants left their sine waves, more sounds were accumulated. During the exhibition period, the collective sound representation was changing from a phase where each sine wave was discriminable to a cluster that consisted of mutually interfering sine waves like a white noise that contains all frequencies.

2.2.1. Foot switch, Fader controller

We employed a foot switch, a fader controller, a rotational controller, and two sets (stereo) of input/output channels of a built-in audio interface to receive the action from participants. The foot switch and the fader controller are connected with a control computer with the audio interface and the rotational controller is directly connected with the control computer through USB. The foot switch has two terminals. The terminals conduct with the push of the foot switch. The fader controller also has two terminals. It behaves as a variable resistor. The move of the fader changes the value of the resister. The foot switch and the fader controller are separately connected with two sets of input/output channels of a built-in audio interface in the same manner of the push button of Monalisa "shadow of sound" (Figure 4).

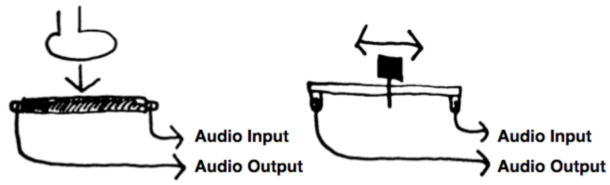

Figure 4: Foot switch (left) and Fader controller (right)

We detected the conduction of the foot switch and the resistance of the fader controller as audio signals on MaxMSP. We assigned a sine wave for each output channel. For the input channel of the foot switch, we set up a gate to identify the conduction in the same manner of the push button of Monalisa "shadow of sound". For the input channel of the fader controller, we set up a volume calculator. It reports the change of the resistance of the fader as the change of volume of incoming audio signal. With the change of the volume, the control computer sends a message for the synthesis computer engine to change the frequency of a sine wave. The sampling resolution for the volume is 16-bit (i.e. 65535 level). Therefore it is sufficient to accomplish subtle changes of the frequency of the sine wave.

2.3. AEO

AEO is a sound performance project consisting of three members: Eye, Taeji Sawai and the author. AEO has performed at several international festivals (e.g. Dutch Electronic Art Festival 2004, Radar 7 at 24 Festival de MEXICO 2008). In the project, each member takes one of three roles: performance (Eye), sound design (Sawai) and instrument design (the author). During AEO performances, the performer holds the instrument in each hand and shakes, sways, or swings them. These movements by the performer produce patterns of sound and light through devices and a computer. The instrument has undergone a transition in function and forms over six iterations [5].

In the latest iteration for the performance at Fujirock Festival 2008 (http://www.fujirockfestival.com/) on 27th July 2008, the instruments consist of two sets of hand-held plastic spheres. Each sphere contains a three-axis accelerometer, a part of distance sensor, a small light bulb, and a switch. One sphere has a transmitter of the distance sensor, while another has a receiver. The accelerometer converts the inclination and acceleration of each axis (i.e. x, y, z) into changes in analog voltage. The distance sensor measures the distance between two spheres by employing ultrasonic sine wave (40KHz) with the transmitter and the receiver (Figure 5). The small light bulb changes its brightness by the change of analog voltage. The switch changes the represented patterns of sound.

The device for the instrument was changed in each iteration. We employed I-CubeX for the first and second iteration, and MakingThings Teleo (http://www.makingthings.com/teleo/) for the third and the fourth iteration. In the fifth iteration, we developed two sets of converter boxes for the accelerometer and the distance sensor. We also employed MAKE Controller Kit to change the brightness of the bulb with PWM (Pulse With Modulation). In the latest iteration, in addition to the converter boxes, we developed an amplified box for the small light bulb. The accelerometer, the distance sensor, and the switch are connected to the convert box with a multi cable, while the small light bulb is connected to the amplified box with other cable.

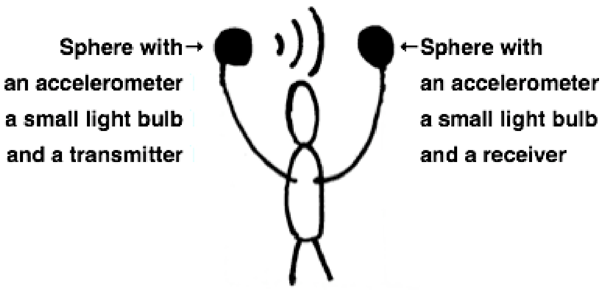

Figure 5: AEO instruments

2.3.1. Accelerometer, Distance sensor, Small light bulb

We developed two sets of converter boxes for the accelerometer and the distance sensor, and an amplified box for the small light bulb. Each converter box has a connector to the sphere, one audio input, three audio outputs, one extra audio input/output, a power socket, and three ring modulation circuits (Figure 6). Every audio inputs/outputs of the converter boxes are directly connected to the outputs/inputs channel of a 56-ch, 192KHz, 24-bit external audio interface, RME Fireface 800 (http://www.rme-audio.com/).

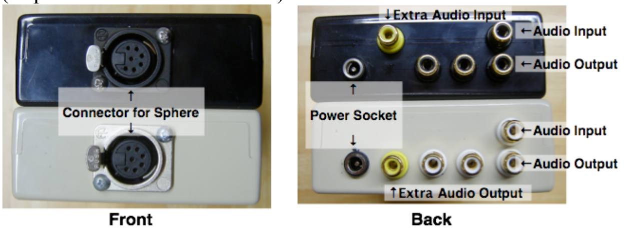

Figure 6: Converter boxes (Black for the sphere with the transmitter, White for the sphere with the receiver)

The amplified box has two audio output, one volume, two audio inputs, a power socket, and two channels amplifier (Figure 7).

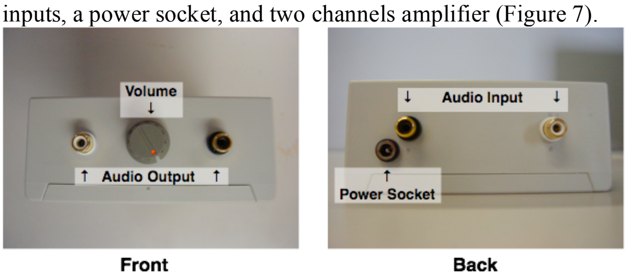

Figure 7: Amplified box

The audio outputs are connected to the sphere and the audio inputs are connected to the output channel of RME Fireface 800. We treat all the audio signals on MaxMSP.

Accelerometer:
For each accelerometer, we employ one audio input, three ring modulation circuits, and three audio outputs of the converter box. We assign an audio signal to the audio input. The audio signal is distributed to the three ring modulation circuits. The amplitude of the audio signal carries the change in analog voltage from each axis of the accelerometer (i.e. x, y, z) by employing ring modulation circuit with two transformers (Sansui ST-75) and four germanium diodes (1K60) (Figure 8). Each audio output produces the change of acceleration in each axis as modulated audio signal.

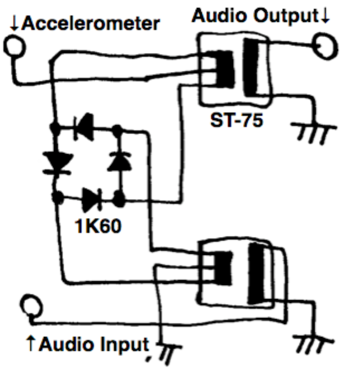

Figure 8: Ring modulation circuit

We detected the inclination and acceleration of each axis of the accelerometer as a separate audio signal. For each signal, we setup a volume calculator. It reports the change in analog voltage from each axis as the change of volume of incoming each audio signal. We employ the reported value to control the parameter of represented sound of the performance.

Distance sensor:
For distance sensor, we employ one extra audio input and one extra audio output of the converter boxes. The extra audio input is connected to the transmitter and the extra audio output is connected to the receiver. We assign an ultrasonic sine wave (40KHz) to the extra audio input. The audio interface enables the utilization of such high frequency with its high sampling rate than the built-in audio interface. The external audio output produces the audio signal from the receiver. We detected the distance between two spheres as an audio signal. For the signal, we setup a volume calculator. It reports the change of the distance as the change of volume of the audio signal. We employ the reported value to control the parameter of represented sound of the performance.

Small light bulb:
For each small light bulb, we employ one audio output, one audio input, and one channel amplifier of the amplified box. By feeding audio signals in appropriate amplifier, it is possible to bring the coil of the bulb in an excited state. We assign the represented sound of the performance to the audio input. We amplify the sound and connect it to the small light bulb. In the performance, the brightness of the bulb reflects the change of the represented sound changes. The volume of the amplified box adjusts the threshold of the brightness.

3. Related works

There are some precursors who also tried to employ audio signals as means to communicate with the physical world. Allison and Place [1] developed the SensorBox. The device accepted six sensor inputs and two audio inputs. The data from each sensor was carried as the amplitude of a sine wave, which was located in the 18KHz to 20KHz, and mixed back on the two audio inputs. They did not provide its technical detail well, however their approach was quite same as ours of the converter box of AEO. Canadian artist artificiel also explored the light bulb as sound source in their electro-acoustic installation "beyond6281 [2]". They feed processed audio signals in powerful amplifiers in the similar way to the small light bulb of AEO. PingPongPlus [4] employed audio signals for position tracking. It detects the position of ping-pong ball with a sets of microphones mounted under the ping-pong table. They detect the time of hit with each microphone and calculate the position with the time difference. Ms. Pinky [9] by Scott Wardle consists of a set of vinyl records and software running on MaxMSP. The vinyl contains special signals and from the signal, the software decodes the velocity, direction, and physical position of the needle on the surface of the vinyl in real time.

4. Discussions

We explored the characteristics of the audio interface as a device for physical computing. The audio interface accomplished higher sampling resolution and sampling rate than the other devices. In the installation Monalisa "shadow of the sound", we simply employ the push button with the built-in audio interface. In The SINE WAVE ORCHESTRA stay amplified, we accomplish subtle change of the frequency of the sine wave with 16-bit sampling resolution. With the latest instrument of AEO, the sampling resolution and sampling rate of the audio interface enables to represent subtle movements of the performer as patterns of sound and light than the other devices. During the latest implementation of AEO instruments, we have conducted informal comparison with the audio interface and the MAKE Controller Kit. We assigned the change of the acceleration into the change of the volume of white noise sound. With the audio interface, when the performer shakes the sphere, the resulted sound could be heard as a kind of maracas percussion with its fine time precision. However, with the MAKE Controller Kit, the resulted change of the sound did not reflect the subtle movement of the performer. As Wessel and Wright pointed out [10], the low latency between gesture and gesture controlled audio output is essential for live computer music performance. We believe that our approach suggests a sensitive way to communicate physical world with its preciousness of measurements and time.

Wessel and Wright argued the possibility to apply existing audio DSP (Digital Signal Processing) modules (e.g. Filters, Fast Fourier Transforms, Linear Predictors) to process the signals from/to the physical world [10]. We have not investigated such possibilities well. However we consider handling multiple sensor signals with one channel of the audio interface by employing band-pass filters.

As a future work, we consider to publish our developments for public. We plan to provide various instructions and examples of the employment of audio signals to communicate with the physical world. We hope to encourage people to stimulate each other to discover the potential of the audio interface as a device for physical computing.

Acknowledgements

We would like to thank for the other member of Monalisa, The SINE WAVE ORCHESTRA, and AEO. Monalisa "shadow of the sound" was developed under the support of NTT InterCommunication Center and FY2005 IPA (Information-Technology Promotion Agency) Exploratory Software Project (Project Manager: KITANO Hiroaki).

References

[1] Allison, J., Place, T., SensorBox: practical audio interface for gestural performance, In Proceedings of the 2003 Conference on New interfaces For Musical Expression, pp.208-210 (2003)
[2] artificiel, beyond6281, Art + Communication 2006: WAVES, RIXC, the Center for New Media Culture, Riga, Latvia, pp.38-39 (2006)
[3] Furudate, K., Jo, K., Ishida, D., Noguchi, M., The SINE WAVE ORCHESTRA stay amplified, Art + Communication 2006: WAVES, RIXC, the Center for New Media Culture, Riga, Latvia, pp.104-105 (2006)
[4] Ishii, H., Wisneski, C., Orbanes, J., Chun, B., Paradiso, J., PingPongPlus: design of an athletic-tangible interface for computer-supported cooperative play, In Proceedings of CHI '99. ACM, New York, NY, pp.394-401 (1999)
[5] Jo, K., Transition of an Instrument: The AEO Sound Performance Project, Leonardo Music Journal No.17, pp.46-48, (2007)
[6] Jo, K., Furudate, K., Ishida, D., Noguchi, M., Transition of instruments on The SINE WAVE ORCHESTRA, ACM Computers in Entertainment, on October (2008) (to be published)
[7] Jo, K., Nagano, N., Monalisa: "see the sound, hear the image", In Proceedings of the 8th International Conference New Interface for Musical Expression, pp.315-318 (2008)
[8] O'Sullivan, D., Igoe, T., Physical Computing, Boston, USA: Thomson Course Technology (2004)
[9] Wardle, S., Ms Pinky, http://www.mspinky.com/
[10] Wessel, D. and Wright, M., Problems and Prospects for Intimate Musical Control of Computers, Computer Music Journal. 26, 3, pp.11-22 (2002)