Audio Engineering Society

Convention Paper 8396

Presented at the 130th Convention
2011 May 13–16 London, UK

Interactive Mixing Using Wii Controller

Rod Selfridge[1], Josh Reiss[2]

[1] Queen Mary University of London, London, E1 4NS, UK
rodselfridge@hotmail.co.uk

[2] Centre for Digital Music, SEECS, Queen Mary, University of London, London, E1 4NS, UK
josh.reiss@eecs.qmul.ac.uk

ABSTRACT
This paper describes the design, construction and analysis of an interactive gesture-controlled audio mixing system by means of a wireless video game controller. The concept is based on the idea that the mixing engineer can step away from the mixing desk and become part of the performance of the piece of audio. The system allows full, live control of gains, stereo panning, equalization, dynamic range compression and a variety of other effects for multi-channel audio. The system and its implementation are described in detail. Subjective evaluation and listening tests were performed to assess usability and performance of the system, and the test procedure and results are reported.

1. INTRODUCTION

There are many benefits of a wireless gesture controlled mixing system. If used in a studio environment, the mixing engineer has the freedom to be intermingled with the performers, enabling a closer relationship and understanding for both. During a live performance the mixing can be carried out in real-time during the performance to rectify any anomalies or create an individual mix for the audience. The mixing can even be the performance itself with sounds evolving and moving throughout. This can be as a standalone performance or possibly along with visual cues.

For novice users the use of the Wii Controller as a mixing device can present a less daunting option compared to the banks of controls on a mixing desk. Basic intuitive gestures can be used to interact with the music to increase and develop understanding of the mixing process. There is also potential to develop the use of the system's basic concepts for music therapy applications.

The use of the Wii Controller as means of an interface to capture the users' gestures and to translate these into control signals for audio is not a new concept. Conducting using the Wii Controller is explored in [1-3], whereas in [4] the process of conducting is examined by use of the controller.

Other projects focus more on the process of sound production. In [5], the interaction between the Wii Controller and the Wii Sensor Bar is manipulated to trigger sound events. In [6], the Wii Controller and Nunchuk accessory are used to simulate drum sticks; producing appropriate sounds as response to accelerometer data.

The work carried out in this paper is closely related to work carried out in [7-9]. In [7] the use of accelerometer

data, from an experimental controller, is processed into MIDI data. This is thereafter used to adjust playback volume, sample triggering and timbre modulation. The use of a Wii Controller for placement of sound in a surround sound system is examined in [8], while the use of auditory cues to provide feedback in gestural control (especially stereo panning) of multitrack audio is evaluated in [9].

The goal of this project is to create an interactive motion controlled mixing system, where users can manipulate various properties and audio effects within a multi-track recording. In this paper we describe a system that implements the following:

- The controller has the ability to select which track in the multi-track recording it modifies, including the Master Track. If required, the selected track may be soloed.
- Additional transport controls (playback, start, stop, and loop on / off) and tempo control are provided on the Wii Controller.
- Motion of the controller is used to adjust the levels of parameters on a variety of digital audio effects.
- The controller has the ability to select which parameters it controls from a preset selection of standard effects and mixing controls.
- The system provides the user with a reset option that can be applied to an individual effect or globally to the whole recording.
- A GUI is provided to assist the user in tracking what they are controlling while using the Wii Controller.

This system, named Kynan, should allow the user to select which track they currently control and allow the user to change which parameters they are controlling within the track. This then offers a wider range of mixing controls and effects that can be applied. it is hoped that a feeling of 'performing the mix' may be instilled in the user as they directly control the sound they hear in real-time.

Once Kynan was completed a number of users were invited to mix a multi-track recording in order to determine if they could achieve the mix they desired and if they found mixing with Kynan easier or harder than conventional methods. Listening tests between pieces of music mixed by conventional methods and Kynan was carried out. Kynan was also tested for accuracy with conventional methods.

2. IMPLEMENTATION

2.1. Overview

The Wii Controller has a number of sensors for gestural indication and buttons that allow the user to control the game being played on the Wii. Inside the Wii Controller is a set of 3-dimentional accelerometer that provide data for *X*-, *Y*- and *Z*-axis acceleration, yaw, pitch and roll. There are a total of 11 buttons on the controller.

The Wii Controller also has an infrared (IR) camera at its front which picks up light sent from an IR source. This provides absolute positioning in the X and Y axis. The signals from these sensors are transmitted to a personal computer using the Bluetooth protocol.

Once the signal from the controller is connected to a computer it has to be manipulated by software packages to produce the required changes in parameters. GlovePIE [10] (Glove Programmable Input Emulator) was used to capture the Bluetooth signal from the Wii Controller and convert this data into a useable format. The output is in MIDI format, which can then be used to manipulate audio. The data passes through an interface program that allows the signal to be used as an input to the sequencing software. The controller signal then communicates with suitable control software. The control software receives this data and uses it to manipulate the audio. The control software works closely with the sequencing software that holds the multi-track recording and allows the user to manipulate the audio. This is depicted in Figure 1.

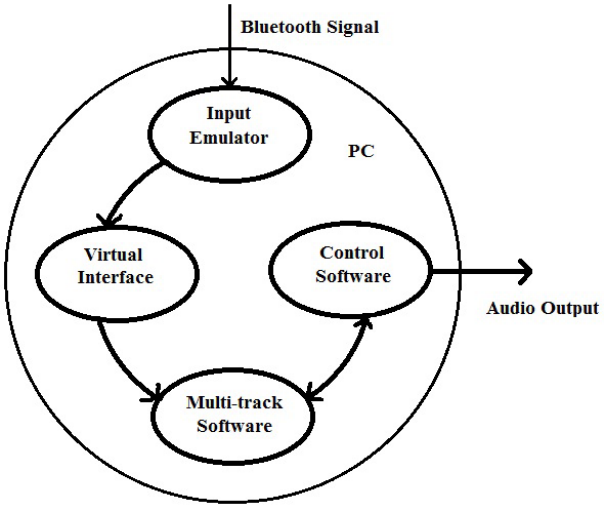

Figure 1. System Data Flow in Kynan.

2.2. Input Emulator

A short program was written in GlovePIE's scripting language to make use of all available sensor and button

outputs from the Wii Controller. These were assigned to represent MIDI data. With the number of sensors available we had to decide which sensors to use in order to accurately represent the gestures of the user.

The IR sensors were used in an initial test program but found to have two limitations which precluded their use in Kynan. The first was the fact that the user would have to be standing facing the IR source and relatively close. One of the desired requirements of this system is that the user is free to move while mixing the audio and this would limit that ability. The second limitation is that the camera already has an established range of motion equivalent to about 90 degrees, again presenting a challenge with a limited range of motion.

X and *Y* acceleration and roll (rotation about the *Z* axis) were used to map gestures of the user holding the controller to MIDI outputs. From Figure 2 it can be seen that using the roll value we can calculate the orientation of the Wii Controller and use this along with *X* and *Y* axis accelerometer values to obtain variables that represent motion in the horizontal and vertical directions. The roll value was necessary since this provides the flexibility of not having to hold the Wii Controller in the same orientation as the accelerometers to obtain horizontal and vertical motion.

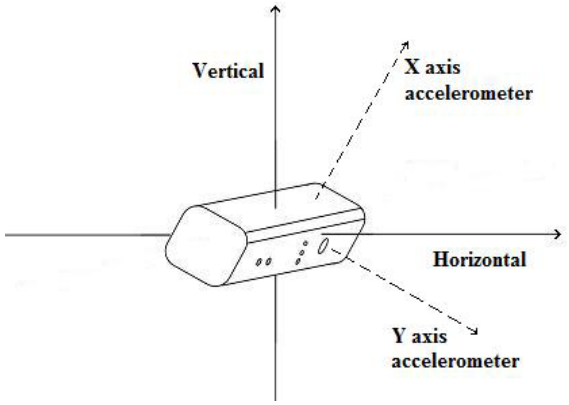

Figure 2. Wii Controller Orientation.

To make the controller more user friendly the assigning of values to the horizontal or vertical variables was placed inside a loop. The condition for the code inside the loop to be executed is that the B button or 'trigger' button on the Wii Controller has to be pressed. This provides the user with the ability to decide when to activate the motion control capabilities and when their movements should not cause any corresponding change. All other buttons are also assigned to MIDI note values, allowing them to be utilized for further control.

2.3. System Configuration

A MIDI virtual interface was required to link the output of GlovePIE and the input of our multitrack sequencing software, Ableton Live (http://www.ableton.com/suite-8). The sequencer looks for a hardware link as a MIDI input. MIDI-OX (www.midiox.com/) was used to create a virtual link between GlovePIE and Ableton Live.

Motion captured by the Wii Controller was able to provide smooth and responsive control over the selected parameters. The parameters chosen within these effects represented the standard parameters the majority of manufacturers make available for adjustment. Variables were calculated for horizontal and vertical motion and then scaled into the MIDI output ranging from 0 to 127. When the values of the un-scaled variables increased or decreased outside set limits, the MIDI output remained at full scale until the un-scaled variable returned to within the limits.

Max for Live (http://cycling74.com/) was used to process the MIDI data from GlovePIE. It allowed us to reference individual parameters with generic path names. The use of wildcards within the path names gave the ability to use the same Max circuitry for all tracks, adding and removing tracks as desired. We could then produce a system that allowed the user to switch between a set of audio effects. One limitation was that the effects had to be in the same set order within the chain of effects for all tracks. In Kynan we used Equalization, Compression, Phaser and Overdrive. The path names were then given to a remote object within Max for Live. The remote object took virtual control of the referenced parameter. These effects were then modified by the data provided by the motion of the Wii Controller.

2.4. Graphical User Interface

The GUI provided for Kynan is shown in Figure 3. It is split into 4 sections; Track Mode, Master Mode, Reset Mode, and Position Display.

The Track Mode is the top left highlighted area, and provides the user with a visual indication of what parameters they are currently controlling within the selected track. The user has control over Pan, Gain, EQ, Compression, Phaser and Overdrive. An exact list of parameters is given in Appendix A. It is possible for the user to turn these effects on and off, which also resets the parameter values. The track may also be soloed allowing the user to hear the track they are controlling, muting the rest of the tracks.

The Master Mode area in the bottom right of Figure 3 is similar to Track Mode except effects here are applied to

all tracks. In this mode we have the ability to adjust Gain, Compression and Reverb. (See Appendix A for further detail).The ability to adjust the tempo has also been provided in this mode, along with turning a loop on and off. (Start and stop points to be set before use).

The Position Display of the GUI indicates the relative position of the Wii Controller within its limits. This was included in response to comments made in [11], where it was found that test subjects wanted visual feedback, along the lines of a laser pointer, to indicate the position of the Wii Controller. The bottom left hand side of this square represents (0,0) values for both horizontal and vertical inputs.

Finally, when Reset Mode is selected the user has the ability to reset all effects that have been previously applied and all effects are zeroed.

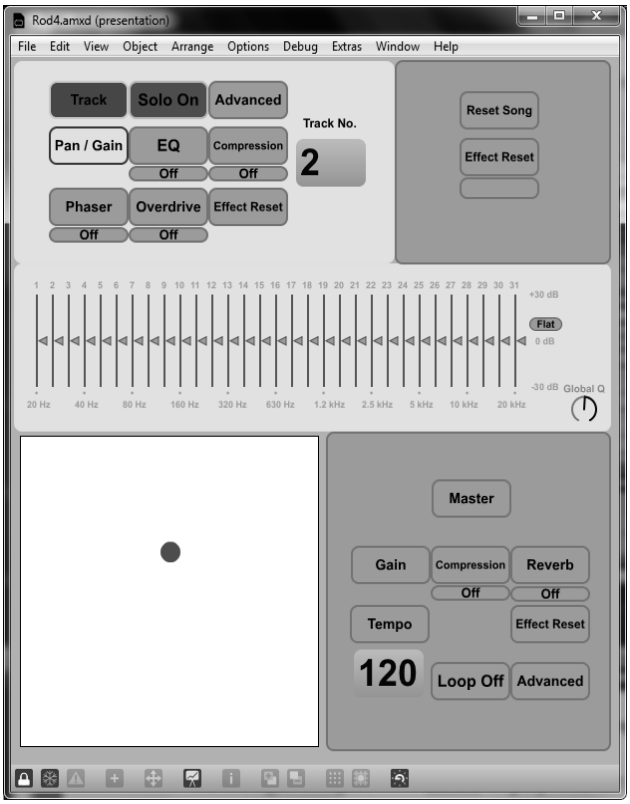

Figure 3. Graphical User Interface for the interactive gesture-controlled audio mixing system.

2.5. Equalization Issues

Of all the effects that were made available to the user, the Equalization (EQ) was the most difficult to accurately control and obtain the desired results. Here, the user draws the shape of the desired EQ curve with the controller. This shape is then applied to a 31 band Graphical Equalizer.

When Kynan is set to adjust Pan and Gain the adjustments can be performed almost in isolation; horizontal motion for Pan and vertical for Gain. When drawing an EQ curve, the motion of the Wii Controller is curved and often diagonal. Horizontal movement is not adjusting a parameter. Instead, it is changing which, which bandpass filter will be adjusted by the vertical motion. Instead of selecting between 0 and 127 discrete levels, as with a normal MIDI controller, the EQ selects a filter between 1 and 31. Vertical motion sets the boost or cut for each filter as the user draws the desired curve.

Our display for the EQ is controlled by a loop which updates each filter slider. When changing one portion of the EQ curve all sliders are updated. This takes a notable amount of time, approximately 0.5 sec to 1 sec. This can be disconcerting when using the device and gives an appearance of a much more problematic and jerky control than actually being applied.

After initial review (and user testing), an alternative equalization implementation was examined. Instead of using a Graphical Equalizer, an equalization curve with 3 adjustable points was used. An example is shown in Figure 4. This implementation gave easy control over the points and more easily obtained the desired result.

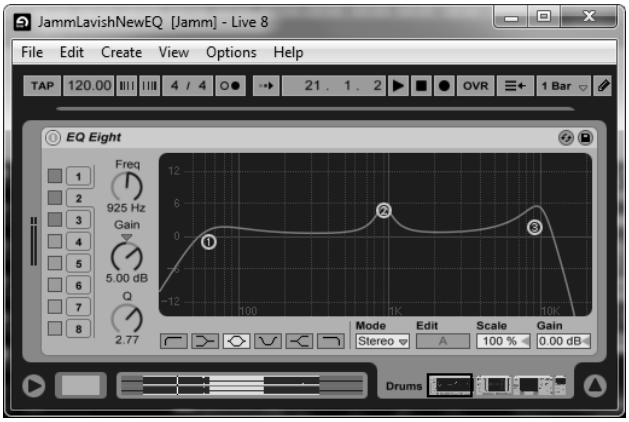

Figure 4 GUI for gesture control of EQ, featuring an equalization curve with 3 adjustable points.

3. TEST RESULTS AND DISCUSSION

A test was carried out to identify the time taken to perform a mix using a mouse to control parameters and compare this to the time it takes to perform a similar mix using Kynan. Accuracy of both conventional mixing and Kynan were also compared.

A test group was utilized to provide feedback and determine to what extent users were able to achieve the desired mixing results. A larger test group were also given a form of Turing test to see if they were able to distinguish between two mixes of the same piece of music; one mixed using Kynan and another using conventional mouse control over parameters.

3.1. Achieving a Desired Mix

An informal test was conducted by the authors to examine the ability of Kynan to achieve a desired mix compared to conventional methods of using keyboard, mouse or mixing desk. In this case the mouse was used to perform the mix. Only the parameters that can be adjusted by Kynan were used. Required values for these parameters were set prior to mixing and the goal was to adjust the parameters to a value as close as possible to the required values. The process of mixing the music using conventional mouse methods and using Kynan were both timed to compare how long it took to achieve this result. The piece of music had five individual tracks to be adjusted, followed by the master. This gave a total of 79 parameters to adjust. EQ was included in the mix but not used in the comparison since setting the 31 values would be extremely difficult on Kynan and would have increased the experimental time greatly.

Mixing the piece using the mouse to adjust parameters took 18 minutes 12 seconds, averaging just under 14 seconds per parameter. The time taken to mix using Kynan was 25 minutes and 15 seconds, averaging at just over 19 secs per parameter.

As expected it took longer to mix using Kynan. If concentrating on Panning for example and moving the Wii Controller to the right, it can often be the case that during adjustment there has been some vertical motion in the movement. This produces unwanted changes in volume which has to be corrected. This is the same for most parameters. Also, very fine adjustments are not easy to achieve with the Wii Controller. It works well with big sweeping movements up and down or left to right but to change a small value takes more of a 'flicking' action or movement of the wrist with the B button pressed. This can be unpredictable and cause unwanted changed in the perpendicular axis.

In both methods it was found that in some cases it was not possible to set the parameter to the required values. This could be because the resolution of the parameter itself does not allow for the exact value to be set. The mouse did manage to set the values closer to the required values than Kynan. Overall, Kynan had a 2.94% error in setting parameters while the conventional method had 1.08% error.

3.2. Subjective evaluation of usability

A group of 7 volunteers were recruited to use Kynan and express their opinions. The volunteers ranged in age from under 15 years to over 60 years. None of the subjects were experienced mixing engineers. Some had very limited experience with mixing music and others had a basic knowledge of the process and had attempted multi-track mixing in the past. A brief demonstration was given along with a printed list of controls. (Appendix A)

One of the first actions the majority of testers did was to point the Wii Controller at the PC monitor and they were informed that this was not required. Some users also moved the Wii Controller very gently causing the accelerometers not to register the motion and no change in parameters will take place.

Following this, all users were able to demonstrate accurate control over Pan and Gain of individual tracks. For example, users were able to move the controller upwards to increase the gain of the selected track, and move the controller to the left to pan the track to the left. This was the simplest control, and a basic goal for the system.

As explained above, equalization was the most difficult effect to control and a large portion of the testers' time was spent trying to come to terms with the type of gestures required to obtain a satisfactory EQ curve. The majority of negative feedback came from the difficulty to control the EQ curve, quoting the difficulty of getting a response in Kynan when moving the Wii Controller in a diagonal motion.

The range of other parameters that testers tried after this was found to be as easy as the Pan and Gain for them to control. The limitations of the users' knowledge on digital audio effects became what limited their progression through the use of Kynan rather than any problems using the Wii Controller. However, one user with little experience of the mixing process noted that Kynan was less intimidating than a mixing desk and the user felt happier learning about mixing through its use.

3.3. Listening Test

21 volunteers participated in a simple listening test. Two different pieces of music were mixed using both Kynan and conventional methods; one piece relatively simple and the other more complicated. To make the two mixes comparable, only the parameters that are

available in Kynan were adjusted, with all other parameters the same for both mixes. The time was also limited to 2 minutes that could be spent adjusting the parameters on each track and an additional 2 minutes for the master track. Although this gives a rough mix it is the same for both methods and assists in comparison.

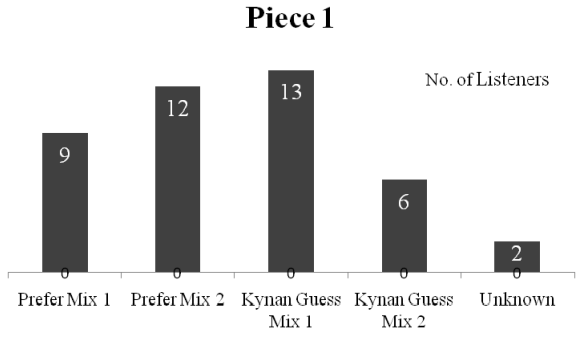

Figure 5 Piece 1 Listening Test Results

The first piece of music was a basic mix of 5 tracks. The listener results for this test are shown in Figure 5. For Piece 1, mix 1 was carried out using Kynan. Here it can be seen that there is not much difference as to which mixing method produces a preferred result, although the traditional method is slightly ahead with 12 listeners preferring it. The number of people who correctly guessed that mix 1 was carried out using Kynan is similar with 13 listeners.

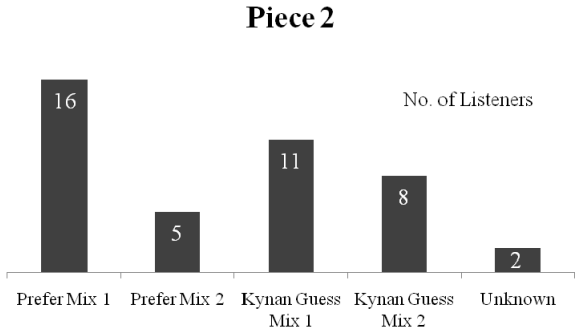

Figure 6 Piece 2 Listening Test Results

The second piece of music was a more complicated mix with a total of 9 tracks. The rhythm section, (percussion and bass) was relatively constant through the mix but there are more chord changes within the piece and an increased movement between the instruments. Here, mix 1 was carried out by conventional mouse control and mix 2 by Kynan. The results are shown in Figure 6. The conventionally mixed track (mix 1) was a clear favorite amongst listeners. The results indicate that the more complex the mix the less pleasing the results from Kynan.

From the listeners who preferred the conventional mixes the reason given most was that there was a crisper sound and better balance between the instruments. Also it was thought the Panning of the instruments was better. Listeners who preferred the Kynan mixes stated they enjoyed the more rhythmic elements. Certainly the melody is further back in this mix.

A split of the listeners' preferences is shown in Figure 7. The number of people who preferred both conventional mixes is eight times the number of those who prefer both Kynan mixes. The results seem to indicate Kynan does not produce as pleasing a mix as conventional methods under the conditions set.

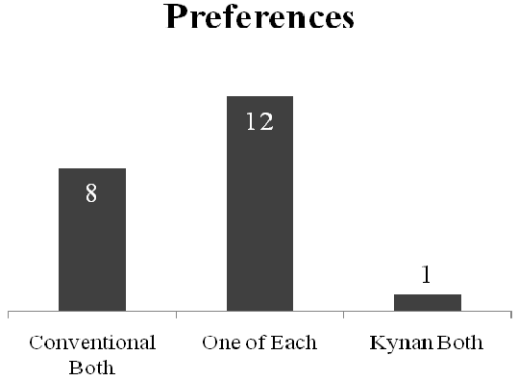

Figure 7 Listener Preference

4. CONCLUSIONS AND FURTHER WORK

The design and use of a motion controlled mixing system provides an innovative and novel method to perform a mix in real-time. It enables the mixing engineer to be released from the conventional hardwired controllers and provides freedom to perform the mix with more physical interaction. The ability to move a track in the air between loudspeakers by way of pointing the Wii Controller is a very intuitive method to positioning tracks within the mix. Extending this method of control to other parameters gives the user the opportunity to physically express their own interpretation when performing a mix that cannot be achieved sitting behind a desk or keyboard.

The implementation of a Kynan is not designed to compete with a mixing desk and the number of different parameters that an engineer has available at the desk, but to complement and enhance the process. It is believed that conventional mixing methods would be stronger at producing a more technically balanced final mix of music. Kynan is meant to give users a new way

of expressing themselves, freedom to walk away from the recording mechanism and immerse themselves in the performance, while still controlling major aspects of the music in real-time.

Listening tests have shown that for simple multi-track pieces of music it is difficult for the layperson to distinguish between conventionally mixed tracks to those mixed with Kynan. When the multi-track piece becomes more complicated the Kynan mixed version appears to be more apparent. Testing indicated that listeners prefered final mixes carried out by the more conventional method. It was also felt that listeners will be more familiar with conventional mixes and that this was not necessarily a reflection on the performance of Kynan.

Future development could include the use of a different controller other than the Wii Controller. With mobile phones now developed with accelerometers built into them, the opportunity exists to control the system using such devices. The user could mix a multi-track recording stored on their mobile phones using the phone itself as the gesture based controller and hear the results through headphones.

An avenue for future development of interactive motion controlled mixing could be to use a controller-free gestural system, such as Kinect (www.xbox.com/en-GB/kinect) . Open source code [12] is already available that may be applied to such as system.

5. REFERENCES

[1] Breugge, B. , et. Al.,"Pinocchio: Conducting a Virtual Symphony Orchestra" In Proceedings of the International Conference on Advances in Computer Entertainment Technology, Salzburg, Austria. 2007

[2] Peng, L. and Gerhard, D. (2009) "A Wii-based gestural interface for computer-based conducting systems". Ninth International Conference on New Interfaces for Musical Expression, Pittsburgh, PA, USA.

[3] Nakra, T., Ivanov, Y., Smaragdis, P., Ault, C.: The USB Virtual Maestro: an Interactive Conducting System, pp.250-255, NIME2009 (2009).

[4] Bradshaw, D. and Ng, K. (2008), "Analyzing a conductors gestures with the Wiimote". In Proceedings of EVA London 2008: the International Conference of Electronic Visualisation and the Arts, British Computer Society, London, UK. 2008

[5] Miller, J. and Hammond, T (2010). "Wiiolin: a virtual instrument using the Wii remote". In Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010), Sydney, Australia.

[6] Wong, E. L., Yuen, W. Y. F., and Choy, C. S. T., Designing Wii controller: A powerful musical instrument in an interactive music performance system. Proc. Int. Conf. Advances in Mobile Computing and Multimedia, New York: ACM, 2008, 82-87.

[7] Demey, M., Leman, M. and Cornelis, O. (2008). "The IPEM_EME: a Wireless Music Controller for Real-time Music Interaction". In Proceedings of the 14th International Conference on Auditory Display, Paris, France. 2008

[8] Quinn, P. Dodds, C. Knox, D. "Use of Novel Controllers in Surround Sound Production," Audio Mostly, Glasgow, 2009.

[9] M. J. Morrell and J. D. Reiss, "Auditory Cues for Gestural Control of Multi-Track Audio," submitted to International Conference on Auditory Display (ICAD), Budapest, June 2011.

[10] Kenner, C., GlovePIE.v. 0.42, June 2010. http://glovepie.org/

[11] Kiefer, C., Collins, N. and Fitzpatrick, G. (2008). "Evaluating The Wiimote As A Musical Controller". In Proceedings of the International Computer Music Conference, Belfast. 2008

[12] Martin, H. libfreenect. OpenKinect, 2010. https://github.com/OpenKinect

6. APPENDIX A

6.1.1. Global Controls

Left/Right buttons: Toggle user mode
Home Button: Toggle start/stop

6.1.2. Track Mode

A Button: Selects effect
1 Button: Toggles Solo
2 Button: Effect reset / toggles effect on/off
Up/Down buttons: Toggles between basic and advanced controls

Pan/Gain	Up and Down	Left and Right
Basic	Gain	Pan
Advanced		

EQ	Up and Down	Left and Right
Basic	Gain	Frequency bands
Advanced	Global Q	

Compressor	Up and Down	Left and Right
Basic	Ratio	Threshold
Advanced	Attack	Release

Phaser	Up and Down	Left and Right
Basic	Mix percentage	Sweep rate
Advanced	Feedback amount	Frequency of first notch

Overdrive	Up and Down	Left and Right
Basic	Tone	Drive
Advanced	Bandwidth of pre-filter	Frequency of pre-filter

6.1.3. Master Mode

A Button: Selects Effect/Tempo
1 Button: Toggles Loop on/off
2 Button: Effect Reset / Toggles effect on/off / Resets tempo
Up/Down Buttons: Toggles between basic and advance controls / increase or decrease tempo

Gain	Up and Down	Left and Right
Basic	Gain	
Advanced		

Compressor	Up and Down	Left and Right
Basic	Ratio	Threshold
Advanced	Attack	Release

Reverb	Up and Down	Left and Right
Basic	Gain of early reflections	Delay before 1st early reflection
Advanced	Gain of defused region	Decay time

6.1.4. Reset Mode

2 Button: Resets all effects and values