A. Georgaki and G. Kouroupetroglou (Eds.), Proceedings ICMC|SMC|2014, 14-20 September 2014, Athens, Greece

Implementations of the Leap Motion in sound synthesis, effects modulation and assistive performance tools

Lamtharn Hantrakul
Department of Music, Yale University
469 College Street, New Haven CT
`lamtharn.hantrakul@yale.edu`

Konrad Kaczmarek
Department of Music, Yale University
469 College Street, New Haven CT
`konrad.kaczmarek@yale.edu`

ABSTRACT

The Leap Motion opens new possibilities for mapping the various degrees of motion of the human hand with musical expression. The Leap Motion is a computer peripheral released in mid 2013 that uses IR cameras to track hand and finger location with unprecedented accuracy. In this paper, we explore implementations of the device in sound synthesis and effects control. The device is interfaced with Max/MSP to provide motion and finger-based control over multiple parameters in a software synthesizer. Next, we implement a 5-grain granular synthesizer where users trigger individual grains by depressing their fingers in mid-air. While triggering grains, users can simultaneously move their hands to dynamically modulate grain length and scrub the sample. The benefits and limitations are discussed in light of recent compositions and performances. The Leap Motion is also used to spatialize the synthesized sound produced from a 6-channel hemispherical speaker. Applications to music composition and music therapy are discussed.

Keywords: Leap Motion, gestural control

1. INTRODUCTION

Advancements in computer hardware and digital signal processing have produced incredible sounding hardware and software synthesizers. However, interactions with these electronic sounds are often limited to the traditional mouse, keyboard, knobs, faders and multi-touch interfaces. Motion-based control over musical parameters allows direct interaction with electronically generated sound.

The use of motion has gained increasing popularity in the audio industry. IK Multimedia, a producer of controllers and mobile interfaces recently introduced the *iRing*. The device is a ring worn on the user's hand that is tracked by a smartphone's front facing camera. X, Y and Z positions can then be used to modulate effects in music applications [1]. In 2013, the software developer Steinberg announced *IC Air*, a gesture-augmentation to the popular DAW *Cubase* [2]. The user can adjust faders, EQ and navigate sessions using gestures instead of a traditional mouse and keyboard. Grammy nominated artist, Imogen Heap has recently debuted her *Mi.Mu* controller gloves, equipped with flex sensors and accelerometers [3]. Elena Jessop at the MIT Media Lab developed a glove shaped controller allowing real time manipulation of vocals [4]. Another instrument is Laetitia Sonami's "Lady's Glove," developed by Sonami and Bert Bongers [5]. The use of motion-based control is not limited to the digital domain; a project successfully funded via KickStarter in 2013 named *Vectr* allows users to control analogue synthesizers using hand gestures [6].

The rapid pace of technological advancements made motion-tracking technology available to the average consumer. When the Microsoft Xbox Kinect was released in 2009, it was immediately adopted by numerous projects interfacing human interaction with art and technology. The Kinect has been used in numerous interactive art installations [7] and novel ways to trigger electronics sounds [8].

This paper focuses on musical applications of the Leap Motion, a new computer peripheral released in mid-2013 that delivers unprecedented accuracy in finger and hand tracking.

2. LEAP MOTION

2.1 Introduction

The Leap Motion is a USB device developed by Leap Motion Inc., released in July 2013 and priced at $79.99. The co-founders, David Holz and Michael Buckwald developed an algorithm that tracks "all 10 fingers up to 1/1000 of a millimeter" [9]. It has a wide 150° field of view, allowing users to interact with their computer via familiar hand gestures like pinching or swiping. The device features an open API for developers and has applications ranging from 3D graphic manipulation to motion games [10].

At the time of writing, a new tracking API was recently released in beta version to developers that solves many of the problems addressed in this paper. These will be discussed in context of a forthcoming Java application built on the implementations in this paper.

Figure 1a. The Leap Motion device

2.2 Existing technologies

There are numerous ways to interface human motion with music software. For example, *SimpleKinect* is an application developed by Jon Bellona that translates limb position into OSC messages [11]. Source Audio's *Hot Hand USB* controller is detected as a generic MIDI device and maps hand tilt to MIDI messages [12]. However, accurate finger tracking has only been possible using systems of gloves and flex sensors [3,4,5].

The Leap Motion provides a *non-invasive* method of independently tracking both hand and finger data. Prior to this device, no commercial sensor could offer this level of accuracy. Recently, Microsoft has released the Xbox One Kinect that boasts a new "time of flight" motion tracking system, a 1080p camera and even finger tracking [13]. Although the device holds great promise, there is currently no way to interface the device beyond the Xbox console at the time of writing.

2.3 Leap Motion in Music

To our knowledge, the most popular commercial application designed to interface the Leap Motion with music software is Geco created by Geert Bevin [14]. The user is able to map hand movement into MIDI messages through an intuitive interface. However, the application has been used primarily to control parameters such as effects dry/wet mix or volume. There is no option to select and trigger notes using hand or finger motion. Moreover, the application is limited to either "open hand" or "closed hand" and provides no option to map individual finger data to musical parameters (Figure 2).

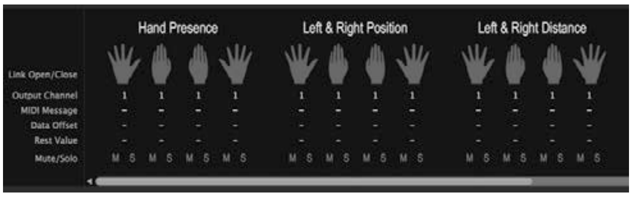


Figure 2. Part of the *GECO*

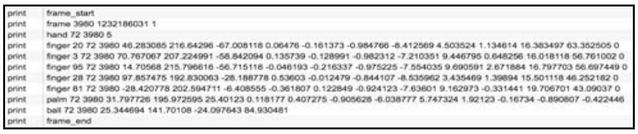
This paper explores ways of controlling pitch in addition to modulation in a performance setting. The paper also investigates implementations that incorporate additional degrees of freedom offered by finger tracking.

3. TECHNICAL IMPLEMENTATION

3.1 Leap Motion and Max/MSP

The Leap Motion is interfaced with Max/MSP using the Max object `aka.leapmotion` created by Masayuki Akamatsu [15]. The object sends coordinates, velocities and accelerations of each hand and finger into Max environment. However, despite the object's wealth of data, it does not distinguish between left and right hands nor does it distinguish which finger is the thumb, index finger etc. Moreover, the object offers no easy method to pick out a particular value - say the position of the left hand's index finger – while preserving other data. This made mapping to parameters in Max and other music software very difficult.

A max patch was developed to sort this stream of data. Briefly, we poll the object every 10ms and use the messages `frame_start` and `frame_end` to trigger comparisons between x positions. The hand ID with smaller x position is assigned the left hand. This method is expanded to assign finger ID's. The x coordinates of each finger are ordered and assigned "thumb, index etc." depending on the hand. For example, the finger with the smallest x position is assigned the pinky on the left hand and thumb on the right hand. We also assume "one finger" means the user is pointing with their index finger while "two fingers" implies index and middle finger and so on.



```
print    frame_start
print    frame 3980 1232186031 1
print    hand 72 3980 5
print    finger 20 72 3980 46.283085 216.64296 -67.008118 0.06476 -0.161373 -0.984766 -8.412569 4.503524 1.134614 16.383497 63.352505 0
print    finger 3 72 3980 70.767067 207.224991 -58.842094 0.135739 -0.128991 -0.982312 -7.210351 9.446795 0.648256 16.018118 56.761002 0
print    finger 95 72 3980 14.70568 215.796616 -56.715118 -0.046193 -0.216337 -0.975225 -7.554035 9.690591 2.671884 16.797703 56.697449 0
print    finger 28 72 3980 97.857475 192.830063 -28.188778 0.53603 -0.012479 -0.844107 -8.535962 3.435469 1.39894 15.501118 46.252182 0
print    finger 81 72 3980 -28.420778 202.594711 -6.408555 -0.361807 0.122849 -0.924123 -7.63601 9.162973 -0.331441 19.706701 43.09037 0
print    palm 72 3980 31.797726 195.972595 25.40123 0.118177 0.407275 -0.905628 -6.038777 5.747324 1.92123 -0.16734 -0.890807 -0.422446
print    ball 72 3980 25.344694 141.70108 -24.097643 84.930481
print    frame_end
```

Figure 3. Data stream from aka.leapmotion object

Once processed, hand and finger ID's are used to extract the desired data (Figure 3). We found the Leap Motion becomes increasingly jittery at the extremes of its field of vision. We limited mappings to "stable" areas of vision nearer the origin. The Max patch built on the aka.leaptmotion object is used throughout this paper and is available for download [16].

4. MODULATION OF EFFECTS

4.1 Leap Motion and Ableton Live

The Max patch in section 3.1 sends MIDI messages to Ableton Live. In this implementation [16] a synthesizer patch was created using Native Instrument's Massive Synthesizer (NI Massive) in the style of a "Dubstep Wobble" frequently heard in electronic music.

Vertical hand distance controls the LFO rate modulating the cutoff of a low-pass filter. Horizontal motion controls a selection of notes defined by the user while forwards motion changes wavetable position. Using this setup, the user can control both note selection and LFO rate using just one hand. The interaction with the electronic synthesizer becomes more direct than conventional knobs, sliders or even multi-touch interfaces.

A touchscreen for example, presents physiological limitations as users can only stretch their index finger a certain distance away from their middle finger. In addition to sonic feedback, an Arduino microcontroller and RGB LED strip were used to provide the user visual feedback.

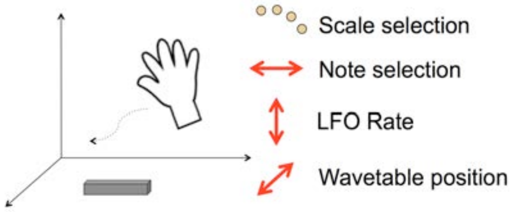


Figure 4a. Motion-based synth

The number of fingers selects the type of scale, octave or patch. We experimented with various ways to allow the user to tap a note with a finger much like an air-piano. In its current form, the Leap Motion's ability to track and retain finger ID's is not yet robust. When a user depresses a finger, the corresponding finger simply "disappears" from the data stream. This lead to mismatches between the ID's of remaining fingers, thus making accurate note selection using fingers very difficult.

The new API released in beta to developers in late May 2014 features Skeletal Tracking, in which the Leap Motion keeps a true model of hand and fingers. Bent fingers are now continuously tracked and made accessible through the API [10].

Although the concept of using motion to control synthesizer parameters has previously been explored, the Leap Motion allows an accurate and non-invasive approach. It does not require the user to wear any kind of device that may hinder performance. This implementation may also have applications in music education. We found that classically trained musicians unfamiliar with electronic music or even non-musicians were able to play the synthesizer and "wobble" it in time with a beat within a few minutes.

4.2 *[A]^2* performance

[A]^2, pronounced "A Squared", is a project that was premiered in December 2013. It explores the concept of "augmented acapella". The vocals from a 5-person acapella ensemble is processed in real time using Ableton Live to produce kick drums, snares, high hats, synthesizer sounds rivaling the finished sound of a recording.

The Leap Motion was used to modulate the effects of a live remix using sampled recordings of a preceding performance. Effects such as reverb, bit-crusher, low-pass and delays were placed at different vertices of an imaginary 3-dimensional cube above the Leap Motion (Figure 4c) [16]. The user can dynamically mix multiple effects by moving one hand while the other hand is free to trigger samples. The set up enabled particularly expressive modulation of effects.

There were times when the Leap Motion was confused by small interferences. For example, a shirtsleeve was enough to make the device jitter between the real hand and an imaginary hand located at the sleeve. Despite these limitations, the author has since adopted this effects cube over conventional interfaces. Controlling multiple effects in 3-dimensions is more intuitive than using multiple XY pads on a flat surface.

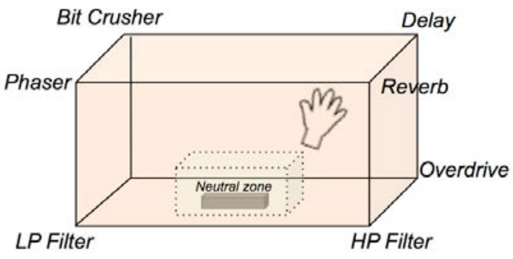


Figure 4c. [A]^2 effects space

Figure 4b. [A]^2 performance

(note that the chopsticks are only used for comical effect)

In its current form, the implementation only supports switching between different presets by changing the number of fingers. We hope to incorporate the new API's finger tracking capabilities to add percussive modulations to the sound when a finger is depressed.

Moreover, the implementation can be combined with tools like the Wekinator [17] to provide non-linear morphing between states of effects. Finger control can then be to alter interpolations between states on-the-fly.

5. MODULATION OF SYNTHESIZED SOUND

5.1 Leap Motion and Granular Synthesis

Granular synthesis is a type of sound synthesis where short fragments of a sample are extracted and then sequenced together to create new textures. These fragments are called "grains", which are often triggered by a periodic signal or random number generator.

A 5-grain granular synthesizer was built in Max/MSP [16]. Unlike conventional granular synthesizers, the user is able to trigger individual grains using finger motion. When a user "depresses" a finger – much like playing piano in the air – the corresponding grain is triggered. At the same time, the right hand's horizontal motion scrubs through the sample while vertical motion controls the grain length. This information is displayed in a GUI (Figure 5a). The red line indicates the play head's current position while the different colored lines represent the play head of each triggered grain. The user can control all these parameters simultaneously using one hand.

To address the shortcomings of finger interaction in 4.1

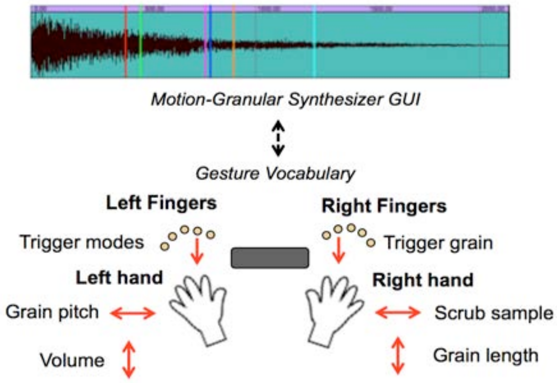


Figure 5a. Motion Granular Synthesizer

and 4.2, we attempt to detect finger depressions using vertical Y velocity. These were found to be more reliable then absolute Y position. When this velocity crosses a user-defined threshold value, a message is sent to trigger the grains. Because each grain performs the same function - as opposed to being distinct notes – small confusions in finger ID were less problematic. We incorporate differentiation between voluntary finger movement and involuntary finger movement due to overall hand movement.

The sound produced from this motion-augmented granular synthesizer was found to be pleasantly organic. For example, when the user triggers grains at regular intervals, human imperfections in timing lead to subtle variations in the trigger rate. This produced a more organic texture than one generated by a periodic signal. Moreover, the user is able to musically playback the sample like an instrument, as will be discussed shortly in 5.3.

This implementation also opens the door for novel ways to "augment" traditional methods of sound synthesis with motion-based control. For example, Physical Modeling synthesizers could use a 3D matrix to mix between sounds with non-linear morphing.

5.2 Leap Motion and Hemispherical Speakers

Hemispherical speakers have been developed and used by P. Cook and S. Smallwood et al [18] for acoustical and musical reasons. Firstly, the hemispherical speaker better mimics the way acoustic instruments propagate sound in all directions. Since the speakers have 6 independent channels, high and low frequencies can be scattered in specific directions. Moreover, the combination of electronics and chamber instruments often results in acoustic instruments being overpowered by PA systems. Cook et al have noted that the hemispherical speakers allow electronically generated sound to have spatial presence akin to an acoustic instrument.

The author was primarily concerned with using motion to dynamically alter spatialization and dispersion of sound from hemispherical speaker. In this implementation, a 6-channel hemispherical speaker was constructed by combining designs documented by the *Stanford Laptop Orchestra (Slork)* [19]. A salad bowl is drilled and fitted with speakers connected to 6 individual amplifiers and a multi-channel audio interface (Figure 5b).

Figure 5b. 6-channel hemispherical speaker

A Max patch maps hand position to a GUI consisting of nodes (Figure 5c). Moving towards a node adjusts the volume of each speaker accordingly while vertical motion controls overall volume, allowing the user to dynamically spatialize sound with one hand.

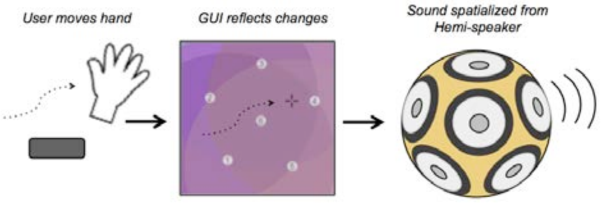


Figure 5c. Hemi interaction schematic

Since these parameters can be controlled with one hand, the implementation was integrated with the motion-augmented granular synthesizer described in section 5.1. The user's left hand controls sound localization while the right hand controls the granular synthesizer. This combination takes full advantage of the multiple degrees of freedom offered by the Leap Motion.

5.3 In Circles composition and performance

In Circles is a composition by the author for cello, Leap Motion and Hemispherical Speaker, premiered at the Yale University Art Gallery. The performance was staged in the Classical Sculpture wing and explored theme of Time, Color, Memory, Fading, and Texture. The recording is available online and is referenced in this discussion [16].

Figure 5d. In Circles performance

The composition In Circles explores the theme of time and memory by recording a live cello solo during the performance and using this material as the basis for processed sound. However, instead of conventional granular synthesis, the author employed the motion granular synthesizer in section 5.1. This allows the author to scrub to specific points in the live recording and musically trigger transients as though playing a "second cello". This represents a duet between the cello's melody and the "memory" of the melody played in the past. In the same way our current experiences can affect our interpretations of memories and vice versa, the author can trigger the granular synthesizer in real time and musically respond to the live performer. The left hand controls overall volume.

The author developed motions that worked in tandem with material recorded live. For example, depressing one hand's fingers consecutively in one sequence while traversing the sample causes transients from different time periods to be triggered. The author found this was particularly effective with musical passages with multiple notes since these produced polyphonic textures not possible on cello (7:08).

The author was able to play and sustain "long notes" by continually depressing fingers over one point in the sample, thereby extending the original note to a new length. The opening of the piece intentionally consists of long notes so the author is able to easily "loop" a portion of the recording using continual depressions (7:42). This was particularly effective when the instrument and granular texture harmonize in thirds (9:45). Audience members were particularly surprised and impressed by the control and expressivity achieved through the device.

During performance however, it became clear that the system in its current iteration has a bug concerning horizontal hand movement and thumb movement. The author found that the thumb's play head would erratically be triggered when the corresponding hand is moved across the origin. This will hopefully be solved with the new tracking API.

Nonetheless, the performance opens new possibilities for using motion-based granular synthesis not only for sound design, but also in a live performance setting. The author was able to play the sampled material expressively like a second instrument.

Future versions will feature a projection of the recorded waveform and play-heads (Figure 5a) so the audience sees this interface during performance. When features such as non-linear interpolations between states of effects and more robust finger tracking are implemented, additional visuals and animations will be projected onto a surface to bridge the gap between the audience's perception of the system and how the performer is actually controlling the sound.

6. INSTRUMENTS SANS FRONTIÈRES

6.1 Vision

An extension of the Leap Motion is usage in a project the author is developing called Instruments Sans Frontières (ISF). The project aims to empower handicapped and disabled patients with musical expression using motion tracking technology and wearable sensors. As previously discussed, many of these technologies exist, but have yet to be applied to people and musicians with disabilities.

Instruments Sans Frontières' first aim to is to contribute to the field of Active Music Therapy in patient recovery and physical rehabilitation. P. Oliveros have developed a musical improvisation interface using webcam tracking for people with severe physical disabilities [20]. The paper noted positive effects such as "increased attention…independence and motivation" in patients using the interface. Secondly, Instruments Sans Frontières aims to create a novel medium for musical expression that enables handicapped patients to perform and improvise with the musicianship of an acoustic instrument.

6.2 Preliminary implementation

Playing an acoustic instrument is practically impossible if the patient does not have motor control over their fingers. However, if the patient is still able to move their elbows, a flex sensor could be used to leverage this range of motion by controlling a parameter such as volume in a software instrument. The Leap Motion will be especially useful for patients who have some control over hand and arm movement. For example, a patient who suffers from cerebral palsy may be shaking constantly but can still control the position of their arms. The Leap Motion's field of detection can be divided into zones. When the patient moves their hand over a zone, a corresponding sound is triggered. This implementation was experimented with a patient through the Yale School of Public Health.

Preliminary testing showed the patient was able to control when the notes in these "zones" were to be triggered. However, the patient voiced an important shortcoming the author overlooked. The patient's remark was "…I don't get how waving my hands in thin air actually makes a sound." Unlike, acoustic instruments, where the sound-producing gesture (plucking a string) is linked with the sound-producing mechanism (the plucked string vibrates), the use of motion tracking produces a discontinuity between these two elements. For first-time users and non-musicians like the aforementioned patient, this can be extremely confusing.

A degree of "tactile" contact with a surface will be incorporated in the future. Perhaps a resonating body could be placed in the patient's hand [21]. The resonating body will vibrate in response to different hand motions over the Leap Motion and also produce the synthesized sound via transducers mounted on the surface.

We also realized that the majority of patients were unable to read musical notation. We implemented an interface similar to the game Dance Dance Revolution [22] where arrows indicate which zone to trigger. We quickly discovered that porting this interface to the Leap Motion left the user very confused. Users were forced to keep track of hand's locations while focusing on the incoming arrows. There was a clear disconnect between the of arrows, virtual zones in the software and the physical space through which the patient interacts.

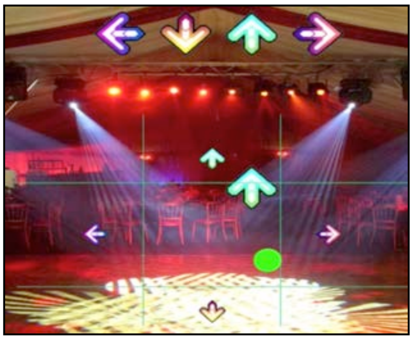


Figure 5e. Zone and arrow interface

Designing for the Leap Motion is as much a problem of interface design as it is a problem of technical implementation. Popular systems for general users like Guitar Hero [23] feature hardware and software that provide a novel interaction, not just an emulation of existing interfaces. Research on existing interface literature will be incorporated in future iterations.

When combined with other mediums, the Leap Motion provides exciting ways to realize the goals of this project. We hope to combine the Leap Motion with a system of projections to make the mapping between user and software more intuitive and imagine new ways of interacting with sound.

7. CONCLUSION AND FUTURE WORK

The Leap Motion opens new and exciting possibilities in motion-based control for musical expression. This paper has demonstrated implementations that take advantage of the device's non-invasive nature and its many degrees of freedom.

The author is particularly excited with the new capabilities offered by the recent beta Skeletal Tracking API for the Leap Motion. The new API provides `hand()` functions that immediately distinguish left and right hands, solving the problem addressed by the author's custom Max patch in section 3.1. The new API also separates tracking data for various fingers, eliminating the need to sort them as previously outlined. In addition, the new API features an integrated model of a hand, so bent fingers are registered as bent fingers and do not simply disappear from view. The user can even turn their hand around and still bend their fingers with tracking. This new feature will hopefully solve the unnatural exaggerations of finger movement the author has employed to ensure detection. This will greatly simplify the finger tracking interactions implemented in this paper.

The author is currently working on a Java progam that integrates the new beta API and performs the implementations prototyped in this paper as a standalone application. For example, the 5-grain granular synthesizer that was prototyped in Max/MSP can be developed into a stand-alone application for live sampling, synthesis and performance. The author aims to make this Java program available to the general public in the near future when a stable Leap Motion firmware is released.

Acknowledgments

The author would like to thank Phillip Wilkinson for his cello performance and feedback on idiomatic cello techniques during the piece *InCircles*.

8. REFERENCES

[1] IK Multimedia. "iRing" [Website] http://www.ikmultimedia.com/products/iring/

[2] Steinberg Multimedia. [Website] "Cubase IC Air" http://www.steinberg.net/en/products/accessories/cubase_ic_air.html

[3] Heap, I. "Gestural Musicware". [Website] 2012. http://imogenheap.com/thegloves/

[4] Jessop, E. "The Vocal Augmentation and Manipulation Prosthesis (VAMP): A Conducting-Based Gestural Controller for Vocal Performance". 2009. *Proc. of the 2009 Conf. on New Interfaces for Musical Expression*

[5] Bongers, B. "Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniques for Real-Time Performance," in *Trends in Gestural Control in Music*, M. M. Wanderley and M. Battier, IRCAM, 2000.

[6] Heins, M. "Vectr". [Website] KickStarter project https://www.kickstarter.com/projects/790206393/theremax-3d-gesture-controller

[7] Bellona, John. "Casting". 2013. Interactive music installation using the Kinect. http://deecerecords.com/music/casting

[8] Little, J. , Hayday J. , Sanderson, P and Delucchi, M. "V-Motion project". 2012. http://www.custom-logic.com/blog/v-motion-project-the-instrument/

[9] Leap Motion Inc. "Device Specifications". [Website] https://www.leapmotion.com/product

[10] Leap Motion Inc. "Developer site" [Website] https://developer.leapmotion.com/

[11] Bellona, Jon. "simpleKinect" http://deecerecords.com/kinect/

[12] Source Audio LLC. "Hot Hand USB". 2013. [Website] http://www.sourceaudio.net/products/hothand/hothand_usb.php

[13] Microsoft. "Xbox One: Get the Facts". 2014. [Website] http://www.xbox.com/en-US/xbox-one/get-the-facts

[14] Bevin, G. "Geco: Multidimentional MIDI expression through hand gestures". [Website] http://uwyn.com/geco/

[15] Akamatsu, M. "aka.leapmotion". 2013. [Website] http://akamatsu.org/aka/max/objects/

[16] For linked media visit (copy-paste into browser): http://lh-hantrakul.com/2014/04/15/linked-media-for-icmc-2014/

[17] Fiebrink, R., D. Trueman, and P. R. Cook. "A meta-instrument for interactive, on-the-fly machine learning." *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2009)*

[18] Smallwood, S., Cook, P., Trueman, D., McIntrye, T. "Don't Forget the Loudspeaker". A DIY guide from http://www.scott-smallwood.com/pdf/delorean.pdf

[19] Wang, G., Bryan, N., Oh, J., Hamilton, Rob. "Stanford Laptop Orchestra". 2009 *Proceedings of the International Computer Music Conference (ICMC 2009).*

[20] Oliveros, P., Miller, L, Heyen, J., Siddall G., Hazard S. "A Musical Improvisation Interface for People with Severe Physical Disabilities". *Music and Medicine July 2011 vol. 3, no.3 172-181.*

[21] Machover, T. "Vocal Vibrations" – an interactive sound installation at Le Laboratoire in Paris, France. http://www.lelaboratoire.org/en/archives-18.php

[22] Konami. "Dance Dance Revolution" [Game] https://www.konami.com/ddr/

[23] Machover, T. "Beyond Guitar Hero – Towards a New Musical Ecology". RSA Journal (London), January-March 2009.