YAŞAR UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

MASTER THESIS

AN AUGMENTED REALITY APPLICATION FOR VIRTUAL INSTRUMENT PLAYING

ONUR AĞBULUT

THESIS ADVISOR: ASST. PROF. DR. KORHAN KARABULUT

DEPARTMENT OF COMPUTER ENGINEERING

PRESENTATION DATE: 13.07.2017

BORNOVA / İZMİR
JULY 2017

We certify that, as the jury, we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Jury Members:	**Signature:**
Asst. Prof. Dr. Korhan KARABULUT Yaşar University	....
Asst. Prof. Dr. İbrahim ZİNCİR Yaşar University	....
Prof. Dr. Aybars UĞUR Ege University	....

Prof. Dr. Cüneyt GÜZELİŞ
Director of the Graduate School

ABSTRACT

AN AUGMENTED REALITY APPLICATION FOR VIRTUAL INSTRUMENT PLAYING

AĞBULUT, Onur
MSc. in Computer Engineering
Advisor: Asst. Prof. Dr. Korhan KARABULUT
July 2017

Augmented Reality has become a trending topic recently in various industries; from automotive to entertainment, telecom, medicine, etc. It provides user a semi-virtual environment by integrating virtual objects on real-world environment based on some arbitrary predefined visual patterns. Besides this, motion sensor devices have also made noticeable progress by enabling end-users to capture their movements without requiring any further setup which makes them portable and easy-to-access. This thesis focuses on creating a virtual instrument with further focus on drums by utilizing the technologies of augmented reality and motion capturing in the comfort of user's home. A setup was created consisting of a user, a computer and a motion sensor device with an integrated camera. The user, drummer in this case, starts to play the virtual drum, in the meantime, the motion sensor device captures the user's movements and the normal color data via its integrated camera and then sends these data to the computer to process. The computer calculates the audio data like the speed of user's arm and leg movements, speed of hits and position of hit drum part based on the input received and plays the appropriate sound for the hit drum part. It also produces the virtual environment and integrates it with the real environment data sent by the motion sensor device and shows the final composed scene environment data to the user on its screen simultaneously with the sound played.

Furthermore, this thesis utilizes both augmented reality and motion tracking which have not been applied together by previous studies. Also it offers more drum parts than previous studies which have only limited number of drum parts. Finally, using dynamic ranges and threshold values for each drum part, the system provides calibration and location-free usage to the players.

This thesis consists of 6 chapters which includes all the mentioned subjects.

Key Words: Augmented Reality, Virtual Instrumentation, Virtual Drum, Motion Sensor, Motion Capturing

ÖZ

SANAL ENSTRUMAN ÇALMA İÇİN BİR ARTIRILMIŞ GERÇEKLİK UYGULAMASI

AĞBULUT, Onur
Yüksek Lisans Tezi, Bilgisayar Mühendisliği Bölümü
Danışman: Yrd.Doç.Dr. Korhan KARABULUT
Temmuz 2017

Artırılmış Gerçeklik son zamanlarda, otomotivden eğlenceye, telekomdan, tıbba, vb. birçok endüstride popülerliği artan bir konu olmaya başlamıştır. Artırılmış gerçeklik, önceden tanımlanmış bazı görsel desenleri temel alarak, sanal görüntüyü gerçek görüntü üzerine entegre ederek kullanıcıya yarı-sanal bir ortam sağlar. Bunun ötesinde, hareket sensörleri de ekstra bir kurulum gerektirmeden kullanıcıların hareketlerini yakalamayı sağlayarak son zamanlarda fark edilebilir gelişme kaydetmiş ve bu durum onları erişimi kolay ve taşınabilir hale getirmiştir. Bu tez; artırılmış gerçeklik ve hareket yakalama teknolojilerini kullanarak kullanıcıların ev ortamında, davul odaklı olmak üzere bir sanal enstrüman oluşturulması üzerine yoğunlaşmaktadır. Bir kullanıcı, bir bilgisayar ve kameralı bir hareket sensöründen oluşan bir kurulum oluşturulmuştur. Kullanıcı, bu durumda davulcu, sanal davulu çalmaya başlar ve bu sırada hareket sensörü entegre kamerasıyla kullanıcının hareketlerini ve renkli kamera verisini yakalayıp işlemesi için bilgisayara gönderir. Bilgisayar, alınan girdiye göre kullanıcının vuruş hızı ve vurulan parçanın pozisyonu gibi verileri kullanarak gereken ses çıktılarını hesaplar ve vurulan davul parçası için uygun sesi oynatır. Bilgisayar aynı zamanda bir sanal ortam oluşturup, hareket sensörü tarafından gönderilen gerçek ortam ile entegre ederek, son oluşturulmuş sahne bilgisini ses ile birlikte eş zamanlı olarak bilgisayar ekranında kullanıcıya gösterir.

Bunların dışında, bu tez çalışması, daha önceki tez çalışmaları tarafından şu ana kadar bir arada kullanılmamış olan artırılmış gerçeklik ve hareket yakalama teknolojilerinin ikisini bir arada kullanır. Ayrıca, daha önceki çalışmalarda sunulan kısıtlı sayıdaki davul parçasına nazaran daha fazla davul parçası sunar. Son olarak, her bir davul parçası için dinamik menzil ve eşik değerleri kullanarak, kullanıcılara sistem kalibrasyonu ve konum bağımsız kullanım sağlar.

Bu tez yukarıda bahsedilen konuların hepsini içeren 6 bölümden oluşmaktadır.

Anahtar Kelimeler: Artırılmış Gerçeklik, Sanal Enstrüman, Sanal Davul, Hareket Sensörü, Hareket Yakalama

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Assist. Prof. Dr. Korhan KARABULUT for his valuable support, guidance, advice and patience on my thesis. This thesis study would not have been possible without him.

I would also like to thank to my sister, for her valuable support as an acting drummer while capturing and observing movements during Kinect phase.

I would like to express my enduring love to my parents, who are always supportive, loving and caring to me in every possible way in my life.

Onur AĞBULUT

İzmir, 2017

TEXT OF OATH

I declare and honestly confirm that my study, titled “An Augmented Reality Application for Virtual Instrument Playing” and presented as a Master’s Thesis, has been written without applying to any assistance inconsistent with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Onur AĞBULUT

……………………………………..

July 31, 2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLS AND ABBREVIATIONS

ABBREVIATIONS:

AR	Augmented Reality
BPM	Beats per Minute
FPS	Frames per Second
IR	Infrared
MIDI	Musical Instrument Digital Interface
SDK	Software Development Kit
VR	Virtual Reality

CHAPTER ONE
INTRODUCTION

1.1. Subject of the Thesis

Virtual instrumentation has been a popular topic for the last decades, notably from the famous head-mounted display work of Ivan Sutherland in 1960s (Sutherland, 1968). Virtual instrumentation offers musicians vast amount of opportunities to produce sound. Initially, the studies were towards to produce sounds by emulating the existing real-worlds instruments, however, over the years, producing completely new sounds that even does not exist in the nature became feasible. While it greatly enhances the creativity for musicians, as it offers an excessive combination of sounds, it also greatly facilitates the production of audio as compared to their original counterpart real instruments, since virtual instruments offer great portability and flexibility.

Along with the improvement of virtual instrumentation, other technologies like augmented reality, virtual reality and motion tracking sensors also made a great progress. Augmented reality (AR) offers a great way to integrate the visualization of computer generated 3D objects into real-world which immerses the user by connecting them to both real and imaginary worlds. Within virtual instrumentation context, AR offers the visual representations of a musical instrument via any peripheral device that is capable of gathering camera data and a computer. Moreover, it also assists to the visual enrichment of these instruments by producing assisting visuals onto real instruments using projection methods. Moreover, studies have been done to enhance the process of learning to play musical instruments by assisting the novice players visually, and projecting visual guides on the instruments.

Similar to AR, virtual reality (VR) offers a totally immersive imaginary world. Besides AR and VR, motion tracking technology has improved greatly as well, so that, the devices that were previously affordable only by corporate firms became available for everyone. In addition to that, it required a lot of work to put them into action, whereas, nowadays they have become more feasible as well as practical to be used by the end-users.

Combining these thriving technologies, virtual instrumentation now has a lot to offer to both professional musicians, as well as, to novice players. Professional musicians can experiment new sounds or customize the instrument to their own needs with much more flexibility and wider sound range than real instruments offers. New learners also can greatly benefit; when combined with motions sensors, augmented reality can offer learning through gamification and visualized aids on the instruments.

Within the light of these topics, in this thesis study, a virtual drum was developed which offers players a more affordable, as well as, portable drum set using a computer and a motion sensor powered by augmented reality. It offers players a traditional drum set in a virtual way, with the benefits of affordability and portability.

Mainly, this thesis study consists of augmented reality, motion capturing, virtual instrumentation and related sub-components. During development, AR technologies and patterns were researched and reviewed. In addition, integration with other systems were researched. Motion capturing devices were researched and experimented. Virtual instrumentation technologies were researched and appropriate methods were selected.

The motion capture devices, motion detectors and motion sensors that are currently used in the industry were reviewed, their capacities and conformities with this thesis subject were experimented and results were examined.

Augmented reality techniques, applications with/without markers were researched and results were examined. Producing a virtual drum via utilizing augmented reality techniques were researched, experimented and applied. Customizations of augmented reality produced scenes, assets were also researched.

Playing sounds by computer, based on the calculations of output data of motion sensor devices was researched and results were examined and appropriate implementation was applied.

Enabling user to customize the drum set, environment and applying various sound filters were researched.

1.2. Aims and Problem Definition

Traditional drum sets have been widely used for a long time, but their size and high-level sound output they produce, make them less likely to be used especially in home environments. With the recent development of virtual instrumentation, namely, augmented reality and motion tracking, a lot of studies have been made to offer players virtual instruments with better sound ranges. In addition, virtual instrumentation helps players learn playing the instrument faster than traditional methods by facilitating the learning process using interactive methods and feedbacks.

However, previous researches generally focused on guitar or other instruments that are similar in size like Dombyra which is a traditional musical instrument of Kazakhstan (Zhaparov & Assanov, 2014). Notable amount of studies has been made to ease the learning curve of it by providing visual assistance, especially for guitar. For example, a study made at Keio University presents a virtual assistant that shows the correct positioning and combination of fingers for the respective chord on the fingerboard of a guitar via a virtual hand (Motokawa & Saito, 2006).

Regarding previous virtual instrumentation studies for drums, there are only a few researches and they mainly lack of a few important points. They either combine augmented reality with traditional drum sets without motion sensors, mainly for feedback purposes or integrated motion sensors without utilizing AR. Thus, they are unable to fully combine augmented reality with motion sensors to create a virtual drum set.

Hence, the aim of this thesis is to produce a virtual instrument, namely drums, which enables a novice player to learn playing in an easier way or an experienced player to improve his/her skills by utilizing missing points from previous studies with the least amount of hardware needed: only a motion capturing device and a computer, without requiring a real drum set. Furthermore, this thesis also aims to achieve this with an acceptable response time, beats per minute (BPM) and efficient playing of sound files powered by AR and motion capture.

During the study, working principles, specialties and capturing area capabilities of motion capture devices were studied. The skeleton tracking system, internal capturing algorithms used in them and their ability to be integrated with other software solutions were also researched.

Firstly, motion capture devices, motion detectors and motion sensors which are used to capture movements of user were researched in detail. Each type of device was experimented, in terms of field capacity of capture, processing and response time and affordability. The results were analyzed and reviewed.

In general, capturing devices are grouped into two types. The first group focuses on capturing user’s skeleton system by identifying only main bones, whereas, the second group of devices focuses only on specific small part of user’s body, like arm bones. Skeleton-wide capture devices offer a wide capturing area for the user. Furthermore, some recent versions can track more than one user at the same time. In contrast, second group of devices offer a very small capturing field, but generally they have more precise tracking. Through this thesis work, three different capturing devices were experimented; two from the first group and the third one was from the second group. After the results were examined, it has been decided to proceed with the device which belongs to the first group.

Methods to do calculations based on user’s movements and detecting areas hit on drum set were also researched and experimented via tracking a real user while playing it.

Playing the appropriate drum sound using pre-recorded sample audio files and software solutions were studied. Techniques and working samples were researched.

Augmented reality techniques both with and without markers to produce a virtual drum set in front of the user were researched. Their integration, applicability on the other functional parts of the system were also researched.

1.3. Context of the Thesis

This thesis is about virtual instrumentation, in particular, creating a virtual drum that has a high response time via utilizing motion capturing and augmented reality. In the initial phase of this thesis work, various motion capture devices were examined, experimented and their corresponding effectiveness were analyzed. Previous works similar to this thesis were also researched, studied and based on their results, improvement areas were determined.

Multiple motion sensor devices including Leap Motion, Kinect version 1 and Kinect version 2 were experimented. As for effectiveness and regarding ability to capture wider area and whole skeleton of humans, Kinect version 2 was selected as the motion capture device. Different software libraries that can support Kinect version 2, augmented reality, and sound playing etc., were examined. Unity was selected as the composer software to process motion capture output, calculate audio data, play the appropriate sounds and display computer generated imagery for realizing augmented reality.

Several experiments were done with a player using Kinect version 2 and Unity in order to create a virtual drum system that tracks and detects user's movements and responds accurately.

Audio solutions were researched and experimented. An algorithm was developed to play appropriate sound files in Unity framework using appropriate calculations.

Augmented reality techniques and corresponding software solutions were researched. Integration with Unity and other parts of the system was studied and a solution was created.

After the whole system was developed and tested, it was experimented by various players. Players' feedbacks were gathered for future improvements.

1.4. Structure of the Thesis

The remainder of this thesis is structured as follows. Chapter 2 gives general information about augmented reality. Chapter 3 gives details about motion capture and motion capture devices that were experimented during the thesis study.

Virtual instrumentation is described in Chapter 4 and Chapter 5 summarizes the previous studies on virtual instruments in the literature. The details of the AR drum set generated in this thesis study is presented in Chapter 6. Finally, summary and future work is given in Chapter 7.

CHAPTER TWO

AUGMENTED REALITY

2.1. Description of Augmented Reality

Augmented Reality (AR) is a trending topic in a lot of industries nowadays; ranging from automotive to medicine, shopping to movie industry. AR integrates virtual 3D objects into a real-life scene using some pre-defined pattern images. This makes AR different from Virtual Reality (VR) which rather creates a fully immersive virtual environment for the user. Also, later on, AR is improved to be capable of integrating virtual objects without using a pre-defined pattern images.

While it has been known since back in 1960s with the invention of head-mounted display by Ivan Sutherland (Sutherland, 1968), AR has become popular in the recent years, especially with the introduction of mobile phones with integrated cameras.

AR has lots of applications in various number of industries. One of the famous examples is the National Geographic's Mall Example (Figure 1). It has been set up using an enormous screen, a camera and a marker to create AR characters when triggered by people standing on it. It contains lots of different virtual interactive things from animals like tigers, dinosaurs, and to weather effects like lightning and storm (Hepburn, 2011).

Figure 1. National Geographic's Mall Example with AR Dinosaurs (Hepburn, 2011)

On the other hand, AR has also been heavily used in navigation systems as well (Figure 2). The directions to destination selected by the driver are shown by integrating virtual markers onto a real-world street-view map which makes it useful for drivers.

Figure 2. AR markers in car navigation system (Route66Training, 2011)

Beyond these applications, AR is also aimed to be used in this thesis study as well. Since the musical instrument is virtual, there is no real drum set and the user is able to see the drum only on the screen which is generated by the AR software. It produces the virtual drum by using the data captured by Kinect's camera and integrating it onto the real-world scene which is seen by the user, as it is fully-integrated onto the scene on the computer display.

2.2. Existing Augmented Reality Frameworks and Libraries

Currently, there are various technologies on the market that focus on AR. Qualcomm, which is well-known for producing microprocessor chips for mobile devices, has made a great investment in AR. Its subsidiary Vuforia developed a framework called Vuforia Augmented Reality Software Development Kit (SDK), which is one of the most widely used proprietary AR solutions in the industry. Although it is a proprietary solution, it is free of charge for use in development projects. Later, it has been acquired by PTC in 2015 (PTC Completes Acquisition of Vuforia, 2015). A sample application developed using Vuforia AR SDK is shown in Figure 3.

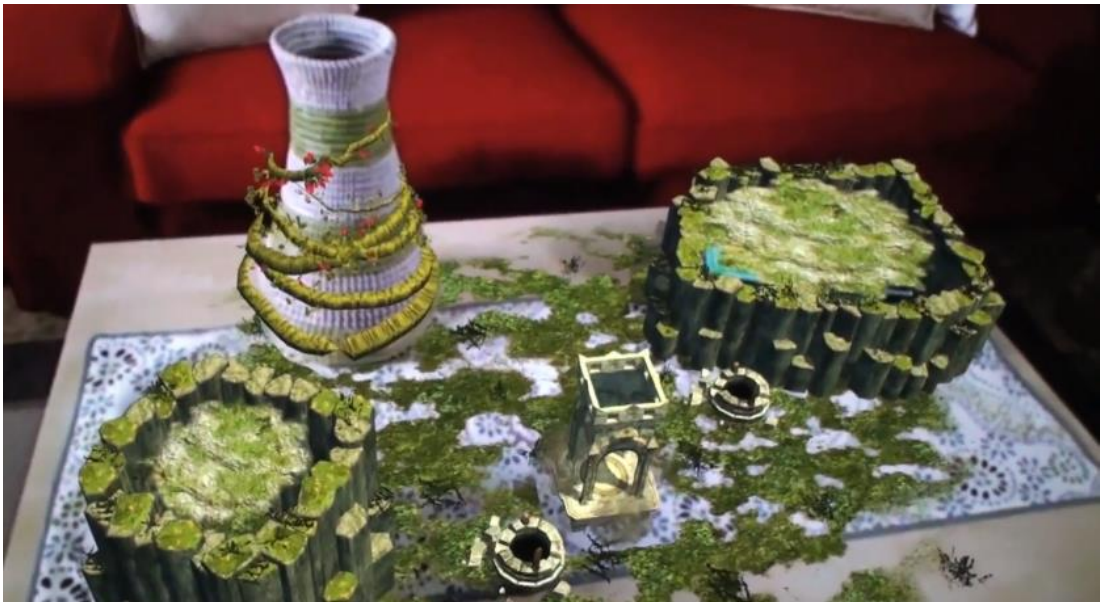

Figure 3. Vuforia’s Smart Terrain Augmented Reality Demo (Gunther, 2013)

Beyond Qualcomm’s Vuforia, there are other proprietary solutions available, like AR-Media SDK by Inglobe Technologies and Wikitude SDK by Wikitude GmbH. However, unlike Vuforia and Wikitude, AR-Media SDK is not free. On the other hand, there several open source solutions available, like Augment, ARToolKit, etc. ARToolKit is one of the most advanced AR frameworks and is available as open source.

Besides these solutions, existing game engines on the market have started to provide support for different AR libraries. Unity, one of the most preferred game engines, especially for mobile platforms, has support for Vuforia and Wikitude. Hence, a developer can easily integrate Vuforia or Wikitude into Unity and start developing AR applications easily. Unreal Engine, another widely used game engine, especially for AAA titles like Deus Ex and Tom Clancy’s series, has its own AR solution called Unreal4AR, which is mainly powered by the open source ARToolKit library.

In addition to these, hybrid and custom tailored solutions for specific systems also exist. Zigfu for instance, offers its SDK called ZDK to support development of Microsoft Kinect applications and has also support for Unity. Thus, AR applications supported by Microsoft Kinect can easily be developed using Zigfu, Kinect and Vuforia.

CHAPTER THREE
MOTION CAPTURING

3.1. Description of Motion Capturing

Motion capturing devices or motion detectors are hardware solutions that track the movements of humans which are within their detection ranges. They are heavily used in gaming and filming industries, as well as in sports, military, computer vision and robotic applications.

The main principle of motion capturing relies on sampling the movements of actors in front of it. Especially in video games and filming industry, the captured movements of actors are mapped onto an animated 3D visual model whose static state is generated by an artist beforehand. The actors generally wear specific tailored suits which hold some markers near on every joint to be identified by the devices (Figure 4).

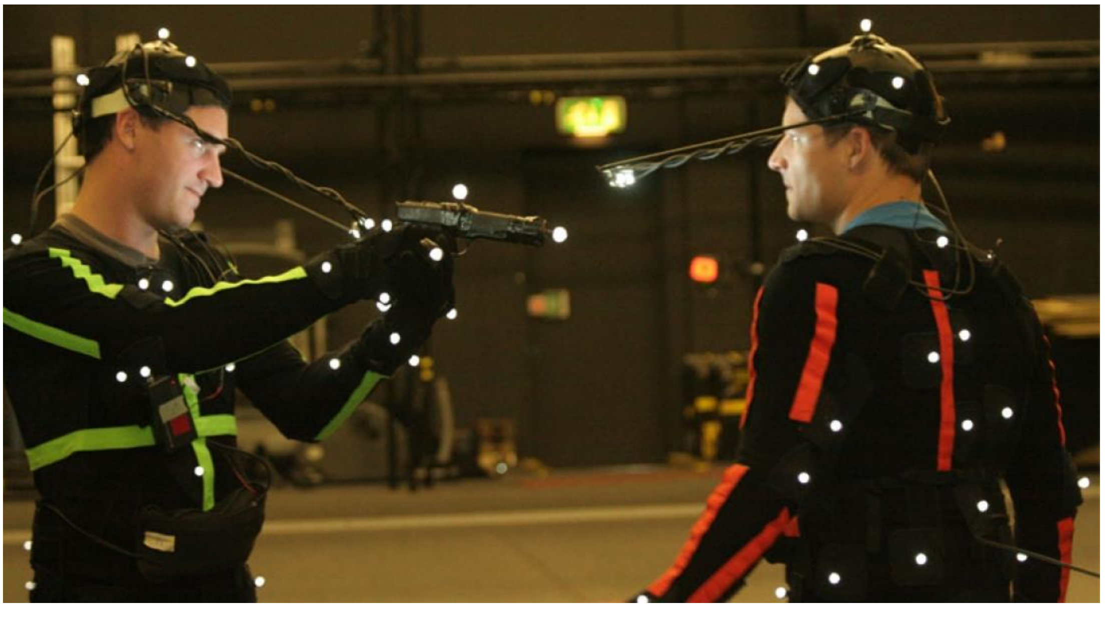

Figure 4. Actors with markers in a motion capture scene (Nieler, 2015)

In today's gaming industry, most of 3D game development studios use this technology in their games, i.e., GTA V by Rockstar Games (2013), Heavy Rain by Quantic Dream (2010), etc. On the other hand, James Cameron's famous movie Avatar is one of the examples in the filming industry that uses this technique in most of its scenes. The whole bodies of actors are captured in general, however, specific parts of the body can be captured as well; like facial capture or capturing of hand bones or ankles.

In the near past, a detective genre video game called L.A. Noire that uses this technology was developed by Team Bondi in collaboration with Rockstar Games. It offers realistic face mimics, where players can come to a decision only by using in-game characters' mimics during the interrogation. During the development of the game, the developer studio heavily used 32 cameras which have been placed in such a way to capture on every aspect of the actor's face to immerse the actor completely with 360 degrees capture and map any mimic on their faces to the 3D models in the game (Peckham, 2011).

These motion capturing applications depend heavily on usage of lots of expensive hardware. However, some affordable hardware solutions were also developed recently; Microsoft Kinect and PlayStation Move, to name a few. These motion tracking devices include an integrated camera and a microphone which help tracking of user's motions and sounds that are in their capture range. Then, this tracked data is sent to the receiver software to process for various purposes by different applications, like gaming, Augmented Reality or Virtual Reality.

As Microsoft and PlayStation's announcements of these devices were primarily for their gaming consoles, these devices were developed with a primary focus on gaming, in order to enable the player to take part in interactive games. Especially Microsoft's Kinect takes a step forward by capturing player's movements without any attachment on the player, while PlayStation's Move requires the players to hold Move devices in one of their hands.

Although these devices are primarily designed for gaming, they can also be used for other purposes as well, such as, 3D virtual modeling, robotics etc. Philipp Robbel from MIT created a robot which reacts to human gestures. By using Kinect and iRobot Create, he also modeled a 3D room where the robot can operate (Wortham, 2010).

Apart from Microsoft's Kinect and PlayStation's Move, there are other devices as well, like Leap Motion from Leap Motion, Inc. Contrary to Kinect and Move, Leap Motion focuses on tracking of hand gestures of user while offering only a small range for user input (Figure 5).

Figure 5. Leap Motion connected to a laptop (Aaron, 2013)

3.2. Motion Capturing Devices / Motion Sensors

3.2.1. Leap Motion

Leap Motion is a motion sensor hardware developed by the company with the same name. It primarily focuses on the capturing of hand movements of the users. The company was founded by David Holz with his childhood friend Michael Buckwald in 2010. The idea was originally developed by David Holz in 2008, while he was studying for his PhD degree. Over 500,000 units were sold up to present (Etherington, 2014).

The size of Leap Motion is quite similar to the size of a USB flash disk. It was actually designed to work on a physical desktop environment and consists of two infrared (IR) cameras and three IR LEDs which give the device a hemispherical tracking area of about one meter. Tracked data is sent to the computer via a cable where this data is processed in Leap Motion software. The accuracy of the device is roughly about 0.7 millimeters (Weichert, Bachmann, Rudak, & Fisseler, 2013).

The small range of tracking and tracking of a specialized part of the body differentiates Leap Motion from Kinect and PlayStation Move, where the former offer full-body tracking and wider area range for tracking. Leap Motion is more suitable to use for general day to day computer usage, like navigation of a website, zooming in and out by pinching, like it is done on the touch screen of mobile devices. Since it has limited area of tracking, capturing two full hand movements can be merely possible, whereas movements of one hand can be captured very accurately.

It also offers an SDK for developers which enable users to develop applications using Leap Motion. An application market has also been presented to the developers, which enables them to publish their applications using its SDK, aimed to be used with Leap Motion sensor. Games and other applications such as playing a piano, controlling Windows operating system, etc. have been developed. Recently, the company announced a new software update to Leap Motion which enables users to use Leap Motion for hand tracking in Virtual Reality environments.

3.2.2. Microsoft Kinect Version 1

Microsoft Kinect is a device that contains a set of motion sensing devices. It was primarily developed for Xbox game console and then made available for PCs. It was first announced at the world's one of the most famous gaming expo, E3, in 2009 (Pham, 2009). Later, it was launched in late 2010 for Xbox 360 to enable interactive gaming. About two years later, in February 2012, its Windows counterpart was released (Eisler, 2012).

Kinect version 1 (Figure 6) includes an RGB camera to capture normal color camera data, a set of microphones to capture sound in the environment and finally, a depth sensor which captures depth information using an IR laser projection. On top of these, a proprietary software runs which was developed by Microsoft Game Studios department of Microsoft. Fundamentally, this proprietary software provides motion capture for full body of a human, as well as facial and voice recognition. It has capacity to track up to two actively moving players. This number can go up to six persons that are static (not moving), by limiting these persons to be in the range of device. It is also able to capture facial mimics of the user. Beyond facial and body tracking, Kinect also offers voice recognition. However, voice recognition works only in some countries including United States, United Kingdom, and Japan etc. On the other hand, this voice recognition feature enables users a headset-free chat with their friends using Xbox.

Figure 6. Microsoft Kinect version 1 (Gell, 2014)

Kinect was primarily announced with focus on gaming audience. Lots of games have been released at the same time with the launch of Kinect and afterwards, as well. After three years from the release, Microsoft launched an SDK for Kinect to allow developers to create applications for it for non-commercial purposes. Following this, after a short time, Microsoft released a program which allows developing commercial software for Kinect. With these recent upgrades, companies began to develop different types of applications for Kinect other than games as well.

3.2.3. Microsoft Kinect Version 2

About four years later, a sequel to Kinect, named Kinect version 2 (Figure 7) is launched for Xbox One consoles. Kinect version 2 offers lots of improvements from its prequel, which include capturing support for more joints and more people, a more powerful integrated camera and some other improvements to the its other subsystems. As with Kinect version 1, after a period of time, a version for Windows was also released for Kinect version 2. Following this release, a much-improved version of the SDK was released as well.

Figure 7. Microsoft Kinect version 2 (Bravo, 2014)

3.2.4. PlayStation Move

As with Microsoft Xbox, Sony, which is a rival company of Microsoft on the gaming console market, has launched its own interactive gaming product PlayStation Move which is a motion sensor device like Kinect. PlayStation Move is only compatible with PlayStation gaming console and cannot be used in PCs. Contrary to Microsoft Kinect, PlayStation Move is more of a game controller and requires the player to hold it whenever they play a game or use an interactive application, much like Nintendo's Wii Remote device. It also requires a PlayStation webcam.

PlayStation Move (Figure 8) consists of two parts, a motion controller and a navigation controller. They fundamentally detect user's motion by using their inertial sensors, which is also tracked by PlayStation webcam. The combined data gathered from these two devices sent to the console for processing.

Figure 8. PlayStation Eye and Motion Controller with a copy of Sports Champions (Koller, 2010)

As with Kinect, PlayStation Move’s primary focus was on gaming and interactive applications. About 15 million PlayStation Move devices were sold by the end of 2012 (Yin-Poole, 2012). Despite high selling numbers, Sony’s PlayStation Move could not make the huge impact the company aimed.

CHAPTER FOUR
VIRTUAL INSTRUMENTATION

4.1. Description of Virtual Instrumentation

Virtual Instrumentation is creating a digital version of a real musical instrument by emulating its sounds completely virtually. This process generally includes utilizing of pre-recorded samples of the considered instrument as well. A standard known as Musical Instrument Digital Interface (MIDI), provides an interface to enable transmission of data between the input device and output sound creating device or software. Its data format includes lots of information like the note, pitch value, etc. MIDI also enables producing various different sounds just by using a keyboard controller (Figure 9).

Figure 9. A MIDI controller connected to a computer (Earl, 2010)

MIDI's history goes back to 1980s, where its first initial idea was generated by Ikutaro Kakehashi in 1981. Two years later in 1983, after a series of experimentations by multiple engineers, Dave Smith created a MIDI connection between two synthesizers. Following a very short time, again in 1983, it has been standardized (Stewart, 2013). MIDI has leaded the creation of different kinds of music genres like house, electronic, etc. Currently, it is widely used in the numerous industries as well.

Besides MIDI devices, there is also another type of hardware devices used for virtual instrumentation called “synthesizers”, which apart from MIDI, have their own integrated audio components. They are able to produce sound without requiring any other sound module.

Apart from hardware devices, with the recent improvements in the technology, software versions of these devices, called software synthesizers became available. Fundamentally, they mimic the audio behavior of hardware synthesizers using a high number of audio samples in their internal libraries. Contrary to the hardware versions, they run on computers and use the resources like CPU, memory, etc., available in the system they execute on. As they do not need any extra hardware devices or have no other costs, they are more portable, as well as, affordable than their hardware counterparts. Moreover, since recent computers have faster processors and more resources available, software synthesizers became capable of processing more advanced algorithms in relatively shorter times compared to their hardware counterparts. Thus, lags during the performances are also minimized which might lead to an unpleasant experience on computers with very limited CPU and memory. Additionally, with the recent advances in the mobile devices’ hardware systems, synthesizers at very basic levels with limited capabilities also became available for mobile operating systems like Android, iOS and Window Phone (“Software synthesizer,” n.d.).

Virtual instrumentation devices, namely MIDI devices, software and hardware synthesizers, are used in a wider areas by sound designers, composers and musicians as well as amateur home musicians. As software synthesizers in particular are more affordable and portable, they are used in music production, as well as to emulate the sounds from traditional instruments, where they are more effective to output the desired sound quality. Besides, they remove the requirement to have actual players to play the instruments, since it can easily be emulated by other people, which decreases the cost of producing music substantially. Beyond that, there are commercial and open-source applications available like Apple Logic Pro, GarageBand, Audacity, Ableton Live and among them GarageBand (Figure 10) comes out of the box with every MacBook or iPad device.

Figure 10. Apple GarageBand Interface (Sethi, 2015)

CHAPTER FIVE
PREVIOUS WORKS

Virtual Instrumentation (VI) has been a great interest for musicians to extend the limits of the existing traditional instruments as well as to invent new sounds. Thus, various studies regarding VI have been done. With the newly developed technologies, together with the relevant technologies like AR and VR, building virtual instrumentation devices became more feasible than before, which leaded to a rise in the studies recently. While majority of the studies focus on string-based instruments like guitar, violin etc., there are also a few studies focusing on other traditional instruments like drum and piano. Besides, there are some other works that experimented some national instruments like Dombyra and SpiderKing.

Studies in the literature generally try to address objectives that can mainly be classified in three categories: decreasing the learning curve for novice players by providing more feedback and aid, extending the traditional instrument by facilitating generation of additional sounds and lastly, inventing new sounds through experimentation and synthesis that are not possible with traditional instruments.

5.1. An AR Drum Kit

This thesis study which was developed in Waseda University is an AR application to teach players how to play drums by offering game-learning methodologies, which is also known as gamification. A system was setup using two mobile projectors, an ambient feedback area, like a wall painted with solid color, a PC and an electronic drum set.

The PC monitor displays vertical lines, where there is a different part of the instrument at the bottom of each vertical bar. There are also small horizontal bars throughout the vertical bars, which refer to notes flowing from top to bottom and user is supposed to hit the corresponding drum part when that specific horizontal line hits at the end of the vertical bar (Yamabe, Asuma, Kiyono, & Nakajima, 2011). The general setup of the developed drum kit is shown in (Figure 11).

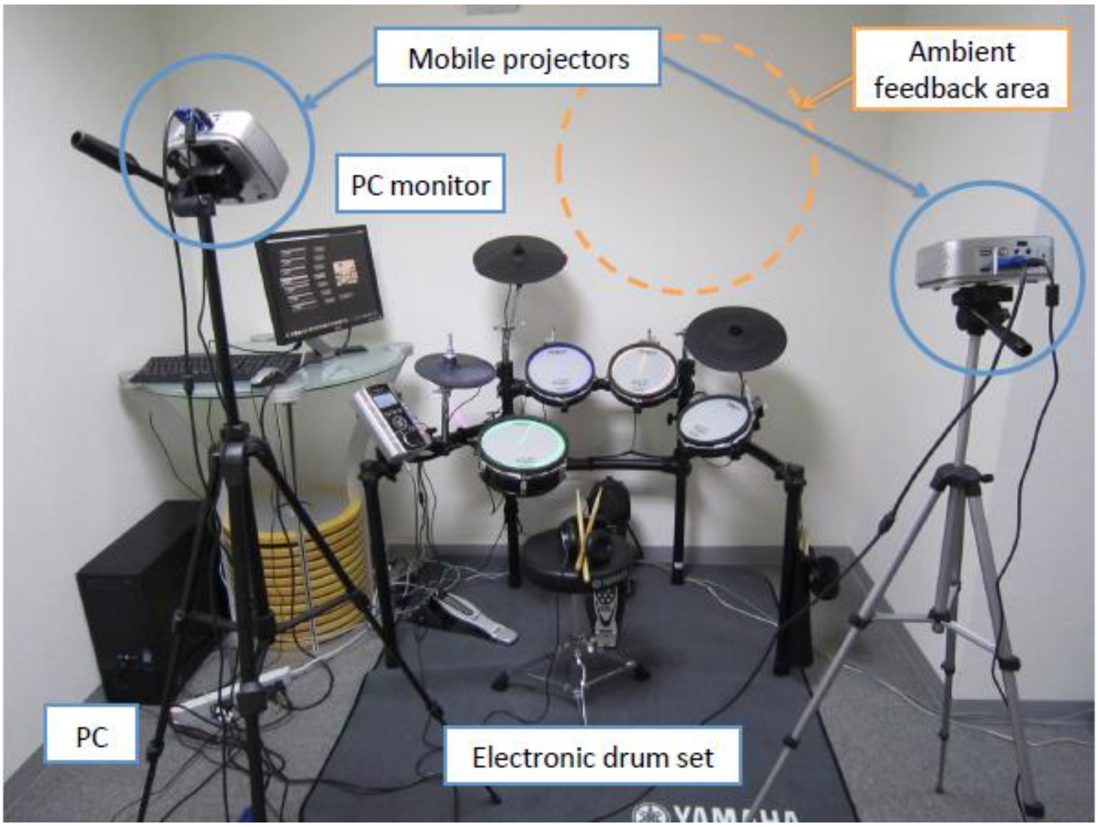

Figure 11. The setup of the AR Drum Kit by Waseda University (Yamabe, Asuma, Kiyono, & Nakajima, 2011)

Drum was placed in front of the feedback wall, and PC monitor was placed behind the drum kit. One of the projectors was aiming the pads of drum while the other projector was targeting to the ambient area. When the player starts to play the drum, the drum sends MIDI data to PC to process. The projector which is directed to drum pads shows feedback on drum parts based on how accurate the user hit the corresponding drum part, whereas the other projector and PC monitor show notes to play as a guide (Figure 12).

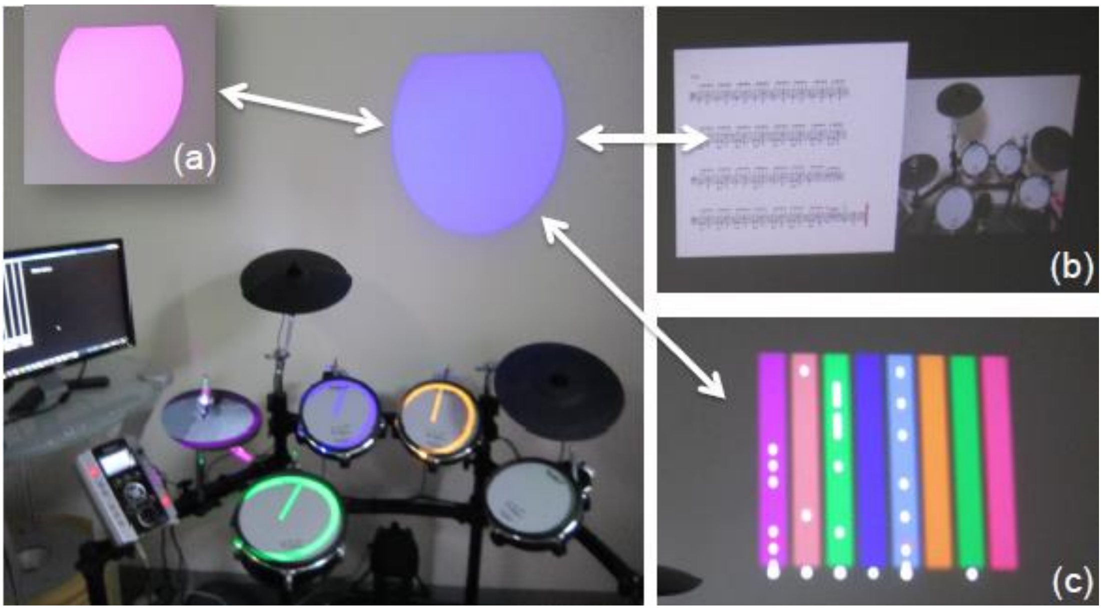

Figure 22. Projection feedbacks on the wall and on the drum parts (Yamabe, Asuma, Kiyono, & Nakajima, 2011)

While this project offers a good combination of a drum set and AR, it only uses AR, and lacks motion sensors. The system requires to have a real electronic drum set in the scene. Moreover, it uses lots of other devices like projectors, PC screen etc., which is not a very practical setup for a home user. Furthermore, since there are lots of I/O devices, player's focus can easily be distracted as well.

5.2. Spider King

Spider King is a thesis work that aims to enable users to play three different virtual musical instruments, namely; drum, guitar and Spider King. It was developed by four people in two universities, National Central University Jhongli in Taiwan and University of Aizu in Japan. The design of Spider King drum over the shoulder of the player is shown in Figure 13.

Figure 13. Design of Spider King drum (Hsu, Kumara, Shih, & Cheng, 2013)

This system utilizes Microsoft Kinect as the motion capturing device and includes numerous software solutions in order to process the Microsoft Kinect data and produce music for different musical instruments. Compared to previous work, this system does not require so many hardware devices in its setup. It only needs a Microsoft Kinect device and a PC to build the system, which makes it easy to setup. Furthermore, a live concert has been organized using the three musical instruments mentioned above (Hsu, Kumara, Shih, & Cheng, 2013). The feedback output and trigger points for different parts of the drum set is shown in Figure 14.

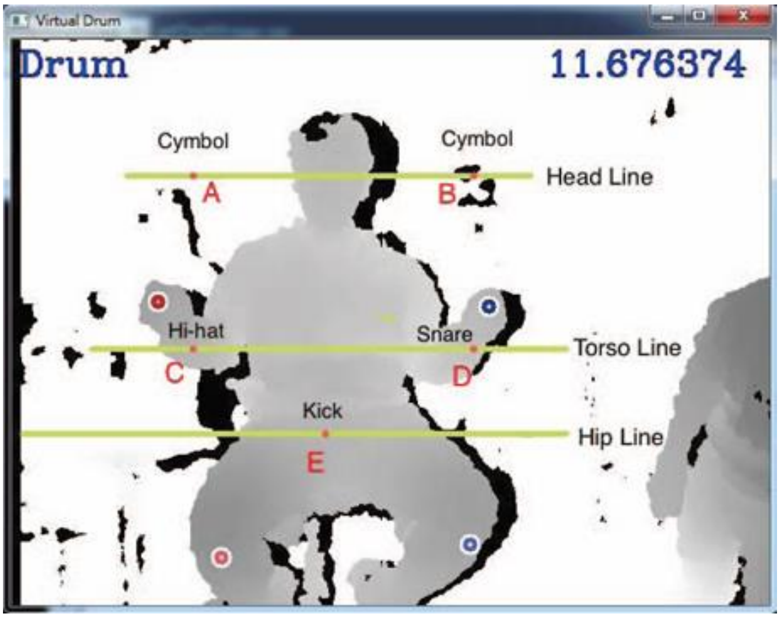

Figure 14. The feedback output on computer display (Hsu, Kumara, Shih, & Cheng, 2013)

In theory, all of three types of instruments seem to work properly, however in practice, performance issues raise due to the very low response rate about 11 frames per second (FPS). Moreover, since Microsoft Kinect can only capture joints, tracking fingertips to play guitar is nearly impossible. Beyond this, although this thesis uses Microsoft Kinect, it does not integrate any AR system or any other technology to give the player a sense of experience of playing the instrument.

5.3. An Augmented Reality Based Learning Assistant for Electric Bass Guitar

This study, developed by three researches at IMAG in France, aims to ease the learning curve of playing electric bass guitar for the novice bass players by utilizing

Augmented Reality. Essentially, it guides the user for the next note to play on the bass guitar and produce the sound beforehand internally from its own sound library, and while in the meantime, it displays the position of the note on the fingerboard of the actual instrument and prompts the user to find the correct note position and play that note. After the user finds and plays the note, the system proceeds to the next note as long as the user plays the note correctly. The system keeps waiting on the same note in case user fails to locate the note on the fingerboard. It determines whether the user played correctly using Augmented Reality by getting the position information of user's finger and neck of bass guitar (Cakmakci, Bèrard, & Coutaz, 2003). Figure 15 shows the markers on the guitar and the player's finger used for recognition and the visual note indicator as a dot on the fingerboard of the bass guitar.

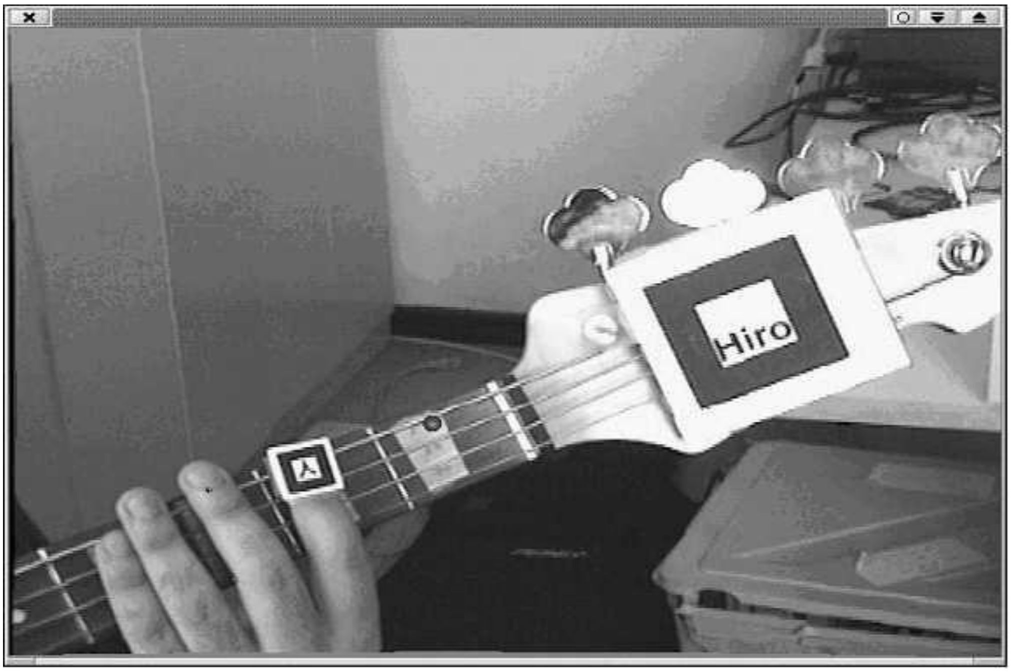

Figure 15. AR markers and visual note indicator on the fingerboard (Cakmakci, Bèrard, & Coutaz, 2003)

Overall, while it is not closely focusing on drums, this project utilizes AR in a teaching manner. Moreover, it decreases the time for new beginners to get used to playing the bass guitar which may become a heavy and time-taking process when utilizing traditional learning methods.

5.4. Support System for Guitar Playing Using Augmented Reality Display

This thesis study works in a similar way with previous study titled “An Augmented Reality Based Learning Assistant for Electric Bass Guitar”, while differentiates itself in terms of the visual style it provides. Instead of displaying a dot as a guide on the fingerboard and playing a sound produced by the system, this thesis study shows a virtual 3D model of a hand placed on the correct chord position, as well as, with the properly placed fingers for the respective chord. It provides an enhanced visual guide compared to the previous study; whereas, it uses a marker on the neck of the guitar to detect the orientation of the device properly in the same way with the previous study (Motokawa & Saito, 2006). Figure 16 shows the developed system that uses the marker and visual guides for playing chords.

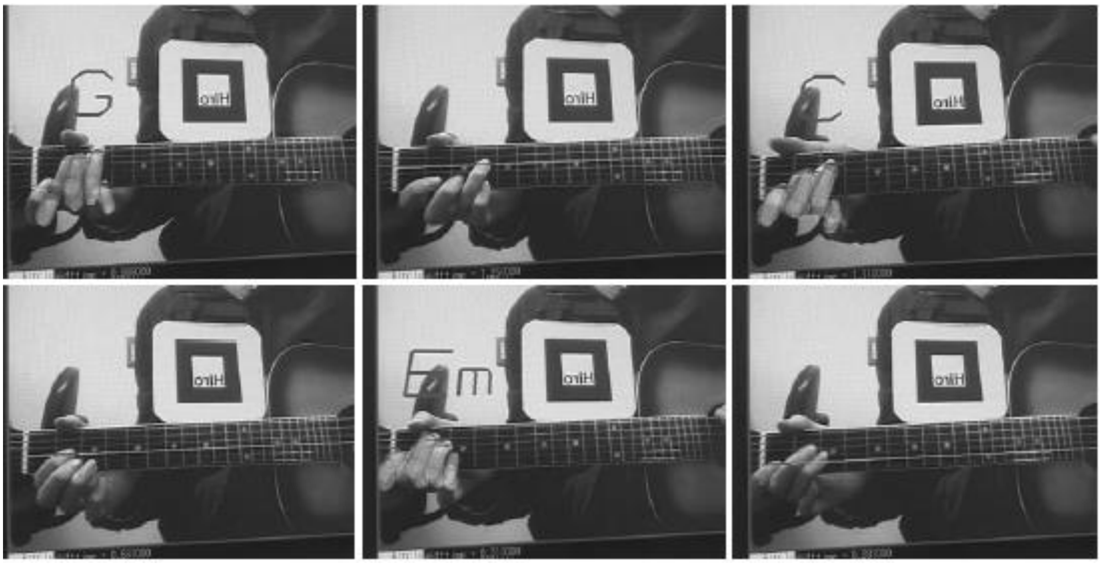

Figure 16. Virtual 3D hand model on the fingerboard (Motokawa & Saito, 2006)

Similar to the previous study, this work also remarkably facilitates learning to play the guitar for beginners. Also, it conforms well to the conditions of the environment the player is in, as well, as the orientation of the guitar facing the camera by utilizing AR.

5.5. Augmented Reality Based on Kazakh Instrument "Dombyra"

As its name suggests, this research focuses on traditional Kazakh instrument called “Dombyra”. Unlike the previous studies, this study aims to enable players to play Dombyra instead of easing the learning process for new beginners, It achieves its goal via utilizing AR on mobile devices using prerecorded sound samples of

Dombyra. It works with predefined markers, similar to the previously described AR project examples. A tablet was chosen to run the application since it is portable and it has an integrated camera on it which is used to capture images from the environment. It captures the marker position on the environment visible to it to feed the AR with it. Then, AR library detects whether it is a valid marker via scanning its predefined set of marker images. Upon detecting validity of marker successfully, application creates a virtual Dombyra over the camera captured image on the screen to enable the user to play it using the visual buttons on the displayed instrument. Along with this, the application has other on-screen buttons to enable the user to control AR generated virtual Dombyra on the screen by translating, rotating and scaling (Zhaparov & Assanov, 2014). Figure 17 shows the computer generated Dombyra on the screen.

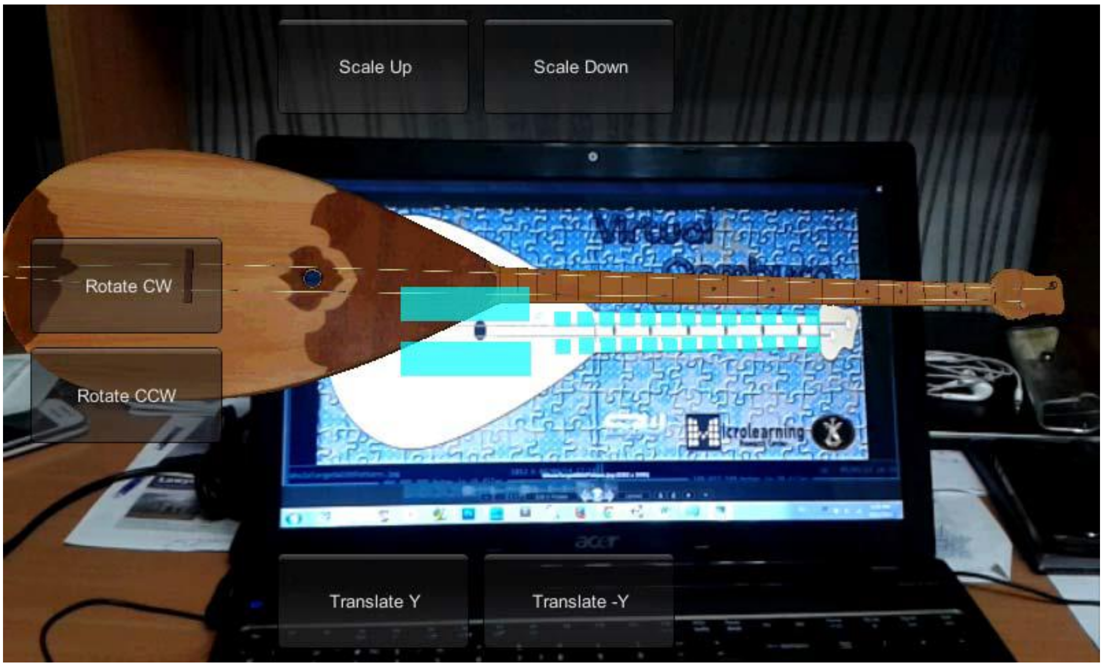

Figure 17. AR generated Dombyra on the tablet screen (Zhaparov & Assanov, 2014)

It provides an easy way to play the traditional Dombyra without the necessity of having the actual instrument, which makes it more effective to have an initial glimpse about Dombyra, especially for the users that want to purchase the actual device.

CHAPTER SIX
DESIGN AND IMPLEMENTATION OF THE VIRTUAL DRUM

6.1. Application of Augmented Reality in This Thesis

Several experiments have been made to achieve the final desired result, in order to integrate AR in this thesis.

First approach was experimenting with existing frameworks and libraries. Vuforia SDK was used together with Zigfu and Unity. Since Vuforia does not have direct support for Kinect, Zigfu was integrated to enable Vuforia to get access to Kinect's camera capture feature. After all, both have been integrated into Unity to achieve the desired result, but since Zigfu and Vuforia's support for Kinect version 2 was not stable enough, final achieved result was not as good as expected.

Next approach was developing a custom solution tailored to this thesis rather than using available frameworks. A minimal solution was developed to experiment with and since it is more flexible to achieve the special needs of this thesis, this method was chosen.

Fundamentally, the system works by placing the 3D drum model in the front of user's captured camera stream based on user's calibrated position value. Initially, pivot positions for the user and drum are defined. After initialization of Kinect, user's new position is captured. The delta amount between user's final position and pivot position is calculated using formula 1 given below:

$$\Delta\vec{P} = \overrightarrow{P_{U1}} - \overrightarrow{P_{U0}} \quad (1),$$

where $\overrightarrow{P_{U0}}: Pivot\ position\ vector\ of\ user$ and $\overrightarrow{P_{U1}}: Captured\ position\ vector\ of\ user$

After the delta amount of user's position has been calculated, the *y* and *z* values of this vector which represent the height and depth values respectively, are applied to their counterpart values of the pivot position vector of the drum respectively, in order to calculate the final position of the drum in the scene. Since the user has to sit exactly in front of Kinect, x value of the delta vector is ignored. Thus, the final position of the drum is calculated using formula 2 as follows:

$$\overrightarrow{P_D} = (\overrightarrow{P_{D0}}.x, \quad \overrightarrow{P_{D0}}.y + \Delta\vec{P}.y, \quad \overrightarrow{P_{D0}}.z + \Delta\vec{P}.z) \qquad (2),$$

where $\overrightarrow{P_{D0}}: Pivot\ position\ vector\ of\ drum$ and
$\overrightarrow{P_D}: Final\ calculated\ position\ vector\ of\ drum$

With the final position of the drum in hand, eventually the drum model was integrated in front of the user in the video stream captured by Kinect which produces a composed video stream of the user with the drum as shown in Figure 18.

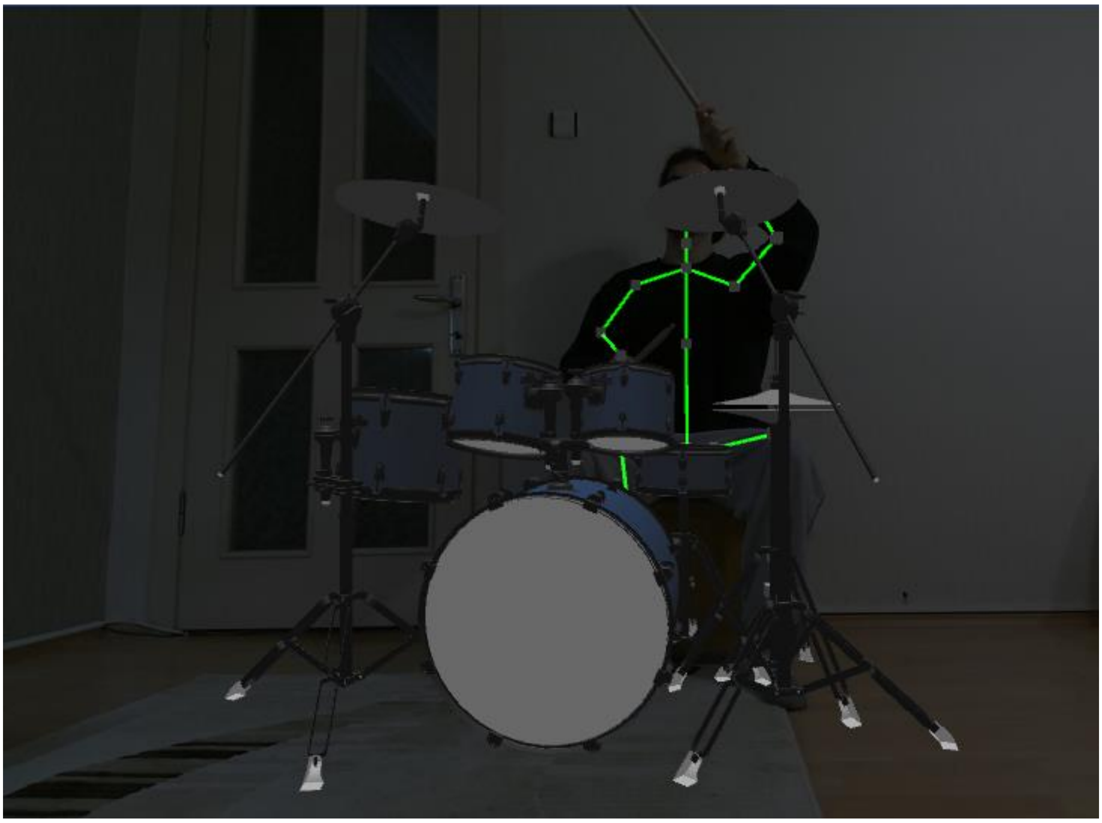

Figure 18. Final composed image of user with AR powered 3D model drum

The integrated 3D drum model also offers some fancy light animations on drum parts to enhance the user experience and give the user more realistic drum playing experience.

6.2. Application of Motion Sensors in This Thesis

During the development of this thesis study, two of the above-mentioned types of motion tracking devices, namely Microsoft Kinect and Leap Motion, were researched and experimented. As a result, Kinect was decided to be used.

Initial idea in the beginning of the thesis study was to produce a virtual guitar by using Leap Motion. Leap Motion has been experimented by using both hands as well as a single hand of a player. Due to the limited capture range of Leap Motion, capturing two hands at the same time did not provide accurate information unless they are both static. Hence, the system was able to be capture only one hand successfully and accurately, however, since one hand is not sufficient to play a virtual guitar or any other string based instrument like bass, ukulele etc., Leap Motion is determined as unsuitable for this thesis' aim of realizing a virtual instrument. In addition to this, the device was not able to capture joint points and finger-tips as accurately as desired.

Since Leap Motion did not meet expectations, Microsoft Kinect was experimented. As mentioned before, Microsoft Kinect offers full-body tracking, which makes playing an instrument more feasible than Leap Motion. Microsoft Kinect is able to track only main joints of the body, so playing virtual guitar is not possible, either. However, playing a virtual drum is feasible, since tracking ankle joints of hands and wrist joints feet are sufficient in order to determine the parts of the drum that are hit by the player. Several experiments were also done with Microsoft Kinect version 1 after this decision has been made. During that time, Microsoft Kinect version 2 has already been released which offers more joints tracking capability, so experiments were continued with Microsoft Kinect version 2.

The general setup used in this study is shown in Figure 19. The system consists of a computer, a display and a Microsoft Kinect version 2 device connected to the computer. Microsoft Kinect is used to detect the position and the movement of the player. The final scene showing the player in front of a virtual drum and sounds generated according to the detected positions of player's hits on the virtual drum are rendered using the developed application.

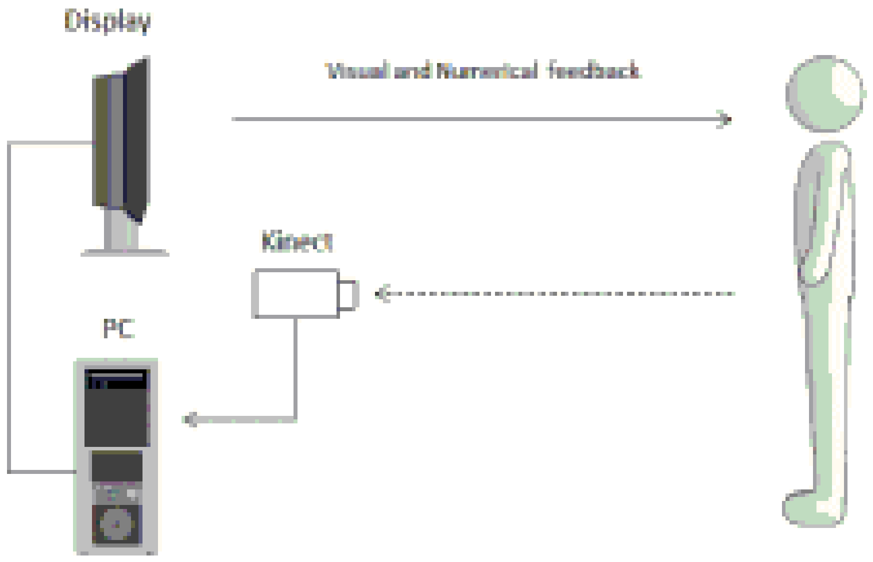

Figure 19. General system design of this thesis work

As a conclusion, desired results have been achieved by using Microsoft Kinect version 2, and the next step in the implementation is to generate a virtual drum set and render the scene with the player sitting in front of the drum set accurately. Figure 20 shows sample computer generated real-time displays using the joints of the player detected by Microsoft Kinect and virtual drum set rendered from different viewpoints. As it can be seen from the figure, the player's position is detected and the virtual drum is placed correctly with respect to the position of the player.

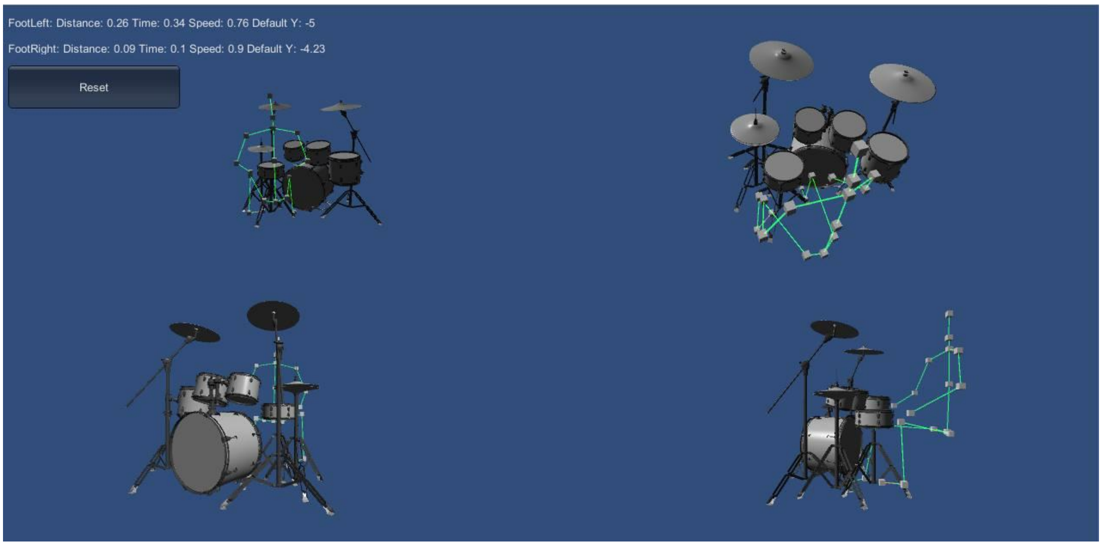

Figure 20. The output on computer display for Kinect experiment

6.3. Virtual Instrumentation in This Thesis

Virtual Instrumentation in this thesis project uses a composer software, Unity, which calculates audio data for the corresponding hit drum part and plays appropriate sound clip based on user data it received from Kinect.

The following procedure is used in order to generate the drum sounds. First, Kinect tracks the user skeleton to grab joint information and sends this data to Unity, to be processed via its official SDK called Kinect SDK. The application developed using Unity SDK processes this data to detect the part of the drum that has been hit along with some other values like the speed of the hit. These values are then used by the same software, to play the appropriate sound according to the received data through Unity. Ultimately, the sound is heard by the player simultaneously when she hits a drum part. Depending on the drum part hit and the speed of hit, the played sound varies as well as. If the speed is not fast enough, it is discarded by the algorithm. This process is repeatedly executed until the system is halted with an external command like a manual shut-down by user. The flow of sound generation is shown in Figure 21.

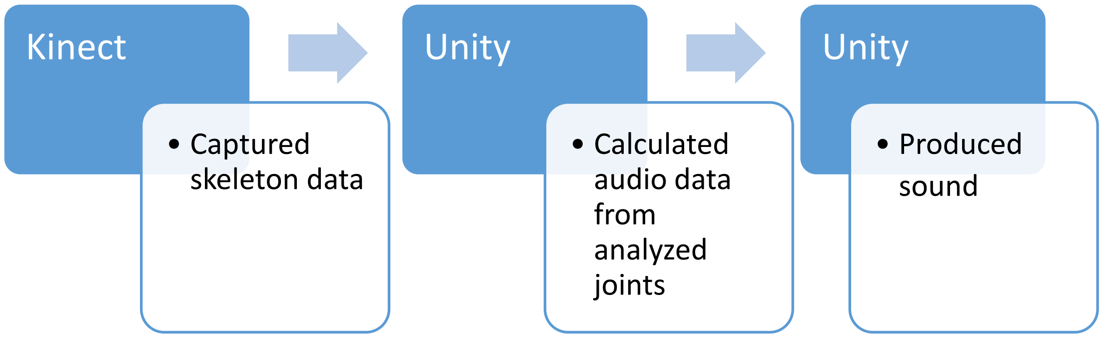

Figure 21. Working schema of sound generation in this thesis

The system runs the following pseudocode (Figure 22) to analyze the joints and play the corresponding sounds.

```
input:  threshold value z,
        threshold y, x_min and x_max values for each drum part,
        threshold_higher_y
        min_limit
previous_y = 0
while user doesn't halt do
        continue ← false
        playSound ← false
        if joint_z >= threshold_z
            category ← close
        else
            category ← far
        for each drum parts dp in selected category do
            if threshold_x_min <= joint_x and joint_x <= threshold_x_max
                continue ← true
                hitDrumPart ← dp
                break
        if continue = true
            if hitDrumPart is located at higher level
                if joint_y > threshold_higher_y
                    playSound ← true
            else
                playSound ← true
            if playSound = true
                if joint_y <= threshold_y of hitDrumPart
                    if ( prev_y – joint_y ) >= min_limit
                        play sound clip for hitDrumPart
                        play lit animation for hitDrumPart
                    else
                        prev_y ← joint_y
                else
                    prev_y ← joint_y
```

Figure 22. Pseudocode for analyzing joints and playing sounds

Virtual drum component, which plays a vital role in this thesis work is composed of different threshold values for different parts of the drum for both wrist and ankle joints. These values also vary for three different axes of 3D coordinate system, x, y, z, respectively. Beyond that, the system is also able to update these threshold values either via manual calibration by the user or automatic calibration if it is enabled. With these threshold values in hand, the system tracks user's joints with a focus on wrist and ankle joints of both hands and feet respectively. Besides that, the system processes wrist and ankle joints separately. For wrist joints, the space in front of the user is divided by multiple parts defined by a z value and multiple x ranges, representing depth and horizontal values respectively. Some parts of the drum like snare drum, floor tom-tom are closer to the user compared to high tom-tom, mid tom-tom, crash cymbals, etc. Hence, the drum parts which are playable by hand are divided into two categories as closer parts and farther parts. Considering this, the system first tracks the z position of the user's wrist joint to detect whether it passes the z threshold value which draws the boundary between the closer parts and the farther parts of the drum. In the meantime, it also checks whether x value of the wrist joint falls in any of the x-value ranges of the drum parts and executes this repeatedly for wrist joints of both hands. If the x value of the wrist joint falls in any one of these ranges, then it tries to detect if there is any hit in vertical axis according to the threshold of that particular drum part.

Exceptionally for some drum parts which are positioned at a higher level like cymbals, the system does one more check for the y value of the joint which indicates its vertical position to see if y value also passes the threshold before checking for vertical hit. After all, if it succeeds to pass all these steps, the second part of the algorithm starts to work which aims to detect any hit occurred in the vertical axis by the user. In order to detect this, it tracks y value of the wrist joint as it represents vertical position. If any wrist joint of both hands exceeds the threshold values defined by each particular drum part, the system checks the previous position of that respective joint to see if there is any difference. If the current re value does not exceed the predefined threshold, the system simply amends the value of the previous position variable of that joint with this value to keep it up-to-date. However, in case it determines any difference in two y values, it then compares the minimum difference limit for that specific drum part with the difference value it calculated to determine if it succeeds to pass that. More to this, the system also does noise filtering in order to eliminate automatic glitches that may be caused by Kinect itself, thus, it ignores if the difference is considerably small. Finally, if it achieves, the system plays the respective sound clip for that drum part that was hit and plays a light animation on the drum part. This process loops until the system is halted.

The threshold values in x and y axis are shown in Figure 23.

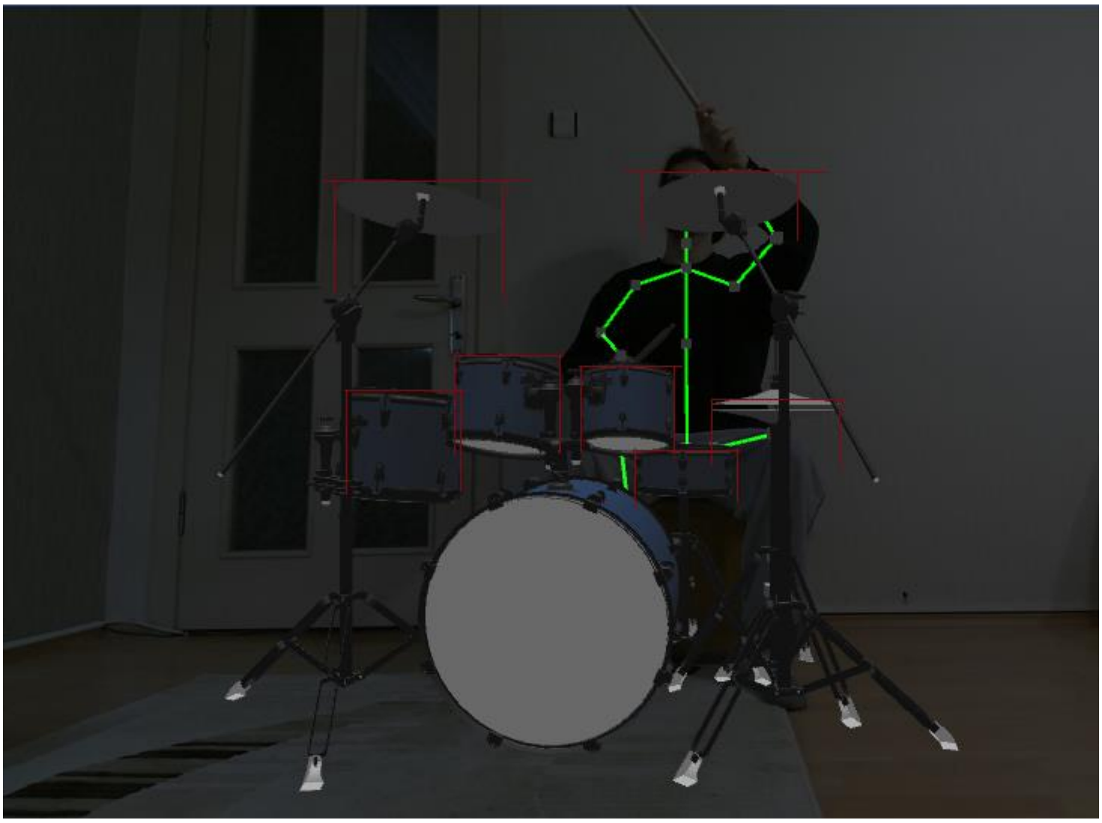

Figure 23. Illustration of Virtual Drum's threshold values in x and y axis

A refined version of the above algorithm is executed for the ankle joints. In contrast to wrist joints, each ankle joint only deals with one drum part, so that left ankle can only play hi-hat cymbals whereas right ankle can play bass drum which simplifies the algorithm considerably. Since ankles are not expected to move in x and z axis remarkably, these values are simply ignored and the algorithm directly proceeds to the second step which checks for difference in the y axis. Similar to the wrist joints, algorithm checks if y value of the ankle joint exceeds y threshold value of the respective drum part and in case of surpasses, it calculates the difference and apply the noise filtering in the same way with wrist joints. Finally, if the difference is large enough, it plays the sound clip of drum part hit and a lit animation. Identical to wrist joints, this process is also executed continuously until the system is halted.

6.4. Comparison to Previous Works

Various studies have been made which focus on virtual instruments, using AR and/or some other technologies. While some of them focus on solely learning purposes, others focus on fun and some other aspects. In terms of instruments virtualized, the majority of works focus on guitar whereas there are only a few studies that virtualize other instruments like piano, drums and other kinds.

Since this thesis study focuses on drums, the comparison is more oriented towards the studies that also focus on drums. In a nutshell, previous studies that aim to create a virtual drum set have some essential missing points as, either they do not operate well as claimed or do not achieve the desired result in practice although they work perfectly well in theory. However, this thesis study improves the work on these points and considers some other aspects, such as the effective functionality of playing a virtual instrumentation, as well as, the look and feel to create best possible user experience.

Firstly, it uses a motion sensor and reduces the required external hardware device amount to minimum, which decreases the cost of hardware to minimum, by only using a Microsoft Kinect device and a personal computer which makes it very affordable for the end-user. Furthermore, this also makes it very portable compared to the previous drum-based works like AR Drum Kit by Waseda University which requires user to have a real drum set and projectors, which makes it ineffective in terms of cost and portability.

Moreover, this thesis study includes a range and threshold based tracking system in order to produce sounds which increases accuracy greatly, compared to similar studies like Spider King by National Central University Jhongli of Taiwan and University of Aizu of Japan. In terms of virtual instrumentation, Spider King seems to achieve a lot of progress and add a lot of value to virtual instrumentation especially in theory, but contrary to theory, in practice, it does not perform as expected. In contrast to it, this thesis work improves the work done by Spider King by adding extra thresholds and ranges in all the three axes, namely horizontal, vertical and depth, which add an extra veracity, and thus the developed drum system addresses user’s hits more accurately. In addition to this, its calibration feature allows users to adjust the threshold values and ranges for their own comfort which increases consistency and usability of the drum.

Finally, integrating an AR powered 3D model of drum into the scene increases the visual quality of the scene and makes the composition look more appealing along with the light animations which also contribute to the user experience.

CHAPTER SEVEN
CONCLUSIONS AND FUTURE WORK

In order to develop a virtual musical instrument, studies and experiments in motion sensors, augmented reality and producing sound were made and issues such as, how the setup and process should be, are resolved in this thesis study.

Firstly, the motion sensor device plays a vital role in determining the virtual instrument to develop. Motion sensor's accuracy and range of capture are the two most important factors. Among the experimented devices, Leap Motion is good and accurate enough for simulating only one hand instruments like maracas or two hand instruments where both hands should be close enough to each other like a keyboard with few keys. PlayStation Move offers a very good range of capture, but requires its motion devices to be held by the player which makes it ineffective for virtual instruments. Consequently, Microsoft's Kinect outshines itself between the three available devices by its capturing range and hands-free capturing, but lack of capturing of fingertips limits it to the instruments that need tracking of joints rather than fingers, like drums.

Space factor also plays an important role when the instrument's volume gets bigger, as Microsoft Kinect will need a wider range of capture to track movements accurately and efficiently. This thesis study supports all the capture values offered by Kinect. It supports 70° field of view for horizontal and 60° for vertical and can capture the player within the range of 0.5mt – 4.5mt where the values are the ones supported by Kinect officially (Duncan, 2013).

Furthermore, the software to be used as a composer between Kinect and computer has an important role as well. It will process the tracking data it gets from Microsoft Kinect and consequently will detect the area hit on the drum and the speed of hit as well. Later, using this data, it plays the appropriate sound clip for the hit drum part.

Regarding drum, a 3D model of classic drum is integrated into the scene using AR. It has 8 parts where 3 of them are cymbals. 7 of its components can be controlled by both hands and 2 of them can be controlled by feet where 1 component, namely hi-hat cymbals can be used by both hands and left foot.

Lastly, software for AR is also required to enable integration of virtual drum in front of the player. It should also be compatible with the composer software. However, in this thesis work, other software for AR is not needed since Unity works as both composer and AR software. Thus, Unity is also used to produce the final rendered output which the player sees.

Comparing with the previous studies as shown in Table 1, this thesis study applies both AR and motion tracking whereas An AR Drum Kit only applies AR and Spider King only utilizes motion tracking. Moreover, this thesis study supports all of the 8 pieces of a classic drum whereas Spider King supports just 5 of them. Furthermore, this study also provides recalculable range and threshold values which enable it to support playing drum in a 3D space, on the other hand Spider King does this only on 2D. Additionally, this support for playing in 3D also enables player to calibrate the system as to their needs and make the system location free. On the other hand, An AR Drum Kit does not offer such flexibility and it needs to be re-positioned as a whole from scratch and Spider King has only predefined fixed threshold values which make the player to sit in a fixed place in front of it, to conform its predefined threshold values.

Table 1. Comparison of features between studies

Feature	An AR Drum Kit	Spider King	This Thesis Study
AR	+	-	+
Motion Tracking	-	+	+
# of Supported Drum parts	8/8	5/8	8/8
Calibration Support	No	No	Yes
Flexibility of movement	No	No	Yes

Regarding future works, a few things are planned to enhance the overall quality of the virtual drum. Firstly, getting the virtual drum to be experimented by more professional drum players is planned to get their feedbacks and enhance the system upon it if needed and adding support for much higher BPM levels for the fast-played music types. Furthermore, customizations to the drum system are also planned to be added, which will enable the player to add, remove or modify drum parts. In addition, the players can alter respective sound filters based on their needs and the genre of music they would like to play. Also players can use different drum tunings to generate wider range of sounds.

Beyond that, producing sound using a MIDI system instead of playing pre-recorded sound clips is also planned to be researched further. Using a MIDI system will generate more realistic sounds. Another addition to the system could be considering the angle of the player's hits in order to produce more accurate sounds using the MIDI software.

Finally, a tutorial system is also planned to be added to the system, which aims to teach the player how to play the drums in an easy way by utilizing songs and gamification.

REFERENCES

Aaron. (2013, May 3). Gesture Control − MYO vs Leap Motion [Online image]. Retrieved from http://outfresh.com/latest-technology/new-tech-devices/gesture-control-myo-vs-leap-motion/

Bravo, S. (2014, March 29). Kinect For Windows v2 Detailed By Microsoft [Online image]. Retrieved from http://acrossthefader.biz/archives/12258

Cakmakci, O., Bérard, F., & Coutaz, J. (2003). An augmented reality based learning assistant for electric bass guitar. In *Proc. of the 10th International Conference on Human-Computer Interaction, Crete, Greece.*

Duncan, G. (2013, December 26). More Kinect SDK v2 from the field... [Blog post]. Retrieved from https://channel9.msdn.com/coding4fun/kinect/More-Kinect-SDK-v2-from-the-field

Earl, S. (2010, July 26). Build A Home Studio for Under $2,000 [Online image]. Retrieved from http://www.homemusicproduction.com/build-your-home-studio-under-2k/

Eisler, C. (2012, January 9). Starting February 1, 2012: Use the Power of Kinect for Windows to Change the World [Blog post]. Retrieved from https://blogs.msdn.microsoft.com/kinectforwindows/2012/01/09/starting-february-1-2012-use-the-power-of-kinect-for-windows-to-change-the-world/

Etherington, D. (2014, March 20). Leap Motion Lays Off 10% Of Its Workforce After Missing On First Year Sales Estimates. Retrieved from https://techcrunch.com/2014/03/20/leap-motion-lays-off-10-of-its-workforce-after-missing-on-first-year-sales-estimates/

Gell, A. (2014, December 31). Microsoft to discontinue original Kinect for Windows in 2015 [Online image]. Retrieved from http://www.hitc.com/en-gb/2014/12/31/microsoft-to-discontinue-original-kinect-for-windows-in-2015/

Gunther, C. (2013, September 5). Qualcomm shows off Vuforia Smart Terrain, makes your living room a game [Online image]. Retrieved from https://www.slashgear.com/qualcomm-shows-off-vuforia-smart-terrain-makes-your-living-room-a-game-05296562/

Hepburn, A. (2011, November 6). National Geographic: Live Augmented Reality [Blog post]. Retrieved from http://www.digitalbuzzblog.com/national-geographic-live-augmented-reality-installation-hungary/

Hsu, M. H., Kumara, W. G. C. W., Shih, T. K., & Cheng, Z. (2013, November). Spider king: Virtual musical instruments based on microsoft kinect. In

Awareness Science and Technology and Ubi-Media Computing (iCAST-UMEDIA), 2013 International Joint Conference on (pp. 707-713). IEEE.

Koller, J. (2010, June 15). This Changes Everything: PlayStation Move Available September 19, 2010 [Online image]. Retrieved from https://blog.us.playstation.com/2010/06/15/this-changes-everything-playstation-move-available-september-19-2010/

Motokawa, Y., & Saito, H. (2006, October). Support system for guitar playing using augmented reality display. In *Mixed and Augmented Reality, 2006. ISMAR 2006. IEEE/ACM International Symposium on* (pp. 243-244). IEEE.

Nieler, P. (2015, November 9). Motion Capture for "Room Scale" Virtual Reality [Online image]. Retrieved from http://www.neurosc.com/2015/11/09/motion-capture-for-room-scale-virtual-reality/

Peckham, M. (2011, February 4). In L.A. Noire Asking Questions Trumps Firing Bullets. Retrieved from http://www.pcworld.com/article/218704/in_la_noire_asking_questions_trumps_firing_bullets.html

Pham, A. (2009, June 1). E3: Microsoft shows off gesture control technology for Xbox 360. Retrieved from http://latimesblogs.latimes.com/technology/2009/06/microsofte3.html

PTC Completes Acquisition of Vuforia. (2015, November 3). Retrieved from http://www.ptc.com/news/2015/ptc-completes-acquisition-of-vuforia

Route66Training. (2011, November 16). *Follow Me TM augmented reality navigation, test drive (ROUTE 66)* [Video file]. Retrieved from https://www.youtube.com/watch?v=VXZ6Hj2IOPE

Sethi, R. (2015, July 2). Apple Updates GarageBand to 10.1 [Online image]. Retrieved from https://ask.audio/articles/apple-updates-garageband-to-101

Stewart, D. (2013, January 29). Technical GRAMMY Award: Ikutaro Kakehashi and Dave Smith. Retrieved from http://www.grammy.com/news/technical-grammy-award-ikutaro-kakehashi-and-dave-smith

Software synthesizer. (n.d.). In *Wikipedia*. Retrieved June 1, 2017, from https://en.wikipedia.org/wiki/Software_synthesizer

Sutherland, I. E. (1968, December). A head mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I* (pp. 757-764). ACM.

Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Sensors*, *13*(5), 6380-

6393.

Wortham, J. (2010, November 21). With Kinect Controller, Hackers Take Liberties. Retrieved from http://www.nytimes.com/2010/11/22/technology/22hack.html

Yamabe, T., Asuma, H., Kiyono, S., & Nakajima, T. (2011, August). Feedback design in augmented musical instruments: A Case Study with an ar drum kit. In *Embedded and Real-Time Computing Systems and Applications (RTCSA), 2011 IEEE 17th International Conference on* (Vol. 2, pp. 126-129). IEEE.

Yin-Poole, W. (2012, November 16). PlayStation 3 hits 70 million units shipped worldwide mark six years after launch. Retrieved from http://www.eurogamer.net/articles/2012-11-16-playstation-3-hits-70-million-units-shipped-worldwide-mark-six-years-after-launch

Zhaparov, M. K., & Assanov, U. (2014, October). Augmented reality based on Kazakh instrument "Dombyra". In *Application of Information and Communication Technologies (AICT), 2014 IEEE 8th International Conference on* (pp. 1-4). IEEE.