

INTERACTIVE MULTIMEDIA SYSTEM BY USING HAND-GESTURE
MOTION SENSOR

AMIRRUL ADHA BIN SULAIMAN

This report is submitted in partial fulfillment of the requirements for the award of
Bachelor of Electronic Engineering (Computer Engineering) With Honours

Faculty of Electronic and Computer Engineering
Universiti Teknikal Malaysia Melaka

June 2015



UNIVERSITI TEKNIKAL MALAYSIA MELAKA
FAKULTI KEJURUTERAAN ELEKTRONIK DAN KEJURUTERAAN
KOMPUTER

BORANG PENGESAHAN STATUS LAPORAN

PROJEK SARJANA MUDA II

Tajuk Projek : INTERACTIVE MULTIMEDIA SYSTEM BY USING HAND-GESTURE MOTION SENSOR TECHNOLOGY

Sesi Pengajian :

1	4	/	1	5
---	---	---	---	---

Saya AMIRRUL ADHA BIN SULAIMAN mengaku membenarkan Laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (\checkmark) :

SULIT*

*(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

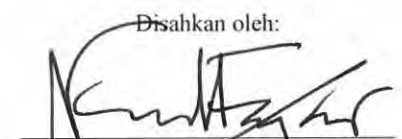
TERHAD**

** (Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

TIDAK TERHAD


 (TANDATANGAN PENULIS)

Tarikh: 12/6/15

Disahkan oleh:

 (COP DAN TANDATANGAN PENYELIA)

Dr. Nurufajar Bin Abd. Manap
 Pensyarah Kanan
 Fakulti Kejuruteraan Elektronik Dan Kejuruteraan Komputer
 Universiti Teknikal Malaysia Melaka (UTeM)
 Hong Tuah Jaya
 76100 Durian Tunggal, Melaka
 Tarikh: 12/06/15

Hereby the author declares that this thesis entitled “Interactive Multimedia System by Using Hand-Gesture Motion Sensor” is the result of my own research except as cited in the references.

Signature

:



Author

:


AMIRRULADHA BIN SULAIMAN

Date

:

12 JUNE 2015

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of the Bachelor of Electronic Engineering (Computer Engineering) with Honors

Signature : 
Supervisor's Name : DR. NURULFAJAR BIN ABD MANAP
Date : 12 JUNE 2015

Dedicated to my beloved family especially for both of my parents, Mr. Sulaiman bin Arshad and Mrs. Adilah binti Md Ramli, my fellow friends and to all my teachers.

ACKNOWLEDGEMENT

First of all, I would like to this opportunity to express my greatest gratitude to project supervisor, Dr. Nurulfajar bin Abd Manap for his guidance, encouragement and endurance during the whole process of this project. It is indeed my pleasure to have him to support me with limitless advices and enthusiastic spirit towards my project. I also would like to offer my special gratitude to my beloved family, exclusively to both of my parents, Sulaiman bin Arshad and Adilah binti Md. Ramli for their continuous support throughout the years of my study in University of Technical Malaysia Malacca (UTeM). They are the main reason for the person that I am becoming today, without their support, this project would not be completed successfully. Not to forget, I would like to appreciate all of my friends especially to my course mates and Sobri for their technical advices. Lastly, to all the people that involve with this project either directly or indirectly, once again, thank you very much.

ABSTRACT

Most of the computer-based industry especially in multimedia industry are pursuing in the search of natural human-computer interaction. There are many concepts and technologies have been developed recently in order to achieve the ultimate goal of human-computer interaction which is to produce a comfortable user interface. This project is about designing an interactive multimedia system by implementing hand-gesture motion sensor technology. The concept of this project is applicable to any multimedia system especially for in-vehicle infotainment system. The proposed system can be integrated with the entire gestural language based on a single hand. Plus, there will be no audio commands required in the system. This multimedia system is integrated with the Leap Motion controller in order to recognize the specified hand-gestures. Only enabled hand-gestures will interact with some specified functions to control the multimedia player application. For example, use a swipe gesture to stop the media file, use a circle gesture to play the media file and so on. Therefore, this system will deliver a new interactive multimedia experience that allows users to control the multimedia player application with hand gestures

ABSTRAK

Kebanyakan daripada industri komputer terutama industri multimedia sedang berusaha dalam pencarian interaksi semula jadi antara manusia dan komputer. Terdapat pelbagai konsep dan teknologi yang telah dibangunkan kini bagi mencapai matlamat utama interaksi antara manusia dan komputer iaitu menghasilkan antara muka yang selesa. Projek ini adalah tentang merekabentuk sistem multimedia interaktif yang menggunakan teknologi pengesan pergerakan isyarat tangan. Konsep projek ini mampu digunakan pada setiap sistem multimedia terutama sistem multimedia dalam kenderaan. Sistem yang dicadangkan ini mampu berintegrasi dengan setiap isyarat sebelah tangan. Tambahan pula, tiada arahan audio diperlukan dalam sistem ini. Sistem multimedia ini diintegrasikan dengan pengawal Leap Motion bagi mengenalpasti isyarat tangan yang ditetapkan. Hanya isyarat tangan yang dibenarkan sahaja akan berinteraksi dengan fungsi yang ditetapkan bagi mengawal aplikasi pemain multimedia. Sebagai contoh, isyarat sapu untuk menghentikan fail media, isyarat bulatan untuk memainkan fail media, dan sebagainya. Oleh itu, sistem ini akan menghasilkan pengalaman multimedia interaktif baharu yang membenarkan pengguna mengawal aplikasi pemain multimedia melalui isyarat tangan.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	PROJECT TITLE	i
	VERIFICATION FORM	ii
	DECLARATION	iii
	DEDICATION	v
	ACKNOWLEDGEMENT	vi
	ABSTRACT	vii
	ABSTRAK	viii
	TABLE OF CONTENTS	ix
	LIST OF TABLES	xii
	LIST OF FIGURES	xiii
I	INTRODUCTION	1
	1.1 Project Background	1
	1.2 Problem Statement	2
	1.3 Objectives	3
	1.4 Scope of Project	3
	1.5 Thesis Structure	5
II	LITERATURE REVIEW	6

2.1	Introduction	6
2.2	History of Human Computer Interaction	7
2.3	Multimedia	8
2.4	Issues on Multimedia Interface	10
2.5	Gesture	11
2.6	Gesture Recognition	13
2.7	Comparison between Gesture and Haptic Inputs	14
2.8	Related Research on Vision-Based Gesture Recognition	15
2.9	Related Research on Gesture Recognition Devices	19
2.10	Summary	22
III	METHODOLOGY	23
3.1	Introduction	23
3.2	Overview of Project Design	24
3.3	Hardware Implementation	25
3.3.1	Leap Motion Controller	26
3.4	Software Implementation	28
3.4.1	Java	29
3.5	Implementation of Hand-Gesture Recognition using Leap Motion Controller	30
3.5.1	Motion Tracking Data	33
3.6	Multimedia Player Application Development	33
3.6.1	Multimedia Player Application using Java	34
3.6.2	JavaFX	35
3.7	Proposed System	36
3.7.1	Proposed Algorithm	37
3.8	Summary	39

IV	RESULTS AND DISCUSSIONS	40
4.1	Introduction	40
4.2	Results and Discussions	41
4.2.1	Implementation of Hand-Gesture Recognition using Leap Motion Controller	41
4.2.2	Multimedia Player Application Development	49
4.2.3	Interactive Multimedia System by Using Hand-Gesture Motion Sensor Technology	52
4.3	Summary	54
V	CONCLUSION	55
5.1	Conclusion	55
5.2	Future Recommendation	56
	REFERENCES	58
	APPENDIX A	60

LIST OF TABLES

NO	TITLE	PAGE
Table 2.1	Issues on current multimedia interfaces	10
Table 2.2	Summary of related research on vision-based gesture recognition	16
Table 2.3	Comparison of features of somatosensory controller	20
Table 2.4	Summary of related research on gesture recognition device	21
Table 4.1	The relationship of performed gestures and the functionality of the multimedia player application	52

LIST OF FIGURES

NO	TITLE	PAGE
Figure 2.1	Timeline approximation of research are was performed on some major technologies involving the human-computer interaction.	8
Figure 2.2	A single hand gesture conveys different interpretations. Gesture (a) may be interpreted as pointing, size holding etc. However, gesture (b) might be interpreted differently as it is slightly different from gesture (a). Same with gesture (c) and (d).	12
Figure 2.3	Comparison of secondary control input performance of drivers using haptic and gesture inputs.	14
Figure 2.4	The ability of IR LEDs to recognize hand gesture in various light condition.	18
Figure 3.1	Overview of project design which consists of several modules.	24
Figure 3.2	Hardware Interconnection between Leap Motion and PC.	25
Figure 3.3	Leap Motion controller.	26
Figure 3.4	The overview of Leap Motion fundamental component which consists of two monochromatic IR cameras and three IR LEDs.	26
Figure 3.5	The overview illustration of Leap Motion interaction area.	27
Figure 3.6	The Structure of Leap-enabled application.	28

Figure 3.7	Java by Oracle logo.	30
Figure 3.8	Eclipse IDE logo.	30
Figure 3.9	Process of implementation of hand gesture recognition using Leap Motion controller.	31
Figure 3.10	The overview illustration of Leap Motion coordinate system.	32
Figure 3.11	The process of developing the multimedia player application using JavaFX.	34
Figure 3.12	The hierarchy of Media classes in JavaFX.	35
Figure 3.13	The integration of both Multimedia Player API and Leap Motion Tracking API is what makes the interactive multimedia system by using hand-gesture motion sensor technology	36
Figure 3.14	The flowchart of proposed algorithm.	38
Figure 4.1	Leap Visualizer shows the tracking images and model of one tool and one hand.	42
Figure 4.2	Java console shows the detection info of hands, fingers, tools and gestures.	42
Figure 4.3	Leap Visualizer shows the tracking images and model of two hands.	43
Figure 4.4	Java console displays the information of both right hand and left hand including its palm position.	43
Figure 4.5	Java console shows the information of tool id to distinguish different tools as long as one does not leave the interaction field.	44
Figure 4.6	Leap Visualizer shows the tracking images and model of the right hand.	45
Figure 4.7	Java console shows the information of each finger and its length and width. (Right-hand).	45
Figure 4.8	Leap Visualizer shows the tracking images and model of the left hand.	46
Figure 4.9	Leap Motion is able to identify each finger and its length and width. (Left-hand).	46

Figure 4.10	The latency values are constant due to no hand tracking performance.	48
Figure 4.11	The latency values spiked due instant hand tracking performance.	48
Figure 4.12	Basic GUI of the developed multimedia player using JavaFX.	50
Figure 4.13	The directory of media file is linked with multimedia player	51
Figure 4.14	The GUI changes as the media file is dropped into the multimedia player application. This example shows that the multimedia player is playing a video file.	51
Figure 4.15	Java console displays pause information when key-tap gesture is recognized to pause the media file.	53
Figure 4.16	Java console displays play information when circle gesture is recognized to play the media file.	53
Figure 4.17	Java console displays stop information when swipe gesture is recognized to stop the media file	53

CHAPTER I

INTRODUCTION

1.1 Project Background

Human-computer interaction is well-defined as a discipline that dedicated on study, design, manufacture and application of human-centric interactive computer systems. Since the introduction of computer, human-computer interaction has been varied in many different ways such as tactile interactions, haptic feedback, and speech recognitions etc. Therefore, user interface plays a main role to ensure precise human-computer interaction. However, even though there are many ways of interaction with the computer, the satisfaction of user is perceptual.

The satisfaction of human-computer interaction may be different based on a few factors such as natural interaction between computers. As mentioned earlier, most of the computer-based industry especially in multimedia industry are on a quest on searching the most natural human-computer interaction. The fundamental goals for this search are to create an easy user interface and to make the user feels comfortable while interacting with it.

There were many concepts and systems of multimedia have been developed which are more complex to design so that the user would feel more natural while interacting with it. However, which of these developments of human-computer interactions is better in terms of easiness and comfort? Recently, the industry came out with the touch screen system. Since it was introduced, it promises significant ease, comfortable and high interactivity compared to conventional button interaction system. However, it also produces some limitations in terms of physical conditions. For example, the user needs to have a physical contact in order to interact with the system. This kind of limitation would be significant if the user's hand is wet. As most of us had the experience of interacting a multimedia application via touch screen systems, the wet condition of our hands is somehow makes it difficult to interact properly with the multimedia application.

Therefore, this project will implement hand gesture as its interface in order to interact with the multimedia system. Compared to touch screen system, hand-gesture based multimedia interface promises significant benefits in terms of easiness, comfortable, and interactivity. This hand-gesture based multimedia would eliminates the limitations in terms of physical conditions since hand-gesture recognition is produced without any physical contact.

1.2 Problem Statement

There are various type of multimedia interface available on the market today such as buttons, touch screens and voice recognition. However, with the rapid pace of technology, most researchers are trying their very best to create a new multimedia interface that is more comfortable, more intuitive and more intelligent that could make our daily life become easier than what we had right now. Some of us might imagine what if we had a computer that could understand us like a friend. A computer that could aid us with our daily life routine with minimal physical interaction.

Over the last few decades, our interaction with the multimedia system has changed since the invention of the button. Many of us had experienced some of the multimedia interfaces that were mentioned earlier. However, there are several issues of current multimedia interfaces regarding interaction and satisfaction have arisen. The conventional multimedia interfaces obviously involves eye attention of the user. For example, buttons and touch screen display. Whenever the user intends to interact with this particular type of multimedia interfaces, the user needs to give out his eye's attention in order to operate with multimedia system properly.

Therefore, in order to overcome this particular issue, there are some approach has been taken by the multimedia-based industry such as speech-recognition system. However, there are several issues from the speech recognition system in order to interact with the multimedia system. For example, language boundaries makes it difficult for the multimedia system to recognize the speech or voice command given by the user. Hence, a new multimedia interface is needed in order to overcome these kind of issues.

To overcome these issues, gesture-based multimedia interface has been proposed. Gesture-based multimedia interface overcomes the language barrier since most of the people use gestures to deliver an information in a way that most of the people could understand. Imagine if we could interact with multimedia system such as Tony Stark in the Iron Man movie only via movement of hands in the air. Gesture-based multimedia interface promises significant benefits in terms of easiness, minimum physical interaction and intuitiveness.

1.3 Objectives

The aim of this project is to produce an interactive multimedia system by using hand-gesture motion sensor technology. Therefore, in order to perform this project, some objectives are recognized as followings:

1. To implement hand-gesture recognition by using Leap Motion controller.
2. To develop a simple multimedia player application.
3. To design an interactive multimedia system by using hand-gesture motion sensor technology.

1.4 Scope of Project

In order to conduct this project, some specifications have to be determined. This project focuses on producing an interactive multimedia systems design that implements hand-gesture motion sensor technology. This system will be designed to respond only to a single hand-gestures. Plus, the will no audio command or any physical interactions will be used.

The hardware to be used for this project is the Leap Motion controller. With this hardware system, the user could interact with the multimedia player within a specific range from 3 centimeters up to 60 centimeters. In order to develop the multimedia player application that uses Leap Motion as its input, Eclipse Integrated Development Environment (IDE) will be used as the tool to develop Leap-enabled application and Java programming language will be implemented as its programming language. Several gestures will be specified in order to be recognized by the Leap Motion controller and to interact with the system such as:

- Circle : Play the media file.
- Swipe : Stop the media file.
- Tap : Pause the media file.

1.5 Thesis Structure

This project thesis delivers the concepts, the methodology, the results and analysis of the designed interactive multimedia system by using hand-gesture motion sensor technology. This thesis consists of five chapters; Introduction, Literature Review, Methodology, Results and Discussion, and Conclusion and Recommendations respectively.

First chapter covers the overview of the project. This chapter aims to give the reader the overall picture about what the project is all about. The background and the problem statement of this project is briefly explained. Then, followed by the objectives and scope of this project.

In Chapter Two, it is all about the literature review that relates to this project. In literature review, all the related works and researches regarding of the project are explained and summarized. Starting with the history of Human-Computer interaction, the definition of multimedia, issues of current multimedia interfaces and so on. The main concept of gesture and gesture recognition are also explained briefly in this particular chapter.

In Chapter Three, it is all about the methodology used to perform this project. The aim of this chapter is to explain all the approaches and procedures in order to complete this project. This chapter includes the hardware and software implementations along with the procedures which consists of three modules.

In the next chapter which is the fourth chapter is all about the showcase of all the results obtained. Based on three modules which are explained earlier in the previous chapter, all the results are discussed this chapter.

Finally, the fifth chapter concludes everything and suggests on future work related to this project.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

This chapter encompasses the study of existing research related to the project. Any reliable information will be analyzed to enhance the understanding of the main concept and terminology which will be implemented all the way through this project. Started with history of human-computer interaction. Follows by the definition of multimedia and the issues of current multimedia interfaces. Later, review on gesture and gesture recognition which later be compared with haptic inputs in terms performance. Finally, related research on vision-based gesture recognition and gesture recognition device are briefly be reviewed respectively.

2.2 History of Human-Computer Interaction

Human Computer Interaction started in 1943 since the introduction of Electronic Numerical Integrator and Computer which is also known as ENIAC. ENIAC was the first general purpose computer built to decipher millions of Nazi's war code. It was controlled by a web of large electrical cables. Since then, we have been experiencing the evolution of human-computer interaction [1].

In 1963, the world was amazed by the development of Sketchpad by Ivan Sutherland. Sketchpad was a major breakthrough of Human-Computer Interaction. It represents the world's first direct manipulation interface. Sketchpad manipulated objects such as grabbing, moving objects on the screen using a light pen [1].

In 1965, the world's first computer mouse was developed at Stanford Research Laboratory to replace the light pen [1]. Even a small innovation like the computer mouse, it does brings us a great contribution in human-computer interaction. It has been decades since the introduction of the computer mouse, the usage of the computer mouse is somehow very much fundamental in our current human-computer interaction.

As illustrated by Myers, Figure 2.1 shows the time lines approximation of where and when work was performed on some major technologies involving the human-computer interaction. From the Figure 2.1 itself, we can foresee the development of human-computer interaction in the future. As Licklider outlined the Man-Computer Symbiosis in 1960, he quoted that, "The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today" [2].

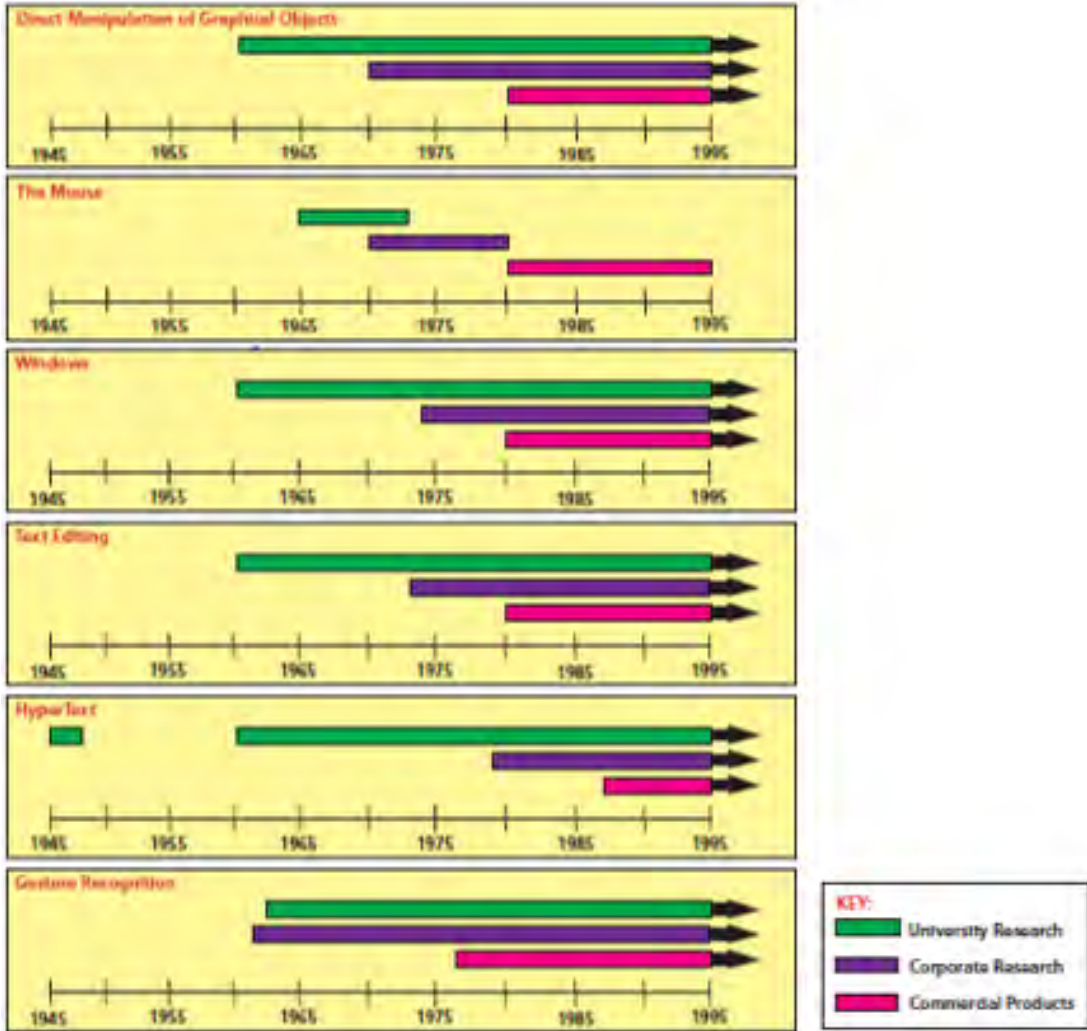


Figure 2.1: Timeline approximation of research area was performed on some major technologies involving the human-computer interaction [1].

2.3 Multimedia

Multimedia comes from the combination of two words; “multiple” and “media”. The word “multiple” refers to the combination of two or more objects and the word “media” is the plural form of the word “medium” which is defined as the dominant material through which sensory impressions are conveyed. Media elements consist of

texts, graphics, audio, images, and videos. Therefore, multimedia is referred to the combination of two or more media elements that conveys any sensory impressions or messages.

Grimes et al (1991) discussed that multimedia can be defined with various meanings. First, he defined that multimedia is an information programming. He stated that multimedia programming happens in three stages. The first stage is to create the multimedia system delivery. The next stage is to create the content of the multimedia itself. Finally, the third stage is to create the links and paths through the content which is also known as the storyline. Content has two distinct kinds which is the information and access paths respectively. As one would experience in video editing, it is difficult to link multiple multimedia contents with suitable access paths in order for user to understand the messages that were intended to be delivered from the video. However, the main point is to possess the tools that allow people to traverse through the information as they see fit [3].

Plus, he defined that multimedia is a multisensory interaction that includes the any combination of human senses such as hearing, seeing, touching, tasting and smelling. For example, watching a movie can be referred as a multimedia interaction as we are able to see the graphics and hear the sounds effects simultaneously in order to receive the message conveyed by the film maker. However, since multimedia is a multisensory interaction, new kinds of interactivity that produces multiple combination of human senses have been introduced lately [3].

Next, he defined that multimedia is the ability to juxtapose media elements. Most people would agree that multimedia is the combination of media elements so that it will produce different kinds of impression rather than dull and gloomy. As an example, a romantic video would more effective if sound effects that suits to the mood are included. However, the inclusion of different media element into another such as audio and video is time-based. Time are categorized into two main parts; real-time and sequential. By referring to the example given, the impression of the video would be different if the sound effects are included in sequence rather than in real-time [3].