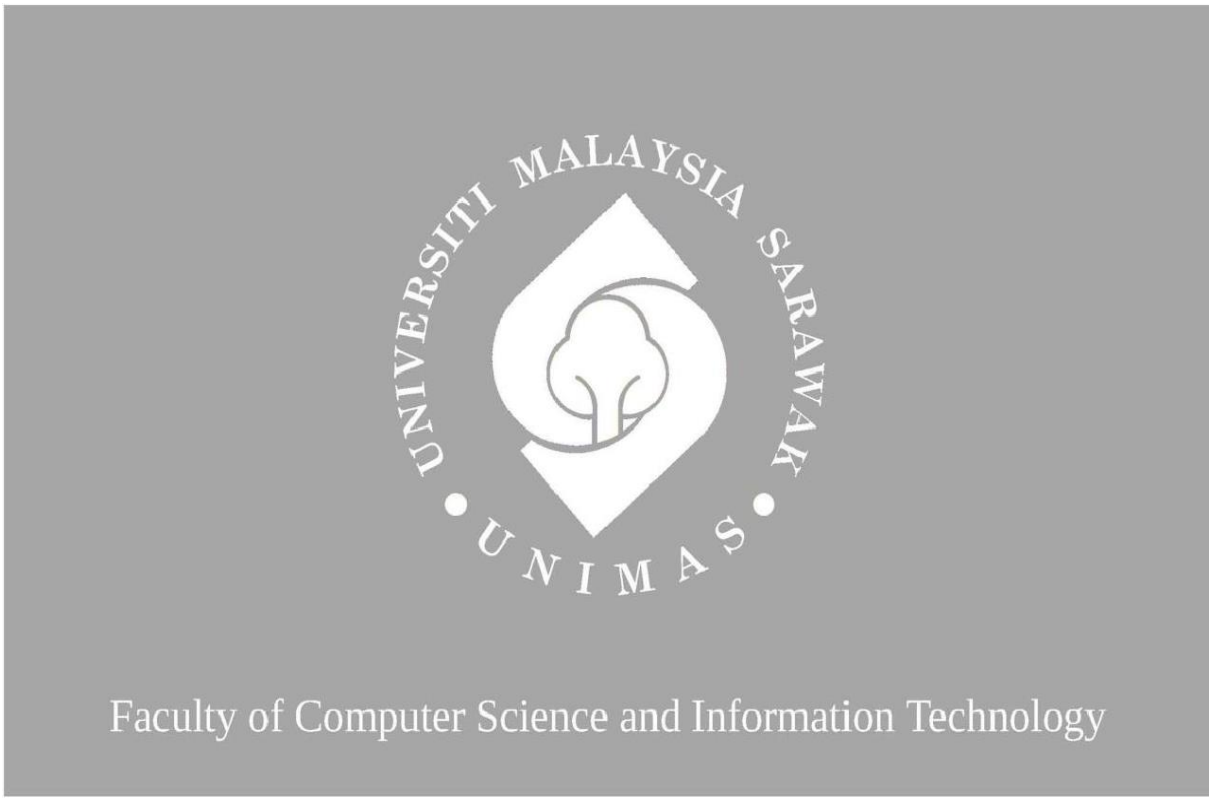

Hand Gestures Interaction in Augmented Reality (AR) Learning Application

IRENE LAW XIN LIN

Bachelor of Computer Science with Honours

(Software Engineering)

2023

HAND GESTURE INTERACTION IN AUGMENTED REALITY (AR) LEARNING APPLICATION

IRENE LAW XIN LIN

This project is submitted in partial fulfillment of the
requirement for the degree of
Bachelor of Computer Science with Honors
(Software Engineering)

Faculty of Computer Science and Information Technology
UNIVERSITY MALAYSIA SARAWAK
2023

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE HAND GESTURE INTERACTION IN AUGMENTED REALITY (AR) LEARNING APPLICATION

ACADEMIC SESSION: 2019/2020

IRENE LAW XIN LIN
(CAPITAL LETTERS)

hereby agree that this Thesis* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [or for the purpose of interlibrary loan between HLI]
5. ** Please tick (√)

☐	CONFIDENTIAL	(Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
☐	RESTRICTED	(Contains restricted information as dictated by the body or organization where the research was conducted)
☑	UNRESTRICTED	

(AUTHOR'S SIGNATURE)

Validated by

(SUPERVISOR'S SIGNATURE)

Permanent Address

BATU 2, JALAN KELUPU
96500, BINTANGOR,
SARAWAK.

Date: 16/7/2023

Date: 17/7/2023

Note * Thesis refers to PhD, Master, and Bachelor Degree
** For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

DECLARATION OF ORIGINALITY

I hereby declare that this project altogether with all of its contents is none other than that of my work, with consideration of the exception of reference or acknowledgement is not made explicitly in the text, nor has any part had written for me by another person.

…….. 28th Jun 2023

(IRENE LAW XIN LIN)

Faculty of Computer Science and Information

Technology, University Malaysia Sarawak

ACKNOWLEDGMENT

First, I would like to express my heartiest gratitude to all people involved directly or indirectly in contributing to my Final Year Project. I want to send my greatest appreciation to my supervisor for Final Year Project, Prof Wang Yin Chai for his assistance, guidance, suggestions and supervision in my Final Year Project. It helps me complete my Final Year Project on time.

Next, I also want to thank University Malaysia Sarawak (UNIMAS) for providing a sturdy platform to gain knowledge and stand a chance to learn a lot of things. The following appreciation goes to my faculty, the Faculty of Computer Science and Information Technology for giving me a chance to experience challenges in my Final Year Project which in turn prepared me to go work in the future.

Deepest thanks and appreciation to my parents, family, and others for their cooperation, encouragement, constructive suggestion and full support for the report completion, from the beginning till the end. Also thank all my friends and everyone that have contributed by supporting my work and helping me during the final year project progress till it is fully completed.

Table of Content

List of Figure

List of Table

ABSTRACT

Augmented Reality (AR) is a technology that has become popular in this recent years. It can mix the virtual objects generated by computers into the real world. Hand gesture recognition is the gesture from human body language which can be recognized by the computer. AR and hand gesture recognition had faced numerous problems such as problems in target recognition, lack of interaction of Augmented Reality (AR), and high cost of hardware used. The project Hand Gesture Interaction in Augmented Reality (AR) Learning Application will be developed to solve this problem. The objective that will be used to solve the problem is to identify hand gesture techniques for learning applications, to design an interactive AR learning application using hand gestures, and to develop a prototype of a learning application with AR using hand gesture techniques. The methodology to develop the application of hand gesture interaction in AR learning is agile methodology. The process of the AR hand gesture by using ManoMotion is stated in the agile methodology. The outcome of the application is the user can use this application to learn astronomy courses such as Sun and Moon through AR. The user also can use the hand gesture to modify the size of the AR to make their learning more interactive. The testing results of the gesture detected to trigger the action reveal that the highest percentages are observed for Right Grab is 36% and Right Pick is 21% while Left Grab attained a rate of 29%, and Left Pick exhibited the lowest percentage at 14%. The limitations of the application include a limited set of hand gestures, the accuracy of detecting hand gestures being affected by lighting conditions and background noise, non-standard hand gesture performance, and hardware limitations. Future work includes exploring new hand gestures and integrating alternative SDKs, improving the ability to recognize various lighting conditions, enhancing hand detection accuracy, and leveraging smartphones with better hardware conditions.

ABSTRAK

Augmented Reality (AR) adalah teknologi yang menjadi popular untuk tahun kebelakangan ini. Ia boleh mencampurkan objek maya yang dihasilkan oleh komputer ke dalam dunia nyata. Pengecaman isyarat tangan ialah isyarat daripada bahasa badan manusia yang boleh dikenali oleh komputer. AR dan pengecaman isyarat tangan telah menghadapi pelbagai masalah seperti masalah dalam pengecaman sasaran, kekurangan Interaktif Augmented Reality (AR) dan kos tinggi perkakasan yang digunakan. Projek Interaksi Isyarat Tangan dalam Aplikasi Pembelajaran Augmented Reality (AR) akan dibangunkan untuk menyelesaikan masalah ini. Objektif yang akan digunakan untuk menyelesaikan masalah adalah untuk mengenal pasti teknik isyarat tangan untuk aplikasi pembelajaran, mereka bentuk aplikasi pembelajaran AR interaktif menggunakan isyarat tangan dan membangunkan prototaip aplikasi pembelajaran dengan AR menggunakan teknik isyarat tangan. Metodologi untuk membangunkan aplikasi interaksi isyarat tangan dalam pembelajaran AR ialah metodologi tangkas. Proses gerak isyarat tangan AR dengan menggunakan ManoMotion dinyatakan dalam metodologi tangkas. Hasil daripada aplikasi tersebut ialah pengguna dapat menggunakan aplikasi ini untuk mempelajari kursus astronomi seperti Matahari dan Bulan melalui AR. Pengguna juga boleh menggunakan isyarat tangan untuk mengubah suai saiz AR untuk menjadikan pembelajaran mereka menjadi lebih interaktif. Hasil ujian pengesanan gerakan untuk memicu tindakan menunjukkan bahawa persentase tertinggi adalah pada Gerakan Grab Kanan sebanyak 36% dan Gerakan Pick Kanan sebanyak 21%, sementara Gerakan Grab Kiri mencapai kadar 29%, dan Gerakan Pick Kiri menunjukkan peratusan terendah pada 14%. Kelemahan aplikasi termasuk set terhadap gerakan tangan, ketepatan pengesanan gerakan tangan yang terjejas oleh keadaan pencahayaan dan bunyi latar belakang, prestasi gerakan tangan yang tidak standard, dan had peranti keras. Kerja masa depan termasuk menjelajah gerakan tangan baru dan mengintegrasikan SDK alternatif, meningkatkan kemampuan mengenali pelbagai keadaan pencahayaan, meningkatkan ketepatan pengesanan gerakan tangan, dan memanfaatkan telefon pintar dengan keadaan peranti keras yang lebih baik.

Chapter 1: Introduction

1.1 Introduction

Augmented Reality (AR) is the technology that can mix the virtual objects generated by computers into the real world (Le & Kim, 2017). It will provide the users with visual elements, sounds and sensory information through devices such as mobile devices (Alexander, 2022). AR applications with the AR development kit can develop the application where ARkit is for IOS devices and ARcore is for Android devices. AR can be divided into three types which are marker-based, markerless-based and location-based (What are the different types of augmented reality, 2021). Marker-based AR includes a QR code or a photo to trigger the AR experience. For markerless-based AR, only the real environment is required to be scanned to trigger the AR experience without any marker (Marker based vs markerless ar, 2021). Location-based AR also known as geo-based AR usually uses GPS technology to get the user's current location and display the 3D object in the real world through mobile devices (Chaudhari, 2019). Nowadays, there are many applications of AR in real life. AR has been used in various fields such as education, manufacturing, medicine and economics (TECH, 2021).

Hand gesture recognition can be thought of as a way for the computer to understand human body language (Huo et al., 2021). It can be divided into two types which are static gestures and dynamic gestures. The static gesture recognizes the hand in a stable shape while the dynamic gesture recognizes the moving hand like handshaking (Oudah et al., 2020). Hand gestures are a more familiar method for a human to interact with mobile devices instead of using hardware such as keyboards and mice (Moberg & Pettersson, 2017). Hand gesture recognition uses image

processing to find patterns trained on data use algorithms. It uses color and depth data to distinguish the hand from the background (Hand tracking and gesture recognition using AI, 2022). Hand gesture recognition can be applied in many fields such as AR, gaming, and home automation (Almaty & Kazakhstan, 2021).

Hand gestures are a natural way of human interaction that plays the important role in collaboration with AR (Huo et al., 2021). Camera-based hand gesturing technology is one of the technologies in hand gesture recognition that is using the camera to capture the hand and finger's movement. Two common methods can be used in this technology. One of the methods is using infrared light to calculate the distance between hands and fingers. The other method is to differentiate the hand from the background by using the hand color. (Moberg & Pettersson, 2017) Both methods can be applied in education for creating an interactive environment for students to learn. AR in education can make learning more interactive because educators can display virtual objects to the students. The interactive learning method can make the students learn and memorize faster (Sinha, 2021).

1.2 Problem statement

Even though the research on hand gesture recognition has made significant progress and attained high recognition rates in several fields, it still faces numerous problems. Hand gesture recognition has a problem with target recognition. Hand gesture recognition is classified into two categories which are static gesture recognition and dynamic gesture recognition. The goal of target detection is to extract the target from a picture stream with a complex environment. It is always a

challenging task in vision-based hand gesture recognition methods to separate the human hand location from other background areas in the image. This is mostly due to a combination of background and unexpected environmental influences (Xu & Dai, 2017).

AR is lack interaction in education because of the lack of interaction in learning (Lee et al., 2020). Education is important for every people because education helps people to gain knowledge, make decisions using logic and fulfill their job qualifications in the future (Abulencia, 2021). However, the ineffective "sitting and listening" learning methods currently used in the school will limit the students in education (Tularam & Machisella, 2018). This method is not suitable for all students because it only focuses on the content that had standardized (Why isn't the traditional education system, n.d). Some students cannot imagine the object they have learned from the book or syllabus, because they never see it in real life such as the Universal, especially Solar System (Ismail, 2020). Besides, the existing learning methods that lack interaction may cause the students to feel bored in their studies (Vafa et al., 2018).

The hand gesture recognition systems need to use high cost of hardware such as a sensor camera and infrared camera the system which is not a user-friendly approach. An example of the infrared camera such as Kinect can produce high quality image and speed efficiency. The gadget is huge and expensive, making it challenging to use in real-world applications. Traditional gesture recognition systems are also inconvenient and limited by the surroundings, such as the use of data gloves. It is not an efficient way for creating a real-world product since customers find it difficult to perform motions when there is simply a clean background or when the face is not moved (Tsai & Tsai, 2021).

1.3 Objective

The objective of the hand gesture interactive in AR learning application is:

i. To identify hand gesture techniques for learning applications.
ii. To design an interactive AR learning application using hand gestures.
iii. To develop a prototype of a learning application with AR using hand gesture technique.

1.4 Methodology

The methodology used in this project is the Agile methodology. The Agile methodology divides the whole app development process cycle into several sub-modules, which are viewed as mini projects. Each submodule goes through every development cycle, from designing to developing, testing, and delivering. It is an incremental and iterative approach to developing mobile applications. Agile methodology is chosen in this project because it is flexible, adaptable to change and has high levels of customers. Agile project management is adopted to shorten the development cycles and increase the frequency of project releases. Agile project management techniques increase the chances of success. There are six phases in this methodology which are requirements, design, development, testing, deploying, and review (Dziuba, n.d).

In the requirement phase, the objective, problem statement, target user and scope of the project will be defined. Starting with the requirement stage is crucial since it establishes the project's initial scope and goal of the project. In the design phase, the requirement should be presented to the supervisor and met with the supervisor (Dziuba, n.d). The prototype AR for Sun

and Moon with hand gesture recognition will be developed to ensure the requirement of the project is included after combining all the requirements needed.

In the development phase, the application will be created after reaching an agreement on the idea discussed with the supervisor. The application is provided in phases over the course of distinct sprints, each of which aims to enhance the previous iteration. The initial release is probably going to go through a lot of change and include new features or functioning (Dziuba, n.d). An application of AR for the Sun and Moon with hand gesture recognition will be developed in this phase. In the testing phase, the application that is developed will be tested to make sure the application is functioning well.

In the deploying phase, the application is working smoothly and the new bug will be fixed at this phase. It is possible to upgrade the application or add new features through additional iterations in the implementation phase. The review phase is the last stage in the agile development cycle (Dziuba, n.d). The finalized application that develops can be presented and the requirement of the project had been met.

1.5 Scope

The project's scope is that the application will be developed using Unity software to model the Sun and Moon with a markerless-based technique. The programming languages used is C# in Unity. The ManoMotion hand-tracking SDK will be used in Unity software to enable hand tracking and hand gesture control. This application will be using the astronomy course which learn Sun and Moon. The application can display models of the Sun and Moon in AR and can be interacted with

using hand gesture recognition. The target user for this application is focused on the students at the primary school who had learned about the Sun and Moon in their school syllabus. Sun and Moon are chosen in this project because Sun and Moon are the star normally will teach in the primary school syllabus. Through this application, the students can have an interactive study and understand the structure of the Sun and Moon.

1.6 Project Schedule

The project schedule for this project will take one year to complete it. It will divide into two phases the first phase is the design phase and the second phase is develop phase. This semester, it will focus on the design phase. The project schedule for the first phase can refer to Appendix A.

1.7 Significance of project

The significance of the project is it applied augmented reality (AR) and hand gesture recognition in the application. Markerless-based AR will be used in this application to display the virtual object of the Sun and Moon. Markerless-based AR is currently used by educators to create a lot of powerful tools and content in their field (Schechter, 2020). For hand gesture recognition, camera-based hand gesture technology will be used in this project. The camera-based hand gesture technology will be used to capture the movement of the hand and the fingers. Camera-based hand gesture is suitable in AR because they can perform the virtual object in the real world without restriction. It also can interact with the camera's field of view (Moberg & Pettersson, n.d). Basic hand gestures such as zoom in and zoom out will be used in this application to move the virtual object on AR.

1.8 Expected outcome

The expected outcome of the project is the application able to work well in performing hand gesture recognition for AR on mobile devices. The application can display the 3D model of the Sun and Moon in markerless-based AR and the virtual visualization of the Sun and Eight Planet can be manipulated under the control of hand gesture recognition. The 3D models can be switched by using different hand gestures in AR. The Sun and Moon models will be animated to make the application more interesting. It can provide an interactive experience for the students to use the application. The virtual object in AR can be zoomed in by using basic hand gestures. Besides, hand gesture recognition will be expected more interactive when dynamic hand gestures are applied in the application. It can reduce the problem of the hand gesture cannot being detected and provided the students with an immersive experience in their study while using this application.

Chapter 2: Background Study

2.1 Introduction

This chapter will discuss the literature review of hand gesture recognition with AR. A successful system must include a literature review because it can be used to identify issues with the current system. Even though the research on hand gesture recognition has made significant progress and attained high recognition rates in several fields, it still faces numerous problems. AR is lack interaction in education because of the lack of interaction in learning. The hand gesture recognition systems also need to use high-cost hardware such as a sensor camera and infrared camera in the system which is not a user-friendly approach. A review of the techniques of hand gesture recognition will list in this chapter. It also will compare the research done by other authors with the self-project. The advantages and disadvantages of the research will be discussed in this chapter.

2.2 Technique of hand gesture recognition

There are a lot of technique will be used in the hand gesture recognition. The techniques such as wearable glove-based recognition, color-based recognition, camera-based recognition, appearance-based recognition, motion-based recognition and skeleton-based recognition will be discussed in this section.

2.2.1 Wearable glove-based recognition

Hand gestures are a form of body language that can be expressed by the position of the fingers, the center of the palm, and the shape of the hand. There have been some research discussions about the wearable glove-based sensor approach as shown in Figure 1. The sensor included inside the wearable glove may be used to record hand movements and location. Additionally, with sensors included in the gloves, they can quickly provide the precise coordinates of the palm and finger location, orientation, and configuration. However, this method prevents easy user and computer interaction since it needs a physical connection between the user to the computer. The devices use for this approach are also quickly high and cannot be afforded by the user. Previously, wearable sensors that were directly linked to the hand with gloves were used to recognize hand gestures. A physical reaction was picked up by these sensors in response to hand movements or finger bending. The data will be obtained through the connection of the computer wire to the gloves (Oudah et al., 2020).

Figure 1: Wearable glove-based sensor approach (Cotoros et al., 2021)

2.2.2 Color-based recognition

The color-based recognition can be categorized into two methods which are color-based recognition using glove marker, and color-based recognition of skin color. Color-based recognition uses a glove marker to detect the motion of the hand by using the glove marked with various colors as shown in Figure 2. This technique has been utilized for 3D model interaction, allowing certain operations, such as zooming, moving, coloring, and typing using a virtual keyboard with high flexibility. The colors on the glove allow the camera sensor to identify the palm and fingers, which enables the derivation of a geometric model of the hand's shape. Color-based recognition of skin color is one of the common methods used in hand gesture recognition applications as shown in Figure 3. This method also can be classified into two methods which are pixel-based skin detection and region skin detection. To represent the color information in an image, color space can be employed as a mathematical model. For skin segmentation, a variety of color space formats are acquired such as RGB, HSV and luminance (YUV) (Oudah et al., 2020). The research by Perimal et al. (2018) presented that there are 14 motions utilizing an HD camera at a close range (0.15 to 0.20 m) and regulated illumination. The gestures were evaluated with three factors, noise, light intensity, and hand size, which directly impact the recognition rate. In research by Pansare et al., the skin was normalized and detected using RGB, and the noise was reduced on the acquired picture by applying a median filter to the red channel (Perimal et al., 2018).