REAL TIME AUGMENTED REALITY BASED TOOL FOR DIGITAL MEDIA PRODUCTION

Project Id: 2021-075

Individual Project Proposal Report

B.Sc. (Hons) Degrees in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

February 2021

REAL TIME AUGMENTED REALITY BASED TOOL FOR DIGITAL MEDIA PRODUCTION

Project Id: 2021-075

Individual Project Proposal Report

A.Vihanga Nivarthana - IT18091380

Supervisor: Dr.Shyam Mehraaj

Co Supervisor: Mr.Thusithanjana Thilakarathne

B.Sc. (Hons) Degrees in Information Technology

Department of Information Technology

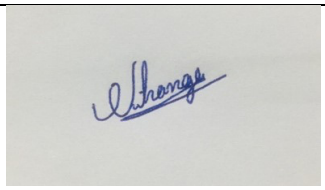
Sri Lanka Institute of Information Technology

Sri Lanka

February 2021

Declaration

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
A.Vihanga Nivarthana	IT18091380	

Name of the Supervisor: Dr.Shyam Mehraaj
Name of the Co-supervisor: Mr.Thusithanjana Thilakarathne

Signature of the Supervisor:
Signature of the co-supervisor:

Abstract

News broadcast is one of the main programs in a television channel or in a radio channel. Both radio and television broadcasting channel measure their popularity separately through the number of viewers tuned that channel during that time. The main difference between the radio broadcast and a television broadcast is the availability of the visual aid alongside with the audio in a television broadcast. The competition between radio channels depend on the content of the news as well as the voice attractiveness of the presenters. But in a television channel, the competition between their channel depend on the visual attractiveness of the news along side with voice delivery. our objective is to increase the visual attractiveness of a television news broadcast. In the past, the whole screen of the television was adjusted to focus only on the news presenter. Because of that the main item on the studio was the table of the presenters. Then with time, different types of screens were placed to add more attractiveness. And they started to improve the studio with more space, screens and used different camera angles to capture the newsroom etc. With these kinds of improvements, channels started experimenting on modern technologies which can be used to add attractiveness. Such technologies are like adding 3D objects and use of Augmented reality. So, these kinds of technical advancements came with challengers for the channels. Such as budget issues, lack of knowledge and expertise and time and planning. Through this point tracking system, a solution for a budget friendly and time efficient will be created. Basically, in this research, points will be identified in a newsroom which are suitable for 3D object placements and according to the different camera angles set on the studio.

Keywords: point tracking, points, news production, camera angles, studio setup, news attractiveness.

Table of Contents

List of figures

List of Tables

1 Introduction

Electronic media is one of the most important tools for transmitting information among the people. Of these, media channels are one of the most important.

Media channels can make a huge difference in the country based on the accuracy of the information transmitted to the public, which can affect either good or bad.
In order to effectively convey accurate information to the public, media channels have a huge role to play in presenting and disseminating information in a way that is easily understood by people of any age.
Unlike audio media, when it comes to presenting broadcasts using visual media, it is essential to use a variety of methods to enhance enjoyment.
Television channels enhance the enjoyment of their broadcasts by displaying various images relevant to the subject, showing short videos made using 3D technology, as well as bringing physical objects to the studio and displaying them to the public.
But beyond that, several media channels in Sri Lanka use the latest technology to amaze the audience and transmit information in a way that is easy to understand.
but, as usual, we still have to watch normal daily broadcasts because the media outlets have to spend a lot of money to activate these systems on a daily basis.
Because the media channels are not able to improve the quality of daily broadcast programs at such a high cost, they operate these systems only for various special broadcasts. These systems, for an example, are used to broadcast election results. The main reason for the high cost of these systems is that there are few people in Sri Lanka who have the technical knowledge to use those systems.
And they should be ready a few days in advance to activate such systems. They bring foreigners with technical knowledge to Sri Lanka and make special broadcasts using their technical knowledge.

This article presents a budget-friendly, efficient and time-saving system that enhances the fun of media broadcasting and makes it easy for even the non-technical person to use.

Our approach is to enable the presentation of broadcasts easily and efficiently using modern technology based on 3D technology and Augmented reality technology.

1.1 Background

Media outlets use a variety of tactics and research to make their broadcasts interesting, but in recent times, in Sri Lanka as well as in other parts of the world, it has become commonplace for media outlets to broadcast inaccurate information. It was also common for people of any age to present information in a way that was difficult to see and understand instantly. For example, it can be pointed out that the values of different graph types are presented incorrectly.
The main reason for this is that graphic designers try to present information creatively rather than accurately.

The aim of this study is to enable media channels to accurately create these types of charts by simply entering values, and to be able to distribute information qualitatively to the public by embedding 3D models that can be easily understood by people of any age when presenting information to the public using Augmented Reality. Setting up an evolving system.

This system not only incorporates 3D models into broadcasts, but also allows the presenter to easily manipulate 3D models using his or her own hand signals, as well as create graphics, as well as present them in 3D.

My responsibility in setting up the system is to identify the points of live broadcast also called point tracking

In this process, it is essential to identify one invariant point in the camera image captured by the camera in order to incorporate three-dimensional models into the broadcast, for example, if you want to place a book on one point on that table, that point will not appear as a different point at all camera angles. Should appear as a single point. This is called point

tracking. What is expected to be done here is to continuously track all the live footage obtained during a live broadcast and identify a point that does not change for each of those screens. We can then use this point to bring this identified point into 3D model broadcasts, just as we would place a book on the table in the previous example. This will enable us to improve the quality of the broadcast and present information that is understandable to all, regardless of age.

1.2 Literature survey

In Literature Survey, I focused on recently published articles on the technology of point pursuit with contemporary reality.

Yin Qin Yu, King Hong Wong and Min Yuan Chong documented the inclusion of unmarked artificial objects in videos under the title "Integrating Artificial Objects Based on Interactive Multiple Formatting with Less-Video Sequences"[1].

I focused on this topic because mark-less technology is well suited to our system, and because it is not suitable for use marker-based augmented reality for the live broadcasts.

They have presented a software solution in their article. Their presentation is a multi-model system that efficiently uses a large number of Kalman filters to track camera movement, positioning 3D models into images.
They point out the differences between the old iterative structure and the multiple model structure they present, and how the multiple model structure performs these processes efficiently.

Genc, S. Riedel, F. Souvannavong, C. Akinlar and N. Navab, "Marker-less tracking for AR: a learning-based approach" [2] in this article They documented a general method for feature-based pose estimation in video streams.
The basic idea of their's article is a vision-based marker-less tracking system that aims at the use of real scene features for estimating the pose of a camera.

where they performed using realistic gestures. They have developed a Lucas-Kanade algorithm in the form of a pyramid and integrated it with the Shi-Tomasi algorithm, using an external tracking tactic. They adjust the posture provided by the external tracker to the 3D points, and then create a display engine that covers the line segments that represent the virtual objects in the wireframe to display the results.

The other article I focused on was U. Neumann and S. You.[3] Research paper related to the research conducted by. It contains a good description of Feature Tracking. Identifying and tracking features using computer vision is a must-have in every augmented reality program. Features here are any object contained in the natural environment.

Another thing I noticed while searching for another topic called Augmented reality is point clouds. Here a large number of points are grouped together to create an object map of the external environment like a cloud Radkowski, Rafael[4]. The research conducted by has been included in the relevant research paper.
When I researched the software, I was able to find similar software. It is called nuke[6]. It is the industry standard compositing software in the field of filmmaking.
It is also a software that can create 3D environments using point clouds.

In a research paper presented by "Khandelwal, Pulkit et al."[5], Augmented Reality technology has been introduced using a depth map of an image. What is done here is to analyze a video from frame to frame and convert the images into a gray scale depth map and display it in 3D. This is a technology that is also used in Facebook. Here we can see objects as 2.5d, not 3d.

These are the profiles I focused on in this topic.

2. Reasearch gap

In this study I discovered some of the problems that exist in performing point tracking with existing systems. These include systems with features that incorporate point tracking into live streaming, but require a large amount of processing power to activate them. Furthermore, the user has little ability to use it easily without technical knowledge. Also, some of the systems identified have to take several days to complete. Some detection systems are only suitable for embedding two-dimensional objects.

But the system we are proposing is a system that can be used even by a person with little technical knowledge, with simple interfaces that can be easily run on a computer that can embed 3D objects into live streams.

In this table, it compares proposed system's features with other systems's features which are already in use.

Features	**Obs**	**After effects**	**vizrt**	**Nuke**	**Our product**
Point tracking feature	✗	✓	✓	✓	✓
Live stream point tracking	✗	✗	✓	✗	✓
Live preview of point tracking	✗	✓	✓	✓	✓
Less processing power	✗	✗	✗	✗	✓
Less Technical knowledge required	✓	✗	✗	✗	✓
Simplified ui	✓	✗	✗	✗	✓
Required higher budget	✗	✓	✓	✓	✗

Table 1 Features comparison

3 Research Problem

Media channels do various things to increase the popularity of their channels. They seek to capture the quality and audience of their broadcasts by presenting summarized information to the public through images / charts, as well as short videos produced using 3D technologies to educate the public on how certain accidents have occurred.

But the problem is that graphic designers sometimes incorrectly incorporate their values in the design of the graphs mentioned earlier. They also spend a lot of time and effort on these activities.

In recent times, misrepresentation by local and foreign channels during their broadcasts has been the subject of much discussion and ridicule on social media. In a recent issue regarding coconut prices, a television channel physically brought some coconuts into the studio and presented prices and other information to the public. But the problem is, is it possible to physically bring everything into the studio like this? To go For an example, the corona virus is a must-see and invisible to the naked eye, suppose they want to present information about the corona virus to the public but they cannot bring a corona virus to the studio and display it. For this, something that does not exist physically with the use of modern technology should be demonstrated to the public as it does physically. We present a system that uses 3D technology with Oggmat Reality to solve these problems.
It also allows channels to easily edit things like graphs. 3D technology allows people to physically display things that do not exist physically.

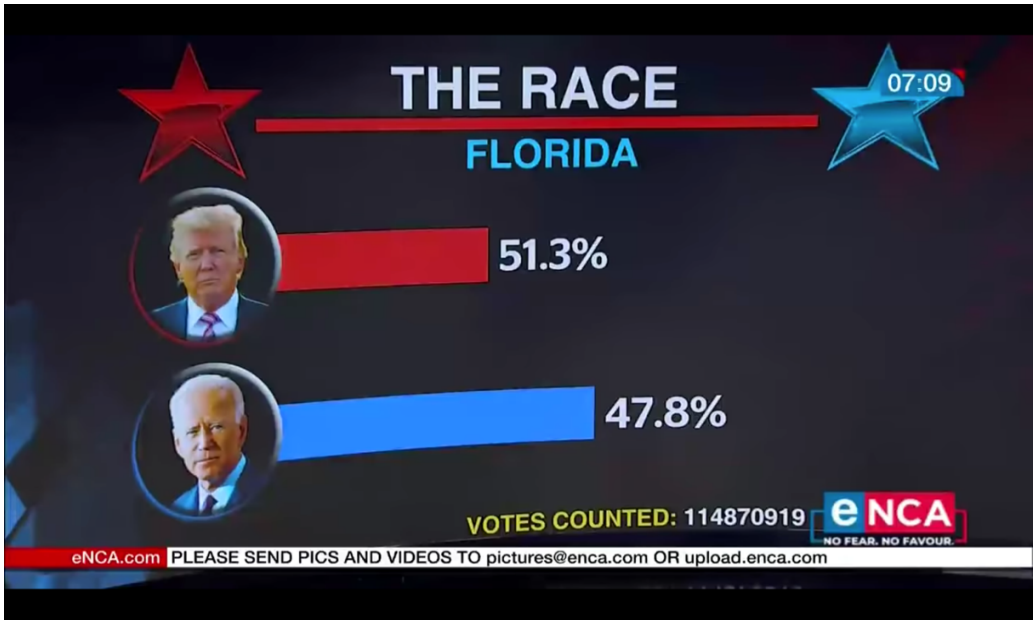


Figure 1 eNCA news broadcast 4/11/2020

4 Objectives

4.1 Main Objective

The main objective of this project is to improve the quality of live broadcasts by incorporating 3D objects. This allows users to easily insert three-dimensional objects at a very low cost, with the ability to easily process graphs, etc. accurately and minimize errors, and to present transmissions that can be easily understood by people of any age.

4.2 Specific Objectives

- Analyze live streaming and set a point where 3D objects must be inserted to fit all camera angles.
- Set up the system as a bug-free and efficient system.
- Ensure that the system is set up at the lowest possible cost.
- Ensure that the system is set up in such a way that the user can operate it easily and quickly (user experience).

5 Methodology

5.1 Research Area

The aim of this project is to develop a system that can improve the quality of live broadcasts using 3D objects. This allows broadcasters to carry out their broadcasts in the highest quality and with the least amount of time and effort. It also aims to enable broadcasters to perform these tasks at very low cost.
The research area of this component used to identify the live broadcast image input point and give the user the function of incorporating it into the live stream as something that exists in the same broadcast location so that it does not change at all camera angles. This allows the system user to insert the desired 3D objects into this identifiable point, thereby creating the

basic environment needed to display the 3D object to the public as a real object in the broadcasting room.

5.2 Architecture of the point tracking

Point tracking is the use of a live streaming image input to identify a single location within which all camera angles remain unchanged. The key elements and basics of this process are as follows.

5.2.1 System diagram

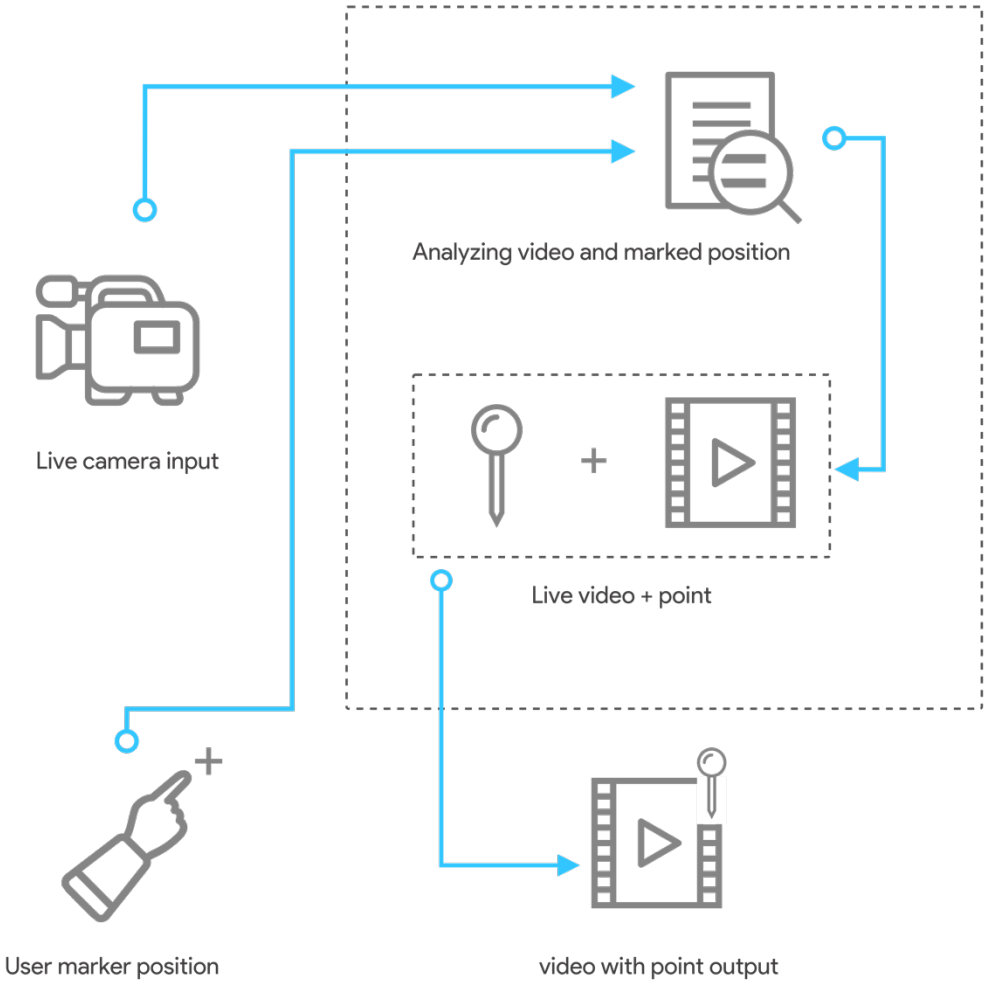


Figure 2 System diagram

5.2.2 High-level diagram

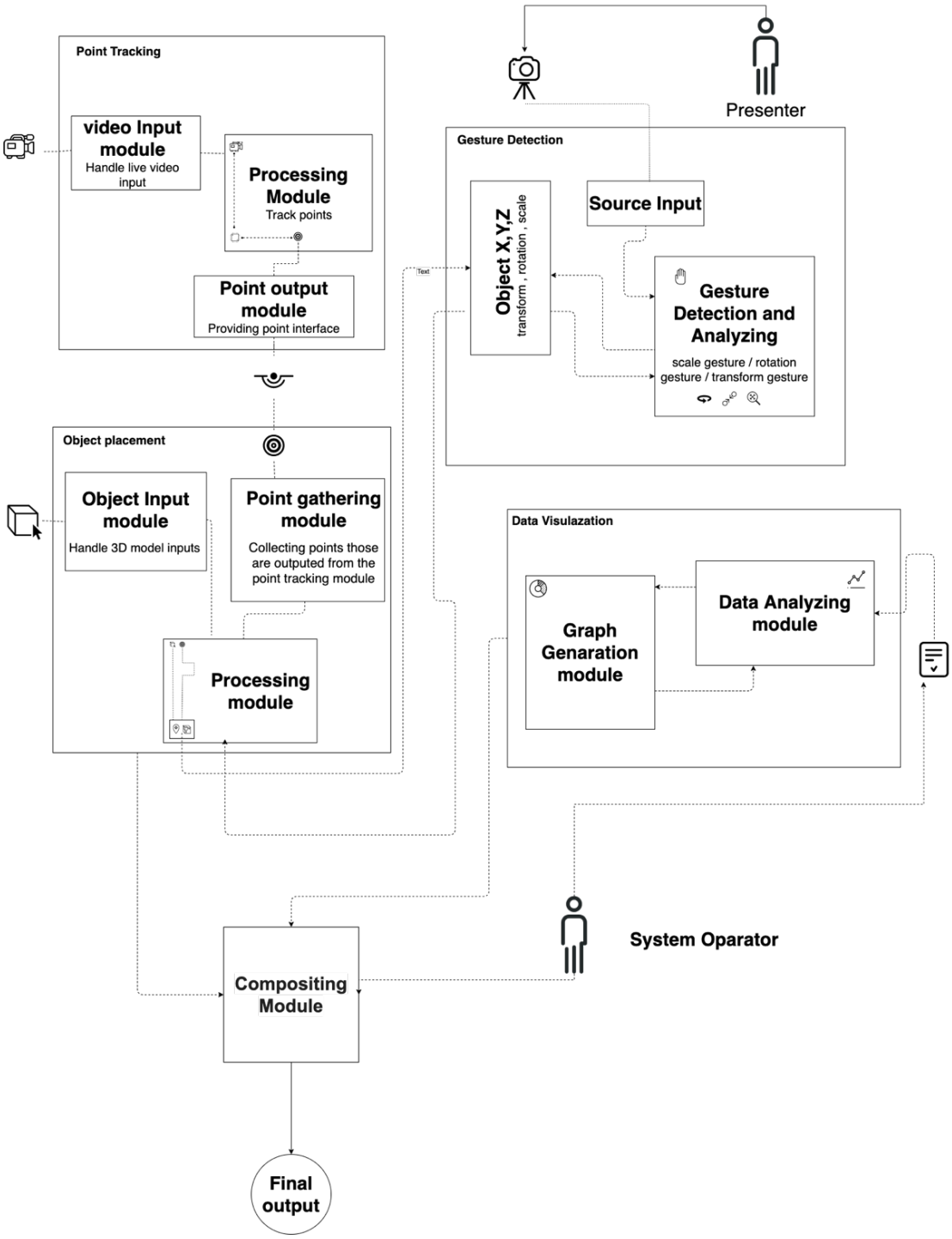


Figure 3 High-level diagram

5.3 Hardware Architecture

Point tracking is done entirely using software and not using hardware. This is done using a computer vision algorithms and therefore does not contain a hardware architecture.

5.4 Software Architecture

We are following waterfall model for the development our system

Waterfall approach was first SDLC Model to be used widely in Software Engineering to ensure success of the project. In "The Waterfall" approach, the whole process of software development is divided into separate phases. In Waterfall model, typically, the outcome of one phase acts as the input for the next phase sequentially.[4]

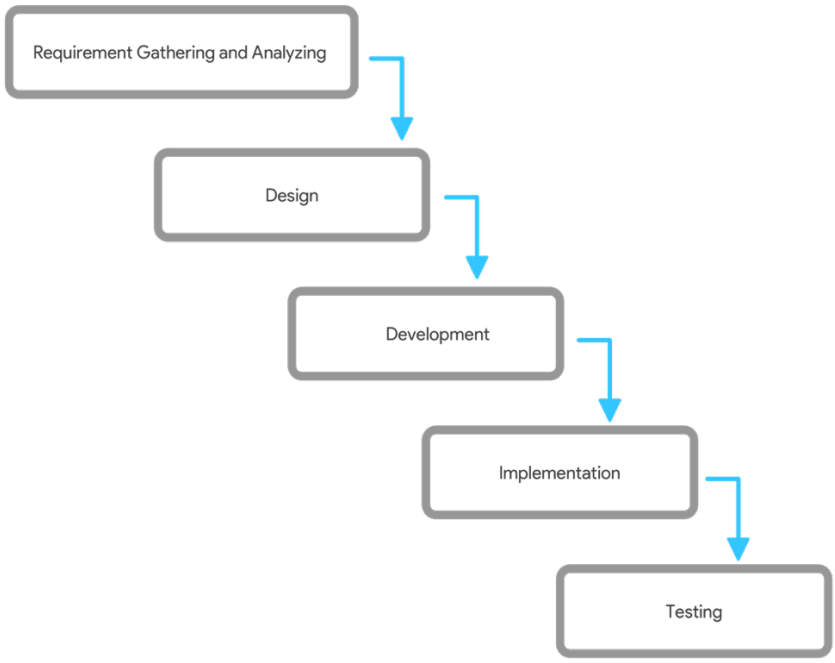


Figure 4 Software architecture waterfall model

5.4.1 Requirement Gathering and Analyzing

The requirements for this test were added by compiling previous test papers and taking a summary of all of them, as well as gaining some insight into the video sources relevant to this subject and their techniques used. We also do research by studying the engineering of existing free software and reverse engineering free software that does not contain any sources.

5.4.2 Design

At this stage we made the basic layout of the system. Discussed the tasks to be performed by the system and how to use the technical components and hardware required to perform those functions. under the guidance of two supervisors, we divided the entire research into four sections, which helped to simplify the project a bit, as we gathered information from

different angles. We designed the perfect layout of the system and how to connect the technical bicycles to each section.

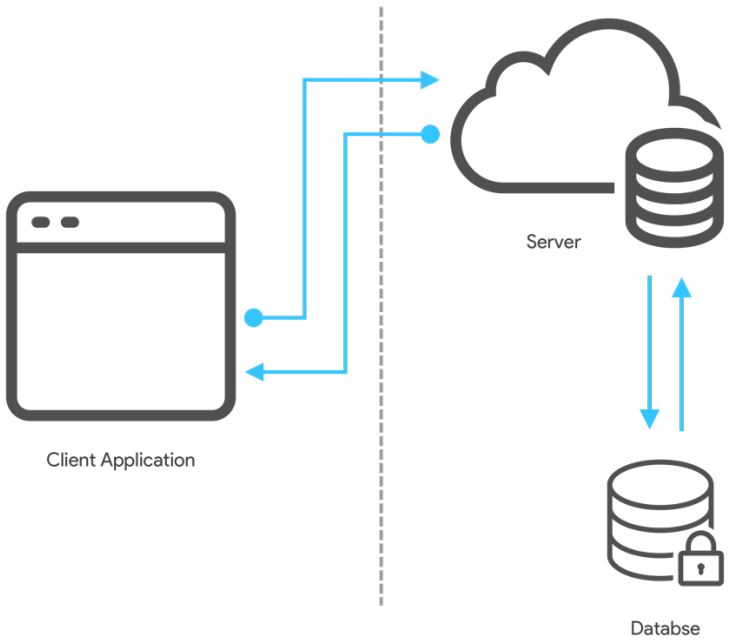


Figure 5 System design diagram

5.4.3 Development and Implementation

At this stage the system is manufactured using all the information and resources planned during the design phase and minimizing the problems that were identified and encountered at that stage. It is hoped at this stage that the individual parts will be implemented as a single system. Team communication and the guidance of supervisors are very important in the process of setting up this system. Once the system is set up, the system is sent back to the testing phase.

Below are the technical aspects related to the components of the system.

Part	Technology
Frontend	React / Electron (Javascript frameworks)
Main Backend API	django (Python) / webrtc / websokets
Server Deployment	AWS Ec2

Table 2 technical components

5.4.4 Testing

After the system is successfully set up. In this section, we will examine several methods to get an accurate idea of the functionality and efficiency of the system. For this, attention should be paid to two main components. They are Beta testing and Backend testing

Before the system is officially released to users, it can be given to a select few users and run in real-time to detect system bugs. It can also be used to get an idea of other features of the system that need to be developed. This is called beta testing.

The other test method is backend testing. All the logic of our system is done on one server. The software used by the user is directly connected to this server, so the software used by the user and all the settings on which the server takes place are separate, so the functionality of the backend is the essence of our system, greatly affects the success of our system. Backend testing is the process of testing through various activities to identify faults in our backend

5.4.5 Commercialization

Lumoz can be sold as a complete package system. The package system can also be used to allow the user to select a package to their liking and to access the cloud library containing pre-created 3D objects.

In the package system, the relevant facilities can be offered to suit each package.

The package can also be introduced for live broadcasts on mobile phones by providing a software version suitable for smartphones. Before purchasing the package, the software can be used for about a month with some features that have been blocked and the ads can be displayed.

5.5 Gantt Chart

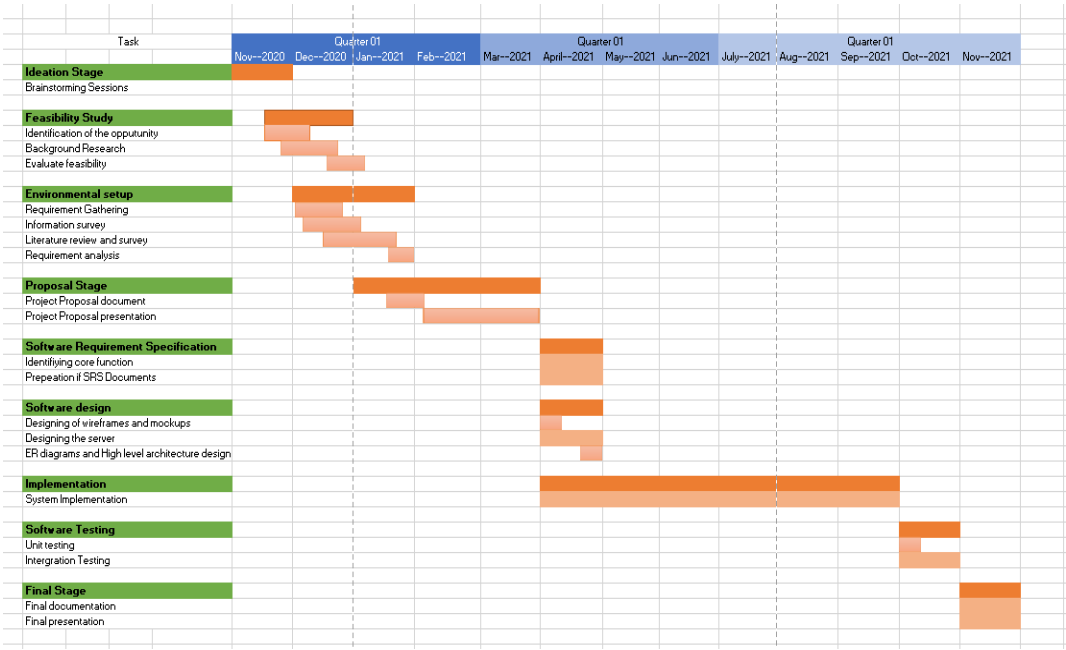


Figure 6 Gantt chart

6. Project requirements

Having a good technical understanding of computer vision and 3D technology is most essential for this project. also need to know about socket broadcast and server knowledge. The main objective of this project is to develop a system that improves broadcast quality by incorporating 3D objects using Ar for live broadcasts. Therefore, the knowledge systems mentioned above are very important requirements.

6.1 Functional Requirements

- Make the input connection between the live video feed with the system
- Identify tracking area
- Continuously track the area and create markers
- Output the markers

6.2 User Requirements

- Knowledge of how to use the user interface
- An understanding of the functions performed by the system

6.3 System Requirements

- Must be efficient.
- Must be able to execute easily and quickly.
- Problems should be kept to a minimum.

6.4 Non-functional Requirements

- Usability - The system should have a simple interface that can be easily understood by the user.

- Performance - The system should be able to run easily even on computers with low processing power.
- Reliability - The system should perform point tracking at the point marked by the user. This should reduce the likelihood of misalignment.
- Availability - The system must be active for the user to use at any time of the day.

6.5 High level user interface wireframe

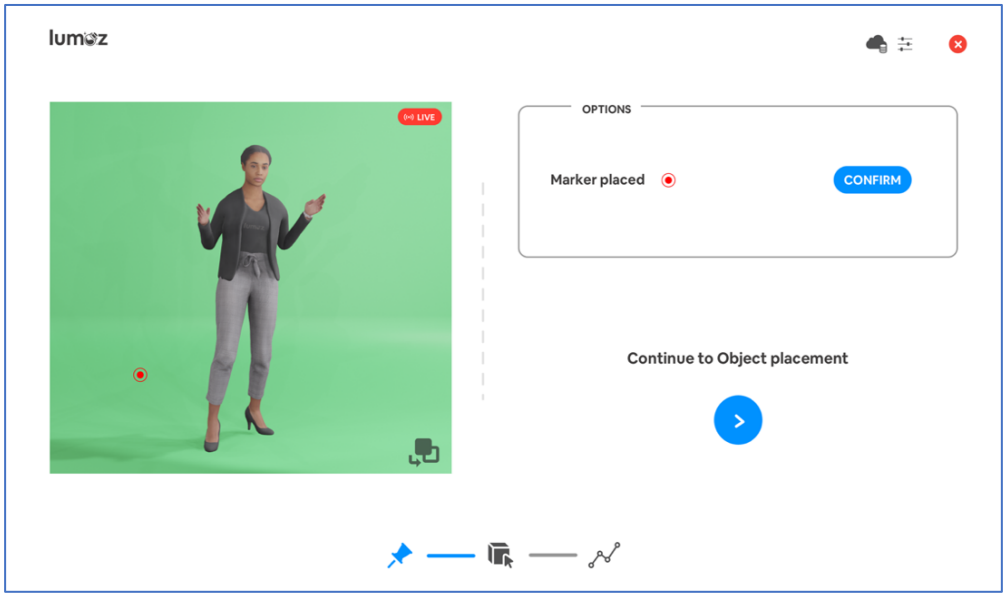


Figure 7 High-level UI Wireframe

7. Budget

Resources	**Price(LKR)**
Electricity	3500
Stationary	3000
Internet	5500
Communication	1000
Paper Publish Cost	4000
Software Purchasing	2000
Total	19000

Table 3 Budget

8. Reference List

[1]Ying Kin Yu, Kin Hong Wong and Ming Yuen Chang, "Merging artificial objects with marker-less video sequences based on the interacting multiple model method", *IEEE Transactions on Multimedia*, vol. 8, no. 3, pp. 521-528, 2006. Available: 10.1109/tmm.2006.870734 [Accessed 24 February 2021].

[2]Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar and N. Navab, "Marker-less tracking for AR: a learning-based approach", *Proceedings. International Symposium on Mixed and Augmented Reality*. Available: 10.1109/ismar.2002.1115122 [Accessed 24 February 2021].

[3]U. Neumann and S. You, "Natural feature tracking for augmented reality", *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 53-64, 1999. Available: 10.1109/6046.748171 [Accessed 25 February 2021].

[4]Radkowski, Rafael. "Object Tracking With A Range Camera For Augmented Reality Assembly Assistance". Journal Of Computing And Information Science In Engineering, vol 16, no. 1, 2016. ASME International, doi:10.1115/1.4031981. [Accessed 6 Mar 2021.]

[5]Khandelwal, Pulkit et al. "Detection Of Features To Track Objects And Segmentation Using Grabcut For Application In Marker-Less Augmented Reality". Procedia Computer Science, vol 58, 2015, pp. 698-705. Elsevier BV, doi:10.1016/j.procs.2015.08.090. [Accessed 6 Mar 2021].

[6] "Nuke, Nukex & Nuke Studio | VFX Software | Foundry". Foundry, 2021, https://www.foundry.com/products/nuke. [Accessed 6 Mar 2021].

[7]*Tutorialspoint.com*, 2021. [Online]. Available: https://www.tutorialspoint.com/sdlc/pdf/sdlc_waterfall_model.pdf. [Accessed 25 February 2021].

9. Appendix

Appendix A: Plagiarism Report

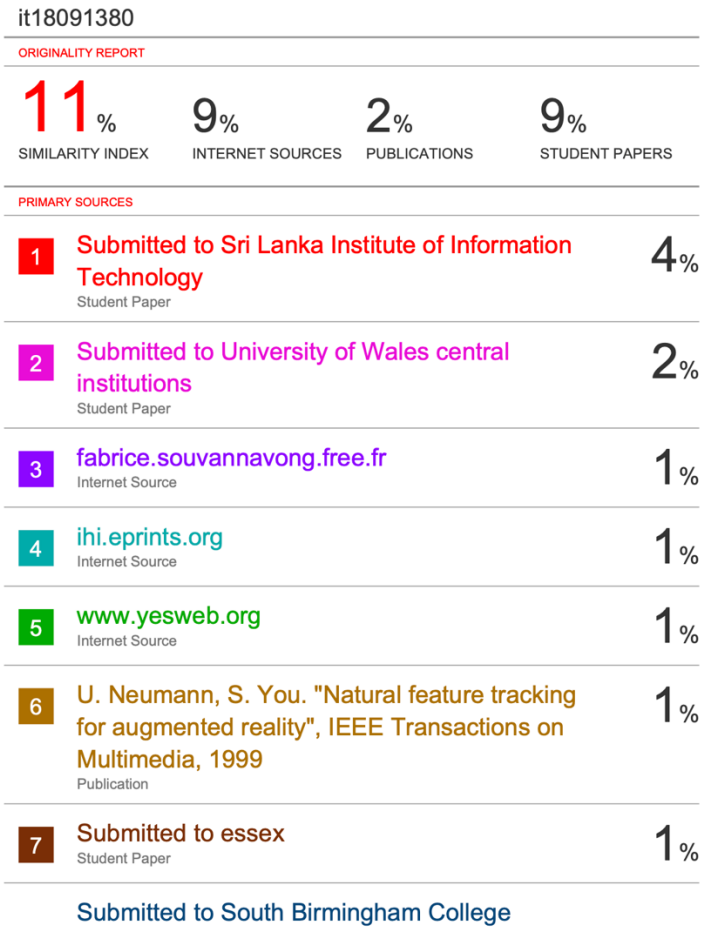
Submission date: 22-Mar-2021 04:56AM (UTC-0700)

Submission ID: 1539251240

File name: it18091380.txt (25.01K)

Word count: 3180

Character count: 22245



it18091380

ORIGINALITY REPORT

11%	9%	2%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Sri Lanka Institute of Information Technology Student Paper	4%
2	Submitted to University of Wales central institutions Student Paper	2%
3	fabrice.souvannavong.free.fr Internet Source	1%
4	ihi.eprints.org Internet Source	1%
5	www.yesweb.org Internet Source	1%
6	U. Neumann, S. You. "Natural feature tracking for augmented reality", IEEE Transactions on Multimedia, 1999 Publication	1%
7	Submitted to essex Student Paper	1%
	Submitted to South Birmingham College	