Handheld AR Device Vision & Scope
February 12, 2021

Handheld Augmented Reality Device

Prepared For
CS 462
Winter 2021

Prepared By
Group 9
Handheld AR Device
Lane Thompson   Christopher Miyasako   Jacob Sowanick

Abstract:
Using OpenCV together with a Raspberry Pi, we will create a program that will allow the user to detect AR markers in their surroundings with a handheld device. It will be capable of taking in a video from a camera module and using computer vision, it will detect AR markers within the video input. The Raspberry Pi will then be used to process the video input and recognize AR markers and will then output corresponding graphics depending on the AR markers recognized. To do this, we have tested receiving video input through the camera module and then using OpenCV code to recognize what is within the input. A blue box was able to be projected onto the corresponding area with regards to the AR markers that were recognized.

Contents

Section 1: Background 3
1.1 Stakeholders 3
Section 2: Vision 3
2.1 Central Hypotheses 3
2.2 High-level Requirements 3
Section 3: Prioritized Project Constraints 4
3.1 Time 4
3.2 Resources 4
3.3 Features 4
Section 4: Scope 5
4.1 Process Flows 5
4.2 User Stories: 5
Section 5: Iteration Plan and Estimate 6
Section 6: Alpha Functionality 6
6.1 Lane Thompson: 6
6.2 Christopher Miyasako: 6
6.3 Jacob Sowanick 6

Section 1: Background

Our project is to use computer vision and the unity engine, run on a raspberry pi, to create an augmented reality device. Our project partner, Dr. Joseph Louis, took inspiration from the EECS AR sandbox on display in the advising lobby. This device is meant to be handheld and easy to use. It will be used like a flashlight, which will help avoid the use of the clunky glasses that most AR systems use.

Every AR system must find a way to see what is in front of it and how to apply an overlay, and that does not change with ours. Our project will need a way to orient itself in a 3d space, and we plan to use QR codes to do that. The QR code will have what the object it is being scanned and how the graphic being returned should be set onto the scene.

1.1 Stakeholders

Our major stakeholders will be construction planners and workers, who will use this device to either design or understand building concepts. This can either be a teaching tool or a tool to help anyone better understand a design.

Section 2: Vision

Upon completion, this device would give users a new and intuitive way to use Augmented Reality. Everyone already knows how to use flashlights, so operation of this device would be as simple as turning it on and pointing. Also, because the digital shapes are shown as a projection instead of through a head mounted screen, it is shareable. One person can operate the AR device, and then anyone standing around them will be able to see the same thing that they are seeing.

The real-life applications that this has extend farther than just fun demonstrations. It could be used effectively during construction as a way for the employees to scan a certain section of wall to see where any studs, pipes, or wires are. It could also be used in an education setting, such as a museum, where guests can illuminate certain displays to see more information in an interesting and interactive way.

2.1 Central Hypotheses

The reasons that this product would be accepted into the current market is because it is a much easier way to access information than the current AR devices out there, which are usually either headsets that are often bulky or uncomfortable, or devices like phones or tablets, which need to be looked through and can be uncomfortable as well. The applications to construction and engineering jobs would allow for faster access to important information that would ultimately make these construction jobs faster and more efficient overall.

2.2 High-level Requirements

The first functional requirement of this project is that this system must be able to capture an image of a specific area in its surroundings and relay the image to the system. After this is the functional requirement that the system will be able to analyze the image taken in by the project to recognize the data being shown to it. Once the data has been recognized by the system, the next functional requirement of the project is to take the appropriate computer information and display it on a surface that the project has been aimed towards.

The first non-functional requirement of this system will be to use a Raspberry Pi Camera Board V2 mounted on the device in order to take in visual data on a surface, specifically a QR code, in order to find the corresponding data required. Using a Raspberry Pi 4 Model B, an Adafruit 9-DOF Orientation IMU Fusion Breakout, and software from the OpenCV library, the system will recognize the components of the image and use it to find relevant data in order to construct a 3d replication of the surroundings based on the data associated with the QR code. A projector mounted on the device will then transmit relevant data depending on the orientation of the device onto a nearby surface.

Section 3: Prioritized Project Constraints

3.1 Time

We are obviously limited by our year time constraint. The class lasts only 3 terms and we would all like to graduate after that. Our project is being broken into three terms with each term getting a specific focus. The first term, the one we are currently in, is going to be spent gaining an understanding of the proposed product and the technologies needed to create. For us, these technologies will mostly be AR, computer vision, and raspberry Pi. We also will be choosing how to organize development and learning how we work as a team. This time spent in preparation will make the second term be much more efficient and productive.

The second term is where most of the work will be done. The technology we spent the first term learning will be used to create working prototypes of our product. These prototypes will allow us to clarify the customers' needs and make a better product.
The third term will be used to finalize our product. Since we may not be able to complete this project, as we finish, we will also clean up our documentation and make sure our code is clean and reusable to make it easier on any future groups.

3.2 Resources

To complete this project, we are given a broad financial range. We have been allowed to choose the items that we believe we need for this project, and can request any additional items in the future, if we do not request things unnecessarily. Currently, we have sent in a request for each team member to receive a raspberry pi and some accessories for the pi like a camera and an IMU. Something that we do not yet have and will need to investigate in the future is something to project images, and maybe a case of some sort. The case can be simple, and we may not even need it if we find some other way to contain the parts in the meantime. The projector, however, is something that will be needed once we get further into the development process.

3.3 Features

The minimum functionality of this program is for the device to be able to take in data about its surroundings using an onboard camera and an inertial measurement system. With the onboard camera, the device will scan a QR code on a nearby surface, using that to know what data will be relevant as well as using it to orient the project device. Using that data, a 3-dimensional scene will be created that will be used to transmit relevant data to be projected onto a nearby surface. For the project, the scope will be centered around the implementation of these features as the minimal deliverable scope and refining these features. If time allows the next features to be implemented will be to add portability by placing these features into the form of a flashlight. Based on the time and resources available to us, a working prototype should be completed by term 3 of the year that will be refined and ready for continued improvement by future groups working on this project.

Section 4: Scope

4.1 Process Flows

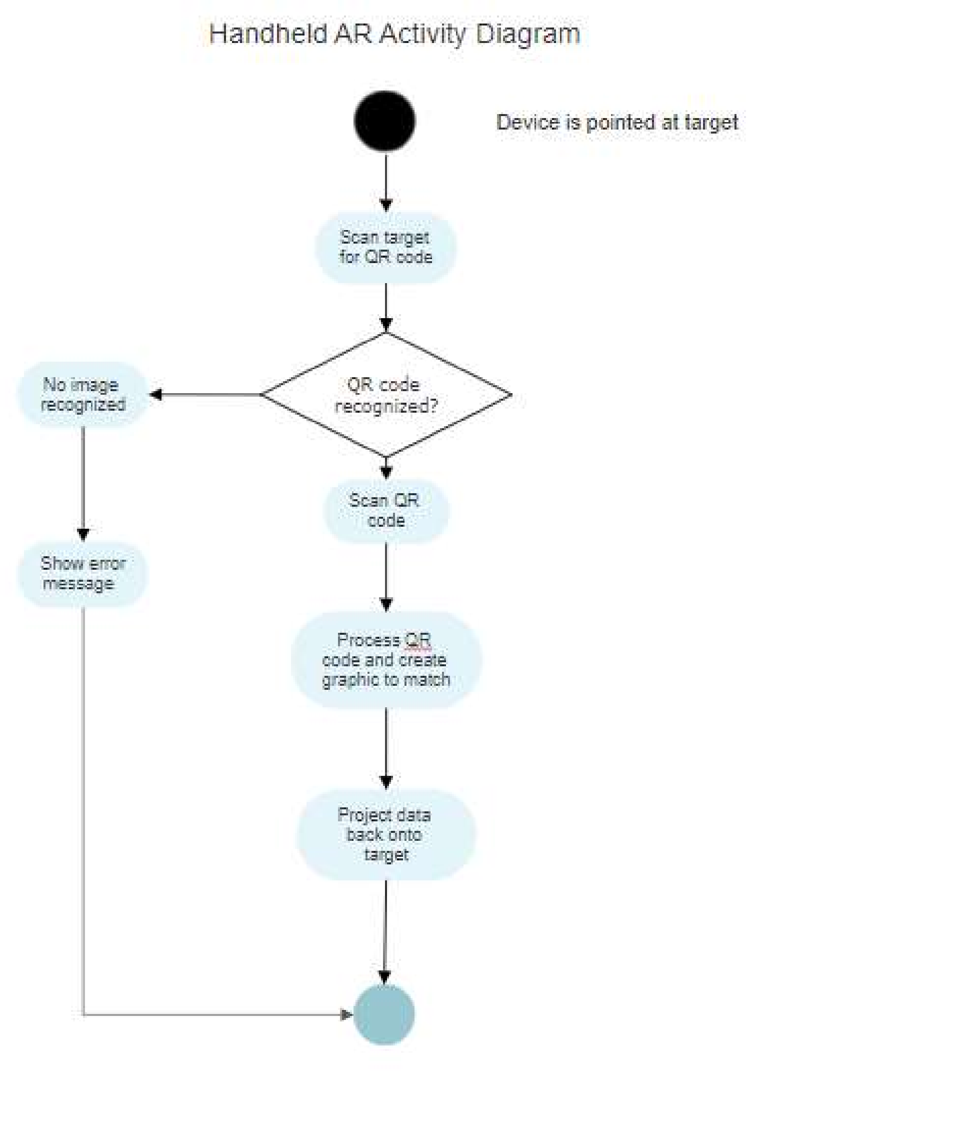

Figure 4.1: The Process Flow Chart for the device. The device will be pointed at a target and use the camera to receive input in the form of an image or video. If the image is not recognized, an error message will be shown. If the image is recognized, it will be identified and used to project relevant data back onto the target.

4.2 User Stories:

As a construction site manager, I need a way to quickly visualize the layout of the house and where different pieces need to be assembled so that I can direct my workers better.

As an architect, I need to be able to show off my designs more easily so that my work can be better understood.

As a construction worker, I need instructions to build the structure so that I can work faster and more efficiently.

Section 5: Iteration Plan and Estimate

Sprint 1 will be spent becoming acquainted with the functionality of the project to gain a better understanding of the knowledge, technology, and resources that will be necessary to make progress with the project. Specifically, Sprint 1 will be spent determining the functionality and details of the project in order to better comprehend the requirements of the project and to judge whether or not it is feasible to accomplish with the limited time and resources provided. For Sprint 2, time will be spent learning about the technology required for this project to better determine the scope and decide on how much is feasible to accomplish in the year given. The technology required for this project will be a Raspberry Pi 4 Model B, an Adafruit 9-DOF Orientation IMU Fusion Breakout, and a Raspberry Pi Camera Board V2 as the hardware equipment used. For software technology, computer vision technology will be researched with a focus on technology such as OpenCV, as well as QR Codes. Other technologies that will be researched include Unity, Python, and Java.

This term will be spent writing code using OpenCV to allow for the device to be able to recognize AR markers within the video input that will be retrieved through the camera. The device will be able to take in video input, process it, and output a corresponding graphic depending on the AR markers that have been recognized.

Section 6: Alpha Functionality

The members of the team include Lane Thompson, Christopher Miyasako, and Jacob Sowanick.

6.1 Lane Thompson:

My name is Lane Thompson, and I am a senior of Computer Science at OSU. This is my sixth year in college, but only my third year as a Computer Science student. I plan to graduate this year with the applied path of Simulation and Game Design.

For this project I have spent most of my time getting the raspberry pi to work in the setup we want and writing code that can be used to recognize a QR code in an image using OpenCV. This portion took a long time for me to get working because of my inexperience with raspberry pi, but my teammate Jacob helped me to finally get the last piece working. I am currently working to connect live video with the code I have written to allow our system to read AR markers from a live feed.

6.2 Christopher Miyasako:

My name is Christopher Miyasako, a college senior that is majoring in Computer Science with an applied option in Simulation and Game Design. I have experience working with C++, Javascript, Java, assembly language, and other programming languages. Over the course of this project, I have spent the first term setting up my Raspberry Pi for the rest of the project, working on getting OpenCV successfully downloaded. I've also been working on the papers that are required for the class, adding to and formatting the milestone papers and retrospectives that were turned in. I have also been looking into using homography to orient the camera depending on the video input, using AR markers to determine the position and orientation of the camera.

6.3 Jacob Sowanick

My name is Jacob Sowanick, and I am a senior set to graduate this Spring with a bachelor's degree in Computer Science: Simulation and Game Design.
My work in this project started with configuring the camera we would use to read a live video. The code I wrote would turn on the camera, take a picture, and then save that photo to a file. Using OpenCV, I am able to project an image over the picture depending on where the AR markers are recognized. Since getting that to work I have put more time into getting OpenCV to work on the raspberry pi with Zbar. I am currently working to be able to read live video from the camera so that I can find AR markers in the feed.