AUGMENTED AUDIO, VOICE RECOGNITION IN MARKERLESS AUGMENTED REALITY

Alawatta Kankanamge Sajith Yugantha Alawatta

IT13030254

Draft submitted in partial fulfillment of the requirements for the B.Sc. Special Honours Degree in Information Technology

Department of Software Engineering

Sri Lanka Institution of Information Technology

October 2016

DECLARATION

I declare that this is my own work and this draft 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 my 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.

Also, I hereby grant to Sri Lanka Institute of Information Technology the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the B.Sc. Special (Hons) degree in IT Dissertation under my supervision.

Signature of the supervisor: Date

Mr. Udara Samaratunge

ABSTRACT

Designing and marketing of furniture according to the exact requirements of the customer has always been a pending problem in the furniture industry. Customers who purchase furniture from dealer have to wait till the furniture arrives at the location in order to check if that products matches with the expectation they had when purchase it. Customer get dissatisfy with the purchased product due to the colors, design or scale not suiting the location where they want it to be placed.

This research study is focused on studying a newly emerging technology called Marker-less Augmented Reality, its techniques used in the present date and selecting an efficient technique in order to develop a system comprising of a mobile application which utilize Marker-less Augmented Reality to place 3D models of the furniture to non-predefined environments. Unity 3D – Game Engine is used for development and the mobile application is created for Android operation system.

In the age where Augmented Reality application are expeditiously taking hold of the market, this study looks into ways of utilizing the said technology to overcome a pending problem in the furniture marketing industry. What methods are current used for augmented Reality furniture applications? What are the possible solutions for mobile marker-less augmented reality? How this solution will help to improve the furniture market? These are some of the questions answered in the pages that follow.

Keywords: Augmented Reality, Marker-less Augmented Reality, Unity 3D, game Engine, Android operating system

ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to our project supervisor, Mr. Udara Samaratunge for his patient guidance, and useful critique of our research work. His willingness to give his time so generously has been very much appreciated. Especially mentioning the moral support and the continuous guidance by providing important feedback enabled us to complete our work successfully. I would also like to thank our lecturer in-charge, Mr. Jayantha Amararachchi for constantly providing us with the guidance and assistance to carry out the research successfully. Also I would like to thank my research group members as well as all the other friends who supported me and encouraged me throughout this project.

Finally, I would like to express my special thank my parents for their support and encouragement throughout my studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

API	Application Programmable Interface
SRS	Software Requirement Specification
AR	Augmented Reality
TTS	Text to Speech
SDK	Software Development Kit
RAM	Random Access Memory
VGA	Video graphic adapter
AAR	Audio augmented reality
Android	Mobile Operating System
OS	Operating System
UI	User Interface
UX	User Experience

1 INTRODUCTION

1.1 Background Context

Furniture manufacture is a vast industry with count-less variations implementation methods and designs. This variations of designs make the marketing of products very complex and cumbersome. Furniture manufacture is a large sub sector within the industries in Sri Lanka. Household furniture is made to look strikingly elegant and comfortable to use; this includes furniture for the bedroom and living room, kitchen furniture as well as any other relevant household or company space.

The marketing aspect of the furniture industry continuously faces a strenuous task of satisfying the customers' demand. Designing furniture according to the exact requirements of the customer has always been a pending problem in the furniture manufacturing industry. Customers who also request ready-made furniture will not know if the selected furniture actually suit them till it has been bought and placed in their home. If they are not satisfied with the furniture they will either refund it or if it cannot be refunded, will have to make do with it. This results customer dissatisfaction as well as the decline of marketability for products.

The system utilizes a technology call Augmented Reality (henceforth referred to as "AR") which uses a mobile device's camera to superimpose a computer-generated image or 3-dimensional (henceforth referred to as "3D") model on to a user's view of the real world. The traditional Augmented Reality tools which were used until now relied upon a pre-defied marker (A image or pattern in the real world) on to which the computer-generated object can be placed. The proposed system will be utilizing the AR technology but will move a step forward in the direct marker-less tracking. This as the name in itself defines will not use pre-defied marker for tracking, instead this technology analyzes the user's surrounding and define track points which can then be used to lock 3D models into place. This method will provide user with greater mobility and allow him/her to view and customize the placed object in any direction and distance without always having to focus on a marker.

1.1.1 Existing Products

- Augment

Figure 1.1.1.1: Augment Vendor Logo

This is an application for sellers to showcase their product catalog in augmented format to customers. After downloading the free Augment app, customers can scan a specific brochure (Marker) to try out all the products offered by the seller in scale in augmented reality before buying.

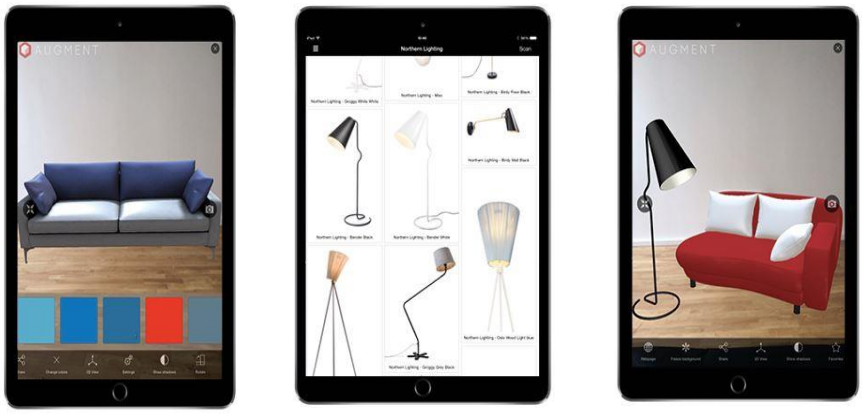

Figure 1.1.1.1: Augment Application

- VIUTEK-uDecore

uDecore is our customisable 3D Room Planner utilising a patent pending Augmented Reality technology for iOS devices. It can visualize multiple photorealistic 3D items to create stunning room planning configurations for kitchens, bathrooms and living rooms.

uDecore can support hundreds of furniture models that can be customised and placed in a virtual 3D room planner. Users can take snapshots of their designs and share them with family and friends. The furniture models can be permanently stored in iOS devices or downloaded from a server.

Figure 1.1.1.2 : uDecore application screenshots

- Sayduck

  Using Vuforia, Sayduck developed a series of product visualization apps enabling consumers to see all kinds of products - from sofas to sinks - within their homes or offices, thereby helping them to make better-informed purchasing decisions, as well as allowing products to be purchased directly from within the app.

  High-end furniture manufacturers and retailers immediately grasped the value of having such virtual showroom apps. Not only would they alleviate the burden of stocking large, expensive items in every color and material, they would make shopping more personal and productive.

Figure 1.1.1.3 : Sayduck application screenshots

- Augmented furniture

The Augmented Reality tool allows you to visualize your selection of objects inside your own room and to make a better evaluation of how it will fit in your own living room.

Once you have found the furniture that fits better in your room you can start configure it choosing your favorite setup, color, material or texture or you can combine it with other objects to complete the décor of the interior.You can then save your configured object in your wishlist and share it with your friends.

Take a screenshot of your furnished interior and send it off via Email, Pinterest, Twitter or Facebook and collect comments and reviews from your friends. They will help you choose the right product for your needs.

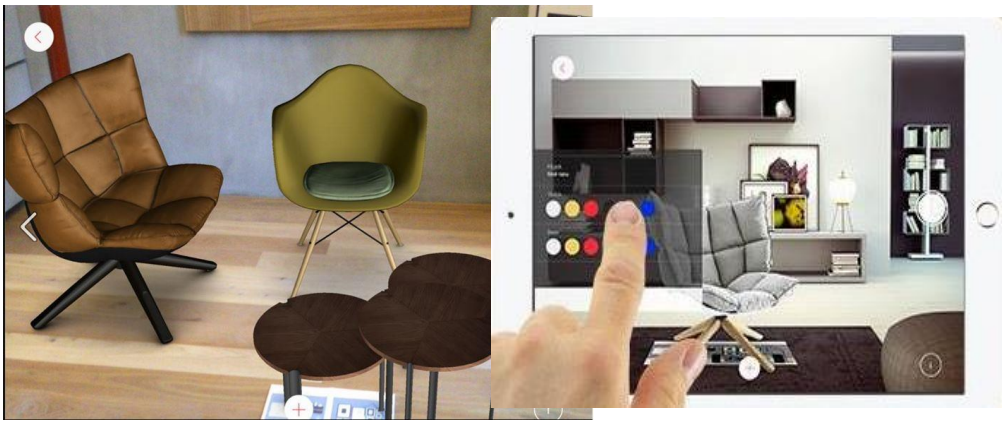

Figure 1.1.1.4 : Augmented furniture application screenshots

1.1.2 Existing products against features comparison

Table 1.1.2.1:Feature Comparison

	Augmnet	VIUTEK	Sayduck	Augmented Furniture	Markless Furniture
Marker less AR	Y	Y	X	X	Y
Support Multiple Objects	Y	Y	X	Y	Y

Support Customization of objects	Y	Y	Y	Y	Y
In-app user feedback	X	X	X	X	Y
Suggestion	X	X	X	X	Y
Voice recognition	X	X	X	X	Y
Augmented Object Details	X	X	X	X	Y
Support Android	Y	X	X	X	Y
Scale detection	X	x	x	X	Y

1.2 Research Gap

Sri Lanka is lacking behind in providing a better customer satisfactory solution for furniture marketing. Several researches have been carried out in fields related to similar problem domain but no practical solution was proposed as of yet. When considering Augmented Reality, solutions proposed were based on marker-based AR which uses predefined markers (An image or pattern of the real world pre-stored in a database) which can be detected in the real world and on which augmented 3D objects are displayed. This use of marker-based AR limited the user to the said predefined markers, as the devices has to detect the markers in order for the augmented reality to work and will always have to be focus on the marker for continuous augmentation. As a general mobile device's camera capability is limited therefore the user is restricted to operate the application only in the vicinity where the marker is visible. To overcome this obstacle marker-less augmented reality was introduced. Marker-less AR used image processing techniques similar to the marker-based AR but in this instant it does not detect markers, hence the name marker-less. Marker-less AR detect and identify the real world itself and calculate the position of device with respect to that calculations this in-turn tackles the operational range limitation of marker-based AR. The next issue which arise in this methodology is the weight of said calculations and image processing that need to be done. Once again as a mobile device's processing capabilities and camera specification limitations the current solutions which are offered by marker-less AR lack the reliability and support necessary to be use in mobile applications.

1.3 Research Problem

Furniture manufacture is a vast industry with count-less variations in implementation methods and designs. This variations of designs make the marketing of products very complex and cumbersome. Furniture manufacture is a large sub sector within the industries in Sri Lanka. Household furniture is made to look strikingly elegant and comfortable to use. This includes but not limited to furniture for the bedroom and living room, kitchen furniture.

The marketing aspect of the furniture industry continuously faces a strenuous task of satisfying the customers' demand. Designing furniture according to the exact requirements of the customer has always been a pending problem in the furniture manufacturing industry. Customers who also purchase ready-made furniture will not know if the selected furniture actually meet their expectation till those said product are brought and placed in their specific location. If they are not satisfied with the furniture they will either refund it or if it cannot be refunded, will have to make do with it. This results in customer dissatisfaction as well as the decline of marketability for products.

1.4 Research objectives

1.4.1 General Objectives

Project "Markless Furniture" objective is to provide an augmented reality mobile application for customers of the furniture market, to help them preview the products they wish to purchase at their location beforehand, in-turn increasing both customer and dealer satisfaction.

- Provide customers with a mobile application that provide 3D models of furniture which can be placed in their home to identify where it suits them before purchasing actual product.
- Provide a means to customize the furniture with the respective customizations provided by the dealer.
- Voice recognition feature to allow customer the ease of customizing 3D models placed in the scene.
- Provide a suggestions feature that identify the real world location's main color and provide furniture suggestion to suit it.

1.4.2 Specific Objectives

- Scan the real world environment and provide data necessary for marker-less AR.
- Identification of position and scale of object with respect to the surrounding.
- Provide means to add multiple object onto the environment.
- Identify and record color histogram to be used in suggestion for furniture.
- Suggestions feature which identify prominent colors and recommend most suitable furniture for the environment.
- Track the object and apply the voice commands for tracked object by identifying customer's voice.
- Provide customizing options via Voice command system.
- Provide details in augmented voice format. (Text to Voice)
- Provide a web portal for dealer to upload the furniture and their respective customization features.
- Provide details of furniture (Prices, Measurements, and any relevant details) next to the respective 3D furniture in the augmented environment.
- Provide services that allow filter furniture search based on budget and other space specifications.

2 LITERATURE REVIEW

In this literature review, the first section will provide a definition for the term "Augmented Reality". Then the next section will review "The Marker-less Augmented Reality". Finally, "Current Mobile Augmented Reality technologies" will be defined with the help of literature.

2.1 Overall Review

The android voice recognition system, according to Yu Zhong, T.V. Raman, Casey Burkhardt, Fadi Biadsy and Jeffrey P. Bigham [1] has an android universal voice control assistance running on devices by synthesizing set of command from on-screen context, and it supports chaining of multiple command in same utterance. It is utilized for commonly used mobile applications. i.e. Sending emails, calling, opening applications, but what our solution utilizes is voice recognition for customization. The voice commands that will be applied in our application to do tasks, within the application's environment and not in global environment.

In Augmented Audio we are focusing on how to apply in augmented audio to Imaginary 3-D model that exists in augmented world for android mobile devices. Augmented reality audio can be used in two main ways one is as a part of aural augmentation, and the other as part of User Interface. As an assembly demo [2] the system will give feedback with beeps. According to LookTel [3], the visually impaired better understand the augmented environment which is smart phone applicable. There taking into account these factors and combining in this application we will be using augmented audio as a means to narrate the details of 3D furniture.

2.2 Why Voice Recognition in AR

Voice recognition is the technology words, phase, sound, spoken by people are converted to digital signal, and pass these signals into coding patterns, which meaning assigned. The concept more generally called "Sound Recognition". We naturally use our voice to communicate within each other in immediate surroundings.

In the context of AR, we suppose to gain the greatest feeling of immersion.

2.3 Speech (voice) Recognition

User input is given as voice command and should be detected via a speech recognition module. Speech (voice) is a continuous audio stream mixed with dynamically changed

states [4]. This is a most challenging part, there are no clearly distinguished parts in an audio recording in the waveform. Words are understood in android (the smallest unit of a waveform) but it is low in true percentage [5].

The common way of speech recognition in android (e.g.: Google Now) is conversational mode. It is taking a waveform, split it by spaces try to identify word such as calendar, alarm. Then apply conversation or open the application. If it couldn't find any matching word convert all to the search query in a google search. [6]

In here we take a waveform, and transfer it to text format, then utterance parsing module process into formalized command, and finally search for actionable objects and manipulate highest marching object /command

2.4 How voice recognition performed

The most popular method is dividing two classes: "Template matching" and "Feature Analysis". Template matching is the simple technique and it is having highest accuracy [6]. but it has a lot of limitations. As with approach to speech recognition, the first step is gathering the input. So that user need to speak phase or words in to microphone. The electrical signal from microphone is digitalized by "Analog Digital Converter (A/D Converter)", and stored in the memory. For the determine meaning of voice phase, the device match the input with digitalized sample, or template. This technique is closed analogy and more suitable for the traditional system that given command from keyboards. The system contains input templates, and attempt to match with actual input using simple conditional statement.

By the way the person's voice is different to each other, so that program cannot possibly contain template for each and every user, so at the beginning the program should train the user's voice before the recognition. During training session, program compute the statistic average of multiple samples. These samples use for templates. To perform the voice recognition, the device should have the "Voice Vocabulary" that is limited to words and commands uses in training sessions. this type of system named as "Speaker Dependent" system. The recognition accuracy can be about 98%.

"Speaker Independent" voice recognition method is process the verbal input into characteristic format using "Fourier Transform Algorithm" or "Linear Predictive

Coding Algorithm (also called as LPC algorithm)". After that finding the similarities between actual digitalized verbal input and expected input. These similarities present for wild range of users. The difference of speech this method can deal with, but for pattern matching it will be failed because of the variety in the voice such as spped, accent, pitch, nationalities. This method will be more difficulty to handle those type of concerns. The overall accuracy of using this method around 90%.

2.5 Text to Voice and Audio Augmented Reality

The input is given as text format by this component it will convert to the speech format. There have been several technologies which have been implemented by basically using AIML based on XML. Google API in Android. The speech covers most instance of making a device for making blind people assistance.

Augmented Audio Reality refers technology that ambled artificially generated audio into user's real environment. According to previously done research on AAR, there is three different geographically definition for head related coordination system [7]

- Origin
- Horizontal plane
- Median plane

In our first iteration, we are going to research about origin based one format perception of distance.

Determine the distance is the challenging part. Distance perspective variance on different perspective/ factors. An important distance factor for sources is the ratio between direct and reverberant sound energy. [8]

For Augmented Audio Reality, we have to convert android TTS APIs matching with android AR.

In our application, we are providing a basic AAR with TTS system to perform augmented reality in mobile technology. For AAR, we have to measure distance according to image factor. The visibility level of the image, we assign according to a scale in origin point factor. These sections are handled by AR assistance module and Narrator will be handled by TTS Agent in Android Platform. According to the scale, we can divide the augmented area and apply augmented audio TTS.

- Combines real and virtual objects in a real environment;
- runs interactively, and in real time;
- Registers (aligns) real and virtual objects with each other.

Although the term *augmented reality* dates back to the early 1990s, the first fully operation and functional AR system originated in the late 1960s [2]. From there onwards AR has been constantly improved and now with increase of computation capabilities of the portable device making its way into it.

3 METHODOLOGY

This research was conducted focusing on two major findings. Implementation of application's voice function part takes 3 different categories.

1. Recognize the customer voice and give relevant result
2. TTS (Text to speech) give the narration of the details
3. Augmented Audio

3.1 Methodology

The software methodology that we used as the process of development of the system is Agile Scrum. Agile Scrum methodology facilitate more flexibility to software development as it allows for requirement change to be handled with ease throughout the process of development. Since research projects are associated with a lot of changes it is hard to follow a set sequence of development such as the waterfall model. Scrum concentrates on task management within a team based environment. As this project is done by a group of four members this methodology will help to team to perform all the formal and informal activities and to solve problems faces by individual members as well as facilitate common understanding of each area of the project by conducting regular meetings. By using this methodology and daily scrum meetings we can found out how many commitments have to made in each sprint., and we could add a sprint backlog (list of task that need to perform or succeed during the sprint).

During the scrum sprint, scrum team takes a set of features from idea and codded, tested those functionalities. The end the feature is successfully tested and done. The scrum model sees daily scrum and team members synchronize the work as they discussed in the sprint, and find the feedback, influences that can be occur in next sprint. In daily scrum meetings we shared what we have done in last sprint and what are the tasks we planned to complete in this day, and divide the tasks that need to implemented among the members, by all these meetings it elaborates and force to complete tasks and finalize the each and every part that successfully completed and tested, integrated to the system.

The development platform used for constructing the application was Unity 3D game development platform version 5.3.1 linked with the Visual Studio 2015 IDE for programming using C# language. Kudan SDK for the augmented placement and identify the camera preview of the object development done by through the unity

development environment. The minimum Android version used was android 4.3 Jelly Bean (API level 18).

'Augmented Reality Furniture Mapping – Voice Commanding part' focusing on covering following tasks.

- Implement a speech recognition API to translate the speech into text format. The speech recognition engine can train focusing a specific person's voice on increasing its accuracy.
- By analyzing input text supplied by speech recognition engine and the process agent can understand the user request
- Agent checks for available commands and method for executing the relevant command that according to provided text details.
- Once generate the output, provide it in proper managed way to user
- Text to voice module gathers the description from web service according to the selected object
- the TTS and AR agent identify the object perspective in mobile camera and module identify the distance of the camera
- AR assistance module provides scale according to the model and divides augmented areas according to the scale of image perspective
- TTS module converts the text into verbal format and according to AR assistance's parameters provide augmented audio experience
- Finally, provide the output of audio and play using pre-configure voice.

'Augmented Reality Furniture Mapping – Voice Commanding part' is the process or set of procedures to be followed to develop a platform to analyze and understand customer's voice and voice command applying suggestions and make necessary changes to the relevant item. The functionality provided by this system it is intended to guide the user to work efficiently and reduce the time consumption.

For enhance the performance and security purposes we used Multi-Threading design for implementation. It means The application is running on multiple threads, i.e. a separate thread is needed to listen for user commands and another thread for the given

commands to be executed, and this is done because if the user wants to stop execution of a command, the user can instruct the system to do so from another thread. This will be impossible to do so if the application ran on a single thread.

Voice recognition Is the technology for identify the sounds, voice, words and phrases spoken by different people in different accent in different regions. The voice signal the analog wave should be transformed into digital signal (1/0 format) that the mobile device can understand. In this part we mostly focused into transfer people voice in noisy environment, with less performance environment. The mobile platform is less in performance, and when we focused on people in Sri Lanka is the main problem is English accent of the people.

For this situation I used google cloud platform, the google API platform provide several access key and authentication code to access cloud API. The cloud API provide different of services regarding a lot of functionalities for the application Cloud Speech API provides the easy integration of user's voice recognition features to our application, it receives the voice code and translate it into the text format from the Cloud Speech API service.

For authentication the application need proper credentials, toauthenticate is identification to services and perform authorization tasks. [4] credentials mechanisms are named in “auth” schemes in android development environment. The API KEY is the simplest authentication method, it is not allowing services authorization, so API KEY is usable on passing data directly to RPC API.

Google cloud voice recognition API contain functions and best practices to access the finalities in cloud, one of main reason for using the cloud platform it will reduce the processing and workload when running the application, the performance is the main task to handle. It recommended use lossless codec to record and transmit, by using we can reduce the loss of accuracy when occurs at the transmission. And we need to use high sampling rate (16000 Hz or Max), use less noisy environment for get inputs those processes is help to get more accuracy output.

The audio pre-processing is the methodology of process the audio before processing the result. It will increase the quality of the audio and clean as much as possible. for processing methods for more accuracy I encoded the audio into Base64 format. The

Base64 encoding algorithm convert the sound clip, to call the REST API Json request to call cloud speech API. [1]

```
// Import the Base64 encoding library.
import org.apache.commons.codec.binary.Base64;

// Encode the speech.
byte[] encodedAudio = Base64.encodeBase64(audio.getBytes());
```

Figure 1.4.2.1 encoding the voice record

word phase hints provide the “hints” to the recognizer and boost the performance of the application, improve the accuracy of words or phrases, that typically user saying words sometimes get same pronunciation (e.g.: “Be”, “Bee”, “B”) so word phase hints methods provide the suggestions all possible suggestions, by filtering those suggestions we can get the exact suggestion that we need.

```
string[] results = data.Split(',');
Debug.Log(" result length " + results.Length);

foreach (string possibleResults in results)
{
    Debug.Log(" possibleResults " + possibleResults);
}

//sample showing the nearest result
string whatToSay = results.GetValue(0).ToString();
logGUI.voice = whatToSay;
getOption(string.Format("Result: {0}", whatToSay));
```

Figure 3.1.2: Getting the selected result

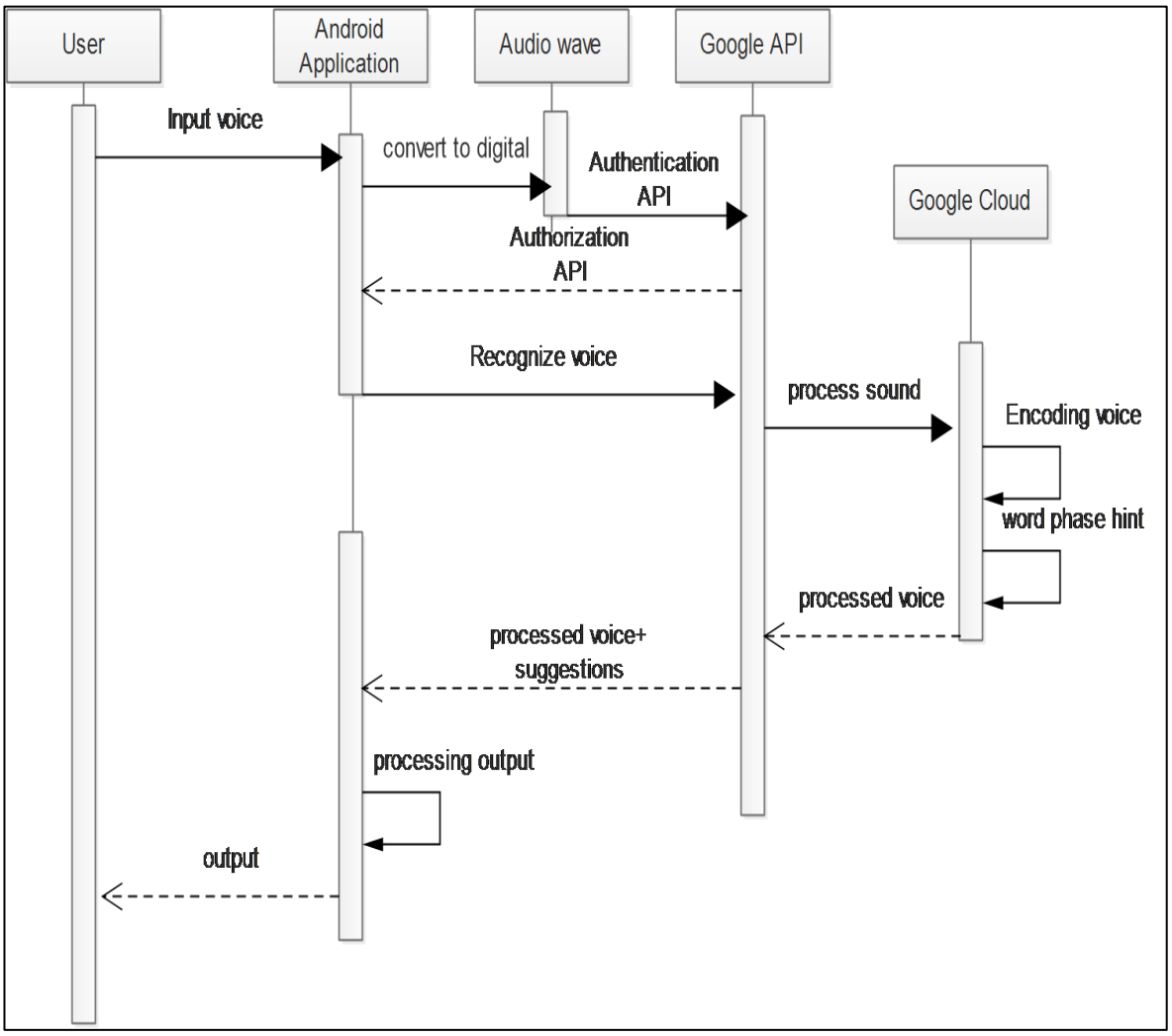

Figure 3.1.3: Sequence diagram: google cloud API

For develop the application we used ANDROID 1.6 version [6]. Android accessibility API provides the Voice recognition plugin for the recognize the voice and get out the relevant results. For that the algorithm convert the voice analog signal to a digital signal. Then it passes the digital signal to acoustic model and language model. In the Language model we can identify the language, pitch, accent and we modify the language package to identify the accent of Sri Lanka according to locale, utterance and pitch. By adding this language package, the system can easily identify the Sri Lankan accent of pronouncing English Language. By using the voice record and its text transcription, acoustic model creates the Statically report of sound that make up each word. This static uses for Speech Recognition engine to recognize the speech. The result is pass to the Speech Engine. The Engine having methods of identifying the speech.

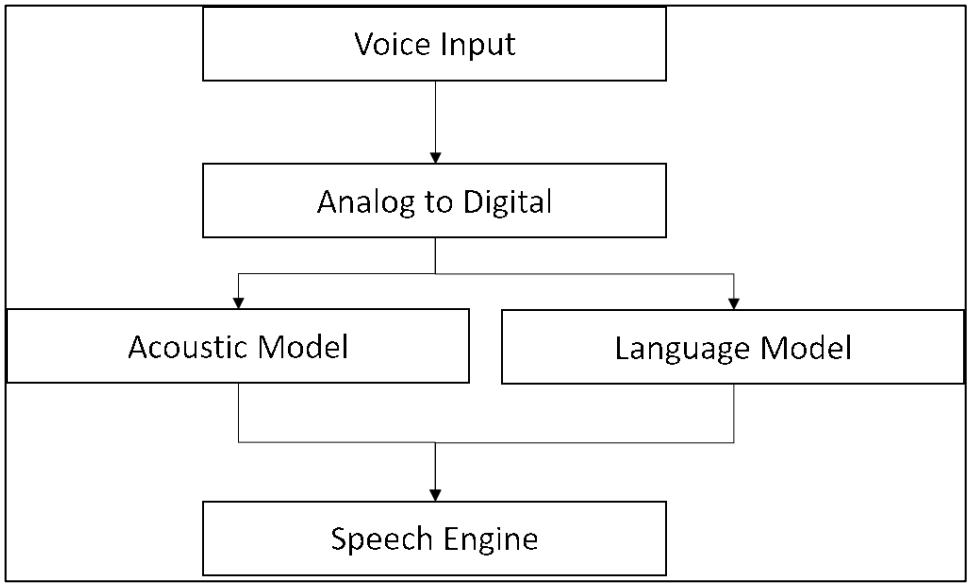

Figure 1.4.2.2: Voice recognition Flow chart

Text to speech is the methodology for Identify the text and convert it into voice output, in ANDROID 1.6 version [6] introduce the TTS function (Text to Speech) to identify the text and convert it into the voice format. [6] We need to define the language before the speak starts, because the Text to speech engine (speech synthesis) doesn't know the exact language that we need to talk, as a word "Mobile" the pronounced in two different ways in UK and US English. So the TTS voice directory is language and accent dependent. So that we firstly need to load the relevant language pack to continue the speech.

Although Android devices (from Android 1.6) support the Text to Speech functionality ship with speech engine, some of devices have less storage. And maybe lack of language packages. So that user need to install those resources from external, TTS API enables system to query the platform for available language files and download if necessary.

```
Intent checkIntent = new Intent();
checkIntent.setAction(TextToSpeech.Engine.ACTION_CHECK_TTS_DATA);
startActivityForResult(checkIntent, MY_DATA_CHECK_CODE);
```

Figure 3.1.5: downloading the necessary packages

The TTS Model Annalise the text and convert it to wave form.

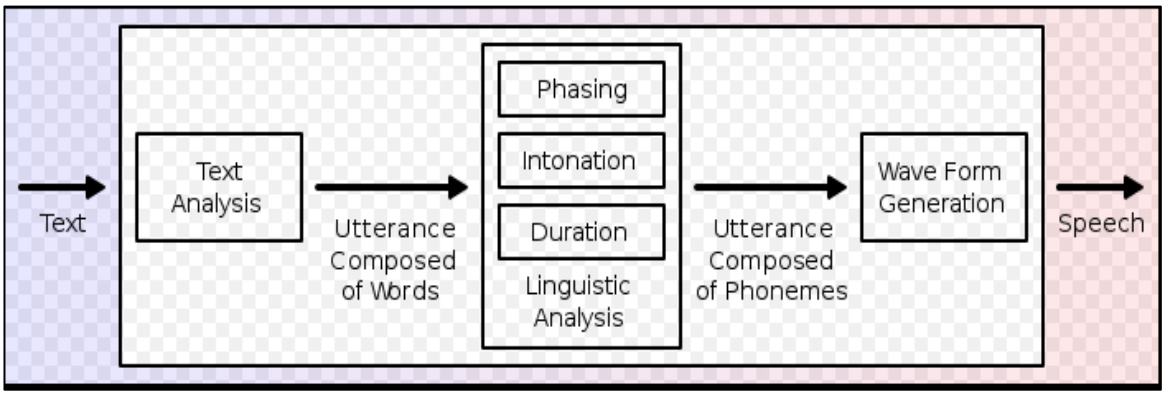

Figure 1.4.2.3: TTS Methodology Source: https://en.wikipedia.org/wiki/Speech_synthesis

The speech synthesis engine manages the queue of entries for synthesize purpose, also called as "Utterances" each and every TTS instance manages its own queue for control which UTTERANCE, that interrupt the current one, and which one is queued. When we call the Speak() , the queue flushed the utterance and new utterance is queued.

Augmented Audio is the methodology of mapping the audio in augmented world. In basic the object become the sound source and it provide the sound source. For that purpose, we need to identify the distance, camera focused object.

For grant the access of the camera we need to modify manifest file as below.

```
<uses-permission android:name="android.permission.CAMERA" />
```

Figure 1.4.2.4: Adding Additional permissions

By accessing camera, we can identify the object that placed in the Augmented environment. For access the augmented environment platform we need to access the Kudan API that holds the augmented reality environment and it's all libraries. So that we access the Kudan camera to get the focused object.

For getting the focused object we send rays from the camera to the object that placed in the augmented area. When the rays hit on the object it identifies the relevant object by using item inventory (that holds the item information that we downloaded in to the environment). By those rays we calculate the vector and the distance for the object then we can get the distance and the focused object details from camera to focused object.

```
Plane targetPlane = new Plane(transform.up, transform.position);
Ray ray = Camera.main.ScreenPointToRay(new Vector3(Screen.width / 2, Screen.height / 2, 0));
//Gets the position of ray along plane
float dist = 0.0f;
//Intersects ray with the plane. Sets dist to distance along the ray where intersects
targetPlane.Raycast(ray, out dist);
//Returns point dist along the ray.
Vector3 planePoint = ray.GetPoint(dist);
//Struct used to get info back from a raycast
RaycastHit hit = new RaycastHit();
```

Figure 1.4.2.5: getting the focused object

By identifying the focused object, we load al the materials that available for that object and customization objects in to material list. And part customization is done by those methods.

3.1.1 More functional requirements for the system are as following.

- **Performance**

If the user enters a valid input to the system, within a few seconds the answer should be generated.

- **Extendibility**

The proposed system is developed in a way that it can be extended by others in future to perform calculations other than addition and multiplication.

- **Reliability**

Reliability is the ability of the system to carry out its normal procedure with minimum failures. To provide reliability to the application testing will be carried to identify and fix each and every possible bug in the system before releasing the final product. Each and every individual component will be tested and finally the integrated system will also be tested under different conditions.

- **Maintainability**

The voice recognition function can be easily maintaining and perform to other languages.

- **Security**

Unauthorized access to the corpus data is restricted.

3.2 Testing and Implementation

3.2.1 Implementation of "Markless Furniture" Application

This sub-topic will explain the high-level structure of implemented application and the features offered for the user and the process through which they operate.

The sequence diagram in Figure 3.2.1 represent the implementation of the voice recognition function, and how it mange to work.

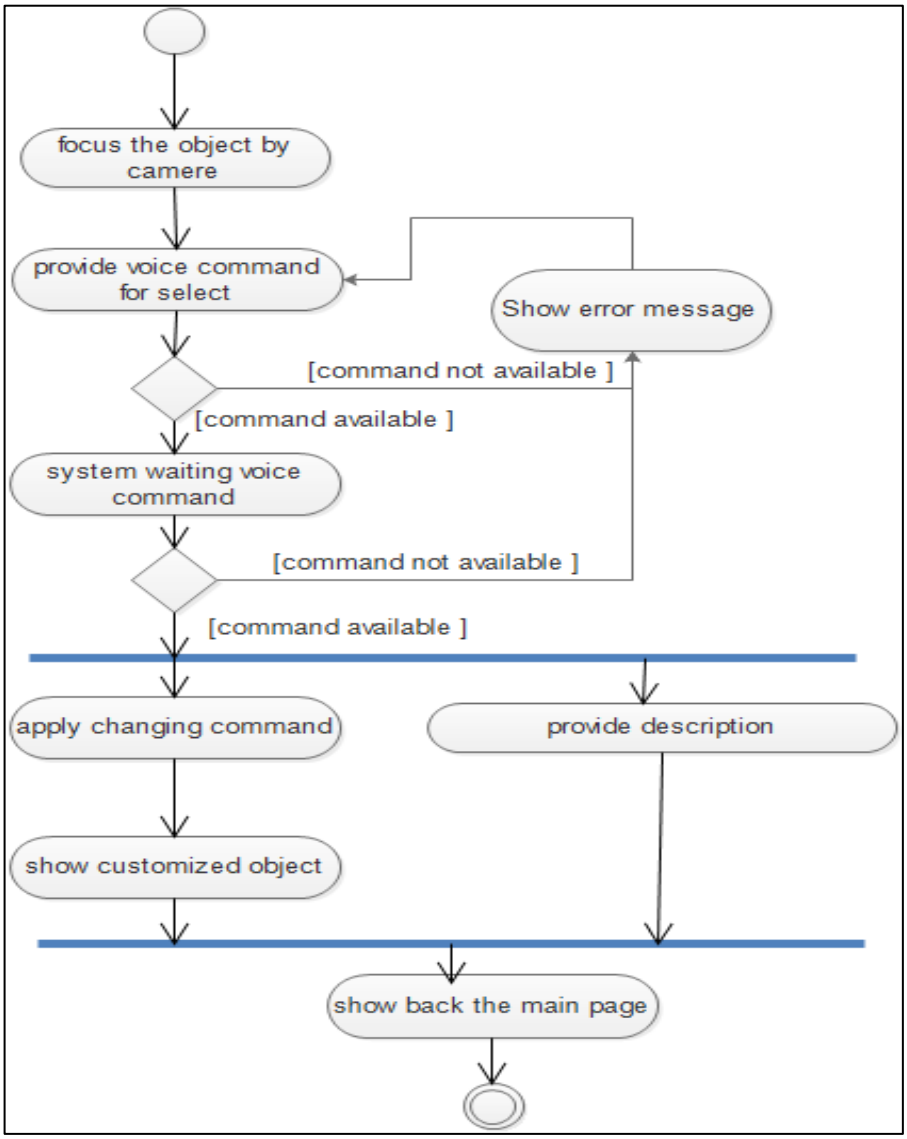

Figure 3.2.1.1.1: sequence diagram

Main Classes implementation as bellows

- Itemdatabase

This class is task with downloading the specific 3D object that the user wish to augment and sort out the information of that downloaded item into item information, the 3D object, the 3D object's color customization information as well as its texture customization information.

- Inventory

The Inventory class contain methods to send information to Itemdatabase regarding the item to be downloaded, and methods to Add or delete the download item. This class is also concerned with inputing the 3D object which was downloaded on the Parent Unity gameObject which contains the MarkerlessTransformDriver in-order for that 3D Object to be augmented onto the real-environment.

- Color Inventory and Texture Inventory

These two classes handle the color or Texture (example for texture if regard to 3D models of furniture are the images of material by which the furniture are been made, Teak or mahogany) information of the downloaded furniture items.

- Color Switcher and Texture Switcher

Once the 3D model of the furniture is being downloaded, its respective color and texture information are downloaded and the model is being augmented, these two class are tasked with identifying selected augmented 3D model and display the color/texture options it contains in order for the user to customize the furniture.

- Kudan Camara

Once the 3D model of the furniture is placed in the camera view the kudan camera need to identify the focused object for the apply algorithm of augmented audio. The distance calculates from the focused object to camera and according to distance voice will be added.

Software interfaces

- Unity 3D

The main integrated development environment on which the mobile application and its respective components are developed.

- Microsoft Azure Cloud Server

This is used for Cloud storage of 3D objects and their respective information.

- Android TTS engine and Optical Character Recognition (OCR Platform)
  For voice narrator need to use android TTS engine and OCR platform. The engine should installed and running properly.

3.2.2 Features Offered to the User

1. Add the augmented furniture model to the marker lees augmented reality environment
2. Identify the real world scale of the environment for the augmentation and utilize it to augmented item to the scale
3. Allow change the position and rotation of the object in augmented environment
4. Allow multiple item at once on the séance.
5. Represent the furniture dealer's description of the furniture on an augmented panel next to the 3D furniture.
6. Allow furniture customization (color and texture) according to the specification given by the dealer
7. Allow narrate the description of the item
8. Allow customize the item via voice commands.

3.2.3 Testing

- **Unit Testing**

Unit testing is the method by which individual units of source code are tested to determine if they work properly. A unit is the smallest testable part of an application. For Speech recognition part it need to identify what we exactly say.

The user's voice input identified correctly and showing it in converted text in GUI for testing purposes. So that we can gerent that the voice recognition part is successfully working without errors.

Figure 3.2.3.1: unit testing

- **Integration Testing**

Integration testing is the phase in software testing in which individual software modules are combined and tested as a group. Integration testing takes as its input modules that have been unit tested, groups them in larger aggregates, applies tests defined in an integration plan to those aggregates, and delivers as its output the integrated system ready for system testing.

After integrating the four separate modules together, system applies the command to the relevant object. That will not have affected to other functionalities of the application. The color of bed change after user input the voice command "Change Color to red"

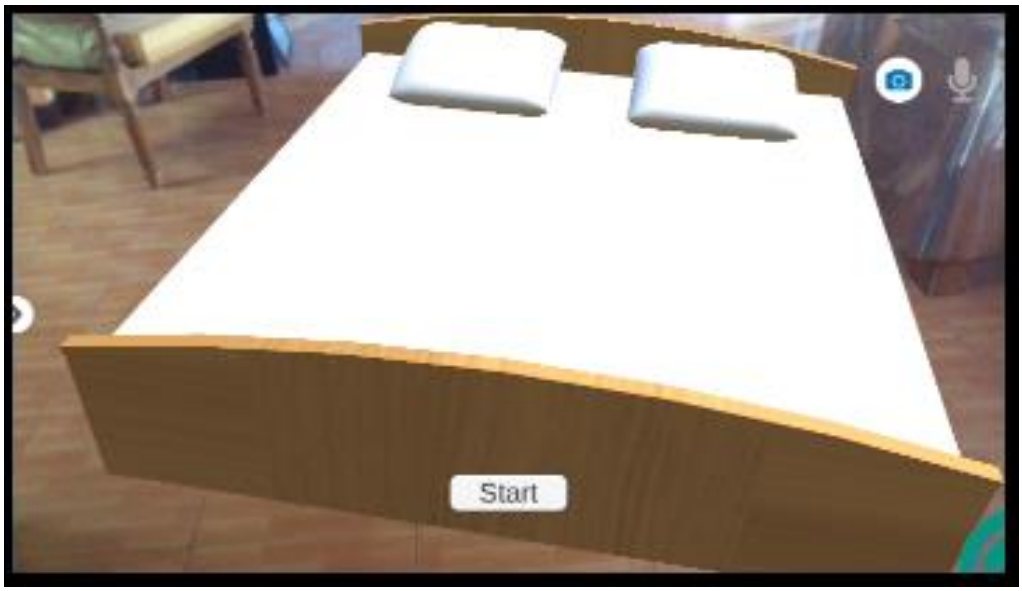

Figure 3.2.3.1: integration testing (before apply command)

Figure 3.2.3.2: integration testing (after apply command)

While integration the system the performance metric need to measure. The internet is need to run voice commands in the application. The test was performed with same mobile device running on normal and power saving performances.

Table 3.2.3.1: Performance Testing

Command	**Method**	**Duration HSDPA+**	**Duration Edge**
Description	Getdescription()	1-2 Seconds	2-3 seconds
Identify users inputs		1-2 seconds	3-5 Seconds
Apply commands to customization	Changecolor()	2-5 seconds	7-10 seconds

3.3 Research Findings

In the current context, Furniture marketing is difficult as customers who purchase furniture may ask for refunds on the product due to dissatisfaction either with the design of the product itself when it was placed in the location that he/she wished, or due to their personal liking changing over time. “Markless Furniture” application was designed in order to overcome this difficulty. The use of this application by customers will help to improve the above mentioned issue, as well as increase the customer user interaction. It will also increase the profit of dealer as the customer satisfaction is increase, while decreasing the risky and burden involving transport of furniture.

The future work for the application include further improvement of the look and feel of the mobile application, improving the immersiveness of the Augment Reality with the introduction and integration with Virtual Reality, improving to identify real-world lighting conditions and applying them to the augmented object.

4 RESULTS AND DISCUSSION

4.1 Results

Designing and marketing of furniture according to the exact requirements of the customer has always been a pending problem in the furniture industry. Customers end up making do with what they get even though that product may not meet their expectation. As of yet no system has been implemented to overcome this problem.

The “Markerless Furniture” mobile application can be used to retrieve 3d models of furniture which can be placed and view through the mobile camera in the location where users wished the real furniture to be place. Allowing them to customize furniture and identify if the furniture they are going to purchase, has the right scale, colour or design to match their home. This new system will drastically reduce the issues that customer face when purchasing, and is more profitable for dealer who will be using this system.

The product has been proven to operate in a wide range of mobile phones which operates on android 4.3 or newer. The application has been tested for operation under different internet connections and performance capabilities.

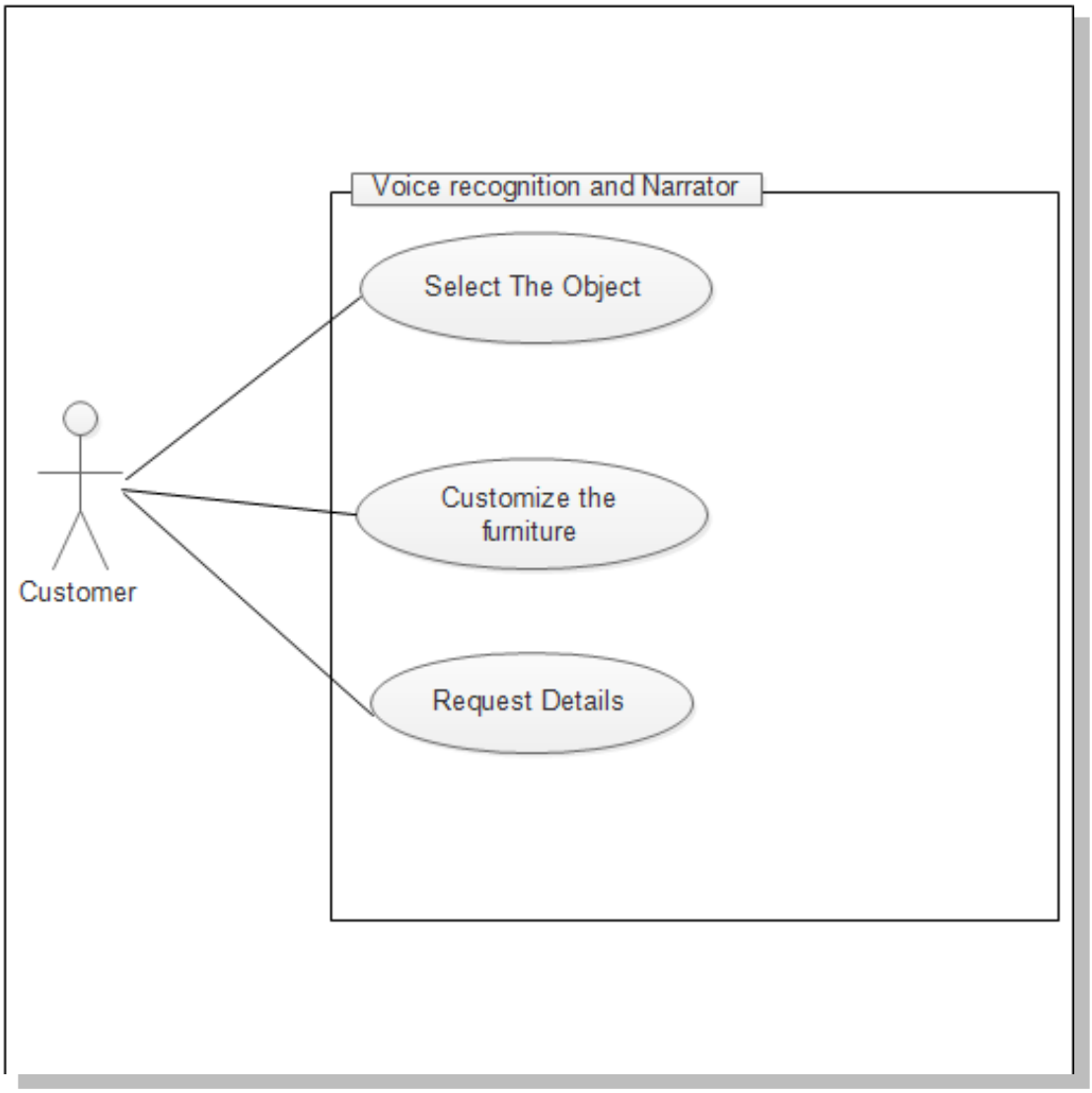

Figure 4.1.1.1: use case diagram

Follow are the test cases with reference to use-case diagram Figure 18

4.1.1 Testing user functionalities

- Select the object

Table 4.1.1.1 Test Case 01

Test Case ID	01
Description	Select the 3D furniture model that camera focused in
Pre-conditions	• Device is connecting to the internet • Application is open

	• The object placed on the environment
Steps	1. Click on the microphone button in left upper Conner 2. Focused the camera to the object
Inputs	
Expected Output	Selected object name appear in view
Actual Output	Selected object name appear in view
Test Status	Passed

Table 4.1.1.2 Test Case 02

Test Case ID	02
Description	Select the 3D furniture model that camera focused in unsuccessful
Pre-conditions	• Device is connecting to the internet • Application is open • The object placed on the environment
Steps	1. Click on the microphone button in left upper Conner 2. Focused the camera to the object
Inputs	
Expected Output	Error name appear in view
Actual Output	“Please focuses the object into the camera view ”
Test Status	Passed

Table 4.1.1.3 Test Case 03

Test Case ID	03
Description	Change of Color of furniture
Pre-conditions	• Device is connecting to the internet • Application is open

	• The object placed on the environment • The item has color customization options • The object is selected for further instruction
Steps	1. Click on the microphone button in left upper Conner 2. Say the valid voice command to customizing with color
Inputs	Voice command “Change color to RED”
Expected Output	The augmented item’s color changes to red
Actual Output	Color changed to red of item (itm003)
Test Status	Passed

Table 4.1.1.4 Test Case 04

Test Case ID	04
Description	Change of Color of furniture unsuccessful
Pre-conditions	• Device is connecting to the internet • Application is open • The object placed on the environment • The item has color customization options • The object is selected for further instruction
Steps	1. Click on the microphone button in left upper Conner

	2. Say the valid voice command to customizing with color
Inputs	Voice command “Change color to Blue”
Expected Output	Voice output error message
Actual Output	Voice output “Sorry invalid option”
Test Status	Passed

Table 4.1.1.5 Test Case 05

Test Case ID	05
Description	Get the description of furniture
Pre-conditions	• Device is connecting to the internet • Application is open • The object placed on the environment • The item has color customization options • The object is selected for further instruction
Steps	1. Click on the microphone button in left upper Conner 2. Say the valid voice command to get description
Inputs	Voice command “get description ”
Expected Output	Voice output , narrator of description
Actual Output	“this bed is made from teak ”
Test Status	Passed

Table 4.1.1.6 Test Case 06

Test Case ID	06
Description	Get the description of furniture
Pre-conditions	• Device is connecting to the internet

	• Application is open • The object placed on the environment • The item has color customization options • The object is selected for further instruction
Steps	3. Click on the microphone button in left upper Conner 4. Say the valid voice command to get description
Inputs	Voice command “description ”
Expected Output	Voice output , error message
Actual Output	“sorry, invalid option ”
Test Status	Passed

4.2 Discussion

This portion mainly focus on discussing about the problems faced during design and implementation and how those issues were resolved. This section also describes how the successful achievements are gained.

The research project “Marker-less Augmented Reality Interactive Furniture Layouting” was started on January 2016, to find a solution for a pending problem faced by customers in the furniture market. The solution was identify as creating a mobile application with the utilization of an emerging field in Augmented Reality called “Marker-less AR” to download and place 3D models of furniture that a customer wish to purchase. The next issue that arose was carried out to identify what kind of technology can be used for the implementation. A literature survey was carried out to identify whether similar types of products were available and if so what limitation were there, and an analysis was carried out to identify and test existing technologies for achieving Marker-less augmented reality in mobile phones. Through this analysis the best technology was obtain on top of which the

implementation of the product was started. During the course of the implementation the application logic was integrated with the framework offered by the previously selected technology to give the user of the application a product which is both easy to operate as well as accurately perform the functionality. During this implementation an issue was identify which limited the functionality of augment reality due to the selected technology relying on a mobile phone gyroscope to give the correct orientation after the 3D object is been augmented. The application could still operate on mobile phone which does not have a gyroscope but had a problem of augmented 3D objects always facing toward the camera no which direction from which the object was viewed. The issue was resolved by allowing the user to correct the orientation when he changes the viewing position.

After the implementation was completed the application was test through a series of test cases of operation and performance. From the test information gathered during the testing process, it is clear that the time taken to download the furniture information depend highly on the network connection quality. Also through the CPU usage testing it was confirmed that the amount of load is minimum.

The end product was then given to several users to operate and provide feedback.

5 CONCLUSION

Designing and marketing of furniture according to the exact requirements of the customer has always been a pending problem in the furniture industry. Customers who purchase furniture from dealer had to wait till the furniture arrives at the location in order to check if that products matches with the expectation they had when purchasing it. Customer get dissatisfy with the purchased product due to the colors, design or scale not suiting the location where they wanted it to be placed. Therefore, this research was carried out in order to find solutions to improve the previously mentioned problem.

This research study provided a solution by developing a system “Markless Furniture” with the use of a newly emerging technology called Marker-less Augmented Reality. This system mainly focused on developing mobile application that help to augmented 3D models of Furniture onto any environment without having to rely on pre-defined markers, and giving the mobility for the user to view the product from different locations after augmenting it, as well as giving the freedom for the user to customize the appearance of the product according to the specifications given by the dealer. The system recognized the real world scale and utilize the information to give the real scale onto the augmented model so that they might also appear in actual scale to the real product. All these functionalities together were used to meet the ultimate goal of this research which is to overcome the difficulties users face when purchasing furniture to suit their expectations.

REFERENCES

[1] " JustSpeak: Enabling Universal Voice Control on Android" by Yu Zhong, T.V. Raman, Casey Burkhardt, Fadi Biadsy and Jeffrey P. Bigham; Google research publication 2004

[2] Siltanen, S., Hakkarainen, M., Korkalo, O., Salonen, T., Sääski, J., Woodward, C., Kannetis, T., Perakakis, M. & Potamianos, A. Multimodal User Interface for Augmented Assembly. Multimedia Signal Processing (MMSP). IEEE 9th Workshop on 1–3 October 2007. P. 78.

[3] LookTel: electronic assistant for visually impaired and blind. http://www.looktel.com [27 April 2012].

[4] Google developer team, "Authenticating to a Cloud API Service", Google cloud platform, [19 Aug 2016]. Available: *https://cloud.google.com/speech/docs/common/auth#setting_up_your_project [Accessed: 02 Sep 2016]*.

[5] Google developer team, "Best Practices", Google cloud platform, [20 Aug 2016]. Available: *https://cloud.google.com/speech/docs/best-practices [Accessed: 02 Sep 2016]*.

[6] Ahamad K. Edly, Mohoammad S. AbdullAll, Shady A. Alefrangy, Waleed E. Aza., "Voice Recognition Service", *Faculty of Engineering publications, The Islamic University of Gaza*

[7] "Enabling technologies for audio augmented reality systems" by Hannes Gamper, Aalto University publication series, doctoral dissertations 39/2014.

[8] Augmented audio reality: Telepresence/VR hybrid acoustic environments by M Cohen, S Aoki, N Koizumi - Robot and Human …, 1993 - ieeexplore.ieee.org