AUGMENTED REALITY SETUP WITH MOBILE PHONES FOR REALTIME

SUMMER INTERNSHIP PROJECT SUMMARY

Submitted by

S SASHANK (2016103060)
APARAJIT K R (2016103505)
S HARINAARAAYAN (2016103029)

COLLEGE OF ENGINEERING GUINDY

## ANNA UNIVERSITY: CHENNAI 600 025

MAY-JUNE 2018

1. PROBLEM STATEMENT:

To detect and recognize objects in an image using the concept reverse image search and use text to speech synthesis to describe about the image with the help of Augmented Reality.

2.FUNCTIONAL DESCRIPTION:

An image of the object is clicked and uploaded to a cloud service and the URL of the image from the cloud service is passes to the Google reverse search API. Since Google reverse search is deprecated we can't directly use the image byte-array from unity and hence we use the cloud service. The API returns a key-phrase which is then passed to the Google custom search engine (or) Oxford dictionary API to retrieve the meaning. The meaning is spelt out with the help of text to speech API and the word is displayed on the focused image with the help of Augmented Reality

3.WORKFLOW:
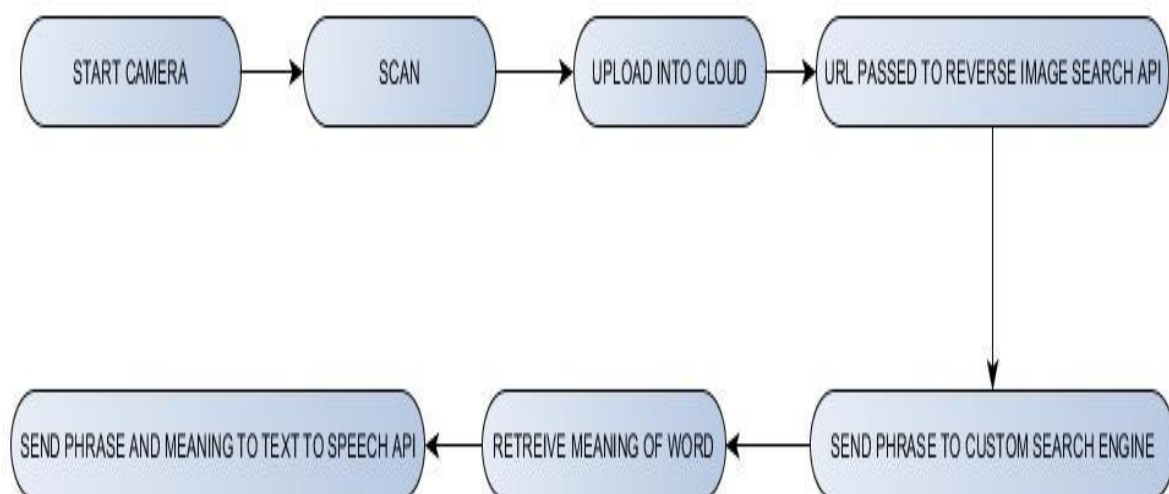
3.1 GENERAL OUTLINE OF WORKFLOW
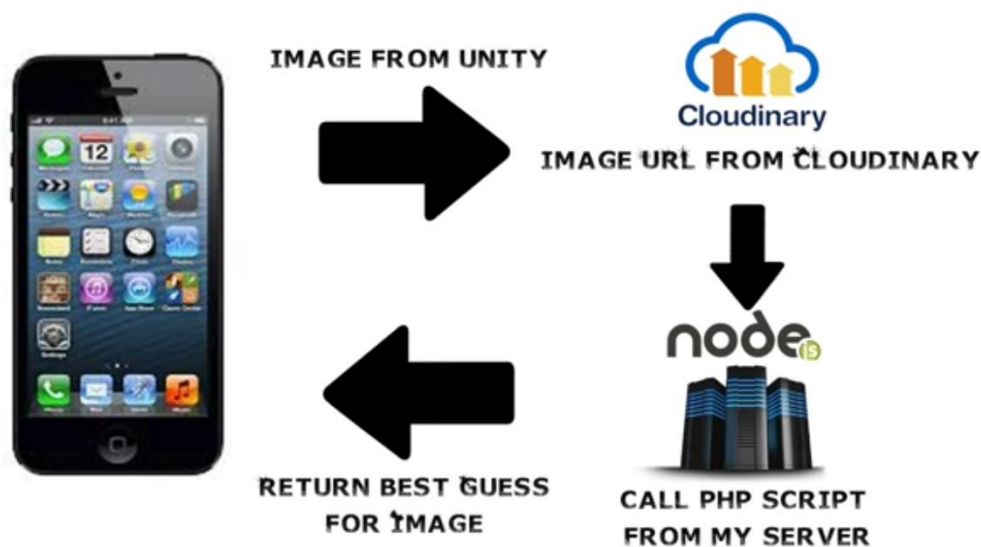
STEPS TO BE FOLLOWED:

1. START CAMERA AND SCAN:
- Select the object for which marker has to be generated
  - Detailed image
  - No repetitive patterns
- Frame the image
  - Scan target
  - Build target
  - Add target to database
- The three colors on top of the camera app indicate the quality of the background in which the image is taken.
  - Red colour indicates its low quality, yellow colour indicates its of medium quality and green colour indicates the background is of high quality (Environment is well-lit).

2. SEND IMAGE TO CLOUD:
- Cloudinary service is essential as the further proper functioning of the project requires image URL
- Google Reverse search works only with Image URLs and not with image files.

3.SENDING CLOUD URL TO REVERSE IMAGE SEARCH API:

IMAGE FROM UNITY

Cloudinary
IMAGE URL FROM CLOUDINARY

node

RETURN BEST GUESS
FOR IMAGE

CALL PHP SCRIPT
FROM MY SERVER

4. USING APIs TO GET THE DESIRED RESULT

The reverse image search API returns the best guess for the image and that best guess is sent to the custom search engine and using web scraping of the Wikipedia links we attain the definition for the word. If there are no such links then oxford dictionary API is used.
The word appears on the focused object using augmented reality and the definition is sent to text to speech API and spelt out.The text to speech synthesis involves conversion of text to phonetic transcriptions which are then split into prosodic units and is converted to output speech by super-imposing the pitch contour.

USAGE OF AUGMENTED REALITY:

- Augmented Reality is implemented in the project with the help of image targets in Unity 3D.
- Image Targets represent images that the Vuforia SDK can detect and track.
- Image Targets do not need special black and white regions or codes to be recognized.
- The SDK detects and tracks the features that are naturally found in the image itself by comparing these natural features against a known target resource database.
- Once the Image Target is detected, the SDK will track the image as long as it is at least partially in the camera's field of view
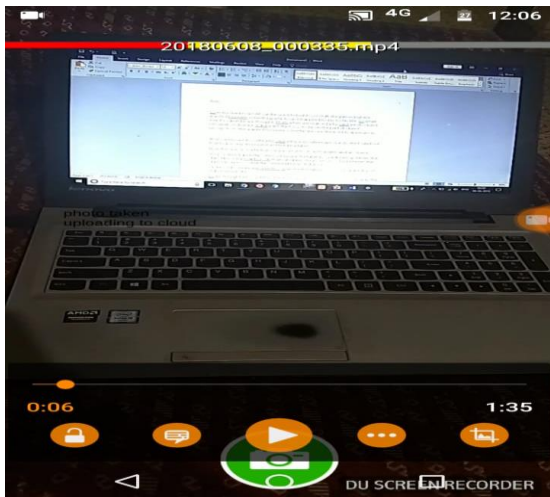
4. SCREENSHOTS OF RESULTS:

The results from our project are screenshotted and displayed as follows. The best guess is displayed in front of the focused object with the help of augmented reality in a three-dimensional fashion. The definition is spelt out with the help of text to speech API as an audio clip.
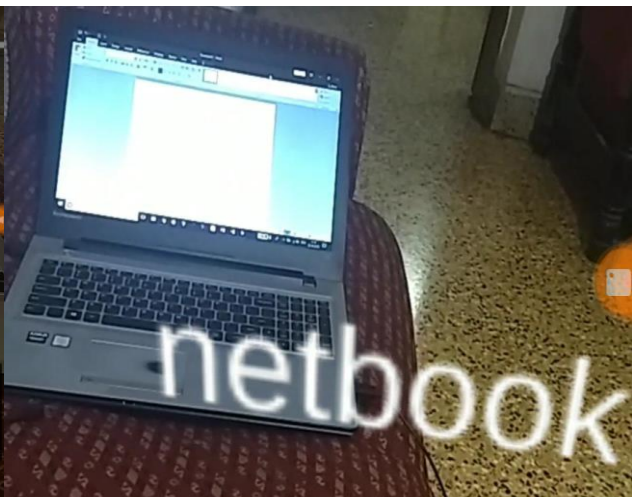
Guess

Capturing Image

Capturing Image

Guess