# Multilingual Digital Signage using Computer Vision and Bluetooth Beacons

A Thesis Presented to

The Faculty of the Computer Science Department

California State University Channel Islands

In (Partial) Fulfillment

of the Requirements for the Degree

Masters of Science in Computer Science

by

Suhas Dwarakanath

Advisor: Dr. Brian Thoms and Dr. David Claveu

October 2018

**APPROVED FOR MS IN COMPUTER SCIENCE**

---

**Advisor: Dr. Brian Thoms and Dr. David Claveu** **Date**

---

**Name** **Date**

---

**Name** **Date**

**APPROVED FOR THE UNIVERITY**

---

**Name** **Date**

**Non-Exclusive Distribution License**

In order for California State University Channel Islands (CSUCI) to reproduce, translate and distribute your submission worldwide through the CSUCI Institutional Repository, your agreement to the following terms is necessary. The author(s) retain any copyright currently on the item as well as the ability to submit the item to publishers or other repositories.

By signing and submitting this license, you (the author(s) or copyright owner) grants to CSUCI the nonexclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video.

You agree that CSUCI may, without changing the content, translate the submission to any medium or format for the purpose of preservation.

You also agree that CSUCI may keep more than one copy of this submission for purposes of security, backup and preservation.

You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. You also represent and warrant that the submission contains no libelous or other unlawful matter and makes no improper invasion of the privacy of any other person.

If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant CSUCI the rights required by this license, and that such third party owned material is clearly identified and acknowledged within the text or content of the submission. You take full responsibility to obtain permission to use any material that is not your own. This permission must be granted to you before you sign this form.

IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN CSUCI, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT.

The CSUCI Institutional Repository will clearly identify your name(s) as the author(s) or owner(s) of the submission, and will not make any alteration, other than as allowed by this license, to your submission.

Title of Item

3 to 5 keywords or phrases to describe the item

Author(s) Name (Print)

Author(s) Signature Date

# Multilingual Digital Signage using Computer Vision and Bluetooth Beacons

Suhas Dwarakanath

September 25, 2018

**Abstract**

Due to globalization and changing lifestyle, more and more people are visiting foreign countries for business and travel. Also lately, a lot of newly arriving refugee families to the U.S face legal consequences. One of the struggles they face is reading documentation they receive through mail; whether bills, court documents or financial assistance documents, they struggle to read and understand them. There are thousands of languages in the world and it is impractical to install signage and print documents in all the languages. In this research, by combining Computer Vision and Bluetooth beacons, multilingual digital information is displayed on the user's smartphone. Smartphone camera allows the user to take a picture of a document. It is then posted to the Google Cloud Vision API which returns the text of that document. It can then be translated to any language using Google Translate API. The system also displays the information of nearby signages (with bluetooth beacons) on the smartphone. This system was implemented in the university campus and the evaluation experiment was conducted by on international students. It was found that the system helps the users to understand their environment better in their native language.

# Contents

# List of Figures

# 1 Introduction

With changing lifestyle and globalization, more people visit foreign countries. Every country is working towards becoming tourism oriented to increase their economy. Most of the visitors use their smartphone to access information. However, when visiting foreign countries, most people face difficulties in obtaining information due to difference in language, which makes it an inconvenience to stay in foreign countries for a longer time.

Recently, a lot of refugee and migrant families from all over the world faced a lot of struggles at the U.S immigrations. A lot of these refugees faced problems with the documentation because they could not understand the content of the legal documents. It is natural for people to understand information better in their home language. Therefore, an effective method of providing and accessing information in multiple languages is required.

In this study, in order to solve such problems, a multilingual information service was developed using Computer Vision and Bluetooth beacons. The information can be accessed from the user's smartphone in most of the languages. We focus on the smartphone's camera to 'see' information in multiple languages. We also use bluetooth beacons to 'push' information to smartphones in proximity. The user can then access the information in multiple languages. By using this method, we expect people visiting foreign countries to access information naturally in the same way they would in their home countries. This paper evaluates various options like GPS, NFC and RFID which could be used to provide information based on user's location. We then describe the required functions and configuration of the prototype system developed.

# 2 Background

Text is highly researched area in the computer vision application to model a smart self-learning system, which involve text associated with shop banners, highway and roadside sign boards or the text on local transport system, such a text provide significant clues about that environment. Text detection helps visually impaired and tourist to convey correct information in more understandable way to user.

As a method of accessing information on public signages, there are two ways of using a smart phone to find information in relevant languages; the user needs to search information by himself, then translate it relevant language to access necessary information. On the other hand, when the digital signage is used, the user can find and acquire information by walking in front of the digital signage without performing any special action. In this study, we explore ways to improve searching and accessing information by using smartphone camera and the method of information provision using the digital signage using iBeacons. This should make users who visit foreign countries get information without making special efforts.

## 2.1 Related Work

A study of multilingual digital signage using iBeacon communication was done, for the purpose of evaluating the technology in an university campus at Japan. [4] In the experiment, the authors installed a total of 25 iBeacon devices at some sightseeing spots, restaurants, souvenir shops, photo shops, etc. in Shirakawa-go, Japan. When vistors look at these heritage spots from the outside, it is difficult for them to understand whether it is a restaurant or a souvenir shop. By using this system, the tourists was able to obtain necessary information in their native languages when they enter each iBeacon area while walking in Shirakawa-go. The paper confirmed that this system can be used effectively for both Japanese and the foreigners [4]. In this system, it was possible to change the displayed language automatically when the user comes near to the digital signage, by using the mechanism of background communication of the iBeacon. **However, the study is dependent on installing and configuring bluetooth iBeacons for the user to be able to understand the signages.**

American MLB has been equipped with iBeacon. Its mobile application can provide navigation system, so the ticket in your hand will pop out automatically and your position will be indicated when you get close to ticket entrance. When iBeacon is connected to the nearest iBeacon base station, the position information of base station will be gained and customers will be provided with high-quality service.[2]

In India, a study of text detection and extraction based on Stroke Width Transform (SWT) and methodology to extract letters was conducted. [6] Major application focus were tourism industry and local transport, to help people to deal with different Indian languages which involve text associated with natural scenes in the local public places. Using SWT method, various Indian languages and their combination with English were detected. **However, this study does not translate the detected texts into the user's preferred languages. This is still a hindrance for foreigners to understand their environment.**

The problem of proximity estimation is found to be difficult in a variety of environment. Existing approaches such as Global Positioning System and Wi-Fi triangulations are insufficient to meet the requirements of accuracy and also requires high cost. A study was made using, Bluetooth which is commonly available in all smartphones to find the proximity over a shorter distance and provide an estimation model to determine the distance based on the Received Signal Strength Indicator (RSSI) values of the Bluetooth [5]. In existing system GPS is used to find the location, it won't work as accurately indoors and inside most commercial building areas. So the authors proposed a system to overcome the problem by using Bluetooth to Bluetooth proximity estimation. By using the signal strength of the Bluetooth device, estimated RSSI value is used to find out exact distance between the devices. This technique helps in tracking the location of nearby user. The study proposed the proximity estimation model by combining Bluetooth RSSI value. The results demonstrate that Bluetooth offers an effective mechanism that is accurate and power-efficient for measuring face-to-face proximity to increase Bluetooth signal length [5]. **While this study does not focus on text detection or digital signage, this is useful for our study to determine the distance of the user/smartphone to the iBeacons**

Optical Character Recognition (OCR) is the electronic conversion of images into machine encoded text. It provides alphanumeric recognition of printed or handwritten characters. OCR has been an active topic of research in the recent past, and has wide applications in banking, healthcare, finance and education. According to the World Health Organization (WHO), around 285 million people around the world are estimated to be visually impaired, out of which 90% live in developing countries. Thus there is a pressing need to develop a system to provide information to the visually impaired. The authors proposed a camera based framework built on the Raspberry Pi, integrated with Image processing algorithms, OCR and Text-to-Speech (TTS) synthesis module [7]. The camera module is used to capture an image of the printed text, and the image was then subject to preprocessing before being fed into the OCR.[7] **This study, however does not focus on language translation or digital signage, we use the adopt image pre-processing techniques like binarization, de-noising, deskewing, segmentation and feature extraction in our research.**

In Taiwan, to encourage learning English, a study was carried out using the iBeacon's micro-positioning function to set the location in museum, restaurant, store, etc. When your mobile phone detects the information of iBeacon's location situation, you can learn by interacting with users' surrounding environment. [2] English obtained from the vocabulary performance includes the single words and expressions commonly used in daily life. With the listing narration and use of straightforward sentence structure and grammar, users can learn English conveniently and rapidly and apply it flexibly to real situations. In this way, users can learn English whenever and wherever possible to enhance English ability, achieve better effects with half efforts and gain language learning fun in the proposed mobile application. With the combination of the proposed application, English learning materials and iBeacon micro-location feature, learners can receive the surrounding related learning materials. In this manner, learners will have chance to build the connection between the learning materials and the environment which can enforce their memory to remember the learning concept. [2] It was found that the system can raise the learning intent of the learners and may improve their learning effectiveness. **However, this study focusses on converting Chinese language to English based on the user's location. It does not support multiple languages and the system is entirely dependent on**

**iBeacons.**

With the rapid increase in data and multimedia services, demand for positioning has increased especially in complex indoor environment which often needs to determine the location information of the mobile terminal. There is a lack of accuracy and robustness in current indoor positioning system. A study was made to design and implement a mobile-based indoor location system which has the mobile applications with the Bluetooth Low Energy technology based on the iBeacon. This study designs and implements an indoor positioning system based on iBeacon [3]. The authors adopted Gaussian filter and Unscented Kalman filter method to robustly extract strong signals from iBeacon device. With the extracted signals, the authors compared them with-in database. They were able to display data based on micro-location on the user's smartphone. It was found that the iBeacon approach has better performance compared with WiFi method [3]. **However, this study does not translate the detected texts into the user's preferred languages. This is still a hindrance for foreigners to understand their environment.**

## 2.2 Outline

With the development of Internet of things (IoT), the user and their location is closely interrelated. The outdoor positioning system based on satellite position (GPS) has matured enough that it provide users with the needs of high-precision positioning. However, GPS positioning satellite signals can only be received in outdoor environment, it difficult to meet the needs of indoor positioning [3]. So, there is an urgent demand for indoor positioning technology and providing micro-location-aware data.

In recent years, various types of emerging short- range wireless communication protocol and related products are presented in the market, where the standards of short-range wireless data communication includes ZigBee, WIFI, Bluetooth, NFC, RFID and Ultra Wideband (UWB) and other types of technology. These technologies play an important role in their respective areas as they meet current application requirements. Let us evaluate some of the options we have to provide data based on micro-location.

### 2.2.1 GPS

The outdoor positioning system based on satellite position (GPS) provides users with the needs of high-precision positioning, accurate to 5 meters. However, GPS positioning satellite signals can only be received in outdoor environment. GPS fails to determine the user's location accurately inside a building or within a room [3]. Hence, we cannot use GPS to accurately provide users with micro-location aware data.

### 2.2.2 Wi-Fi

The coverage of radio wave is very wide, it can reach 100 meters, and achieve coverage with comprehensive. Also, the speed of network transmission is higher, it's useful for real-time interactive facet of the application. The Wi-Fi network construction has a low cost, easy to maintain and very suitable for use in mobile phones. However, the core drawbacks are data security and performance slightly less compared to Bluetooth [3]. It is also difficult to accurately triangulate the user's position inside the network.

### 2.2.3 Near Field Communication (NFC)

NFC is a set of short-range wireless technologies that enable two electronic devices, one of which is usually a portable device such as a smartphone, to establish communication by bringing them in close proximity. However, the range of NFC is less than 10 centimeters and it gets very expensive to install a lot of NFC tags. Thus, it is not feasible to use NFC in our study.

### 2.2.4 Bluetooth

Bluetooth is a radio technology which supports short distance communication from each other, the positioning of Bluetooth technology is the measuring of radio wave signal intensity values for targeting. The main advantage of Bluetooth indoor positioning technology is that, Bluetooth chips are small, easy to integrate in mobile phones and even in smaller devices. On the other hand, the shortcomings of Bluetooth technology are; more expensive equipment, relatively high power consumption and lack of stability when the system in complex space environment [3].

### 2.2.5 iBeacon

The communication protocol used in iBeacon is the Bluetooth Low Energy (BLE), maximum Bluetooth version 4. 0 devices supports Bluetooth BLE version [3]. iBeacon communication frequency consumption is the 2.4GHz band, which is as fast as the Wi-Fi. iBeacon has proximity sensing technology of BLE, it can transfer Uniform Code of unique ID (UUID), APP intelligent terminal obtained the information about UUID and RSSI, and it can be converted into a physical location, which triggers location-aware applications. iBeacons are also low-cost devices and typically run for an year on a single coin CR2032 battery. This makes it feasible for us to implement iBeacons into out study of providing data based on user's micro-location.

## 3 Understanding iBeacon Technology

iBeacon is Apple's implementation of Bluetooth low-energy (BLE) wireless technology to create a different way of providing micro-location-based information and services to nearby smartphones supporting BLE. This allows mobile applications on a smartphone to determine when it has entered or left the region, along with an estimation of proximity to a beacon.

Bluetooth 4.0 is proposed according to the Bluetooth technology standard set by SIG to achieve two-way transmission in technology. Compared with traditional Bluetooth, Bluetooth 4.0 has the advantages of lower cost and lower energy. One button cell can make the Bluetooth Low Energy be operated for 1 to 2 years. [2]

iBeacon devices simply broadcasts its configured advertising packets in specific time intervals [4]. Apple has standardized the format for BLE Advertising. Under this format, an advertising packet consists of four main pieces of information.

**UUID**: This is a 16 byte string used to differentiate a large group of related beacons. For example, if National Park Services (NPS) maintained a network of beacons among all national parks, all national park beacons would share the same UUID. This allows NPS dedicated smartphone app to know which beacon advertisements come from NPS owned beacons. Example: E6FA79C7-804F-4742-863B-BD4D282ED9BA

**Major**: This is a 2 byte string used to distinguish a smaller subset of beacons within the larger group. For example, if NPS had fifty beacons in a particular national park, all fifty would have the same Major. This allows NPS to know exactly which park its visitor is in. Example: 1101

**Minor**: This is a 2 byte string meant to identify individual beacons. Keeping with the NPS example, a beacon at the front of the park would have its own unique Minor. This allows NPS's dedicated app to know exactly where the visitor is in the park. Example: 9901

**Tx Power**: This is used to determine proximity (distance) from the beacon. TX power is defined as the strength of the signal exactly 1 meter from the device. This has to be calibrated and hardcoded in advance. Devices can then use this as a baseline to give a estimate the distance between the beacon and the smartphone.

## 3.1 The algorithm to calculate distance based on RSSI

The rate of signal strength will be unstable because of the volatility of RF signal in actual data collection method. So there was no longer accurate correspondence between the RSSI and distance. Consequently, we need to purify the sample data in order to reduce the relevant positioning errors.

We place iBeacon at some fixed position in the test environment before positioning, the mobile terminal detects the response of iBeacon signal strength, and we build RSSI vector set based on it, this set represents the RSSI vector is R=($R_1$, $R_2$.....$R_i$.....$R_P$),where i is the signal strength of i-node RSSI, P is the total number of iBeacon.

We then repeatedly measured RSSI and averaging the values, the average value as the iBeacon characteristic value r=($r_1$, $r_2$ ,,,, $r_i$ ,,, $r_m$), the formula is as follows:

$$RSSI = 1/m \sum_{i=1}^{m} RSSI_{\mathrm{i}} \tag{1}$$

where m is the total number of collected RSSI vector at the coordinate points. This approach would correct the rather unstable distance of the beacon from the user.

# 4 Google Vision API

Google Cloud Vision is an image recognition technology that allows us to remotely process the content of an image and to retrieve its main features. By using specialized REST API, called Google Cloud Vision API, developers exploit such a technology within their own applications [1]. By using Google's Cloud Platform to compute the content of an image through advanced machine learning processes, this solution allows developers to extract some relevant information from visual data, including image labeling, face and landmark detection, optical character recognition (OCR), and tagging of explicit content. It is possible to interact with Google's Cloud Vision platform by using specialized REST API, called Google Cloud Vision API [1].

We propose to exploit such cloud-based software resources in order to achieve a system for people with understand signages and their environment in general. In particular for users who are visiting foreign countries, our solution may help them to interact with the environment and the things around them. In this paper, we focus on the OCR functionality of the Google Cloud Vision. We also discuss the implementation and integration of the Google Vision into our software solution, a mobile application we developed. The smartphone requires an Internet connection to submit the captured images to Google's cloud platform. The response from platform is processed and displayed on the user's smartphone.

The algorithm shown in Figure 1 has been implemented in Javascript. In particular, here the request for Google Cloud Vision includes the TEXT DETECTION feature. A sample request and response is shown below:

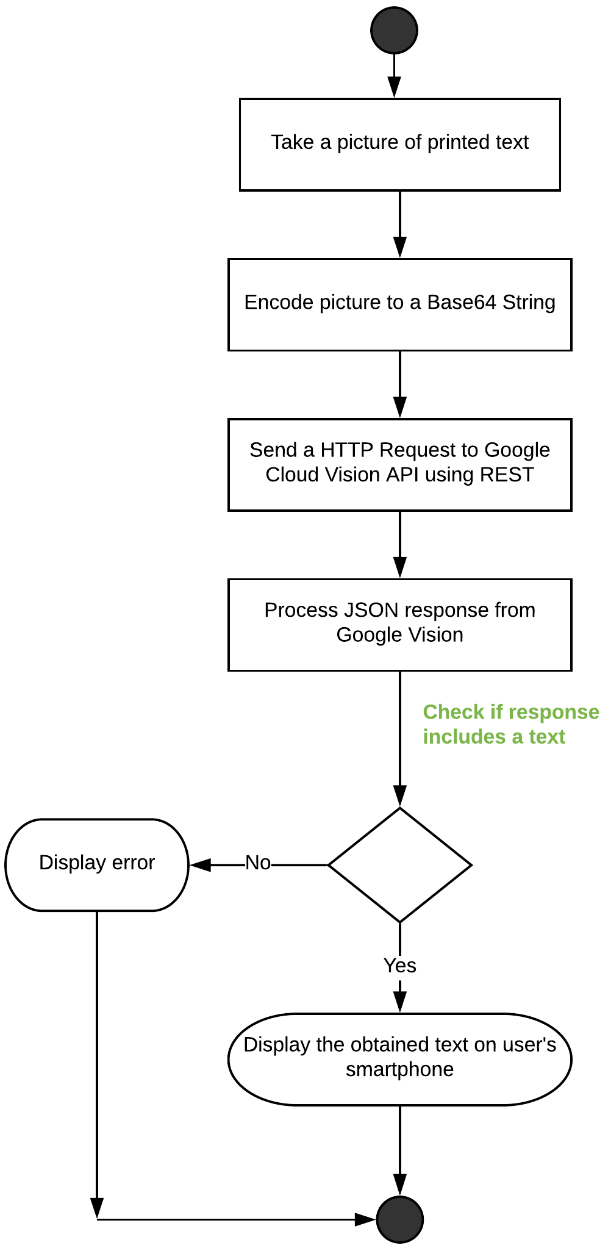


Figure 1: UML activity diagram for Text Detection using Google Vision API

## 4.1 Sample Request

Let us consider the image shown in Figure 2 as the source of the image. We use the Algorithm mentioned in Figure 1 to extract the text from the image.


Figure 2: Example image for testing Google Vision API

We convert this image to a base64 encoded string using Javascript and specify in the headers of the REST API to use 'Text Detection' feature of Google's Cloud Vision API. A sample request is as shown below:


```
{
        "requests":[{"image":{
        "content":"base64_encoded_string_of_image"
},
```

```
        "features":[{"type":"TEXT_DETECTION",
        "maxResults":1}]}]
}
```

## 4.2 Sample Response


```
{
"responses": [{
"textAnnotations": [{
"locale": "en",
"description": "Lorem Ipsum",
"boundingPoly": {
"vertices": [
{"x": 327,"y": 412},{"x": 1158,"y": 412},
{"x": 1158,"y": 680},
{"x": 327,"y": 680}]}}]}]
}
```


In this section, we discussed the capabilities of Google Cloud Vision API and how we implement it into our system.

# 5 Google Cloud Translation API

The Translation API provides a simple programmatic interface for translating an arbitrary string into any supported language using state-of-the-art Neural Machine Translation. It is highly responsive, so websites and applications can integrate with Translation API for fast, dynamic translation of source text from the source language to a target language (such as French to English). Language detection is also available in cases where the source language is unknown. The underlying technology is updated constantly to include improvements from Google research teams, which results in better translations and new languages and language pairs.

Translation API supports more than one hundred different languages, from Afrikaans to Zulu. Used in combination, this enables translation between thousands of language pairs. Translation API can be accessed using REST APIs. We extract text from the printed signages using Google Cloud Vision API, we then send the text to Google Translate API and convert it into the user's desired language.

The figure 3 describes the process of language translation in our system. We obtain the text from the printed media using Google Vision API as mentioned in section 4. We then pass the obtained text to Google Cloud Translation API with the preferred language in the headers of REST API. The response from the Google Translate is then displayed in the user's smartphone. A sample request and response is shown below:

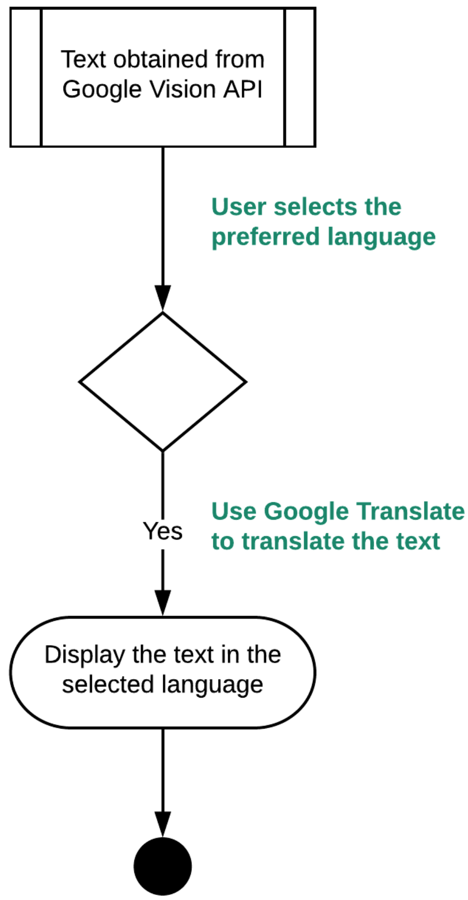


Figure 3: UML activity diagram using Google Translation API

## 5.1 Sample Request

Below code segment shows how to use Google Cloud Translate API using REST APIs using Javascript. Developers should replace their own Google API Key at API-KEY-HERE. The target language to be translated is passed under 'target' parameter of the REST API.

We use javascript to interact with Google Translation API as follows:

```
    fetch('https://translation.googleapis.com/language/translate/v2
    ?key='API_KEY_HERE', {
      method: 'POST',
      headers: {
        'Accept': 'application/json',
        'Content-Type': 'application/json',
      },
      body: JSON.stringify({
        q: "text_to_be_translated",
        target: "target_language"
      })
    })
    .then(data => data.json())
    .then(res => this.cleanText(res))
    .catch(err => console.log(err))
  }
```

## 5.2 Sample Response

Google processes the requests and responds with a JSON object containing the translated text. A sample response is shown below:


```
{
  "data": {
    "translations": [
      {
        "translatedText": "translated_text_here"
      }
    ]
  }
}
```

# 6 System Construction

In this study, a prototype of multilingual information service system was designed and constructed. 3 beacons was configured and placed next to each of the 3 signage in a park. These beacons continuously broadcast their UUID, major and Minor. A mobile application is developed to listen to these beacons. The mobile application leverages the smartphone's bluetooth technology to listen to the beacons.

The relevant information of the signage for each beacon is stored in a database in the cloud. Once the mobile application detects a beacon, it makes a call to the cloud, passing the detected beacon configuration to get the appropriate information of the signage using REST APIs. The information returned from the cloud is then translated to required languages using the Google Translate API and displayed on the user's smartphone.

## 6.1 Hardware

In this study, we used the iBeacons from www.estimote.come. The parameters of the beacons, which are discussed in the section 3 can be configured only by the administrator of the mobile application. To evaluate the system, we use 3 beacons. Below are the configurations of each beacons used in the study.

UUID: E6FA79C7−804F−4742−863B–BD4D282ED9BA
Major: 1111
Minor: 9991

---

UUID: E6FA79C7−804F−4742−863B–BD4D282ED9BA
Major: 1111
Minor: 9992

---

UUID: E6FA79C7−804F−4742−863B–BD4D282ED9BA
Major: 1111
Minor: 9993

### 6.2 Mobile application

## 7 Conclusion and future work

Your work goes here

# References


[1] Maria Fazio Massimo Villari Davide Mulfari, Antonio Celesti and Antonio Puliafito. Using google cloud vision in assistive technology scenarios. *IEEE Workshop on ICT solutions for eHealth 2016*, 2016.

[2] Y. Jeng, M. Hong, and H. Lu. The development of the mobile situated learning application based on microlocation technology. In *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, pages 715–718, July 2017.

[3] X. Li, D. Xu, X. Wang, and R. Muhammad. Design and implementation of indoor positioning system based on ibeacon. In *2016 International Conference on Audio, Language and Image Processing (ICALIP)*, pages 126–130, July 2016.

[4] T. Ogi, K. Ito, and S. Konita. Multilingual digital signage using ibeacon communication. In *2016 19th International Conference on Network-Based Information Systems (NBiS)*, pages 387–392, Sept 2016.

[5] S. Rajkiran and C. Balakrishnan. Rssi estimation using bluetooth in smart phones. In *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*, pages 1431–1434, Feb 2015.

[6] P. Salunkhe, S. Bhaskaran, J. Amudha, and D. Gupta. Recognition of multilingual text from signage boards. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 977–982, Sept 2017.

[7] S. Sonth and J. S. Kallimani. Ocr based facilitator for the visually challenged. In *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, pages 1–7, Dec 2017.