POLITECNICO DI TORINO

Master degree course in Computer engineering

Master Degree Thesis

DEVELOPMENT OF AN ANDROID FRAMEWORK FOR VISUAL NAVIGATION AND ORIENTATION

Supervisors

Prof. Paolo Prinetto

Mr. Giuseppe Airò Farulla (PhD)

Candidate

DEJEN SHIMELIS TADDU

Mat. 192593

February 2018

Abstract

Leveraging on the modern advances in technology, machine to human interaction is becoming essential and both useful in our daily life. A friendly interaction between an automated systems and a human is becoming as important as the functionality provided by these systems. Specially, concerning navigation and orientation systems, and especially when dealing with visually impaired people, the ease of interaction with these automated [navigation] systems is very crucial. Many algorithms and technologies have been proposed and implemented to address the problem related to visual navigation and orientation yet they are implemented using a specific or few set of technologies for a particular domain environment hence they have limitations. This work of thesis presents a hybrid navigation system that combines different technologies such as Camera, GPS, RFID, IBeacon tags and Maps into a single app [ORIENTOMA] to determine the user's location and provide valuable information as per the user request. The proposed system integrates all the above mentioned navigation and orientation systems and technologies into one single application to provide an available system that makes a choice among the provided services according to the context of the environment and provide location based information for the visually impaired user.

The primary objective of this work is to implement a hands-free android application that exploits the different navigation systems or location information providers to allow a visually impaired person to interact independently in both indoor and outdoor environments. The visually impaired person activates the android app once and is then allowed to choose among the different location information through an intuitive voice-based command system. The possible set of response from the proposed system would constitute the direction to a destination, current location of the person, the obstacles detected near the person, the nature of an object detected, and others will be presented to the user in real time. The conversion from speech to text and vice versa is done by the Google Speech API and TTS engine. The object or obstacle detection is done by OpenCV library. The whole system is packed in a single APK and mounted or integrated on a user’s android mobile device during its entire usage.

Acknowledgements

First and foremost, I would like to take this opportunity to express my sincere appreciation and gratitude to my supervisor Mr. Giuseppe Airò Farulla, for his close guidance, suggestions and encouragement over the past years. He has given me continuous help and support when I faced obstacles I did not know how to deal with.

I would like to thank my thesis professors Prof. Paulo Prineto for his tolerance and understanding of my inconveniences and for the long time I took to communicate or update him about my thesis work.

I would like to thank also my friend Luel Mieraf Kiros for formatting the layout this thesis work; Endris Ayub Yimer, Derbew Alemayehu Lishane, and my work colleagues Lorenzo Schiattarella, Donatino Di Carlo, Sabrina Marino for their close collaboration and help during the difficulties in development of this work. In general I would like to acknowledge any persons that either directly or indirectly have supported me in this work.

Last but not least, I would like to thank my parents for their support without which I would not be the person who I am, and I would be not here to present my thesis.

SHIMELIS TADDU, DEJEN

Contents

Abstract II

Acknowledgements III

List of Figures VIII

List of Acronym X

1. Introduction 1

1.1 Introduction 1

1.2 Objective and Scope 2

1.3 Main Contribution for outdoor-indoor localization 3

1.3.1 Outdoor localization 3

1.3.2 Indoor localization 3

1.3.3 Key features 4

1.4 Thesis Outline 4

2. Navigation and Localization Systems 5

2.1 Introduction 5

2. 2 Visual Navigation and Orientation System 5

2.3 Location based Services 6

2.4 Uses of location-based services 11

2.5 Outdoor Navigation 13

2.6 Indoor positioning[31] 13

2.6.1 How Does Indoor Positioning Work?[31] 14

2.6.2 Indoor Positioning, Tracking and Indoor Navigation with Wi-Fi[31] 15

2.6.3 How Does Indoor Positioning with Wi-Fi Work?[31] 15

2.6.4 Indoor Positioning, Tracking and Indoor Navigation with Beacons[32] 17

2.6.5 Functionality of Bluetooth Low Energy Beacons[32] 18
2.6.6 Client-based indoor positioning using BLE[32] 19
2.6.7 Server-based indoor positioning using BLE[32] 19
2.6.8 Parameterization and Interferences of Beacons[32] 20
2.6.9 Use Cases for Indoor Positioning with Beacons[32] 21
2.6.10 Indoor Positioning with Ultra-wideband[33] 22
2.6.11 Functionality of Indoor Localization with Ultra-wideband[33] 23
2.6.12 Setting up an UWB-Based System: Two Approaches 23
2.6.13 Indoor Localization with RFID[34] 25
2.6.14 Camera Systems in Addition to Indoor Positioning[35] 27
2.7 Indoor navigation with Wi-Fi [31] 28
2.8 Indoor Navigation with Beacons[31] 30
2.9 Indoor Navigation with Ultra-wideband[31] 30
2.10 Application of indoor navigation[31] 31
3.Image Recognition and Object Detection 34
3.1 A Brief History of Image Recognition and Object Detection 34
3.2 Components of the Viola-Jones framework 34
3.3 Image recognition and object detection using OpenCV 35
3.4 Object detection in a digital image 35
3.4.1 Object detection 35
3.5 Object tracking in a digital video 37
3.6 Object detection using Single Shot Detectors and MobileNets 38
3.7 Integrating MobileNets and single shot detectors for fast, efficient deep-learning based object detection 40

3.8 Deep learning-based object detection with OpenCV 40
3.9 Their Results 41
4. Related Works and Proposed Method 44
4.1 Related Works 44
4.2 Navigation with voice recognition 44
4.3 Google's searchable, interactive indoor maps for large venues 45
4.4 Proposed Method 47
5. Implementation of the proposed Algorithm 50
5.1 Introduction 50
5.2 General Structure 50
5.3 GPS (Global Positioning System) 52
5.4 Ultrasonic 53
5.5 Text to Speech 53
5.6 Speech to text 54
5.6.1 Speech recognition: 54
5.6.2 Google voice recognition API (speech to text) 54
5.5 GPS (Global Positioning System) 54
5. 6 RFID 55
5.6.1 Tags 55
5.6.2 Readers 56
5.6.3 Signaling 56
5.7 Camera 57
5.8 Bluetooth LE IBeacon detection 58
5.8.1 Difference between IBeacons and RFID 59

5.9 Indoor mapping: 59
5.9.1Google's Indoor mapping made flexible and easy to manage 59
6. Results and Performance Analysis 60
6.1 Results 60
6.2 Performance Analysis 64
6.2.1 Detection time vs. quality of object matching 64
7. Conclusion and Future works 66
7.1 Conclusion 66
7. 2 Advantages 66
7.3 Future works 67
BIBLIOGRAPHY: 68

List of Figures

Figure 1: Geofencing in an Indoor environments[31] 7
Figure 2: Infsoft automation process of Geofencing[31] 8
Figure 3: Basic components of LBS systems 9
Figure 4: The different components of LBS 10
Figure 5: Client-based and server-based indoor positioning[31] 14
Figure 6: Indoor positioning via client[31] 14
Figure 7: Indoor positioning using Wi-Fi[31] 15
Figure 8: Restrictions for server-based detection via Wi-Fi[31] 16
Figure 9: Comparison of different technologies for client-based indoor positioning[31] 17
Figure 10: Beacons for client –and server-based localization[32] 17
Figure 11: Exemplary deployment of a beacon infrastructure[32] 18
Figure 12: Client-based indoor position using BLE[32] 19
Figure 13: Server-based indoor positioning using BLE.[32] 20
Figure 14: Indoor positioning with Ultra-wideband[33] 23
Figure 15: Server-based positioning of Ultra-wideband[33] 24
Figure 16: Client-based positioning of Ultra-wideband[33] 24
Figure 17: Application of RFID for tracking and asset management[34] 26
Figure 18: Object identification using RFID.[34] 26
Figure 19: Combination of RFID and UWB for identification and localization of goods[34] 27
Figure 20: Indoor camera to locate mobile devices, objects and persons[35] 28
Figure 21: Camera system in addition to indoor positioning[35] 28
Figure 22: Application of Beacons tag in indoor environments[31] 30
Figure 23: Application of Indoor navigation in airports[31] 31
Figure 24: Application of Indoor navigation in railways[31] 32

Figure 25: Application of Indoor navigation in offices[31] ... 32

Figure 26: Application of Indoor navigation in retail[31] ... 33

Figure 27: Objects represented by class ... 37

Figure 28: (Left) Standard convolutional layer with batch normalization and ReLU. (Right) Depthwise separable convolution with depthwise and pointwise layers followed by batch normalization and ReLU (figure and caption from Liu et al.) ... 39

Figure 29: Two Toyotas on the highway recognized with near -100% confidence using OpenCV, deep learning, and object detection. Their first result shows cars recognized and detected with near -100% confidence. ... 42

Figure 30: An airplane successfully detected with high confidence via Python, OpenCV, and deep learning. ... 42

Figure 31: A person riding a horse and two potted plants are successfully identified despite a lot of objects in the image via deep learning-based object detection. ... 43

Figure 32: Deep learning + OpenCV correctly detect a beer bottle from an in input image. ... 43

Figure 33: Block diagram of proposed system ... 47

Figure 34: Block Diagram of Proposed Architecture ... 51

Figure 35: Object detection steps using camera + OpenCV ... 58

Figure 36: Bluetooth LE version and symbol used in the proposed solution ... 59

Figure 37: Overlap of functionality between Bluetooth LE and RFID ... 59

Figure 38: Application deployed on android s4 mini ... 60

Figure 39: Application in white theme. ... 61

Figure 40: IBeacon Bluetooth LE feature enabled and in use. ... 62

Figure 41: Location [positioning] setting using audio command ... 63

Figure 42: NFC feature enabled and in use to detect RFID tags ... 64

Figure 43: Object detection steps ... 65

List of Acronym

API Application Programming Interface

APK Android Package Kit

ATM Automated Teller Machine

BLE Bluetooth Low Energy

Bluetooth LE Bluetooth Low Energy

DNN Deep Neural Networks

GLONASS Russia's version of GPS (Global Positioning System)

GPS Global Positioning System

HSV Hue Saturation Value

IBeacon Apple's Bluetooth

LBS Location Based Services

NFC Near Field Communication

OpenCV Open Source Computer Vision

ORIENTOMA Orientation application

RFID Radio frequency Identification

RGB Red Green Blue color

TTS Text To Speech

Webcam Web Camera

Wi-Fi IEEE 802.11x wireless fidelity

YUV Luminance (Y), blue-luminance (U), red-luminance (V) (SECAM and PAL color spaces)

Chapter 1

1. Introduction

1.1 Introduction

The need to assist visually impaired people in navigating and orienting both in the indoor and outdoor environment is increasing. However, there are numerous challenges that can limit this. Among them are, the need for privacy, user friendliness, and dependability to guide or orient the user has been broadly considered. In particular, RFID/NFC supported location based information aims at guaranteeing the handiest way to describe the localization information in real time, even off-line without the need of any connectivity. Indoor navigation and orientation is among those areas that need the usage of the RFID/NFC technology. The amount and size of location data to be processed could be large depending on the geographical area covered by a visually impaired people and the importance of the data under consideration.

The Orientoma project aims at developing a versatile hands free android app that seamlessly integrates Speech recognizer and TTS, Camera, RFID/NFC reader, Bluetooth LE and Map so that visually impaired people could manipulate the provided service alternatively by voice.

The camera feature with speech and TTS support guides the user in navigating indoor and outdoor, in tracking objects, identifying obstacles (e.g., trees, bicycles, pedestrians, traffic lights) and informing the user in real time leveraging on few specific voice commands and text-to-speech. The Near Field Communication (RFID/NFC) feature including IBeacon tag detection is used to seamlessly identify and exploit the locality information from RFID/NFC and IBeacon tags around the user. The Orientoma map with speech and TTS support, provides a location context information such as direction to a destination in navigating in an indoor and outdoor to a destination.

The most important aspect of a visual navigation system is object detection and identification in real world. Object detection and identification makes a navigation and orientation system practical in modern visually impaired navigation and orientation system since it enables a safe and reliable self-assistance for the user to reach his/her area of interest. Object/obstacle detection is an expensive and complex process in that it needs to read raw data from sensors and analyze the geometry of the nearby target object to identify it's type to inform the user the nature, interest or criticality of the object in a real time. The key point in object/obstacle detection and identification is categorizing the nature of the object in real time while minimizing the amount of data processed and the complexity in time.

The difficulty of communicating and interacting with the outside environment for the visually impaired people is the main reason for implementing a navigation and orientation system. In a real time object or obstacle detection, the natural visual scene is sampled spatially and temporally. The main aim of sampling images is to identify the nature of the object with less amount of data and processing time. It consists of a time sequence of images, called frames, displayed at certain frame rate. A lot of object detection techniques have been proposed to clearly and quickly identify the object and tell its nature and criticality to the user. A third party

1.2 Objective and Scope

The scope of this thesis is providing a versatile location information provider for visually impaired people. Among the means of getting location information this thesis primarily focuses on RFID tags, IBeacon tags, Maps and sensors like camera for Object detection. The problem addressed in this work is how to identify and tell the nature of an object from the input sensors, how to extract the valuable information from the sensors and present the meaning of the information in real time under many circumstances for the user to understand it clearly and take action accordingly.

The nature of the information provided by the different input sensors are different and requires a different mechanism to process it. For example, the information provided by an rfid reader could be a tag for a hyper link. Hence the application has to recognize and preprocess that hyper link to extract its content and present to the user in real time accordingly.

The performance of object detection with camera algorithms are limited by the resolution of the image sequences. The luminosity of the environment and the direction of light entering the camera also determines the reliability and detection time of an object/obstacle. As a result, a trade of between computational complexity and reliability based on the color quality of the video sequence or image must be played. The main objective of this thesis is to develop a navigation and orientation application for an android framework which adopts a certain algorithm with minimal complexity that serves as a reference for future works to enable the visually disabled people interact with the environment they live in.

The time required to activate the input sensors and switch among them according to the nature of the environment and start capturing/scanning an input could be more significant than the time it takes to detect/identify an object. Hence, it may degrade the performance in real-time object/obstacle detection and location information access. Most of existing navigation and orientation algorithms are based on exhaustive and iterative model or pattern search algorithms that require complex and lots of time to analyze the patterns. However, the intended purpose of this thesis is to develop an android framework for visual navigation and orientation with a reasonable computational complexity and cost but with a higher availability, reliability and dependability.

1.3 Main Contribution for outdoor-indoor localization

The main contribution of this thesis is the proposed object detection algorithm based for outdoor-indoor environments and the implementation of the proposed object/obstacle detection and location information providing service for an android framework. The framework is composed of four main components: Object/obstacle detection module, RFID tags detection module, IBeacon tags detection module, Map support unit and Text To Speech module for voice interaction. In places where providing a voice based interaction with the user is inconvenient (impossible), such as in noisy areas, the application optionally delivers a Bluetooth signal to a Bluetooth enabled smart watch or headphone (if available) so that the visually impaired user could be triggered with a vibration for critical notifications such as traffic lights. This android framework is briefly described in section 5 of this report.

1.3.1 Outdoor localization

This work of the thesis will have a significant contribution for outdoor localization in that it makes use of multiple sensors to collect location information in the outside environments.

1.3.2 Indoor localization

This work of thesis will have a significant contribution for the indoor localization in offices to identify meeting rooms, cafeteria, location-based advertisements, to access calendar in hospital with treatment schedules, and in general to access and navigate in public areas.

1.3.3 Key features

• Voice-based hands-free real-time interaction.

• Camera tracking for navigation and obstacles identification.

• Locality information using Bluetooth LE and RFID/NFC tags.

• Map support for indoor and outdoor navigation.

1.4 Thesis Outline

This work of thesis is organized in different chapters. The first three chapters reviews literature about visual navigation and orientation and provides relevant background information for this thesis, the need to have an android framework is illustrated. the next three chapters describes the designed system , together with its performance and the outcomes of a preliminary validation campaign.

Chapter 2:- introduces the basic concepts behind object detection and analyze some object detection methods. Moreover it describes relevant object detection layers and at the end it will discuss about the subject matter visual navigation and orientation in an android framework.

Chapter 3:- reflects the fundamental of face detection and recognition focusing on OpenCV libraries. After describing the concept of face detection in OpenCV it discusses the techniques and algorithms which are in use in different object detection standards.

Chapter 4:- Provide an overview of the designed object/obstacle identification algorithms and describes the limitation of other related works.

Chapter 5:- Presents the developed framework for the proposed algorithm detailing each components.

Chapter 6:- Presents the performance of the developed framework based on real-time test results

Chapter 7:- Concludes the thesis work and give recommendation for the future works.

Chapter 2

2. Navigation and Localization Systems

2.1 Introduction

The past decades has been witness to huge advances in technology and navigation systems. While the GPS is the undisputed king amongst satellite systems, it is by no means the only one. GLONASS, created and maintained by the Russian Federation, also provides users with accurate position fixes and the European Community is actively considering another system to be totally independent of the other two. The large-scale manufacture of microchips has enabled the production of low-cost equipment with capabilities that could only have been dreamed about a decade ago. This reduction in size and cost has also brought sophisticated navigation equipment within reach of small-boat owners[14].

Navigation and orientation systems are the most significant part of a modern transportation systems that influences the interaction of a person or an object in a given environment. The main objective of a navigation and orientation systems are to disseminate location information or location based information in a given environment during the movement of a person or an object. In this chapter, a sample of navigation methods are given.

2. 2 Visual Navigation and Orientation System

An autonomous visual system, such as a mobile robot, typically comprises some type of sensor system for sensing an environment. Preferably, the sensor system has the capability to detect obstacles within the path of the robot so that appropriate action may be taken. Although several approaches have been investigated to detect an obstacle in an environment, the method that has worked best in practice is a simple approach called object detection. As a result, the object detection algorithm has been adopted by many visual navigation and orientation systems.

To support inter working and coherence, it has been necessary to define standard methods of object detection to allow products from different manufacturers to communicate effectively. This has led to the development of a robust and most commonly used library for object detection, including the OpenCV libraries. The libraries are particularly relevant to image/video tracking and identification since video may be used in some video surveillance applications. By providing libraries, developers or end users are able to use the library to develop applications related to image processing and identification.

The Open Source Computer Vision Library (OpenCV) is a library for computer visions designed to analyze, process and understand the objects in videos and photos aiming to produce information. OpenCV is the most used libraries in robotics to detect, track and understand the surrounding world captured by image sensors. It is an open-source library leveraged for extension and update for whom interested to develop and add new functionalities.

The term face recognition or object detection is usually related with surveillance in videos. The usage of a robust face recognition or object detection algorithm on videos taken in real time through the usage of image sensors such as webcam is one the crucial. There are plenty of applications, and methods that shows us how to do face recognition in videos each with its own advantage and disadvantage.. For face detection a cascade classifier is used while face recognizer is used for face recognition. Fisher faces method is also used for face recognition because it is robust against large changes in illumination.

The most important and complex part of object/obstacle detection process is the time required to detect or recognize an object with a reasonable amount of confidence value from the raw camera input. Majority of real-time object detection algorithms have been proposed since the introduction in 1981[1]. Most of the object detection architecture have been implemented in libraries such as OpenCV so as to meet the power and size constraints. Most of them was aimed to have lower complexity with high performance.

2.3 Location based Services

We all remember when GPS's were the newest and best thing to happen for our driving. We could type in an address, and our personal assistant would verbally take us there. So convenient. Then, they got even better. We could tap a box for restaurants and get a long list of those that were nearby. Pretty soon after those devices were commonplace, design engineers would build GPS systems into cars. This all seems rather primitive to us now. And of course, if we were not in our cars, GPS systems were useless. It was only a matter of time before location based services technology would evolve, and ideas would form for even better location tools. Of course, these have been in the form of location based service apps for both Android and iOS devices. This was a natural outgrowth of the technology, because users want more than just directions and restaurants, and they want information on the go. Ideas to create a location based app for such things as ATM's, finding one's car in a parking lot, movies and concerts, fitness centers or jogging paths – the possibilities are endless – are generated every day. And developers are busy creating custom location based service apps for people with these great ideas.[30]

Location-based services are location-specific mobile services. They provide information or functions to smartphone users, depending on their location. Various methods can be used for indoor positioning within buildings to implement location-based services. A distinction is made here between reactive and proactive services. For reactive location-based services, the user searches for locations in the vicinity directly on his/her device, e.g. for ATMs at the airport. Proactive services "recognize" when a user enters a specific area and trigger an action – e.g. sending information or an offer to the user's smartphone. This is referred to as location-based marketing. An app is required for these services.[29]

Geofencing in an Interior Spaces [29]

The triggering of an action when taking a specific path is called geofencing (combination of geography and fencing). This can also be used inside buildings without GPS reception.[29]

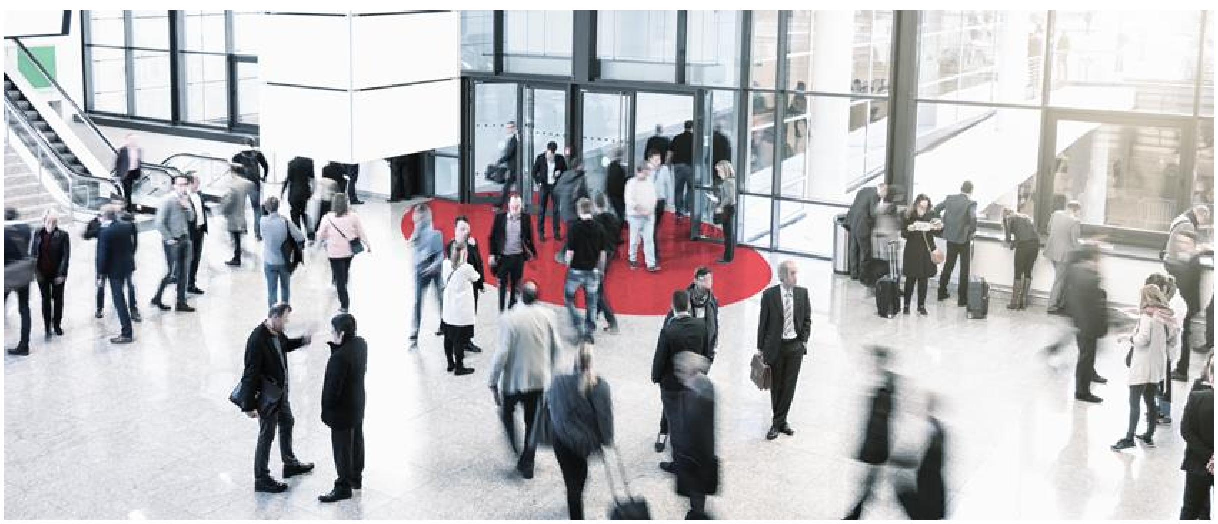

Figure 1: Geofencing in an Indoor environments[31]

Figure 2: Infsoft automation process of Geofencing[31]

Geofencing combines awareness of the user's current location with awareness of the user's proximity to locations that may be of interest. To mark a location of interest, it is necessary to specify its latitude and longitude. To adjust the proximity for the location, we add a radius. The latitude, longitude, and radius define a geofence, creating a circular area, or fence, around the location of interest. We can have multiple active geofences, with a limit of 100 per device user. For each geofence, we can ask Location Services to send we entrance and exit events, or we can specify a duration within the geofence area to wait, or dwell, before triggering an event. We can limit the duration of any geofence by specifying an expiration duration in milliseconds. After the geofence expires, Location Services automatically removes it. [30]

Geofencing is quite interesting for marketing in shopping centers, airports, train stations and at trade shows, among others. Users can be assigned anonymous "tags" that are based on their interests and behaviors. This could include characteristics such as age, gender, visited areas / stores and length of stay. This helps to provide users with information and offers that are only relevant for them.

A location-based service (LBS) is a software application: such as Dark Sky, City Hour, MapMyFitness, Glympse and others for mobile device that requires knowledge about where the mobile device is located. Location-based services can be query-based and provide the end user with useful information such as "Where is the nearest ATM?" or they can be push-based and deliver coupons or other marketing information to customers who are in a specific geographical area. Location-based services are location-specific mobile services. They provide information or functions to smartphone users, depending on their location. Various methods can be used for indoor positioning within buildings to implement location-based services. A distinction is made here between reactive and proactive services. For reactive location-based services, the user searches for locations in the vicinity directly on his/her device, e.g. for ATMs at the airport. Proactive services "recognize" when a user enters a specific area and trigger an action – e.g. sending information or an offer to the user's smartphone. This is referred to as location-based marketing. An app is required for these services [16].

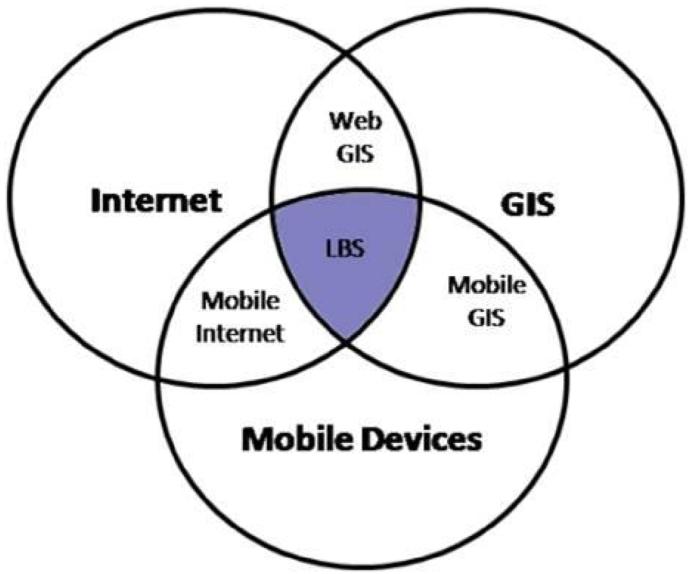

Figure 3: Basic components of LBS systems

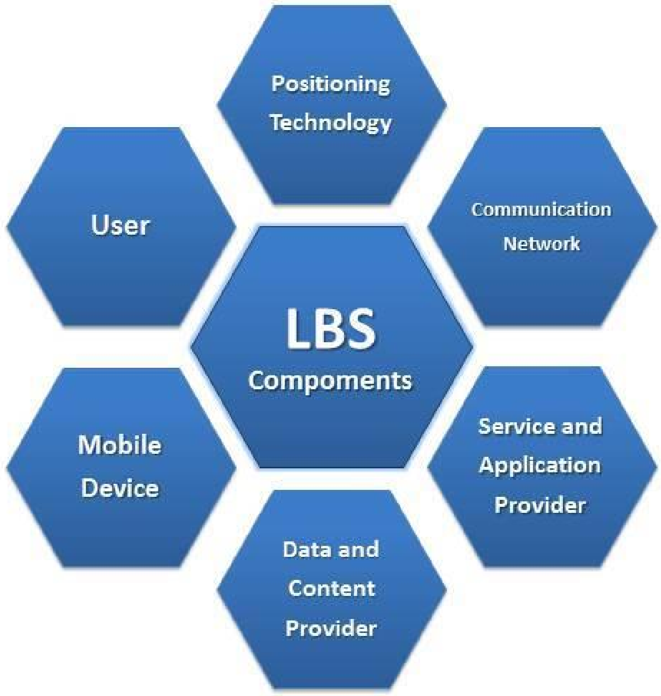

Figure 4: The different components of LBS

An LBS system as shown in figure 2.3.1 above has many components: the service provider's software application, a mobile network to transmit data and requests for service, a content provider to supply the end user with geo-specific information, a positioning component (see GPS) and the end user's mobile device. By law, location-based services must be permission-based. That means that the end user must opt-in to the service in order to use it. In most cases, this means installing the LBS application and accepting a request to allow the service to know the device's location. Although location-based services have been around since 2000, they have mostly been used in commerce with a subscription-based business model. The release of Apple's 3G iPhone and Google's LBS-enabled Android operating system, however, has allowed developers to introduce millions of consumers to LBS. According to the 2008 fourth-quarter report from Nielsen Mobile, a division of The Nielsen Company, location-based services account for 58 percent of the total downloaded application revenue for mobile phones in North America [16].

Location-based services (LBS) use real-time geo-data from a mobile device or smartphone to provide information, entertainment or security. Some services allow consumers to "check in" at restaurants, coffee shops, stores, concerts, and other places or events. Often, businesses offer a reward — prizes, coupons or discounts — to people who check in. Google Maps, Foursquare, GetGlue, Yelp and Facebook Places are among the more popular services. Location-based services use a smartphone's GPS technology to track a person's location, if that person has opted-in to allow the service to do that. After a smartphone user opts-in, the service can identify his or her location down to a street address without the need for manual data entry [16].

2.4 Uses of location-based services

Applications of location-based services[28]

Location-based services may be employed in a number of applications, including:

- Recommending social events in a city[23]
- Requesting the nearest business or service, such as an ATM, restaurant or a retail store
- Turn-by-turn navigation to any address
- Assistive healthcare systems[26]
- Locating people on a map displayed on the mobile phone
- Receiving alerts, such as notification of a sale on gas or warning of a traffic jam
- Location-based mobile advertising
- Asset recovery combined with active RF to find, for example, stolen assets in containers where GPS would not work
- Contextualizing learning and research
- Games where your location is part of the game play, for example your movements during your day make your avatar move in the game or your position unlocks content.
- Real-time Q&A revolving around restaurants, services, and other venues.
- Tracking a NASA lunar lander.[27]

For the carrier, location-based services provide added value by enabling services such as:

- Resource tracking with dynamic distribution. Taxis, service people, rental equipment, doctors, fleet scheduling.

- Resource tracking. Objects without privacy controls, using passive sensors or RF tags, such as packages and train boxcars.
- Finding someone or something. Person by skill (doctor), business directory, navigation, weather, traffic, room schedules, stolen phone, emergency calls.
- Proximity-based notification (push or pull). Targeted advertising, buddy list, common profile matching (dating).
- Proximity-based actuation (push or pull). Payment based upon proximity (EZ pass, toll watch), automatic airport check-in.
- Store locators. Using location-based intelligence, retail customers can quickly find the nearest store location.
- Proximity-based marketing. Local companies can push ads only to individuals within the same geographic location. Location-based mobile marketing delivers ads to potential customers within that city who might act on the information.
- Travel information. An LBS can deliver real-time information, such as traffic updates or weather reports, to the smartphone so the user can plan accordingly.
- Roadside assistance. In the event of a blown tire or accident, many roadside assistance companies provide an app that allows them to track your exact location without the need for giving directions.
- Mobile workforce management. For logistics-dependent companies that employ individuals out in the field or at multiple locations, an LBS allows employees to check in at a location using their mobile device.
- Fraud prevention. An LBS creates another level of security by matching a customer’s location through the smartphone to a credit card transaction. Tying the smartphone’s location to a credit card allows you to flag transactions made across several geographic locations over a short time. Mobile goods trigger an alarm signal as soon as they leave a defined area (anti-theft protection)
- Goods, machinery or vehicles send a message when they enter or leave an area, or a defined driving route is not observed (asset tracking)

- Security is notified when persons enter security-relevant or locked areas, e.g. in airports (person tracking)
- Customers, travelers or patients with specific characteristics are sent useful information to a smartphone as soon as they enter an area.

2.5 Outdoor Navigation

Outdoor Navigation involves in the navigation of an object or a person in outdoor environments such as road, outdoor markets, or airports and others. Outdoor Navigation is takes place mainly by outdoor enthusiasts such as bikers, sketchers, or outdoor workers such emergency workers, among which visually impaired people are part of it. Even if an outdoor navigation is different from a pedestrian navigation, in this thesis context an outdoor navigation also includes pedestrian navigation. Maps are very specifically street and road oriented. Outdoor navigation is where a person or an object takes a route from a specific start location to a specific destination location where in between a complex set of entities participate specially in the case of complex urban world. Advances in technology are allowing blind people and visually impaired people to literally interact in the outdoor environments due to the numerous technologies and algorithms developed to relieve such limitations. Technology available for navigation of the blind is not sufficiently accessible some devices rely heavily on infrastructural requirements.

The goal is to give an easy and free application that will exploit the onboard camera and location Map permit visually impaired individuals explore freely or independently in the outdoor environment. To achieve these goals, we focused the work on developing camera based and map-based application for the outdoor navigation that assists the blind and visually impaired individuals interact seamlessly in the society. Our main objective is to detect obstacles and provide appropriate voice based instructions or route information to the user through the interactive the embedded phone speaker in real time.

2.6 Indoor positioning[31]

Indoor navigation with automatic positioning is normally used as a client-based application. This means that the position is determined directly on the smartphone of the user and thus requires an app. The location is determined usually via Wi-Fi or beacons. A feedback channel is also available, for example, for sending push notifications. A server-based approach is also possible, but this comes with technical challenges. Indoor positioning systems (IPS) enable you to locate the position of objects, people and animals within buildings. Since GPS is unreliable in interior spaces because there is no visual contact with the GPS satellites, an IPS (indoor positioning system)must use other positioning methods. In most cases, this includes Wi-Fi or Bluetooth beacons in combination with the internal sensors of a smartphone. The first and most important step in the implementation of indoor positioning systems is the selection of the indoor positioning method and technology.[31]

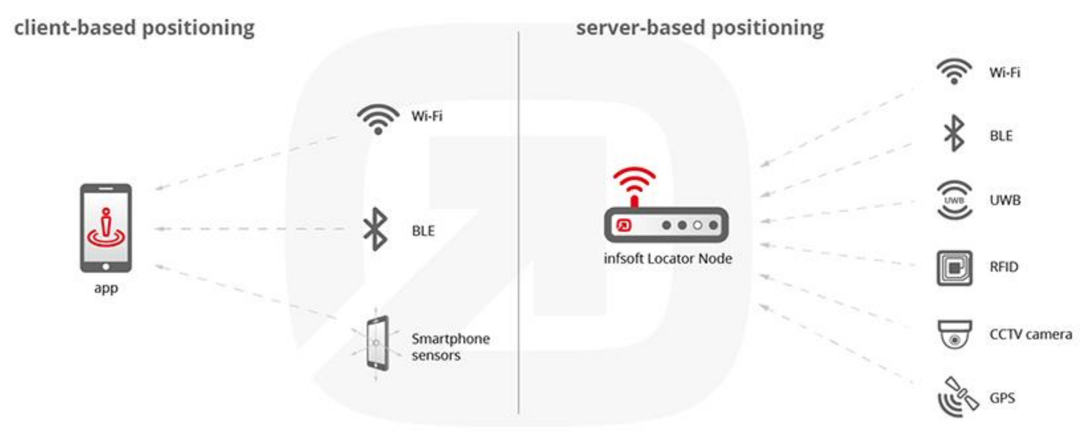

Figure 5: Client-based and server-based indoor positioning[31]

2.6.1 How Does Indoor Positioning Work?[31]

A) INDOOR POSITIONING VIA CLIENT[31]

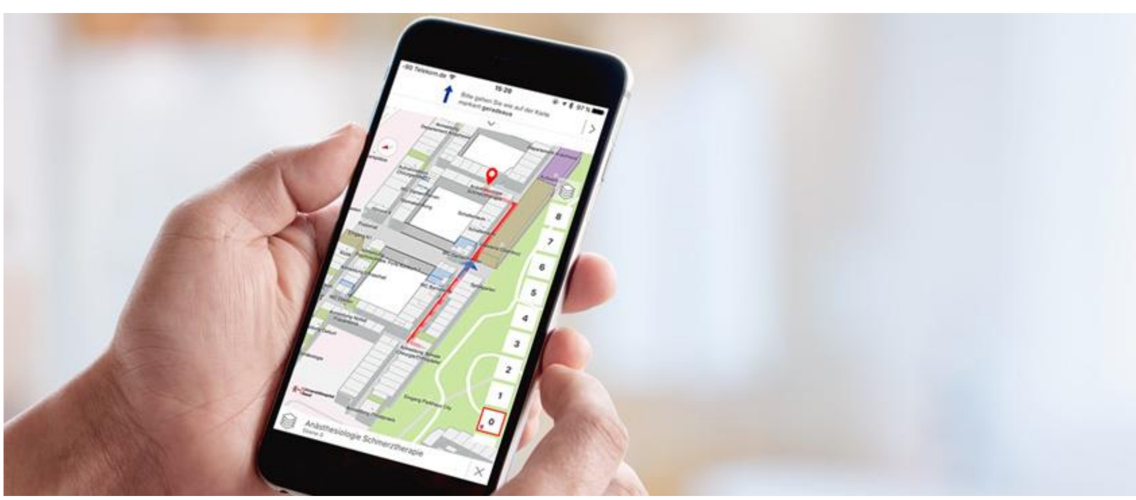

Figure 6: Indoor positioning via client[31]

Indoor positioning via client means that the position is determined using user's mobile device ("client" – usually a smartphone). A mobile application is required to implement this feature. It is most often used for indoor navigation projects and wherever you need to communicate with the user.

2.6.2 Indoor Positioning, Tracking and Indoor Navigation with Wi-Fi[31]

Indoor positioning and indoor navigation based on Wi-Fi are used in many projects. The reason: A wide range of different types of existing Wi-Fi hotspots can be used for this. In addition, positioning already functions when users have enabled Wi-Fi on their smartphones, no login is required. However, there are some disadvantages to indoor positioning with Wi-Fi: Accuracy is 5-15 meters - less than with Bluetooth Low Energy. Furthermore, iOS devices are excluded from client-based positioning.

2.6.3 How Does Indoor Positioning with Wi-Fi Work?[31]

Each Wi-Fi access point, whether customer hotspot, router or Internet-capable point of sale system, transmits specific data. Using a RSSI (Received Signal Strength Indication) and MAC address (Media Access Control), an app can calculate the current location of the end user device (client-based positioning). This requires a database with information about the locations with which this data can be compared. This method is called fingerprinting. It only functions with Android devices due to technical restrictions. iOS devices cannot be used for Wi-Fi indoor navigation.

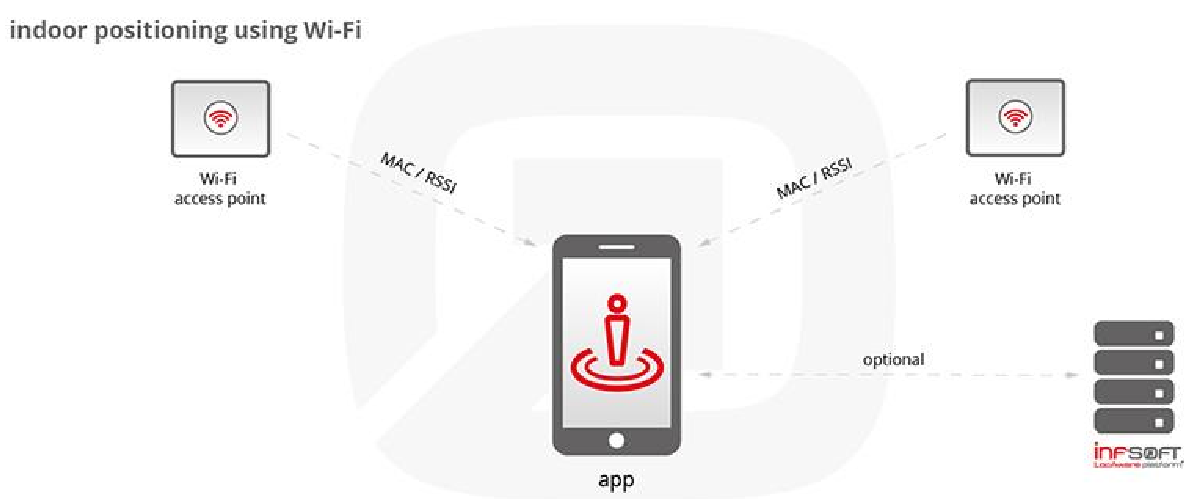

Figure 7: Indoor positioning using Wi-Fi[31]

Depending on whether the device is connected to a Wi-Fi network or not, localization is affected by various restrictions. If the Wi-Fi client is not connected, latency times of up to 5 minutes result from the irregular sending of "probe request" (scanning for available access points), which is controlled by the mobile operating system and cannot be influenced. In addition, due to randomly generated, changing MAC addresses ("MAC randomization"), the smart device or the actual number of devices in an area cannot always be determined unambiguously. If the smartphone is connected to a Wi-Fi network, latency times are significantly lower as long as traffic is generated over the network. In addition, there is no MAC randomization, which means that the device can be uniquely assigned. The position accuracy is not affected by the connection to an access point.[31]

Restrictions for server-based detection via Wi-Fi

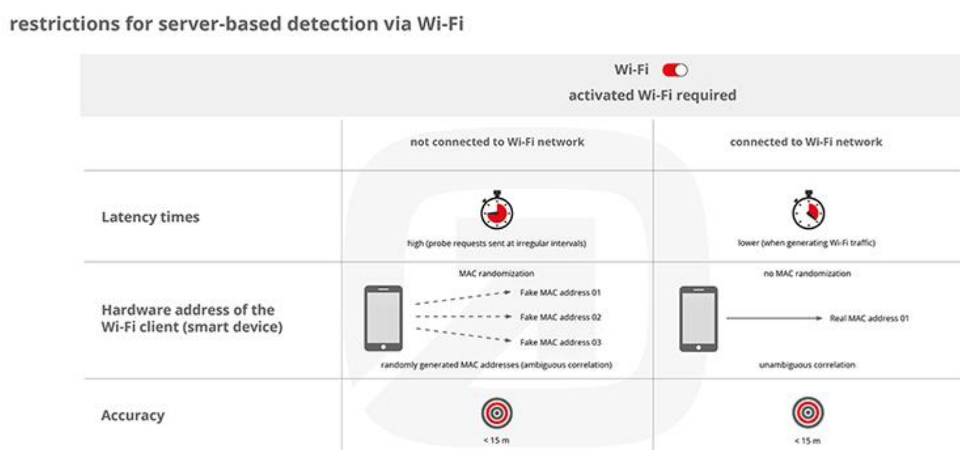

Figure 8: Restrictions for server-based detection via Wi-Fi[31]

The accuracy of Wi-Fi for indoor positioning is typically 5-15 meters because access points are usually used whose position has been optimized for data communication. This precision depends on the shielding through walls, ceilings and people, as well as the number of access points. The use of smartphone sensors can improve the results and the determination of the floor level is also possible.

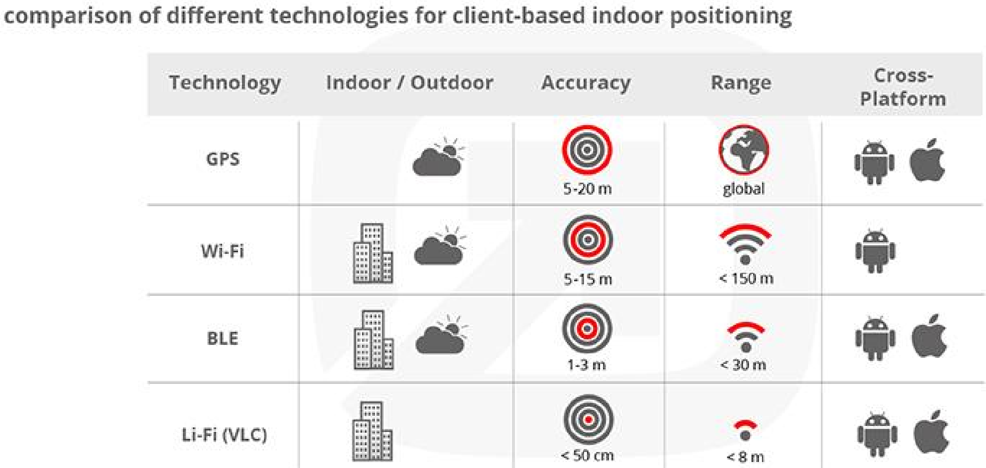

Figure 9: Comparison of different technologies for client-based indoor positioning[31]

2.6.4 Indoor Positioning, Tracking and Indoor Navigation with Beacons[32]

Indoor navigation with beacons (the most common types are the iBeacon and Eddystone) offer decisive advantages for projects that are dependent on high accuracy and want to include Apple devices. Beacons are typically transmitters in client-based methods. For example, they enable indoor navigation for airline passengers using the app - cross-platform and with an accuracy of up to 1 meter. The server-based beacon tracking of persons or goods is only possible with third-party components (e.g. infsoft Locator Node, Cisco, Aruba).

Figure 10: Beacons for client –and server-based localization[32]

When we use beacons for indoor navigation, temporary installations, e.g. for exhibitions and in unusual locations are also possible. Mounting the beacons is simple and extremely flexible. The housing is available in different colors so that devices can be installed discreetly. The Bluetooth Low Energy (BLE) technology enables battery operation over two to eight years without having to access an external power supply. Beacons can also be connected to the power supply or use the power supply from the lighting. Paper-thin beacons with a printed battery are ideal for very discreet and space-saving installations, for example, on ISO cards or on advertising materials. They have a battery life of three to four days. The Beacon Management tool integrated into infsoft systems makes beacon management and monitoring easy.[32]

2.6.5 Functionality of Bluetooth Low Energy Beacons[32]

Bluetooth beacons are small radio transmitters that send out signals in a radius of 10-30 meters (interior spaces). The advantages of beacons are obvious: They are cost-effective (three to thirty euros), can be installed with minimal effort, determine a position accurately up to 1 meter and are supported by many operating systems and devices. The new BLE (Bluetooth Low Energy) standard is also very energy efficient. Beacons can be used for both client-based as well as server-based applications. With beacons it is possible to detect the current floor.

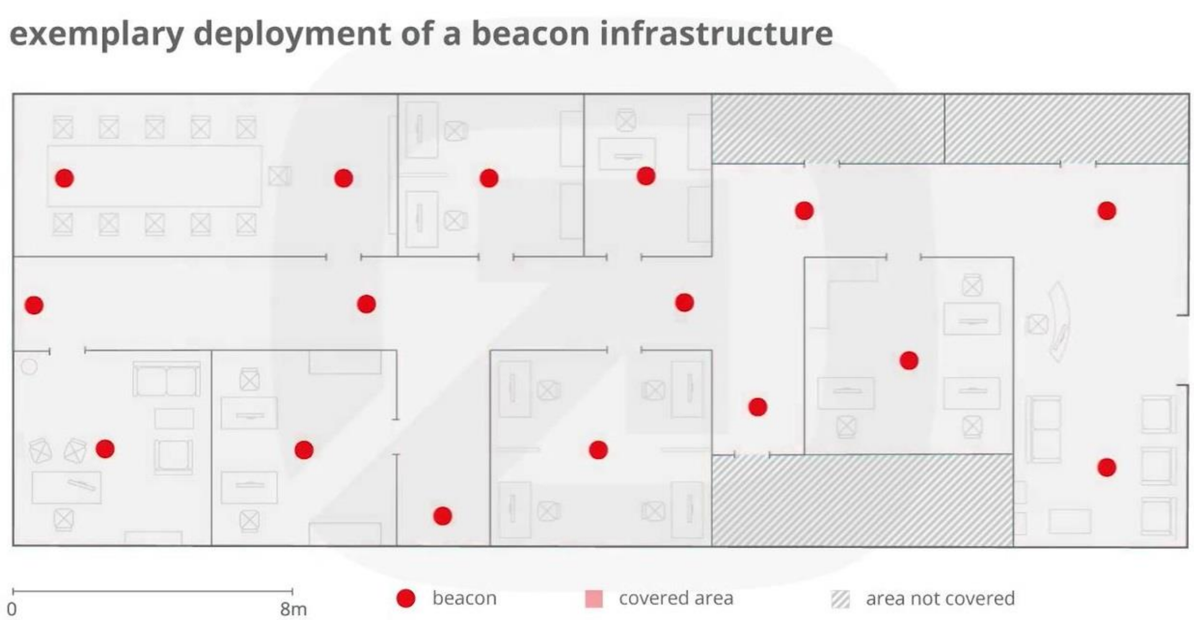

Figure 11: Exemplary deployment of a beacon infrastructure[32]

2.6.6 Client-based indoor positioning using BLE[32]

If the positioning data should be directly displayed on a mobile device (smartphone), the smartphone receives Bluetooth signals from beacons installed in the building and uses the signal strength measurement for localization. The device can also transfer the data to the infsoft LocAware platform®, where it is intelligently processed. The accuracy of a client-based procedure is 1-3 meters, the range up to 30 meters.

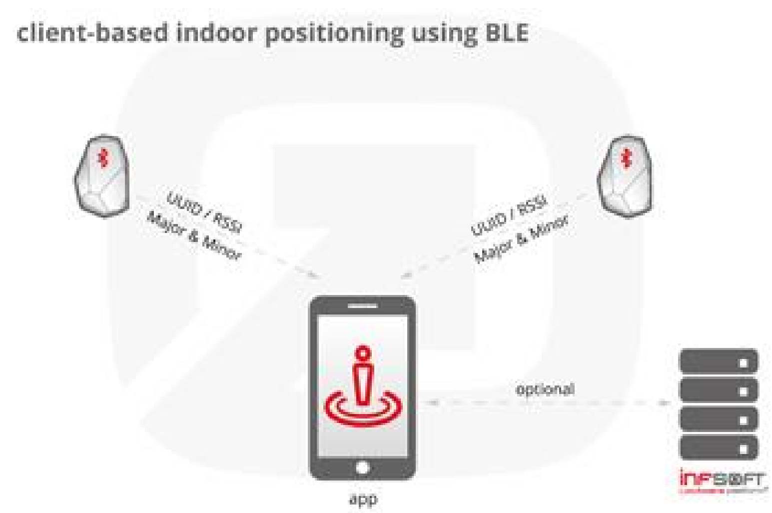

Figure 12: Client-based indoor position using BLE[32]

2.6.7 Server-based indoor positioning using BLE[32]

In the case of a server-based positioning, the beacon attached to the asset to be tracked sends BLE signals to infsoft Locator Nodes that are installed in the building. The Locator Node processes the provided data and sends it via Wi-Fi, Ethernet or UMTS to the infsoft LocAware platform®. Here, the position is displayed on a map and motion statistics can be retrieved. The accuracy of a server-based procedure is less than 8 meters, the range up to 75 meters.

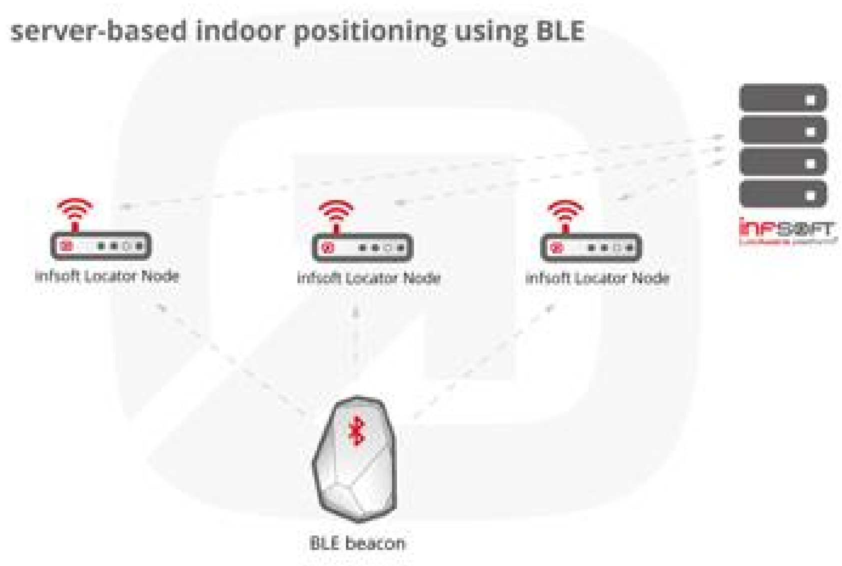

Figure 13: Server-based indoor positioning using BLE.[32]

Bluetooth Low Energy beacons are available from numerous vendors and come in various shapes and sizes for different use cases. Using the industry standard Bluetooth Smart, the solutions of infsoft are compatible with beacons of all manufacturers.

2.6.8 Parameterization and Interferences of Beacons[32]

Beacons have specific indicators that can be used for allocation, differentiation and parameterization. For iBeacons, these include UUID (association to a superordinate group, e.g. "Frankfurt Airport"), Major (association to a specific group, e.g. "Terminal 1"), Minor (identification of several beacons, e.g. "Beacon #10"), TX Power (signal strength) and transmission interval. Eddystone beacons have several "frame types" that differentiate them from the iBeacon: Eddystone UUID: (16-digit ID that indicates the group association and identity of an individual beacon), Eddystone URL (freely selectable URL) and Eddystone TLM (current status of the beacon).

Bluetooth beacons normally do not affect other radio networks (**interference**) and they also do not interfere with medical devices. However, if you install beacons in a space with lots of Wi-Fi signals (for example, at a trade show), then interference can occur because BLE and Wi-Fi share the same frequency range (2.4 GHz). The problem can be easily avoided here by not using channels 2, 3, 4, 13 and 14 when configuring the Wi-Fi and using 1, 6, 7, 8, 9, 10, 11 and 12 instead. Bluetooth uses the remaining available channels to capacity in a uniform manner (frequency hopping). Advertising channels that are used for positioning are marked in red in the graphic. The blue-colored channels are reserved for additional functions such as a temperature sensor.

There are attenuations in the signal dispersion within buildings for BLE beacons. Corresponding characteristics are taken into account when installing and during parameterization.

Low attenuation properties: Wood, synthetic materials, glass

Medium attenuation properties: Brick, marble

High attenuation properties: Plaster, concrete, coated glass surfaces (bulletproof versions, etc.)
Extremely high attenuation properties: Metal, water (this includes people and groups of people)

2.6.9 Use Cases for Indoor Positioning with Beacons[32]

There are multiple application scenarios for indoor positioning with Bluetooth Low Energy beacons. Here are some examples:

1. Theft protection

   BLE beacons can be attached to valuable or expensive objects. When a beacon leaves a defined area (geofencing), an alarm is triggered or security receives a notification.

   Software solution used: *Indoor Tracking*, *Indoor Location Based Services*

   Sensors used: *Bluetooth Low Energy (beacons)*

   Example of use: *Investigation of Bicycle Theft*

2. Evacuation

   Employees and visitors in large buildings can be equipped with a beacon (ID card format). In case of an emergency it is then possible to determine the number and position of people still in the building and to evacuate them.

   Software solution used: *Indoor Tracking*

   Sensors used: *Bluetooth Low Energy (beacons)*

   Example of use: *Evacuation of Employees and Visitors*

3. Inventory management

   Each object that should be tracked is equipped with a tiny beacon sticker. The position and characteristics of the objects (department, purchasing date, value, etc.) can be displayed in a web interface at any time.

   Software solution used: *Indoor Tracking*

   Sensors used: *Bluetooth Low Energy (beacons)*

   Example of use*: Inventory of Economic Assets on Company Premises*

4. Indoor Navigation at airports / railway stations

   Beacons enable indoor navigation for airline or rail passengers. Using an app, they can easily find the right gate/track or important points of interest such as cash machines, rest rooms and ticket counters. This works with Android and iOS devices. *Software solution used: Indoor Navigation*

   Sensors used: *Bluetooth Low Energy (beacons)*

2.6.10 Indoor Positioning with Ultra-wideband[33]

Indoor tracking in industrial environments often requires very precise localization. For this purpose we offer solutions based on the short-range radio technology Ultra-wideband (UWB). It does not work with consumer standards such as Wi-Fi and Bluetooth Low Energy, but offers an accuracy better than 30 cm.

Figure 14: Indoor positioning with Ultra-wideband[33]

Indoor positioning with Ultra-wideband has some significant advantages: The accuracy is 10-30 cm, which is considerably better than when working with beacons (1-3 meters) or Wi-Fi (5-15 meters). Latency time is very low (position request up to 100 times/second). Height differences can be measured accurately. However, the technique is a special solution which requires appropriate components and thus is mostly suitable for special industry applications.

2.6.11 Functionality of Indoor Localization with Ultra-wideband[33]

Ultra-wideband is a short-range radio technology which can be used for indoor positioning. In contrast to Bluetooth Low Energy and Wi-Fi, positioning is done with transit time methodology (Time of Flight, ToF) instead of the measurement of signal strengths (Receive Signal Strength Indicator, RSSI). This method measures the running time of light between an object and several receivers (anchors – infsoft Locator Nodes). For the exact localization of an object at least 3 receivers are necessary (trilateration).

The object or the person that should be tracked (asset) is equipped with a small tag (infsoft Locator Tag) which runs on battery power or draws its power via a forklift, for example. It sends data (ID, ToF, timestamp) to the infsoft Locator Nodes. The nodes have a fixed position in the infrastructure and can use the running time of light to calculate the distance of the asset. Combining the data of 3 Locator Nodes or more results in a positioning accuracy of 10-30 cm.

2.6.12 Setting up an UWB-Based System: Two Approaches

A) Server-based positioning of Ultra-wideband[33]

Visualization of positioning data in the backend, no display on a client (smartphone).

In order to localize assets, the infsoft Locator Tag measures the distance to several infsoft Locator Nodes and sends the data back to them. infsoft Locator Nodes are modularly structured and have an integrated UWB module, which lets them act as a receiver. The infsoft Locator Node processes the data it receives and sends them to the infsoft LocAware platform® via Wi-Fi, Ethernet or UMTS. Here, the position is displayed for example on a map.

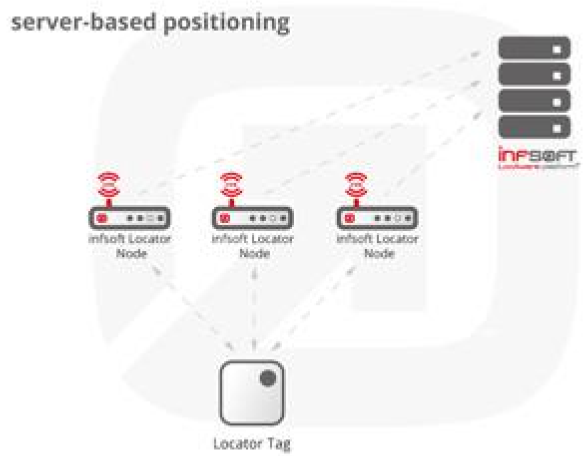

Figure 15: Server-based positioning of Ultra-wideband[33]

B) Client-based positioning of Ultra-wideband[33]

Visualization of positioning data on a client (smartphone) and if necessary transmission to the backend.

If the positioning data should be immediately displayed on a mobile device (smartphone), the infsoft Locator Tags can communicate with the smartphone via Bluetooth or USB interface directly. Time of Flight is measured, just like explained above, but the Locator Tag exchanges the data with the device via Bluetooth or USB port. The data can be sent to the infsoft LocAware platform® from the device, if necessary.

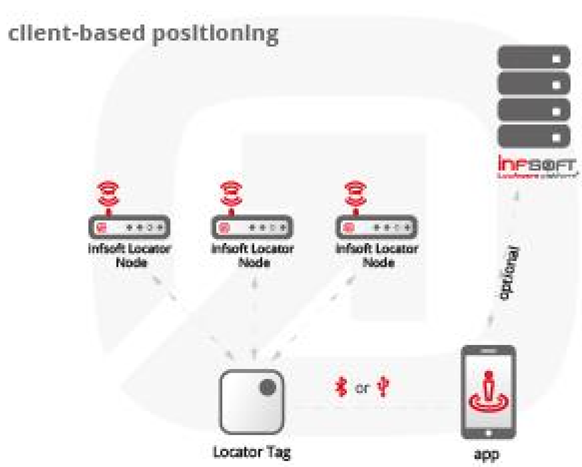

Figure 16: Client-based positioning of Ultra-wideband[33]

Technical Features of Ultra-wideband[33]

- Usage of extremely wide frequency bands with a bandwidth of at least 500 MHz
- Almost no interferences
- Frequency bandwidth 3,1 – 10,6 GHz
- Transmission power 0,5 mW / −41,3 dBm/MHz
- Reach 10 – 150 m (depends on the use case)
- Data rate 110 kbit/s – 6.8 mbit/s

Use Cases for Ultra-wideband Indoor Positioning

Indoor positioning with Ultra-wideband can be used in various application scenarios. Here are some examples:

1. Asset tracking

   Pallets in extensive warehouse systems can be recorded and their location, service life and movement history can be retrieved. With pallet tracking, it is possible to prevent wrong deliveries or incorrect inventories.

   Software solution used: Indoor Tracking

   Sensors used: *Bluetooth Low Energy (beacons), Ultra-wideband, RFID*

2.6.13 Indoor Localization with RFID[34]

RFID (radio-frequency identification), which uses radio waves to wirelessly transmit the identity (e.g. serial number) and other characteristics of an object, is an emerging positioning technology that allows for mobility tracking of objects or people. As it offers a limited range of less than a meter, RFID is not suitable for area-wide positioning, but rather for a selective object identification. It is cost-effective, easy of maintenance and provides both identification and location. This makes localization via RFID particularly suitable for tracking solutions in industrial environments (e.g. asset management).

Figure 17: Application of RFID for tracking and asset management[34]

How Does Indoor Tracking with RFID Work?[34]

An RFID indoor positioning system typically contains of transponders (attached to objects/people) and a reader. In a passive RFID system, the Locator Node functions as a power source and transfers radio-frequency energy to the transponder at short distance (remote-coupling). ID and data from the transponder are then captured by the Locator Node and forwarded to the infsoft LocAware platform®, where the data is processed.

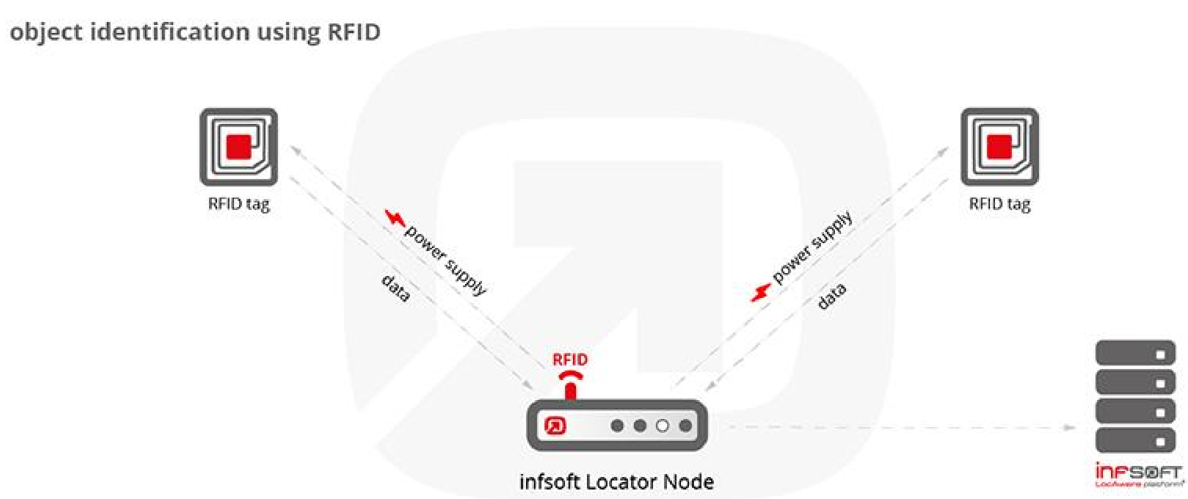

Figure 18: Object identification using RFID.[34]

In contrast to area-wide localization technologies, e.g. via Wi-Fi, Bluetooth Low Energy (BLE) beacons or Ultra-wideband (UWB), RFID tracking only enables spot localization due to its highly limited range of less than a meter. However, if an application requires more than spot detection, RFID can be combined with real-time positioning technologies such as BLE, Wi-Fi or Ultra-wideband. For example, if you provide a forklift with an infsoft Locator Node which sensors respond to RFID and UWB, you can make a connection between the location data of the forklift and the detection times of RFID-tagged goods.[34]

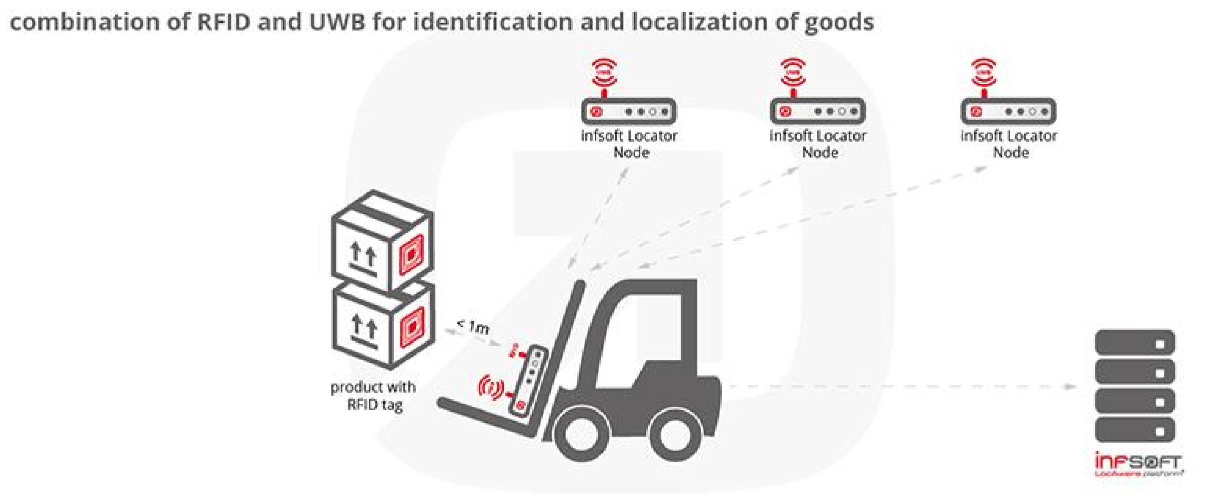

Figure 19: Combination of RFID and UWB for identification and localization of goods[34]

2.6.14 Camera Systems in Addition to Indoor Positioning[35]

Technologies for indoor positioning can be augmented with an imaging system, e.g. dome cameras. Especially in logistics and the security sector, the connection of an external camera system and thus the establishment of a correlation between camera image and position data can provide significant advantages. Not only does it enable to locate mobile devices, objects and persons, but also to allocate optical features (e.g., clothes, colors, or numbers as identification criteria). Furthermore, a corresponding recognition of difference images offers benefits when supplementing positioning technologies such as Bluetooth Low Energy do not provide sufficient accuracy on their own.[35]

Figure 20: Indoor camera to locate mobile devices, objects and persons[35]

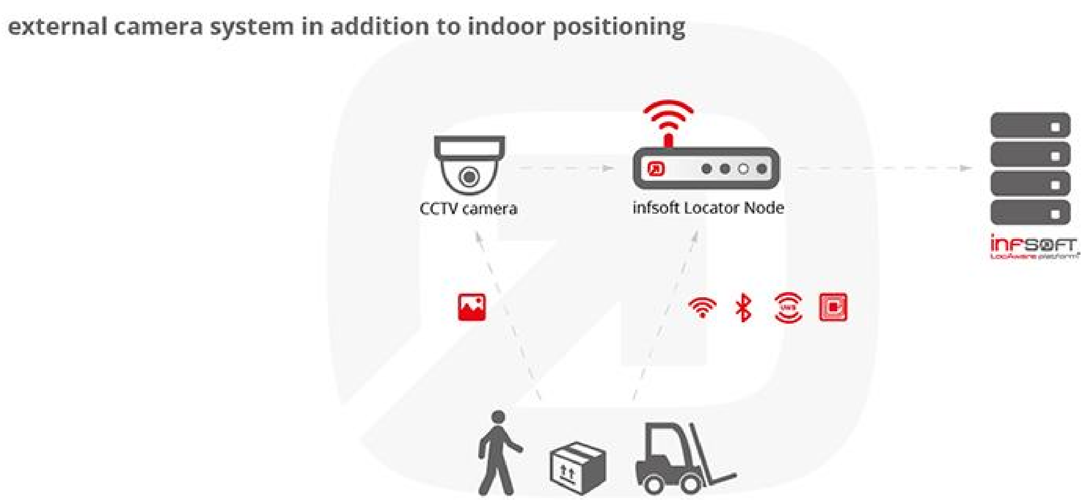

Figure 21: Camera system in addition to indoor positioning[35]

B) INDOOR POSITIONING VIA SERVER

2.7 Indoor navigation with Wi-Fi [31]

Indoor navigation with Wi-Fi has an accuracy of 5-15 meters. The various signal strengths of several Wi-Fi access points are evaluated for this purpose. Precise positioning within the building, even over multiple floors is made possible through specific shielding characteristics. The advantage of Wi-Fi is that the available infrastructure can be used (e.g. customer hotspots, Wi-Fi-capable point of sale systems, routers) - the user only has to activate Wi-Fi on his/her smartphone, a connection is not required. However, client-based positioning via Wi-Fi is not supported by Apple devices - so if we only use Wi-Fi as a positioning technology we will be excluding a large portion of users. Beacons are the alternative.

The popularity of indoor navigation has significantly increased during the last decade. Nowadays, many studies aim to develop new indoor navigation systems and improve the accuracy of the already existing ones. Unfortunately, no definitive method for indoor navigation has been approved yet, which is why this type of systems are not as accessible as those used for outdoor navigation. The purpose of this thesis is to introduce the concept of indoor navigation and demonstrate how common electronic devices can be used to create a simple indoor navigation system. To do this, a mobile Android application that helps users navigate inside a building was developed. Signal readings and other information broadcasted by Bluetooth Low Energy (BLE) devices, known as beacons, were used as reference for this application. In addition to this, adapters and sensors integrated into the mobile device were used to track the user's movement. The application is also capable of providing information to the user related to their location using additional data attachments associated to the beacons.[22]

An indoor positioning system (IPS) is a system to locate objects or people inside a building using radio waves, magnetic fields, acoustic signals, or other sensory information collected by mobile devices.[21]

Indoor navigation takes place within blocks of a building. Because GPS reception is normally non-existent inside buildings, other positioning technologies are used here when automatic positioning is desired. Wi-Fi or beacons (Bluetooth Low Energy, BLE) are often used in this case to create a so-called "indoor GPS". Contrary to GPS, however, they also enable us to determine the actual floor level. Most applications require an "indoor routing" functionality that guides people precisely through a building using an indoor navigation app and in this way, automatically determines their position - very similar to the navigation systems that we use in our cars. A typical application is turn-by-turn navigation in an app (displaying directions on a digital map) used for train stations, airports, shopping centers and museums. This kind of application can also include many other useful services. However, indoor navigation is not possible without automatic positioning - for example, when a digital building map is integrated into a website or in a digital signage system (multi-touch kiosk/interactive terminal). In this case, no location hardware is required (Wi-Fi, beacons).

2.8 Indoor Navigation with Beacons[31]

Indoor navigation using beacons is quite widespread because Bluetooth transmitters function across platforms and have an accuracy of 1-3 meters. The most well-known types are called iBeacon (from Apple) and Eddystone (from Google). Both operate using the BLE standard (Bluetooth Low Energy) and thus are very energy efficient. Numerous hardware manufacturers market these small devices. They have a battery life of 1-8 years and cost between EUR 3 and 30 depending on the functional scope and battery type. A beacon management platform from infsoft simplifies the maintenance and replacement of the transmitters. A beacon should be placed every 7-10 meters depending on the desired accuracy. Beacons are the most popular hardware for indoor positioning due to their high level of flexibility and accuracy. Read more about beacons and indoor positioning here.

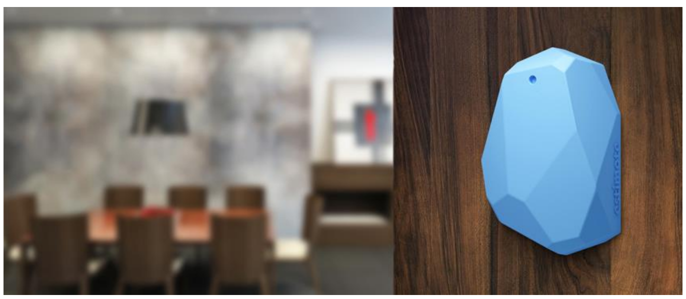

Figure 22: Application of Beacons tag in indoor environments[31]

2.9 Indoor Navigation with Ultra-wideband[31]

Indoor navigation with Ultra-wideband has some significant advantages in industrial environments: The accuracy is 10-30 cm, which is considerably better than when working with beacons (1-3 meters) or Wi-Fi (5-15 meters). Latency time is very low (position request up to 100 times/second). Height differences can be measured accurately.

For the client-based installation info soft Locator Tags are required. They transmit their position directly to the smartphone - either via a USB dongle which is directly plugged into the smartphone or via Bluetooth. On the smartphone, there is an app installed which contains a digital map determining and displaying the current position.

However, the technique is a special solution which requires appropriate components and thus is mostly suitable for special industry applications. Compared to Wi-Fi or BLE, UWB is more cost-intensive when installed on the same area. One possible use case are floor conveyors whose drivers should receive precise turn-by-turn directions. Because of their high speed, latency must be kept to a minimum.

2.10 Application of indoor navigation[31]

Indoor navigation is always called for when people want to navigate through complex buildings. However, there are also some interesting visitor services available along with the navigation function.

- **In airports,** for example, this can include the display of current wait times, the display of detailed information on important points of interest (POIs), the transmission of personalized coupons and/or messages, and support for the entire intermodal travel chain.

Figure 23: Application of Indoor navigation in airports[31]

- **Train stations** can inform visitors regarding current delays and track changes.

Figure 24: Application of Indoor navigation in railways[31]

- **In offices**, the booking of meeting rooms, the cafeteria layout, a colleague finder or invitation management can be useful added functions.

Figure 25: Application of Indoor navigation in offices[31]

In the retail industry, a car finder as well as location-based advertisements and coupons (location based marketing) are useful features.

Figure 26: Application of Indoor navigation in retail[31]

Hospitals can integrate the booking of additional services, a calendar with treatment schedules and barrier-free routing into their patient app.

- Trade shows score points with their visitors with information on speakers, exhibits and programs, as well as topic-related tours of the premises.
- Navigation in industrial areas for the drivers of floor conveyors and facility management teams.

Chapter 3

3. Image Recognition and Object Detection

3.1 A Brief History of Image Recognition and Object Detection

The Viola–Jones object detection framework is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones[36][37]. They show demonstrated faces being detected in real time from a webcam video input and it was the most stunning demonstration of computer vision and its potential at the time. Soon, it was implemented in OpenCV and their name became synonymous with face detection. Every few years a new concept and algorithm comes along that overtake the preceding ones. In object detection, that idea came in 2005 with a paper by Navneet Dalal and Bill Triggs. Their feature descriptor, histograms of Oriented Gradients (HOG), significantly outperformed existing algorithms in pedestrian detection.

The problem to be solved in Viola-Jones algorithm was detection of faces in an image. A human can do this easily, but a computer needs precise instructions and constraints. To make the task more manageable, Viola–Jones requires full view frontal upright faces. Thus in order to be detected, the entire face must point towards the camera and should not be tilted to either side. While it seems these constraints could diminish the algorithm's utility somewhat, because the detection step is most often followed by a recognition step, in practice these limits on pose are quite acceptable.[38]

3.2 Components of the Viola-Jones framework

Feature types and evaluation

The characteristics of Viola–Jones algorithm which make it a good detection algorithm are:

- Robust – very high detection rate (true-positive rate) & very low false-positive rate always.
- Real time – For practical applications at least 2 frames per second must be processed.
- Face detection only (not recognition) - The goal is to distinguish faces from non-faces (detection is the first step in the recognition process).

Since then many algorithms have been proposed and implemented that improve the recognition rate and other parameters such as performance from the preceding algorithms. Deep Learning is a recently dominating idea that supersedes its predecessors. Deep Learning algorithms had been around for a long time but it became mainstream in computer vision with its resounding success at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) of 2012. In that competition, an algorithm based on deep learning by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton shook the computer vision world with an astounding 85% accuracy which is 11% better than the algorithm that won the second place. On the next year, all winning entries were based on Deep Learning and in 2015 multiple Convolutional Neural Network (CNN) based algorithms surpassed the human recognition rate of 95% [1].

3.3 Image recognition and object detection using OpenCV

Object detection is nothing but a process of comparing two or more objects and tell if their geometry is identical or not. And it is one of an important step in object detection process that will determine whether the pattern of an input image matches with an existing trained model image. The technical term for this process is ’pattern recognition’ but is often referred to as ’object detection’ or 'obstacle identification'. In this portion steps and different layers of object detection is addressed. There are different representations of image in a computer system such as bitmap, vector, rastered form. The most common form to represent digital images or other forms of graphics that are rich in feature in a computer system is bitmap or raster images. The selection of an appropriate image representation method in an open source library is important because most object detection algorithms make use of a model train images represented in a form library. To recognize for example a human face in videos, we need some list of model faces of a human to train the model or learn a face recognizer. Images from the gender classification class in OpenCV library are reused to test experiment also.

3.4 Object detection in a digital image

3.4.1 Object detection

Object detection is a computer technology related to a computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. It is a process of comparing two or more objects and tell if their geometry is identical or not. And it is one of those important steps in a face detection algorithm that will determine whether the pattern of an input image matches with an existing sample image. The technical term for this process is “pattern recognition”, but is often referred to as “object detection” or “obstacle identification”. object detection algorithms deal with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance. Well-researched domains of object detection include face detection and pedestrian detection.

Most object detection algorithms make use of a model train images represented in some format on a computer system. To recognize for example a human face in videos, we need some list of trained model faces of a human for a face recognizer. The selection of an appropriate image representation method in an open source library is important because the detail of information contained in an image could degrade the processing time and hence the performance of a detection algorithm.

Every object class has its own special features that helps in classifying the class – for example all circles are round. Object class detection uses these special features. For example, when looking for circles, objects that are at a particular distance from a point (i.e. the center) are sought. Similarly, when looking for squares, objects that are perpendicular at corners and have equal side lengths are needed. A similar approach is used for face identification where eyes, nose, and lips can be found and features like skin color and distance between eyes can be found.

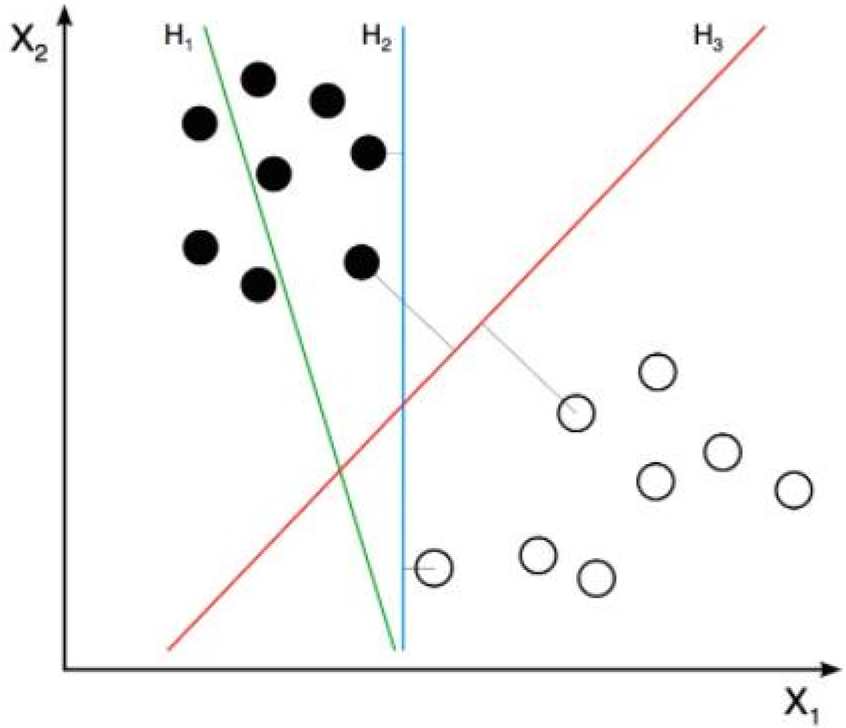

Figure 27: Objects represented by class

The 2D points on figure 27 represent two classes (let's say for example cats and background). The two classes are represented by two different kinds of dots. All black dots belong to one class and the white dots belong to the other class. During training, we provide the algorithm with many examples from the two classes.

Different learning algorithms figure out how to separate these two classes in different ways. In the figure above, H1, H2, and H3 are three lines in this 2D space. H1 does not separate the two classes and is therefore not a good classifier. H2 and H3 both separate the two classes, but intuitively it feels like H3 is a better classifier than H2 because H3 appears to separate the two classes more cleanly.

3.5 Object tracking in a digital video

Digital video is a discrete representation of images sampled in both spatial and temporal domain. A pixel has one or more component according to its color space representation. The common types of color spaces used in an image signals are RGB and YCrCb(YUV). RGB color space describes the relative proportions of Red, Blue, and Green in a pixel and they are measured in a range of 0-255 (8bit) integer value.

OpenCV is a robust open source library and most used in robotics to detect, track and understand the surrounding world in real time and detect and recognize objects in a digital video captured by a webcam or image sensors. For example, ball detection / tracking using OpenCV in a digital video involves to go through the following steps:

1. Capture a video and take a fixed amount of frame per second from the video.
2. Converts the input image from RGB color space to HSV space. HSV (hue saturation value) space gives us better results while doing color based segmentation. Then it detaches the input image into its three component color spaces i.e: H, S, V each of which are a one dimensional image or intensity image. The three color components are used to get intensity values in the image to get a binary image.

   i.e Let's say we took the H intensity image. If our ball is red color then in this image we will find that the values of the pixel where the ball is present, lies in a specific range. So we define a condition for every pixel. If pixel is greater than threshold minimum and pixel is less than threshold then max pixel of output image would be one else it would be zero. At this point there are three binary images only black and only white color. Which has the region of ball as 1's and everything else which has the intensity values greater or less than threshold .The pixels that do not pass this conditions will be zero.

3. Combine all the above three binary images. All the pixels that are white in the three images will be white in the output of this step. So there will be regions which will have 1's but with lower areas and of random shapes.
4. Use Houghs transform on the output of the last operation to find the regions which are circular in shape. Finally, we draw the marker on the detected circles as well as display the center and radius of the circles.

3.6 Object detection using Single Shot Detectors and MobileNets

When used together these methods could be used for super-fast, real-time object detection on resource constrained devices such a smartphones. Then, the OpenCV's deep neural network module could be used to load a pre-trained object detection network model and apply the pattern matching process. This will enable us to pass input images through the network and obtain the output boundary coordinates of each object in an image. Finally, we'll look at the results of applying the MobileNet Single Shot

Detector to example input images. When it comes to deep learning based object detection there are three primary object detection methods that we'll likely encounter:

1. Faster R-CNNs (Girshick et al., 2015)
2. You Only Look Once (YOLO) (Redmon and Farhadi, 2015)
3. Single Shot Detectors (SSDs) (Liu et al., 2015)

Faster R-CNNs is the method for object detection using deep learning; however, the technique can be challenging to train and also hard to implement. Furthermore, even with the "faster" implementation R-CNNs where "R" stands for "Region Proposal" the algorithm can be quite slow, on the order of 7 FPS. If we are looking for pure speed then we tend to use YOLO as this algorithm is much faster, capable of processing 40-90 FPS on a Titan X GPU. The superfast variant of YOLO can even get up to 155 FPS. The problem with YOLO is that it leaves much accuracy to be desired.

SSDs, originally developed by Google, are a balance between the two. The algorithm is more straightforward than Faster R-CNNs. Depending on which variant of the network to use, we can also enjoy a much faster FPS throughput than Girshick et al. at 22-46 FPS. SSDs also tend to be more accurate than YOLO.

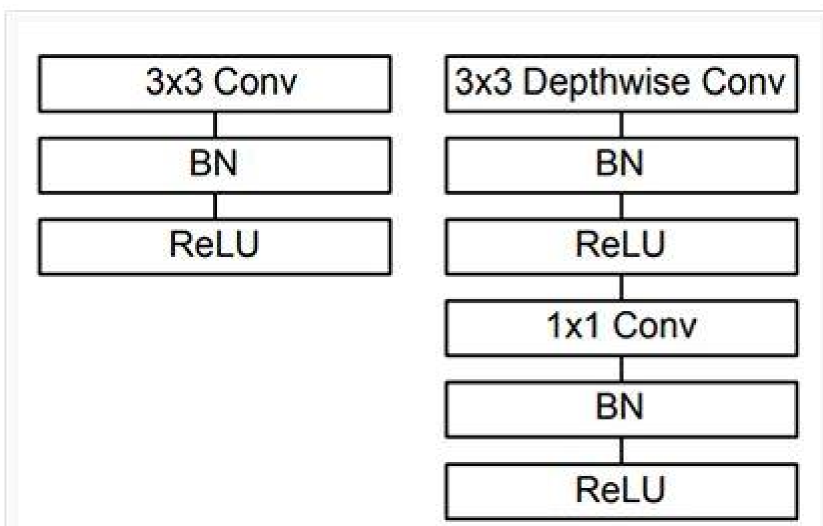

Figure 28: (Left) Standard convolutional layer with batch normalization and ReLU. (Right) Depthwise separable convolution with depthwise and pointwise layers followed by batch normalization and ReLU (figure and caption from Liu et al.).

When building object detection networks we normally use an existing network architecture, such as VGG or ResNet, and then use it inside the object detection pipeline. The problem is that these network architectures can be very large in the order of 200-500MB. Network architectures such as these are unsuitable for resource constrained devices due to their sheer size and resulting number of computations. Instead, we can use MobileNets (Howard et al., 2017), another paper by Google researchers. We call these networks "MobileNets" because they are designed for resource constrained devices such as smartphone. MobileNets differ from traditional CNNs through the usage of depthwise separable convolution.

The general idea behind depthwise separable convolution is to split convolution into two stages:

1. A 3×3 depthwise convolution.
2. Followed by a *1×1* pointwise convolution.

This allows us to actually reduce the number of parameters in our network. The problem is that we sacrifice accuracy. MobileNets are normally not as accurate as YOLO but they are much more resource efficient.

3.7Integrating MobileNets and single shot detectors for fast, efficient deep-learning based object detection

If we combine both the MobileNet architecture and the single shot detector (SSD) framework, we arrive at a fast, efficient deep learning-based method to object detection.

The model we'll be using in this blog post is a Caffe version of the original TensorFlow implementation by Howard et al. and was trained by chuanqi305. The MobileNet SSD was first trained on the COCO dataset (Common Objects in Context) and was then fine-tuned on PASCAL VOC reaching 72.7% MAP (mean average precision). We can therefore detect 20 objects in image (+1 for the background class), including airplanes, bicycles, birds, boats, bottles, buses, cars, cats, chairs, cows, dining tables, dogs, horses, motorbikes, people, potted plants, sheep, sofas, trains, and tv monitors.

3.8 Deep learning-based object detection with OpenCV

Deep learning algorithm does not show where an object is located on an image. The algorithm can classify an image into one of ImageNet's separate model class labels. To get the approximate boundary coordinates for a target object in an image, we need to instead apply object detection algorithms first.

Object detection does not only state us what object is on an image but it also shows where that object is located on the image.

The deep learning neural network makes use of an input image, a caffe prototxt, which is a great and very widely used framework for deep learning, offering a vast collection of out-of-the-box layers, a pre-trained model and a confidence (the minimum probability threshold to filter weak detections. Usually 20% is used as a default value). The first step is to initialize class labels and random colors for boundary corner values. After initializing the class labels and colors, it loads the model and the query image and prepare a blob which will be fed forward through the network. Then it extracts the height and width and calculates a fixed dimensional pixel blob from the input image. Next, it heavy lifts the blob and passes it through the neural network. It sets the input to the network and compute the forward pass for the input, storing the result as detections. Then, it computes the forward pass and associated detections which could take a while depending on our model and input size. Finally, it iterates over the detections and determine what and where the objects are in the image.

During the iterations multiple objects can be detected in a single image. A check is also applied to the confidence (i.e. probability) associated with each detection. If the confidence is high enough (i.e., above the threshold), then it will display the predicted value on the terminal as well as draw the predicted value on the image with text preferably and with a unique color border. During the looping processes on the detections, the first thing to do is extract the confidence value. If the confidence value is above the minimum threshold value defined, we extract the class label index and compute the border around the detected object. Then, we extract the (x, y) coordinates of the border box which we will use shortly for drawing a rectangle and displaying text.

Finally, they build a text label containing the class name and the confidence. Using the label, they print it to the terminal, followed by a coloured rectangle around the detected object using the extracted (x, y) coordinates. In general, they want the label to be displayed above the rectangle but if there is no place to hold it, it will be shown below the rectangle with an overlay of coloured label on the image using the calculated y-value.

3.9 Their Results

OpenCV and deep learning object detection results:

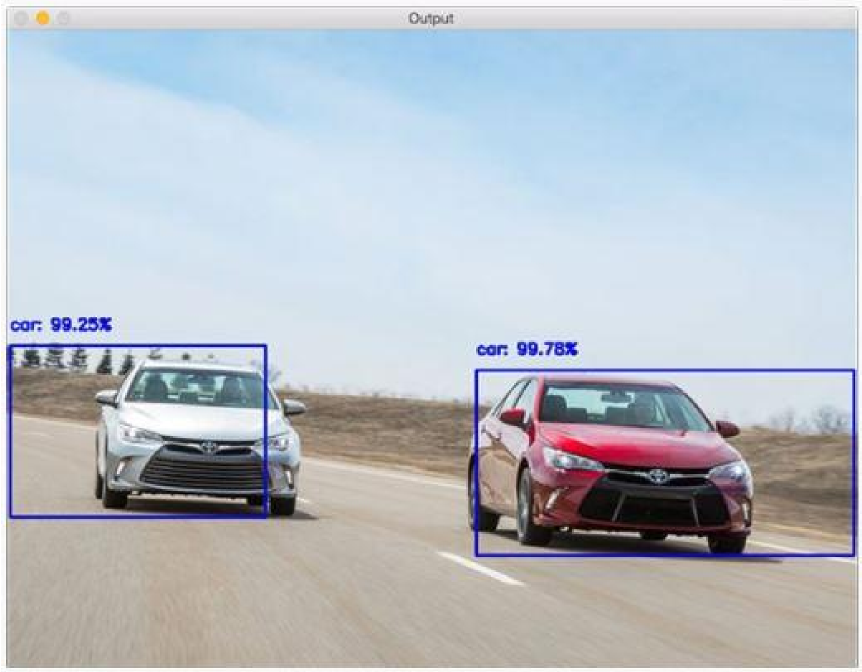

Figure 29: Two Toyotas on the highway recognized with near -100% confidence using OpenCV, deep learning, and object detection. Their first result shows cars recognized and detected with near -100% confidence.

Figure 30: An airplane successfully detected with high confidence via Python, OpenCV, and deep learning.

.

Figure 31: A person riding a horse and two potted plants are successfully identified despite a lot of objects in the image via deep learning-based object detection.

On their another example, they detected a beer bottle with an impressive 100% confidence.

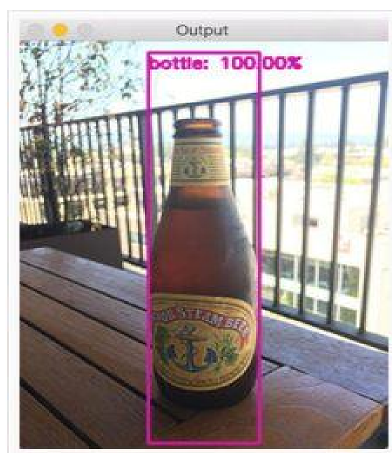

Figure 32: Deep learning + OpenCV correctly detect a beer bottle from an in input image.

Specifically, they used both MobileNets + Single Shot Detectors along with OpenCV 3.'s brand new (totally overhauled) dnn module to detect objects in images.

I owe a lot to the main contributor of the dnn module, Aleksandr Rybnikov, for making deep learning so accessible from within the OpenCV library.

Chapter 4

4. Related Works and Proposed Method

4.1 Related Works

There are plenty of works and contributions both for an indoor and outdoor navigation that aim at guiding visually impaired people to navigate in both indoor and outdoor environments. Still many algorithms and approaches are being added to improve an existing navigation and orientation systems in today's world. There is a vast space to explore new algorithm and solutions for the navigation and orientation system. However, they all are concerned with tread offs between performance and quality of service provided on one specific technology.

These visual navigation and orientation systems mostly rely on one or two technologies to exploit the location information for a navigation or orientation purpose. Most commonly used technologies are GPS, WIFI, Bluetooth, Radio-frequency identification (RFID), ultra-wide band (UWB), or sensor networks as infrastructure for localization. While infrared communication is alternatively used for indoor localization.

4.2 Navigation with voice recognition

The authors of [3] improve accuracy of a Wi-Fi indoor positioning system using Bluetooth dongles, but do not integrate RFID Tags. Another related reference [9] integrates Google Maps and Android into an indoor positioning system, tracking indoor objects using ZigBee module locations and outdoor objects using GPS. NFC-based indoor positioning and navigation systems targeting smartphones have been proposed by [10], and [5]. The indoor navigation application presented in uses NFC tags positioned at landmark locations for indoor routing. The client app downloads custom maps from a server and processes location information, computing navigation routes. This concept is extended in [10] by adding QR code support for localization and implementing a series of dedicated navigation services (e.g., finding the nearest toilet or car park). However, these publications use special-purpose maps which do not integrate with existing outdoor mapping services. We consider aggregation of various location sources as proposed by [5] to be required for determining an approximate location while not in the proximity of NFC tags or QR codes.

Beyond technological infrastructure, indoor positioning requires sophisticated positioning techniques and algorithms. Liu [15] and Al Nuaimi [1] categorize available techniques and algorithms into three groups: geometric approaches, scene analysis and proximity. Gu [8] mentions vision analysis as an additional group.

Main contribution of this paper is on one hand a generic architecture which seamlessly integrates indoor and outdoor location in an Android application based on Google Maps. The proposed IIPS architecture supports required indoor maps to be provided and location information processed by third-party service providers, offering interesting business opportunities. On the other hand, the presented server-based architecture aggregates data originated by a wide variety of location sensor technologies. Capturing of sensor-technology-specific properties supports algorithms in sensor prioritization, eventually increasing accuracy and offering reliable mechanisms for location consistency verification.

We argue that appropriate handling of complex, heterogeneous location sensor information, which mobile devices can acquire, requires a huge amount of context information for accurate interpretation. Examples include, but are not limited to, WLAN access point positions, signal strength, and building-specific signal propagation models. Therefore, IIPS prefers server-based processing of client-provided raw sensor information over client-only-processing or providing context information to mobile devices.

4.3 Google's searchable, interactive indoor maps for large venues

Google's MapsIndoors has revolutionized the indoor positioning by proposing and implementing an easy and cost-effective way of offering up-to-date way finding.

Everyone has, at some point, been lost inside a large convention center, struggled to find their gate at the airport or been late to a lecture, because they couldn't find the right auditorium at the university. But getting lost inside large and complex venues could soon be a thing of the past. Meet Maps Indoors – the next generation of indoor navigation and way finding. Maps Indoors is built on Google Maps, which makes the transition from outdoor navigation to indoor navigation completely seamless. All the known functionality and design of Google Maps is brought inside and the user can get directions from any point outside the venue and all the way inside. Maps Indoors can be integrated into existing applications for mobile and desktop or as a standalone service. And with the Maps Indoors Content Management System, updating and customizing the indoor maps has never been easier. That is indoor navigation with Maps Indoors[3].

MapsIndoors is built to fit in all industry needs such as in airports to indicate the directions to the gate, a specific shop or to a restaurant by which improving the traveller's experience and reducing the number of people missing their flights.. MapsIndoors offers seamless navigation from the outdoor world and into vast campus areas. In shopping centers to guide shoppers to specific shops and restaurants, which makes the shopping experience more fun and satisfying. Could collect information on how shoppers move around and use it to optimize advertising and rent pricing. In hospitals to improve a hospitals staff's efficiency and reduce patient and visitor stress by providing a familiar and easy-to-use way finding tool and many more.

Google Maps MapsIndoors can be interfaced with the most suitable indoor positioning system at your venue e.g., Wi-Fi positioning, Bluetooth (BLE) or positioning based on magnetic fields. The accuracy of the indoor positioning system depends on the chosen positioning technology. It is also influenced by the type of building that you are mapping. By interfacing your indoor navigation platform with a positioning system, you will be able to use location-based services such as push messaging and you will enhance the overall visitor experience. MapsIndoors works with leading positioning technologies such as Wi-Fi positioning, Bluetooth beacons and magnetic fields. Across the floors, you can manage your system and easily make changes. If you want to push information or coupons using push-notifications, you determine which content your Bluetooth beacons or Wi-Fi positioning system should push, at which locations and at what time. Use push-notifications to deliver a warm welcome to customers entering your shop or to notify them of a special offer at a certain location[3].

MapsIndoors CMS is an easy and cost-effective way of always offering users up-to-date way finding. It is intuitively built to make sure that anyone in an organization can use Maps Indoors CMS with only very little training. We can easily add, update and delete points of interest on your map. With a few clicks, we can add new locations and delete the old. Also change descriptions, opening hours and contact information for every location on our indoor map. By adding photos to a location's information page, we make sure that users can easier identify their destination. It is also possible to add icons to every venue, location or point of interest on the indoor maps. It can be company logos to highlight certain locations or any other icons of your choice. Also decide at which zoom level icons and content should appear. An indoor mapping solution can be customized to a preferred look and feel. The color scheme can be edited and route style to match the colors of a company or users facility[3].

4.4 Proposed Method

The proposed algorithm which is implemented here in this thesis work employs multiple technologies together as shown in the figure 4.4.1 below which can be used depending on the nature of the different location a user can be in at any time.

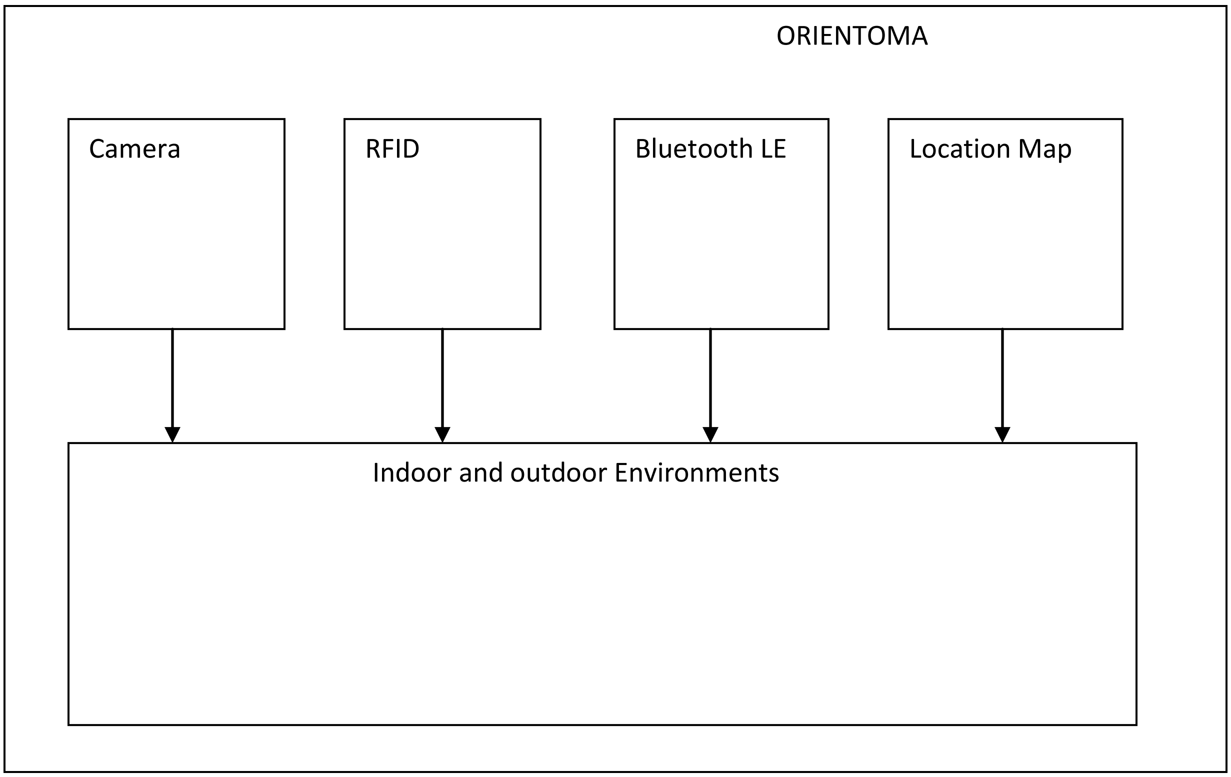

Figure 33: Block diagram of proposed system

But the location of a user is dynamically changing depending on both the direction of the motion activity and the speed of the motion activity. The location of a visually impaired person should be tracked dynamically during his/her movement in a day. This is handled by using the suitable alternative technologies provided by ORENTOMA application. A suitable feature with a voice command is chosen among the four different technologies shown in figure 4.7 depending on the nature of the environment in context. The detection of an object or obstacle or a location information depends on the type of service/technology selected. For an outdoor environment, the camera and map service are used mostly to detect obstacles or get location information, resulting in better quality results due to their suitability for outdoor environments. If the environment changes from outdoor to indoor (less light detection by the camera or missing signal from GPS), and the current selected technology should be changed from map/camera to for example to RFID/Bluetooth LE for the sake of exploiting indoor navigation technologies. This dynamic adjustment of the ORIENTOMA features increases the robustness of the proposed algorithm to deal with any kind of object detection or to get location based information during the daily activities of a visually impaired people.

Figure 4.2 shows the proposed algorithm which mainly includes different modules. The software and hardware modules used in the proposed algorithm includes OpenCV library, OpenCV manager, Google TTS API engine for English and Italian languages, GPS, Android phone with camera for object/obstacle detection, Bluetooth LE, and passive RFID tags.

The architecture of the proposed algorithm consists of a layered phase for each technology. The camera option has the capture phase, pattern matching phase, voice response with matched object name and optional information or orientation or warning. The RFID tags option has a Launch/initialization phase, Signaling phase, scan/reading phase, voice response of detected tag/ bulk read. The Bluetooth LE option has a GAAT command initialization phase, connect to server phase, exchange info phase, voice response of detected tag/ bulk read to the user. Map option has a load map phase, locate latitude and altitude of current location phase, and voice response of location information to the user.

The text to speech module provides a hands free and user interface independent command to the system directly from the blind and partially sighted people. To support voice command, the system integrates the Google's Text To Speech module for voice interaction and the Google Voice Recognizer API. Text To Speech synthesizes speech from text. It is used by the system to internally process information to be spoken to the user.

A server client approach is followed by voice module, where Voice Recognizer part communicates with the server to processes information and send it back to the user as text. 3) Address query translate: The address query translates geographic to coordinate this includes the geocoding part, converting the destination address into latitude and longitude, this in turn provides the detail information of the destination address the information module comprises all the data related to points of interest outside the building that is outdoor environment. Every location, bus stand, counters, schools, colleges is associated to one specific category.

If a visually impaired or blind person wants to go from one place to another place, the route must be predefined. This information must be clear and concise, once it will be used for the selection of services it will be delivered to the user by synthesized speech. 4) Route Query: Route query takes the blind user current coordinate from GPS and the destination coordinate, and compute the routes. The localization module is designed to constantly monitor the position of a user using GPS module. Different Places have different latitude and longitude hence as the person move from one places to another place the latitude and longitude changes according to the place. Because of this, many new designs can take the advantage of being tractable. 5) Route transversal: Route transversal provides audible instructions to user in the form of speech so that the blind person can travel independently. Once the destination address is obtained this address must be translated to geographic point. The destination address will be geocoded using the GEO-CODER module and then passed to text to speech synthesizer to generate a pedestrian route. 6) Object detection: For obstacle detection ultrasonic sensor has been used, this ultrasonic sensor emits ultrasonic beams to the environment, which are reflected by the object; the system calculates the distance from the object according to the time difference between the emitted and received beam. The stereo-vision systems use the object tracking

Chapter 5

5. Implementation of the proposed Algorithm

5.1 Introduction

There are approximately 38 millions of people across the world mainly in developing countries who are blind and visually impaired, over 15 million are from India. Blind persons most of the time are withdrawn from the society because they feel that people and the society are prejudiced and they may not be welcomed most of the time. The remarkable achievement, which is the outcome of persistent struggle and hard work between “Anne Sullivan” – the teacher and “Helen Keller “-the blind student resulted in a revolutionary method of learning and communication, which ultimately culminated in the development of Braille language. Blind person do not need pity, but require empathy, so as to mingle in the society and be independent for their routine chores (activity). Hence blind people need an assistive device that will allow blind user to navigate freely and this requirement has become crucial. Most of the blind people depend on other individuals, white cane or guide dogs to travel freely. Currently, there are several visual information that helps visually impaired people to move in a right way (e.g. takes a right direction, take left, move forward, move backward and avoid obstacles,) but they all limit the freedom of the user[12].

The necessity for visually impaired people to undertake the daily activities in both an indoor and outdoor environment without any constraint and to have access to location information has been limited by the versatility of application/algorithms that provides such services. As a result, implementation of a hybrid technology that integrates the functionality of different domain services such as Bluetooth and RFID for near field communication preferably in indoor environments, while camera and map services for outdoor navigation and orientation system is necessary to fully utilize their time instead of getting concerned about how they would reach from a specific location to some destination. Several algorithms are proposed to address the problem of navigation and orientation.

5.2 General Structure

The Proposed algorithm is implemented in a software and hardware by employing five different technologies and hardware infrastructure. These are Camera, RFID, Bluetooth LE, Map and OpenCV from software library. The usage of a specific technology at any time is managed by the user who anytime decides which infrastructure of the application best suits for the current context or environment to use. As it is described in the previous section, different quality of service is delivered by each technology and the user is responsible to choose and use appropriate technology feature from the proposed algorithm in the context of the environment. The block diagram for the architecture is shown in the figure 5.2 below:

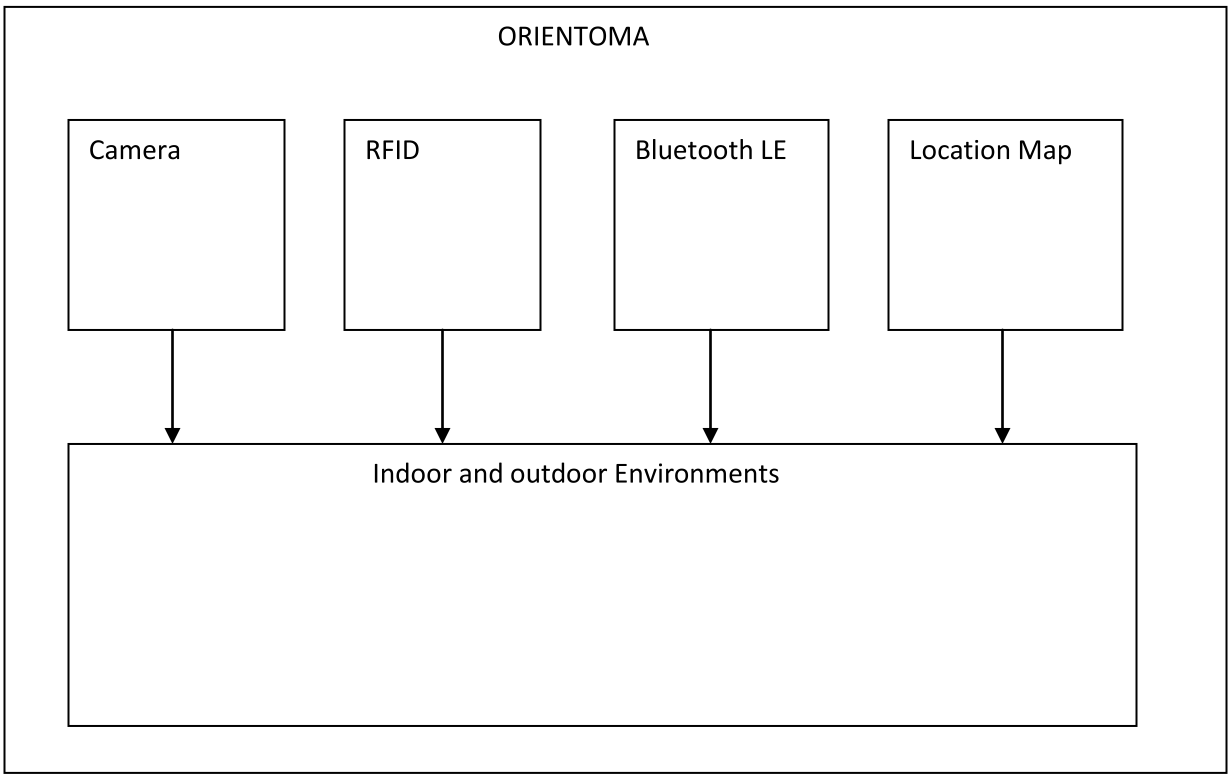

Figure 34: Block Diagram of Proposed Architecture

In the following sections the Camera, RFID, Bluetooth LE and MAP as well as the other modules which are described above are discussed in detail.

This days in visual navigation and orientation systems the critical issues are not only throughput and memory usages but the quality of the encoded video signals also matters. This is very critical when dealing with higher resolution video. Here in this work the proposed algorithm which is described in previous chapter is implemented in a hardware considering all the issues mentioned above including throughput, memory and qualities in higher resolution videos. In the following sections the proposed architecture is described in detail.

Walking securely and unhesitatingly with no human help within urban environment is a troublesome undertaking for visually impaired and blind individuals. The fundamental goal is to give an ease or financially savvy approach that will permit visually impaired individuals to explore freely or independently in the outdoor environment. Based on this real context or condition we focused the work on developing assisting technologies that may help blind individuals bringing them back to the society. Our main objective is to make a compact, self-sufficient system that will permit these blind people to travel through an environment. This voice based route navigation system can provide solution to this problem. This System is based on embedded system and provides navigation instructions to the user by giving audio instructions through speaker. This navigation system will detect an obstacle using camera sensor and guide blind person by providing audio instructions.

5.3 GPS (Global Positioning System)

The Global Positioning System (GPS) is a space-based satellite navigation system that provides location and time information in all weather conditions, anywhere on or near the earth where there is an unobstructed line of sight to four or more GPS satellites. The system provides critical capabilities to military, civil, and commercial users around the world. The United States government created the system, maintains it, and makes it freely accessible to anyone with GPS receiver. The Global Positioning system (GPS) is a network of 30 satellites orbiting the earth at an altitude of 20,000km Whenever you are on the planet, at least four GPS satellites are ‘visible’ at a time .Each one transmits information about its position and the current time at regular intervals. These signals, travelling at the speed of light, are intercepted by your GPS receiver, which calculates how far away each satellite is based on how long it took for the message to arrive. The GPS (Global Positioning system) receiver continuously receives the latitude and longitude values for every position of the system and it is interfaced by the GPS receiver. The Global Positioning System (GPS) offers the capability to accurately determine location anywhere on earth in addition to speed, altitude, heading, and a host of other critical positioning data. The GPS receiver requires a successful lock onto at least four GPS satellites to gather an accurate signal for calculating position and velocity. The module triangulates its position with relation to three satellites, using a fourth satellite as a clock source. $GPRMC,121413.000,V,2400,0000,N,12100.0000,E,000.0, 000.0,280606,,,N*7C Where: RMC Recommended Minimum sentence C 121413Fix taken at 12:14:13 UTC A Status A=active or V=Void.

2400, 0000, N Latitude 24 deg 00.00' N 12100.0000, E Longitude 12 deg 10.000' E 022.4 Speed over the ground in knots 280315Date - 28th of March 2015 *7C the checksum data, always begins with *

5.4 Ultrasonic

The human ear can hear sound frequency of around 20HZ ~ 20KHZ. And ultrasonic is the sound wave beyond the ability of which humans can hear is of 20KHZ and are not harmful for human being. The Ultrasonic Transmitter which will send a signal out into its surrounding area. The Ultrasonic Receiver will detect this signal once it bounces off from an object/obstacle. Ultrasonic sensors are basically used to measure the distances between the obstacle object and the sensor. The ultrasonic sensor works based on the principle of Doppler Effect. It consists of an ultrasonic transmitter and a receiver. The transmitter transmits the signal in one direction and this transmitted signal is then reflected back whenever there is an obstacle and it is received by the receiver. So the total time taken by the signal to get transmitted and to received back will be used to calculate the distance between the ultrasonic sensor and the obstacle. Ultrasonic sensor provides a very low-cost and easy method of distance measurement. Many animals have the ability to hear ultrasonic frequency range for example bats.

5.5 Text to Speech

The function of a TTS system is to convert the given text into a spoken waveform. In order for us to give verbal instructions to the user, we need to convert our text instructions into audible speech. We decided to use Espeak. The major benefit of using Espeak is that it is open source and it allows you to output speech in many different languages. Therefore we sent a string of data to Espeak with the instruction which we wanted to tell the user and this text to speech synthesiser converted the text into speech data, which was then played for the user. The database has been created for different addresses which include bus stops, colleges, hospitals, etc. This conversion involves text processing and speech generation processes. This approach seeks to develop strategies for concatenating stored speech segments as a means of Synthesizing speech. Sub-word units, such as syllables or diaphones, in which articulation between adjacent phonemes are preserved, are considered as satisfactory units, under this approach to synthesizing speech.

5.6 Speech to text

5.6.1 Speech recognition

We used the open source speech to text engine, pocket sphinx [2] developed by Carnegie Melon University. We choose this particular open source engine because it allowed us to easily add new words to the dictionary of required words and also allowed us to train the engine to better recognize the speech of a particular user. To use pocket sphinx we needed to create input files for the engine .To "Teach" the engine a new word, we had to generate a phonetic dictionary containing the word and a language model containing the word as well .we also needed an acoustic model, but the one that came with pocket sphinx would be sufficient. The phonetic dictionary file is a simple mapping of the word to its corresponding phones. The phoneme for a word is just the distinct set of units of sound that describes that word and distinguishes it from other words.

5.6.2 Google voice recognition API (speech to text)

Speech recognition can be achieved in many ways on Linux (so on the Raspberry Pi), the accuracy is very good, and has a strong accent; it starts recording and saves the audio in a flac file format. The audio file is then sent to Google for conversion and text will be returned and saved in a file called “stt.txt”.

5.5 GPS (Global Positioning System)

We used the open source speech to text engine, pocket sphinx [2] developed by Carnegie Melon University. We choose this particular open source engine because it allowed us to easily add new words to the dictionary of required words and also allowed us to train the engine to better recognize the speech of a particular user. To use pocket sphinx we needed to create input files for the engine .To "Teach" the engine a new word, we had to generate a phonetic dictionary containing the word and a language model containing the word as well .we also needed an acoustic model, but the one that came with pocket sphinx would be sufficient. The phonetic dictionary file is a simple mapping of the word to its corresponding phones. The phoneme for a word is just the distinct set of units of sound that describes that word and distinguishes it from other words.

For example, the light computations on image detection provides a performance gain in real time. As a result implementation of integrated architecture is necessary for such application. Several architectures are proposed to integrate multiple algorithms together and a minor performance loss is observed.

As the complexity of object detection is highly influenced by the train model set and search algorithm and the search range, most of the proposed solutions are mainly focus on the throughput and performance of memory for the increased search range. Image matching requires a number of memory access that leads to high power dissipation, so there is always a trade of between power consumption and internal memory usage.

This days in object detection process the critical issues are not only throughput and memory usages but the quality/confidence of the detected object also matters. This is very critical when dealing with poor resolution image in areas of low visible light and indoor environments. Here in this work the proposed algorithm which is described in previous chapter is implemented in a software considering all the issues mentioned above including throughput, memory and confidence in low resolution images/videos. In the following sections the proposed architecture is described in detail.

5. 6 RFID

Radio-frequency identification (RFID) is a technology that uses electromagnetic fields to automatically detect, identify and track RFID tags attached to objects. The tags are used to store information. Tags are of two types: passive tags which collect their energy from a nearby RFID reader's interrogating radio waves whereas active tags have a local power source (such as a battery) and may operate hundreds of meters from the RFID reader. Unlike a barcode, the RFID tags need not be within the line of sight of the reader, hence are suitable methods to embed them in specific areas for the sake of blind and visually impaired people to track location and some commonly used objects in their day to day life.

5.6.1 Tags

An RFID system uses labels or commonly known as tags attached to objects to be identified. Bi-directional transmitter-receivers radio signal interrogators or readers send a signal to the tag and read its response. The tags could be either passive (no battery connected), or active or battery-assisted passive. An on-board battery is available for the active tags hence it can periodically transmit its ID signal. A battery-assisted passive (BAP) has a small battery on board and is activated when in the presence of an RFID reader. In terms of cost, the passive tag is cheaper and smaller because it has no battery; but it uses the radio energy transmitted by the reader. To operate a passive tag, there must be a power stronger than for signal transmission. The greater power makes a difference in interference and in exposure to radiation. Tags may either be read-only, or may be read/write, or write-once read-multiple; in which case object-specific or location specific data can be written into the tag by the system user.

An RFID reader transmits an encoded radio signal to interrogate the tag. The RFID tag receives the message and then responds with its identification and other information. This may be only a unique tag serial number, or may be product-related information such as a stock number, lot or batch number, production date, or other specific information. Since tags have individual serial numbers, the RFID system design can discriminate among several tags that might be within the range of the RFID reader and read them simultaneously.

5.6.2 Readers

RFID systems can be classified by the type of tag and reader. A Passive Reader Active Tag (PRAT) system has a passive reader which only receives radio signals from active tags (battery operated, transmit only). The reception range of a PRAT system reader can be adjusted from 1–2,000 feet (0–600 m)[citation needed], allowing flexibility in applications such as asset protection and supervision.

An Active Reader Passive Tag (ARPT) system has an active reader, which transmits interrogator signals and also receives authentication replies from passive tags.

An Active Reader Active Tag (ARAT) system uses active tags awoken with an interrogator signal from the active reader. A variation of this system could also use a Battery-Assisted Passive (BAP) tag which acts like a passive tag but has a small battery to power the tag's return reporting signal.

Fixed readers are set up to create a specific interrogation zone which can be tightly controlled. This allows a highly defined reading area for when tags go in and out of the interrogation zone. Mobile readers may be hand-held or mounted on carts or vehicles.

5.6.3 Signaling

Signaling between the reader and the tag is done in several different incompatible ways, depending on the frequency band used by the tag. Tags operating on LF and HF bands are, in terms of radio wavelength, very close to the reader antenna because they are only a small percentage of a wavelength away. In this near field region, the tag is closely coupled electrically with the transmitter in the reader. The tag can modulate the field produced by the reader by changing the electrical loading the tag represents. By switching between lower and higher relative loads, the tag produces a change that the reader can detect. At UHF and higher frequencies, the tag is more than one radio wavelength away from the reader, requiring a different approach. The tag can backscatter a signal. Active tags may contain functionally separated transmitters and receivers, and the tag need not respond on a frequency related to the reader's interrogation signal.[18]

An Electronic Product Code (EPC) is one common type of data stored in a tag. When written into the tag by an RFID printer, the tag contains a 96-bit string of data. The first eight bits are a header which identifies the version of the protocol. The next 28 bits identify the organization that manages the data for this tag; the organization number is assigned by the EPCGlobal consortium. The next 24 bits are an object class, identifying the kind of product; the last 36 bits are a unique serial number for a particular tag. These last two fields are set by the organization that issued the tag. Rather like a URL, the total electronic product code number can be used as a key into a global database to uniquely identify a particular product.[19]

Often more than one tag will respond to a tag reader, for example, many individual products with tags may be shipped in a common box or on a common pallet. Collision detection is important to allow reading of data. Two different types of protocols are used to "singulate" a particular tag, allowing its data to be read in the midst of many similar tags. In a slotted Aloha system, the reader broadcasts an initialization command and a parameter that the tags individually use to pseudo-randomly delay their responses. When using an "adaptive binary tree" protocol, the reader sends an initialization symbol and then transmits one bit of ID data at a time; only tags with matching bits respond, and eventually only one tag matches the complete ID string.[20]

5.7 Camera

The users cell phone embedded camera is used to capture and track objects. The detection of an object/obstacle involves feature extraction, feature matching with the model images from a local database, and confidence level evaluation for each matching steps. The diagram 5.1 below shows the steps performed during an object detection using the camera technology.

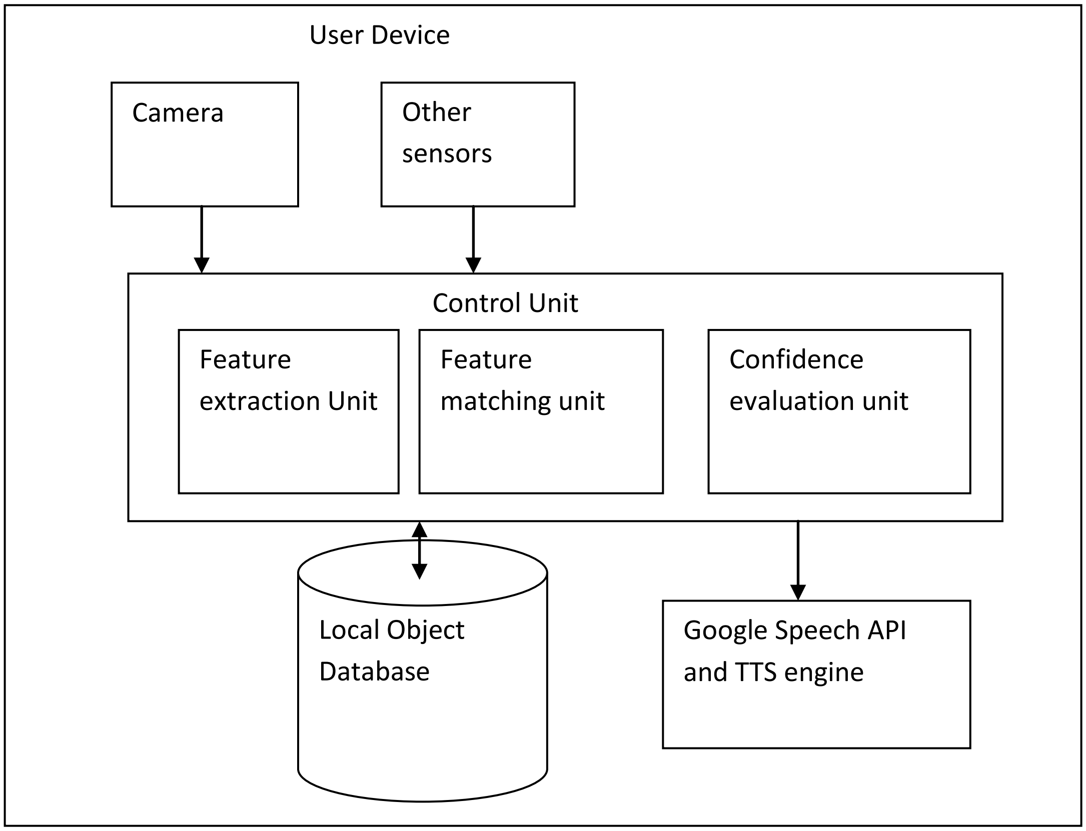

Figure 35: Object detection steps using camera + OpenCV

5.8 Bluetooth LE IBeacon detection

Though Android does not have native IBeacon support, to use IBeacon on Android, we have created code that parses BLE packets to find IBeacon data. Bluetooth low energy devices can operate in an advertisement mode to notify nearby users or devices of their presence. IBeacon is a Bluetooth low energy device that emits advert following a strict format that is IBeacon prefix, followed by a variable UUID, and a major, minor pair. Figure 5.1 shows the version and symbol used as a representation for Bluetooth LE.

Bluetooth™ 4.0 Low Energy

Figure 36: Bluetooth LE version and symbol used in the proposed solution

5.8.1 Difference between IBeacons and RFID

BLE (or active RFID) systems operate by a tag "beaconing," or sending out transmission, to a reader, and then transmitting that location to the cloud. iBeacons are a type of active RFID that use BLE..

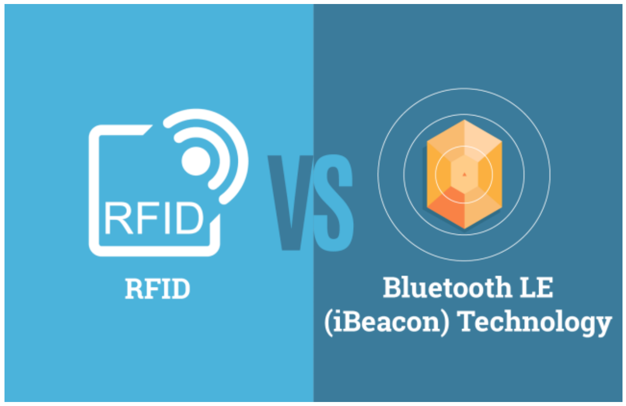

Figure 37: Overlap of functionality between Bluetooth LE and RFID

5.9 Indoor mapping

5.9.1Google's Indoor mapping made flexible and easy to manage

Maps Indoors CMS developed by Google is custom-built content management system that allows anyone to add, edit and control location-based data in his/her indoor mapping solution. The system does not require any programming skills – it can be used by anyone in an organization. It is cloud-based so that no server resources are required from a user. Data can either be put in manually or via an automated bulk upload. This makes it easy to carry out everything from a single adjustment to large changes.

Chapter 6

6. Results and Performance Analysis

6.1 Results

To evaluate the proposed algorithm as part of applicable real-world navigation and orientation system, the proposed system was installed and deployed on android mobile phone. The deployed system was set on camera mode and used to detect different object sequences. Twenty different test images that are listed in the table 6.1 are used to compare the proposed algorithm with earlier algorithms. The test is carried on different types of images with different resolution. The image detection algorithm is used from the OpenCV library.

Screenshots shown below are taken during the usage of the application in real time. Figure 6.1 shows the main or home page of the application in a black theme. The theme option is not necessary as the application is mostly used in voice-based interaction.

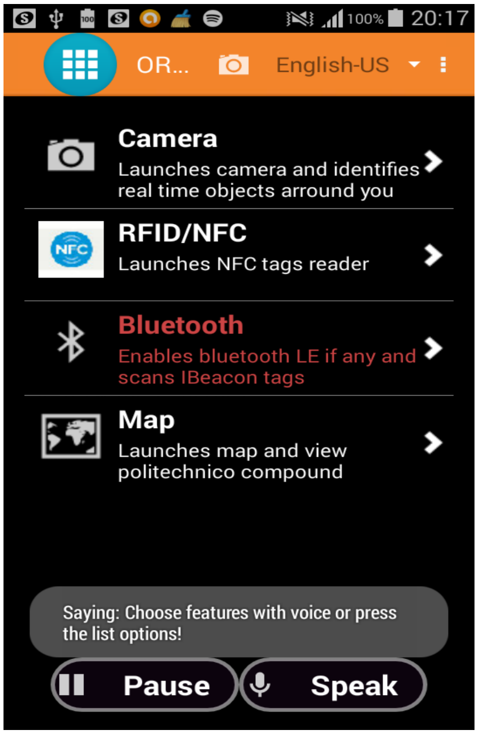

Figure 38: Application deployed on android s4 mini

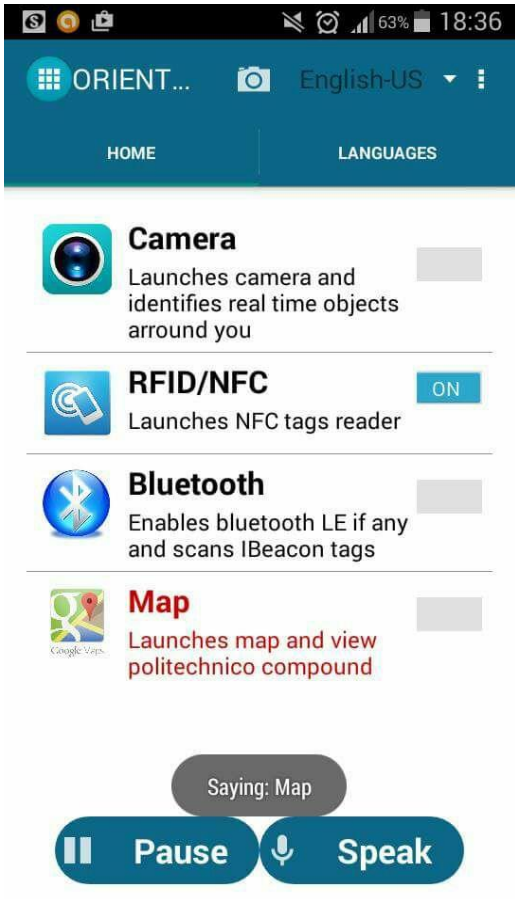

Figure 39: Application in white theme.

Figure 40: IBeacon Bluetooth LE feature enabled and in use.

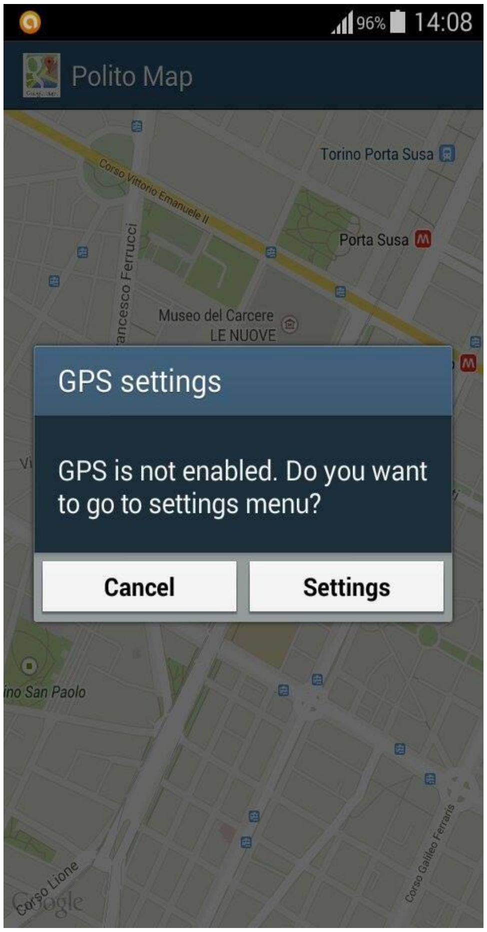

Figure 41: Location [positioning] setting using audio command

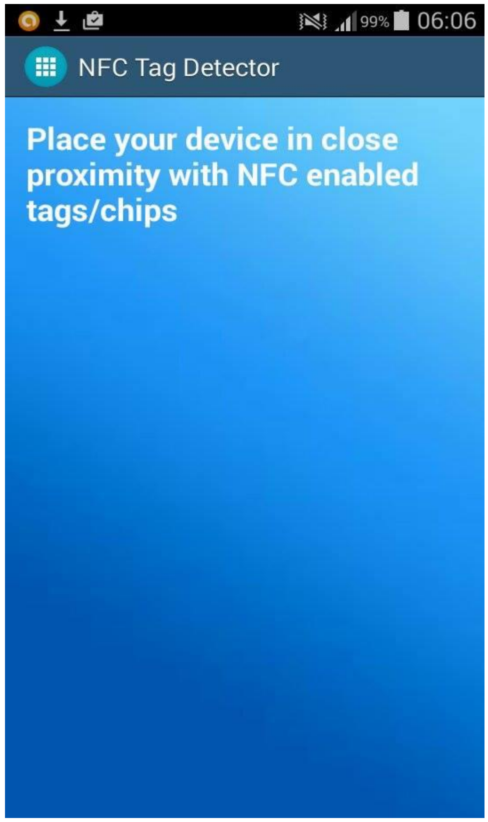

Figure 42: NFC feature enabled and in use to detect RFID tags

6.2 Performance Analysis

In this section the result for quality of detection and performance for the proposed algorithm as well as related algorithms are presented and discussed. The performance in terms of object detection quality is measured by a relative comparison with the previous existing systems. Whereas the performance in terms of computational cost is compared by the response time.

6.2.1 Detection time vs. quality of object matching

The efficiency of detection is measured by its real-time performance achievability. Features of an image is extracted and used to recognize an object from the test domain. The preliminary steps to image matching includes rotation, scaling, translation, adjusting brightness, resolution and others. Image marching should look for an interest points and the match has to be unambiguous in other images.

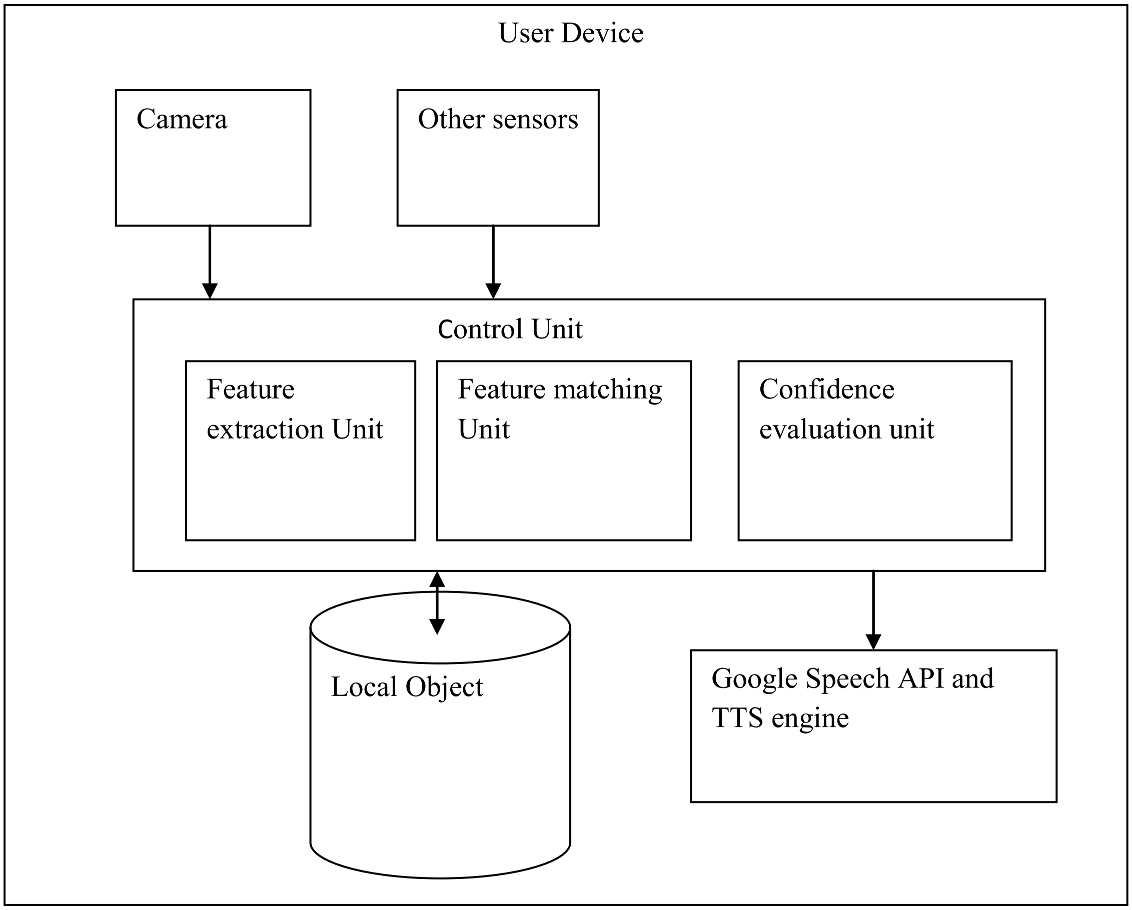

Figure 43: Object detection steps

Chapter 7

7. Conclusion and Future works

7.1 Conclusion

To improve the quality of life for visual impaired or blind people, in this work we focused on developing a hybrid technology to help these persons to access the outdoor-indoor environments in particular such as universities, hospitals, and other public utility. Therefore this work plays a special role in these fields providing as much information as possible for visually impaired or blind people, which allows them to take a comfortable navigation. To build a prototype we focused on users and their interests. The thesis work aims to build a system that assists people with disabilities in providing the navigation and location information.

Overall the project has been a success with the entire project requirement. In the proposed architecture integration of different technologies that detect an object/obstacle or provide location information is employed however still other new technologies and techniques could be added to improve the navigation and orientation system to support a hazardous environment where it is impossible to utilize the voice communication. The other scope of this work is the restricted usage of technologies for a certain domain area. The proposed algorithm is implemented with a camera feature for the day time of an outdoor environment where there is enough visible light to detect objects. However, this work can be extended to work with the camera feature in night mode where there is few visible lights by employing a flash light from the camera or using a complex algorithm to detect objects from a low quality images.

7. 2 Advantages

- Easy to use application
- Free to use
- Applicable for both the indoor and outdoor environments such as hospitals and colleges
- Easy to set up

- Less space
- Low power consumption application
- This system can be used in both the known and unknown environments like airports, malls and public parks etc.

The main goal of this thesis was the development of visual navigation and orientation system for visually impaired people. The computational complexity and delay of switching between application feature suitable for a particular environment is very important. The complexity gets worst while trying to switch a feature with a voice command in noisy areas where it takes a number of trials to get the voice command detected by the user. As a result different approaches can be employed to alleviate the this problem. The proposed algorithm was implemented in a software and hardware to the best way possible to increase the performance and reduce the unnecessary computational cost on the users mobile phone.

7.3 Future works

In the proposed architecture we focused on integration of different technologies that provide location information or detect an object/obstacle which has an overhead on performance due to the different libraries used. The other scope of this work is the restricted usage of technologies for a certain domain area. For example, the proposed algorithm is implemented with a camera feature for the day time of an outdoor environments where there is enough visible light to detect objects. However, this work could be extended to support camera feature in night mode where there is few visible lights by employing a flash light from the camera or using a complex algorithm to detect objects from a low quality images. Also, the application could be extended to support a remote areas where internet connectivity and accessibility to RFID/beacon tags are limited so that the same information could be delivered in a reliable way. For example, an offline map could be employed for such purposes. The map of a specific environment could be downloaded at once during an access to internet and then the main points/nodes on the map could be labeled (with their latitude and longitude) and relate the points/nodes with a pre-recorded information to be played to the user during the entrance of the user in that range of the latitude and longitude without the need to use any RFID or beacon tags or camera inputs. For example while the application is active, when the user enters in a specific latitude and longitude range, the application could invoke (speak to the user in this case) the stored or pre-recorded information about that location.

BIBLIOGRAPHY

[1] EVOLUTION of Mobile Location-based Services, by BHARAT RAO AND LOUIS MINAKAKIS, 2003.

[2] Joachim Fabini, Reinhard Russ, and Lukas Wallentin, *Architecture and Prototypical Realization of NFC-based Indoor Positioning with Android and Google* Maps.

[3] "Searchable, interactive indoor maps for large venues", 2017

[4] Nandish M S, Mr. Chetan Balaji, Prof. Shantala C P, *An Outdoor Navigation With Voice Recognition Security Application For Visually Impaired People*, Volume 10 Number 10 - Apr 2014

[5] Troy Coverstone, Christine Cronin, Sofie Kniazeva, *GPS TECHNOLOGY TO AID THE BLIND AND PARTIALLY SIGHTED IN COPENHAGEN, DENMARK*, May 7, 2007

[6] Manav Singhal, Anupam Shukla, *Implementation of Location based Services in Android using GPS and Web Services*, Vol. 9, Issue 1, No 2, January 2012 ISSN (Online): 1694-0814

[7] Louise Barkuus, and Anind Dey, *Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns*, IRB-TR-03-024, July, 2003

[8] Jae Sung Cha, Dong Kyun Lim, Yong-Nyuo Shin, *Design and Implementation of a Smart Phone Navigation Application for Blind Persons*, ISA 2013, ASTL Vol. 21, pp. 149 - 151, 2013

[9] Andrew G. Howard Menglong Zhu Bo Chen et al, *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,* 17 Apr 2017.

[10] Harsha Gawari, Prof. Meeta Bakuli, *Voice and GPS Based Navigation System For Visually Impaired,* ISSN : 2248-9622, Vol. 4, Issue 4(Version 6), April 2014, pp.48-51

[11] Jae Sung Cha, Dong Kyun Lim and Yong-Nyuo Shin, *Design and Implementation of a Voice Based Navigation for Visually Impaired Persons,* Vol. 5, No. 3, June, 2013

[12] Anushree Harsur Bangalore, India, *Voice Based Navigation System for Blind People Using Ultrasonic Sensor,* ISSN: 2321-8169 Volume: 3 Issue: 6 4117 – 4122 June 2015.

[13] Satya Mallick, *Image Recognition and Object Detection*, NOVEMBER 14, 2016.

[14] L.Tetley and D.Calcutt, *Electronic Navigation Systems,* 2000

[15] "*Object Class Detection"* 2013-10-09.

[16] *"location-based service (LBS)",* May 2009

[17] "*Automatic Identification and Data Collection (AIDC)*" Archived May 5, 2016

[18] Daniel M. Dobkin*, The RF in RFID: Passive UHF RFID In Practice, Newnes 2008 ISBN 978-0-7506-8209-1, chapter 8*

[19] John R. Vacca *Computer and information security handbook*, Morgan Kaufmann, 2009 ISBN 0-12-374354-0, page 208

[20] Bill Glover, Himanshu Bhatt *,RFID essentials, O'Reilly Media*, Inc., 2006 ISBN 0-596-00944-5, pages 88-89

[21] Curran, Kevin; Furey, et.al, *An Evaluation of Indoor Location Determination Technologies,*2011.

[22] Milan Herrera Vargas *Indoor Navigation Using Bluetooth Low Energy (BLE) Beacons, 2016*

[23] Quercia, Daniele; Lathia, Neal; Calabrese, Francesco; Di Lorenzo, Giusy; Crowcroft, Jon (2010). *Recommending Social Events from Mobile Phone Location Data (PDF). 2010 IEEE International Conference on Data Mining. p. 971.* 10.1109/ICDM.2010.152. ISBN 978-1-4244-9131-5.

[26] Pourhomayoun; Jin; Fowler (2012). *"Spatial Sparsity Based Indoor Localization in Wireless Sensor Network for Assistive Healthcare Systems" (PDF). EMBC2012.*

[27] "CrowdOptic Enhances Launch and Tracking of Lunar Lander". NASA. Retrieved 26 October 2016.

[28] "Location-based service", 2018.

[29] "Geo-Based Processes & Services", 2017

[30] "10 Great Examples of Location-Based Apps", 2017.

[31] " Indoor Positioning, Tracking and Indoor Navigation with Wi-Fi", 2017

[32] "Indoor Positioning, Tracking and Indoor Navigation with Beacons" , 2017

[33] "Indoor Positioning with Ultra-wideband", 2017

[34] "Indoor Localization with RFID" , 2017

[35] "External Camera Systems in Addition to Indoor Positioning", 2017

[36] "Rapid object detection using a boosted cascade of simple features"

[37] “Viola, Jones: Robust Real-time Object Detection”, IJCV 2001 pages 1,3.

[38] “Viola–Jones object detection framework”, 2018