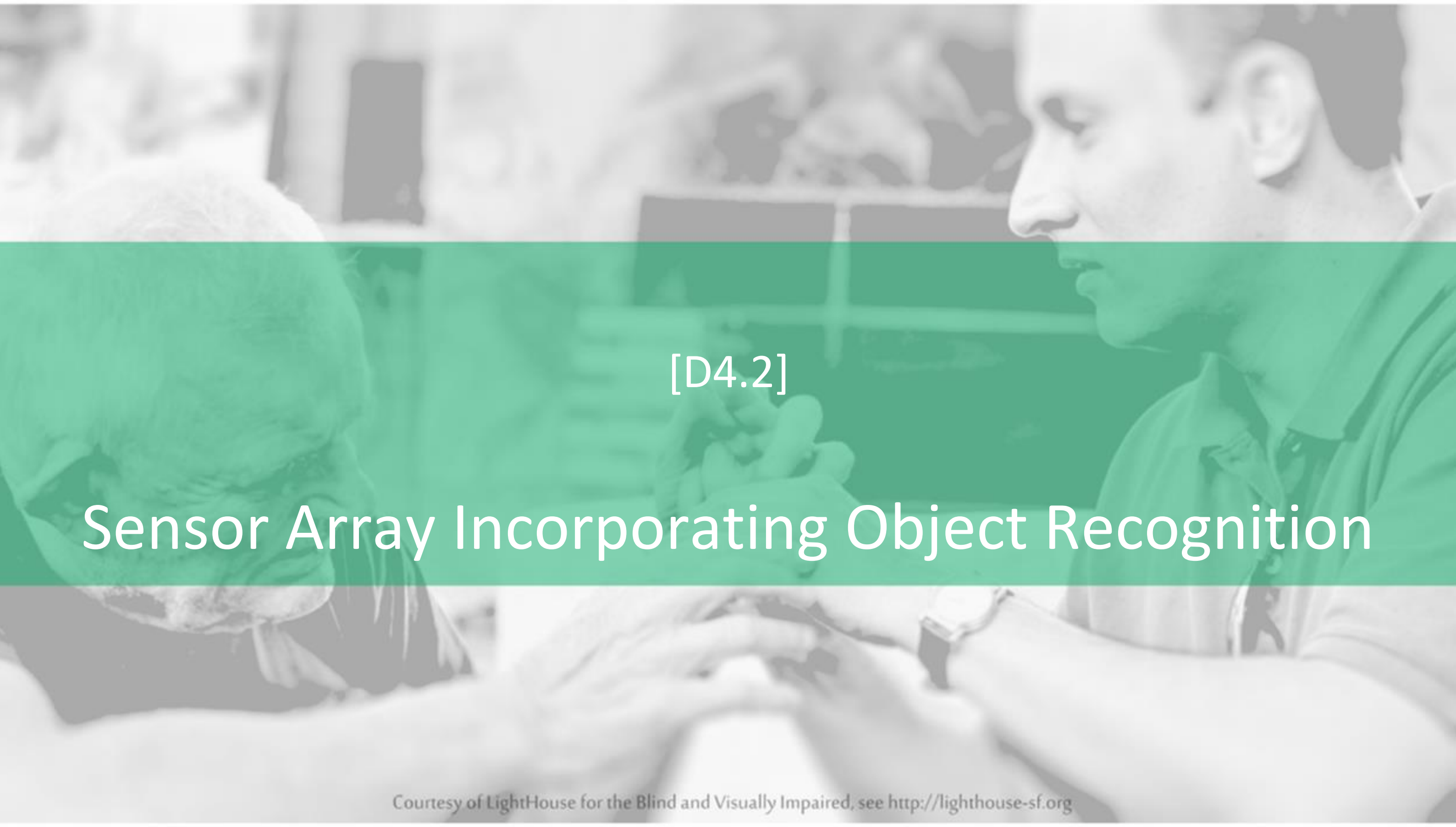
SUITCEYES

1 Jan 2018 - 31 Dec 2020

Smart, User-friendly, Interactive, Tactual, Cognition-Enhancer, that Yields Extended Sensosphere
Appropriating sensor technologies, machine learning, gamification and smart haptic interfaces

[D4.2]

# Sensor Array Incorporating Object Recognition



Courtesy of LightHouse for the Blind and Visually Impaired, see http://lighthouse-sf.org

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780814.

| Dissemination level | | |
|---|---|---|
| **PU** | PUBLIC, fully open, e.g. web | X |
| **CO** | CONFIDENTIAL, restricted under conditions set out in Model Grant Agreement | |
| **CI** | CLASSIFIED, information as referred to in Commission Decision 2001/844/EC. | |

| Deliverable Type | | |
|---|---|---|
| **R** | Document, report (excluding the periodic and final reports) | |
| **DEM** | Demonstrator, pilot, prototype, plan designs | X |
| **DEC** | Websites, patents filing, press & media actions, videos, etc. | |
| **OTHER** | Software, technical diagram, etc. | |

| Deliverable Details | |
|---|---|
| **Deliverable number** | 4.2 |
| **Part of WP** | 4 |
| **Lead organisation** | UNIVLEEDS |
| **Lead member** | Raymond Holt |

| Revision History | | | |
|---|---|---|---|
| **V#** | **Date** | **Description / Reason of change** | **Author / Org.** |
| **v0.1** | - | Structure proposal | - |
| **v0.2** | 10/12/2019 | Draft for Internal Review | Zhengyang Ling / UNIVLEEDS |
| **v1.0** | 19/12/2019 | Submission to the EU | T. Bebis / HB |

| Authors | |
|---|---|
| **Partner** | **Name(s)** |
| **UNIVLEEDS** | Raymond Holt |
| **UNIVLEEDS** | Brian Henson |
| **UNIVLEEDS** | Zhengyang Ling |

| Contributors | | |
|---|---|---|
| **Partner** | **Contribution type** | **Name** |
| **CERTH** | Reviewer, Panagiotis Petrantonakis | |
| | | |
| | | |

| Glossary | |
|---|---|
| **Abbr./ Acronym** | **Meaning** |
| **AAEON Up Board** | A tiny computer intended for use in robotics and Internet of Things projects. |
| **Actuator** | An actuator is a component of a machine that is responsible for moving and controlling a mechanism or system. |
| **BLE Beacon** | Bluetooth Low Energy Beacon – a beacon that broadcasts information using the Bluetooth Low Energy protocol. Using its identifier and signal strength, the proximity of such beacons to a device can be determined. |
| **HIPI** | Haptic intelligent personalized interface – the goal of SUITCEYES and built as a textile structure. |
| **Inertial Measurement Unit (IMU)** | A unit measuring acceleration and rotation. Normally comprises an accelerometer to detect linear acceleration, gyroscope to detect rotational velocity, and may include a magnetometer to detect rotation relative to the earth's magnetic field. |

| | |
|---|---|
| **RGB-D Camera** | A camera capturing both a conventional (RGB) image, and a depth map such that both 3D spatial information is available. |
| **Raspberry Pi** | A tiny computer made for teaching computer science. Widely used in development projects. |
| **ROS** | Robot Operating System. An operating system that provides methods for capturing and processing sensor information and delivering instructions to actuators. |

# Table of Contents

## Contents

# Executive Summary

This deliverable describes the extension of the initial sensor system outlined in Deliverable 4.1 to incorporate the reasoning ontology from Work Package 3. During the user interviews in Work Package 2, the ability to recognise and locate objects and to navigate unfamiliar environments were determined to be an area in which the proposed Haptic Intelligent Personalised Interface could usefully improve the quality of life of many people with deafblindness. This deliverable addresses this need by exploring the use of the depth camera and bluetooth low energy beacons incorporated into the initial sensor for actively locating and identifying objects in the environment and for navigation of unfamiliar indoor environments. The deliverable is divided into two corresponding parts.

The first part concerns the use of computer vision and bluetooth low energy beacons to support people with visual impairments to identify environmental information. As the computer vision algorithms require more computing power than can reasonably be achieved with the sort of portable mini-computer required to make the system wearable, the RealTime messaging framework incorporated into the system in Deliverable 4.1 is used to send visual and location information to a remote server for analysis by the ontology developed in Work Package 3 and receive responses about objects nearby and in view. When an object to be identified is specified, the system checks to see if a BLE beacon corresponding to that object is in range, and the visual information provided by the depth camera is queried to see if it is in view. The user is then notified if the object is in range but not view, and how far away the object is (so that they can turn to determine if the object comes into view, or move to verify whether the object becomes closer or further away). If the object is in view, then the user is told whether it is towards the left or right of the image captured by the depth camera, and whether it is near or far – this allows the user to begin exploring the area with their hand to attempt to locate the object.

The second part describes the use of the system for indoor navigation using the proximity of BLE beacons in the environment. In this case, maps are defined using a graph structure, allowing the system to infer the current location based on the proximity of beacons, and how this has changed as the system has moved. In this way, the system can convey information about where the user is and also about where they should move next to reach their next direction.

An initial evaluation of both systems was carried out by a single user with deafblindness to assess their potential for the future. The next step is to develop a more comprehensive tactile signaling strategy as part of Deliverable 4.3, as the sensors are integrated into the wider garment.

# Part 1 Active Object Search

## 1. Introduction and Aims

Based upon the user interviews conducted in Work Package 2, it was identified in Deliverable 2.2 that active object search and identification is a common requirement for the daily life activities that many people with Deafblindness want to do on a regular basis. Examples include: “finding lost items in the house” or “Recognizing important things in the house”. The ability is valuable for improving people with deafblindness’ ability to live independently, and such information about objects also assists people with deafblindness to be aware of unfamiliar situations.

Computer vision techniques have been widely used in object recognition assistive systems for visually impaired people (Tapu, Mocanu et al. 2017, Weiss, Luck et al. 2018). This could be further developed for people with deafblindness, by using a proper user interface such as a haptic device. Haptic interfaces can be very effective means of communication for individuals with impaired vision and hearing and essential for people with no vision or hearing. Early work on computer vision based object recognition explored the use of handcrafted visual features (e.g., SIFT and HOG), such as the PalmSight system (Yu, Horvath et al. 2016) or the Third Eye system (Zientara, Lee et al. 2017). Recent advances in machine learning have improved object recognition at a more robust and higher success rate. However, this kind of algorithm usually has a high cost of computation, and is hard to implement on a smartphone or a low cost embedded computer such as the Raspberry Pi 3. In an attempt to alleviate this problem, we applied a real-time communication interface between a local controller and a remote server, as outlined in Deliverable 4.1. This approach can enable the system to run deep learning based visual analysis algorithms (such as those developed in Work Package 3) while retaining system mobility.

Object detection can also be achieved by using Visual Tags (Al-Khalifa 2008), RFID (Nassih, Cherradi et al. 2012) and beacon sensors (Trongwongsa, Chankrachang et al. 2015) such as Bluetooth Low Energy (BLE) beacons. This is a further benefit of the BLE Beacons incorporated into the initial sensor system (Deliverable 4.1), as it can not only provide indoor location data where GPS is not available, but has also been used to develop cognitive assistive systems for visually impaired people’s navigation and situation awareness (Cheraghi, Namboodiri et al. 2017). In this deliverable, we explore its use in our system to extend the detect region of a narrow field of view camera for object searching.

Although there is a large body of studies related to object recognition, few studies address the problem of assisting people with deafblindness to find an object. How to guide the user using haptic information to reach the target is not yet well developed. The purpose of this research primarily concerns increasing the independence of people with deafblindness by improving their active object search ability. We applied real time wireless communication to enable system mobility during object search, and used state of art deep learning based computer vision object recognition developed in Work Package 3 on a stationary computer. In cooperation with the BLE beacon technology, the search region was enlarged. A method was developed to transfer information about the location of the detected object into guidance information using a haptic interface. Our approach is particularly beneficial for guiding people with deafblindness to an object efficiently.

# 2. Related Work

Poggi and Mattoccia (2016) used a custom passive stereo vision based RGBD sensor for a mobility aid system. Detected obstacles could be semantically categorized using machine learning techniques with a smartphone or wireless connection. Both audio and tactile feedback were used. This study focused on the recognition of the closest obstacle to help the user to be aware of the possible hazards on the scene instead of searching for an object.

Pintado, Sanchez et al. (2019) developed a prototype using the Raspberry Pi3 and a convolutional neural network (CNN) based object recognition algorithm to assist visually impaired people to shop. However, no experiments are presented in their paper to validate the system. Furthermore, the system is not able to detect objects.

Lock, Cielniak et al.'s (2019) work emphasized actively guiding the user towards specific points of interest. The Markov Decision Process based algorithm was developed to generate real-time instructions to guide a user towards a target object. They demonstrated this with a smart phone using QR codes to represent objects. However, the guidance region is limited to one image.

The most similar system to our work is DLWV2: Deep Learning-based Wearable Vision-system with feedback developed to guide Blind and Visually Impaired (BVI) people to reach objects (Shih, Chen et al. 2018). The system consists of "Perception" and "Guidance" modules. The "Perception" module utilized a fish-eye camera to detect the object and estimate the distance from the size of the object's 2-D bounding box, the camera was mounted to the right hand wrist. A HTC Vive 3-D Object Tracker was used to trace the camera when the object was out of the view. The detected object was mapped into 5 regions and five corresponding micro vibration motors were actuated in a wrist band, with the duration of vibration period indicating the distance of object related to the hand. User studies were also conducted. The limit of this system is the guidance region is a desk sized region; it is not a mobility system.

# 3. System Design

## 3.1 System Structure Overview

The system built upon the Initial Sensor System described in Deliverable 4.1 Two kinds of sensors are used in for the object recognition elements described in this report: (1) a commercial intel Realsense R200 RGB-D camera and (2) Bluetooth Low Energy (BLE) modules. The data processing is completed in two kinds of processors: (1) a local controller Raspberry Pi3 and (2) a station computer as a remote server. The feedback interface is a vibro-tactile belt consisting of 6 micro vibration motors.

The local controller receives point cloud data, depth images and RGB images from the RGB-D camera and obtains the Received Signal Strength Indication (RSSI) from the BLE module. These data are processed in “Image processing”, “Proximity detection”, “Object location” and “Haptic control” nodes. These nodes communicate using the Robot Operation System (ROS) in the local controller. The utilization of ROS enables us to design system modularly and insert other functions in the future. We do not process “object and people detection and recognition (visual analysis)” in the local controller instead of sending RGB-D images to a remote server in real time, this configuration helps us make use of the server computation resource to process state of the art deep learning based image analysis algorithms. We estimate the location of a detected object related to the user by using registered depth images. Between the local controller and remote server, a real time communication interface is implemented and convert the ROS format message into Json format message which is more popularly used in the web-based application.

An ontology engine plays the role of a “brain” in our system to infer the actions that the user should take by fusing the sensor information. The generated action command or object location are mapped to a vibration pattern on the waist to guide the user. The flow diagram of the system is shown in Figure 1. In the following, we introduce our sensor data processing pipeline development, “Real time communication” and “Ontology engine” modules.

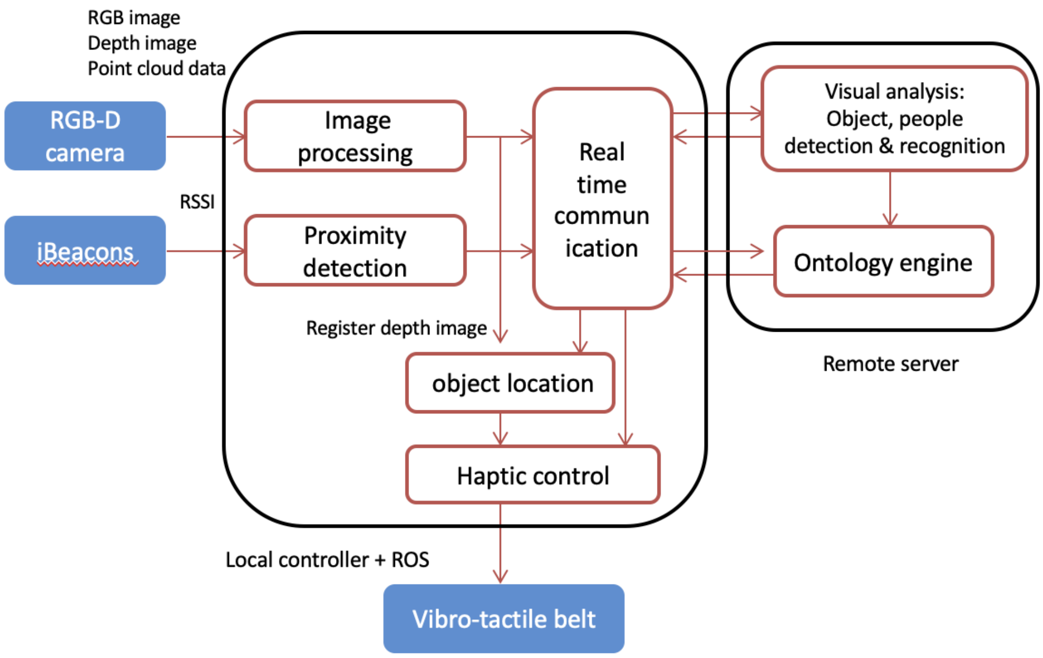



Fig 1: Overview of system structure

## 3.2 Proximity Detection

A BLE Beacon is used as a proximity detection sensor. A BLE Beacon attached to an object can be detected by receiving a message when in proximity to it. In this case the beacons use the iBeacon protocol developed by Apple, in which each beacon needs to be allocated a Universally Unique Identified (UUID), and Major and Minor identifiers before use and this information is contained in the message and recorded in a database (on the Raspberry Pi3 in this study) related to their represented object name.

Our system consists of “BLE scan”, “Signal filter”, “Distance estimation” and “Proximity detection” modules. “BLE scan” detects all BLE devices in the environment, the detect region can reach 100 meters without shelters. “Signal filter” only keeps those BLE modules whose address meets the iBeacon message format, and smooths the signal using an exponential moving average method. “Distance estimation” depends on the Received Signal Strength Indication (RSSI) values from the iBeacon, which are highly fluctuating and influenced by implementation and environmental conditions. We applied a Log-Normal Shadowing Model (LNSM) to fit the RSSI data and conducted this calibration process before use. Then the controller queries a database installed in the controller (Raspberry Pi3) to map the received address to an object. The outputs of proximity detection are objects and their proximity distance. The proximity distance is usually categorized into 3 different ranges: Immediate (within 2 meters), Near (from 2-5 meters), Far (from 5-10 meters).

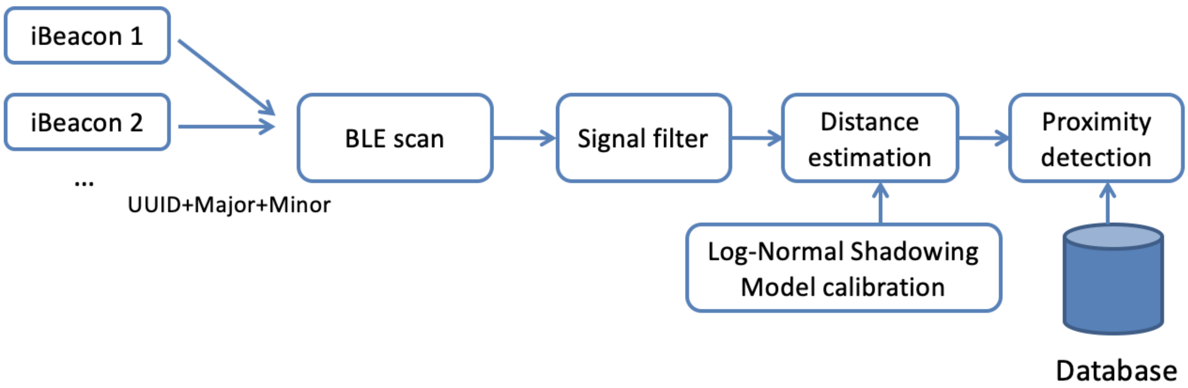



Fig 2: Proximity detection process

## 3.3 Object and People Detection and Tracking (WP3)

This part of work was developed in Work Package 3. Visual analysis constitutes a core component in SUITCEYES. In order to extract knowledge using the visual information that a wearable camera can capture, several computer vision tasks are carried out within the component, such as object detection and tracking; face detection and tracking. All tasks have been chosen and designed to work in line with SUITCEYES principals and intentions to support the users with deafblindness, and are generally accomplished using state-of-the-art, robust, and efficient computer vision algorithms, i.e. deep convolutional neural networks.

## 3.4 Object Location

If an object is identified by the "object and people detection and recognition" module, and an appropriate bounding box established, it is possible to identify the distance of the object from the user by projecting its bounding box in the RGB image to the depth image, to estimate the distance and orientation of the object. The RealSense is capable of measuring targets from the range between 1.0–2.5 m reliably. This is the range in which object recognition is most likely to be needed, so this fits well for our active object search task.

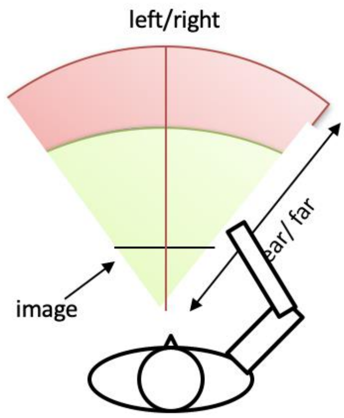



Fig 3: Location of an object in four regions in front of the user using depth image

With separation of image into "left" and "right", and estimated distance "near" and "far", we create four regions in front of the user and locate the object in these regions, the reduced location information of object can be represented by haptic signals. This information can assist the user in reaching the object.

## 3.5 Active Object Search Function

Figure 4 illustrates the flow chart of the active object search function. An object is specified by the user and is processed by the object and people detection and recognition module firstly to check if the object has been detected in the captured image. If not, the iBeacon database will be searched to check if the object is attached to an iBeacon sensor. If the object is listed in the database, the proximity detection will provide information about the object's distance form the system. If the distance is immediate, this means that the object is around user, but out of range of the camera, so the user could rotate his or her body until the camera detects the specified object. If the object's distance is near or far, it will require the user to move in order to search for the object, and inform the user whether they are moving away from or closer to the object. The user can use an ascent match strategy or follow a touch path with walls in the room to continue searching for the object until the iBeacon shows it is immediate. This algorithm has been implemented in an ontology engine in the remote server (as part of WP3).

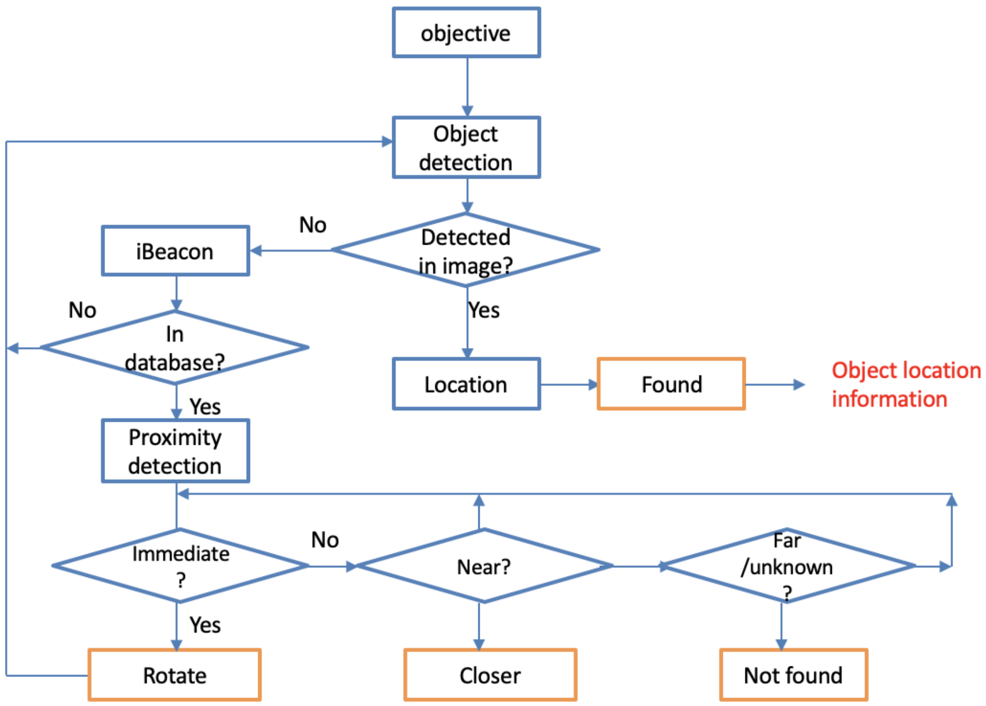



Fig 4: Flow chart of the active object search function process

## 3.6 Real-time Communication

As described in Deliverable 4.1, the Realtime Framework (https://framework.realtime.co/messaging/) is used to build our cross-platform system that requires real-time communication between a local controller and a remote service. We developed a message converter to integrate Realtime Frames Json message with the ROS message successfully. Images are uploaded to server using package CPR (https://github.com/whoshuu/cpr) which is a simple wrapper around libcurl inspired by the Python Requests project.

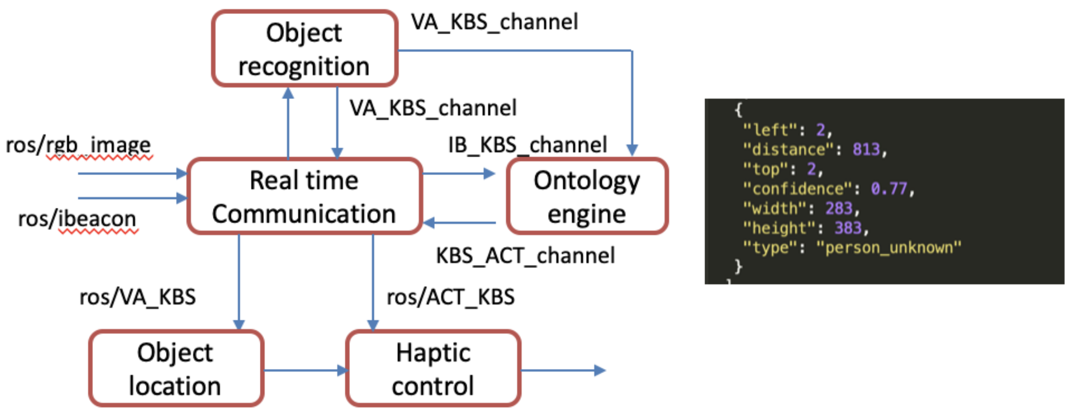



Fig 5: Message flow communication between different modules

Figure 5 shows the message flow communication between different modules in our cross-platform system. “ros/” means ROS format message and “_channel” are channel names to listen messages in RealTime Framework. An example of Json format message published in “VA_KBS_channel” is shown in the same figure. There are three main message pipelines:

(1) RGB images “ros/rgb_image” are sent to the “Object recognition” module, and the processed data are published in “VA_KBS_channel”. Messages on the “VA_KBS_channel” are converted into ROS type message “ros/VA_KBS” to ROS node “Object location”.
(2) iBeaocn messages “ros/ibeacons” are passed to the “Ontology engine” using channel “IB_KBS_channel”.
(3) Based on the results from “Object recognition” and “Proximity detection”, Ontology engine sends action commends to the local controller by using channel “KBS_ACT_channel”. Action messages are converted into ROS type message “ros/ACT_KBS” to ROS node “Haptic control”.

## 3.7 Haptic Feedback Interface

We used a custom vibro-tactile belt as haptic feedback interface, which consists of 6 vibration motors around the user waist. This was designed to represent different information as shown in Figure 6. When the vibro-tactile belt is mute or three back motors vibrate simultaneously, this means that the object is not in the range of the camera and it may be detected by iBeacon “unknown”, “far” or “near”. The user should move and search until the “immediate” message appears (three front motors work), that suggests that the user should rotate his/her body to allow the camera to capture the object. Once the object was detected in the image, all vibration motors are actuated and it follows a “location mode” message, that represents “Orientation” and “Distance” using side motors (orientation) and intensity of motor vibration (distance).

| | Ontology | Message | Viro-tacile belt motors work |
|---|---|---|---|
| Search mode | VA not find<br>PD not find or far | Not find | Mute |
| | VA not find,<br>PD find near | Closer | Three back motors |
| | VA not find,<br>PD find immediate | Rotate | Three front motors |
| | VA find<br>PD find not far | Find | All |
| location mode | Orientation | | Right motor<br>Left motor |
| | Distance | | Intensity of motor vibration |
| VA: Visual analysis, PD: Proximity Detection | | | |

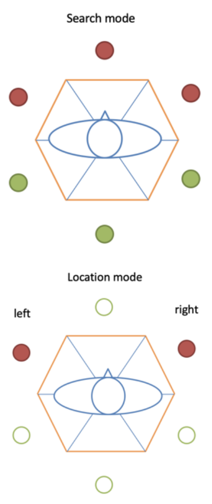



Fig 6: Vibration pattern of the vibro-tactile belt

# 4. Pilot Experiment Results

We conducted several pilot experiments using the system as shown in Figure 7. Realsense R200 RGB-D camera and Bluetooth iBeacon were used in the experiments. The resolution of the RGB-B image is 480x640 and the scan rate is 30 fps. Several DSD TECH HM-10 BLE iBeacon were attached to the objects for proximity detection. The thresholds of the iBeacon distance classification were adjusted: Immediate (within 1.5 meters), Near (from 1.5-3.5 meters), Far (from 3.5-10.0 meters) to meet the search region in a room (~4x6 m2). A customized vibro-tactile belt with 6 micro vibration motors is used as a haptic feedback interface. All controller, motor driver and battery are packed in a portable bag. These experiments are carried out in the Affective Engineering lab at the University of Leeds.

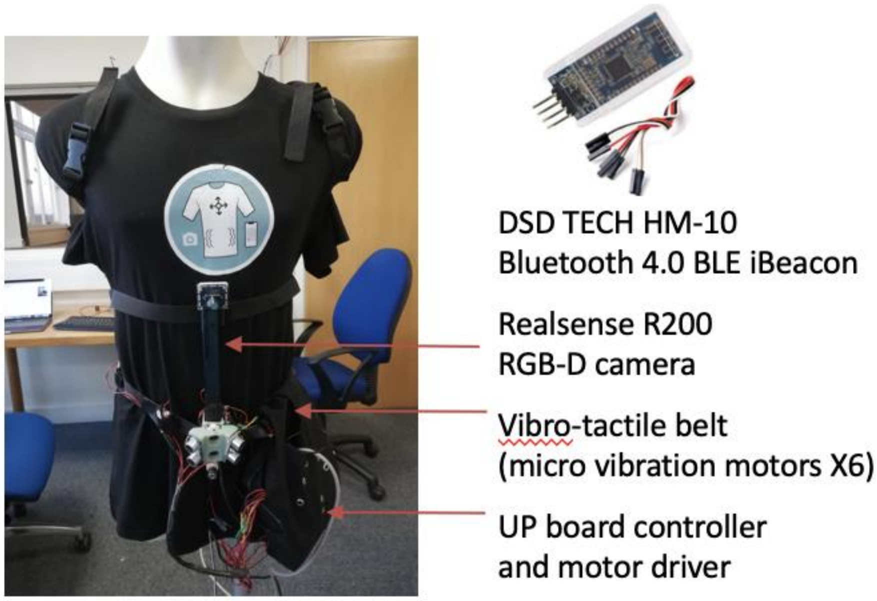



Fig 7: Initial sensor system developed in D4.1

## 4.1 Realtime Communication and Location Estimation Validation

In the first test, we validated the performance of real-time communication when visual analysis and proximity detection were working and evaluated the accuracy of object location estimation. The detected objects are used to show that these results could feedback to the local controller in real-time. The algorithm from WP3 has been incorporated into processing algorithms. All of these images are recorded in Leeds.

We placed several objects in front of the sensor system: a table, a chair, a laptop, a mug, a computer and a phone. The images are uploaded to the server in CERTH and the detected results are sent back to Leeds in real-time. Figure 8 shows the results of this process: a bounding box on the RGB image has been specified from visual analysis, and the distance for that bounding box is identified from the corresponding points on the registered depth map. The center position of the bounding box in the image was classified into left and right.


Fig 8: Results of static object detection. Images were captured in the local computer in Leeds, the image color changed due to image display format

The recorded average scan rate of object detection loop (update image, detect objects, send back message) can reach 1Hz. Figure 8 shows instances of object detection from different distances. Larger objects are detected from large distances whereas for smaller objects the camera has to be closer to them. This is mainly due to the low resolution of the camera. If the resolution is low, as is the case in our system, i.e., 480X640 pixels, very few pixels represent a small object (i.e. a mug) when “seen” from a large distance. Small objects like “mug”, “computer” “cell phone” and “laptop” are successfully detected from distances <1.6m, which is enough for the user to reach the object. In all cases the different objects were correctly classified. The experiment verifies also the reliable real-time communication interface.

The estimated distance and orientation are also shown in the figure 8. Generally, system can estimate the distance correctly, however, it failed in some cases, for example, it estimated a wrong distance of “mug_cup” at 0.7 m, this could be caused by over-saturation of the image by the active IR illumination of RGB-D camera, and results in a low resolution of the depth image. On the other hand, we estimated the distance using all values in the bounding box region, this leads to some background noise. For example, the table bounding box also includes the part of the chair, this results in a short distance value. At a far distance (>2.5m), the distance measurement is out of the range of the depth camera capability. Distance estimation can be further improved by comparing the size of the bounding box. The closer the camera is to an object, the larger the size of the object is to be captured in the image.

Four iBeaons were attached to the objects “laptop”, ”mug_cup”, ”table” and ”chair”. The proximity detection results are also shown in the figure 8. It can be seen the estimated distance changed from ‘’near” to “immediate”, when the sensor system was approaching to the objects. A change of message alternatively between two modes “immediate” and “near “could be observed in the range of 1.3-2 m, this could be caused by highly fluctuating RSSI signals in the boundary. However, when small objects are out of the sight of camera, iBeacons still can detect them and show these objects are near the user.

.

## 4.2 Active Object Search Test

The second test was conducted in a large room space (~6mX4m) and both camera and beacon sensors are applied. The purpose is to validate the system dynamically and test the ontology logic presented in section 3.5. The experimental process is illustrated in Figure 9. An object was placed randomly in the room and a participant was asked to search for it according to the haptic signals that he received. We did not consider obstacles avoidance in this test. In order to ensure the safety of participant, we emptied the room, and an assistant followed the participant to provide protection for the participant. The participant had a short training to be familiar with the vibro-tactile signals.

During the test, we set the target as “table” and asked a sighted and hearing participant to close eye and stand far from it in the beginning. Then he moved along the wall by touching the wall, the back three vibration motors were actuated when iBeacon detected the object “near”. Participant continued to move forward. Near the corner, iBeacon detected the object “immediate” and the front three vibration motors were running, but camera did not face to the object, system delivered a command” rotate” to require user to rotate the body until the camera can detect the object, then completed the test. Figure10 shows a sequence of snapshots showing the participant in different modes. Currently, we are not able to record the trajectory of the participant’s movements, so that we cannot compare the effects of the system quantitatively.

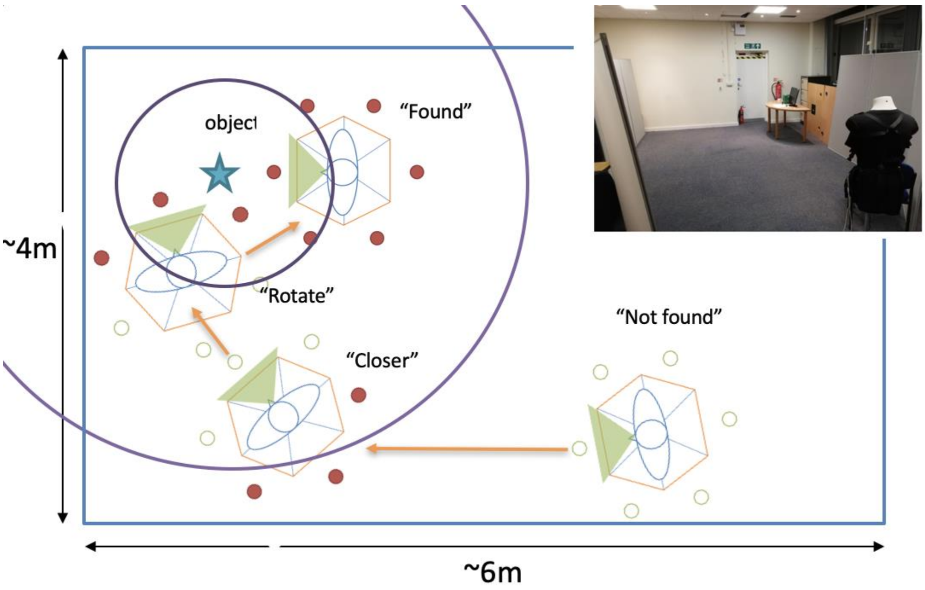



Fig 9: Illustration of the setup of active object search experiment


Fig 10: A sequence of snapshots of the participant in different modes

Several issues are found during the experiment:

(1) The delay of iBeacon signals. This is due to the slow change of received signal strength index (RSSI), and an exponential moving average signal filter which records the previous time data. The filter parameters need to be optimized in the future.

(2) The highly fluctuating RSSI signals caused the change alternatively between two modes when the user is on the boundary, such as a jump between "rotate" and "closer". This could be improved by adding more information for logic inference, such as whether participant moving information using the inertial measurement unit included in the sensors system (as described in Deliverable 4.1).

## 4.3 Initial User Study

We carried out an initial user study with one participant with deafblindness, who uses cochlear implants, during a workshop at a consortium meeting on 5$^{th}$ December 2019 in Offenburg, Germany. For the test, the user wore the sensor system prototype and searched for a laptop on a table. Since the haptic feedback interface was not available at that time, we described the commands generated by the sensor system to the user orally. We also offered some explanation of these commands as this was the first time for the user to use it. Based on this information, user was able to find the laptop, a sequence of snapshots from the sensor system during the user searching for the laptop was shown in Figure 11. Some suggestions were given by the user during the workshop:

(1) A clearer description of the command, such as "closer" could be represented by the intensity of vibration motor
(2) Make the sensor system adjustable so it is suitable for different body shapes and can be wore easily, this will be addressed in the D4.3 integrating sensors into garment
(3) Install a camera on the head for easy rotation and searching for objects.


Fig 11: A sequence of snapshots of the user searching for the laptop

# 5. Conclusion

In this part of the report, we described the development process of active object search function by incorporating object recognition into the sensor array. The characteristics of this system are:

(1) Realtime communication between a portable controller and a remote server, which enables the system to run state of the art deep learning-based computer vision developed by WP3

(2) Distance and orientation estimation of detected objects in a certain range

(3) Inference of search actions based on visual analysis and proximity detection information to guide the user to reach a target

(4) All sensors, haptic feedback interface (vibro-tactile belt) and controller are packed for mobility

We conducted a series of pilot experiments to validate the above system functions, and understand the capability of different sensors that detect objects at different distances.

# Part 2 Indoor Navigation

## 6. Introduction and Aims

Indoor navigation is important for people with deafblindness to navigate, explore and access unfamiliar public spaces such as supermarkets, hospitals and transports (Dias, Teves et al. 2015). Furthermore, most O&M (Orientation and Mobility) training progresses from the easily controlled indoors environment to the more complex outdoors environment (Goldschmidt 2018). Therefore, an indoor navigation system was chosen for development for deaf blind people in this study.

From knowledge of cognitive psychology studies for visually impaired people and people with deafblindness (Loomis, Klatzky et al. 1993, Golledge, Klatzky et al. 1996, Golledge 1999, Kalia, Schrater et al. 2013, Velázquez Guerrero 2017, Jeamwatthanachai, Wald et al. 2019), we identified three kinds of information that an assistive navigation system should provide to deafblind users: (1) a sequence of segmented line paths represented by “turn by turn direction” at specific choice points for path integration, (2) landmarks or cues to update spatial environment and help people with deafblindness to create a “mental map” and “cognitive anchor points”. (3) provide people with deafblindness with “confirmation cues” during travel using orientation and path trace on the “mental map”.

According to this knowledge, we designed a navigation system based on a topological map to replace metric high precision maps. A behavioral approach is applied to guide people with a series of acts. We build a graphic map from floor plans, with each node representing elements of the environment such as rooms, corridors or halls. Then, the navigation route was generated using a graphic searching method. Navigational behaviors were inserted between two nodes to guide users from one place to another by using egocentric coordinate, such as “enter/leave a room”, “turn right/left” and “go straight”. Detailed turn by turn direction information were included in these behaviors.

Path integration navigation strategy may be not sufficient for people with deafblindness to comprehend the environment (Golledge, Klatzky et al. 1996). Landmarks and cues are used by blind people to establish and maintain orientation. However, this landmark information is often not available to people with deafblindness in accessible formats (Hersh 2016). People with deafblindness may only use haptic perception; however, tactile information has a small perception region (Bradley and Dunlop 2005). Therefore, we proposed to develop “tactile landmarks” which are represented by haptic feedback using sensor input information such as a camera or beacon sensors, to help people with deafblindness with building a cognitive map and anchor points. Several types of landmarks are identified from related literature and are grouped into “Architecture”,” Function”, “Objects”.

We used Low Energy Bluetooth beacons to deliver “turn by turn” information between two structural nodes and landmark signals in the real environment. A goal-oriented action plan and execute method was applied to monitor the implementation of these behavior actions. Since the haptic interface will be integrated into the whole system in D4.3, we would not discuss haptic interface in this report.

# 7. Related Work

Current research on indoor navigation systems mainly focuses on indoor positioning. GPS (Global Positioning System) has been widely used in outdoor environments, but it cannot provide accurate temporal and spatial estimates indoors. Early work on alternatives to GPS for indoor environments have explored using specialized beacons: for example, ActiveBadge (Want, Hopper et al. 1992) uses infra-red beacons; Cricket (Priyantha, Chakraborty et al. 2000) uses a combination of radio frequency and ultrasound signals. Nowadays, Indoor positioning has focused more on methods that rely on existing short-to-mid range communication techniques between mobile devices (Bluetooth), wireless local area network (WLAN) (Fang and Lin 2010) and wireless sensor network (Zigbee) (Luoh 2014). iBeacon proposed by Apple Inc attracts our attention due to its ability to provide a solution for context-awareness with low energy consumption, low cost and large area applications. It also has been used to develop cognitive assistive systems for visually impaired people. iBeacon can deliver landmark information allowing users to be aware of the environment they are in, for example, at street crossings (Liao and Rakauskas 2010). Furthermore, iBeacon can also be used to develop indoor location systems. Two iBeacon based indoor navigation systems have been demonstrated to guide visual impaired people in a large region.

“NavCog” is developed by researchers in CMU and IBM (Ahmetovic, Gleason et al. 2016). The “turn by turn” information is delivered online after locating the user in a map using a “fingerprinting” method by measuring signal strengths at different locations on the map. This successfully demonstrated the system in a large-scale space in the CMU campus and in a shopping mall. However, the first issue with “fingerprinting” method is that it requires a large effort to calibrate “fingerprinting”. In order to solve this issue, researchers developed a robot machine to collect “fingerprinting“ and this machine was located using a LIDAR senor. However, the accuracy of “fingerprinting” also could be decreased by the position of the receiver’s position handhold or pocket (e.g. smart phone). By leveraging its accurate localization capabilities, in the following study, the presence of semantic features and points of interest in the vicinity (e.g., doorways, shops) are added to improve user’s spatial awareness of environment (Sato, Oh et al. 2017, Guerreiro, Ohn-Bar et al. 2018). The second issue are missed turns where the user turns too early or too late. Researchers discovered that participants were more precise in performing “ample turns” (60◦–120◦) than “slight turns” (about 22.5◦– 60◦) (Ahmetovic, Oh et al. 2018).

“PERCEPT” is developed at University of Massachusetts Amherst (Ganz, Schafer et al. 2012), and detailed landmark-based navigation instructions to the selected destination are provided to the BVI user. A simple “Log-distance path loss propagation” model based location approached was used(Yang and Ganz 2018). This system is successfully deployed in in a large venue such as North train station, although location, moving direction and landmark proximity computation are not very accurate(Ganz, Schafer et al. 2018).

Both studies used “location system + metric maps” strategy and a large number of iBeacons were used to improve precision of positioning. (218 Bluetooth beacons positioned every 5-10m in “NavCog” and 255 Bluetooth beacons in ”PERCEPT” project), and landmarks are emphasized as being important for providing users with information of their surroundings ,as well as its ability to leverage the error caused by position and direction inaccuracy.

# 8. Indoor Navigation System Design

## 8.1 Learning from Cognitive Psychology Studies for Deafblind People's Navigation

In order to design an adaptable assistive navigation system for people with deafblindness, we reviewed literature from cognitive and developmental psychology and our user studies (Work Package 2) to identify the suitable information that should be provided to people with deafblindness by the navigation system.

Golledge, Klatzky et al. 1996 discussed both the cognitive mapping process and wayfinding skills by adults without vision. They confirmed that there was no significant difference in spatial task performance between those with and without vision under some circumstances, even for congenitally blind people. According to their study, basic components of wayfinding by blind individuals include: representing current location, progressing on a target route, spatial updating, updating by path integration and computing a novel path. During progressing on a target route, blind people would often follow linear features "shorelining" in the environment such as the edge of a sidewalk, a fence, a curb, or the walls or hallways of a building. Blind travelers usually segment a route by focusing on key points and account them and also remember turn angles. However, this form of wayfinding strategy does not help to develop an understanding of the environment and a capability to search alternative routes. Spatial updating capability is an important component for wayfinding. This can be undertaken by recognizing landmarks, mentally retracing a path by reconstructing previous movements, object-to-object strategy (i.e. anchor points (Couclelis, Golledge et al. 1987)). Two different types of information are needed for navigation aids: proximal environment for obstacle avoidance and larger scale geographic space. This information can be represented and stored in one's cognitive map (Golledge 1999). Learning spatial layouts and remembering specific paths through them are two difficulties for many blind travelers and prevent them from traveling to unfamiliar routes. Critical landmark cues used as anchor points for the cognitive representation of the environment are preferred instead.

Pissaloux and Velázquez 2018 concluded that path integration, landmark-based strategy and geometry-based strategy were three main human mobility strategies, and emphasized the importance of visual cues, clues and landmarks for sighted people navigation. As far as visually impaired people's mobility is concerned, the main challenge is how to adapt mobility models and navigation strategies and assist visually impaired people to understand and explore the space with their perceptual capabilities. Authors comment that a successful locomotion requires several elements: (1) path integration and space integration and drift correction; (2) data on landmark and obstacles acquisition; (3) orientation and path trace building; (4) socio-urban data perception and space integration.

Hersh 2016 interviewed 28 people with deafblindness in six different countries and studied their travel behavior. She found that the use of tactile information was particularly important for them during their travel. Some show good capability of spatial understanding and representation skills; however, spatial information is often not available to them in accessible formats to help them understand environment. People with deafblindness use a number of different techniques to represent space: "mental map", "Physical memories", and tactile maps. A wide variety of different types of landmark are used for people with deafblindness depending on their available senses, the nature of the environment and personal factors. After they lose vision, they may prefer tactile cues, such as lampposts, curbs and different surfaces. Hersh suggested "*there is a need for an international standard for accessible crossing points including vibrotactile indicators, tactile arrows to facilitate alignment with the crossing direction, and tactile tiles to indicate the best straight path across the road. This would facilitate long-distance and tourism travel by deafblind people, as well as encouraging best practice*"

We also learned the navigation strategies and issues of people with deafblindness from our user study interviews with such individuals, conducted by the Work Package 2 team. The need for more accurate and detailed information about the environment was a very common theme in the interviews. By way of example, one of the participants described her/his travel navigation experience as follows:

"*I once went to Amsterdam to the Dovenontmoetingscentrum [meeting centre for people with hearing impairments], but I did not know exactly where it was. In advance, I studied which tram to take, where to get off the tram, etc. But it was in the dark, so I had to remember everything, how far to walk, where to go to the left, where to cross the street, to the right. But the dog walked on the bike lane, because for some reason, there was a long queue of people waiting in front of a supermarket. So he went around it. So then I asked, could you please make some room for us, and that is what they did. Then we had to pass a very long street and then we stopped. Then I thought that we had arrived, but in the dark, I could not see the number. I think I had to be at number 98. I asked someone for number 98, and then I found out that we already had passed it. In such cases, more accurate navigation would be helpful. Also in the station, it sometimes doesn't work. For example, in Rotterdam or Amsterdam, when you walk to the train or the metro, I sometime loose mobile phone connection, and then I do not know where I am. So then I have to ask someone where is the station?*"

-Hanneke, Netherlands

Here It can be seen that path integration was the main approach that was used and this was common to participants. However, the remembered path could be interrupted by a change in the environment such as a queue on the street, and a new path had to be planned. Confirmation information such as "arrive at destination" was required during the navigation. In an indoor environment such as a train station, the mobile phone signal could be lost and the GPS based navigation system may fail to work.

From both the literature reviews and interviews, we can learn that path integration and landmark, cues are used by people with deafblindness to follow a routine and understand the environment. For some, haptics may be the only effective way for them to interact with environment. We use this knowledge from cognitive psychology and interviews to guide our design of an assistive navigation system. We consider the following to be the most important three kinds of information that the assistive navigation system should deliver to people with deafblindness:

(1) a sequence of segmented line paths represented by "turn by turn direction" at specific choice points for path integration;

(2) landmarks or cues to update spatial environment and help people with deafblindness to create a "mental map" and "cognitive anchor points";

(3) locate a position on the "mental map" and provide people with deafblindness with confirmation cues during travel.

In order to cover a large population of people with deafblindness, all this information should be represented by use of tactile formats. The next section will introduce the navigation system structure and its components.

## 8.2 Navigation System Structure

Our navigation system consists of two pipelines, which are intended to generate two kinds of different information, as identified in the previous section:

The first pipeline is a behavior navigator including "Graphic map building", "Route planner" and "Task manager", the output is a sequence of "turn by turn" directions such as "go straight", "turn right" to guide the user from one waypoint to the next one and complete the path.

The second pipeline is tactile landmarks. It includes selection, detection and representation of landmarks. This pipeline generates information consisting of some features of the environment such as "elevator", "stairs" to assist the user to build "anchor points" in the mental map.

Ideally, we should have the third pipeline that will track the position of the user to produce "confirmation cues" and alert the user if they are on a wrong path. This will be content of a subsequent deliverable, we will leave this part in the future work to discuss, and focus on the behavior navigator and landmarks pipelines in this report.

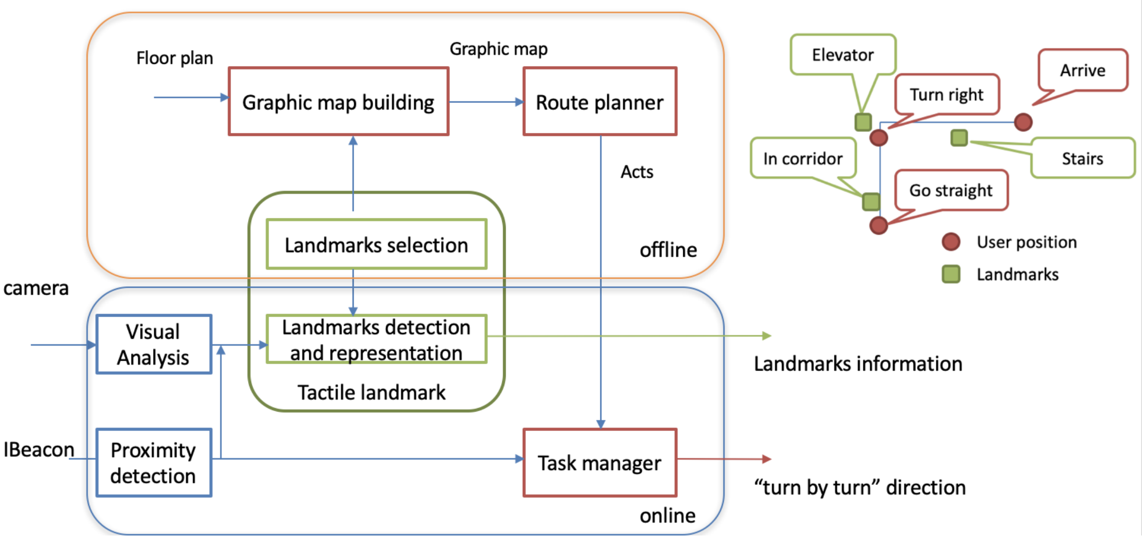



Fig 1: Overview of navigation system and example of navigation information

## 8.3 Behaviour Navigation

Most navigation systems rely on a precise real time location system and map generate “turn by turn” information to users such as “NavCog”(Ahmetovic, Gleason et al. 2016) and “PERCEPT”(Ganz, Schafer et al. 2018). The location of the user is detected by a location system (such as a GPS), and is explicitly shown on a metric map (a geometric model of the word). Then a route can be planned using search algorithms. Turn by turn acts are generated to guide the user to arrive at the destination.

However, an accurate indoor position system is not yet well developed on the market. In our previous test using iBeacons with “Log-distance path loss propagation” model method, the accuracy of the system was estimated on the level of 2.5 m, assuming there is no interference. Because the model is affected by multipath losses or shadowing due to walls or people’s walking. Usually, fingerprinting can achieve high positioning accuracy compared to other methods (~1.3m error in NavCog). However, fingerprinting methods require a high burden of calibration (training process) to build a fingerprint database manually (Gleason, Ahmetovic et al. 2017). Furthermore, this process must be repeated when there is a significant change in the target environment.

However, humans do not use precise location information to navigate indoor environments. Sepulveda, Niebles et al. 2018 applied a behavior approach to indoor autonomous navigation for a robot using deep learning. They encoded an indoor environment as a graph with semantic location as nodes and navigational behaviors as edges. Those navigational behaviors can be considered as a map from visual inputs to produce motor controls. This approach enables the robot to “*avoid explicit computation of its precise location or the geometry of the environment, and enables navigation at a higher level of semantic abstraction*”.

Inspired by this kind of behavior approach (Werner, Krieg-Brückner et al. 2000), we replaced a metric map with a graphic map to represent the indoor environment. This graphic map concept is similar to the cognitive map that people constructed their environment (Jefferies and Yeap 2008). We built a graphic map from floor plans, with each node representing elements of environment structure such as rooms, corridors or halls. Then the navigation route was generated using graphic searching method, navigational behaviors are inserted between two nodes to guide the user from one place to another using egocentric coordinates, such as enter/leave an office, turn right/left and go straight, detailed turn by turn direction information are included in these behaviors. Compared to navigation using a metric map based, this method offers two advantages: (1) It is robust to localization errors, and can even work without a location system; (2) the behavior description of navigation is easier for integrating more functions in the system. In the following sections, we will explain how to build a graphic map from a floor plan.

## 8.4 Graphic map building

In order to illustrate the process of building a graphic map from a floor plan, we took the fourth floor of the Mechanical engineering building at University of Leeds as an example. This is shown in Figure 2. Six structural elements of the environment could be selected and defined from this floor plan, namely “room (R)”, “hall (H)”, “corridor (C)”, ”elevator (E)”, ”stairs (S)”, “toilet (W)”. A guidance for constructing a topological representation of large environments from labeled floor plan could be found in (Whiting 2006). We labeled the same elements to discriminate between them. For instance, “460-R”,” 461-R”. An open source annotation tool (Aydemir, Jensfelt et al. 2012) was used to implement this design. The input is a “.png” format floor plan and the output is a “.xml” format file.

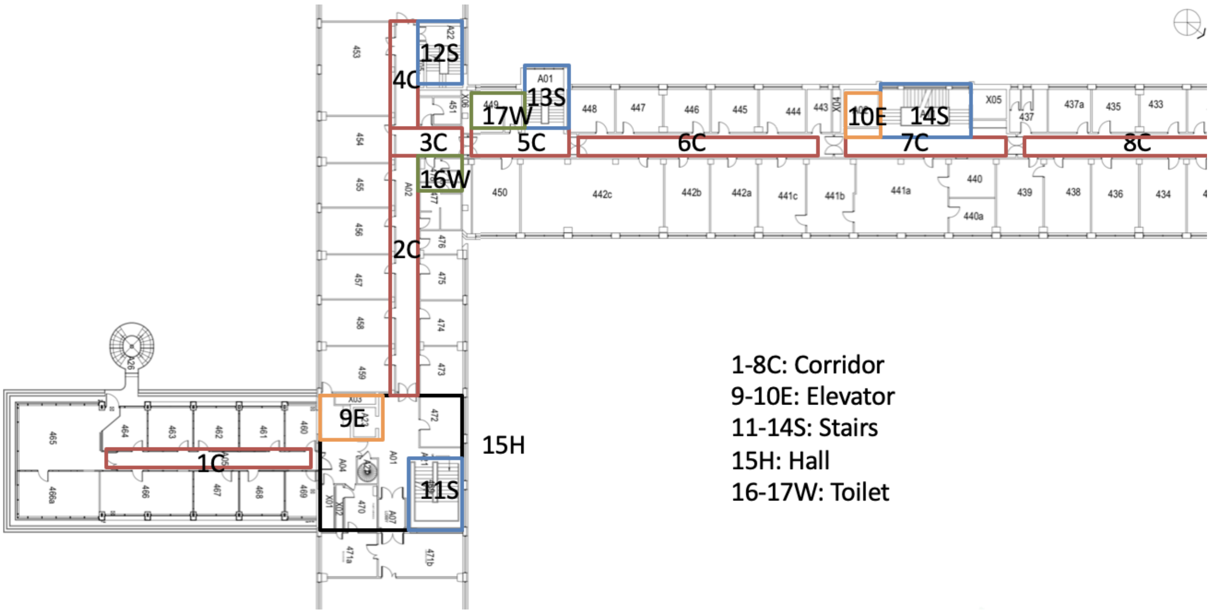



Fig 2: Floor plan of the fourth floor of the Mechanical Engineering building at University of Leeds and selected structural elements

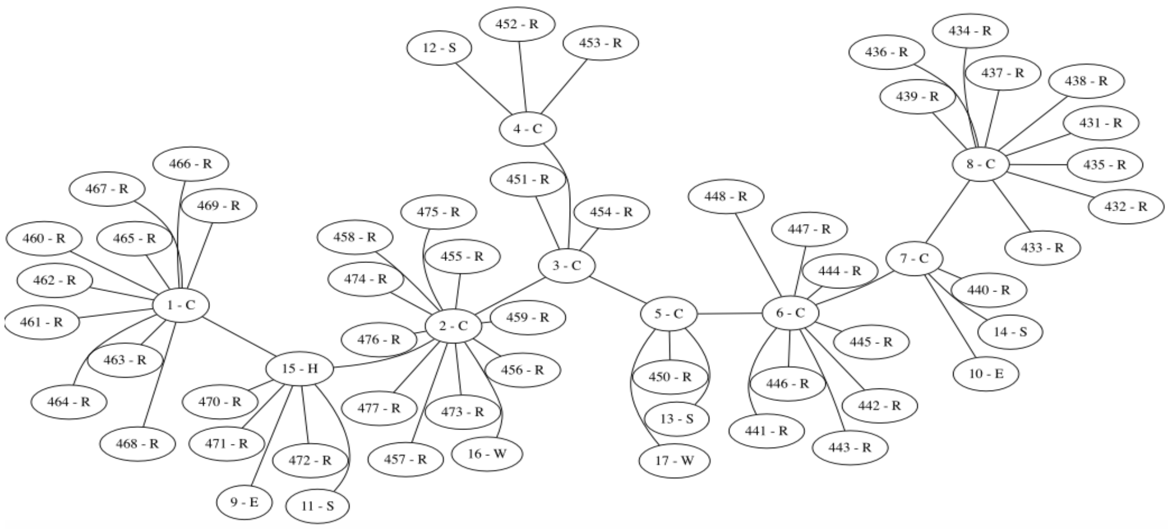



Fig 3: Graphic map representation of the floor plan in Fig 2

## 8.5 Tactile Landmark

A landmark can be defined as an object or location external to the observer, which has key characteristics that make them recognizable, memorable in the environment and in serves to define the location of other objects or region (Millonig and Schechtner 2007, Ishikawa and Nakamura 2012). In our navigation system, we expect that landmarks can help people with deafblindness to:

(1) structure the whole spatial information during cognitive mapping;
(2) establish and maintain orientation; and
(3) allow them to confirm the progress of their travel when they repeated the route

The landmark module requires the selection of a set of appropriated landmarks, and the use of appropriate techniques to detect and represent them in a haptic form. In the next two sections, we explain how this happens.

### 8.5.1 Landmarks selection

Sorrows and Hirtle 1999 developed a comprehensive framework of landmarks for real and electronic spaces. They considered singularity, prominence and content as important features of the landmark, and proposed three categories of landmarks: visual, cognitive, and structural. Ohm, Müller et al. (2014) assessed the pedestrian visual attraction using an eye tracker and selected landmarks in large scale indoor environments and classified indoor landmarks into four groups: “Architecture”, “Function”, ”Information”, “Furniture”. Gangaputra (2017) conducted a case study for the indoor landmark identification issue. From the questionnaire feedback, it was found that most of the participants could differentiate between an indoor and an outdoor landmark and indoor landmark could be classified into three different types such as figural level, vista level and environmental level. Fellner, Huang et al. (2017) developed landmark-based route instructions for indoor navigation, an approach identification of landmark categories was developed for landmark extraction and weighing, which enables a landmark selection algorithm. Landmark are scored based on spatial feature categories. The landmarks selected in these studies here are listed in Table 1 and we adopted similar Müller’s method to calssify them. All these studies were conducted with sighted people.

There are two studies that mentioned the indoor landmarks that visual impaired people often used during navigation. Dias, Teves et al. (2015) highlighted the use of landmarks and environmental clues to inform visually impaired people in indoor navigation and provided a list of environmental clues and landmarks used by visually impaired people when navigating indoor spaces. The latest version of “NavCog3” provides information about nearby landmarks and point of interests to improve users’ walk comfort and confidence (Sato, Oh et al. 2017).

Table 1: Indoor landmarks selection for people with deafblindness based on lit. review

| | Müller et al. (2014) | Gangaputra (2017) | Fellner, Huang et al. (2017) | Dias, Teves et al. (2015) | Sato, Oh et al. 2017 | Selected in this study |
|---|---|---|---|---|---|---|
| Architecture | Pillars<br>Fronts | | | Intersections<br>Turns<br>Corners<br>Walls | Tactile paving<br>Ramp<br>Step<br>slope | T Intersections<br>Cross Intersections<br>Corners |
| Function | Doors<br>Stairs<br>Elevators | Doors with door numbers or names<br>Staircases<br>Toilet<br>Elevators | Stairs(1),<br>Entrance/Exit(2)<br>Elevator(4)<br>Toilet(5)<br>Auditorium (7)<br>Meeting room(9)<br>Seminar room(10)<br>Bridge(11)<br>Door(12)<br>PC room(13)<br>Project room(14)<br>Lounge(15)<br>Front office(16)<br>Library self-checkout(17)<br>Study area(19)<br>Library search terminal(20) | Doors and doorways<br>Stairwells<br>Elevators<br>Water fountains<br>Information or front desks | Door type of doorway (automatic, open or close)<br>Elevator<br>Escalator<br>Stairs (straight, u-shape) | Doors (close/open)<br>Stairs<br>Elevators<br>Toilets |
| Information | Signs<br>Posters | Information board | | | | |
| Object (Furniture) | Tables<br>Chairs<br>Benches<br>Vending Machines | Trash boxes<br>Fire extinguishers<br>Paintings<br>Snack automat machines | Locker(3)<br>Vending machine(6)<br>Scanner(18)<br>Computer terminal(8) | Tables<br>Vending machines<br>Fire alarm<br>Bulletin boards | Trash can | Tables<br>Chairs<br>Vending machines<br>Trash boxes<br>Bulletin boards |

All typical indoor landmarks we found in the literature are listed in Table 1. The last column reports the landmarks selected in this study. In the third column, numbers are the rank of landmarks. The derived indoor landmarks in this study are examples of landmarks in a university building. However, landmarks could be different in a supermarket and an airport. It can be found that most indoor landmarks are similar between sighted and unsighted people except for "information". Visually impaired people rely more on tactile features of the environment. These landmarks can be labeled in the graphic map. An example was shown in Figure 4. The next section explains the detection and representation of landmarks.

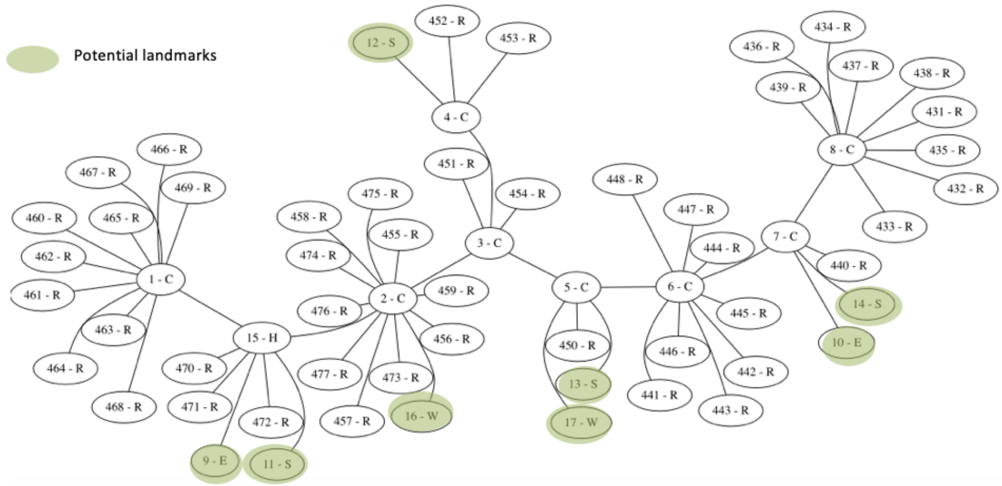



Fig 4: Refine the graphic map after adding landmarks

### 8.5.2 Landmarks detection and representation

Those landmarks information that are hard to be detected by a white cane can be interpreted by using assistive technique. We propose to apply a "tactile landmarks" approach to our navigation system for people with deafblindness, that uses computer vision or beacon sensors to detect landmarks in the environment, and represents these landmarks using tactile signals.

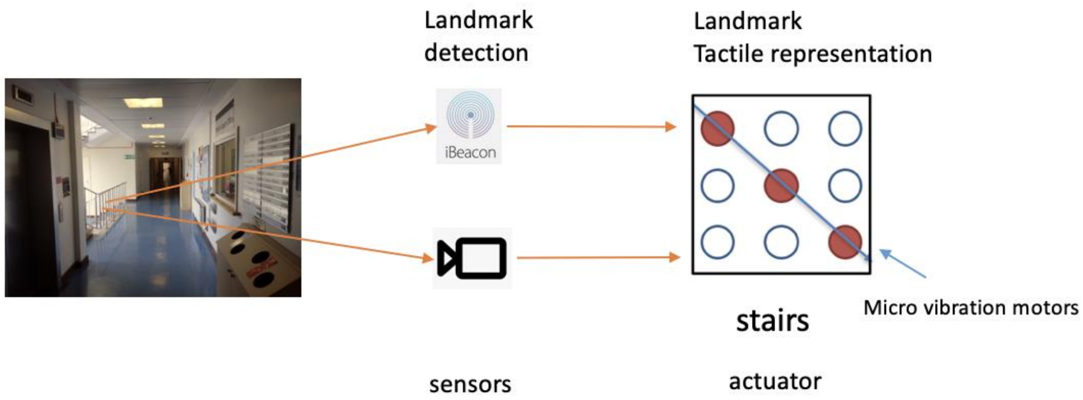



Fig 5: The concept of "Tactile landmark". iBeacon or camera detects landmark (stairs) and represents it in the 3X3 vibro-tactile interface

The landmark detection can be achieved by using the similar methods that we have presented in the "active object search" part with computer vision or iBeacons. The tactile signals could be created by using the abstract or symbolic approach (Ternes and Maclean 2008). The abstract representation is about manipulating a stimulus' characteristics, whereas the symbolic approach focuses on the semantic association of stimuli with known metaphors. A study of using wearable tactile to display landmarks was investigated by Srikulwong and O'Neill (2011) for an urban pedestrian navigation system. However, Srikulwong and O'Neill (2011) 's study is not related to people with deafblindness or indoor navigation, so that landmarks selection is different. Design of tactile feedback pattern to represent landmarks is out of the scope of Work Package 4. Colleagues from Work Package 6 are developing haptograms to transfer semantic meaning of objects to user by using a vibro-tactile matrix interface, that potentially could be used to represent the "tactile landmark". An example of tactile landmark of stairs was shown in Figure 5.

## 8.6 Route Planner

Based on graphic map and landmark information, a feasible route could be planned, here we implemented the Dijkstra's algorithm in the Boost Graph Library (Siek, Lumsdaine et al. 2002) to find the shortest path. However, this route method could be refined in the future by adding the following considerations (Ugulino and Fuks 2015, Koester, Awiszus et al. 2017):

(1) Select shoreline routine and less open space, for instance, more corridor, less hall
(2) Select a route with elevators rather than that with stairs

(3) Select a route with more landmarks and cues

Between two nodes, we used edges to present behavior, such as <node|behavior|node>. These behaviors consist of one or a series of acts, such as: enter/ leave a room, turn right/left, go straight. Currently, the behaviors and landmarks are added on the auto-generated route manually. At some nodes, the node connects to several potential landmarks. In this case, we choose the highest rank landmark (Fellner, Huang et al. 2017). An example was shown in Figure 6, Supposing that a user wants to walk from Room 465 to Room 444. The corresponding route in the floor plan was shown in Figure 7.

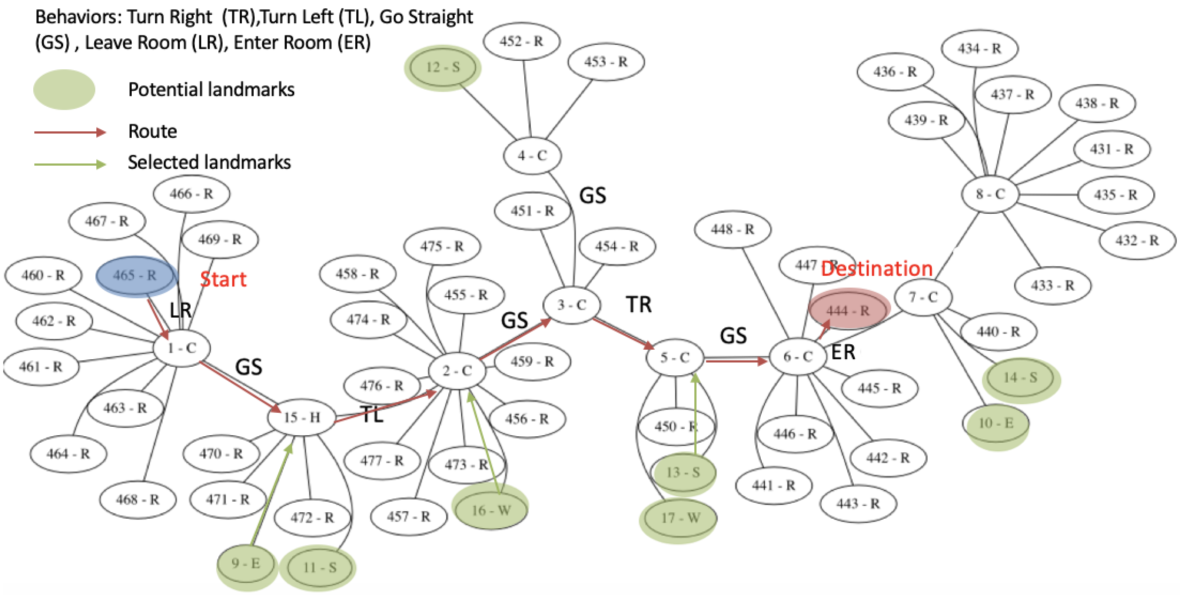



Fig 6: Auto-generated route using Dijkstra's algorithm with manually added behaviors and landmarks


Fig 7: Generated route in Figure 6 on the floor plan

## 8.7 Task Manager

Task manager includes two parts: an action planner and an executor. The code was developed based on a ROS package "task_planning" (https://github.com/UTNuclearRoboticsPublic/task_planning ). The Goal Oriented Action.Planning (GOAP) approach was used to generate a plan by giving a series of actions. Each action consists of three items: Entitles, Propositions, and Effects (Russell and Norvig 2016). For example, we defined the following two sentences:

The first action is "Move to waypoint a", with preconditions that are "not at waypoint a", the effects that are "at waypoint a" and "not at other waypoints". The second action is "Turn at waypoint a", with preconditions that are " at waypoint a" and "not Turn", the effects that are "Turn" .

```
Planner.addAction("Move(a:waypoint)",{"~Docked()",''~At(a)"},{"At(a)","~At(~a:waypoint)"}
)

Planner.addAction("Turn(a:waypoint)", {"At(a)",''~Is_Turn(a)"},{"Is_Turn(a)"})
```

The planning algorithm is based on an A* graph search over the action space. We selected this approach instead of using the results directly from the route planner, because this approach helps us to implement new behavior with a minimum of re-coding. This is important for us to integrate other functions, for example, when we call for the "leave room" function, it may require the system to call the active search object function to find the door first.

Executor is used to handle execution of plans produced by action planner. In executor, we defined call back functions and related them to each action. For example, the "Move to waypoint a" was bound to the function Move_callback(), this function monitors the iBeacon proximity detection results, once the target iBeacon is detected (usually at the joint of two structure nodes), the Move_callback function would stop and executor would run the next function.

ROS server-client structure (Newman 2017) was used to communicate between proximity detection and task manager. Task manger sends a request message to the proximity detection module, and the message includes the name of the target iBeacon. Once the target iBeacon was detected in the "immediate" region, the proximity detection module responds to the task manager, and functions in the task manager would execute relevant instructions. This method also could be used to communicate task manager to any other functions, such as the visual analysis module.

# 9. System Validation

In the previous sections, we have explained how to generate the navigation information:" turn by turn" directions and landmarks from a floor plan, and implement these methods with sensors. We took an example of the fourth floor of Mechanical Engineering building at University of Leeds to illustrate the process. In this section, we used the generated route and iBeacon sensors to validate the navigation functions. Eight iBeacons were placed in the environment as shown in Figure 8. Five iBeaons were used for navigator (with numbers), they should be installed at the joint of two structure nodes, numbers could be allocated randomly. Three iBeaons were attached to the selected landmarks: ”elevator”, “stairs”, “toilet”. The participant stated from the Room 465 and walked to the Room 444. A visualization interface was developed to monitor the information delivered to the user as shown in the same figure. During the test, navigation system was able to execute the right actions: “move” and “turn” after detecting the iBeacons, and deliver the correct “turn by turn” direction. An example is that the user arrived at the position near the iBeacon 3 and received a message “turn right” and the next iBeaon is 4. The system can also recognize the selected landmarks. Currently, we have not added the behavior “enter/ leave door”, which needs to use functions we developed in “active object search” for searching for a door.

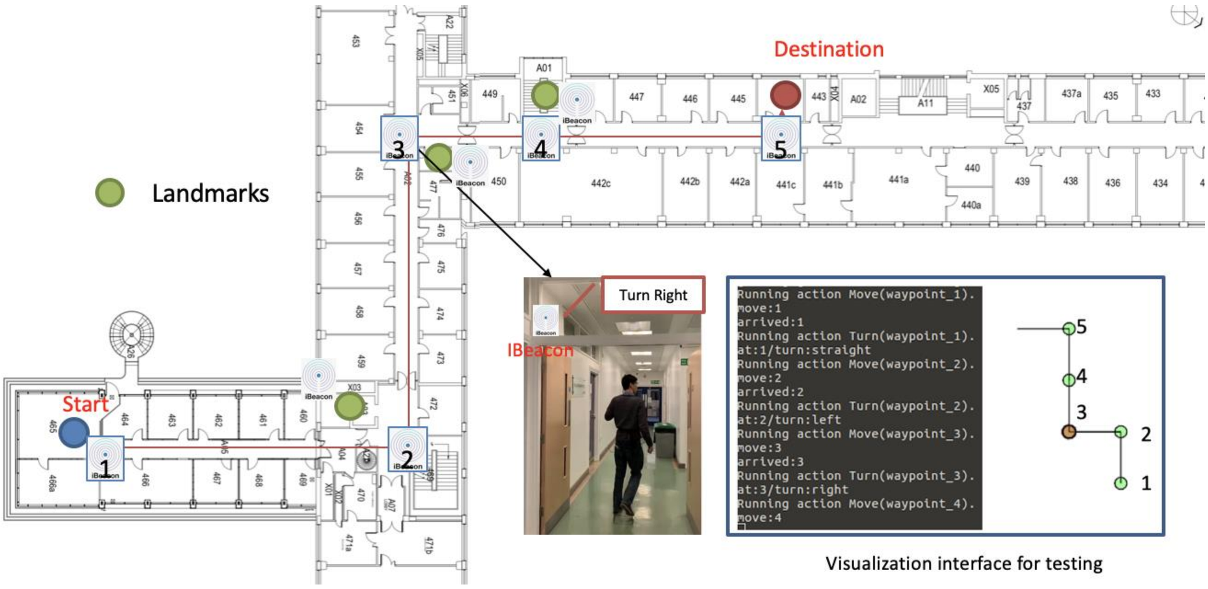



Fig 8: Positions of iBeacons in the real environment and a snapshot of the visualization interface during the test. The user was at the position near iBeacon 3 and received a message: “turn right” and moved to the next iBeacon 4

We also carried out an initial user study with one participant with deafblindness who uses cochlear implants during a workshop at a consortium meeting on 5th December 2019 in Offenburg. Because the haptic feedback for navigation will be integrated in the D4.3. During the test, we described the commands generated by the controller to the user orally. We placed several iBeacons on the corners of the meeting room to simulate turns. In the first round, we only gave “turn by turn” direction information, and in the second round, we added simulated landmarks information such as “stair”, “elevator”. The participant emphasized that the correct “turn by turn” information was the most important for him to complete the route, also it would be better to describe turn more accurate, such as turn 45 degrees, 90 degrees, or 135 degrees. It was hard to evaluate the landmark now due to the fact that the simulated navigation route was too short and without a tactile interface feedback. At this stage, we validated that people with deafblindness can accept the navigation information generated by the system.

# 10. Conclusion and Future Work

In this part of the report, we described the process of development of indoor navigation. The characteristics of this system are:

(1) The navigation information was defined by considering navigation strategies and issues of people with deafblindness

(2) A behaviors navigation approach was adopted to work without an indoor location system

(3) “Tactile landmark” was proposed to enhance the user’s awareness of environment during navigation. This method will be applied by incorporating visual analysis into the haptic interface

In the future work, we will integrate the system with haptic feedback interface and carry out more navigation experiments with users in the real environment. The potential improvements of this navigation system are:

(1) Add a “Position tracking” pipeline to help navigation system check if user on the right route. This could be achieved by applying Inertial measurement unit (as described in Deliverable 4.1) based pedestrian dead reckoning (Park and Teller 2014, Xiao, Wen et al. 2015, Lu, Uchiyama et al. 2019) or using “situation awareness” developed in Work Package 3.

(2) Currently, the behaviors and landmarks are added manually based on the route, these could be done automatically in the future.

(3) Add more behaviors functions in the navigation system, such as “enter/leave room” and “obstacle avoidance”.

# 11. References


Ahmetovic, D., et al. (2016). NavCog: a navigational cognitive assistant for the blind. Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, ACM.

Ahmetovic, D., et al. (2018). Turn right: Analysis of rotation errors in turn-by-turn navigation for individuals with visual impairments. Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, ACM.

Al-Khalifa, H. S. (2008). Utilizing QR code and mobile phones for blinds and visually impaired people. International Conference on Computers for Handicapped Persons, Springer.

Aydemir, A., et al. (2012). What can we learn from 38,000 rooms? Reasoning about unexplored space in indoor environments. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE.

Bradley, N. A. and M. D. Dunlop (2005). "An experimental investigation into wayfinding directions for visually impaired people." Personal and Ubiquitous Computing **9**(6): 395-403.

Cheraghi, S. A., et al. (2017). GuideBeacon: Beacon-based indoor wayfinding for the blind, visually impaired, and disoriented. 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom), IEEE.

Couclelis, H., et al. (1987). "Exploring the anchor-point hypothesis of spatial cognition." Journal of environmental psychology **7**(2): 99-122.

Dias, M. B., et al. (2015). "Indoor navigation challenges for visually impaired people." Indoor Wayfinding and Navigation: 141-164.

Fang, S.-H. and T.-N. Lin (2010). Accurate indoor location estimation by incorporating the importance of access points in wireless local area networks. 2010 IEEE Global Telecommunications Conference GLOBECOM 2010, IEEE.

Fellner, I., et al. (2017). "“Turn Left after the WC, and Use the Lift to Go to the 2nd Floor”—Generation of Landmark-Based Route Instructions for Indoor Navigation." ISPRS International Journal of Geo-Information **6**(6): 183.

Ganz, A., et al. (2012). "PERCEPT indoor navigation system for the blind and visually impaired: architecture and experimentation." International journal of telemedicine and applications **2012**: 19.

Ganz, A., et al. (2018). "PERCEPT navigation for visually impaired in large transportation hubs."

Gleason, C., et al. (2017). Luzdeploy: A collective action system for installing navigation infrastructure for blind people. Proceedings of the 14th Web for All Conference on The Future of Accessible Work, ACM.

Goldschmidt, M. (2018). Orientation and Mobility Training to People with Visual Impairments. Mobility of Visually Impaired People, Springer**:** 237-261.

Golledge, R. G. (1999). Wayfinding behavior: Cognitive mapping and other spatial processes, JHU press.

Golledge, R. G., et al. (1996). Cognitive mapping and wayfinding by adults without vision. The construction of cognitive maps, Springer**:** 215-246.

Guerreiro, J., et al. (2018). How context and user behavior affect indoor navigation assistance for blind people. Proceedings of the Internet of Accessible Things, ACM.

Hersh, M. (2016). "Improving deafblind travelers' experiences: an international survey." Journal of Travel Research **55**(3): 380-394.

Ishikawa, T. and U. Nakamura (2012). "Landmark selection in the environment: Relationships with object characteristics and sense of direction." Spatial Cognition & Computation **12**(1): 1-22.

Jeamwatthanachai, W., et al. (2019). "Indoor navigation by blind people: Behaviors and challenges in unfamiliar spaces and buildings." British Journal of Visual Impairment **37**(2): 140-153.

Jefferies, M. E. and W.-K. Yeap (2008). Robotics and cognitive approaches to spatial mapping, Springer.

Kalia, A. A., et al. (2013). "Combining path integration and remembered landmarks when navigating without vision." PloS one **8**(9): e72170.

Koester, D., et al. (2017). Mind the gap: Virtual shorelines for blind and partially sighted people. Proceedings of the IEEE International Conference on Computer Vision.

Liao, C.-F. and M. Rakauskas (2010). Accessible Traffic Signals for Blind and Visually Impaired Pedestrians, University of Minnesota, CTS Report.

Lock, J., et al. (2019). "Active object search with a mobile device for people with visual impairments."

Loomis, J. M., et al. (1993). "Nonvisual navigation by blind and sighted: assessment of path integration ability." Journal of Experimental Psychology: General **122**(1): 73.

Lu, C., et al. (2019). "Indoor Positioning System Based on Chest-Mounted IMU." Sensors **19**(2): 420.

Luoh, L. (2014). "ZigBee-based intelligent indoor positioning system soft computing." Soft Computing **18**(3): 443-456.

Millonig, A. and K. Schechtner (2007). "Developing landmark-based pedestrian-navigation systems." IEEE Transactions on Intelligent Transportation Systems **8**(1): 43-49.

Nassih, M., et al. (2012). Obstacles recognition system for the blind people using RFID. 2012 Sixth International Conference on Next Generation Mobile Applications, Services and Technologies, IEEE.

Newman, W. (2017). A Systematic Approach to Learning Robot Programming with ROS, Chapman and Hall/CRC.

Park, J.-g. and S. Teller (2014). "Motion compatibility for indoor localization."

Priyantha, N. B., et al. (2000). The cricket location-support system. Proceedings of the 6th annual international conference on Mobile computing and networking, ACM.

Russell, S. J. and P. Norvig (2016). Artificial intelligence: a modern approach, Malaysia; Pearson Education Limited.

Sato, D., et al. (2017). Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, ACM.

Shih, M.-L., et al. (2018). DLWV2: a Deep Learning-based Wearable Vision-system with Vibrotactile-feedback for Visually Impaired People to Reach Objects. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE.

Siek, J., et al. (2002). The boost graph library: user guide and reference manual, Addison-Wesley.

Tapu, R., et al. (2017). "DEEP-SEE: Joint object detection, tracking and recognition with application to visually impaired navigational assistance." Sensors **17**(11): 2473.

Ternes, D. and K. E. Maclean (2008). Designing large sets of haptic icons with rhythm. International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, Springer.

Trongwongsa, T., et al. (2015). Shopping navigation system for visual impaired people based on proximity-based technology. 2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE.

Ugulino, W. and H. Fuks (2015). Landmark identification with wearables for supporting spatial awareness by blind persons. Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, ACM.

Velázquez Guerrero, R. (2017). "Mobility of visually impaired people: fundamentals and ICT assistive technologies." OPENAIRE.

Want, R., et al. (1992). "The active badge location system." ACM Transactions on Information Systems (TOIS) **10**(1): 91-102.

Weiss, M., et al. (2018). "A Survey of Mobile Computing for the Visually Impaired." arXiv preprint arXiv:1811.10120.

Werner, S., et al. (2000). Modelling navigational knowledge by route graphs. Spatial cognition II, Springer**:** 295-316.

Whiting, E. J. (2006). Geometric, topological & semantic analysis of multi-building floor plan data, Massachusetts Institute of Technology.

Xiao, Z., et al. (2015). "Robust indoor positioning with lifelong learning." IEEE Journal on Selected Areas in Communications **33**(11): 2287-2301.

Yang, Z. and A. Ganz (2018). "A Sensing Framework for Indoor Spatial Awareness for Blind and Visually Impaired Users." IEEE Access **7**: 10343-10352.

Yu, Z., et al. (2016). "PalmSight: an assistive technology helping the blind to locate and grasp objects." Carnegie Mellon University, Pittsburgh, PA, Tech. Rep.: 16-59.

Zientara, P. A., et al. (2017). "Third Eye: a shopping assistant for the visually impaired." Computer **50**(2): 16-24.