Technology
Arts Sciences
TH Köln

Faculty of Information, Media and Electrical Engineering

Javier Rodríguez Puigvert

Matrikelnummer: 11093408

Obstacle Collision Avoidance System based on Tango

Forschungsprojekt

27.04.2017

Contents

Motivation

According to the World Health Organization about 285 million people of all ages are estimated as visually impaired of whom 39 million are blind and 246 million have low vision [1]. Visual impaired people have to face a number of challenges everyday as reading a document, identifying what the food's expiration date is and even to know where they are. Simple everyday tasks become a challenge when one is deprived of visual sense [2]. While navigating through space, a metal representation of the space is naturally built based on what it is being seen or heard. The corresponding movements within the space are related to the close surroundings [3].

Navigating inside unfamiliar buildings is one of the major difficulties that visually impaired must go through. In the context of mobility, the information of how far the obstacles are allows us to move around safely. Unfortunately, visual impaired are unable of acquiring data from the visual sensoric system, hence they need to rely on another source such as touch or hearing senses and avail themselves of external aids. There are two very well established.

- When it comes to mobility, the *white cane* is the symbol par excellence. It helps visually impaired to move freely in unknown surroundings, to identify any irregularity on the floor or even to differentiate objects like chairs, walls or tables. By moving this mechanical object one can find out whether there is anything in front only by bumping the cane. Nowadays the white cane is still the most used travel aid not only because the ease of use but also its weight, size , folding and even the direct vibration feedback that is transferred when the cane touches the floor.

  However, Mobility and Orientation (M and O) training is needed for effective protection, the typical learning process has four stages: During the first stage, one learns the basics about how to use the white cane. In the second stage, full concentration is required to use it. When the subject moves correctly the cane without a need of concentration is on the third stage. Finally, at the end of the adaptation process, in the fourth stage, the cane reliably provides information and protection without involving that much effort [4].

- *Guide dogs* are mobillity aid for visual impaired, they help during the navigation and the obstacle avoidance proccess. Without a doubt, these animals provide companionship and affect as well as enhance confidence and independence. Since the animals are puppies, it takes around twelve months to acquire the required abilities to guide a blind person. Besides mobility and orientation training is mandatory to learn how to interact with the guide dog [5]. Nevertheless the effort of taking care of them, price and number of dogs available are reasons why they are not so often used [6].

Researchers have been aware of mobility difficulties for visual impaired, in consequence, Electronic Travel Aids (ETA) devices have shown up in the last decades to assist visual impaired during the mobility proccess. To replace these the traditional tools represents still today quite a challenge [7].

Smartphones have already proved to be very successful in solving a series of challenges in serving as help for people with visual disability, e.g. helping identifying objects, reading emails out loud or even using the camera to read medication labels [8]. But nowadays user - device interaction is still limited by the screen having the exception of recording video or shooting photos. With Tango enabled devices users can interact with the real world. Tango is a technology based on computer vision that provides the device the ability to understand the real world thanks to additional integrated hardware [9].

With the next generation of smartphones supporting Tango technology new possibilities arise. This research project investigates whether these devices could be the new game changer for visual impaired.

Goals

The goal of this project is to present a brief review of the current available solutions, describing on which technology they are based and concretely focuses on assessing whether Tango enabled devices are able to assist the visual impaired during obstacle avoidance process in unknown indoor environments.

To this extent, the necessities of the target group will be collected to assess the viability of using current technologies to solve them. Afterwards, a solution will be implemented that can solve these necessities, based on Tango technology.

State of the art

In the mobility context, the approaches should always focus on the real user necessities. How to improve the quality of life from visual impaired is being a large research field in the last decade. As techlonology progresses, the improvements have been making lives easier. Extensive research has already dealt with assisting visual impaired in several tasks as an extension of the white cane [10][11], as a robot for mobility indicating in which direction should walk [12] or even as object recognition [13].

Current approches are able to detect the environment around using different sensor technologies:

- Ultrasonic sensors:

  Ultrasonic sensors are sensors able to measure distances by sending a sound wave at a specific frecuency and receiving the bounced wave. Ultrasonic sensors are insensitive external factors e.g. fumes, humidity, ambient noise, object color. Detecting surfaces at more than 45 degrees where the waves incidence provoke specular reflexion. Even the nature of the surface influences in the mesurement up to the sound energy absorbed or reflected by the material [14].

  Smartcane is based on this technology. This electronic cane is able to detect objects between 0.5 m and 3 m in its path. This ETA device couples on the top fold of white cane extending the angle of detection. The feedback is sent out as a pattern of vibrations. These vibrations provide information about how far obstacles are[10].

- Infrared sensors:

  Infrared sensors are widely used in ETA devices. This kind of sensors cover a range between 5 cm to 10 m and in close distances the response time is faster than in ultrasonic sensors.

  An ETA device based in this techlonogy is Tom Pouce, an infrared proximiter. This device can be clipped on the white cane transforming it on an electronic cane. It sends infrared beams in different directions

Figure 3.1: Tom Pouce device[11]

and at different emission powers, to cover the protected field. This device can detect obstacles up to 2 m high and can reach between 2 m and 15 m ahead [11].

- Camera as information source:
  When it comes to ETA devices, the use of cameras is as source of information very common. The challenge of using this technology is to build suitable information from the image itself.

  Stereo camera are cameras based on two or more lends which acquire images independently. Having different perspectives of the same scenario, disparity can be achived by estimating the relative displacement of the same point in the different samples. The concept of disparity in binocular vision is what provides humans the ability of depth perception and thereby the creation of 3D images.

  Projected- light 3D Camera is a combination of a camera with the projection of structured light. By adding a depth sensor based normally in infrared projector to an RGB or stereo camera, each pixel can be mapped with the distance measured. These kinds of cameras have a range between 0.5m and 3.5m but the sensing range can be adjustable. Microsoft kinect is the most famous example of this technology. Kinect cane is a research project that tries to assist visually impaired users in finding objects in everyday environment[13]. Moreover, the existing Time of Flight cameras can be attached to any smartphone in order to give them the hability of depth perception. e.g. Pico flex family from

Pmd technologies and Infineon [15] [16]. The ToF and stereo cameras used on tango devices.

Project Tango

In 2014 Google Tango started as a research project from Google ATAP (Advanced Technology And Projects) [17]. Tango is an innovative mobile technology which allows users to interact with the real world through their smartphone, not only shooting photos or recording videos but also making use of it e.g., in augmented and virtual reality. Tango enabled devices can estimate where they are located and what is around them using computer vision and without any necessity of external entitiy[18]. One of the main goals of Tango is to unify the way people think about space and how devices think about space. During 2016 Google Tango evolved as an independent project called Tango.(see figure)

Figure 4.2: Google Tango evolved as Tango in 2016

At the moment of writing, this young technology is integrated only in three devices:

- Penaut phone was the first Tango device, released by Google. This smartphone has been used on different kinds of reseach projects mostly based on robotics and computer vision. A good example of that is a research project based on an autonomous robot able to navigate in different enviroments, developed by NASA and tested on the International Space Station[19][20].
- The Yellowstone tablet was the official "Project Tango Developer Kit" that integrated a fisheye camera for motion tracking and an infrared

Figure 4.3: Tango Developer Kit

Figure 4.4: Lenovo phab 2 Pro

projector with an RGB-IR camera for depth perception.

- In 2016 Lenovo Phab 2 Pro, the first consumer device was launched with 4 GB RAM, Qualcomm Snapdragon 652 processor built for Tango and 16 MP PDAF fast-focus rear camera[21].
- On the CES 17 ASUS announced the ZenFone AR, with 8Gb RAM, which will be the first smartphone able to support both Tango and Daydream technologies[22].

Three main abilities characterize the technology: motion tracking, area learning and depth perception.

- With motion tracking, a device is able to understand its relative position within the world. The ability of tracking how far or in which direction the device is moving is possible, thanks to the fusion of the visual inertial odometry and the inertial measurement unit[23].

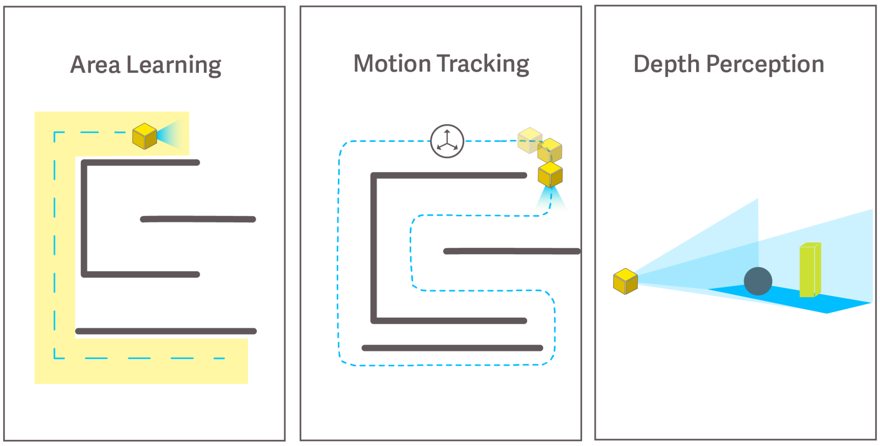

Figure 4.5: Area Learning, Motion Tracking and Depth Perception

- Area learning gives the ability to see and distinguish characteristics of the space, so that a device can recognize whether it has been in that place already before[24].
- Depth perception provides devices the ability to understand distances to objects in the real world[25].

Tango API provides six degrees of freedom, so that it can track movements in three axis of motion and three axis of orientation.

In the presented approach, the ability used is depth perception to acquire data about how far the objects around are.

4.1 Depth Capabilities

The depth perception on Tango requires a minimum distance to the obstacle to detect. Actually, a Tango enable device is able to acquire suitable data between 0.5 and 4 meters. In addition, we have to pay attention to how fast the sensor reacts, as Tango technology was not optimized for detecting sudden changes in the environment. The recording of point cloud data is currently limited by frame rate and cannot be used to detect sudden changes in the environment.

Areas or objects where infrared are not reflected or areas lit with light sources high in infrared like sunlight or incandescent bulbs can not be detected[25]. Light conditions play an important role for the accuracy of acquired data.

"Tango and Android are not hard real-time systems because the Android Linux Kernel cannot provide strong guarantees on the execution time of the software running on the device. As a result, Tango is considered a soft real-time system." [23]

Obstacle Collision Avidance System

The Obstacle Collision Avoidance System pretend to assist visually impaired people during the obstacle avoidance process using the depth perception of a Tango enable device as information source. In order to move freely around the space, a user could use this system as a virtual white cane. The user should touch the screen to discover what the depth sensing is receiving. Touching on the left side of the screen, causes receiving depth information on the left side of the space, analogously, in the right side of the screen. The feedback is given in form of smartphone's vibration according how far the detected obstacles are. Using ~~the~~ Depth Perception functionality, the distance between the device position and the obstacles around can be measured. Applying this procedure, the objective is to provide a kind of three-dimensional feeling.

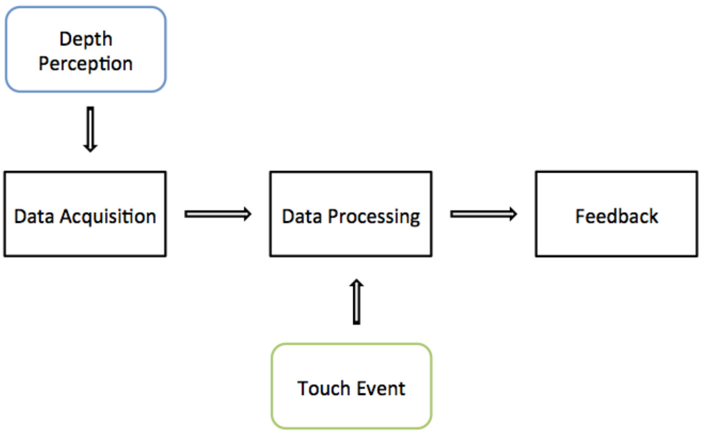

Figure 5.6: OCAS architecture

Device	Max. Data Rate (MB/s)
Tango development kit	28.125
Lenovo Phab 2 Pro	84.375

Table 5.1: Maximum Data rates

Figure X shows the OCAS architecture. Three main modules compose the architecture: Data acquisition module, data processing module and feedback module.
Data acquisition module is in charge of acquiring pointcloud data~~,~~ from the depth perception service of Tango. In the data processing module, the data is classified considering where the user want to be focused. The Feedback module is responsible of transmit the corresponding signal to the user.

5.1 Data Acquisition

With depth perception data from Tango, distances between the current device's position and objects in the real world can be measured. This data is acquired in point-cloud format, where each point is described in 3 dimensions (x, y and z). Each point from the point cloud corresponds to a point in the relative space at concrete time related to the device's position[25].

In the Tango development kit the point cloud resolution is max. 12000 points per measurement while the Phab 2 Pro the resolution is max. 36000 points. Since Tango service sends new point cloud information, this leads to:

$$Max.DataRate = Resolution * 200ms * 3\frac{Floats}{Point} * 4\frac{Bytes}{Float} \tag{5.1}$$

The table X presents the maximum data rate up to device.

The resolution on each device is related to the Field Of View (FOV). The FOV in Tango describes the extent of the observable world that Tango is seeing on the current scene at a specific time.

$$FOV_H = 2 * \arctan(\frac{0.5 * width}{F_x}) \tag{5.2}$$

$$FOV_V = 2 * \arctan(\frac{0.5 * height}{F_y}) \tag{5.3}$$

width: The width of the image on the image sensor in pixels.

height: The height of the image on the image sensor in pixels.
Fx: Focal length, x axis, in pixels.
Fy: Focal length, y axis, in pixels.[26]

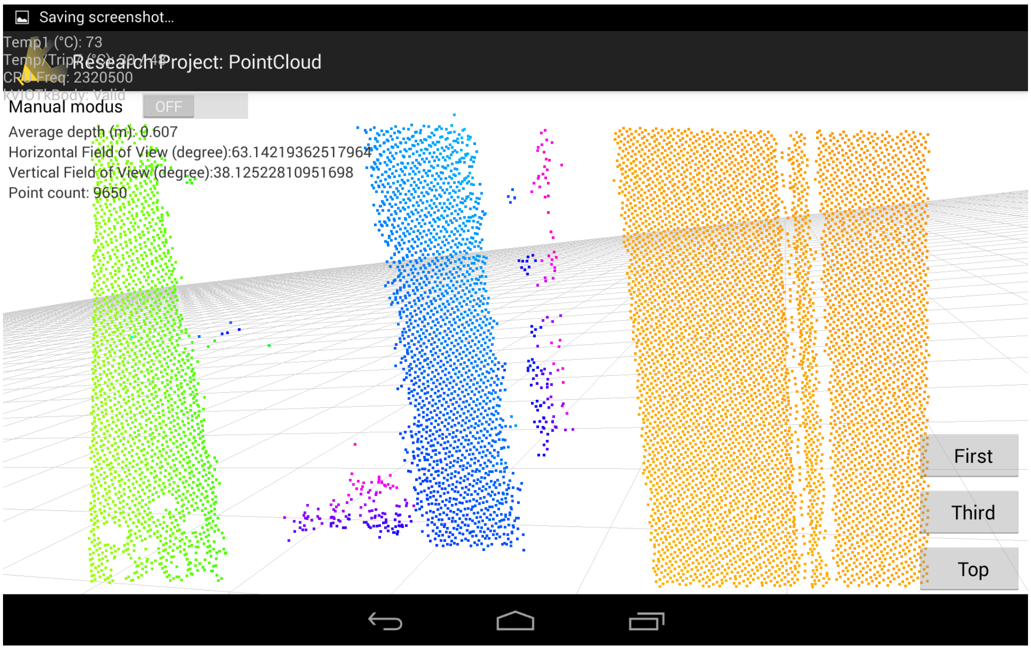

Figure 5.7: Data adquisition example

5.2 Data Processing

The fact of placing the white cane on the floor, implies that a signal is transmitted a long the cane by each bump. This means that a visually impaired person receives a real-time feedback while discovering the surroundings. To unify touch sense with depth perception as the white cane does, we map the screen area with a 3D area in the relative space, according to the point cloud acquired. The system will be focused in that particular space of the real world where the user is pointing on the screen. Having a relation between the screen width to the horizontal FOV and the screen height vertical FOV we can find out which range of the FOV the user wants to focus on.

The mapping process between FOV and screen size is done as follows:

1. Get touch position in pixels when the user touches the screen.

2. Normalize the values regarding the screen size.

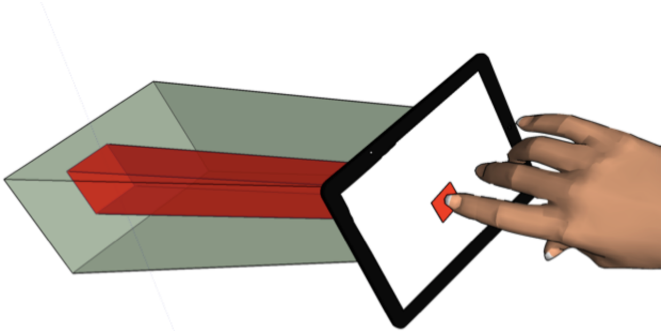

Figure 5.8: Mapping Field of View and Screen

3. Shift origin of coordinates to the device center.

4. Multiplying the horizontal FOV with the x coordinate of the normalized point and the vertical FOV with the y coordinate of the normalized point, the result are the angles corresponding with x axis and y axis. The intersection of these areas is the 3D space that the user was looking for.

$$DIR_X = (\frac{Touch_X}{Screen_{Width}} - 0.5) * FOV_H(degrees) \tag{5.4}$$

$$DIR_Y = (0.5 - \frac{Touch_Y}{Screen_{Height}}) * FOV_V(degrees) \tag{5.5}$$

5. For each point from of the pointcloud, the angle in relation to x axis and y axis is calculated.

6. Comparing whether the angles, it can be known whether the point is in the particular space.

Once it is known which points of the pointcloud should be consider, the proximity is evaluated.

Three areas have been defined taking into account the length of a white cane, people height and step length related to the proximity of the objects. The first area covers the maximum range of a white cane until the half of the maximum. The second area until the minimum distance that a Tango device detects and the third area starts where Tango is not able to detect any obstacle. Points inside of the three defined areas are significant for the obstacle collision avoidance system.

In different researches for visual impaired people it has been demonstrated that users don't want to rely their lives on automatic solutions where they don't even know what is happening[3]. Moreover, overstress to the user with much information can cause a bad user experience, that is why in this approach, the user controls which area should be explored.

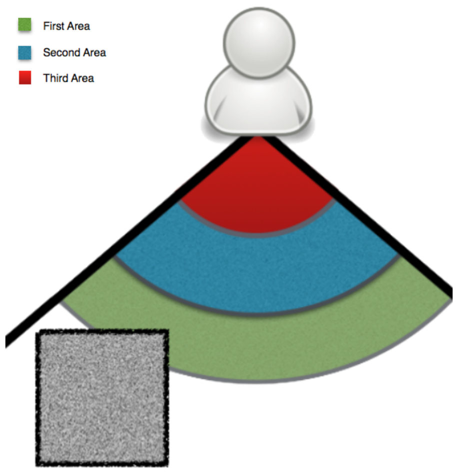

Figure 5.9: Three detection areas

5.3 Feedback

The user touches the screen to find out what is in front of him or even to find the outlines of an obstacle. As soon as something impedes the progress, a vibro-tactile feedback signal will be sent.

When an obstacle trespasses the first area, the user receives a series of soft vibrations according how far the object is. This feedback can be useful if the user wants to follow a wall using the same procedure as a real white cane bumping against the wall.

In case that any obstacle reaches the second area a continuous vibration is sent, warning a possible collision. To avoid the collision the user must rotate itself and touch the screen to find a free path.

Walking through two close obstacles is feasible if the user can find manually the outlines of the obstacles, he could figure out which is the correct way, as is represented schematically in Figure 5.8.

Most of the ETA devices uses vibrotactile feedback because is not so invasive as sound feedback because hearing sense is important during the

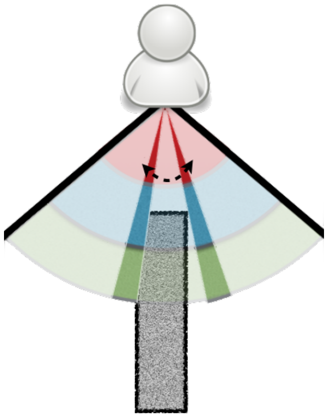

Figure 5.10: Findind the outlines of obstacles

orientation proccess. However, in case of an emergency situation a voice message is sent warning and explaining the danger:

- An obstacle is too close and nothing is detected
- The user is walking too fast.
- The light conditions are not good enough to acquire data.
- The battery of the device is too low.

5.4 Requirements

5.4.1 Device holding

To guarantee suitable data readings, the device must be placed on landscape orientation and pointing forward, with a slight inclination between 20-45 degrees on the Z-axis. Sensors and camera should always face forward looking at what is in front of the user. Figure 5.9 describes the optimal device position.

The system is intended for moving forward at any time. If the user moves laterally or holds the devices in inappropriate position, the feedback acquired will not correspond with the expected movement. Another possible setup is to hang the device to the users neck like in this experiment [27]. It is very important that no obstacle should impede the reception of data. The cameras and the IR projector should not be cover by the hands or fingers.

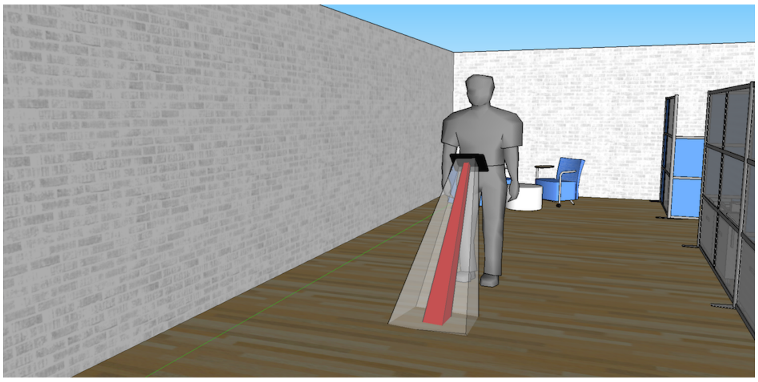

Figure 5.11: Holding the device

5.5 Specification as a User Story

- Case 1:
  When the user is walking.
  And the device detects an obstacle.
  Then the device vibrates.

- Case 2: Detection Error: speed limit
  When the user exceeds the speed limit.
  Then the device sends a voice notification Too fast

- Case 3: Detection Error: Nothing to detect
  When the sensors are covered
  Or an obstacle is too close
  Then the device cannot detect any obstacle
  And the device sends a voice notification Nothing to detect

Evaluation

To evaluate the effectiveness of the presented system during the obstacle avoidance proccess two experiments has been carried out.

6.1 Experiments focused on mobility

The first experiment consisted on walking inside a building with no previous orientation during 90 seconds. The goal was to observe whether the subject is able to walk around without any collision.

The second experiment was carried out in the same environment but including obstacles in the path, to evaluate whether the subject was able to avoid them.

Experiments have participated sighted people and visually impaired people. The volunteers were instructed about how to hold the device and how to use the system properly.

6.2 Result

The results of the experiments were positive. In general, the subjects were able to control the situation and use the obstacle avoidance system as expected. Unfortunately, a couple of collisions occured during the experiments: when devices were directly pointing to a non reflecting material and nothing could be detected and also when the subject rotated suddenly and an undetected obstacle was too close.

After the experiment, the participants were invited to give their impressions about the approach. A summary of the received feedback is:

The position in which the device must be hold was considered as uncomfortable for some of the participants. With the white cane, blind people is able to distinguish obstacles by bumping e.g.whether the obstacle was a wall, a chair or a table and in comparison the device is too heavy, a blind participant pointed.

Figure 6.12: Volunteer participating in the experiments

Figure 6.13: Pointing to a non reflecting material

When an object is too close the vibration was too hard which, might give the false impression that there is an imminent danger.

Conclusion

The present paper presented an obstacle collision avoidance system based on ~~a~~ Tango technology for visual impaired people. The system has been implemented as a virtual cane, using depth perception concept from Tango. So that the user is able to detect whether an obstacle is in front. The user detects by touching the smartphone screen in which direction the obstacles are and by vibration how far they are.

The experimental results indicate that the system is feasible of assisting visual impaired as a mobility aid.

That the system is developed on a smartphone is a clear advantage because a lot of features can be added without more hardware e.g. object recognition.

Although there are technical mobility aid on the market, a minory of the visually impaired population are making use of these aids[3]. Taking into account the huge acceptance of smartphones in visually impaired, to combine the concepts of an ETA device and a smartphone could be a great solution to reach more interest.

Outlook

The need of a companion is normally a must when visual impaired is in an unfamilliar building. In this research, it was used a relative coordinate system up to the user position but Tango technology is able go further, developing an indoor navigation system, so that the visually impaired users could navigate by themselves.

Future work should be focused on improving the orientation using motion tracking and area learning of Tango. That would be the natural way to progress with this research since in the presented approach has been focused only on mobility.

Object recognition is the most requested feature of visually impaired. Being able to recognize objects within in the space or even read product labels means a huge independency. So that, the presented approach is developed on an Android device and an object recognition could be achived within the same application. Moreover, making relations between the current position and objects in the space could also be helpful to achieve an allocentric orientation.

Bibliography

[1] W. . W. health organization, "visual impairment and blindness," in *EEE Systems Journal* (media centre, ed.), p. fact sheet no. 282.

[2] E. B. M. R. M. Y. Z. S. White and J. P. Bigham, "Visual challenges in the everyday lives of blind people," in *Proceedings of the 2013 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI).*, pp. 2117–2126.

[3] L. F. C. E. A.-V. C. C. P. M. Gori, "From science to technology: Orientation and mobility in blind children and adults," in *Neuroscience & Biobehavioral Reviews* (E. Ltd, ed.), p. Pages 240251.

[4] VisionAware, "For independent living with vision loss." `http://www.visionaware.org/info/everyday-living/essential-skills/an-introduction-to-orientation-and-mobility-skills/how-do-i-learn-to-use-a-cane/1235`, 2017.

[5] C. E. M. Anna Chur-Hansen, Lucy-Kate Werner and S. Hazel, "The experience of being a guide dog puppy raiser volunteer: A longitudinal qualitative collective case study," in *EEE Systems Journal*, pp. 1–11, l.

[6] G. E. for the Blind. `https://www.guidingeyes.org/learn/orientation-mobility-seminars/`, 2017.

[7] K. L. Li, "Electronic travel aids for blind guidance,an industry landscape study." `http://scet.berkeley.edu/wp-content/uploads/INDENG290-Project-Report-Electronic-Travel-Aids-for-Blind-Guidance-Kun-Li.pdf`, 2015. IND ENG 290 Project Report.

[8] P. T. Neha Bari, Nilesh Kamble, "The experience of being a guide dog puppy raiser volunteer: A longitudinal qualitative collective case study," in *Android based object recognition and motion detection to aid visually impaired*, p. Volume 3 No.10, International Journal of Advances in Computer Science and Technology.

[9] Google, "Tango." https://get.google.com/tango/, 2017.

[10] M. D. G. A. S. V. J. S. A. C. G. A. G. S. S. B. M. P. K. R. P. M. D. Singh Vaibhav, Paul Rohan, "'smart cane for the visually impaired: Design and controlled field testing of an affordable obstacle detection system," in *Department of Computer Science and Engineering, Indian Institute of Technology (IIT), Delhi,*.

[11] Visioptronic, "Tom pouce." http://www.visioptronic.co.uk/sites/default/files/bd_fiche_techtom_pouce_gb.pdf.

[12] H. J. D. P. S. K. F. T. Wakita, K., "Human-walking-intention-based motion control of an omnidirectional-type cane robot.," in *IEEE/ASME Trans. Mechatronics 18*, p. 285296.

[13] Y. S. A. M. E. N. M. S. Takizawa, H., "Kinect cane: an assistive system for the visually impaired based on three-dimensional object recognition.," in *IEEE/SICE International Symposium on System Integration (SII)*, p. 740745.

[14] T. Instruments, "Ultrasonic distance measurement boosterpack." http://www.ti.com/lit/ug/tiduai8a/tiduai8a.pdf, 2017.

[15] P. technologies, "Pico flexx camera." http://pmdtec.com/picofamily/assets/datasheet/Data-sheet-PMD_RD_Brief_CB_pico_flexx_V0201.pdf, 2015.

[16] I. T. AG, "Real3 image sensor family." http://www.infineon.com/dgdl/Infineon-REAL3+Image+Sensor+Family-PB-v01_00-EN.PDF?fileId=5546d462518ffd850151a0afc2302a58, 2015.

[17] Google, "Google advanced techonology and projects." https://atap.google.com, 2017.

[18] Google, "Tango developer overview." https://developers.google.com/tango/developer-overview, 2017. January 19.

[19] PCWorld, "Google's project tango headed to international space station." http://www.pcworld.com/article/2110660/googles-project-tango-headed-to-international-space-station.html, 2014.

[20] N. Jet Propulsion Laboratory, Computer Vision Group. https://www-robotics.jpl.nasa.gov/groups/ComputerVision/, 2014.

[21] Lenovo, "Lenovo phab 2 pro specifications." http://www3.lenovo.com/us/en/smart-devices/-lenovo-smartphones/phab-series/Lenovo-Phab-2-Pro/p/WMD00000220, 2017.

[22] Assus, "Zenfone ar." https://www.asus.com/us/Phone/ZenFone-AR-ZS571KL/, 2017.

[23] Google, "Tango motion tracking." https://developers.google.com/tango/overview/motion-tracking, 2017.

[24] Google, "Tango area learning." https://developers.google.com/tango/overview/area-learning, 2017.

[25] Google, "Tango depth perception." https://developers.google.com/tango/overview/depth-perception, 2017.

[26] Google, "Tango camera intrinsics and extrinsics." https://developers.google.com/tango/overview/intrinsics-extrinsics, 2016.

[27] L. P. A. Aladr en, G. L opez-Nicol as and J. J. Guerrero, "Navigation assistance for the visually impaired using rgb-d sensor with range expansion," in *EEE Systems Journal*, pp. 1–11, l.