UPPSALA UNIVERSITET

IT 18 029
Examensarbete 30 hp
Juni 2018

Feasibility Study of a "Smart" Aid for the Visually Impaired and Blind's Independent Mobility in Outdoor Environments

Qian Lu

Institutionen för informationsteknologi
Department of Information Technology

UPPSALA UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
http://www.teknat.uu.se/student

Abstract

Feasibility Study of a "Smart" Aid for the Visually Impaired and Blind's Independent Mobility in Outdoor Environments

Qian Lu

This paper investigates feasibilities of technological solutions for empowering visually impaired and blind people to travel in outdoor environments independently. Through related work on users' spatial models, usage of long white canes, and existing technological aid tools, an in-depth understanding of users and problems of independent travel is developed. The main constituents of the independent travel are object detection, navigation, and orientation. While long white canes are very useful for supporting visually impaired and blind people to travel independently, the usefulness is limited to near-space environments. This understanding is later validated using face-to-face interviews, online surveys, and field observations. Based on the understanding, I proposed a design to utilize machine learning mechanisms and computer vision in order to provide real-time and customized support for independent travel in conjunction of use of long white canes. With feedback from user evaluation of a lo-fi prototype of the design proposal, the design concept showed potential abilities of empowering users with travelling on their own. The proven concept initiated a start of a series of future work for possible implementations.

This paper is focused only on support for individuals with visual impairments to navigate outdoors. Thus, indoor navigation is not specifically addressed, and deafblind people are not covered in the target audience.

Handledare: Lars Oestreicher
Ämnesgranskare: Mats Lind
Examinator: Anders Arweström Jansson
IT 18 029
Tryckt av: Reprocentralen ITC

ACKNOWLEDGEMENTS

I would like to extend sincere thanks to all the people who helped me during the thesis work:

Mikael Ståhl, Henrik Götesson, Lennart Karlsson, and all participants in the user studies for their time and suggestions;

Huria Hussein for assisted me with user evaluation on the design proposal;

Lars Oestreicher and Mats Lind for supervising and reviewing the thesis;

Yunfeng Xu and Sanbao Xu for advising me on feasible technologies and technical requirements.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The number of people who have lost their vision completely or have different severities of visual impairment is increasing. According to the World Health Organization (WHO) estimate of 253 million people living with vision impairment, 36 million are blind and 217 million have moderate to severe vision impairment (World Health Organization, n.d.). Globally, chronic eye diseases are the main cause of vision loss. Among the 253 million people with different levels of visual impairment, 81% are aged 50 years and above (World Health Organization, n.d.). In Europe, about 65% of the visually impaired are 50 years of age or older, among which about 90% are over the age of 60 (World Health Organization, n.d.). With an increasing population of older people, more people will be at risk of vision impairment due to chronic eye diseases (World Health Organization, 2012). Also, refractive errors cause an estimated 12 million children below age 15 to lose their vision partially or completely. Around 1.4 million children have irreversible blindness, requiring access to vision rehabilitation services to optimize functioning and reduce disability (World Health Organization, 2012).

Independent mobility is a precondition for people with visual conditions to live a quality life (Golledge, R.G. et al., 1996). However, due to vision loss, mobility can be very challenging for blind people. As a result, the life quality of blind and visually impaired people is affected negatively. Accidents happen especially in mobility-related tasks. Manduchi and Kurniawan (Manduchi & Kurniawan, 2011) conducted a survey interview in 2010 on accidents that blind or legally blind people have experienced during their mobile tasks. In their findings following interviews with 300 individuals, around 34% of legally blind and more than 45% of blind participants experienced head-level accidents once a month or less while about 18% of legally blind and 9% of blind interviewees experienced the same type of accidents more often, on a monthly basis. 23% of these accidents had medical consequences. In total, 7% experienced falls in walking once a month or more. The supportive tools long white cane and guide dogs could not prevent these accidents, due to their limitations in object detection above knee-level, navigation, and orientation.

Brambring's travel model (Brambring, 1985) as shown in Figure 1 demonstrates that locomotion, also known as mobility, is realized by perception of objects and perception of orientation. Visually impaired and blind people use the identification of landmarks to determine their specific locations on a route for determining orientation and navigation.

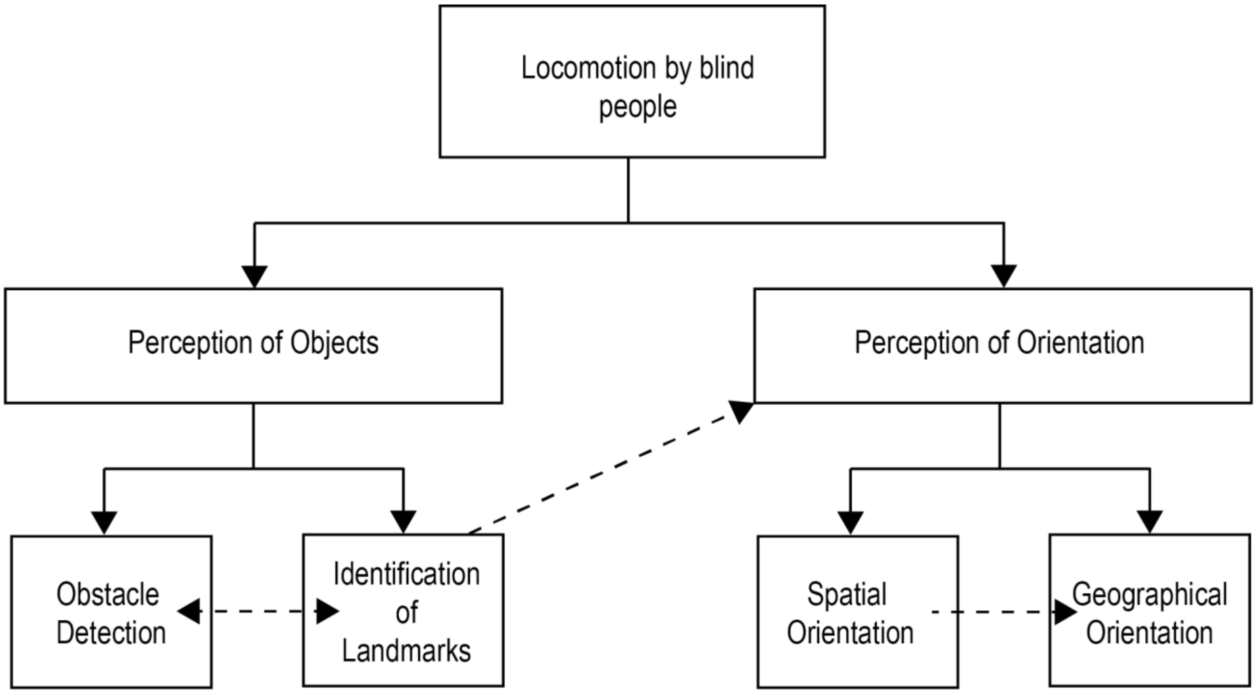

Figure 1 Brambring's decomposition of travel by blind people (Keating, 2008)

In order to address these difficulties and empower blind and visually impaired individuals in their independent mobility, many technological assistive tools have been introduced since World War II. However, these tools are not widely used and adopted by the blind group. The factors that hold back the widespread usage of technological mobility aids can be attributed to poor satisfaction of users' needs such as user experience, affordable purchase price and usability. Roentgen, et al.'s literature review on 146 electronic support tools indicates that none of the reviewed tools can substitute white canes or guide dogs but serve as secondary tools besides the use of white canes or guide dogs as primary support (Roentgen, Gelderblom, Soede, & Witte, 2008).

This paper intends to unveil needs of the visually impaired in their daily mobility tasks and to propose a design of a feasible smart aid that can fulfil their needs together with functions provided by regular long white canes in order to encourage them to travel independently. It is inspired by project work with the same general theme done by Huria Hussein and me for course Non-Excluding Design and Evaluation at Uppsala University.

The discussion of this paper is structured upon the people, activities, contexts, and technologies (PACT) design framework. It starts with a review of the current status of related work and an inventory study of existing technological solutions and technologies for supporting blind and visually impaired people in terms of independent mobility, with an aim to understand them and their obstacles in independent travel. User studies present what activities they do and what needs they have during independent travel, touching on the contextual environment.

Following the discussion on the people and activities, an in-depth discussion is placed on relevant contexts and technologies for exploring feasibilities of a design for the “smart” aid. Based on these discussions, a design proposal is made. After user testing on a low-fidelity (lo-fi) prototype, the results are presented and discussed. Finally, necessary future work is suggested.

1.1 TERMINOLOGY

The following terminology is used in this paper:

1. Azimuth

   Horizontal coordinate

2. Blindness

   The term blindness refers to visually impaired and visually challenged people. Blindness can be defined as the loss of sight either partially or fully. It can be temporary or permanent either damage to any portion of the eye, the optic nerve, or the area of the brain responsible for vision can lead to blindness ("Medical Definition of Blindness," n.d.). In this paper, the term "blind" is used to encompass all forms of visual impairment.

3. Congenital blind

   Congenital blind people were blind at birth or lost their vision soon after birth (Wan, Wood, Reutens, & Wilson, 2010).

4. Early Blind

   Early blind people became blind before age 14 (Wan et al., 2010).

5. Elevation

   Vertical coordinate

6. Geographical orientation

   A far-space activity. It refers to the ability of travelers to determine a position in the geographical space of the entire journey. It is determined by information and strategy about how to get from place A to place B. Navigation is the term often used in this context (Keating, 2008).

7. GPS and GNSS

   Global positioning system (GPS) is one of many satellite constellations that are collectively known as Global Navigation Satellite System (GNSS). While GPS is developed by the US military, there are many other developers of satellite constellations, such as Russia and China. In this paper, GPS is used interchangeably with GNSS, due to its great prevalence of use ("GPS vs GNSS - What Engineers Need To Know For Their Design," n.d.).

8. Late Blind

   Late blind people lost their vision after age 14 (Wan et al., 2010).

9. Legal blindness

   A level of vision loss that has been legally defined to determine eligibility for benefits (American Foundation for the Blind, n.d.).

10. Spatial orientation

A near-space activity. It refers to the ability of travelers to acknowledge a position in their immediate environment. For example, finding an optimal position on a path and the ability to walk without veering from a desired course (Keating, 2008).

11. Visually impaired

    A visually impaired person has been clinically determined to have a visual acuity of 20/70 or less in the stronger eye, while a person who is legally blind is defined to have a visual acuity of 20/200 or less in the stronger eye. In the US, people whose visual acuity is at either of these levels receive government benefits, such as the right to possess a white cane or own a guide dog (American Foundation for the Blind, n.d.-a).

2 PURPOSE

The purpose of this thesis is to find technologically feasible ways to improve safe and independent mobility for visually impaired and blind people.

2.1 RESEARCH QUESTIONS

In order to design a technological aiding system that supports visually impaired and blind people for independent mobility, I will need to look into the following research areas:

1. Investigate the validity of previous findings on the needs of visually impaired and blind individuals in independent mobility.
2. Uncover additional requirements for supporting visually impaired and blind individuals to complete daily mobility tasks.
3. Find feasible and robust technologies for supporting independent mobility cost-effectively.
4. Evaluate conceptual designs.

2.2 DELIMITATION

Due to time and budget constraints, only a limited number of individuals with complete vision loss and severe level of visual impairments could be recruited.

This paper is focused only on support for individuals with visual impairments navigating outdoor. Thus, indoor navigation is not specifically addressed, and deafblind people are not covered in the target audience.

3 RELATED WORK

Passini, et al. (Passini, Dupré, & Langlois, 1986) stated that "Spatial orientation and wayfinding are the foundations of mobility". In order to support people with visual impairments, electronic solutions have been developed to assist or replace the conventional tools for people's mobility, navigation, and object detection (Csapó, Wersényi, Nagy, & Stockman, 2015; Keating, 2008). Many research groups and organizations have investigated and proposed supportive tools to empower visually impaired and blind people in terms of orientation and navigation to support independent travel and free mobility (S. Y. Kim & Cho, 2013; Roentgen et al., 2008). In a literature review on then existing electronic support aids for mobility conducted by Roentgen et al. from October 2007 to March 2008 (Roentgen et al., 2008), an inventory of 146 different devices was classified into two categories based on their purpose: 1) object detection and orientation; 2) navigation. A majority of the aids for object detection and orientation are no longer available or never productized. The failure to enter in the production phase and to the market is a consequence of imprecise positioning, insufficient output interactions, and users' dissatisfaction regarding how the product addressed their needs (Kammoun et al., 2012). High prices (Abdel-Wahab & El-Masry, 2011; Mau & Melchior, 2008). Poor user experience (Lacey, Namara, & Dawson-Howe, 1998) , and problematic usabilities (S. Y. Kim & Cho, 2013) also attributed to the limited adoption of the electronic aids. (Roentgen et al., 2008)Roentgen et al. (Roentgen et al., 2008) discovered that good orientation and mobility (O&M) skills are essential for using sensor-based devices. However, good O&M skills can only be obtained through extensive trainings. Furthermore, all the 146 sensor-based products in their inventory study emit either acoustic output, such as ultrasound or optic output, such as laser and infrared beams. The sensory systems detect and process the beams that are reflected by objects in their detection range in order to construct perceptible information for users (Farmer & Smith, 1997). There are many factors that affect the detection range. For example, sensor numbers, their positions, and mounting angles. So far, there has not been a successful electronic substitute to the traditional white cane during the 40 year development of electronic mobility aids (S. Y. Kim & Cho, 2013).

In the following sections, I will firstly investigate important aspects of how the visually impaired and blind generate spatial models. Secondly, I will review the long white cane and consider its advantages and disadvantages. Thirdly, I will present an inventory of existing assistive technologies in order to discuss design possibilities in an up-to-date context, given that other technologies than sensory systems have been developed and proposed for designing electronic mobility aids since the study of Roentgen et al. (Roentgen et al., 2008).

3.1 USER VISION AND PERCEPTION

The word vision is usually related to seeing with eyes. Nevertheless, Cattaneo and Vecchi argued in the book Blind Vision (Cattaneo & Vecchi, 2011) that the visual modality is not a necessity for generating mental representations. The mental representations can be generated from information that is acquired from haptic, auditory, and olfactory modalities or semantic knowledge in long-term memory, even though visual information acquired by a single gaze is undeniably massive and in great detail.

They also stressed that semantic knowledge includes visual characteristics. For example, even though people who lose their vision at birth never see the color yellow, they still know the color of bananas is yellow.

For early blind people, their memory is developed from other sensors except visual sensory modality. In a study conducted in 2001 (Ogden & Barker, 2001), tactile, auditory, and spatial (excluding visual) information were reported to exist in the memories of the early blind. The same is found in recent memories of late blind people, although their early memories contain visual information besides tactile, auditory, and spatial details. In contrast, visual information is dominant in memories of sighted participants. However, without vision to discover detailed characteristics of what is perceived while capturing a huge amount of information, blind individuals are still capable of generating vivid mental images and making use of them just like sighted counterparts. In several experiments conducted by Cornoldi, et. al. for rating vividness of imagery nouns, blind people demonstrated performance similar to sighted participants (Cornoldi, Calore, & Pra-Baldi, 1979; Tinti, Galati, Vecchio, De Beni, & Cornoldi, 1999). In particular, blind participants outperformed sighted participants in words related to auditory while the two parties had a similar performance at words related to visual. In another experiment investigating participants'ability to create images of scenes based on descriptions of objects that are visible and hidden in the scenes, both groups of participants remembered more visible objects in the described scenes (Zimler & Keenan, 1983).

Although visually impaired and blind people rely primarily on auditory, tactile, and olfactory sensory modalities, it is not assertive that these modalities are significantly enhanced for them, as there is not a consistent definition of the sensory compensation. Therefore, I will not consider enhancement of auditory, tactile, and olfactory sensories of the visually impaired and the blind as the design input.

3.1.1 Sound localization

A sound source can be located in the azimuth, elevation and range spatial dimensions that build up spatial localization. The term sound localization is defined as the result of a perceptual process of finding azimuth, elevation and range spatial dimensions of a sound. Sound localization is an important part of auditory experience (Cattaneo & Vecchi, 2011). The process of sound localization is described as follows in sequence:

1. At first, multiple acoustic sounds, such as differences of interaural intensity, are integrated.
2. Secondly, the integrated sounds are used to find the azimuth (horizontal coordinate) of the sound that uses the head as a relative base coordinate.
3. Thirdly, the pinna of the ear provides the analysis of spectral shape cues for locating the elevation (vertical coordinate) of the sound. Simultaneously, the sound can be distinguished from front and rear sources in range spatial dimension.

A few experiments have suggested that visually impaired and blind people have the same performance as sighted individuals in terms of locating sound that is located in the center of the space while sighted people were outperformed at locating peripheral sound in the space (Gougoux et al., 2004). Studies also show that visually impaired and blind people utilize echo cues for navigation. The term echo is defined as a sound that is heard after it has been reflected off a surface such as a wall or a cliff (Cambridge English Dictionary, n.d.). It is the perception of two sounds that are originated from one single sound source because they arrive at different time due to different transmission paths. Evidently, blind people utilize echo cues for navigation (Cattaneo & Vecchi, 2011). Nevertheless, the finding that absolute errors made in orienting heads to sounds by people with visual impairment and blindness are much greater/more frequent than for sighted individuals suggests that calibration in elevation plane localization relies more on visual input. The same finding is not discovered in azimuth plane localization using sound (Lewald, n.d.; Zwiers, Van Opstal, & Cruysberg, n.d.).

3.1.2 SPATIAL AUDITORY MAPS GENERATION

The fact that sighted newborn babies can orient themselves to sounds before their visual orientating abilities develop indicates that visual input is not essential for generating spatial auditory maps (Morrongiello, 1994; B. Röder et al., 1999). This is further supported by an experiment about spatial auditory capabilities of the blind conducted by Lessard et al. (Lessard, Pare, Lepore, & Lassonde, 1998). In their experiment, both groups of early blind and blindfolded sighted people demonstrated high performance when porting their hands to the perceived sound source, which implies that the lack of visual information does not affect the generation of three-dimensional spatial maps. In general, people with visual impairment and blindness can generate mental images from auditory, tactile, and olfactory input that accurately encloses spatial relations and non-visual details (Cattaneo & Vecchi, 2011).

However, studies also show that visually impaired and blind people have trouble perceiving unfamiliar environment and objects at the first encounter. Nevertheless, once they have experienced unfamiliar places and objects, their perception of the places and objects is recoded. As a result, they will have fewer problems with perception/perceiving the environment when visiting the same places and experiencing the same objects, which is the same as sighted subjects (Cattaneo & Vecchi, 2011). In the theory of Cattaneo and Vecchi (Cattaneo & Vecchi, 2011), blind people perceive spatial information and environmental object shapes mostly in analogy except for certain differences caused by their naturally dominant sensory modalities. For example, lines are perceived just as lines by them.

Besides spatial relations, the rules of perspective are also important for constructing spatial auditory maps. The term perspective refers to the way that objects appear smaller when they are further away and the way parallel lines appear to meet each other at a point in the distance (“Perspective,” n.d.). In some experiments about visually impaired and blind people using the rules of perspective when perceiving objects, both blindfolded sighted people and late blind people showed the ability of using the rules of perspective in the mental representations, while early blind and congenital blind individuals do not use the perspective rules in their mental imagery (Arditi, Holtzman, & Kosslyn, 1988; Vanlierde & Wanet-Defalque, 2005). The findings that blind people do not apply perspective rules by nature implies that telling distances may be difficult for people who are congenital blind and early blind. Nonetheless, they can learn about rules of perspective from the perspective descriptions of sighted individuals and then use their mental images in great details to visualize them (Cattaneo & Vecchi, 2011). This is in fact in line with their similar performance at rating imagery words (see discussion in 3.1).

Although people who have vision impairment and vision loss have demonstrated the same ability as sighted individuals for generating mental representations that contains spatial information but with non-visual input, they are not able to develop the mental images all at once, but instead based on objects that are gradually encountered within their reachable range and are not aware of objects that are beyond reach. In contrast, an enormous amount of information can be perceived by just a glance, including those in a far distance, with the help of vision (Cattaneo & Vecchi, 2011). In a word, sighted people are able to gain a preview of the surrounding environment within a short time while visually impaired and blind people cannot. As a result, they need to continuously recode perceptual experience of objects as they move, which is very cognitively demanding (Cattaneo & Vecchi, 2011).

3.1.3 PERCEPTION POSITION

There are two approaches of locating objects. One is ego-centered (i.e. body-centric), the other is external reference centered, that is using an object in the perceived environment as a reference point. People with visual impairment and blindness, especially congenitally blind and early blind people, adopt mostly the ego-centered approach in all sorts of tasks of mobility travel: way-finding, object location, spatial information encoding, as they cannot take the information about all objects in the environment all in at once (Cattaneo & Vecchi, 2011; Kammoun et al., 2012; Keating, 2008). For other tasks that are concerned with exploration, the external reference centered approach is more robust (Keating, 2008).

3.1.4 MEMORIES AND CAPTURE OF SPEECH

In terms of short-term memory, such as word span, studies have shown that visually impaired and blind people have better capacities compared to sighted individuals. This is also true with respect to long-term memory. They also can differentiate meaningful sounds, such as speech, at a very fast tempo. Usually for sighted people, temporal discrimination of sounds is around 8-10 syllables per second, which is outperformed by people with visual impairment and blindness who can understand up to 25 syllables every second (Bull, Rathborn, & Clifford, 1983; Hull & Mason, 1995; Brigitte Röder & Neville, 2003; Brigitte Röder, Rösler, & Neville, 2001).

3.2 ADVANTAGES AND DISADVANTAGES OF A WHITE CANE

The white cane originated in Europe in 1921 when James Biggs, a photographer who had lost his vision, began to paint his walking cane white to alert others to his presence (Lions Club International, 2014). When veterans of World War II returned to America with vision impairment and blindness they wanted to have the same level of independence as they had before the war. Because of this, the white walking cane was altered into the long cane form that is still prevalent today (Strong, 2019).

A white cane is often carried by the blind and visually impaired to give more freedom to the individual. The two main functions of the cane are identification and safety: it should notify the seeing pedestrians and drivers that the user has some degree of vision loss and also alert the user to obstructions and changes in their path (Canadian National Institute for the Blind, n.d.).

The use of a white cane usually focuses on two major aspects: grip and arc. For outdoor use, where a person’s pace is faster and more regular, the proper grip used to hold the white cane is the palm facing up at waist height with the index finger pointing along the cane and the remaining fingers and thumb wrapping around the cane lightly. When indoors or in a more congested environment, such as a crowded city street, the grip changes in such a way that the user holds the cane as if it were a pencil: upright, at sternum height, and closer to the body. With both grips, the elbows are kept tucked close to the body (Harada, Kaneko, Hirahara, Yanashim, & Magatani, 2004). The second component, the arc, refers to the sweeping motion of the cane performed by the user. The user sweeps the cane over an area just larger than shoulder width, tapping the ground on the opposite side of the foot currently taking a step-in order to prepare for the next step (for example, tapping the ground to the left of the body when stepping forward with the right foot.)

There are several types of white canes available for different uses as elaborated in Table 1. Although the white cane was introduced around 1940, only 2% of visually impaired people use long white canes as a navigation tool for performing daily activities independently (Roentgen et al., 2008). This can be attributed to that most of visually impaired people do not travel alone outside of their own residence due to risk of accidents as well as lack of preview and knowledge of environments and information for orientation (Manduchi & Kurniawan, 2011; Roentgen et al., 2008).

Table 1 Different Types of White Cane

White Cane Type	**Length**	**Advantage**	**Disadvantage**
Long cane	It extends from the floor to the user's sternum. The shaft is generally 115 to 165 cm long (“Whitecane,” n.d.).	Navigation and object detection	Walking support; Hard to bring when travelling due to its length
Folding cane	70 - 135 cm (“Ambutech,” n.d.)	Navigation and object detection Easy to bring when travelling due to its length	Walking support
Support cane	80 cm - 105 cm (“Ambutech,” n.d.)	Walking support	Navigation, object detection, and mobility
ID cane	Shorter than long canes	Alerting others of users’ visual impairment	Navigation, object detection, and mobility
Guide cane	Generally extending from the floor to the user's waist	Navigation and object detection	Walking support

Identification canes are short (reaching only to the user’s waist), provide little to no protection, and are generally more popular with the visually impaired who only want to alert others of their impairment. Support canes have the same purpose as identification canes, except that they provide more support and balance for the legs and body of the user. Long canes, the type of cane chosen to be modified into a Smart Cane, reach the user’s sternum and provide the most safety for the user, alerting them of terrain and height changes, walls, doors, and obstacles. They are also the most visible to others (“Why Would Someone Need a White Cane?,” n.d.).

Advantageous features that are attributed to different types of white canes are scattered and distributed, meaning that there is no single cane to meet all sorts of needs in terms of navigation, orientation, and object detection.

In order to use the white cane properly, people need to undergo intensive training that is provided by orientation and mobility (O&M) instructors (Sander, Bournot, Lelièvre, & Tallec, 2005). Walking with the white canes is widely concerned with two techniques: Two-point touch technique and constant contact technique (D. S. Kim, Wall Emerson, Naghshineh, & Auer, 2017).

Although the different walking techniques together with cane properties such as length and weight have influences on walking efficiency (D. S. Kim et al., 2017), the important determinant of the walking efficiency is preview of surroundings. For example, with the preview of the environment provided by a Sonic Pathsounder (Kay, 1974), an electronic mobility aid that was used together with a long white cane, the visually impaired can walk 18% faster than merely with a long white cane (Clark-Carter, Heyes, & Howarth, 1986).

Besides limited mobility provided by a white cane, it is also impossible for the visually impaired to detect objects that are above knee level and gain a preview of object placements in surrounding environment merely with a white cane as a navigation and orientation tool (Sander et al., 2005). The use of a white cane may also be stigmatized. The more serious visual impairment that people have, the lower socioeconomic status they will gain (Lansingh, Carter, Ulldemolins, Valencia, & Eckert, 2012). Many have felt ashamed to use a white cane, because normally sighted people do not think a white cane is a normal tool for the visually impaired to gain independence (Nichols, 1995).

Further, the condition of an environment also contributes to the efficiency of using the white cane. As a white cane has a certain length, regardless of the cane type, neither of the two-point touch or constant contact techniques can be utilized to its full extent where space is limited in a surrounding environment that may not accommodate the use of a white cane (Nichols, 1995). The low confidence of visually impaired and blind people will have a negative impact on mobility, as Clark-Carter et al. (Clark-Carter et al., 1986) discovered.

3.3 INVENTORY OF EXISTING TECHNOLOGICAL SOLUTIONS

In this section, an inventory of existing technological assistive aids is presented and discussed in terms of the purposes served, usability, affective experience, ethical considerations, and stigmatization. The discussion is focused on both primary aids that are used on their own for safe mobility and secondary aids that are supplemented to primary aids (Keating, 2008).

The methods that are used to make the inventory of the existing travel aids include search in academic libraries, such as Uppsala University library and IEEE, and online search on Google with keywords “assistive technology”, “electronic travel aids”, “electronic orientation aids”, “navigation for the visually impaired”, “orientation for the visually impaired”, “object detection for visually impaired”, and “support tools for the visually impaired” as well as variants of these keywords (Cattaneo & Vecchi, 2011; Csapó et al., 2015; Keating, 2008; S. Y. Kim & Cho, 2013; Roentgen et al., 2008).

As there is no published scientific paper concerned with these specific solutions, the solutions are compared and classified by their adopted technologies. The review of the inventory is made based on academic papers about the employed technologies and their general implementations as well as online resources that include contents on their office websites and their product descriptions. Generally, there are two types of solutions: environmental adaptation-independent solutions that are used individually without support from other systems or facilities in the environment, and environmental adaptation-dependent solutions that are used together with other systems and facilities in the environment that are deployed as part of the solutions, as shown in Table 2.

Table 2 Types of Assistive Technologies

Category	**Assistive Technology**
Environmental adaptation-dependent	Bluetooth beacons; Internet of Things (IoT) Radio-frequency identification (RFID)
Environmental adaptation-independent	Augmented Reality for Visual Impairment (AR4VI); BlindSquare; Camera-based computer vision; GPS Ultrasound; Laser light; Infrared; Seeing AI;

3.3.1 AUGMENTED REALITY FOR VISUAL IMPAIRMENT

Augmented reality (AR) can provide accessible annotations and auditory, haptic, tactile and visual enhancements to objects in the surrounding environment (Coughlan & Miele, 2017; Ichikari, Yanagimachi, & Kurata, 2016; Peli & Woods, 2009). James et al (Coughlan & Miele, 2017) coined the terminology “Augmented Reality for Visual Impairment” (AR4VI) for these kinds of AR tools that allows people with visual impairment and vision loss to explore their surroundings by natural gestures with implementation of a gestural interface. They also classified AR4VI tools into two categories: global applications that augmented the surrounding environment in near and far space and local applications that people can touch and point in their near space.

Several such AR tools have been developed to provide visual enhancement and auditory feedback for visually impaired and blind people (Coughlan & Miele, 2017; Pyun, Kim, Wespe, Gassert, & Schneller, 2013), one of which is eSight. eSight is a headset that is similar to virtual reality (VR) gear and uses a HD camera to capture visual information in the surroundings as input and two OLED displays as show the captured visual information as enhanced output. The input video is enhanced in terms of contrast and quality using a video processing algorithm (eSight, n.d.; Gartenberg, n.d.).

However, AR4VI-based solutions are focused primarily on employing text-based information for environment augmentation, which means they are useful for object detection for the visually impaired but not fully capable of providing environment preview that is important in ensuring safe navigation.

3.3.2 Beacon Technologies

The employment of beacon technologies enables people to trigger actions on their mobile phones, point-of-sale terminals, digital displays, lighting, or any appliance that is available on the Internet, as they move around in the physical world (Statler, 2016b). Through pre-installed beacons, Wayfindr is an audio-based system that includes a mobile app to give people who have visual impairments instructions about navigation and orientation that are triggered at significant waypoints of their journeys ("wayfindr," n.d.). Installed beacons send small Bluetooth pulses and set intervals that are received by a mobile phone.

The reception of the Bluetooth pulses will trigger an app to instruct the user where and how to proceed in the journey ("wayfindr," n.d.). When Kevin, a blind, participated in the testing of Wayfindr, he described this experience as his first steps into empowered independence and felt Wayfindr was a sighted guide walking side by side with him to go through the journey ("How Wayfindr Guided My First Steps to Independence on the Tube," n.d.). Wayfindr early testing hasshownthat a good coverage of beacon signals plays an important role inindependent navigation and orientation for the visually impaired, which requires collective work on environment adaptation that can affect its official rollout plan.

Furthermore, as this solution is merely focused on navigation and orientation but not object detection, users with visual impairments will encounter difficulties as when merely using the white cane in detecting objects that are above knee level. If the users do not use long white canes but other types of white canes, the detection ability will decrease due to their lower detection radius as discussed previously. Since generally beacon signal strength can reach a 50-meter range (Statler, 2016a), theoretically it's possible to install beacons on objects that are above knee-level in environment to provide a preview of the surroundings to the blind. However, economic- and time-wise, it's not possible, as there are many more objects existing in the world other than the manmade. Therefore, this solution should be integrated with object detection solutions discussed in the previous sections that do not require environmental adaptations.

3.3.3 Camera-Based computer vision

BMVA (The British Machine Vision Association and Society for Pattern Recognition) defines computer vision: "Humans use their eyes and their brains to see and visually sense the world around them. Computer vision is the science that aims to give a similar, if not better, capability to a machine or computer. Computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images. It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding." ("What is computer vision?," n.d.)

Currently in the market, there are two kinds of camera-based computer vision solutions: one uses high speed cameras to capture images of surroundings and enhance the contrast and quality; the other uses cameras to capture images of the contextual environment and analyze the images to inform users with respect to objects and directions using audio.

Through the enhanced contrast and quality of environmental images, legally blind people can better perceive the world. Most of them have lose their central sight but their peripheral vision is not harmed.

However, this approach still uses vision to communicate the information to people who are visually impaired (Gartenberg, n.d.). Therefore, it is not suitable for people who have lost their sight completely. For them, information can be perceived via auditory modality, which implies that they can listen to what the assistive tools tell them about the environment after the tools have captured objects that may affect their safe mobile travel.

Another example of camera-based solution is Seeing AI ("Seeing AI," n.d.). It is a mobile app that is designed for people with low vision and describes what appears in the camera in real time and also can indicate what is in the photos that are taken by the camera of their phones. It is intended to help people know who and what is around them and read textual information. Seeing AI is only available free of charge in iOS. One more instance of camera-based solutions is OrCam. It is a standalone device, compared to Seeing AI that utilizes mobile phone cameras. It supports reading texts, recognizing faces, and identifying products in real time ("OrCam User Stories," n.d.).

Based on the aforementioned characteristics and serving purposes of camera-based solutions, it is possible to use them for detecting objects and assisting with orientation and navigation based on processed images. With them, people with visual impairments and blindness will not have as much cognitive workload as they have today. However, their current usage is limited: not every aid is intended for people with vision loss. For example, eSight and Seeing AI; Not everything can be detected in real time except being detected in static photos.

Although camera-based vision technologies have not yet received much attention in commercial markets, academic research is making progress in developing the potential of such systems for detecting objects and supporting orientation and navigation to support independent travel for people with vision loss. In the research project NAVIG (Navigation Assisted by embedded VIsion and GNSS, 2009–2012) (Kammoun et al., 2012), computer vision was explored for perceiving the current environment with visual object locations and providing accurate information about locations in conjunction with DGPS and GIS. They made design prototypes and evaluations, although I could not find information about the prototypes being productized. There are also papers that discuss using technologies developed for autonomous vehicles for designing supportive tools for visually impaired and blind people. NAVIG discoveries and driverless vehicle technologies will be discussed in greater detail in later sections.

3.3.4 GPS

Global Positioning System (GPS) is an accurate worldwide navigational and surveying facility based on the reception of signals from an array of orbiting satellites (Oxford Learner's Dictionaries, n.d.). It is now a standard feature on smartphones that most users take with them every day. There are many navigation and location tools based on GPS. So far, not a single solution can replace GPS, although GPS has an obvious shortcoming in that it cannot be used in indoor settings (Coughlan & Miele, 2017).

BlindSquare is a mobile app that is used for navigation and travel planning by most of the participants in the user studies. It claims to be the world's most popular GPS based app for visually impaired and blind people. It provides information about points of interest, descriptions about the environment as well as intersections of streets "on the go" along your route via voices. Users can interact with the app via voice control (BlindSquare, n.d.).

However, GPS still cannot sufficiently provide information for safe and effective navigation, due to the following reasons: Differentially corrected GPS (DGPS) is not available in many locations; Signals from satellites can be blocked in narrow built-up areas; GPS-based systems are based solely around the GPS function not the contextual information, which results in lacking support for dynamically changed environment and context-awareness; GPS positioning precision is insufficient for pedestrians (Keating, 2008).

3.3.5 INFRARED

Compared to GPS, infrared technology Remote Audible Infrared Signage (RIAS), is a viable navigation alternative in indoor environments (Coughlan & Miele, 2017). Besides usage in navigation, RIAS also has proven ability in object detection. For example, Tom-Pouce, an example of RIAS adoption, emits infrared beams that allow the detection of head-level obstacles (Damaschini, Legras, Leroux, & Farcy, 2005). Nevertheless, as infrared can be disturbed by sunlight and dark objects, the infrared adoptions/adaptations do not work accurately outdoors (Kanagaratnam, 2009).

3.3.6 IOT

The Internet of Things (IoT) refers to objects increasingly become smart and connected (Statler, 2016a). Handisco is a secondary device that is attached to a white cane used by a blind person. It interoperates elements, such as traffic lights, in an environment that is essential in navigation and object detection as well as orientation on network ("HANDISCO," n.d.).

The device uses ultrasonic waves and GPS technology to gather data from traffic lights, bus locations, and crosswalks to provide-real time navigation for the blind. Tagged sensors at different shops and restaurants communicate with the device, telling the user what the shop sells, where the entrance is, and the business hours ("A Smart Stick Puts Digitization in Service to the Blind," n.d.). Thus, it inherits the features of the host white cane with enhancement of object detection using IoT and adds on support for orientation. However, as discussed about beacon technologies, environmental adaptation requires economic- and time-wise contributions, which is an obstacle to rolling out the solution to mass public (Keating, 2008).

With a look at the implementation details, the Handisco website claims that the device has a relatively long battery life span of 10 hours per charge, which can satisfy daily use. However, due to the use of GPS (Brabyn, Alden, Haegerstrom-Portnoy, & Schneck, 2002). The usage scenarios can be limited to outdoor mainly and the device has a spatial resolution of approximately 10 meters in urban environment when it comes to navigation.

3.3.7 LASER LIGHT

As an alternative for object detection, the drawback of adopting laser light is a narrow detection range. To overcome this shortcoming, users should continuously use two-point touch technique with which users tap the tip of the cane side by side (S. Y. Kim & Cho, 2013), which can be energy-consuming.

3.3.8 ULTRASOUND

In contrast to infrared and laser light, ultrasound has the fewest limitations. It has a wide detection range that can stop users from constant need of tapping their white canes. Ultrasound sensors detect obstacles up to head-level and at long distances (S. Y. Kim & Cho, 2013). In Kim et al's study, all of test participants who are visually impaired gave the highest rating to the smart cane prototype that uses ultrasound sensors (S. Y. Kim & Cho, 2013).

3.3.9 SUMMARY

Navigation, object detection, and orientation are very important factors for ensuring safety and independence of the blind (Alkhanifer & Ludi, 2014; Cattaneo & Vecchi, 2011; Keating, 2008; Lahav, Schloerb, Kumar, & Srinivasan, 2011; Marin-Lamellet, Pachiaudi, & Le Breton-Gadegbeku, 2001; Schinazi, Thrash, & Chebat, 2016). Every existing solution is not inclusive of all the three important factors and only focused on one or two of these factors. Generally, the solutions that require adaptation of environment need more time and financial support, compared to those that do not depend on environment to adapt.

4 METHODS

As the research questions are concerned with users, their activities, and tools and technologies of independent mobility tasks that happen in different contexts, I employed the PACT framework to modularize this study and to find answers to the research questions in the form of investigations that are conducted in three consecutive phases, as the PACT framework is useful for finding the proper integration of technologies to support users during their activities in different contextual scenarios (Benyon, 2014).

To answer the following research questions 1-2 that are concerned with people and activities of the PACT framework, I conducted user interviews and field studies with recruited participants of various types of blindness to understand needs of visually impaired and blind people in mobility tasks.

1. Investigate the validity of previous findings on the needs of visually impaired and blind individuals in independent mobility.
2. Uncover additional requirements for supporting visually impaired and blind individuals to complete daily mobility tasks.

To answer the following research question 3 that is concerned with technologies of the PACT framework, I made an investigation into assistive technologies in order to find feasible technologies that can be used to implement the discovered improvements.

3. Find feasible and robust technologies for supporting independent mobility cost-effectively.

To answer the following research question 4, I proposed a design for an improved technical aiding system to solve users' needs in mobility activities, which is followed by user evaluation on a lo-fi prototype, as evaluation is essential to interactive system designs (Benyon, 2014).

4. Construct and evaluate conceptual designs.

4.1 USER STUDIES

User studies were conducted using face-to-face interviews, online survey, and field studies.

4.1.1 FACE-TO-FACE INTERVIEWS AND ONLINE SURVEY

Due to different causes, people become visually impaired at different stages: congenital blind, early blind, late blind (Wan et al., 2010) (See Section 1.1). In order to understand how their spatial models are generated as well as discover new user needs in mobility work, I contacted Synskadades Riksförbund (The Swedish Association of the Visually Impaired), also known as SRF, and recruited 11 individuals who are of different blindness levels for face-to-face interviews and field studies with help from SRF.

In addition, the interview questions were distributed in the form of online survey with aim reach as many different blind individuals as possible. Online surveys are flexible, stress-free, and convenient that users can answer at their own pace (Evans & Mathur, 2005). SurveyMonkey® was used instead of other survey tools due to its good accessibility, as suggested by one contact person at SRF.

In the interviews and online survey, participants were asked the following questions:

1. Age
2. Gender
3. Which country/city do you live?
4. Is your blindness congenital or acquired later in life?
5. How long have you been blind?
6. What support tool(s) do you use? For example, a white cane.
7. For what do you use the support tool(s) in your everyday life? [For example, navigation, object detection, and orientation.]
8. How do you think the ability of the support tool(s) for detecting objects? [bad (0)-fine (5)-good (10)]
9. How do you get your orientation?
10. If you lose your orientation, how do you regain it?
11. What do you need to know in order to find your way to a certain place?
12. What kind of phone do you have?
13. How do you use your phone in daily activities?
14. What do you think about the map on your phone?
15. What technological services for visual impairment do you know? [For example, Blindsquare, Be My Eyes.]
16. What is your personal experience of the smart devices, if you have used them?
17. What is your opinion about them?
18. If you were to be asked to design a white cane or another support tool, what features would you design?

4.1.2 Field Observation

In the field observation, I will follow individual participants from a starting point to a destination of their respective travel that are chosen by themselves. For instance, a participant can choose to walk from home to a supermarket nearby in a field observation.

During the observation, the observed participants will wear an action camera in their chests to capture the movements of long white canes and the encounters it would have in the environment, while I will use a camera and an audio recorder for note taking.

4.2 Feasible Technologies

The related work indicates that there are basically two types of technical aids: 1) Environmental adaptation-dependent; 2) Environmental adaptation-independent. As Keating, et al. (Keating, 2008) suggested, the first category requires investments of money and time which prevents it from being implemented to great number of people in need in a reasonable time frame.

Also, from the related work, long white canes are good at providing tactile feedback via their tips for people with visual impairments and blindness to know about near-space environments. Although the visually impaired and blind people can generate mental images from auditory, tactile, and olfactory input (Keating, 2008), it is difficult to use input from these sources to get to know far-space environments. Moreover, the inventory of existing technological solutions (see Section 3.3) shows that camera-based assistive tools have more advantages and fewer drawbacks.

Thus, the focus of searching for feasible technologies will be placed on the second type and vision-based technologies.

The methods that are used to look for feasible technologies include search in academic libraries, such as Uppsala University library and IEEE, and on search engine, such as Google with keywords “assistive technology”, “self-driving technologies for assistive tools”, “computer vision for the blind”, “vision-based object detection”, “vision-based navigation”, “orientation for the visually impaired”, and as well as variants of these keywords.

4.3 DESIGN PROPOSAL

The design was constructed using a social model of disability that is intended to remove social and physical barriers by facilitating the interaction between a person’s health and contextual factors, which is in contrast to the medical model of disability (Cameron, Swain, & French, 2003). The medical model sees impairments from within individuals and as the cause of disadvantage that results in occupational therapies and rehabilitations (WHO, 1980). For the social model, Keating et al. think that the impairments are caused by the dysfunctional interaction between people and contextual factors (Keating, 2008). Based on the social and physical barriers that were discovered in user studies of independent mobility of visually impaired and blind people, the design was proposed to facilitate the aforementioned interaction with feasible technologies in order to assist people with visual impairments and blindness in independent travel.

The design was proposed using Adobe Illustrator to visualize the design concept.

4.4 DESIGN EVALUATION

A low-fidelity (lo-fi) prototype was made and evaluated to verify that the design concept meets users’ needs when they do mobility activities independently, as lo-fi prototype translates high-level design concepts into tangible and testable artefacts with focus on functionality.

The participants who signed up for the future test in the phase of studies were contacted to evaluate the design. Due to financial and time constraints, the number of participants was as small as four, which is similar to the phase of user studies. Moreover, the testing venue was chosen by the participants for their convenience.

4.4.1 TEST PURPOSE

The test is intended to find out users’ acceptance and feedback on the form, the content, and the level of detail of the output that is given by the “smart” white cane system.

4.4.2 TEST PLAN

In each test, a participant walked through a route towards a pre-selected destination in an outdoor environment with the long white cane, since the design is targeted at an outdoor environment. For a lo-fi prototype evaluation, testers can act as a computer system and change the interaction as designed to simulate the interaction between users and the product. Therefore, in the design evaluation, I acted as the proposed aiding system to interact with participants with use of mock-up hardware and voice instructions prepared and recorded by a male voice in their native language. The voice instructions offered some or all of the following points of interest:

- Where I am
- Where I am exactly at this moment
- What the area looks like
- How far it is from the destination in meters
- How far it is from the destination in minutes
- How far is left from here to the destination in meters
- How far is left from here to the destination in minutes
- Which direction I will take
- What I have passed
- What I will pass
- Anything/anyone in my way?
- How should I react?

Before each test, participants were given a brief introduction to the design proposal.

During each test, a GoPro motion camera was mounted on the chest of each participant to record the movements in the front of participants when they move with the long white canes. In addition, photos and videos were taken around them in order to capture overall movements. Throughout the whole testing, a voice recorder was mounted near their neck to record their thoughts when they did think-out-loud during the test, as well as to record the interviews and conversations before, during, and after the tests.

After each test, a post-interview was conducted using NASA Task Load Index (NASA-TLX) questionnaire. The NASA-TLX is a common workload assessment tool to provide subjective and multidimensional assessment on effectiveness of a system or a task or other aspects of performance ("NASA-TLX," n.d.), as follows:

- How mentally demanding was the task?
- How physically demanding was the task?
- How hurried or rushed was the pace of the task?
- How successful were you in accomplishing what you were asked to do?
- How hard did you have to work to accomplish your level of performance?
- How insecure, discouraged, irritated, stressed, and annoyed were you?

In addition, I asked the following questions that are modified based on standard post-test questionnaires ("Computer System Usability Questionnaire," n.d.; Finstad, 2010):

- How helpful was the information/instruction provided?
- How much would you like to use this system?
- Do you feel confident in independent travel with the system?
- Do you feel it would be easy to learn to use the system?
- What are your suggestions?

4.4.3 TEST PARTICIPANTS

Four individuals who signed up for future test phase participated in the design evaluation. All of them are male. Three of the participants finished their respective test in the same environment based on their own selection of testing venue. They can be considered as "expert users" who are familiar with assistive technologies, due to their great exposure to technology products and frequent use of technological assistive tools at work and in everyday life. Table 3 gives more details about their background.

Table 3 Participants in the Design Evaluation

Participant	Age	Background
#1	Above 65	The participant #1 is a male with an age of over 65. He has been visually impaired for 10 years and has used the long white cane since then. He is not completely blind as his left eye has very weak sight and can sense some light. For example, he can see buildings very vaguely. He is a Ph.D. in political science. He uses a landline phone at home while he uses a mobile phone for calling taxi services in outdoor environments. His phone is not a smart phone so he has not used applications like BlindSquare for navigation, although he had an iPhone before but tossed it because he found it useless.
#2	55-64	The second participant works as an administrative officer at SRF with focus on disability policies and communications with politicians. He has been visually impaired for 17 years and has used the long white cane since then. He uses an iPhone and other assistive tools as well as mobile applications such as BlindSquare and Seeing AI.
#3	35-44	The third participant works as an administrative officer at SRF with focus on investigations, surveys, statistics, municipal services - companion, vision and hearing instructors, eye care, general disability policy. He was born with low vision but it was degraded to zero. He uses a long white cane because he thinks it gives him good predictions and makes him feel safe. He uses an iPhone and a smart watch, as well as other assistive tools, such as VarioUltra Braille display, and mobile applications, such as BlindSquare.
#4	35-44	The fourth participant works as an administrative officer with focus on rehabilitation, Braille, Adult, IT Aid, New Technology, Culture – Synthesis. He was born blind. He uses a long white cane. He uses an iPhone, an Apple Watch, and other assistive tools, such as VarioUltra Braille display, and mobile applications, such as BlindSquare.

5 Results

5.1 User Studies

Four face-to-face semi-structured interviews and one online survey together with one field observation were conducted. In this section, the results of the interview and online survey are reported, which is followed by the result of the field observation. At the end, the results are analysed.

5.1.1 Interview and Survey Results

In the end of collecting qualitative data that spanned for a month, there were a total of 11 responses from the interviews and online survey, among which 2 of 11 were from female participants and 9 of 11 were from male participants as shown in Figure 2. As shown in Figure 3, Figure 4, and Figure 5 respectively, the interviewee group consisted of different ages groups, durations of being visually impaired, and covered all three blindness types. Figure 5 shows that participants live in Stockholm, Uppsala, and Strömstad, although 3 respondents did not specify their residence locations.

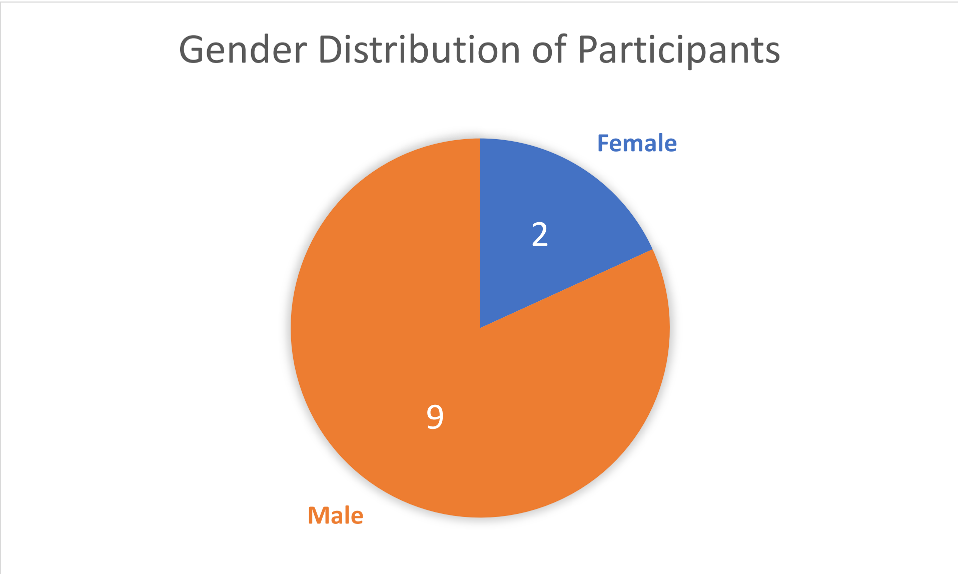

Figure 2 Gender Distribution of Participants

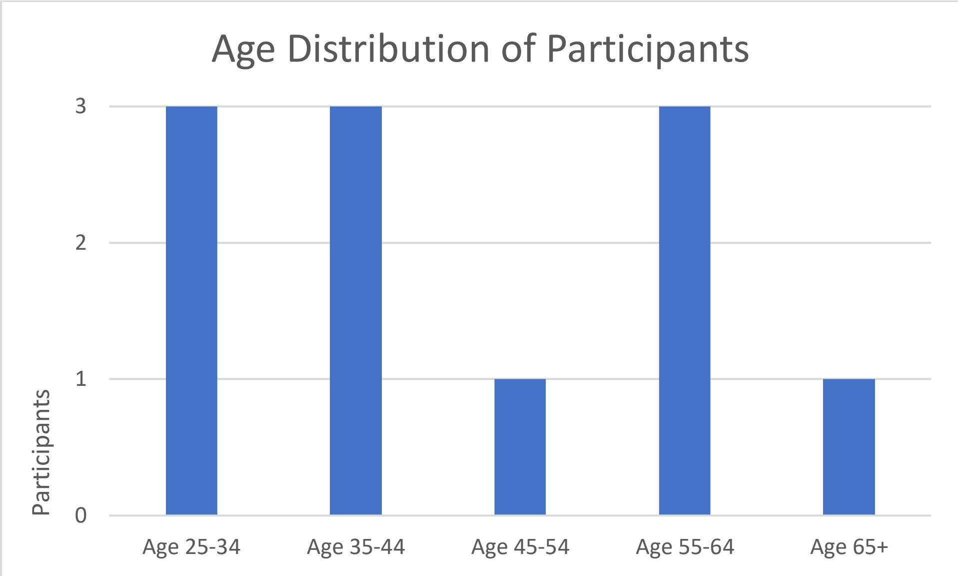

Figure 3 Age Distribution of Participants

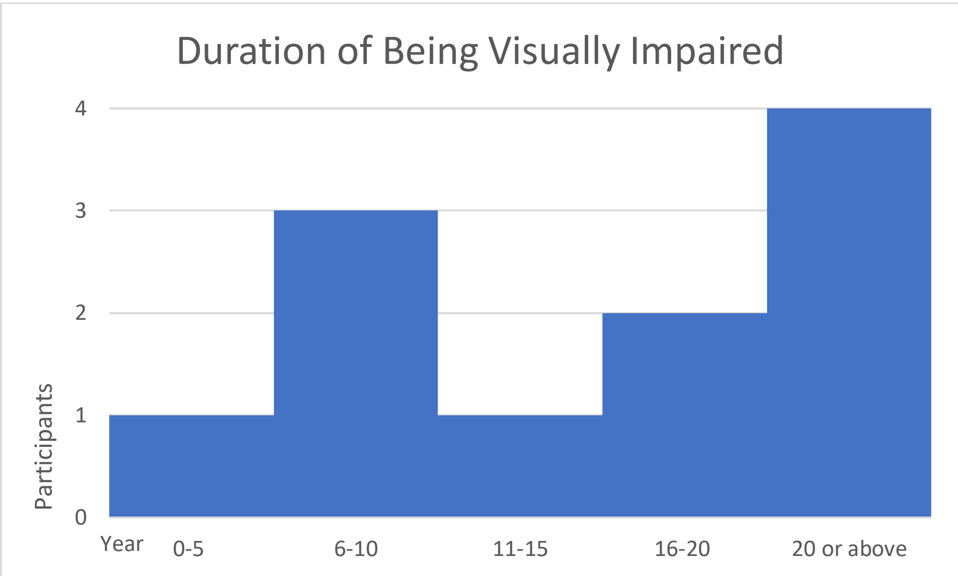

Figure 4 Duration of Being Visually Impaired

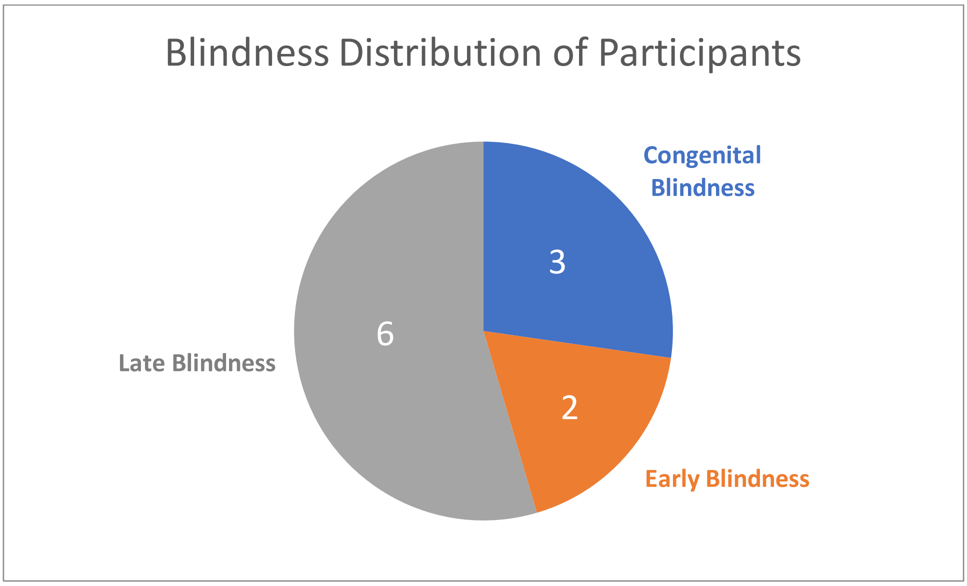

Figure 5 Blindness Type Distribution

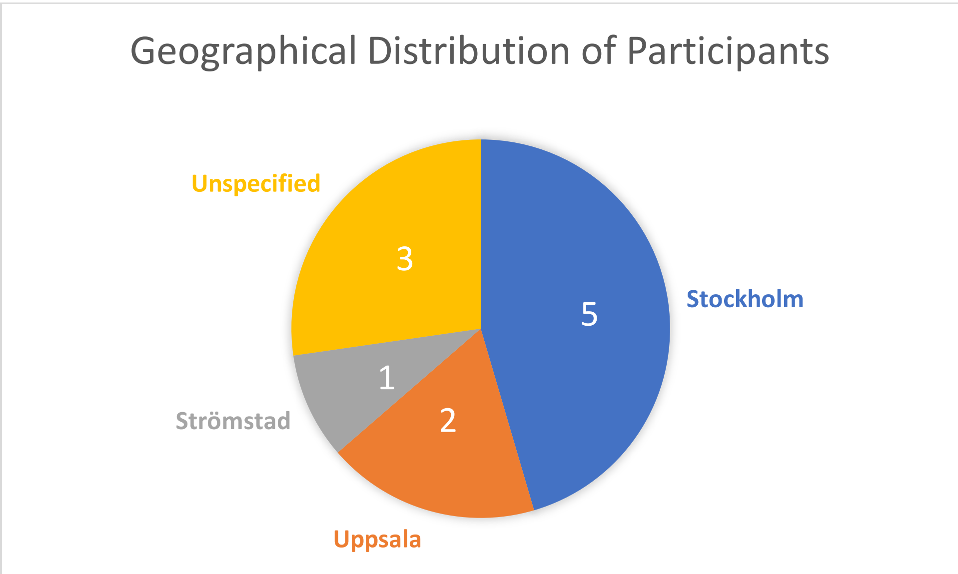

Figure 6 Geographical Distribution of Participants

Their answers to interview and survey questions no. 6-11 are presented as follows:

What support tool(s) do you use? For example, a white cane.

All the 11 participants use long white canes as their regular travel aid besides other support tools such as guide dogs and electronic devices such as Track Breath.

For what do you use the support tool(s) in your everyday life? [For example, navigation, object detection, and orientation.]

9 of 11 specifically stated that they use the support tools for navigation, orientation, and object detection regardless of their blindness type, while one of the rest 2 of 11 mentioned using a guide dog together with a GPS service working on voiceover in the mobile phone is a very good combination and the other one participant answered to run with the cane. 2 of 11

interviewees mentioned holding the cane makes them feel comfortable and secure. In addition, using the white cane as ID of being blind was mentioned by 2 of 11 people.

How do you think the ability of the support tool(s) for detecting objects? [bad (0)-fine (5)-good (10)]

Although white canes are used by all of the participants, the average score of object detectability of a white cane is 4 with a scale from 0-10 where 0 indicates bad, 5 means fine, and 10 indicates good. 3 of 11 participants mentioned they have been injured by objects at least once because the tools are unable to detect objects at head-level and knee-level or above.

How do you get your orientation?

For orientation, 10 of 11 reply on echoes and white canes to detect objects, while 1 of 11, a female late blind, asks for company of a sighted person for guidance if she goes to new or unfamiliar places. Two individuals shared that concentration is necessary for maintaining orientation and thus quite tiresome.

If you lose your orientation, how do you regain it?

If they happen to lose their orientation, 7 of 11 replied that they will ask people for help besides listening to surrounding sound, trying to visualize where they are by detecting surrounding objects, such as street curbs, and starting mobile navigation apps like the other 4 of 11 participants.

What do you need to know in order to find your way to a certain place?

In order to find their way to a certain place, they need to know the information as I summarize as follows:

- Current location
- Distance to destinations
- Directions
- Area layout
- Location of obstacles
- Points of interest within travel, such as turning point, stop name, and how far left away from the destination
- Actions needed during travel, such as making turns

One participant (1/11) said someone has to show him for the first time if he goes to a place that he has never visits before. Also, another participant said it would be nice to be told how to get to the destination by some device. In addition to guidance, one participant (1/11) pointed out he usually asks people when he is a few meters away from his destination or travel indoor, due to low GPS precision and limited use in indoor environment. Although somewhat unrelated to the original question, 1 participant mentioned that traffic lights are problematic when travelling. In Uppsala, traffic light sounds were changed due to an EU regulation. However, now the sounds are hard to recognize. Unfortunately, he does not remember what the EU regulation is and thus it is unclear why the sounds were changed.

What kind of phone do you have?

In terms of mobile phones, 10 of 11 participants use iPhones, although they are of different models. The other one participant, a male late blind, mainly uses a landline phone and uses a mobile phone for calling for taxi service. According to the participants, the reason why they use iPhones is because the screen reader on iOS is very satisfying. In fact, 3 interviewees told me that Apple products have very good accessibility and it is not surprising all 3 of them have been using Apple Watch.

How do you use your phone in daily activities?

Only 5 of 11 participants answered this question. Two of them were interviewed together in one session and they share a similar way of phone use in daily activities. They use smart phone every day. For example, they use it for navigation. They also connect their phones with Vario Ultra via Bluetooth. They usually use screen readers to activate apps on their phones.

Another participant uses his smart phone for everything from dawn to dusk. Other participants also use it for all sorts of activities including mobile banking, emails, social media, SMS, phone calls as well as Apps like BlindSquare and GPS on the phone.

What do you think about the map on your phone?

In contrast to using Blind Square for navigation, Google Maps and Apple Maps are not popular due to poor accessibility.

What technological services for visual impairment do you know? [For example, Blindsquare, Be My Eyes.]

Regarding technological services, 9 of 11 participants find BlindSquare to be satisfactory on mobile phones for navigation while one participant responded that he would install and use it soon as it has good reviews. The only individual who is not aware of any technological services mainly uses a landline phone except using his mobile for calling taxis. Several other apps are also used or heard of by the participants: Be My Eyes, Wayfindr, Seeing AI, Color detector, and Tap Tap See, KNFB reader, and Qwenda. For example, Busclip-vibration works with ultrasound and can detect objects in the front and at the head-level. One person (1/11) mentioned he does not tend to use Be My Eyes, because he does not know or trust the person who is going to help him.

What is your personal experience of the smart devices, if you have used them? And what's your opinion about them?

When asked about their opinions on technological tools, 9 of 11 participants showed a welcoming attitude in general with different expectations on functions. On the other hand, one participant (1/11) responded that he did not see points in using them. In addition, another person was critical about the tools and thought it is not beneficial to use them. In his clarification, he first pointed out that there were so many tools to carry everyday nowadays and then explained that the tools always require hearing signals and wearing glasses, which may lead to blind individuals being less aware of their surroundings/environment, which is extremely important for their independent mobility. Furthermore, he pointed out there may be blind people who do not use smart phones according to an investigation done by SRF.

If you were to be asked to design a white cane or another support tool, what features would you design?

As for what features they would like to see on smart tools if they use them, they think the features are helpful as I summarize as follows:

- Lightweight and easy to carry when travelling
- Able to provide preview of front view
- Able to detect objects and help navigate and orientate
- Able to give instructions and warnings, such as how many steps and which floor it is

Interestingly, 2 of 11 respondents mentioned robot/smart guide dogs, one of whom clarified that he has used the white cane for more than 20 years and does not see what can be done with it to make it better.

5.1.2 Field Observation

In the only field observation that I conducted, the observed participant is a male around age 40. He was born with complete vision loss. He works with assistive technology and digitalization. Therefore, it is not surprising that he is a tech savvy and uses a few technological devices to facilitate his travel.

Figure 7 The Participant Used Handles Instead of the White Cane

I followed the participant when traveling from his office in south of Stockholm via the metro to Odenplan Station and then changing to the commuter train to Uppsala station. Since this is his daily routine route, all the environment along the way is familiar to him. The trip was comprised of three parts. In the first part from his office to the metro station, the participant started with the long white cane using swiping techniques. A few minutes later when he was on to the path sided with street fences on his right, he put his right hand on the fences and used them as a guidance instead of the cane. As he approached the station, he switched back to using the cane and entered the station. When he was climbing up the stairs to the platform, similarly he used handles instead of the cane to guide him as shown in Figure 7. After he passed the entrance door to the platform, he walked to the right side of the platform and walked straight towards the end immediately. There was no problem finding a door when the metro came, thanks to the sound. However, he did make a big step when entering the train to avoid the gap between the platform and the train. In the second part from Odenplan metro station to Odenplan commuter train station, the switch from one station to another was problematic and troublesome. Since it was during rush hours of commuting, the platform was crowded, and people were standing everywhere including tactile paths. Due to the cane length, the long cane became hard to use in the crowds, as it was tapped at people's feet quite often. The environment noise was too much for him to detect the sound of the escalator that leads to the commuter train station. After a while, he found a way to the escalator using the long white cane and the sound of a running escalator. He attempted to tap the cane tip against the bottom of the wall enclosing the escalator in order to find the entrance to the escalator. However, as it was very noisy, he did not notice the escalator he found at the first attempt was actually going upwards until he detected it with the cane. According to his experience, he turned around and walked towards an opposite direction to look for the downwards-direction escalator. By using the same approach, he found the right escalator. The problem continued. Due to the design of the station, the escalator was only going to the middle floor between the commuter train platform floor and the metro platform floor. After getting of this escalator, he still needed to concentrate and utilize the long white cane and surrounding sounds to find the next escalator. When he reached the commuter train platform, he checked time on his Apple Watch by listening to the watch to respond. He remembered the schedule of the commuting train to Uppsala, thanks to daily commuting, so he knew how long he would wait for the train to arrive. However, he still used the app Tåg on his iPhone to check if there were delays. He continued to walk until the entrance areas specifically designed for people with additional needs, as shown in Figure 8. When the coming train stopped, and the door opened, he got on the train without waiting for other passengers to get off first. Because he could not see that people were getting off.

Figure 8 Entrance Area for People with Needs

In the third part from Uppsala platform to the station taxi waiting area, the participant and I took some time to adjust camera equipment onboard the train after its arrival in Uppsala. As a result, there were only a few people on the platform when we got off the train. Even though he was familiar with the environment, he still went off the desired route for a moment and almost hit a trash bin. However, thanks to environmental sounds such as automatic door opening and closing, he found his way back to the route. The rest of the walk was very smooth and at a fast pace, except the cane hit feet of some passengers who were sitting on the bench on the first floor in the station

5.1.3 ANALYSIS

The findings in user interviews and field observation are consistent with the viewpoints about vision and perception of visually impaired and blind people found in the literature review (see Section 3.1). As shown in the result analysis, long white canes are the primary support tools for people with visual impairments. Together with awareness of the surrounding environment, such as echoes and sounds in the context, they can navigate and orientate at the same time of avoiding obstacles. However, undeniably, lost vision requires blind people to expend more effort than sighted controls while navigating, as the blind people need to concentrate in order to keep up finding cues from the surroundings for constructing a visualization of the environment, which is reportedly tiresome. Furthermore, due to object detection limits of their current tools, it is easy to get injured by objects at head-level and knee-level or above. All these obstacles prevent blind people and visually impaired from moving freely, independently and safely as much as needed. The insufficient object detection also indirectly prevents the visually impaired from following social conventions in some context. In the field observation, the participant did not wait to get on the commuter train until other people got off, as is the social convention in Stockholm. Because he was not able to gain a preview of objects including people in this context and thus did not know what to react. The passengers who were getting off looked a bit offended, possibly by the breach in social convention.

Table 4 lists problems to solve for independent travel based on user responses and requirements.

Table 4 Problems to Solve for Independent Travel

Phase	Problem
Travel planning	• Where is the destination; • What the destination looks like; • How to get there; • Points of interest during travel.
Travelling	• Where I am; • Where I am exactly at this moment (for example, which floor?); • How far is it from the destination; • Which direction will I take; • What does the area look like; • What I have passed; • What will I pass; • How many steps left; • Anything/anyone in my way? How should I react?

In the answers from all participants, it is not observed that genders cause differences in blind individuals' needs and problems during their daily travel. However, as the number of participants is small, it is also difficult to conclude that gender difference does not have impact on how the visually impaired perceive their needs and problems. As there are rarely papers that discuss the difference of blind male and female individuals in needs and problems perception in the context of daily independent mobility, I will proceed with this paper based on what is revealed in the interview and survey responses. Similarly, the user responses do not indicate that congenitally blind, early blind, and late blind people have different needs and problems during their daily tasks related to mobility, which is also suggested by Lewis et al. (Lewis, Sharples, Chandler, & Worsfold, 2015). Although people of different blindness types have differential performance on sensory perception in varied mobility tasks, due to different onset time, its duration of visual deficit, personal expertise, and motivation that can affect perception of people with visual impairments and vision, they still use the same sensory modalities to acquire information for generating internal images of the external, which implies that their needs on the support tools are not distinct. Thus, I will not take gender difference and blindness types into consideration in the following discussions.

5.2 FEASIBLE TECHNOLOGIES

Feasible technologies were found by search using keywords listed in Section 4.2. Before presenting the findings, I will discuss the contexts in which the technologies should work robustly and effectively. The findings of feasible technologies are presented with respects to three important components of independent mobility of the visually impaired and the blind: navigation, object detection, and orientation.

5.2.1 CONTEXTS

As stated in the introduction, this paper is only focused on independent mobility of visually impaired and blind people in an outdoor environment. By ‘independent mobility’, it is meant that blind and visually impaired people can walk, take public transportations, work out, go to cafés and do other recreational activities on their own. The contexts of their experience can be streets, subway stations, parks, forests, parks, cafés, and other outdoor locations, which are subject to all sorts of weather and light conditions including rains, strong sunlight, and low light as well as conditions in which there are no batteries or available mobile networks. Furthermore, long white canes need a certain radius to be used efficiently due to their length, no matter if two-point touch technique or constant contact technique is used. Hence, technological solutions must be robust to support people with visual impairment in carrying out activities irrespective of the contexts, even when the environment does not allow efficient use of long white canes.

5.2.2 TECHNOLOGIES

Technologies used in existing assistive aids have not been able to cover all three aspects of mobile travel: navigation, orientation, and object detection. Their focus is usually limited on one of two aspects. Also, there are technological assistive aids that require being used in conjunction with deployments in the environment, which has very high expense as Keating, et al. (Keating, 2008) have pointed out in their discussion about establishing location identifiers in the environment that can transmit the information in speech to users when the sensors on the users detect the identifiers remotely via infrared, Bluetooth beacons, and so on.

As the long white cane has very good performance on near range and under knee level object detection, I will keep the long white cane as a primary tool and discuss feasible technologies for secondary tool that will be coupled together with the long white cane. The focus of the secondary tool is placed on far range and above knee level object detection, navigation, and orientation in all the needed contexts.

Orientation, Far-space and above knee level object detection

A glance with the eyes encompasses a tremendous amount of information about the environment and objects within the environment at near space and far space (Cattaneo & Vecchi, 2011; Schinazi et al., 2016). This glance gives an individual a preview of the surroundings that is essential for safety and efficiency of mobility and facilitates people to gain context awareness of the environment (Armstrong, 1975; D. S. Kim et al., 2017). Thus, it is necessary for visually impaired and blind people to have knowledge of the environment itself that consists of different objects as well as their positions in the environment in order to make decisions and proceed in their independent travel (Roentgen et al., 2008).

For people to become context aware, multi sensing technologies can be used. These everyday needs of visually impaired and blind people are shared by autonomous vehicles that also must understand what is in the surroundings and the contexts in order to react accordingly through so called computer vision (Martinez & Koester, 2017; Pendleton et al., 2017).

Similarly to pedestrians with visual impairment and blindness, it is challenging and crucial for self-driving vehicles to deal with unexpected situations to ensure travel safety. For a safe driving experience, a self-driving vehicle must be able to perceive the environment, plan the moves, and control its actions (Pendleton et al., 2017), which is similar to the essential components of a safe walking experience for people with visual impairments (Cattaneo & Vecchi, 2011; Keating, 2008).

However, the technologies for these activities are not developed well for people. Although, existing technological solutions often make use of ultrasound, near and short wavelength infrared, and laser light for detecting distant objects, they do not work in a few particular scenarios: near and short wavelength infrared signals require an external infrared illuminator to emit infrared signals and reflect on the objects, during which interference by sunlight and dark objects may occur; laser light has a narrow detection range that will reply on the users to tap the long white cane side by side in order to input in a wider range; ultrasonic beams can be reflected by boundaries and thus cause interference on detecting objects.

Thanks to booming interest in autonomous vehicles, the perception of driverless vehicles is continuously improved for operations in all weather conditions including raining and foggy days. New technologies have emerged and been developed to solve problems that are related to computer vision (Katz, Truillet, Thorpe, Jouffrais, & Jouffrais, 2010; Martinez & Koester, 2017; Pendleton et al., 2017). Therefore, I can look into those technologies to get inspiration.

As discussed previously, camera-based technologies have been applied for developing assistive aids for the visually impaired and the blind. Similarly, camera-based systems are used by driverless vehicles to realize their vision. Usually there are cameras mounted on the vehicle front and images are processed in real time. During image processing, the vehicle systems that are trained with deep learning methods will identify the environment and objects, read the situation, and react accordingly, which requires 3D perception and semantic understanding of the environment (Schneider et al., 2016).

Figure 9 Stixel World representing the three-dimensional environment in front of the vehicle. Source: Stixel-world

In an experiment conducted by Martinez, et al. (Martinez & Koester, 2017), the blindfolded participants wore the camera in their chest, and the images were processed using a fast implementation of the stixel segmentation approach. Stixels refer to rectangular columns named super-pixels ("Stixel World," n.d.). Each of the rectangular columns, that is, each stixel stands on the ground vertically with a certain height. It is defined by its 3D position that is relative to the cameras. As shown in Figure 9, they are used to segment images frame by frame in order to track 3D points and to determine positions in traffic scene with relevant information intact. A segmentation of the input image is extracted from several objects per image column, which supports an optimal stixel grouping. Consequently, a 3D traffic scene is perceived. Figure 9 demonstrates the representation of a traffic scene where the red color indicates near space and green color indicates far space.

Given that stixels limit free space and approximate object boundaries, the advantage of using the stixel representation is in reducing computation demand for image processing significantly by only analyzing approximately a few hundred stixels instead of hundreds of thousands 3D points for every image (Cordts et al., 2017; Erbs, Schwarz, & Franke, 2013; Pfeiffer, Erbs, & Franke, 2012; Pfeiffer & Franke, 2010). As a result, all dynamic objects are detected easily. (Cordts et al., 2017; Enzweiler, Hummel, Pfeiffer, & Franke, 2012; Erbs et al., 2013; Pfeiffer & Franke, 2010; "Stixel World," n.d.). Among all the objects, it is important to recognize those objects that are semantically significant to decision making for people in motion as shown in the bottom image of Figure 10, which usually requires employment of convolutional neural networks for deep learning (Martinez & Koester, 2017).

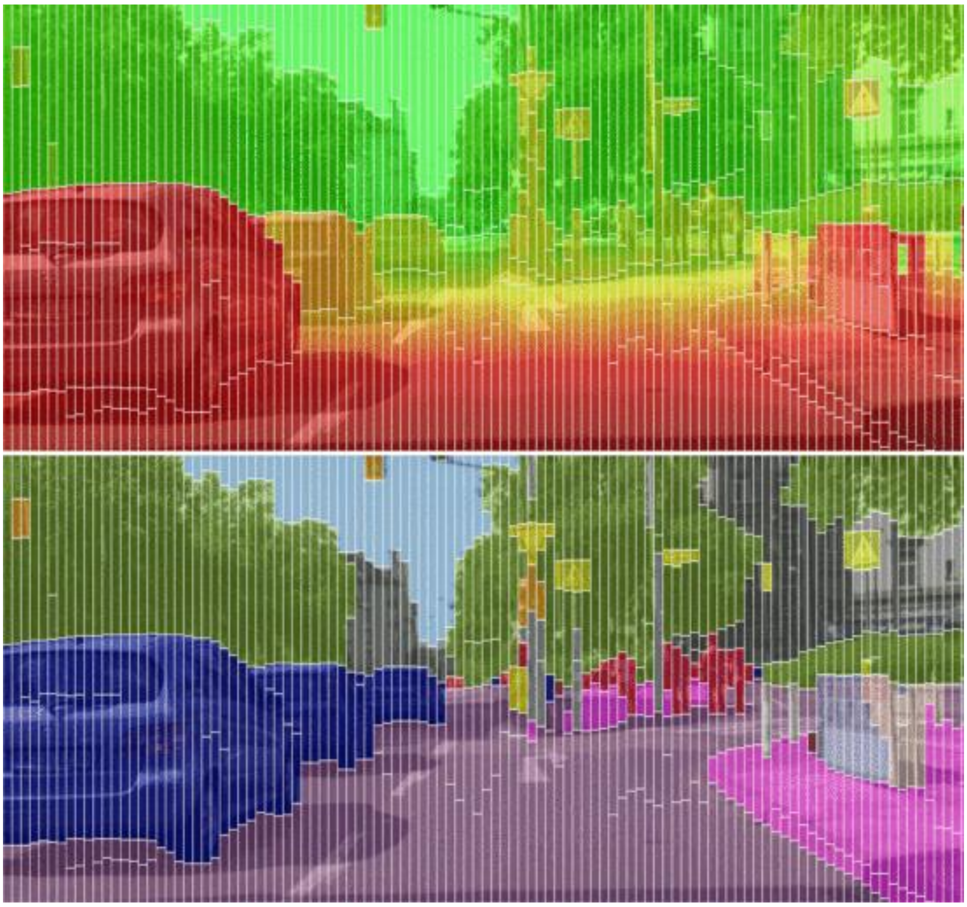

Figure 10 Top: stixels obtained from a driving scene. Bottom: Their semantic segmentation. Source: Stixel-world.

Although Martinez et al. suggested that usually the best performing neural networks are slow for mobile applications (Martinez & Koester, 2017), there are several possibilities for leveraging between the computation efficiency and detection accuracy in conjunction with use of mobile specific neural networks such as MobileNets (Howard et al., 2017). At first, for visually impaired and blind people to decide where to go, free space, other pedestrians, and obstacles including dangerous objects should be identified. However, identifying details of all the objects that must be recognized is not necessary. For example, what pedestrians look like, such as faces and body shapes, is not needed for identifying them as humans. Therefore, the image processing for 3D perception of the environment may not need visual input of high resolutions, which would consequently reduce the demand for high resolution cameras and computation power. Nevertheless, this does not imply that low resolutions of visual input can facilitate image processing to meet the quality requirement of image recognition. Instead, it is necessary to discover the leverage between the resolution level and quality of image recognition. The latter is determined by latency tolerance in real-time detection and classification accuracy of objects. Since human brains averagely need 0.25 to 0.5 seconds to process and react on visual input that is taken in from the eyes, object detection that is done by the computer vision should not take longer time than the average response time of human brains so as not to restrict the pace of people's motional activities. As mentioned previously, the ability of classifying and identifying free space where the visually impaired and blind can walk, humans, obstacles such as trees and traffic light status will provide sufficient accuracy for required quality of image recognition.

Complementary support for orientation and object detection

Visible cameras refer to When it comes to vision, human eyes and visible cameras are restricted by the same limitations in the conditions of bright light like oncoming headlight and sunlight, glare, and night time (Saito, Hagihara, Hatanaka, & Sawai, 2008). In these limited conditions, far infrared sensors are used quite often in addition to cameras for object detection in autonomous vehicles (Händel, Yao, Unkuri, & Skog, 2008; Saito et al., 2008; H. Wang, Cai, Chen, & Chen, 2016).

Far infrared sensors have been long adopted by night vision technologies. For example, night vision goggles. This is because objects including humans in the environment emit heat that can be captured by far infrared sensors to fill in the detection void left by the cameras, which also implies that there is no need for an additional illumination source. Far infrared sensors detect objects by capturing their portions of infrared light spectrum that are emitted as heat from objects (Sapkota, 2013). This passive infrared approach does not meet the same drawbacks as near infrared approach with the need for a source to actively emit infrared signals. Hence, there have been a few commercial products and studies on using far infrared cameras for driverless vehicles in low light and night scenarios.

Nevertheless, the employment of a far infrared camera system poses a question: how does image processing handle sources from far infrared camera and optical cameras? One of the answers is provided by so called image fusions for multi-sensory systems (Fendri, Boukhriss, & Hammami, 2017; J. H. Lee et al., 2015). Image fusion is defined by Wikipedia as "In computer vision, Multisensor Image fusion is the process of combining relevant information from two or more images into a single image. The resulting image will be more informative than any of the input images". Among existing studies on image fusions for far infrared and visible images, J. Lee et al (J. H. Lee et al., 2015) proposed an image processing system in which axes of far infrared and visible light cameras are placed in parallel horizontally and the same people from visible images are mapped onto the people that are obtained from far infrared images with background extraction in the far infrared images. Although Lee et al claims that dual camera systems have a lower image processing speed than single camera systems, as the former needs to process images from two sources, the total image processing of detecting people takes 27.04 milliseconds, which falls into the range of acceptable latency in real time scenarios according to the discussion in section 0. However, whether the total image processing time is acceptable must be experimented together with a suitable image recognition method, which is out of scope of this thesis.

Another approach to dealing with images from far infrared and visible light is to keep the processing of images from these two sources separately and independently, which means far infrared system can be used as an add-on on top of the selected visible camera solution and its associated image processing and recognition. Since visible cameras are good at capturing rich visual information about the environment, not to mention that digital image processing is capable of handling images captured in low light conditions to a certain extent when charge coupled device (CCD) image sensor-adopted cameras are used, visible cameras can be used as main source of images and turned on by default. The switch between which source of images to process and use as input for image recognition depends on pre-defined environment settings. When the pre-defined environment settings are met, the image processing of far infrared can be turned on while the one of visible images is turned off. Consequently, the image recognition for images from far infrared must be developed besides the one for visible images, which will be explored in the future work instead of this thesis.

Navigation

As discussed in section 3.3.4, GPS has still not been replaced by a single solution for the purpose of navigation, despite the fact that it does not position a spot within 10 meters (Coughlan & Miele, 2017; Li, Wang, & Li, 2013). The navigation mobile app BlindSquare, used by most of participants in the user studies, is based on GPS. BlindSquare provides an application programming interface (API) that can be integrated with other services. The integration can be made by using an application directed URL scheme. With the URL scheme, BlindSquare by be launched and the navigation information can be passed to the integrating service. Similarly, Google Maps and Apple Maps can also be integrated into other services through their own API. The common problem is the positioning precision.

Existing studies show that it is possible to utilize existing GPS systems for approximate positioning but to rely on computer vision for a more granulate positioning (Brilhault, Kammoun, Gutierrez, Truillet, & Jouffrais, 2011; Y. H. Lee & Medioni, 2016; Pinho, Carvalho, & Carreira, 2015; Rituerto, Fusco, & Coughlan, 2016; Vishal, Jawahar, & Chari, 2015; J. Wang et al., 2004). There are two main methods: 1) combining GPS with map-matching images; and 2) combining GPS, other inertial sensors, and probability maps with depth information extracted from images (Barański & Strumillo, 2012; Li et al., 2013; Y. Wang, Virrantaus, Pei, Chen, & Chen, 2012).

As computer vision can be used for orientation, far range and above knee level object detection, an approach is to reuse the same resource of vision to complement GPS positioning (Li et al., 2013). Although there is enormous database of images to be used as references for mapping with real-time images, such as Google Street View, it requires further investigation to decide whether either of the methods can be employed and which one works more efficiently with the same image resources used for object detection. Nevertheless, the method that combines GPS and map-matching images depends on a database of geo-referenced images, whose growth does not necessarily increase by time but is also dependent on financial and human resource factors. To free the database growth from such factors, machine learning method can be used so that it gains learnability of recognizing places for navigations based on received images, which have been discussed by researches that focus on vision-based navigation merely or mainly use images as input for navigating with machine learning ability (Jaderberg et al., 2016; Mirowski et al., 2016, 2018; Zhang et al., 2016; Brahmbhatt, Pai, & Singh, 2017). This machine learning driven vision-based navigation methodology imitates human's capabilities of navigating without the guidance of a map (Mirowski et al., 2018). Consequently, this allows a system to expand its ability to navigate without relying on GPS, which may provide possibilities of navigating using obtained spatial recognitions under circumstances where network connections are not available. Hence, in the future study, both combination of GPS and vision-based and merely vision-based approaches together with a supporting machine learning mechanisms should be investigated further in order to provide users robust and precise navigation.

5.3 DESIGN PROPOSAL

Keating et al. (Keating, 2008) introduced a social model of framework for designing assistive technology systems that focus more on removal of societal and physical barriers and discrimination than the medical model that is concerned with rehabilitation for physically and cognitively impaired people. This is intended to empower physically and cognitively impaired people and employ user-centered and user participatory design approaches. The social model suggested by Keating et al. proposes that the interaction between a person’s health and contextual factors is the cause of impairments, which has been recognized by WHO classification system although the classification is still mainly driven by a person’s health conditions (Keating, 2008). Therefore, the design should facilitate the aforementioned interaction in order to assist people with visual impairments and blindness in independent travel.

Figure 11 gives an overview of the design proposal that will be elaborated in-depth in the following sections. As it demonstrates, visually impaired and blind people will still use long white canes but mainly for their advantageous aspects in near-space navigation and object detection. For the far-space spatial perception and object detection, an eyewear will be used in order to capturing environmental information as extensively as possible. The interaction between the eyewear and the user is flexible and in real time. The information between users and the eyewear is transmitted via a speech API, such as Google Cloud Speech API. It is controlled by voice commands from users.

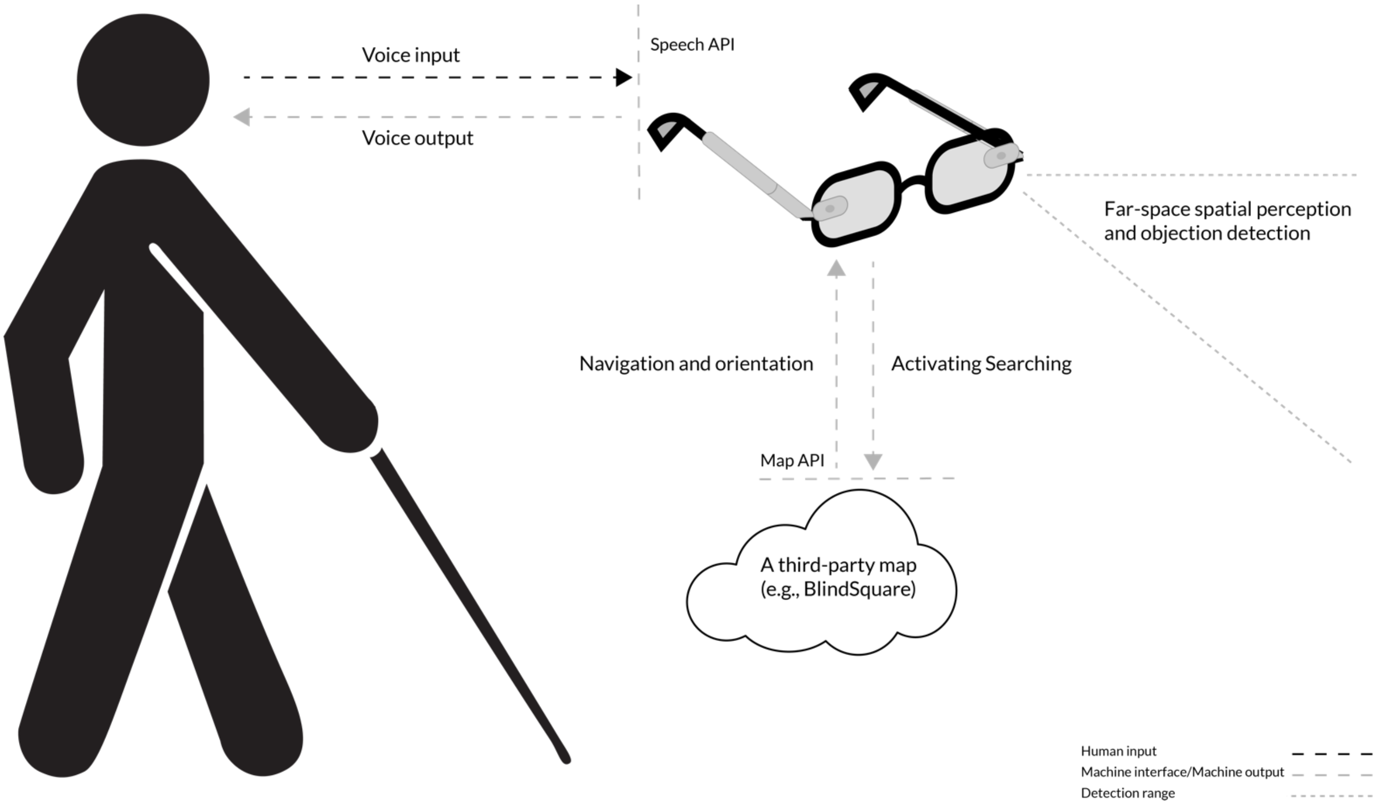

Figure 11 Proposed Aiding System for Navigation, Orientation, and Object Detection

Upon navigation requests, the eyewear will initiate a search in a third-party map via its provided API and return the information via bone conduction speaking in real time. With bone conduction, the information will be communicated via sound conducted to the inner ear through the bones of the skull ("Bone conduction," n.d.). Thus, ear canals of users will be free for catching environment noises that are important for mobility activities of visually impaired and blind people.

For spatial information, it is up to users whether they would like to receive an overview and other information about the far-space environment before approaching to that area. On the other hand, the eyewear will also learn users' behaviors as well as use histories to adapt its interactions with users in order to provide tailored services with the underlying machine learning mechanisms. For example, if a user stops the eyewear from giving certain information about a place, the eyewear will not provide such information next time except when requested by the user. However, with respect to safety of independent mobility, the eyewear will proactively give warnings of dangerous objects whenever it detects any, whether that the objects are static or dynamic.

5.3.1 Interaction of sensing the environment

During mobility, people with visual impairments and vision loss interact with the environment in order to make sense of it. Without the visual modality, the interaction between the people and the environment is supported by other modalities such as the auditory sensor as discussed in section 3.1, which perceives a smaller amount of information about the environment and takes longer time than using the visual modality. The interaction is comprised of the following processes as people move:

- Detect objects as well as navigate and orientate in the near space
- Preview objects and decide navigation and orientation in the far space

As discussed previously in section 3.2, the long white cane cannot reach over 1.65 m, due to its maximum length of 165 cm (“Whitecane,” n.d.), in the far space and it is unable to detect above knee level horizontally and vertically. To improve the interaction quality, I propose to use camera-based computer vision as the main source for real time environment perception beyond the capabilities of long white canes. It will serve as a secondary aid that complements to the use of long white cane. Therefore, as shown in Figure 12, the long white cane will be used as a near space tool for navigation and orientation and object detection under knee level in the immediate environment, while the camera-based computer vision will assist people with acquiring preview of objects and environment in the far space for proceeding with their travel.

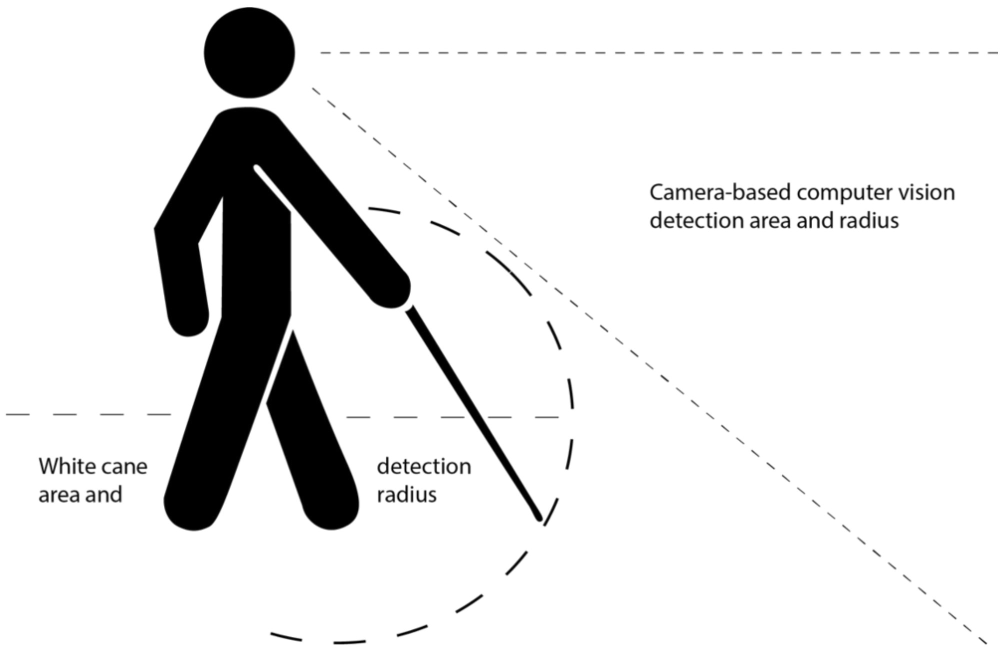

Figure 12 White cane and camera-based computer vision detection area and radius

The camera-based source should be mounted on the head for the following two reasons:

- As long white canes usually reach individuals’ lower chest, their horizontal perception range is low. A head mounted source will obtain the possibly maximum perception in the horizon plane.
- Eyes that provide visual input are located on heads.

One interviewee commented on technology aids during user studies, “We have already needed to bring a lot of devices with us every day. I don’t want to bring an extra one.” Thus, it will be helpful to integrate functions in a single device so that users will not need to use more devices than necessary.

This interviewee together with other participants are technology enthusiasts. For instance, they use Apple watches and other cutting-edge products. Hence, if the device looks cool and provides satisfyingly robust and useful integrated functions, it is possible for them to use a new device, let alone 9/11 participants who have shown interests in using technology support tools. Therefore, to meet the requirements from participants in user studies, I propose to design futuristic-looking glasses for providing camera-based computer vision as shown in Figure 13, with inspirations from the look and feel of Google Glass ("Google Glass," n.d.).

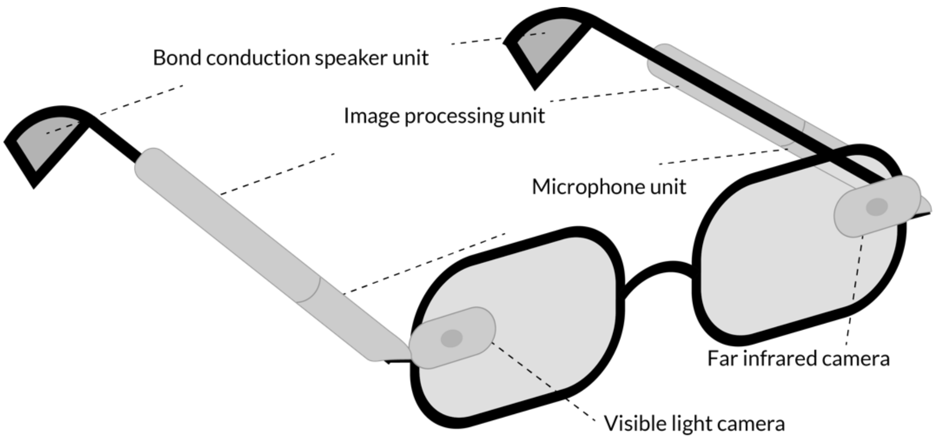

Figure 13 Prototype of Eyewear

A visible light camera and a far infrared camera are placed respectively on top of each glass of this eyewear, close to the glass leg that is also known as temple side. The visible light camera is used when the light is sufficient for capturing images, while the far infrared camera is used when there is not much light like in the night. Images captured using the cameras will be processed by the image processing unit with image processing and recognition methods discussed in Section 5.2.2 in order to give users real time information based on the recognized images.

The earpiece of each temple falls behind the earlobe for the built-in speaker to transmit sound through bones using their natural sound conducting properties. Using bone conduction will keep ear canals open to environment noise that is significant for people with visual impairments and loss during their mobile activities.

When users navigate in the environment, they can speak to the glasses if they need to know information about directions and walking routes, etc. In real time, the glasses will proactively give information on points of interests and warn them of dangerous objects in the environment, etc. Moreover, users will be assured of walking on the right route, according to findings in a user testing on Wayfindr prototype ("VALIDATING WAYFINDR," n.d.). The voice instructions provided by the eyewear will cover everything in Table 4.

5.3.2 Interaction of Navigating Around the Environment

As Brambring demonstrated in the travel model (Brambring, 1985) (See Figure 1), visually impaired and blind people identify landmarks to determine their immediate locations in the environment in order to gain orientation. In fact, it is similar to how sighted people use visual cues of contexts in terms of navigation when they do not use a map (Mirowski et al., 2018). For the visually impaired and the blind, they locate landmarks and generate spatial 3D maps by sound, although their perception of landmarks and map generation are not smooth in new and unfamiliar environment. See section 3.1.2 for detailed discussion on spatial auditory maps. The landmarks are not necessarily limited to architectures, public facilities, and organizations that are in a certain environment (Keating, 2008). Rather, they are generally points of interest that people with visual impairments and blindness perceive, generate, and record routes on which they travel. For example, a turning point and a traffic intersection. In addition, spatial and geographic orientation is another essential for people to orientate successfully and safely.

Hence, visually impaired and blind people will be provided with such information in auditory cues like speech in real time. Geographical location and navigation about architectures, public facilities, and organizations like stores and bus stops can be provided by a map service through its map API, such as Google Maps. Such information will be searched and retrieved by the eyewear upon requests of users' voice input. In return, the retrieved information will be provided to users using spoken words that are communicated via bond conduction speaker unit.

In terms of expressions, BlindSquare tells directions using clockwise instructions. However, whether this expression is sufficient, or another expression is preferred must be verified by further studies in the future.

5.4 DESIGN EVALUATION

In order to verify that the design meets users' needs and assist them with their independent travel, a lo-fi prototype was made and evaluated. A lo-fi prototype translates high-level design concepts into tangible and testable artefacts with focus on functionality (Babich, 2017).

5.4.1 TEST EXECUTION

Four participants from the pre-design user interviews took part in this evaluation. All the four tests were conducted in outdoor environment. The first test was conducted near Ekonomikum, Uppsala. Another three tests that were conducted close to Synskadades Riksförbund (The Swedish Association of the Visually Impaired), also known as SRF, at Enskede, Stockholm. It is considered that the effectiveness of the "smart" white cane system may be influenced by their high familiarity with the test environment and the efficiency of the testing may be affected by differentiated testing environment. However, given that the test purpose is to unveil their acceptance and feedback on the output, these factors are disregarded as long as the testing device is consistent.

Figure 14 Participants wearing testing devices

Pre-test introduction

Before the test, I introduced briefly the proposed “smart” white cane system. I also mentioned that this was a lo-fi prototype and thus I would work as the system while walking along. To interact with the “system”, they can speak to me as if I were the system. For example, they can ask me to stop the instructions if they already know the environment or they can ask me to give more information if the one they receive is not enough.

Participant #1 was excited to know the system can interact with him and give him instructions and other information as needed. Since he has worn a pair of glasses to protect his eyes and has experience with having glasses while walking, I did not provide him the test glasses.

As the testing with participants #2 and #3 were planned at the same time slot, the short introduction to the proposed “smart” white cane system was given to them at a joint pre-test session. They asked a few questions about how it will work in details. For participant #4, the pre-test introduction was made in a different time slot.

Test near Ekonomikum

This test happened from 4pm to 5pm on a Monday afternoon near Ekonomikum, Uppsala.

In the test, the participant walked from the main entrance of Ekonomikum to the bus stop Kyrkogårdsgatan by Kyrkogården, as shown in Figure 15.

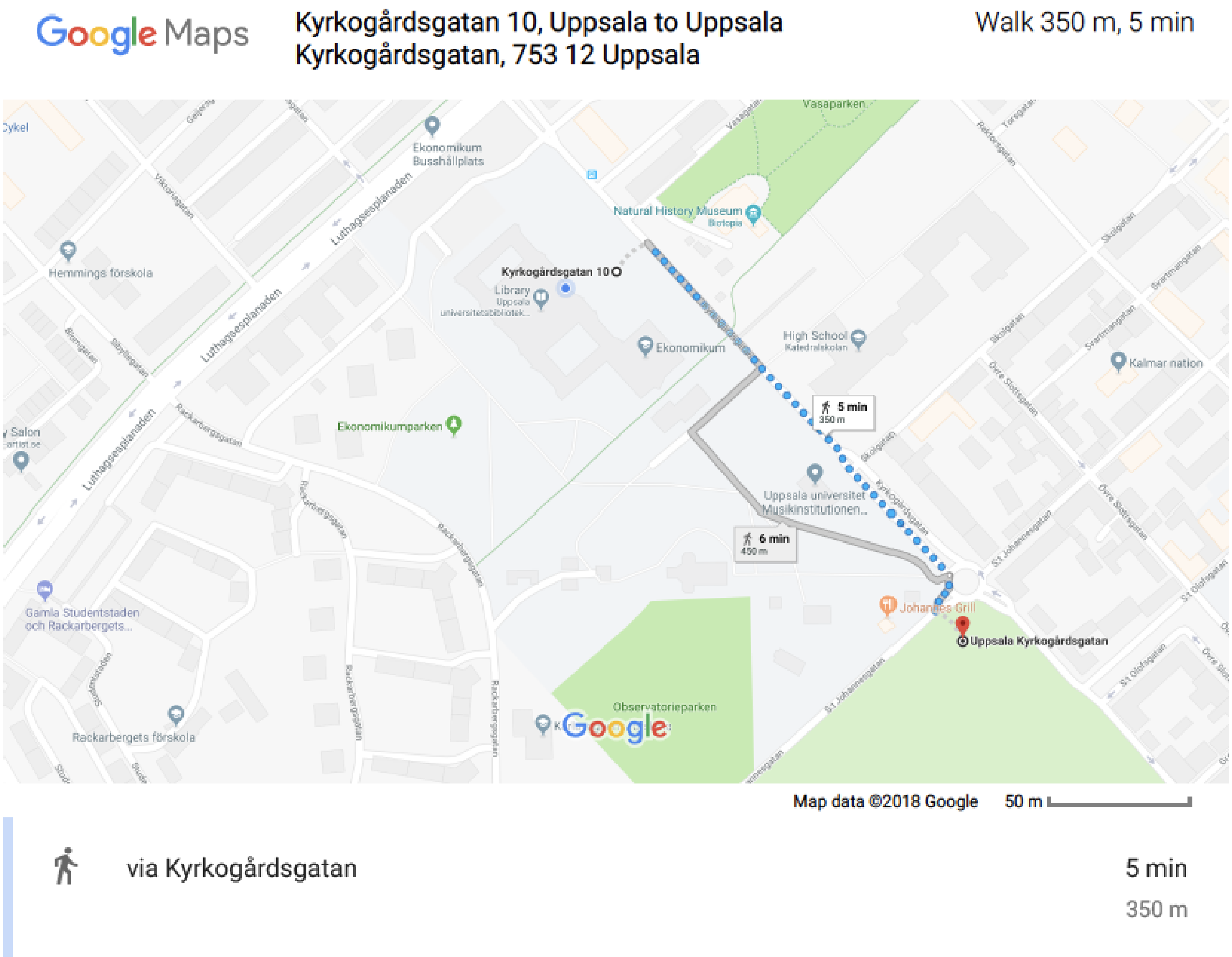

Figure 15 Test Route near Ekonomikum, Uppsala

From the beginning to the end of the test, voice instructions as shown in Table 5 were played selectively at points of interests.

Table 5 Voice Instructions Played during Test near Ekonomikum

Point of Interest	**Instruction (Swedish)**	**Instruction (English)**
Where I am	(Du är vid) Ekonomikums huvudentré.	In Ekonomikum main entrance.
Where I am exactly at this moment	Du är på bron till entŕedörren i mitten, på framsidan av Ekonomikum mot Kyrkogårdsgatan.	On the bridge to the entrance door in the front middle of Ekonomikum facing Kyrkogårdsgatan.
How does the area look like	Det finns tre entrédörrar mot Kyrkogårdsgatan. Korta broar leder mellan entrédörrarna och gångvägen längs Kyrkogårdsgatan. Marken under broarna är täckt med gräs. Det går att gå ner under bron genom att gå nedåt från vägkanten.	There are three entrance doors facing Kyrkogårdsgatan. The doors are connected by short bridges to pedestrian path along Kyrkogårdsgatan. Under the bridges, it’s all covered in grass. You can go under the bridge by walking

		downwards from pedestrian curb.
How far is from the destination in meters	Det är 350 meter härifrån till destinationen.	350 meters from here to the destination.
How far is from the destination in minutes	Det är 5–6 minuter härifrån till destinationen.	5–6 minutes from here to the destination
How far has left from here to the destination in meters	Det är 200 meter kvar till destinationen.	200 meters minutes left to the destination
How far has left from here to the destination in minutes	Det är 2–3 minuter kvar till destinationen.	2–3 minutes left to the destination.
Which direction I will take	Ta till höger Ta till vänster Vänd tillbaka Gå rakt fram	Turn right Turn left Turn back Go straight
What I have passed	(Du) har passeratpasserar Musicum, musikinstitutionen vid Uppsala universitet, musikinstitutionen på din högra sida	Passing by Uppsala universitet Musikinstitutionen Musicum on your right
What I will pass	En vägkorsning	Traffic intersection
Anything/anyone in my way?	Cyklister på vänster sida Cyklister på höger sida	People biking on your left; People biking on your right
How should I react?	Gå rakt fram; Invänta trafikljuset Trafikljuset är grönt; Trafikljuset är rött.	Go straight; Wait for the traffic light; The traffic light is Green; The traffic light is Red.

In the test, several unplanned situations occurred due to false play of some voice instructions, which is expected. In those situations, I gave the correct instructions instead, as if I were the "smart" white cane system.

As the participant was very familiar with the test environment, he spoke out loud what he may pass along the way including the points of interests that were included in the voice instructions. For example, Uppsala University Music Institute and traffic intersection. Nevertheless, he was still glad to be notified of these points of interests since he could prepare his reactions. Moreover, he thought he would not need a guide dog anymore if he could use the proposed system, before I reminded him that a dog will not need a battery to run while the system battery life is a challenging problem to solve.

An unexpected take from this test is the participant’s appreciation of the spoken male voice with which the instructions were given. He thinks it is very pleasant to hear, not to mention it is a human voice.

Test near SRF at Enskede

This test happened from 1pm to 3pm on a Monday afternoon near SRF, Enskede.

In the test, the participants walked from the entrance of SRF to the subway station Skogskyrkogården in Enskede, as shown in Figure 16. All participants were tested in their respective testing sessions.

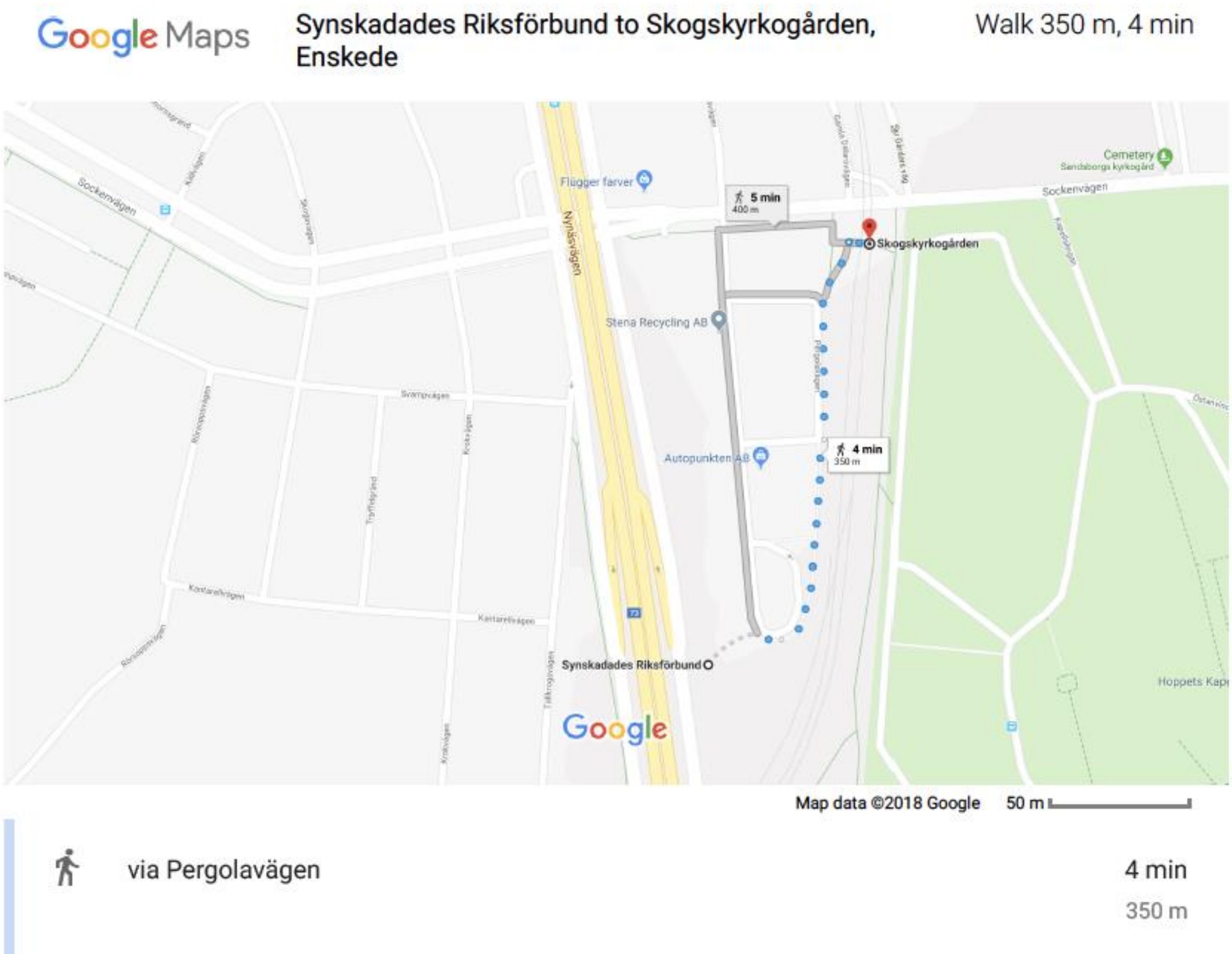

Figure 16 Test Route near SRF, Enskede

From the beginning to the end of the test, voice instructions as shown in Table 6 were played selectively at points of interests.

Table 6 Voice Instructions Played during Test near SRF

Point of Interest	Instruction (Swedish)	Instruction (English)
Where I am	Du befinner dig utanför Synskadades Riksförbund.	In front of Synskadades Riksförbund.
Where I am exactly at this moment	Du står med ryggen mot ingången. Du är vänd mot Sandsborgsvägen.	Now your back is facing the entrance door. You're facing Sandsborgsvägen.
How does the area look like	Du står med ryggen mot ingången. Till vänster finns en ingång till restaurangen Matborgen, och till höger en gångväg till tunnelbanestationen Skogskyrkogården. Framför dej finns ett räcke. Du är vänd mot Sandsborgsvägen.	Now your back is facing the entrance door, on your left is a entrance to the restaurant Matborgen, and on your right is a walking path towards subway station Skogskyrkogården. In front of you is a handrail. You're facing Sandsborgsvägen;
How far is from the destination in meters	Det är 350 meter härifrån till destinationen.	350 meters from here to the destination.
How far is from the destination in minutes	Det är 4 minuter härifrån till destinationen.	4 minutes from here to the destination
What I have passed	Du har passerat Pergolavägen på din vänstra sida.	Pergolavägen on your left
What I will pass	Begravningsbyrån FONUS, på din vänstra sida	FONUS on your left
How far has left from here to the destination in meters	Det är 200 meter kvar till destinationen.	200 meters minutes left to the destination
How far has left from here to the destination in minutes	Det är 2–3 minuter kvar till destinationen.	2–3 minutes left to the destination.
Which direction I will take	Ta till höger Ta till vänster	Turn right Turn left

	Gå rakt fram	Go straight
How does the area look like at the station	Snabbköpsbutiken Myway ligger till höger. SLs biljettkontor ligger till vänster.	Myway the convenient store is on your right. SL ticket office is on your left.
After entering the station	Det finns en trappa 2 meter framför dig; Det finns en trappa 4 steg framför dig; Det är 34 trappsteg upp till plattformen.	A staircase is 2 meters in front of you; A staircase is 4 steps in front of you. It has 34 steps till the platform.
How many steps to the platform	Det är 34 trappsteg upp till plattformen; Det är 13 trappsteg kvar upp till plattformen.	34 steps up till the platform; 13 steps left till the platform.
Entering the platform	Dörren ut till plattformen är 4 steg framför dig; Dörren ut till plattformen är 2 meter framför dig.	The entrance door is 2 steps from you; The entrance door is 2 meters in front you.
Anything/anyone in my way?	Människor till vänster Människor till höger	People on your left; People on your right
How should I react?	Biljettspärren till vänster framför dej är för inpassering. Börja gå; Stanna; Fortsätt framåt; Fortsätt gå framåt.	The ticket machine on your front left is for entrance. Start; Stop; Keep going (ahead/forward); Keep on walking/going ahead/forward.

In the test, as I were the "smart" white cane system, a few voice instructions were not played based on considerations of the high familiarity of the participants with the surroundings. For example, voice instructions that describe the environment outside the SRF office were not played. Also, several unplanned situations occurred, which is expected because the test environment is dynamic. In those situations, I also gave the correct instructions.

During the test, participant #2 thought out aloud that he was doing this test and thought it was useful to know where he had passed. Participant #3 was surprised to hear how exactly the voice instructions were given at points of interest. Participant #4 thought the instructions about steps of the staircase to the platform in subway station were incorrect.

Post-test interview

After the test, we went through the following NASA Task Load Index (NASA-TLX) questionnaire and additional questions.

Similarly to the pre-test, I conducted a joint post-test interview for participants #2 and #3.

5.4.2 TEST RESULT

Not surprisingly, the test workload that was assessed using NASA-TLX was very low for all the four participants, due to their familiarity of the test environment. In an unfamiliar environment, participant #2 thought that the physical and mental workload depend on the environment and frequency of doing independent travel. If an individual needs more information for completing the task, the task is more demanding. However, participants #2 and #3 emphasized that they would prefer if the instructions were given on demand and customized for individuals so that they would have the control. Participant #4 also commented that the information given in this test was excessive, as he had known the environment for two years. Nevertheless, he thought it was useful to know what he had passed and what he would pass during the test. Participant #2 thought it would be useful to choose what information to receive: navigation, orientation, moving objects, or etc. Furthermore, participant #4 suggested providing different user preferences of measurement means when giving instructions in terms of distance, time, and directions. For example, using minutes or meters to measure the distance.

All of the participants thought the information given during the tests was helpful. They also shared the voice instructions given through the bond conduction earphone were clear and did not prevent them from listening to environment noise. Moreover, all of them expressed that they prefer a human voice to interact with them except participant #4. Participant #4 thinks it is OK to hear a synthesized voice, as he hears synthesized voices from all sorts of devices that he uses on a daily basis. All of the participants thought that the human voice with which the instructions were given in the testing was pleasant.

Nielsen suggests that expert users usually focus more on problems at hands (Nielsen, 2010). While participant #1 showed high interests in using the system, participant #4 thinks he would use the system but a combination with a guide dog would provide a more secure feeling. Moreover, participants #2 and #3 shared their concerns on the usefulness of the proposed system, especially in a crowded environment that would be detected when visually impaired and blind people listen to environment noise. Similarly, participants #2 and #3 suggested the proposed system may have problems with detecting moving objects and warning people of them, as people need to process and react to the spoken instructions, in which scenario vibrations may work more effectively. According to their behaviors in mobility, they usually do not travel unaccompanied to unfamiliar environment. Hence, they thought the proposed system would not be needed. Nevertheless, they agreed that it could be useful when they travel to an unfamiliar environment. On the system battery life, they thought it would be necessary for people to be alerted of a decreasing battery life at different battery life levels.

Furthermore, participant #2 suggested there should be an alternative for the glasses, as a cool appearance of glasses is not necessarily the determining factor of people’s decision making.

Not to mention there may be people who do not like wearing glasses. Nevertheless, he shared an interesting perspective that the adaptability of smart glasses and goggles may be increasing, thanks to people’s increasing use of VR and AR products.

One more suggestion from participant #2 was about the test setting. He pointed out that there were too many testing devices like chest-mounted camera and the body-mounted recorder as well as the mock-up glasses, although he understood the devices were for testing purposes. He also suggested that the test should be conducted in a pre-defined environment with a few objects placed at certain spots. In response to his feedback, I explained the test was conducted near their office or home instead of in a pre-defined environment with a certain object setting is for the convenience of the participants, which would help user recruiting under limited financial and time budgets. Also, as explained before, the test setting is sufficient to serve the test purpose at this stage of work.

In summary, all participants expressed that the instructions received during the test journey were clear and helpful. Using bone conduction to communicate the information, they could hear the instructions while listening to the environment noise. All of them thought it was nice to hear a human voice but one also thought it was ok if a synthesized voice was used. Nevertheless, they also shared concerned on usefulness of the proposed design in crowded environment and its battery life as well as the physical artefact of an eyewear. Moreover, they provided suggestions on users' interactions with the proposed system. For instance, users' control on what instructions to receive.

6 Discussion

This paper studied users' needs and spatial models and investigated feasible technologies, based on which it proposed a technological aid for empowering visually impaired and blind people for performing mobility tasks independently.

For visually impaired and blind people, navigation, orientation, and object detection are important tasks of mobility. Based on data collected from the participants, it was not discovered that gender difference and differences in blindness types (congenital blind, early blind, and late blind) introduce differential perceptions and support requirements for mobility tasks. Hence, this paper was focused on the collected data and does not take gender difference and blindness types into consideration. However, the sample size of this study was small so it is also undeniably difficult to assume the findings will be the same if a significantly large group of blind and visually impaired people would have participated in the study.

The proposed aid is intended to utilize machine learning mechanisms to support computer vision-based navigation, orientation and object detection in real time. It aims to capture, process and interpret the environment in order to provide suggestions to users. The suggestions will aid users in decision making related to taking actions at any given temporal and spatial point of their travel. However, the majority of participants in these user studies had high awareness and acceptance of technologies. Consequently, the proposed design must be evaluated by a large group of users with different backgrounds and various levels of awareness and acceptance of technologies. In this way, a more thorough investigation of usefulness and effectiveness of the design can be accomplished, which will help discover possible adjustments of the design.

Nevertheless, even with the small user base included in this study, users' needs and perception models of the essential mobility tasks could still uncovered and validated. This discovery was used to construct the conceptual design for empowering people with task execution. The user evaluation of the proposed design indicated that the general idea will be able to provide support to users for travelling on their own, even though it is still in an early stage of the design. To continue with the design, it is anticipated that further studies are required in several aspects.

7 Future Work

The proposed conceptual design provided a PoC, which has started a series of studies that require further work in the future. The current design proposal does not involve use of a mobile phone, as not every blind and visually impaired individual uses a modern smart phone. Nevertheless, there are several factors that may lead an alternation of the presented design proposal.

I will conclude and elaborate the future work in three aspects: design, technical implementation, and user testing.

7.1 Design

As the current design proposal is a high-level concept, based on current understanding of users and needs, user personas and use scenarios can be created in order to generate detailed requirements and unveil scenario-specific problems using scenario-based design method at the next round of the design iteration (Benyon, 2014).

Also, it is necessary to look for an alternative for the eyewear in order to engage users who do not like eyewear. Google recently announced the mobile app Lookout for object detection is supposed be used with a mobile phone worn around a user's neck or in a shirt pocket (Moon, 2018), which implies the potential alternative should be used at least on the above-chest level. To fulfil the requirement of positioning, an alternative to the eyewear can be a necklace or any other accessories that can be worn around neck or above.

As the proposed system is intended to learn user behaviors and cater for users based on interactions between users and the system with the underlying machine learning mechanisms, the interactions initiated by the system, such as instructions, should be provided based on the immediate context or completely on-demand depending what control mode users choose. However, despite the control mode, the system should alert users of potential harmful obstacles autonomously to avoid human errors, although users can stop the alerts if they are already aware of the given situation. Regarding the user preference of measurement units, different options should be provided so that users can choose whichever does not require cognitive efforts to interpret.

Furthermore, on detection of moving objects and crowds, a new proposal should be given in details that cover both technology and interaction aspects for providing guidance in navigation in crowds and in avoidance of moving objects.

In terms of expressions, BlindSquare gives directions using clockwise instructions. However, whether this expression is sufficient, or another expression is preferred must be verified.

For the future work, all of the considerations and constraints introduced by those factors must be experimented with and validated in different use scenarios in order to find the best suitable design.

7.2 Technical Implementation

There are several problems to solve before making technical implementation of the design:

- Which part of the aiding system will handle computation for image processing and recognition besides other computational demands?

- What is needed to support the aiding system when its battery runs out and network is not available?
- What third party service will be used?

This thesis discussed image processing in visible light and low light environment, for which further studies are expected to investigate which of the following approaches is more efficient and effective: the image processing of far infrared can be turned on while the one of visible images is turned off; image fusions for far infrared and visible images. Moreover, the image recognition for images from far infrared must be developed besides the one for visible images, which will be explored in the future work instead of this thesis.

In order to develop efficient image and processing recognition that facilitates real-time support for users on the fly, the leverage between image quality including size and resolutions and sufficient accuracy for required quality of image recognition should be located. In order to find the leverage, research on algorithms of mobile neural networks and corresponding experiments are necessary.

The implementation needs to be robust and useful even when the product runs out of battery and cannot detect mobile network reception.

As computer vision can provide complementary input for navigation in addition to GPS, this allows a system to expand its ability to navigate without relying on GPS, which may provide possibilities in navigating using obtained spatial recognitions under circumstances where network connections are not available. Hence, in the future study, both a combination of GPS and vision-based and merely vision-based approaches together with a supporting machine learning mechanisms should be investigated further in order to provide users robust and precise navigation.

In terms of the system physical artefacts, such as long-lasting batteries and alerts of decreasing battery life, these would be part of the physical design within the next iteration. As image processing and recognition of vision-based object detection and navigation is concerned with huge computation. A potential supply for long-lasting batteries in an outdoor environment may be from Sunlight. The eyewear glasses could be made of semi-transparent organic solar cells that are based on hydrocarbons (Landerer et al., 2017) .

If the battery that can be integrated into the eyewear cannot keep pace with the power consumption in relation to its battery life, a mobile phone may be employed to provide support in this situation. Furthermore, in terms of cost-effectiveness, a further investigation must be conducted to verify if vision-based object detection and navigation can be powered by the eyewear at the same or a similar cost of a smart phone that supports the required computation. The hardware of such as phone is theoretically capable of handling the required computation by the proposed aiding system, although the use of mobile phone for computation will need to take data transmission latency between the eyewear and the phone into consideration.

Moreover, a third-party map service is not necessarily free of charge for users. For instance, BlindSquare costs 419 Swedish Kronor to download and use the application ("App Store Preview BlindSquare," n.d.). Therefore, whichever map service will be activated by the eyewear must be validated against users' registration before the activation.

For the future work, all of the considerations and constrains introduced by those factors must be tested and validated in different use scenarios in order to find the best suitable design.

7.3 USER EVALUATION

It is interesting and expected to hear experts' insights on the potential problems of the design proposal and their thoughts on the design details. Hence, for future testing, a test plan will start with clear and brief introduction to the tested design, how the proposed system works, and the testing method, such as that testers act as the system during the test. In the future testing, when there is a working prototype as well as sufficient financial and human resources to exploit, a pre-defined environment with pre-set objects should be used to maintain a consistent testing setup.

ACKNOWLEDGEMENT

I would like to extend sincere thanks to all the people who helped me during the thesis work:

Mikael Ståhl, Henrik Götesson, Lennart Karlsson, and all participants in the user studies for their time and suggestions;

Huria Hussein for assisted me with user evaluation on the design proposal;

Lars Oestreicher and Mats Lind for supervising and reviewing the thesis;

Yunfeng Xu and Sanbao Xu for advising me on feasible technologies and technical requirements.

REFERENCE

A Smart Stick Puts Digitization in Service to the Blind. (n.d.). Retrieved January 6, 2018, from https://www.cisco.com/c/en/us/about/csr/stories/social-innovation-in-france.html

Abdel-Wahab, A. G., & El-Masry, A. A. A. (2011). *Mobile Information Communication Technologies Adoption in Developing Countries: Effects and Implications*. https://doi.org/10.4018/978-1-61692-818-6

Alkhanifer, A., & Ludi, S. (2014). Towards a Situation Awareness Design to Improve Visually Impaired Orienta tion in Unfamiliar Buildings : Requirements Elicitation Study. *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, 23–32. Retrieved from https://pdfs.semanticscholar.org/a5d6/1848518247a10e5996539e9354c7004c333a.pdf%0Ahttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6912244

Ambutech. (n.d.).

American Foundation for the Blind. (n.d.). Vision Terms. Retrieved March 25, 2018, from http://www.afb.org/info/blindness-statistics/key-definitions-of-statistical-terms/25

App Store Preview BlindSquare. (n.d.). Retrieved March 2, 2018, from https://itunes.apple.com/app/blindsquare/id500557255

Arditi, A., Holtzman, J. D., & Kosslyn, S. M. (1988). Mental imagery and sensory experience in congenital blindness. *Neuropsychologia*, *26*(1), 1–12.

Armstrong, J. . (1975). Evaluation of Man-machine Systems in the Mobility of the Visually Handicapped. In T. T. Picket RM (Ed.), *Human Factors in Health Care* (pp. 331–343). Lexington Books.

Babich, N. (2017). Prototyping 101: The Difference between Low-Fidelity and High-Fidelity Prototypes and When to Use Each. https://doi.org/https://theblog.adobe.com/prototyping-difference-low-fidelity-high-fidelity-prototypes-use/

Barański, P., & Strumillo, P. (2012). Enhancing positioning accuracy in urban terrain by fusing data from a GPS receiver, inertial sensors, stereo-camera and digital maps for pedestrian navigation. *Sensors (Switzerland)*, *12*(6), 6764–6801. https://doi.org/10.3390/s120606764

Benyon, D. (2014). *Designing interactive systems: A comprehensive guide to HCI, UX and interaction design* (Thrid Edit). Harlow: Pearson Education.

BlindSquare. (n.d.). BlindSquare.

Bone conduction. (n.d.). Retrieved June 16, 2018, from https://en.wikipedia.org/wiki/Bone_conduction

Brabyn, J., Alden, A., Haegerstrom-Portnoy, G., & Schneck, M. (2002). GPS Performance for Blind Navigation in Urban Pedestrian Settings. In *Proceedings of Vision 2002*. Göteborg,

Sweden.
Brahmbhatt, K., Pai, A. R., & Singh, S. (2017). Neural network approach for vision-based track navigation using low-powered computers on MAVs. *2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017*, *2017–Janua*, 578–583. https://doi.org/10.1109/ICACCI.2017.8125902
Brambring, M. (1985). Mobility and orientation processes of the blind. *Electronic Spatial Sensing for the Blind*, *99*, 493–508. https://doi.org/https://doi.org/10.1007/978-94-017-1400-6_33
Brilhault, A., Kammoun, S., Gutierrez, O., Truillet, P., & Jouffrais, C. (2011). Fusion of artificial vision and GPS to improve blind pedestrian positioning. *2011 4th IFIP International Conference on New Technologies, Mobility and Security, NTMS 2011 - Proceedings*, (March). https://doi.org/10.1109/NTMS.2011.5721061
Bull, R., Rathborn, H., & Clifford, B. R. (1983). The voice-recognition accuracy of blind listeners. *Perception*, *12*(2), 223–226.
Cambridge English Dictionary. (n.d.). echo. Retrieved March 28, 2018, from https://dictionary.cambridge.org/dictionary/english/echo
Cameron, C., Swain, J., & French, S. (2003). *Controversial Issues in a Disabling Society*. Buckingham, UK: Open University Press. https://doi.org/10.1017/S0047279403337195
Canadian National Institute for the Blind. (n.d.). The White Cane. Retrieved January 3, 2018, from http://www.cnib.ca/en/living/safe-travel/white-cane/Pages/default.aspx
Cattaneo, Z., & Vecchi, T. (2011). *Blind Vision: The Neuroscience of Visual Impairment*. MIT Press.
Clark-Carter, D., Heyes, A., & Howarth, C. (1986). The efficiency and walking speed of visually impaired people. *Ergonomics*, *29*(6), 779–789. https://doi.org/10.1080/00140138608968314
Computer System Usability Questionnaire. (n.d.). Retrieved April 25, 2018, from http://garyperlman.com/quest/quest.cgi
Cordts, M., Rehfeld, T., Schneider, L., Pfeiffer, D., Enzweiler, M., Roth, S., … Franke, U. (2017). The Stixel World: A medium-level representation of traffic scenes. *Image and Vision Computing*, *68*, 40–52. https://doi.org/10.1016/j.imavis.2017.01.009
Cornoldi, C., Calore, D., & Pra-Baldi, A. (1979). Imagery rating and recall in congenitally blind subjects. *Perceptual and Motor Skills*, *48*(2), 627–639.
Coughlan, J. M., & Miele, J. (2017). AR4VI: AR as an Accessibility Tool for People with Visual Impairments. *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*, 288–292. https://doi.org/10.1109/ISMAR-Adjunct.2017.89
Csapó, Á., Wersényi, G., Nagy, H., & Stockman, T. (2015). A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research. *Journal on Multimodal User Interfaces*, *9*(4), 275–286. https://doi.org/10.1007/s12193-015-0182-7
Damaschini, R., Legras, R., Leroux, R., & Farcy, R. (2005). Electronic travel aid for blind people. In A. Pruski & H. Knops (Eds.), *Assistive Technology: From Virtuality to Reality* (pp. 251–255). IOS Press.
Enzweiler, M., Hummel, M., Pfeiffer, D., & Franke, U. (2012). Efficient stixel-based object recognition. *IEEE Intelligent Vehicles Symposium, Proceedings*, 1066–1071. https://doi.org/10.1109/IVS.2012.6232137
Erbs, F., Schwarz, B., & Franke, U. (2013). From stixels to objects — A conditional random field based approach. In *Intelligent Vehicles Symposium (IV), 2013 IEEE* (pp. 586–591). https://doi.org/10.1109/IVS.2013.6629530
eSight. (n.d.). eSight. Retrieved April 15, 2018, from

https://www.esighteyewear.com/technology
Evans, J., & Mathur, A. (2005). The value of online surveys. *Internet Research*, *15*(2), 195–219.
Farmer, L. W., & Smith, D. L. (1997). Adaptive technology. In B. B. Blasch, R. L. Welsh, & W. R. Wiener (Eds.), *Foundations of orientation and mobility* (pp. 231–259). AFB Press.
Fendri, E., Boukhriss, R. R., & Hammami, M. (2017). Fusion of thermal infrared and visible spectra for robust moving object detection. *Pattern Analysis and Applications*, *20*(4), 907–926. https://doi.org/10.1007/s10044-017-0621-z
Finstad, K. (2010). The Usability Metric for User Experience. *Interacting with Computers*, *22*(5), 323–327. https://doi.org/https://doi-org.ezproxy.its.uu.se/10.1016/j.intcom.2010.04.004
Gartenberg, C. (n.d.). The eSight 3 is an augmented reality headset designed to help the legally blind see. Retrieved April 15, 2018, from https://www.theverge.com/circuitbreaker/2017/2/16/14637804/esight-3-augmented-reality-headset-legally-blind-see
Google Glass. (n.d.). Retrieved April 25, 2018, from https://x.company/glass/
Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. (2004). Neuropsychology: Pitch discrimination in the early blind. *Nature*, *430*(6997), 309.
GPS vs GNSS - What Engineers Need To Know For Their Design. (n.d.). Retrieved April 25, 2018, from https://www.semiconductorstore.com/blog/2018/GPS-vs-GNSS-What-Engineers-Need-To-Know-For-Their-Design-Symmetry-Blog/3083
Händel, P., Yao, Y., Unkuri, N., & Skog, I. (2008). Far Infrared Camera Platform and Experiments for Moose Early Warning Systems*. *JSAE Transactions*.
HANDISCO. (n.d.). Retrieved January 6, 2018, from https://handisco.com/#documentation
Harada, T., Kaneko, Y., Hirahara, Y., Yanashim, K., & Magatani, K. (2004). Development of the navigation system for visually impaired. In *Proceedings of the 26th Annual International Conference of the IEEE EMBS*. San Francisco, CA, USA. https://doi.org/https://doi.org/10.1109/IEMBS.2004.1404355
How Wayfindr Guided My First Steps to Independence on the Tube. (n.d.). Retrieved January 9, 2018, from https://www.wayfindr.net/blog/wayfindr-guided-independence-tube
Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., … Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. https://doi.org/arXiv:1704.04861
Hull, T., & Mason, H. (1995). Performance of blind children on digit-span tests. *Ournal of Visual Impairment & Blindness*, *89*(2), 166–169.
Ichikari, R., Yanagimachi, T., & Kurata, T. (2016). Augmented Reality Tactile Map with Hand Gesture Recognition. In *Lecture Notes in Computer Science* (Vol. 9579, pp. 123–130). Springer, Cham. https://doi.org/https://doi-org.ezproxy.its.uu.se/10.1007/978-3-319-41267-2_17
Jaderberg, M., Mnih, V., Czarnecki, W. M., Schaul, T., Leibo, J. Z., Silver, D., & Kavukcuoglu, K. (2016). Reinforcement Learning with Unsupervised Auxiliary Tasks, 1–14. https://doi.org/10.1051/0004-6361/201527329
Kammoun, S., Parseihian, G., Gutierrez, O., Brilhault, A., Serpa, A., Raynal, M., … Jouffrais, C. (2012). Navigation and space perception assistance for the visually impaired: The NAVIG project. *Irbm*, *33*(2), 182–189. https://doi.org/10.1016/j.irbm.2012.01.009
Kanagaratnam, K. (2009). Smart Mobility Cane : Design of Obstacle Detection Smart Mobility Cane : Design of Obstacle Detection, (April).
Katz, B. F. G., Truillet, P., Thorpe, S. J., Jouffrais, C., & Jouffrais. (2010). NAVIG: Navigation Assisted by Artificial Vision and GNSS. *Workshop on Multimodal Location Based Techniques for Extreme Navigation*, (1), 1–4. Retrieved from

http://www.irit.fr/~Philippe.Truillet/publications/2010/2010_Pervasive10.pdf
Kay, L. (1974). A sonar aid to enhance spatial perception of the blind: engineering design and evaluation. *Radio and Electronic Engineer*, *44*(11), 605. https://doi.org/10.1049/ree.1974.0148
Keating, D. (2008). *Assistive Technology for Visually Impaired and Blind People*. (M. A. Hersh & M. A. Johnson, Eds.) (1st ed.). Springer.
Kim, D. S., Wall Emerson, R., Naghshineh, K., & Auer, A. (2017). Drop-off detection with the long cane: effect of cane shaft weight and rigidity on performance. *Ergonomics*, *60*(1), 59–68. https://doi.org/10.1080/00140139.2016.1171403
Kim, S. Y., & Cho, K. (2013). Usability and Design Guidelines of Smart Canes for Users with Visual Impairments, *7*(1), 99–110. Retrieved from http://www.ijdesign.org/index.php/IJDesign/article/view/1209
Lacey, G., Namara, S. Mac, & Dawson-Howe, K. M. (1998). Personal Adaptive Mobility Aid for the Infirm and Elderly Blind. In *Assistive Technology and Artificial Intelligence* (Vol. 1458, pp. 211–220). https://doi.org/https://doi.org/10.1007/BFb0055980
Lahav, O., Schloerb, D. W., Kumar, S., & Srinivasan, M. A. (2011). A Virtual Map to Support People Who Are Blind in Navigation through Real Spaces. *Journal of Special Education Technology*, *26*(4), 41–57. https://doi.org/10.1177/016264341102600404
Landerer, D., Bahro, D., Röhm, H., Koppitz, M., Mertens, A., Manger, F., … Colsmann, A. (2017). Solar Glasses: A Case Study on Semitransparent Organic Solar Cells for Self-Powered, Smart, Wearable Devices. *Energy Technology*, *5*(11), 1936–1945. https://doi.org/10.1002/ente.201700226
Lansingh, V., Carter, M., Ulldemolins, A., Valencia, L., & Eckert, K. (2012). Social inequalities in blindness and visual impairment: A review of social determinants. *Indian Journal of Ophthalmology*, *60*(5), 368. https://doi.org/10.4103/0301-4738.100529
Lee, J. H., Choi, J. S., Jeon, E. S., Kim, Y. G., Le, T. T., Shin, K. Y., … Park, K. R. (2015). Robust pedestrian detection by combining visible and thermal infrared cameras. *Sensors (Switzerland)*, *15*(5), 10580–10615. https://doi.org/10.3390/s150510580
Lee, Y. H., & Medioni, G. (2016). RGB-D camera based wearable navigation system for the visually impaired. *Computer Vision and Image Understanding*, *149*, 3–20. https://doi.org/10.1016/j.cviu.2016.03.019
Lessard, N., Pare, M., Lepore, F., & Lassonde, M. (1998). Early blind human sunjects localise sound sources better than sighted subjects. *Nature*, *395*(September), 278–280.
Lewald, J. (n.d.). Vertical sound localization in blind humans. *Neuropsychologia*, (12), 1868–1872.
Lewis, L., Sharples, S., Chandler, E., & Worsfold, J. (2015). Hearing the way: Requirements and preferences for technology-supported navigation aids. *Applied Ergonomics*, *48*, 56–69. https://doi.org/10.1016/j.apergo.2014.11.004
Li, X., Wang, J., & Li, T. (2013). Seamless positioning and navigation by using geo-referenced images and multi-sensor data. *Sensors (Basel, Switzerland)*, *13*(7), 9047–9069. https://doi.org/10.3390/s130709047
Lions Club International. (2014). White cane, 1.
Manduchi, R., & Kurniawan, S. (2011). Mobility-related accidents experienced by people with visual impairment. *Insight: Research and Practice in Visual Impairment and Blindness*, *4*(2), 44–54.
Marin-Lamellet, C., Pachiaudi, G., & Le Breton-Gadegbeku, B. (2001). Information and Orientation Needs of Blind and Partially Sighted People in Public Transportation BIOVAM Project. *Transportation Research Record*, (1779), 203–208.
Martinez, M., & Koester, D. (2017). Using Technology Developed for Autonomous Cars to Help Navigate Blind People 2 . Synergies between Autonomous Cars and.

https://doi.org/10.1109/ICCVW.2017.169
Mau, S., & Melchior, N. (2008). *BlindAid: An ElectronicTravel Aid for the Blind*. Retrieved from http://swing.adm.ri.cmu.edu/pub_files/pub4/mau_sandra_2008_1/mau_sandra_2008_1.pdf
Medical Definition of Blindness. (n.d.). Retrieved January 4, 2018, from https://www.medicinenet.com/script/main/art.asp?articlekey=20629
Mirowski, P., Grimes, M. K., Malinowski, M., Hermann, K. M., Anderson, K., Teplyashin, D., … Hadsell, R. (2018). Learning to Navigate in Cities Without a Map. Retrieved from http://arxiv.org/abs/1804.00168
Mirowski, P., Pascanu, R., Viola, F., Soyer, H., Ballard, A. J., Banino, A., … Hadsell, R. (2016). Learning to Navigate in Complex Environments. https://doi.org/10.1007/s10845-014-0909-4
Moon, M. (2018). Google's Lookout will help the blind navigate their environment. Retrieved May 15, 2018, from https://www.engadget.com/2018/05/08/google-lookout-app/
Morrongiello, B. A. (1994). Effects of colocation on auditory-visual interactions and cross-modal perception in infants. In *The Development of Intersensory Perception: Comparative Perspectives* (pp. 235–263).
NASA-TLX. (n.d.). Retrieved April 25, 2018, from https://en.wikipedia.org/wiki/NASA-TLX
Nichols, A. (1995). Why Use The Long White Cane.
Nielsen, J. (2010). Testing Expert Users. Retrieved April 28, 2018, from https://www.nngroup.com/articles/testing-expert-users/
Ogden, J., & Barker, K. (2001). Imagery used in autobiographical recall in early and late blind adults. *Journal of Mental Imagery*, *Journal of*(25), 107–130.
OrCam User Stories. (n.d.). Retrieved March 26, 2018, from https://www.orcam.com/en/blog/category/testimonials/
Oxford Learner's Dictionaries. (n.d.). GPS. Retrieved January 8, 2018, from https://www.oxfordlearnersdictionaries.com/definition/english/gps
Passini, R., Dupré, A., & Langlois, C. (1986). Spatial mobility of the visually handicapped active person: A descriptive study. *Journal of Visual Impairment & Blindness*, *80*(8), 904–907.
Peli, E., & Woods, R. L. (2009). Image enhancement for impaired vision: the challenge of evaluation. *Int J Artif Intell Tools*, *18*(3), 415–438. https://doi.org/10.1142/S0218213009000214
Pendleton, S., Andersen, H., Du, X., Shen, X., Meghjani, M., Eng, Y., … Ang, M. (2017). Perception, Planning, Control, and Coordination for Autonomous Vehicles. *Machines*, *5*(1), 6. https://doi.org/10.3390/machines5010006
Perspective. (n.d.). Retrieved April 8, 2018, from https://dictionary.cambridge.org/dictionary/english/perspective
Pfeiffer, D., Erbs, F., & Franke, U. (2012). *Pixels, Stixels, and Objects*. (A. Fusiello, V. Murino, & R. Cucchiara, Eds.), *Computer Vision – ECCV 2012. Workshops and Demonstrations. ECCV 2012. Lecture Notes in Computer Science* (Vol. 7585). Springer, Berlin, Heidelberg. https://doi.org/https://doi-org.ezproxy.its.uu.se/10.1007/978-3-642-33885-4_1
Pfeiffer, D., & Franke, U. (2010). Efficient representation of traffic scenes by means of dynamic stixels. *IEEE Intelligent Vehicles Symposium, Proceedings*, 217–224. https://doi.org/10.1109/IVS.2010.5548114
Pinho, F., Carvalho, A., & Carreira, R. (2015). Improving geolocation by combining GPS with image analysis. *Lecture Notes in Geoinformation and Cartography*, *214*, 213–225. https://doi.org/10.1007/978-3-319-11463-7_15

Pyun, R., Kim, Y., Wespe, P., Gassert, R., & Schneller, S. (2013). Advanced Augmented White Cane with obstacle height and distance feedback. *IEEE International Conference on Rehabilitation Robotics*. https://doi.org/10.1109/ICORR.2013.6650358

Rituerto, A., Fusco, G., & Coughlan, J. M. (2016). Towards a Sign-Based Indoor Navigation System for People with Visual Impairments. *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '16*, 287–288. https://doi.org/10.1145/2982142.2982202

Röder, B., & Neville, H. J. (2003). Developmental functional plasticity. In J. Grafman & I. H. Robertson (Eds.), *Handbook of Neuropsychology: Plasticity and Rehabilitation* (pp. 231–270). Amsterdam: Elsevier Science,.

Röder, B., Rösler, F., & Neville, H. J. (2001). Auditory memory in congenitally blind adults: A behavioral-electrophysiological investigation. *Brain Research*, *11*(2), 289–303. https://doi.org/https://doi.org/10.1016/S0926-6410(01)00002-7

Röder, B., Teder-Salejarvi, W., Sterr, A., Rösler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, *400*(6740), 162–166.

Roentgen, U. R., Gelderblom, G. J., Soede, M., & Witte, L. P. de. (2008). Inventory of Electronic Mobility Aids for Persons with Visual Impairments : A Literature Review. *Journal of Visual Impairment & Blindness*, *102*(11), 702–724. Retrieved from http://ezproxy.its.uu.se/login?url=https://search-proquest-com.ezproxy.its.uu.se/docview/222058014?accountid=14715

Saito, H., Hagihara, T., Hatanaka, K., & Sawai, T. (2008). Development of pedestrian detection system using far-infrared ray camera. *SEI Technical Review*, (66), 112–117.

Sander, M.-S., Bournot, M.-C., Lelièvre, F., & Tallec, A. (2005). Les personnes ayant un handicap visuel. Les apports de l'enquête Handicaps - Incapacités - Dépendance [The people with visual impairment. Contributions of the survey handicap-incapacities-dependencies]. *DREES, Etudes et Resultats*, *416*, 1–12. Retrieved from http://www.sante.gouv.fr/drees/etude-resultat/er416/er416.pdf%5Cnhttp://www.sante.gouv.fr/drees/etude-resultat/er416/handicapvisuel.pdf

Sapkota, N. (2013). *Real time digital night vision using nonlinear contrast enhancement*. *ProQuest Dissertations and Theses*. Retrieved from https://search.proquest.com/docview/1428873298?accountid=45153

Schinazi, V. R., Thrash, T., & Chebat, D. R. (2016). Spatial navigation by congenitally blind individuals. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(1), 37–58. https://doi.org/10.1002/wcs.1375

Schneider, L., Cordts, M., Rehfeld, T., Pfeiffer, D., Enzweiler, M., Franke, U., … Roth, S. (2016). Semantic Stixels: Depth is Not Enough, (Iv). Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7535373&tag=1

Seeing AI. (n.d.). Retrieved March 26, 2018, from https://www.microsoft.com/en-us/seeing-ai

Statler, S. (2016a). Choosing the Right Beacon. In *Beacon Technologies* (pp. 81–119). Apress, Berkeley, CA. https://doi.org/https://doi.org/10.1007/978-1-4842-1889-1_7

Statler, S. (2016b). Introduction. In *Beacon Technologies*. Apress, Berkeley, CA.

Stixel World. (n.d.). Retrieved April 24, 2018, from http://www.6d-vision.com/aktuelle-forschung/stixel-world

Strong, P. (2019). The History of the White Cane. Retrieved January 3, 2018, from http://www.acb.org/tennessee/white_cane_history.html

Tinti, C., Galati, D., Vecchio, M. G., De Beni, R., & Cornoldi, C. (1999). Interactive auditory and visual images in totally blind persons. *Ournal of Visual Impairment & Blindness*, *93*, 579–583.

VALIDATING WAYFINDR. (n.d.). Retrieved April 8, 2018, from

https://www.ustwo.com/blog/validating-wayfindr/
Vanlierde, A., & Wanet-Defalque, M. C. (2005). The role of visual experience in mental imagery. *Journal of Visual Impairment and Blindness*, *99*, 165–178.
Vishal, K., Jawahar, C. V., & Chari, V. (2015). Accurate localization by fusing images and GPS signals. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2015–Octob*, 17–24. https://doi.org/10.1109/CVPRW.2015.7301390
Wan, C. Y., Wood, A. G., Reutens, D. C., & Wilson, S. J. (2010). Early but not late-blindness leads to enhanced auditory perception. *Neuropsychologia*, *48*(1), 344–348. https://doi.org/10.1016/j.neuropsychologia.2009.08.016
Wang, H., Cai, Y., Chen, X., & Chen, L. (2016). Night-time vehicle sensing in far infrared image with deep learning. *Journal of Sensors*, *2016*. https://doi.org/10.1155/2016/3403451
Wang, J., Garratt, M., Lambert, A., Wang, J. J., Han, S., & Sinclair, D. (2004). Integration of Gps/Ins/Vision Sensors To Navigate Unmanned Aerial Vehicles. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XXXVII*(Part B1), 963–970.
Wang, Y., Virrantaus, K., Pei, L., Chen, R., & Chen, Y. (2012). 3D personal navigation in smart phone using geocoded images. *Record - IEEE PLANS, Position Location and Navigation Symposium*, 584–589. https://doi.org/10.1109/PLANS.2012.6236931
wayfindr. (n.d.). Retrieved January 6, 2018, from https://www.wayfindr.net/,
What is computer vision? (n.d.). Retrieved April 3, 2018, from http://www.bmva.org/visionoverview
Whitecane. (n.d.). Retrieved January 8, 2018, from https://en.wikipedia.org/wiki/White_cane
WHO. (1980). *International Classification of Impairments, Disabilities and Handicaps*. Geneva, Switzerland.
Why Would Someone Need a White Cane? (n.d.). Retrieved January 8, 2018, from https://www.dhs.wisconsin.gov/blind/whitecane/information.htm
World Health Organization. (n.d.). Blindness and visual impairment. Retrieved March 25, 2018, from http://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment
World Health Organization. (2012). *GLOBAL DATA ON VISUAL IMPAIRMENTS 2010*.
Zhang, F., Duarte, F., Ma, R., Milioris, D., Lin, H., & Ratti, C. (2016). Indoor Space Recognition using Deep Convolutional Neural Network: A Case Study at MIT Campus. Retrieved from http://arxiv.org/abs/1610.02414
Zimler, J., & Keenan, J. M. (1983). Imagery in the congenitally blind: How visual are visual images? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(2), 269–282.
Zwiers, M. P., Van Opstal, A. J., & Cruysberg, J. R. (n.d.). A spatial hearing deficit in early-blind humans. *Journal of Neuroscience*, *21*(9), RC142: 1-5.