WASEDA UNIVERSITY

DOCTORAL THESIS

Assisting Blind People's Independence in Public Spaces

Author:
Seita KAYUKAWA

Supervisor:
Dr. Shigeo MORISHIMA

Graduate School of Advanced Science and Engineering
Department of Pure and Applied Physics
Research on Image Processing

July, 2022

Waseda University

Abstract

Graduate School of Advanced Science and Engineering
Department of Pure and Applied Physics

Doctor of Engineering

Assisting Blind People's Independence in Public Spaces

by Seita Kayukawa

This thesis explores how computational technologies can assist blind people to gain independence in public spaces and the challenges faced by blind people using assistive technologies in the public domain. Public spaces such as shopping malls, airports, and museums are dynamic environments with complex structures and various facilities that are shared by other people. Blind people in public spaces must rely on assistance from others. The ability to travel independently through public spaces is a fundamental dream for every blind person.

To achieve its goals, this thesis proposes three approaches. First, we assist blind visitors' *"exploration tasks"* in public spaces. When sighted visitors with specific purposes, interests, and knowledge explore unfamiliar public spaces, they rely on the visual information available in public spaces, such as doorways, maps, signage, and surrounding facilities. To help blind people explore public spaces, we introduce a museum navigation robot, a wayfinding system that recognizes surrounding signage in public buildings, and smartphone-based walking assistance that avoids obstacles and recognizes intersections in indoor corridors.

Second, we present assistive systems that facilitate *"social behaviors"* by blind people in public spaces. In the dynamic environments of public spaces shared by other people, sighted people decide their appropriate behavior in real-time while visually sensing the movements and behaviors of surrounding people. For example, sighted people avoid collisions with nearby pedestrians by continuously adapting their speed and direction. While waiting in line, sighted people find the end-of-line position and follow the preceding person with intermittent movements. The assistive systems designed in this thesis will help blind people to adopt the social behaviors of sighted people.

Third, this thesis investigates the *"social acceptance"* of autonomous navigation robots for blind people in public buildings. Assistive technologies deployed in public spaces must be acceptable not only to blind users but also to other users of the space and to facility managers. To assess the acceptance level of a navigation robot for blind people, this thesis survey the opinions of people in general, facility managers, science museum visitors, and blind users. Based on the results, we discuss the concerns toward the navigation robot.

Based on our user evaluations and investigation results, we finally discuss the future challenges and opportunities for blind people with assistive systems gaining independence in public spaces. The next stage of accessibility research will be conducted in real-world rather than laboratory scenarios and will improve the flexibility and social acceptance of assistive systems to the general public.

Acknowledgements

This thesis is the result of six years of research supported by many advisors, researchers, and friends in multiple countries. My thanks are extended to all individuals (too numerous to mention) who have influenced my work.

My deepest gratitude goes to my Ph.D. thesis advisor, Professor Shigeo Morishima, who provided smart and intellectually stimulating advice throughout this study. Over the past six years, I have learned much from his guidance, insights, and vision. I also thank him for supporting my projects and providing me with many opportunities. I am grateful to the thesis committee members, Prof. Shigeo Morishima, Prof. Hideyuki Sawada, and Dr. Chieko Asakawa, who offered thoughtful and constructive suggestions for improving my thesis.

I especially thank Dr. Keita Higuchi, who introduced me to HCI research (and craft beer) and taught me how to conduct successful and valuable HCI research projects. Meeting him shortly after embarking on my research was a most fortunate occurrence. I also especially thank Dr. Chieko Asakawa and Dr. Hironobu Takagi for their professional advice as experts in accessibility research. Discussions with them have always been fruitful and have helped me to expand the boundary of my research by considering the future direction of present ideas.

I thank my many collaborators, Masaki Kuribayashi, Yutaro Yamanaka, Daisuke Sato, Tatsuya Ishihara, Masayuki Murata, Akihiro Kosugi, Jayakorn Vongkulbhisal, João Guerreiro, Ken Sakurada, Yoichi Sato, and Kris Kitani. This thesis is based on papers I wrote with them.

I have spent considerable time with past and present members of Morishima Laboratory, especially Ryota Natsuma, Tomoya Yamaguchi, Kazuki Inoue, Takahiro Itazuri, Tatsuya Yatagawa, Shugo Yamaguchi, Masanori Nakamura, Ayano Kaneda, Takayuki Nakatsuka, Qi Feng, Ryo Shimamura, Yukitaka Tsuchiya, Shoki Miyagawa, Keitaro Tanaka, Asuka Hirata, Taichi Higasa, Yusuke Miura, Hideki Tsunashima, Tsukasa Nozawa, Shintaro Yamamoto, Yoshihiro Fukuhara, and Satoko Matsuda. Our discussions on interesting research topics, including deep learning, computer vision, computer graphics, and sound processing, are fondly remembered. I also thank the members of Miraikan Accessibility Lab., especially Xiyue Wang, Takashi Suzuki, Kanami Kamoshita, and Fuki Miyazaki.

I express much gratitude to my parents, Akio and Shoko, for all their support. Thank you for always encouraging me to keep moving forward. I also thank my sister, Aoba, for her help and support.

Finally and most importantly, my peerless partner, Rie, has constantly supported me, cheered me, encouraged me, and remained with me. These few words cannot express my debt to you, but I just want to thank you for all of your continuous support. I am so excited to start the next chapter of our lives together. Thank you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> *"I want to do window shopping while listening to surrounding information at a shopping mall. I can do some shopping if someone helps me, but I feel sorry asking when I do not have specific purpose. By using the suitcase [our suitcase-shaped navigation robot], I want to walk freely in a shopping mall by myself and do window shopping."*
>
> Participant 2
> IBM Japan Ltd
> March 25, 2021

Independently traveling through public spaces is a dream of blind people. By definition, a public space is open to people. Typical examples are shopping malls, airports, stations, museums, and hospitals. These spaces should be socially inclusive for all visitors, including those with various disabilities. Blind visitors face significant and unique challenges when navigating public spaces. Whereas sighted people rely on the visual information in public spaces, such as doorways, obstacles, shops, and signage, blind people must navigate and move through the real world without this semantic information.

To improve real-world accessibility for blind people, previous researchers proposed various systems that improve the orientation and mobility skills of blind people. Blind navigation systems provide turn-by-turn instructions based on *static* topological route maps and localization methods to navigate users to the front of a destination (Section 2.1.1). Collision-avoidance systems allow blind users to avoid or find *static* objects such as walls, chairs, boxes, and poles. These systems detect objects using various sensors or computer vision techniques (Sections 2.2 and 2.1.2). Although user studies with blind participants have evaluated the effectiveness of these systems, the experimental environment was typically a *static and controlled* environment reproduced in the laboratory space or controlled to prevent the entry of other pedestrians. Participants were asked to walk along *pre-fixed routes* with the system or to avoid or find *static* objects.

However, wild public spaces are *dynamic* environments occupied by other people. Therefore, the surrounding situation is changing dynamically, and people modify their behavior in consideration of the surrounding people. Blind pedestrians must avoid collisions with walking or standing pedestrians. If blind visitors want to join a standing line, they must find the changing end-of-line position and follow the preceding person while moving intermittently. In addition, people visiting public spaces have different purposes and interests. Public spaces such as shopping malls, airports, and science museums have complex structures and a variety of shops, facilities, and exhibitions. Visitors explore public spaces at their own pace driven by their purpose, interests, and knowledge. Such exploration tasks present a major challenge to blind people [20, 21]. Therefore, the current stream of research is removed from the real situations of blind people visiting public spaces.

The goal of this thesis is to explore how computational technologies can provide independence to blind people in public spaces and to illustrate how blind people with assistive

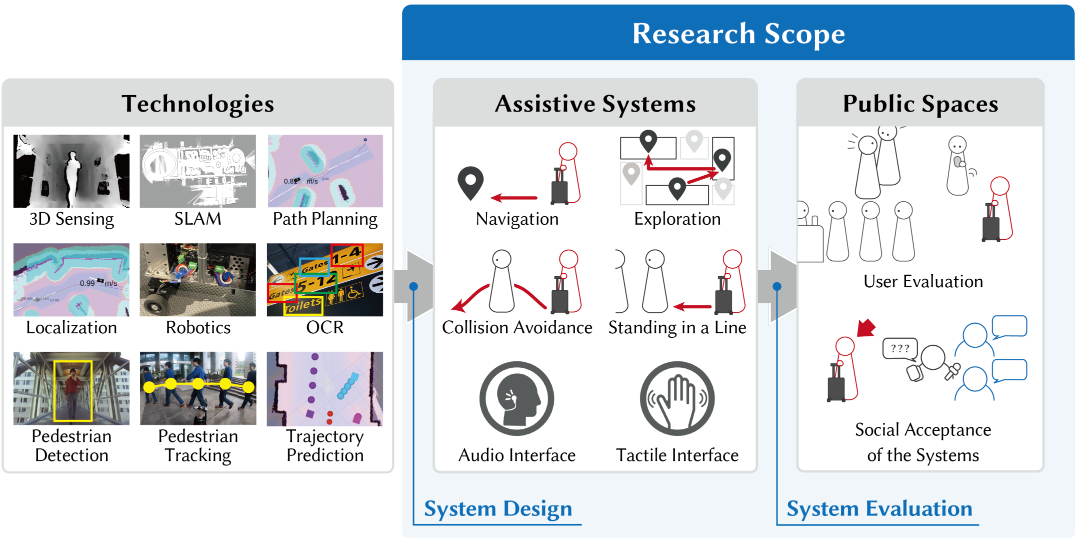


Figure 1.1: Research scope of this thesis. Assistive systems combining technologies such as computer vision and robotics are developed and evaluated. The acceptance and popularization of the assistive systems in public spaces are then discussed.

technologies can advance in public spaces. To this end, it develops assistive systems that help blind people in public spaces with tasks such as obstacle avoidance and navigation. The system combines various computational technologies (e.g., computer vision and robotics) (see Figure 1.1). Moreover, considering that assistive systems are likely to be widely used in future public spaces, we conduct user evaluations in public spaces and investigate the social acceptance of the systems.

1.1 Thesis Statement

What challenges do blind people face in public spaces? To address this question, we consider a scenario in which an unassisted blind person visits a shopping mall. Blind people may arrive at the shopping mall entrance using a global navigation system such as Google Maps [80].

1.1.1 Exploration in Public Spaces

The first challenge of a blind visitor is finding a place to visit. Public spaces such as shopping malls, airports, and science museums are large and complex structures housing a variety of facilities, shops, and exhibitions. The routes of visitors depend on the purpose, interests, and knowledge of the visitors as individuals. Sighted visitors can explore unfamiliar shopping malls at their own pace. They can check the mall map, surrounding shops, and signage, and enjoy window shopping (i.e. visiting different shops with no specific destination). Despite previous efforts to improve the situation, blind visitors cannot easily explore public spaces at their own pace, as they depend on assistance from family, friends, or facility personnel. Blind people prefer not to completely rely on such assistance because they are conscious of burdening sighted assistants [20]. Although blind visitors can select a place to visit by referencing the webpage of the place, independent window shopping is a difficult task for them because they require assistance from another person. Increasing the autonomy of blind visitors is important for allowing their independent experience in public spaces.

Part I of this thesis introduces three projects that provide exploration assistance to blind

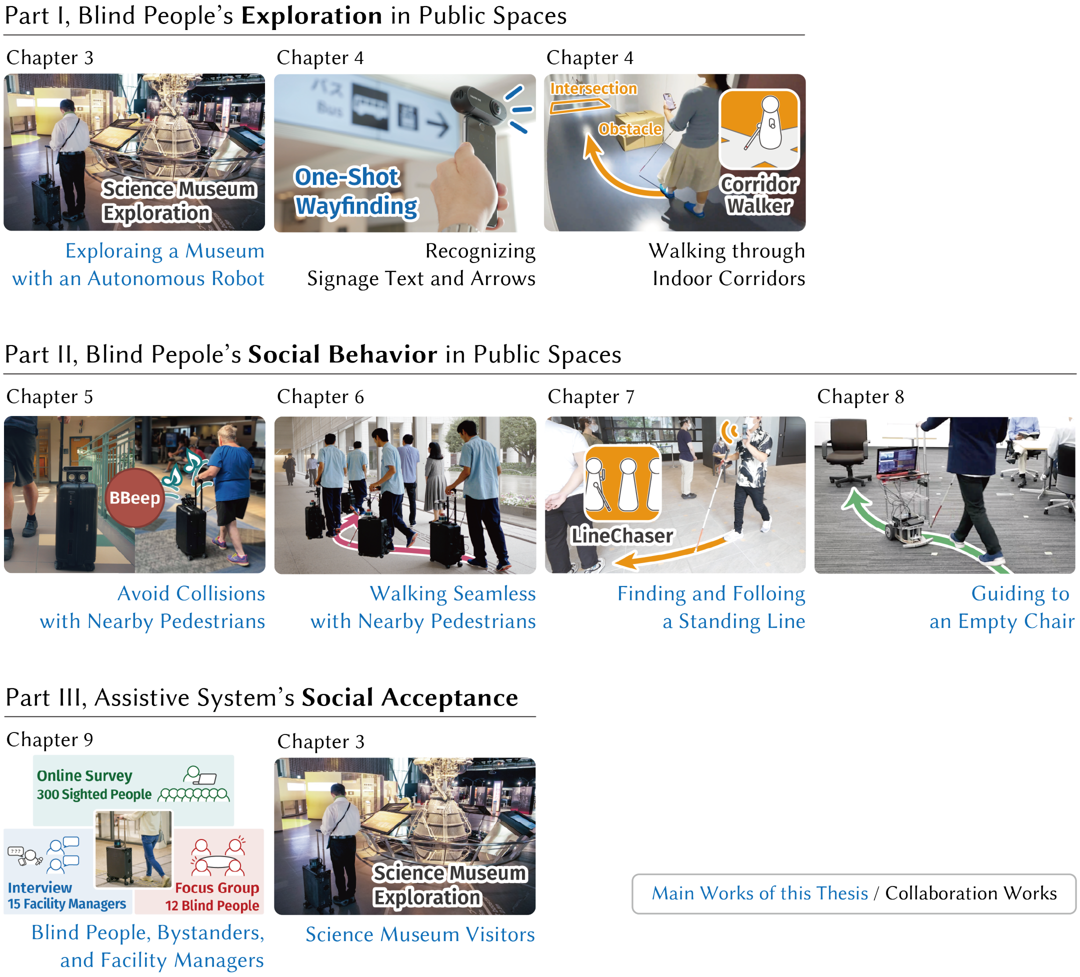


Figure 1.2: Overview of the works presented in this thesis. The main contributions are described in Chapter 3 and Chapters 5–9. Two collaboration works are overviewed in Chapter 4.

people (Figure 1.2). Chapter 3 describes a science-museum exploration system that combines the power of a navigation robot, an audio guide, and the intelligence of human assistants. Blind users can control the robot to navigate them to the desired exhibits while listening to audio descriptions of those exhibits. Users can also browse the detailed explanations on their smartphones and call museum staff if interactive support is needed. This work is currently under review as a paper entitled "Assisting Blind Visitors' Science Museum Exploration Using an Autonomous Navigation Robot". This work was conducted in collaboration with Daisuke Sato from Carnegie Mellon University, Masayuki Murata and Tatsuya Ishihara from IBM Research, Hironobu Takagi and Chieko Asakawa from IBM Research and Miraikan - The National Museum of Emerging Science and Innovation, and Shigeo Morishima from Waseda University.

Chapter 4 introduces two collaboration works that assist blind people to recognize surrounding information using a smartphone. First is the *"One-shot wayfinding system"*, by which blind people can determine the correct direction. The system recognizes surrounding signage and analyzes the relationship between detected text and arrows on the signage to estimate the correct direction to the user's destination. This work has been published as "One-Shot Wayfinding Method for Blind People via OCR and Arrow Analysis with a 360-degree Smartphone Camera" in *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '21)* [254].

This work was conducted in collaboration with Yutaro Yamanaka and Satoshi Kurihara from Keio University, Hironobu Takagi from IBM Research, Yuichi Nagaoka from Tokyo Independent Living Support Center for the Visually Impaired, and Yoshimune Hiratsuka from Department of Ophthalmology, Juntendo University School of Medicine.

The second collaboration project is *"Corridor-Walker"*, a smartphone-based system that helps blind people to avoid obstacles and recognize intersections. This system constructs a 2D occupancy grid map of the surrounding environment using a light detection and ranging (LiDAR) sensor equipped with a smartphone. The system generates an obstacle-avoiding path and detects upcoming intersections on the grid map. This work has been published as "Corridor-Walker: Mobile Indoor Walking Assistance for Blind People to Avoid Obstacles and Recognize Intersections" in *Proceedings of the 24th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '22)* [134] and "Designing a Smartphone-Based Assistance System for Blind People to Recognize Intersections and Obstacles in Indoor Corridors" in *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '21 Poster)* [133]. This work was conducted in collaboration with Masaki Kuribayashi and Shigeo Morishima from Waseda University, Jayakorn Vongkulbhisal, Chieko Asakawa, and Hironobu Takagi from IBM Research, and Daisuke Sato from Carnegie Mellon University.

1.1.2 Social Behavior

While blind visitors are exploring the shopping mall or walking to their destinations, their surrounding environment is constantly altered by the activities of other people. Using their sense of vision, sighted pedestrians continuously adapt their speed and direction to achieve seamless walking among nearby pedestrians [171]. Visitors making a purchase will queue at the cashier stations. When joining a queue, the purchaser finds the end-of-line position and follows the preceding person, which requires intermittent movement. Sighted people visually sense the movements and behaviors of surrounding others and decide their appropriate behavior in real-time. Such behaviors that consider the presence of surrounding people are called *"social behaviors"*. We aim to assist blind people to adopt the same social behaviors as sighted people.

Part II introduces four projects that assist the social behaviors of blind people (Figure 1.2). Chapter 5 presents an assistive suitcase system, *BBeep*, which supports blind people walking through crowded environments. BBeep uses pre-emptive sound notifications that alert both the user and nearby pedestrians of imminent collisions to be avoided. BBeep triggers notifications by tracking pedestrians, predicting their future position in real-time, and providing sound notifications only when a future collision is anticipated. This work has been published as "BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians" in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)* [115]. This work was conducted in collaboration with Keita Higuchi and Yoichi Sato from University of Tokyo, João Guerreiro, Kris Kitani, and Chieko Asakawa from Carnegie Mellon University, and Shigeo Morishima from Waseda University.

Chapter 6 proposes a guiding system that helps blind people to seamlessly walk through public spaces occupied by nearby pedestrians. Blind users carry a rolling suitcase-shaped system containing two RGBD cameras, an inertial measurement unit sensor, and a LiDAR sensor. The system senses the behavior of surrounding pedestrians, predicts the collision risk, and alerts users of pending collisions. Collision avoidance is operated in two modes: the "on-path" mode that reduces the user's walking speed without changing the path, and the "off-path" mode that navigates the use along an alternative path. This work has been published as " Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians" in *Proceedings of the ACM on Interactive, Mobile, Wearable and*

Ubiquitous Technologies (IMWUT), 4, 3, Article 85 [117]. This work was conducted in collaboration with Tatsuya Ishihara, Hironobu Takagi, and Chieko Asakawa from IBM Research and Shigeo Morishima from Waseda University.

Chapter 7 presents a smartphone-based assistive system named *LineChaser*, which navigates a blind user to the end of a line and continuously reports the distance and direction to the last person in the line, who will then be followed by the user. LineChaser uses the RGB camera in a smartphone to detect nearby pedestrians and a built-in infrared depth sensor to estimate the positions of pedestrians. From the pedestrian position estimates, LineChaser determines whether nearby pedestrians are standing in line and notifies the user with audio and vibration signals for start/stop movements. In this way, users can remain in the queue while maintaining the correct social distance. This work has been published as "LineChaser: A Smartphone-Based Navigation System for Blind People to Stand in Line" [135] in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)* and "Smartphone-Based Assistance for Blind People to Stand in Lines" [119] in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20 LBW).*. This work was conducted in collaboration with Masaki Kuribayashi and Shigeo Morishima from Waseda University, Hironobu Takagi and Chieko Asakawa from IBM Research, and João Guerreiro from Carnegie Mellon University.

Chapter 8 proposes an assistive robot, *BlindPilot*, which guides blind users to landmark objects such as an empty chair using an intuitive handle. BlindPilot detects the positions of target objects using an RGB-D camera and builds a 2D map of the surrounding area from LiDAR data. Based on the sensing results, BlindPilot then generates a path to the object and guides the user safely along that path. This work has been published as "BlindPilot: A Robotic Local Navigation System that Leads Blind People to a Landmark Object" in *Extendet Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20 LBW)* [116]. This work was conducted in collaboration with Masaki Kuribayashi and Shigeo Morishima from Waseda University and Hironobu Takagi and Chieko Asakawa from IBM Research.

1.1.3 Social Acceptance of Assistive Systems

At the shopping mall, blind visitors will interact with other visitors and shop staff. How blind visitors with assistive systems are viewed by other people is an important issue for blind users. Koelle et al.considered that "a human–machine interface can be considered socially acceptable if its presence or the user's interactions with it are consistent with the user's self-image and external image, or alter them in a positive way. Human–machine interfaces that cause a negative change to self and external image show a lack of social acceptability" [127]. The social acceptability of an assistive technology for people with disabilities depends not only on functionality, but also on the appearance, privacy, and security of the technology. How can assistive technologies be made acceptable to both users and the public? Addressing this question is imperative to popularizing assistive systems and encouraging blind people into the public sphere.

In Part III of Chapter 9, we investigate the acceptance and concerns regarding autonomous navigation robots for blind people in public buildings. Robots deployed in public buildings must be accepted not only by blind users, but also by other users and facility managers of the buildings. Therefore, we investigated the acceptance and concerns of our prototype, which looks like a regular suitcase, from three perspectives. We first conducted an online survey, which revealed wide acceptance of a navigating robot for blind users. Second, we interviewed facility managers, who expressed concern that the robot's camera would compromise customers' privacy. Finally, we conducted focus group sessions with blind participants experiencing the robot. The discussions revealed that collision risk between other people and an inconspicuous robot may increase if sighted people are unaware of the users' blindness. Nevertheless, the robot's design is easily assimilated into the surrounding environment. This feature was liked by many participants. This work has been published as "How Users, Facility Managers, and Bystanders Perceive and Accept a Navigation Robot for Visually Impaired People in Public Buildings " in *Proceedings of the 31st IEEE International Conference on Robot & Human Interactive Communication (IEEE RO-MAN '22)* [118]. This work was conducted in collaboration with Daisuke Sato from Carnegie Mellon University, Masayuki Murata, Tatsuya Ishihara, Akihiro Kosugi, Hironobu Takagi, and Chieko Asakawa from IBM Research, and Shigeo Morishima from Waseda University.

1.1.4 Increasing the Independence of Blind People in Public Spaces

Finally, we discuss future challenges and opportunities for blind people using assistive systems to gain independence in public spaces (Part IV, Chapter 10). We focus on the safety, privacy, and visibility concerns (see Chapter 9). We finally discuss future needs for flexibility and social acceptability of assistive systems to both users and non-users.

Chapter 2

Related Work

This chapter reviews existing works related to real-world accessibility by blind people. We illustrate the main achievements and remaining challenges of applying these approaches in public environments. We then briefly discuss how this work can bridge the gap between current research approaches and the real world. First, we review the literature on orientation and mobility (O&M) assistance, including blind navigation systems and obstacle avoidance systems. We then summarize the types of hardware design used in assistive technologies for O&M assistance, from smartphone-based to robot-based systems. We also introduce audio and tactile interfaces as nonvisual interfaces for blind people. Finally, we review the social acceptance of assistive technologies, including potential conflicts and benefits for visually impaired people.

2.1 Navigation Systems

2.1.1 Global Navigation

As many blind people already use smartphones [160, 170], commercial navigation apps (e.g., Google Maps [80], Ariadne [48], SoundScape [166], and BlindSquare [168]) have become gradually popularized in the blind community. Various types of global navigation systems for blind people have been proposed [120, 132]. Most of these apps and systems use turn-by-turn navigation to guide blind users to their destination. Navigation is enabled by localization technologies such as global positioning systems (GPSs) [33, 190, 201], magnetic information [77, 207], visual features [144, 259], radio frequency identifiers tags [15, 47, 62, 64, 72, 189], visible-light communication (VLC) [174, 175], and bluetooth low energy (BLE) beacons [6, 44, 57, 122, 172, 215]. Using static topological route maps, these technologies navigate users to the front of a destination with acceptable accuracy (e.g., 1.7m average [172]).

The user effectiveness of most blind navigation systems is evaluated through user studies in which blind participants walk with the system along *pre-fixed routes*. However, visitors to public spaces such as shopping malls and museums have different purposes, interests, and knowledge, and typically explore the spaces at their own pace. Such exploration tasks present major challenges to blind people [20, 21], which have not been adequately explored. This thesis proposes three assistive systems for blind people wishing to explore public spaces (Part I). In Chapter 3, we propose a science-museum exploration system for blind visitors that combines the power of a navigation robot, an audio guide, and the intelligence of human assistants. In Chapter 4, we introduce two collaboration works: (1) Corridor-Walker, a mobile indoor walking assistance for obstacle avoidance and intersection recognition, and (2) the One-shot Wayfinding System, a mobile wayfinding assistance for signage recognition in public buildings.

In addition, although the existing systems can help users reach a *fixed* destination, the destination can change in the dynamic environments of real public spaces occupied by other people. For example, people must often join queues in regular public spaces. For this purpose, people must find and follow the end of a line, which changes over time. Our system is designed to complement the existing navigation system in line standing tasks (Chapter 7).

2.1.2 Local Navigation

While existing global navigation systems can help users reach a *fixed* destination, they do not completely support local navigation, which requires blind users to interact with other pedestrians or certain objects. For example, a global navigation system might guide a blind person safely to a lounge. Once inside the lounge, the user must identify an empty chair using their hands or a white cane. To overcome this limitation, the system must support both global and local navigation.

Most blind people use white canes to identify landmark objects. Canes are useful but their sensing range is limited by their length (approximately 1 m) [124]. To augment the sensing range of blind users, previous studies have proposed local navigation systems that detect landmark objects (e.g., doors [67, 101, 126], chairs [99, 101, 243], and artworks in museums [136]) using an optical head-mounted display [67], a smartphone camera [214], or a stereo camera [99, 101, 243]. These systems also provide navigational information of the detected objects (e.g., direction and distance) via audio [67, 99, 126, 136, 214] or tactile [243] feedback to blind users. Although blind users can approach a certain object while correcting their direction based on the given feedback, such orientation requires continuous fine adjustments. To simplify this situation, we propose a robotic local navigation system that directly leads blind users to a landmark object (Chapter 8).

2.2 Obstacle Avoidance Systems

2.2.1 Obstacle Detection

Besides guide dogs, white canes [31, 247] are the most common tool by which blind people find obstacles and avoid collisions. Although canes are efficient, they must physically hit an obstacle before the object is detected. This is undesirable, especially when the obstacle is a pedestrian. Various supportive technologies allow blind users to detect obstacles with non-contact sensing [155, 185, 236, 241]. These systems detect surrounding obstacles (e.g., walls [114, 243], boxes [104, 114, 197], chairs [144, 243], and poles [197, 210]) using various types of sensors (e.g., lasers [130], ultrasonics [102, 224], time-of-flight distance sensors [114], phone speakers and microphones [234], or depth sensors [38, 68, 85, 95, 98, 100, 143, 191, 197, 243, 257]). The information of the detected obstacles (e.g., distance [98, 100, 102, 130, 153, 191, 197, 243], shape [38, 100, 144], or category [98, 191, 257]) is then conveyed to the user. Users must avoid obstacles based on the information provided. When several obstacles are present, understanding the situation and avoiding all obstacles may be difficult, necessitating path generation for obstacle avoidance.

2.2.2 Safe Path Generation

Some systems plan a path around the detected obstacles and navigate users through different interfaces. For instance, sound feedback [144, 197, 261], tactile feedback [140, 141], a cane connected to a wheeled robot [235], a mobile robot [85, 150, 195], and a leashed aerial robot [23] have been proposed. Blind users can follow the system feedback to avoid static obstacles such as chairs, desks, and walls. Although these systems avoid collisions with inanimate objects, they often regard pedestrians as dynamic obstacles that will self-avoid a blind user. For this reason, they do not support pedestrian collision avoidance [144, 235]. However, pedestrians in public spaces may be unaware of blind users, especially when using mobile devices or talking to others.

One common assumption is that sighted people will notice and avoid a blind person. However, sighted people who are distracted by looking at smartphones or talking to others can easily collide with blind people. To overcome this problem, we propose two types of pedestrian-avoidance systems. BBeep is a sonic collision warning system that alerts both the blind user and nearby sighted pedestrians of a potential collision risk (Chapter 5). The second system is a guiding system that helps blind people to adapt their walking speed to avoid collisions with approaching pedestrians. The system also supports the avoidance of standing pedestrians (Chapter 6).

2.3 Hardware Design for O&M Assistance to Blind People

The performance and usability of an assistive system largely depends on the hardware design of the system. Various assistive systems for visually impaired people, including several commercial solutions, have been developed through mobile phones, wearable devices, and robots [132]. The systems presented in this thesis use smartphones, suitcase-shaped devices, and automated robots as hardware. This section reviews previous hardware designs, such as smartphones, wearable devices, and automated robots.

2.3.1 Wearable Systems

Wearable systems provide hand-free assistance to blind users. Previously proposed wearable systems include head-mounted devices [195], shoulder or chest-mounted devices [140, 141], and wristband devices. Wearable cameras can easily capture images in the first-person view of the user, but significant motion-induced blur in the captured data decreases the system's detection performance. Here we designed two suitcase-shaped systems (Chapter 5 and Chapter 6). The rolling suitcases provide storage and attachment spaces for sensors, power, actuators, and computing resources. Blind users can easily walk with the system through flat spaces and the images captured by the system are free of significant motion-induced blur.

Another limitation of wearable systems is their scarce availability to blind people, which hinders the adoption of the technology [222]. In contrast, smartphones have become a common device in the blind community and are often used in daily-life activities of blind people [160, 170, 183, 209]. To ensure accessibility, our assistive systems require only a smartphone (Chapters 4 and 7). For example, our LineChaser system using only a smartphone with a depth sensor allows blind users to stand in a line (Chapter 7).

2.3.2 Smartphone-based Systems

The expansion of smartphone usage among blind people [160, 170, 183, 209] has popularized the adoption of computer vision-based assistive applications by blind people. Commercial applications include Seeing-AI [165], Or-Cam [182], TapTapSee [49], Aipoly [9], and Envision [42]. In addition, researchers have proposed assistive systems that provide blind users with information on their surroundings (e.g., object [9, 29, 30, 42, 49, 126, 137, 149, 165, 214, 265], text [9, 29, 30, 42, 137, 149, 165, 265], face [263], and signage [214, 220]). These systems recognize information in captured images with technologies such as computer vision [9, 42, 49, 126, 165, 214, 220], crowdsourcing [30, 137, 265], and remote guides [10, 29, 149]. These systems can recognize and read printed letters and even provide simple captions to pictures taken by a blind user, but are not designed to detect surrounding people and obstacles with sufficient accuracy for avoidance and queuing behavior by blind people. In contrast, our system adopts the depth sensor of a smartphone to obtain the positions of surrounding obstacles (Chapter 4) and/or pedestrians (Chapter 7).

2.3.3 Automated Robot-based Systems

The navigation systems developed for previous research and commercial solutions assist visually impaired people through mobile phones, smart canes, wearable devices, and robots [132]. Navigation robots can potentially further increase the mobility and independence of visually impaired people by guiding them along a route toward a destination while avoiding obstacles and passersby, similarly to a guide dog [23, 24, 46, 85, 131, 161, 169, 177, 188, 213, 230, 232, 252, 258, 260]. Users can perceive and follow the directional changes in this type of system. For example, Guerreiro et al.proposed an autonomous navigation robot that guides blind users to their destination while avoiding obstacles along the path [85]. Blind users hold the handle of the robot, which actively guides and influences their trajectory. The effectiveness of this robot was investigated in a user study in which blind participants walked along pre-fixed routes with the robot. However, blind visitors exploring museums will be influenced by their interests, posing a major challenge to navigational robots [20, 21]. To meet this challenge, we developed a robot-based museum exploration system and tested it's usefulness in a real-world user study (Chapter 3). During the evaluation, blind participants freely explored a science museum during its opening hours.

Prior works evaluated their new navigation robots only on blind users. However, to assess whether an autonomous navigation robot will be accepted by wider society, the robot must be evaluated not only by blind users but also by facility managers and surrounding people. We therefore investigated the social acceptance of navigation robots in public buildings through an online survey of sighted people, interviews with facility managers, and three focus groups with blind users (Part III, Chapter 9).

2.4 Nonvisual Interfaces for Blind People

Various nonvisual interfaces have been designed and evaluated for blind people. In this thesis, we developed text-to-speech (TTS), beep sounds, sonification, spatialized audio, vibration patterns, shape-changing interfaces, and a handle for the robot. This section reviews the existing interfaces for blind people.

2.4.1 Audio Interfaces

The audio interfaces in the existing blind-assistance systems include TTS [62, 67, 116, 135, 144, 215, 233, 254], sonification [4, 50, 67, 197, 259], spatialized audios [32, 33, 50, 146, 147, 165, 166, 210, 250], and beep sounds [115, 117, 210]. Although TTS provides a range of clear instructions to blind users, these instructions should be minimized because they block the ambient sounds on which blind people often rely [36], and which may not be heard in a noisy area [25]. Therefore, TTS risks increasing the cognitive load of users [159]. Moreover, although TTS conveys a variety of instructions, it does not easily allow slight orientation adjustments [215] (e.g., "rotate 4° to the right"). Lock et al. [146] reported that bone-conducting headphones with spatialized audios effectively allow slight orientation adjustments of a blind user. Audio interfaces can provide clear, easily learned information but may disrupt the ability of blind users to hear the ambient sounds on which their safety depends.

2.4.2 Tactile Interfaces

To overcome or supplement the limitations of audio interfaces, some researchers have developed tactile interfaces that detect vibration patterns [114, 119, 135, 141, 187, 195, 211, 243, 256], shape changes interfaces [227], and thermotactile signals [123, 178]. Interfaces that detect vibration patterns can approximate the turning directions [114, 119, 135, 187, 211, 243, 256], whereas shape-changing interfaces provide refined directions. Animotus [227] is a cube-shaped interface that conveys heading directions when its top half is rotated. During an evaluation study, the route provided by Animotus was successfully followed by blind participants [228]. Audio and tactile modalities have distinct characteristics and the efficiency of each interface depends on the task and environment. Some research has shown that blind users prefer tactile-based navigation because it does not interfere with ambient sounds [159]. Other studies have shown that blind users prefer an audio interface because it minimizes the need for wearable devices [202].

2.5 Social Acceptance of Assistive Technologies for Blind People

Koelle et al.considered that "a human–machine interface can be considered socially acceptable if its presence or the user's interactions with it are consistent with the user's self-image and external image, or alter them in a positive way. Human–machine interfaces that cause a negative change to self and external image show a lack of social acceptability" [127]. Socially acceptable assistive technologies for people with disabilities must satisfy not only the functionality requirements, but also the appearance, privacy, and security demands. The social acceptance [127] of assistive technologies for visually impaired people have been variously reported. Examples are computer-vision-based assistance [3, 12], a wearable camera [11, 138, 198], and a drone [22]. Many of these studies investigated the privacy issues related to camera-based assistance [3, 11, 138, 198]. Other studies assessed users' self-perceptions and public perceptions of the assistive technology [11, 12, 22, 112, 138, 186, 222]. Studies have revealed that when the assistance technology is conspicuous, unusual, or non-mainstream in appearance, users feel deviant, stigmatized, and othered [186, 218, 266]. Although cameras and image recognition technologies are promising assistance tools for visually impaired people, they expose both users [12] and bystanders [11, 139] to privacy and security risks. The willingness of people to be recognized by an assistance technology for blind people or to share their information with such technology has also been assessed [11, 12, 22, 112, 138, 186, 222]. When people were aware of a user's disability, they felt more comfortable with being recognized by the assistance technology. Profita et al. [198] named this effect the AT-effect. Such considerations can influence the adoption and usage of assistive technologies [53].

Studies on the social acceptance of robotic assistance reported similar outcomes. Users desire discreet, inconspicuous, and safe robots [22, 91]. One interesting exception was reported by Azenkot et al. [24]. Members of their design team, some of whom were visually impaired, agreed that function was more important than form. Their robot was designed not as a personal device but as a resident in a building. Nevertheless, the robot was expected to flow with traffic, avoid disruptive or attention-seeking actions, and exhibit other socially acceptable behaviors.

The social acceptance of our navigation robots for blind people is discussed in Chapters 3 and 9. In Chapter 3, we investigate our robot-based museum exploration system in a real museum during its opening hours. Simultaneously, we investigate the social acceptance of the robot-based museum exploration system by interviewing 108 sighted visitors who saw blind participants walking with the robot in the museum. In Chapter 9, we investigate how a blind navigation robot is perceived by both blind users and sighted people. Our ultimate goal is to implement a socially acceptable robot in public buildings (e.g., shopping malls, hospitals, and museums). We therefore conducted interviews with 15 facility managers from six organizations and explored their concerns regarding robots in their buildings.

Part I

Assisting Blind People's Exploration in Public Spaces

Chapter 3

Science Museum Exploration System

"Just like sighted people who enjoy museums, I could walk around the exhibits by myself at my own pace and request an SC when I wanted a guide. It was a fun experience that I've never had."

Participant 6
Miraikan - The National Museum of Merging Science and Innovation
August 23, 2021

3.1 Introduction

Museums should be socially inclusive for all visitors regardless of their disabilities. Existing regulations (e.g., Article 30 of the United Nations Convention on the Rights of People with Disabilities [179] and the Americans with Disabilities Act, titles II and III [109]) state that all appropriate measures should be taken to ensure that people with disabilities enjoy access to cultural facilities, such as museums.

Consequently, museums are improving the accessibility of their exhibits to blind people through specialized tours [19, 94] and access to tactile representations of artworks [173]. In addition to providing such nonvisual museum experiences, increasing visitors' autonomy is also important to realize quality museum experiences for blind visitors. Researchers have revealed two main challenges in making nonvisual museum experiences accessible for the visually impaired: (1) enabling people to navigate a museum safely [18, 20, 162, 239, 240] and independently according to their interests [20, 21] and (2) representing visual artifacts nonvisually through touch or audio [18, 20, 41, 88, 238, 239, 240]. To overcome these challenges, many researchers have proposed assistive technologies for museums, such as museum navigation systems [21, 73, 74, 79, 105, 136, 162, 203] and nonvisual representation systems for exhibit objects [5, 7, 16, 203, 238].

This project focuses on the first challenge of enabling blind people to navigate and explore a science museum safely and independently and to increase their autonomy in socially inclusive ways. Choosing a series of sub-exhibitions at their own pace based on personal knowledge, interest, and comprehension speed of a science topic is an imperative part of a science museum experience. By walking around a science museum floor, blind visitors can listen to the sound at various locations, sense the size of the sub-exhibits, and feel the atmosphere of the museum [20]. Despite the effort in previous studies, it is still challenging for blind visitors to explore a science museum at their own pace, given the dependency on assistance from family, friends, or museum personnel. Blind people prefer not to rely on such assistance all the time because they are concerned about the burden on sighted assistants [20]. According to Small et al. [225], it is important for a better tourist experience to consider their various travel arrangements such as independent travel or travel with friends, family, professional attendant, professional carer, or commercial specialist.

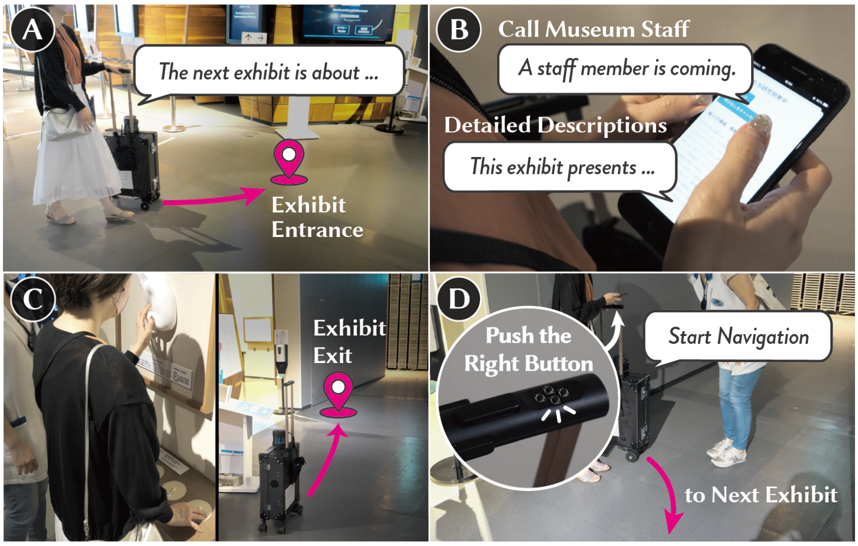


Figure 3.1: Science museum exploration with a suitcase-shaped autonomous robot to assist blind visitors. (A) The robot safely guides a blind user to an exhibit while narrating a short description of it. (B) At the exhibit, the user can listen to detailed descriptions and call museum staff if additional assistance is required. (C) While the user and staff member interact at the exhibit, the robot moves and waits. (D) The user can continue the rest of the exploration by pushing a button on the robot's handle.

We have developed a prototype of an independent museum exploration assistance for blind visitors by effectively combining the power of a navigation robot, audio guide, and the intelligence of human assistants, with the specific design target of a science museum (anonymized). Our system consists of a navigation robot based on an open-source project[1] and our newly developed smartphone app. While the robot is navigating, the user can use the smartphone app to listen to a short description of the next exhibit, as shown in Figure 3.1-A. Once the user arrives at the exhibit, they can browse detailed descriptions via the app if interested. In addition, the user can call a museum staff member if interactive support is required (Figure 3.1-B). While the user and the museum staff interact at the exhibit, the robot moves to a waiting area (Figure 3.1-C). Then, the user can command the robot to go to the next exhibit by pressing a button on the robot's handle, without needing to manipulate the app (Figure 3.1-D).

The system evaluation entailed three components: (1) a real-world user study with eight blind participants at a science museum, (2) a questionnaire on the robot's social acceptance by nearby sighted visitors, and (3) two focus groups with the blind participants. The user study was conducted during the museum's opening hours. Each participant was asked to explore one floor of the museum for 90 minutes. The floor had 10 themed exhibit areas in a space of around 2,100 m^2 (see Figure 3.2). During the study, the participants were allowed to go to any exhibit area in any order by selecting destinations according to their own pace and interests. In the user study, the robot took the participants to the exhibit entrances, and the museum staff then took turns guiding them inside the exhibit areas. Simultaneously, we asked 108 sighted visitors who observed the robot nearby to answer questions. For better technology adoption, it is important to consider social interactions between the blind visitors who use the robot and other visitors [222]. Questions include their acceptance of such a robot in a museum, as well as their safety and privacy concerns. The two focus groups [81] with the eight blind participants were conducted after the user

[1] https://github.com/CMU-cabot

study was completed. In these sessions, we discussed the participants' needs and challenges in depth to facilitate a more independent museum experience and extend our system to other use cases.

The results of the user study showed that the blind participants could explore the museum safely and independently. All participants appreciated that they could choose exhibits according to their interests and enjoy the museum at their own pace, which had been impossible when they depended on sighted assistants all the time. They also commented that it was a great experience because the robot avoided collisions with obstacles and other visitors. The questionnaire on social acceptance revealed that the sighted visitors accepted the presence of a navigation robot assisting a blind visitor at a museum. Most of these respondents said that (1) they did not feel that the robot moving with the blind visitor was disruptive or dangerous, (2) the robot's movements looked natural, and (3) they did not mind being captured by the robot's camera. In the focus groups, all participants mentioned that they enjoyed moving around alone and calling museum staff only when needed. They also expressed enthusiasm for being more independent in this situation. That is, the participants would prefer to enter exhibit areas by themselves with the robot and hear more detailed descriptions in front of each museum object, and they would prefer to call museum staff only when they want to communicate interactively. Furthermore, participants mentioned airports, shopping malls, hospitals, and other public spaces as potential areas in which the combination of a navigation robot and a human assistant may improve their experience.

3.2 Related Work

We specifically focus on assisting blind visitors' science museum exploration using an autonomous navigation robot. When blind people visit museums, one of the biggest accessibility issues is mobility and orientation. Many blind people rely on help from their families or friends [18, 240]. Previous researchers proposed various types of mobile or wearable systems to guide visually impaired people in indoor public spaces, for applications such as providing navigation instructions to a destination (Section 2.1.1) and helping people avoid obstacles (Section 2.2). In this section, we review related work specific to this chapter that were not detailed in Chapter 2.

3.2.1 Mobility Assistance

Several systems have targeted navigating visually impaired people in specific locations, especially museums [21, 105, 162]. Asakawa et al.proposed a smartphone-based system for navigating blind people in museums [21]. The system offers seamless interaction for artwork appreciation by using the user's orientation. The system's app reads the description of artwork only when the user orients toward it, and the app seamlessly resumes navigation to the next artwork when the user changes their orientation. However, the system cannot help users avoid obstacles and other visitors. Meliones et al.proposed a similar system combined with an obstacle avoidance system, and they tested it in museums [162]. Although either of these navigation systems can guide visually impaired users to their destinations, a user may veer away from a navigation path, which leads to a longer navigation time and requires a higher cognitive load for mobility.

Several researchers have tested navigation robots for visually impaired people, with the robot either in front of the person or to the side (Section 2.3.3). These robots have sensors to detect and avoid surrounding obstacles, as guide dogs do. In addition to obstacle avoidance, some robots also have functions to locate their position and navigate toward a user's destination. Visually impaired people can follow such robots by holding a handle or a leash; as a result, they do not need to pay as much attention to orientation and can walk in a more relaxed way. Nevertheless, to the best of our knowledge, no robot designed specifically for visually impaired people in museums has been studied.

CaBot is a state-of-the-art open-source blind navigation robot [85]. The robot's effectiveness has been investigated in a user study in which blind participants walked on prefixed routes with the robot. For blind visitors in museums, exploring museums according to their interests is one of the major challenges [20, 21]. Thus, we developed a smartphone app that allows blind users to control the robot's destination, listen to exhibits' descriptions, and call museum staff if additional assistance is required. We also conducted a real-world user study in which blind participants freely explored a science museum during its opening hours. Furthermore, we asked sighted visitors about their impressions of such assistive robots.

3.2.2 Exhibit Accessibility

Increasing the accessibility of exhibits is also essential to enhance the museum experience of blind visitors [88]. Especially at science museums, many exhibits use diagrams, photos, videos, interactive displays, untouchable objects, and other visual mediums to communicate scientific content. Museums have made efforts to increase accessibility by introducing objects such as tactile replicas or reproductions and audio descriptive guides [164]. Some museums provide specialized tours or workshops [19, 94]. In addition, various assistive technologies have been proposed [203, 238, 240], such as tactile models created by 3D printing [251], tactile reproductions augmented with touch sensing and audio descriptions [16], and touch screens for visual artwork exploration [7]. For the work described here, we created text descriptions of exhibits and made them available to blind visitors through our smartphone app. However, because the primary focus of this paper is mobility assistance, the exhibits in the museum for our user study, such as inaccessible interactive touch displays, were used "as is." Instead, we relied on the museum staff for help when the blind participants required such access.

3.2.3 Technologies for Communication in Museums

Communication with museum staff (e.g., curators and science communicators) is a valuable experience to increase visitors' comprehension of exhibits, especially in science museums for unfamiliar topics, academic details, or recent updates. However, visually impaired visitors have difficulties finding available staff on site. The Brooklyn Museum provides a chat application that allows visitors to communicate with museum staff online [56]. Especially for visually impaired visitors, assistance and interpretation by museum staff and volunteers are a high-priority service need [88]. Even though blind visitors require sighted assistance, they prefer not to rely on such assistance all the time, because they are concerned about the burden on sighted assistants [20]. Thus, we implemented a calling function so that blind users can ask museum staff for communication only when needed.

Additionally, robots are a promising technology to guide visitors on behalf of museum staff. Researchers have deployed a variety of autonomous robots in museums for navigating and guiding visitors. Earlier works mainly focused on safe navigation, robust localization, and advanced automation (e.g., automatic recharging) in the deployment environment [40, 181, 231]. These robots have interactive displays for guidance with multimodal content and robotic faces to attract visitors' attention. Later works mainly focused on social interaction with visitors via human-like robots used as museum guides[14, 223, 255]. For example, to attract visitors' attention, a pair of humanoid robots talked to each other about exhibits [223], and another humanoid robot observed visitors' faces to adjust its head motion [255]. Recently, robots have used advanced speech recognition and text analysis technologies to answer visitors' questions [14]. By contrast, we mainly focused on the navigation function of

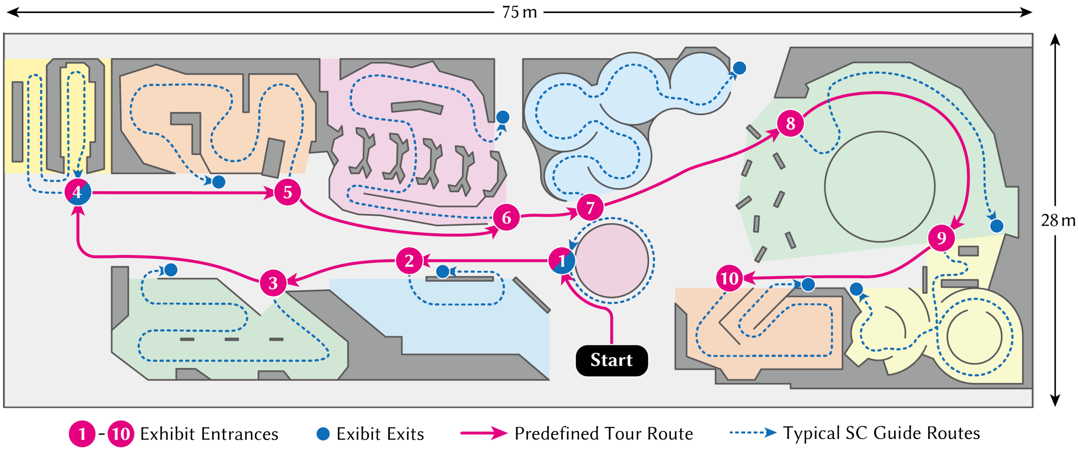


Figure 3.2: Floor map of the science museum, the predefined tour route, and the typical science communicator (SC) guide routes.

robots while relying on museum staff for visitors' questions.

3.3 System Design

3.3.1 Science Museum Experience for Blind Visitors

For blind visitors, museums are one of the most challenging places to walk and experience independently. Significant challenges lie in navigating a large space and appreciating exhibits. Previous works mostly attempted to solve those problems with mobile devices or audible icons in the environment [21, 73, 74, 79, 105, 136]. However, blind visitors have difficulty focusing on the appreciation of an exhibition while navigating the space and avoiding other visitors and obstacles [21]. Navigation with mobile robots is a promising way to reduce such stress in museum visits [85]. Therefore, in this study, we designed a user experience of science museum appreciation for blind visitors with a mobile robot and a smartphone, and we obtained feedback on the system through a user study and a focus group study.

We designed the proposed system for a science museum (anonymized) that has scientific exhibits for all ages in a multistory building. The museum staff includes science communicators (SCs) with whom visitors can talk about the exhibits. We targeted one floor of the building, with an area of about 2,100 m^2, that includes 10 of the museum's 25 themed exhibit areas, as shown in Figure 3.2. The users can command the mobile robot through a smartphone app, which will provide a brief description of the destination exhibit while navigating and also provide detailed descriptive content for the exhibition at the entrance of each exhibit. If the user wants to see more details of the exhibition, an SC can be called from the app. In each area, visitors may pass through from an entrance to an exit to see the area's artworks and interactive displays. Although making exhibit fully accessible (e.g., artwork descriptions, interactive displays, narrow paths and steps) is one of the fundamental factors in developing an accessible, independent museum experience [88, 164]. As mentioned earlier, this factor is outside the focus of this study; accordingly, we relied on the SCs to support the participants to pass through the exhibit and fill this gap. The SCs usually communicate with museum visitors to increase their comprehension of the exhibits. Before our user study, we had a training session for the SCs on interacting with visually impaired people.

3.3.2 System Overview

The museum exploration system consists of two major components: a navigational robot and a smartphone app. The robot guides a blind visitor to specified destinations while safely avoiding obstacles and other visitors. The visitor uses the smartphone app to control the robot's destination and explore the exhibits through detailed descriptions.

Museum Exploration App

The smartphone app is the main interface of the museum experience for blind visitors. The smartphone connects to the robot to control its navigational behavior and get feedback from its navigation status (e.g., avoiding obstacles). The app is designed to be fully accessible for blind visitors by working with the smartphone's screen reading software. In our user study (Section 3.4), we used the smartphone speaker so that experimenters also listened to the audio interface. We note that, if the experimenters did not require the use of the speaker, the users would be able to use their favorite hearing devices (e.g., a bone conduction headset or open ear headset) during museum exploration.

The user manages destinations by selecting a predefined tour or building a custom tour from the list of exhibits. In this study, we provided the predefined tour shown in Figure 3.2. The app also shows a detailed explanation of each exhibit as a text-based web page so that users can browse the content at their own pace using the screen reading software. The use of text-based content allows the user to change the speaking rate, which visually impaired people often want to configure according to their preferences [106]. Users can also use gestures such as flicks or multi-finger taps on the screen to move the focus and read text information.

The app is synchronized with the connected robot's navigation status. For example, the app speaks when the robot is trying to avoid other visitors. Detailed content is automatically opened when the user and the robot arrive at an exhibit entrance. If the user becomes interested in the exhibit and wants to explore it further, they can call an SC through the app for further communication. The museum's SCs also have a smartphone app to be notified promptly of visitors' requests. While supporting a visitor, an SC can command the robot to move and wait at a designated waiting area near the exhibit's exit.

Autonomous Robot for Blind Navigation

The robot is a mobile navigation robot with a handle for the user to hold and visual sensors for awareness of surrounding obstacles and other visitors to guide the user safely. The handle controls the robot's speed and navigation state and has four directional buttons to provide user shortcuts (the haptic handle in Figure 3.3). The up and down buttons control the speed. The robot's maximum speed is 1.0 m/s, but we set the default speed to 0.5 m/s for the museum setting. The right and left buttons start and stop navigation. The handle also has a touch sensor for the robot to detect whether the user is holding the handle. The robot proceeds only while the user holds the handle unless an SC commands it to move and wait.

3.3.3 Exploration Scenario

The following is a typical scenario of museum exploration for a blind visitor using the system.

1. The blind visitor borrows a navigation robot (and a smartphone if needed) at the museum's reception area. For the user study, we specified a designated start area on the floor (Figure 3.2, Start).

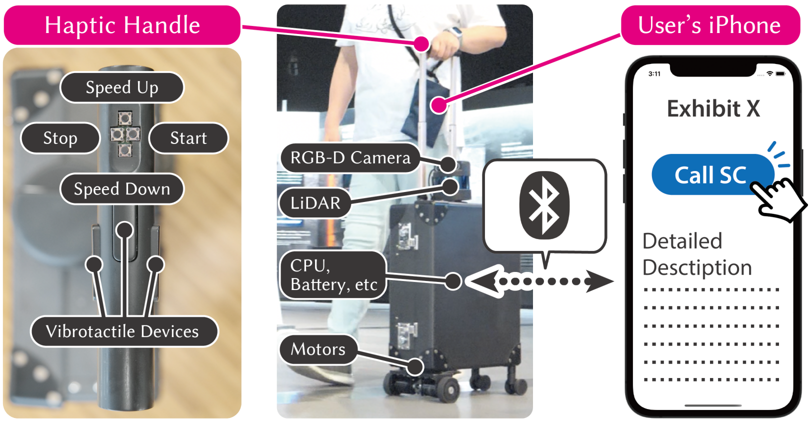


Figure 3.3: Overview of our prototype robot. The suitcase-shaped robot's handle has four directional buttons to provide user shortcuts. The robot and the smartphone in the user's shoulder pouch are connected via Bluetooth, and the user can call an SC or listen to the detailed description of an exhibit.

2. The visitor opens the app on the smartphone and selects a tour from a list of tours or an exhibit from the list.
3. The visitor holds the handle and pushes the right button on the handle to command the robot to proceed (Figure 3.1-D).
4. The robot moves toward the next exhibit in the tour while the smartphone app narrates a brief description of the exhibit (Figure 3.1-A).
5. When the robot arrives at the exhibit entrance, the app automatically pops up a browser to show the detailed content of the exhibit.
6. The visitor can browse the content on the app, call an SC for further communication about the exhibit (Figure 3.1-B), or command the robot to proceed to the next exhibit.
7. If an SC is called, the SC arrives at the visitor's location, commands the robot to move and wait at the exhibit's exit, and takes the visitor into the exhibit (Figure 3.1-C).
8. In the exhibit, the SC explains the visual features of the exhibit, navigates the visitor to touchable features, helps the visitor interact with interactive displays, and answers the visitor's questions.
9. At the exhibit's exit, the visitor can resume the rest of the tour by pushing the handle's right button (Figure 3.1-D).

Blind visitors walk with the robot while holding their white cane with their right hand, touching the robot's handle with their left hand, and listening to the description of an exhibit from the smartphone in the user's shoulder pouch. When they want to manipulate the smartphone, they can stop the robot's navigation by releasing their left hand from the robot's handle.

3.3.4 Implementation

Hersh et al. reported that visually impaired people prefer a navigation robot that is inconspicuous and discreet but attractive and elegant, and that does not draw attention to the user [92]. Accordingly, we selected a navigation robot that looks like a suitcase so that

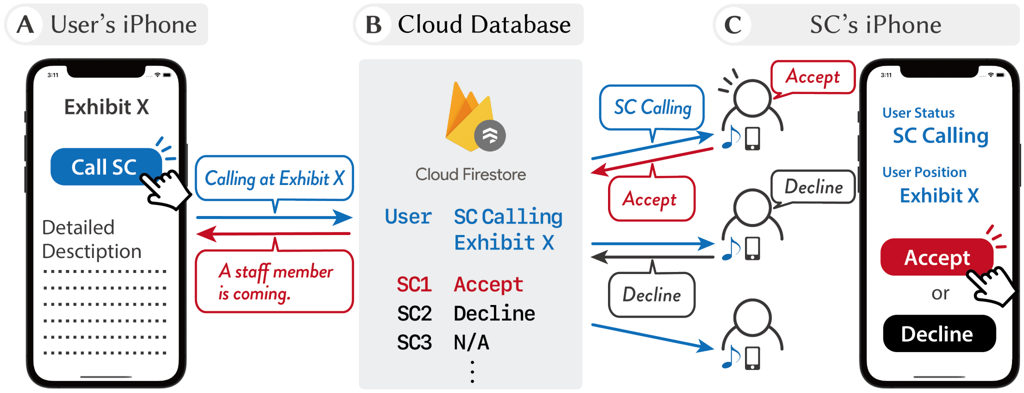


Figure 3.4: Overview of the science communicator (SC) calling system. (A) The user can call an SC or listen to the detailed description of an exhibit. (B) The user's current status is maintained in a cloud database. (C) When the user calls an SC, the SCs' smartphones receive a notification, and each SC can accept or decline the call.

it can assimilate into the environment, as shown in Figure 3.3. The robot's hardware and software were based on an open-source project[2]. The robot has a handle for the user to hold. The suitcase portion has sufficient space for hardware resources, such as the battery and CPU. The robot also has a LiDAR scanner, a 3D laser range finder to accurately estimate its position by using Cartographer ROS[3] with Bluetooth Low Energy beacons placed in the museum, and a stereo camera to detect people in front of it by using the YOLOv4 [34] image recognition engine.

We developed an iOS app to connect with the robot via Bluetooth by using the SwiftUI[4] and Google Cloud Firestore[5] frameworks. The app was installed on an iPhone 12 Pro (Figure 3.3). We customized the robot app to enable synchronization between the app and robot. The short descriptions for narration during navigation and the detailed content were edited by the museum's SCs. By using a database on Google Cloud Firestore, the app can manage user status information, including whether the user is calling an SC and where they are navigating (Figure 3.4-B). VoiceOver, the built-in screen reader on iOS, is activated to enable the users to control the robot and browse content on the app by themselves.

3.4 User Study and Focus Groups

To evaluate the effectiveness of the proposed museum exploration system and its acceptance by museum visitors, we conducted a real-world user study at the science museum during its opening hours. We asked eight blind participants to freely explore and experience one floor of the museum for 90 minutes by using our system. During the study, we also asked sighted visitors to complete a short questionnaire about the robot's social acceptance. Finally, after the user study, we conducted two focus group sessions with the blind participants [81]. In these sessions, we discussed their in-depth needs and issues for developing a more independent museum experience.

3.4.1 Participants

We recruited blind participants via an e-newsletter to which 80 visually impaired people had subscribed. The conditions were as follows: totally or legally blind people, between 20

[2]https://github.com/CMU-cabot
[3]https://github.com/cartographer-project/cartographer_ros
[4]https://developer.apple.com/xcode/swiftui/
[5]https://cloud.google.com/firestore/

Table 3.1: Demographic information of the user study participants.

ID	Age	Gender	Eyesight	Museum Visits
P3.1	49	Male	Blind since age 14	1–2 times/year
P3.2	25	Male	Blind since age 4	2–3 times/year
P3.3	35	Female	Blind since age 14	A few times
P3.4	29	Male	Blind since age 10	A few times
P3.5	51	Male	Blind since age 8	A few times
P3.6	29	Male	Blind since age 5	Once every 2–3 years
P3.7	53	Male	Blind since birth	A few times
P3.8	43	Female	Blind since age 3	A few times

and 70 years of age, able to manipulate an iPhone using VoiceOver, able to go to the meeting place of the experiment (a station ticket gate) by themselves, and $70 compensation. We recruited the first eight blind participants to arrive (six men, two women) with ages ranging from 25 to 53 years (mean 39.25, SD 10.41), as listed in Table 3.1. All participants were totally blind and primarily used a cane. As seen in Table 3.1, three participants (P3.1, P3.2, and P3.6) visited museums once in a while, but the others had only visited a museum a few times in their life.

3.4.2 Procedure and Metrics

Pre-Interview and Training Session

After obtaining an Institutional Review Board-approved (the Ethics Review Committee on Research with Human Subjects of Waseda University, 2020-039) informed consent from the participants, we conducted a pre-interview session of 10–15 minutes in which we asked about their experiences and challenges in museums. Then, we provided them approximately 20 minutes of training to familiarize them with the system. Specifically, they practiced walking with the robot, setting destinations with the smartphone app, and calling an SC.

Main Session

After the training session, the participant moved to the starting position at the floor's entrance (Figure 3.2). We told the participant, *"Please explore and experience the museum freely for 90 minutes with the robot."*[6] We did not specify which destinations to choose (the predefined tour and specific exhibits); instead, the participants chose the destinations according to their strategy and interests. The participants were informed that a researcher would walk behind them (5–10 m away) to assist them immediately if they required support. We also informed the participants that they could take a break whenever they liked.

Post-Interview Session

After the main session, we conducted a post-interview session, which took approximately 30 minutes. We first asked the participants to answer a set of questions (Table 3.3, Q3.1–Q3.6) consisting of items answered on a 7-point Likert scale (1: strongly disagree; 4: neutral; 7: strongly agree). Then, we asked open-ended questions about the advantages and

[6] All communication with the participants was in their native language. In this paper, we present any translated content in the form of *"translated content."*

issues of our system, their strategy for exploring the museum with it, and suggestions for improvement.

3.4.3 Questionnaire on Robot's Social Acceptance by Sighted Visitors

For this part of the study, we asked sighted visitors who saw the participants walking with the robot to rate five sentences on a 7-point Likert scale. We gave a ballpoint pen with the museum logo to the sighted visitors who answered our questionnaire.

3.4.4 Focus Groups

After finishing the user study, we organized two online focus group sessions with four participants each. The sessions were semi-structured to focus on further requirements for our museum exploration system. Specifically, we first asked the participants for suggestions to improve their museum experience, namely, *Can you recommend any new functions to improve our museum exploration system?"*. We further inquired about possible solutions to comprehend the exhibits' contents while comparing with our current system, which relies on SCs to guide blind participants in an exhibit. Finally, we asked the participants, *"In what other scenarios would you like to use our robot-based exploration system?"* The two focus groups covered the same topics, and each took approximately 60 minutes. Each session was audio-recorded and transcribed for further analysis.

3.5 Results

3.5.1 Overview of Exploration Activity

Table 3.2 lists the orders of the exhibits visited by each participant. All participants started their museum exploration by following the predefined tour. Six participants (P3.1–P3.6) first completed the predefined tour without calling an SC, while two participants (P3.7 and P3.8) occasionally called an SC during the tour. After arriving at Exhibit 10 and finishing the predefined tour, six participants (P3.2–P3.4 and P3.6–P3.8) visited some of the exhibits again by repeatedly navigating to a specific exhibit from the exhibits list and calling an SC. The other two participants (P3.1 and P3.5) followed the predefined tour again and then called an SC when they arrived at an exhibit of interest.

Table 3.2 also summarizes the activity duration times, including A) the activities of walking with the robot, B) using the smartphone, and A+B) spent alone without SCs' supports. During their 90-min sessions, the participants walked with the robot for about 9 min on average (Table 3.2-A). While we set 0.5 m/s for the default robot speed, all participants changed the speed (P3.1: 0.8 m/s, P3.2–P3.6: 1.0 m/s, P3.7: 0.9 m/s, and P3.8: 0.75 m/s). Three participants (P3.3, P3.4, and P3.8) did not hold their cane and explored the museum floor only using the robot, and the other five participants used the robot while having their cane in their right hand The walking style of the participants also differed (Table 3.2-Walking Style). The participants spent approximately 17 min on average operating the app, including selecting destinations, browsing descriptions, and waiting for SCs (Table 3.2-B). They all called SCs 5 or 6 times and spent about 1 hour with SCs. During the session, 2–4 SCs helped support this study. On average, the participants spent 26 min (about 30% of their time exploring the museum) alone.

As an example, Figure 3.5 shows the routes of P3.3 and the robot. P3.3 first followed the predefined tour and explored the whole floor. Then, she visited five exhibits (Exhibits 4, 6, 7, 3, and 1) and called an SC at each exhibit, because she wanted to learn more and ask questions about these exhibits. We describe the participants' comments on their museum exploration strategies later, in Section 3.5.4.

Table 3.2: Order of exhibits visited by the participants in the study (* indicates an exhibit where the participant called an SC), together with their walking style (with the robot and their cane or with the robot only), the durations of A: walking with the robot, B: using the smartphone, and A+B: the time spent alone.

The order of exhibits they visited (*: Called an SC)	Walking Style	A) With Robot	B) Using Smartphone	A + B
P3.1 Tour (1–10), Tour (**1***, 2, **3***, **4***, 5, **6***, 7, **8***, **9***)	Cane & Robot	08:59	19:48	28:47
P3.2 Tour (1–10), **4***, **6***, **7***, **3***, **1***	Cane & Robot	11:53	17:31	29:24
P3.3 Tour (1–10),**9***, **3***, **1***, **8***, **4***, **5***	Robot only	08:25	16:25	24:50
P3.4 Tour (1–10), **3***, **5***, **8***, **9***, **10***, **6***	Robot only	08:31	11:11	19:42
P3.5 Tour (1–10), Tour (1, 2, **3***, 4, 5, **6***, **7***, 8, 9, **10***), **5***	Cane & Robot	10:27	13:43	24:10
P3.6 Tour (1–10), **1***, **3***, **4***, **8***, **9***, **10***, **6***	Cane & Robot	07:41	28:24	36:05
P3.7 Tour (**1***, 2–5, **6***, 7–10), **3***, **7***, **8***, **4***	Cane & Robot	08:38	14:23	23:01
P3.8 Tour (1, 2, **3***, **4***, 5, **6***, 7–10), **5***, **7***, **10***	Robot only	07:25	16:15	23:40
	Average	9:00	17:13	26:13

3.5.2 Subjective Ratings

Table 3.3 summarizes the results for the six Likert-scale questions (Q3.1–Q3.6). All the participants agreed (score greater than 5) that they enjoyed experiencing the museum with the robot (Q3.1); that they could explore the museum independently at their own pace (Q3.2); that they did not feel any danger while walking with the robot (Q3.3); and that calling the museum staff was effective (Q3.6). For usability (Q3.4) and the effectiveness of the exhibits' short descriptions (Q3.5), all the participants except P3.6 gave positive scores. While we gave participants only 20 minutes as the training session, in the main session, all participants were able to operate the system with little or no assistance from us.

3.5.3 Social Acceptance of the Robot

Through the four days in which the study was conducted (two participants per day), an average of 272 people per day visited the museum (SD 52.7)[7]. We obtained questionnaire responses from 108 visitors in total, for an average of 13.5 visitors per blind participant. The age distribution of the sighted visitors was as follows: teens and younger: 22 (20.4%); 20–29: 21 (19.4%); 30–39: 30 (27.8%); 40–49: 27 (25%); 50–59: 4 (3.7%); 60–69: 2 (1.9%); 70 and older: 2 (1.9%). Figure 3.6 shows the questionnaire results (Q3.7–Q3.11). For all the questions, the value of the first quartile was more than 5 points, and we observed that more than 75% of the visitors accepted the user and the robot. The details of the results are as

[7]Before the COVID-19 outbreak, around 4,000–5,000 people per day typically visited the museum.

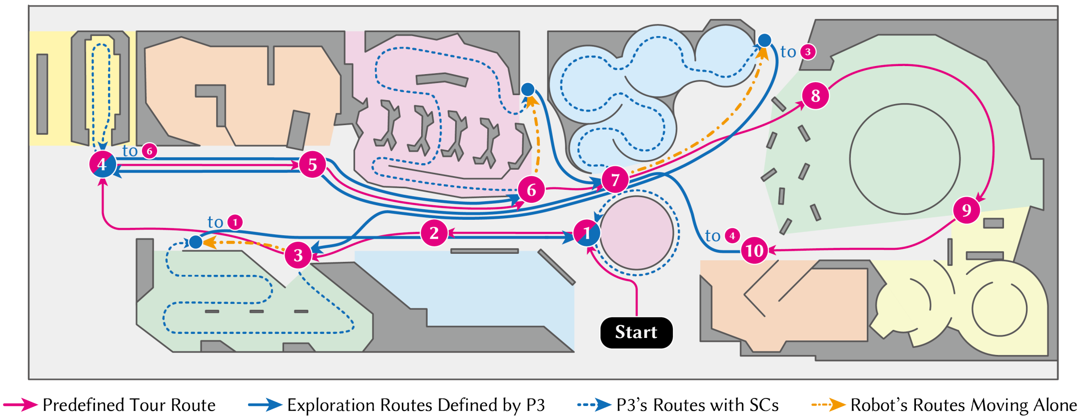


Figure 3.5: Example showing the routes of P3.3 and the robot in the user study. P3.3 first followed the predefined tour and explored the entire floor (exhibits 1 to 10). Then, P3.3 visited five exhibits again and called a science communicator (SC) at each one (Exhibits 4, 6, 7, 3, and 1).

Table 3.3: Summary of Likert-scale responses (1: strongly disagree; 4: neutral; 7: strongly agree). M: Median.

Question	P3.1	P3.2	P3.3	P3.4	P3.5	P3.6	P3.7	P3.8	M
Q3.1: I enjoyed exploring the museum with the robot.	7	7	7	7	6	7	7	7	7
Q3.2: I could explore the museum independently at my own pace.	7	7	7	7	6	7	7	7	7
Q3.3: I did not feel any danger while walking with the robot.	6	6	7	7	7	5	7	7	7
Q3.4: The system was easy to use.	5	7	7	7	7	3	7	6	7
Q3.5: The narration of the exhibits' short descriptions was effective.	7	5	7	7	7	2	7	7	7
Q3.6: Calling the museum staff was effective.	7	7	7	7	7	7	7	7	7

follows: 99.1% agreed that the robot for blind visitors should be introduced in museums (Q3.7); 78.7% felt that the movements of the blind visitors and the robot were natural (Q3.8); 86.1% did not feel that the blind visitors and the robot were disruptive (Q3.9); 88.9% did not feel any danger from the blind visitors and the robot (Q3.10); and 78.7% accepted the robot's camera capturing them (Q3.11).

3.5.4 Qualitative Feedback

This section summarizes the participants' comments, both positive and negative, from the post-interview session.

Overall Experience

All participants appreciated that our system enabled them to explore the exhibits independently at their own pace.

> **A3.1:** *"Just like sighted people who enjoy museums, I could walk around the exhibits by myself at my own pace and request an SC when I wanted a guide. It was a fun experience that I've never had."* P3.6

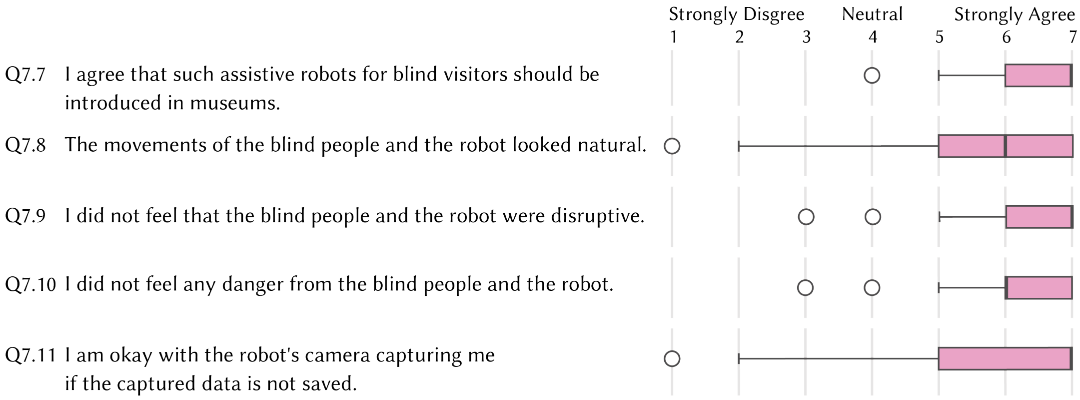


Figure 3.6: Questionnaire results of the social acceptance of the robot.

> **A3.2:** *"I could go around the exhibits at my own pace in my preferred order. When I go with a friend, I do not want to spend a long time on exhibits that my friend is not interested in, even if I want to go. With this robot, I could go around my favorite exhibits as much as I wanted."* P3.8

> **A3.3:** *"When I ask my friends or museum staff to guide me, I need to be with them all the time while walking. I don't want to bother them by asking them to wait for a long time. Using this system, I did not have to worry about that, and I could freely walk around the exhibits I am interested in as much as I wanted."* P3.4

We conducted the user study during the museum's opening hours. Although other visitors constantly came and went on the floor (272 people per day on average during the study), five participants (P3.3–P3.5, P3.7, and P3.8) did not feel any danger while walking with the robot.

> **A3.4:** *"(When I used the robot,) I did not have any stress while moving. I could focus on the voice summary of the next exhibit and listen to sounds from the exhibits. I could get into the atmosphere of the exhibits."* P3.3

> **A3.5:** *"Even in the environment with people, the system properly stopped or avoided people, so I could walk with confidence."* P3.7

Museum Exploration Strategies

The orders of the exhibits that the participants visited varied among them. All the participants started their museum experience by following the predefined tour to explore the whole floor. After completing the predefined tour, six participants (P3.2–P3.4 and P3.6–P3.8) went to specific exhibits according to their interests, while two participants (P3.1 and P3.5) followed the predefined tour again.

> **A3.6:** *"At first, I explored the floor by following the tour. It was good because I could grasp the rough structure and size of the floor by myself while walking."* P3.5

> **A3.7:** *"After grasping the whole structure of the floor by following the tour, I went to the places that I was interested in and that had high priority for me one by one. It was good to grasp the whole structure of the floor not only by listening to the voice guidance for the exhibit list but also by listening to the sounds from exhibits and feeling the atmosphere while walking."* P3.2

> **A3.8:** *"At first, I went around the floor and grasped what I was interested in. Because I could not remember the exhibits only by listening to their titles, I followed the tour again [by not selecting a destination from the list] and remembered the exhibit contents by listening to the summary while walking. Then, I requested an SC at the exhibits I was interested in."* P3.5

Narration of Exhibits' Short Descriptions

Three participants (P3.1, P3.4, and P3.7) gave positive comments on the function of narrating the exhibits' short descriptions.

> **A3.9:** *"The function of reading the short descriptions was good because I could figure out the next exhibit."* P3.7

> **A3.10:** *"(When I walked with the robot,) the feeling was close to being guided by a person. When I moved, I trusted the robot and could walk while listening to the short descriptions and surrounding sound and thinking about the next exhibits."* P3.1

On the other hand, P3.6 commented that the narration function was not effective.

> **A3.11:** *"I was not used to the robot yet, and it was difficult to concentrate on listening to the short descriptions."* P3.6

Calling an SC

As seen for Q3.6 in Table 3.3, all the participants greatly appreciated that the system could call museum staff when needed.

> **A3.12:** *"The function to call an SC is valuable. Because I cannot find help in large spaces such as museums, it was nice to call for help when I wanted to."* P3.2

> **A3.13:** *"I did not want to bother the SCs by asking them to guide me all the way. The function to call an SC only when needed was good."* P3.8

> **A3.14:** *"Compared to walking with a staff member all the way, I'm glad I could talk to various SCs."* P3.3

System Usability

Although seven participants rated the system as easy to use (Table 3.3, Q3.4), some participants provided suggestions for improving the user interface.

> **A3.15:** *"When the robot stopped, I could not understand whether it had stopped because there was a person in front of us or I had held the handle incorrectly. I would like to know why the robot stops when it stops."* P3.6

3.5.5 Focus Groups

This section summarizes the participants' comments in the two online focus group sessions (Group A: P3.2, and P3.5–P3.7; Group B: P3.1, P3.3, P3.4, and P3.8).

Toward a More Independent Museum Experience

When we asked the participants for suggestions to improve their museum experience, all the participants commented that they wanted to listen to audio guidance while walking through an exhibit area.

> **A3.16:** *"Because there were exhibits that played sound, I wanted to be close to them and listen to descriptions while also listening to the sound from the exhibits."* P3.3

> **A3.17:** *"Among the exhibit areas that the SC introduced, there were some exhibits where I could experience the sizes of museum objects by walking around them (e.g., anonymized). Rather than just listening to guidance in front of the entrance, it would be nice if I could listen to the descriptions while walking inside with the robot and experiencing the size."* P3.2

> **A3.18:** *"(Compared with just listening to voice guidance at the entrance to an exhibit,) if the voice guidance told what is displayed in each location while walking with the robot in an exhibit, it would be nice because I could understand what kinds of items are where in the exhibit."* P3.7

In the proposed system, we rely on SCs to guide blind participants in an exhibit. Five participants (P3.1 and P3.5–P3.8) wanted to call SCs to have them discuss the contents of an exhibit rather than explain the exhibit. In addition, six participants (P3.1, P3.3–P3.6, and P3.8) commented that they wanted to comprehend the exhibits' contents more independently.

A3.19: *"I called an SC every time I wanted to enjoy an exhibit in this experiment, but I would prefer to call an SC only when I have questions after I understand an exhibit as much as I can with the robot."* P3.1

A3.20: *"If the robot explains an exhibit, I would prefer to call an SC to discuss the exhibit with them rather than ask them to explain it."* P3.7

A3.21: *"Rather than listening to predefined voice guidance, it would be nice if I could talk with the staff and ask questions about exhibits through video calls, for example."* P3.5

A3.22: *"Because I come up with questions as I walk around exhibits, it would be nice if the AI [Artificial Intelligence] could respond. If the system explanation is only one-sided (like the current system), I would prefer to call an SC and enjoy exhibits while asking questions."* P3.6

Prospective Scenarios for a Museum Exploration System with a Robot

When asked about scenarios in which the participants would want to use our exploration system, they suggested shopping malls (P3.1–P3.8), airports (P3.1, P3.3–P3.5, and P3.7), train stations (P3.4–P3.6), universities (P3.1, P3.2, and P3.6), hotels (P3.2 and P3.7), hospitals (P3.1 and P3.6), and amusement parks (P3.3 and P3.4).

A3.23: *"In airports, I ask staff to take me to my gate, and I wait there the whole time. I would like to go to the bathroom or explore shops and restaurants with a robot at any time while waiting."* P3.7

A3.24: *"For the use case in a shopping mall, the robot would be useful because I could go to shops that I am interested in, just as I walked around the exhibits in the museum. I could also call a clerk in a shop when I have a little difficulty, just as I called an SC when I wanted an explanation."* P3.2

A3.25: *"Theme parks are fun not only for riding the attractions that I plan to ride but also for exploring while feeling the atmosphere. I would like to take time exploring alone by using the robot."* P3.4

3.6 Discussion

3.6.1 Independent Museum Experience

The proposed system successfully enabled the blind participants to explore the science museum by effectively combining the power of a navigation robot and the intelligence of human assistants. All the participants agreed that they enjoyed exploring with the robot (Table 3.3, Q3.1) and could explore the museum independently at their own pace (Table 3.3, Q3.2). As seen in Table 3.2, the participants chose to visit a variety of exhibits according to their own interests and strategies (A3.6–A3.8). In the user study, participants interacted with the system for about 26 minutes on average, meaning the system-human ratio in time was roughly 30%–70%. (Table 3.2). Nevertheless, the participants appreciated the independent museum experience in the user study because such experiences have not been possible for them when they visited a museum alone or with their families and friends (A3.1–A3.3).

The participants rated the predefined tour highly because they could grasp the rough structure and size of the floor by walking through it rather than just listening to a long description of it (A3.6 and A3.7). The narration of the short description for each exhibit during navigation was effective for independent museum exploration (Table 3.3, Q3.5). The participants could gain an overview of the upcoming exhibit while walking (A3.9 and A3.10).

The participants unanimously agreed on the effectiveness of calling a museum staff member (SC) via the smartphone app (Table 3.3, Q3.6). They commented on their need to ask the museum staff for support, but at the same time, they typically hesitate to take the staff's time for an entire visit, or they feel uncomfortable being accompanied by a human assistant. In contrast, our system enabled the blind visitors to call staff for support only when they needed it. They felt comfortable asking for support from multiple assistants, and they enjoyed communicating with multiple assistants. (A3.12–A3.14).

3.6.2 Safety Concerns

Maintaining a sense of safety for both blind and sighted visitors is one of the critical challenges in deploying navigational robots in a real-world museum setting. The user study was conducted during the museum's opening hours when other visitors constantly came and went on the floor. Even in such a real-world situation, the participants could walk with the robot through the museum floor without any incidents or safety concerns (Table 3.3, Q3.3). They could rely on the robot's navigation and focus on the museum experience (A3.4 and A3.5). Three participants especially explored the museum without having their cane (Table 3.2-Walking Style). In addition, 86.1% of the sighted questionnaire respondents who saw the participants and the robot also did not feel any danger from their movements (5–7 points in Figure 3.6, Q3.10). We cannot generalize these results to other situations without more data, especially for the case when the museum is crowded. Nevertheless, both the blind and sighted visitors did not have any safety issues or concerns in the study environment. This is an encouraging result that can provide a beachhead for designing our next exploration system for real-world deployment.

3.6.3 Social Acceptance

In the user study, the surrounding sighted visitors accepted the navigation robot well (Figure 3.6): 99.1% of them agreed that assistive robots for blind visitors should be introduced in museums (5–7 points in Figure 3.6, Q3.7), and the percentage of visitors who felt that the blind people and the robot were disruptive was only 2.8% (1–3 in Figure 3.6, Q3.9). Researchers have previously reported on blind users' considerations of their image and the public perception of assistive technologies [11, 12, 22, 112, 138, 186, 222], but in this study, 78.7% of the sighted visitors regarded the presence of a blind user and robot as natural (5–7 in Figure 3.6, Q3.8).

Privacy concerns about camera-based technologies have been discussed in previous studies [3, 11, 138, 198]. We mounted the camera on top of the robot and expected that this feature might invoke privacy concerns. However, we found that the sighted visitors generally (78.7%) accepted the robot's camera capturing them (5–7 in Figure 3.6, Q3.11). Privacy concerns are usually considered the most serious challenge for the practical deployment of camera-based assistive technologies. Instead, we found that the use of a camera may not be a deal-breaker for social acceptance and practical deployment in science museums if the robot works to assist blind people.

3.6.4 System Usability

Although seven of the eight participants rated the system as easy to use (Table 3.3, Q3.4), we also found opportunities to improve the user interface. P3.6 negatively rated the system's usability (Table 3.3, Q3.4) and the effectiveness of the exhibits' short descriptions (Table 3.3, Q3.5). His point concerned the transparency of the robot's actions. For example, he commented that he wanted to be aware of the reason for each "stop" whenever the robot stopped (A3.15). We could have the robot explain the reason for each stop, but it might overwhelm users with too much information. We thus need to carefully balance the information presented. In addition, P6 commented that he could not pay attention to the short descriptions because he had to pay attention to the robot (A3.11). This might improve with more usage of the system. In this study, only P6 gave negative ratings, but we should expect more diverse feedback, given the variety of skills and experiences among the blind population. Accordingly, we should carefully design and evaluate interface options and personalization features to meet each user's preferences (e.g., walking and smartphone skills).

3.6.5 Toward a More Independent Museum Experience

The proposed system successfully improved the independence of participants by focusing on navigation and exploration among sub-exhibitions (A3.1–A3.3). We found a strong preference to be independent. Most participants commented that they would like to spend more time without being dependent on SCs (A3.19–A3.22). Six participants commented that they would like to understand each exhibit's content as much as they can with the robot and then call an SC only when they have questions (A3.19). All the participants wanted robot navigation "inside" each sub-exhibition, rather than just listening to guidance at the entrance as was the case in this study. The following system should support finer navigation among panels/exhibits inside each sub-exhibition. Each sub-exhibition consists of several panels/exhibits every 1 to 2 m, so the robot should navigate a user among panels/exhibits along a typical route on a finer scale.

Beyond this finer navigation and explanation, blind visitors strongly prefer independence in their science museum experience, but enabling nonvisual science communication is the challenge. One possible future direction is the use of automated or remote question-answering (Q&A) technologies. Five participants agreed that science communication was a valuable experience in the museum because it enabled them to ask questions (A3.19–A3.22). They also commented that they preferred even this question-answer part to be automated; for example, this could be done with a remote assistance system [10, 29], a chat system with museum staff [56], or an AI-based Q&A system [14, 167, 237] on the robot or the app.

At this moment, the appropriate balance between independent exploration with an assistive robot and human assistance for science communication is not clear. Blind visitors require human assistance, but they also would prefer an independent experience, as we observed in the focus group sessions. We hope to improve our system to navigate inside sub-exhibitions and integrate automated or remote Q&A technologies.

3.6.6 Prospective Scenarios for a Science Museum Exploration System with a Robot

The proposed system allows blind users to explore a science museum floor freely at their own pace and according to their own interests, and they can call a human assistant when needed. In the focus group sessions, the participants commented that the system has potential usage not only in museums but also in various other places, such as airports, shopping malls, hospitals, and amusement parks (A3.23–A3.25). The participants hoped to be able to explore large areas when finding a favorite restaurant, for example, and feel the surrounding atmosphere, while asking for interactive human support only when necessary, as sighted people usually do. In our future work, we want to expand the coverage of our robot and exploration system to offer new independent experiences in various situations.

Acknowledgement

We would like to thank all participants who took part in our user study. We would also like to thank the anonymous reviewers for their helpful comments. This work was supported by JST-Mirai Program (JPMJMI19B2), JSPS KAKENHI (JP3.20J23018), and Consortium for Advanced Assistive Mobility Platform.

Chapter 4

Smartphone-based Environment Recognition System (Collaboration Works)

4.1 One-shot Wayfinding System[1]

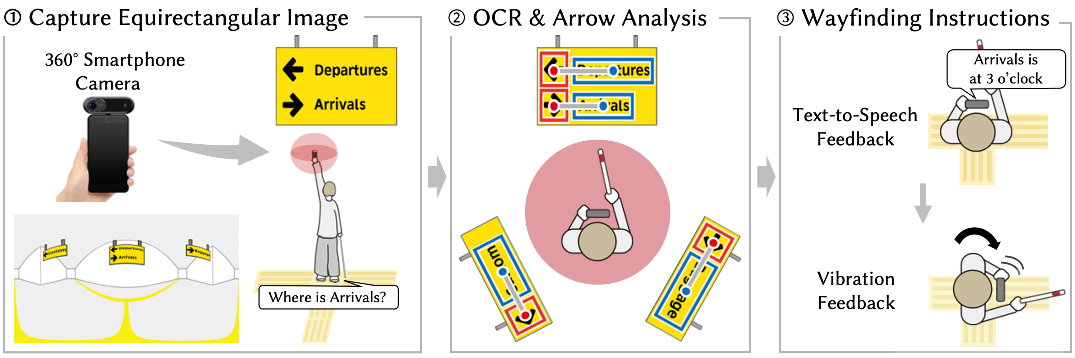


Figure 4.1: Overview of our *one-shot wayfinding* system for blind people. 1) When approaching an intersection in a public building, a blind user takes a picture with a 360° camera attached to a smartphone. 2) The system detects the text and arrows on surrounding signage and links them to estimate the destination direction. 3) The system provides wayfinding instructions (TTS and vibration feedback) for the estimated direction.

Many public buildings provide signage showing the directions toward points of interest [66, 86]. Such signage can orient sighted visitors but offers little benefit to blind people. Recent studies have proposed assistive technologies that recognize information on signage (e.g., text or pictograms) by combining a smartphone camera with computer vision technologies such as optical character recognition (OCR) [9, 42, 165, 214]. However, these signage recognition systems require the fast and accurate capture of pictures with the appropriate framing, which is a difficult task for blind people [1, 107, 154, 264]. Thus, blind users cannot always obtain the required information from these systems. Moreover, whether failure to capture the correct information is caused by lack of signage or incorrect camera framing is difficult to discern.

To overcome this limitation, we developed a *one-shot wayfinding* system with a 360° smartphone camera that captures all signage around a user in a single shot. The user does not need to adjust the camera aim. To decide the correct direction to a destination, blind people must understand the directions of the arrows on signs. For example, Figure 4.1 shows a sign with a right arrow indicating a right turn at the intersection. Therefore, our wayfinding system recognizes not only text but also the arrows pointing the way to a destination. The system recognizes text, arrows, and text–arrow associations on surrounding signage and converts the direction of each arrow into an egocentric direction (i.e. a direction relative

[1]Project Page: https://yutaroyamanaka.netlify.app/publication/oneshot/

to the user's body). It then verbalizes the egocentric direction as a clock position, which is a standard way of presenting directions to blind people. In other words, our wayfinding system for blind people requires no environmental modifications such as markers, BLE beacons, or Wi-Fi beacons. Nor does it require the preparation of data such as maps and points-of-interest datasets.

The proposed system first detects the text and arrows from a captured equirectangular image [226] using an OCR system and a convolutional neural network object detector. It then links the detected text to each detected arrow via a minimum spanning tree. When setting the edge weights linking the text and arrows, we considered the relationships between text and arrows (e.g., the text above an arrow typically shares a weaker correspondence with the arrow than the text below the arrow). In Figure 4.1 (2), the system links "Departures" and "Arrivals" to the left and right arrows, respectively. It then estimates the directions to the destinations in egocentric coordinates relative to the user's current orientation. For instance, when the user stands beside an "Arrivals" sign linked to a right arrow, the system utters *"'Arrivals' is at 3 o'clock."* To further convey the estimated direction, the system gives vibration alerts when the user faces the correct direction [Figure 4.1 (3)].

The usability of the system's interface was evaluated through a user study involving eight blind people. As a baseline system, we implemented a simple signage reader system that uses an RGB camera built into a smartphone (not a 360° camera). We asked the participants to find the correct direction to a destination using either the proposed system or the baseline system. Participants determined the correct direction after fewer rotations of the proposed system than of the baseline system. The participants' feedback also supported our hypothesis that the proposed system can assist wayfinding tasks in public buildings. Based on our findings, we discussed the development of a more flexible and comfortable wayfinding system for future use in public buildings.

4.2 Corridor-Walker[2]

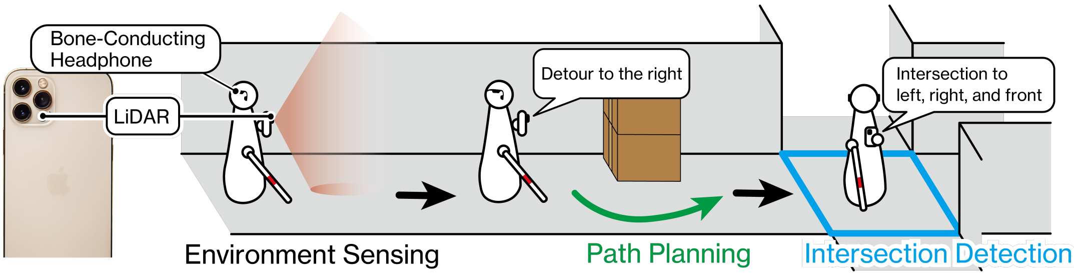


Figure 4.2: Corridor-Walker helps blind people to recognize obstacles and intersections. Using the system, a blind person detects an upcoming intersection and recognizes the paths leading from the intersection while avoiding obstacles.

Navigation through indoor corridors is a challenging task for blind people. In such environments, blind people usually navigate using the surrounding walls [13]. During this process, they can encounter various obstacles such as wall-mounted furniture and objects [108]. Collisions with these obstacles can damage the objects and/or injure the person. Ground obstacles are detected by white canes, which are commonly carried by blind people. In addition to white canes, users can employ various systems that alert them to the existence of obstacles at or above ground level [50, 187, 197]. As these systems detect only the existence of such objects, the user must determine the path that avoids the obstacle (obstacle-avoiding path). Although guide dogs can help users to navigate an obstacle-avoiding path, they are not preferred by all blind people because they require caretaking [206, 246]. Another drawback is the small number of guide dogs. For example, only 5,000 dogs versus 360,000 legally blind people, comprising approximately about 1.4% of the United Kingdom population [206]). To alleviate this situation, researchers have developed assistive technologies that navigate blind people along an obstacle-avoiding path using mobile robots [85, 117, 150] and wearable devices [141, 144, 243]. However, these solutions use special hardware which is not commonly available to blind people, thus hindering the technology adoption [222].

Besides avoiding obstacles, blind people navigating through a corridor must know the corridor's geometric structure [61, 219], , such as intersections. Walking past an unnoticed intersection or turning into an incorrect intersection can disrupt the navigation path. Correct navigation requires reliable position and shape information of each intersection in the corridor. A white cane alone might be insufficient for locating an intersection, resulting in walking past an unnoticed intersection [84]. In addition, the white cane does not fully support shape recognition of intersections because it has a limited range of contact. Guide dogs can help blind people to locate an intersection [84] but cannot convey its shape to the user. In contrast, indoor turn-by-turn navigation systems convey the correct information on intersections [62, 215], allowing blind users to reach their destinations without becoming lost. However, such systems require static route maps or additional infrastructure, which are unlikely to be available in every building [60, 229].

We present *Corridor-Walker* (Figure 4.2), a mobile indoor walking assistance system that supports obstacle avoidance for blind people and recognizes intersections (i.e., locates and grasp the paths leading from the intersections). The system is aimed for use in the indoor corridors of apartments, offices, or hospitals, where static route maps and infrastructure are unavailable but in which the user knows the turns required to reach the destination. Such users may be familiar with the environment from prior travel or gain knowledge from

[2]Project Page: https://sites.google.com/view/corridor-walkerproject/

tactile maps [83, 194] or interactive devices [97]. Many blind people already use smartphones [160, 170] for messaging and obtaining assistance systems available in app stores, which recognize items, read printed letters, and guide navigation along a recorded route (e.g., , Seeing AI [165], Tap Tap See [49], and Clew [259]). Therefore, our system was designed for compatibility with a single smartphone to improve the technology adoption. The system provides obstacle avoidance by navigating the user along an obstacle-avoiding path using both spatialized audio and TTS feedback. The system alerts the user to the presence of an intersection and informs the user of its shape through vibration and TTS feedback. To achieve these functionalities, the system first constructs a 2D occupancy grid map [55, 140, 141, 195] of the surrounding environment using a LiDAR sensor equipped with an iPhone 12 Pro [17], which supports accurate grid-map construction. The system then plans an obstacle-avoiding path on the grid map using the A* path-planning algorithm [89]. Simultaneously, it detects upcoming intersections using the "you only look once (YOLO) v3" detector [205]. As the system uses only real-time sensing results, it accomplishes these functionalities without requiring a static route map or additional infrastructure.

The usability of our system was evaluated through a user study involving 14 blind participants. The participants were asked to perform three tasks. In the first task, the participants entered different types of intersections and were asked to list all directions leading from each intersection. In the second task, the participants walked through a straight corridor while avoiding several placed obstacles. In the last task, the participants navigated a corridor containing both obstacles and intersections. The study revealed that Corridor-Walker enabled the participants to avoid obstacles with less reliance on the wall. Participants also acquired good shape knowledge of the intersections.

Part II

Assisting Blind People's Social Behaviors in Public Spaces

Chapter 5

BBeep: Sonic Pedestrian Avoidance System[1]

"People were noticing that I was approaching and people were moving away... giving me the path."

Participant 4
Pittsburgh International Airport
August 30, 2018

5.1 Introduction

Blind people face significant challenges when navigating public spaces due to the lack of visual sensing. Recent research using computer vision aimed to assist blind users' orientation and mobility skills for avoiding potential obstacles or hazards [38, 68, 98, 100, 144, 191, 235, 257, 261]. While these systems are often able to detect static obstacles, the detection and avoidance of collisions with dynamic elements, in particular pedestrians, is still relatively unexplored in the literature. Technical challenges aside, one possible reason for a lack of work on dealing with dynamic elements is the assumption that sighted pedestrians are aware of blind people and therefore will always clear the path for them. However, this is not always the case as sighted people may be looking at their smartphone, talking with others, or facing another direction (looking at a board or TV). In such scenarios, blind people face significant risks of collision with other pedestrians.

We present an assistive suitcase system, BBeep, that uses a sonic collision warning system to alert both the blind user and nearby sighted pedestrians about potential risks of collision (Figure 5.1). This approach extends common sonic warning systems that are used to clear the path for moving vehicles, such as airport carts driving through crowded terminals or large trucks driving in reverse. This work leverages the simple fact that sighted persons can quickly get out of the way of a blind person who is walking, if they are given appropriate information about a blind person's presence. However, our work aims to go beyond the paradigm of constantly playing a sound to convey the user's presence, as the constant emission of loud alarm sounds can be social disruptive and make the blind user feel overly self-conscious. Instead, we present an adaptive sonic warning system that only emits sounds when needed. More specifically, BBeep is designed to consider the motion of nearby pedestrians, predicts future collisions, and gives sonic feedback only when necessary.

Although we explicitly target the navigation of blind people in airports, we believe that the form factor of a travel suitcase is also appropriate in many other real-world crowded environments such as train stations or shopping malls. The use of a suitcase has several benefits in such environments. For the blind user, a suitcase can often act as an extended sensing mechanism for identifying changes in floor texture or as a form of protection from collisions in very dense crowds. Even without any smart sensing, a suitcase can be used as an assistive device. In many cosmopolitan environments, a suitcase is a common object and

[1]Project Page: https://wotipati.github.io/projects/BBeep/BBeep.html

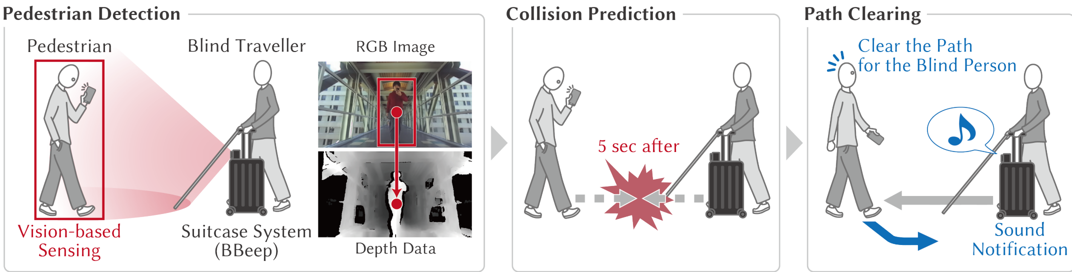


Figure 5.1: We present an assistive suitcase system, BBeep, that uses a sonic collision warning system to alerts both the blind user and nearby sighted pedestrians about potential risks of collision.

does not draw unnecessary attention to the user when it is not in use as a sonic warning system. As a robotic sensing system, the suitcase form factor also provides a convenient place to store and attach sensors, power and computing resources.

BBeep uses a RGBD camera to detect, track and predict the motion of nearby pedestrians. The RGB image is used to detect people using a convolutional neural network and the depth channel is used to estimate the distance to pedestrians. The tracker greedily associates the nearest pedestrian detections to track a pedestrians motion. Averaged position estimates are used to estimate pedestrians' velocity and linear extrapolation is used to predict their future path. Depending on the proximity of the predicted path with the user, an appropriate sound is emitted by the system.

To investigate how to convey sonic feedback effectively, we performed an observational study where the suitcase-shaped system emits alarm sounds of different types and timings. The results suggest that sound emission is an effective method to change the pedestrians' walking direction away from the platform, and that the timing of sound emissions has more impact in changing their trajectories than the sound types. Based on these findings, we designed the sonic notification interface of BBeep that used three stages of sound emissions to notify potential collision risks and imminent collisions with pedestrians.

In order to evaluate the effectiveness of BBeep for preventing collision with pedestrians, we performed a user study where six blind users walked with the suitcase in crowded areas of an international airport. We observed that BBeep reduced the number of situations of imminent collision risk, when compared to only notifying the blind user. Participant feedback also supported our hypothesis that BBeep is useful for collision avoidance in crowded public spaces. Based on our findings, we discuss future requirements for directions towards a more flexible and personalized solution that is able to adapt to different scenarios and users.

5.2 Related Work

We specifically focus on alerting both the blind user and nearby sighted pedestrians about potential risks of collision. In this section, we review related work specific to this chapter that were not detailed in Chapter 2.

5.2.1 Sound Alert for Urgent Notifications

Beep sounds have been used as an auditory alert to notify people of urgent situations, such as in hospital intensive care units [163], nuclear power plants [158], and aviation [35]. Audio notifications can also alert drivers of an eminent collision or assist in navigation [157]. The relationship between user perception and different types of alert sounds plays a vital role in their usability. Several works have found that auditory parameters of beep sounds (e.g., fundamental frequency, pulse rate, and intensity) affect perceived urgency levels [59, 87, 157]. Other studies also observed a trade-off between perceived urgency and annoyance levels of alert sounds [59, 78, 157].

As described above, emitting beep sounds is a common approach to notify users of urgent situations. Thus, we use this type of alert sounds to make pedestrians aware of a blind user. We investigate what types of beep sounds are effective for collision avoidance, and design a sound emission policy for our prototype system.

5.3 System Design: Path Clearing System

Our main goal is to ease the mobility of blind people in crowded environments. We argue that collisions with pedestrians can be avoided if both the blind user and sighted pedestrians are made aware of the collision risk. For this purpose, we developed BBeep, a sonic collision avoidance system that aims to clear the path for blind users, by also alerting nearby sighted pedestrians. BBeep relies on its ability to track and predict the future position of pedestrians in real-time, and on the use of sound notifications only when there is a risk of collision.

5.3.1 Limitations of Notifying only the Blind User

Prior research on obstacle detection for blind people focuses on notifying the user alone about the presence of obstacles, prompting them to change their orientation [144, 235, 261]. Such approach increases the user's knowledge of the surroundings, but also comes with significant limitations in this context. First, actively changing the walking direction of blind users may be unsafe (taking the user through a different/unknown path); second, a group of pedestrians may block the entire route of a blind user; and finally, notifying users about all pedestrians in crowded environments may require complex feedback, which may be cognitively demanding to users. In addition, by focusing on obstacles in general these approaches do not take advantage of sighted pedestrians' ability to cooperate in collision prevention. Therefore, we argue that conveying feedback only to the user may not be effective to avoid collisions with other pedestrians, in particular in crowded environments.

5.3.2 Sound Notifications for Users and Pedestrians

BBeep uses sound notifications due to the ability of sound to attract people's attention even when they are focused on something else. Although other modalities, such as visual stimuli (e.g., Vection Field [70]), can also impact pedestrians' walking direction, it may not be as effective in several scenarios. For instance, the prevalence of smartphones significantly reduces sighted pedestrians' awareness of the surroundings, resulting in potential collisions [45]. Moreover, (groups of) people talking or looking at a different direction may not notice a blind person until their white-cane hits them. For that reason, our approach is inspired in the common use of beep sounds to notify pedestrians of urgent situations prompting them to clear the path. A few examples include carts in crowded airports, large motor vehicles driving in reverse, or automated guided vehicles in factories[2]). However, such approach may also come with significant challenges, since frequent emission of loud alarm sounds can be socially disruptive and make the blind user feel uncomfortable.

5.3.3 Collision Prediction to Reduce Sound Emissions

To enhance social acceptance of sound emissions, it is important to emit alert sounds only when absolutely necessary. Moreover, a collision prediction technique is required in order to decrease as much as possible the number of sound emissions, while maintaining its effectiveness. For this reason, BBeep relies on real-time pedestrian tracking and collision

[2]https://reports.nissan-global.com/EN/?p=9747

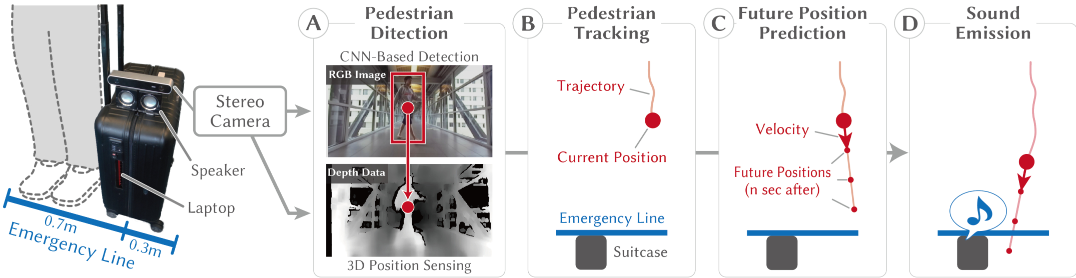


Figure 5.2: Overview of the sonic alerting process. The stereo camera is mounted on a suitcase and records RGB images and depth data. A) The system detects pedestrians using the RGB images and B) tracks their position using the depth data. Based on these tracking results, C) the system predicts the future positions of each pedestrian. Finally, D) the system emits an audible alarm signal if there is a risk of collision with the blind user.

prediction to provide notifications only when there is a potential risk of collision. This is beneficial to reduce both collision risks and social disruption in public spaces.

5.4 Implementation

We developed a vision-sensing system for tracking the motion of pedestrians and predicting their future positions in order to generate an audible warning signal that will clear the path in front of a blind user (Figure 5.2). A stereo camera is attached to a suitcase to capture RGB images and collect depth data. One advantage of this setup is the ability to capture images without significant motion-induced blur and to perform the necessary computations *in situ*. The system detects pedestrians using RGB images and tracks their position using the depth data in real time. Based on these results, the system predicts the future positions of pedestrians and determines their risk of collision with a blind user. The system can then emit an audible alarm if necessary.

We note that there is prior work on pedestrian trajectory forecasting [121, 125, 253] and the aim of this work is not to advance the state of the art in this respect. Instead, our contribution is the analysis and development of effective sonic feedback mechanisms based on such predictive input. To this end, our main challenge is to develop a real-time forecasting technique with sufficient accuracy for collision prediction.

5.4.1 Pedestrian Detection: Requirements and Design

The proposed system detects pedestrians located in front of the suitcase. We define the following requirements for image sensing and pedestrian detection to accurately predict collisions with a blind user or the suitcase in real time:

- **Long- and wide-range image sensing:**
  We require long- and wide-range image sensing for both the RGB and depth streams. A limited sensing range does not allow enough time for the system to predict pedestrian movements until just before the collision. A wide field of view is also important to capture nearby pedestrians.

- **Running at a high image sampling rate:**
  We require that all of the detection process be done rapidly to allow running at a high frame capture rate. This helps increasing the collision prediction accuracy.

- **Robust pedestrian detection:**
  We require a robust vision-based method for detecting individual pedestrians even if

the camera does not capture their full body (in particular, when a pedestrian is near the suitcase and the camera captures only the lower part of their body).

Many existing systems designed for avoiding static obstacles use Kinect v2 as a depth sensor [38, 98, 191]. However, Kinect v2 is limited in terms of its range and the FPS of image sensing (depth distance: 0.5 to 8.0 m, horizontal field of view: 70 degrees, and FPS (frame per seconds): up to 30 Hz). It also has a limited range of pedestrian detection (up to 4.0 m), and often fails to detect pedestrians when their whole body is incompletely captured.

To address these shortcomings, we implemented a novel system combining stereo image sensing and a CNN-based generic object detector (YOLOv2 [204]). We use a ZED™2K Stereo Camera[3], as it has a wider horizontal field of view (90 degrees), a longer depth range (0.3 to 20 m), and a higher FPS (up to 100 Hz) than Kinect v2. The stereo camera also supports a 3D odometry API that provides access to 3D movements from the camera in real time. Our system used such odometry information to remove the influence of suitcase rotation.

We used YOLOv2 to detect pedestrians using the RGB streams. The method robustly detects individuals even if their body is not completely included within the camera images. We confirmed that the combination of the ZED camera and YOLOv2 can detect bounding boxes of pedestrians from a distance of 10 m. We used the central area of the bounding boxes to obtain the 3D positions of the detected pedestrians in the camera coordinate system.

Updating positions using a high FPS is important for accurately predicting pedestrians' future positions. We used a laptop computer (Intel Core i7-7700HQ CPU, NVIDIA GeForce GTX 1060 GPU) to process object detection at a rate of 15 fps, but this was insufficient to obtain satisfactory prediction accuracy. We therefore used a given bounding box for obtaining 3D positions, and updated it once a new detection result was available. The system thus tracks pedestrians at a frame rate in excess of 40 fps. Note that the detection and tracking processes run simultaneously on different threads.

5.4.2 Pedestrian Tracking

The system processes pedestrian tracking based on the detection results. We propose an algorithm to track individual pedestrians in real time. The system uses the following procedure to update a tracking list of pedestrians based on the detection results at each frame:

1. The tracker generates a set of bounding boxes from pedestrian detection and computes a set of 3D positions in the camera coordinate system based on the central area of the bounding boxes.

2. The tracker repeats steps (3) and (4) for each detected 3D position (the current position).

3. If there are no existing pedestrians in the list from the current position within a distance α, the tracker adds the point to the list as a new pedestrian.

4. Otherwise, the tracker updates the position of the nearest pedestrian in the list to the current position, and saves the previous position as a record of its trajectory.

5. The tracker removes pedestrians from the list if their position has not been updated by the tracker in $\beta\ frames$ of the tracking process.

Based on our observations, we set the parameter values $\alpha = 1$ m and $\beta = 5$ frames for all of our studies.

[3]https://www.stereolabs.com/zed/

5.4.3 Position Prediction and Sound Emission

The system predicts the future positions of the pedestrians in the tracking lists derived from the tracking process. The system uses the 3D positions of the pedestrians in the camera coordinate system to predict the relative speed and direction of displacement between the suitcase and each pedestrian using their current position and their trajectory. To improve the stability of the pedestrian position measurement, the system first compensates for rotations of the suitcase (camera) by rotating the detected pedestrian positions using 3D odometry information. The system then computes the expected future position $\hat{\mathbf{p}}_{t+s}$ of each pedestrian after s seconds using the $N-1$ most recent points of its trajectory as follows:

$$\mathbf{p}_\mu(i,n) = \frac{1}{n}\sum_{j=i-n+1}^{i} \mathbf{p}_j \tag{5.1}$$

$$\hat{\mathbf{p}}_{t+s} = \frac{\mathbf{p}_\mu(t, \frac{N}{2}) - \mathbf{p}_\mu(t-\frac{N}{2}, \frac{N}{2})}{\Delta t(\mathbf{p}_{t-\frac{N}{4}}, \mathbf{p}_{t-\frac{3N}{4}})} s + \mathbf{p}_t \tag{5.2}$$

Let $\mathbf{p}_t$ be the position of a pedestrian in the camera coordinate system at time t, $\mathbf{p}_\mu(i,n)$ the average over n previous positions (from $\mathbf{p}_{i-n+1}$ to $\mathbf{p}_i$), and $\Delta t(\mathbf{p}_t, \mathbf{p}_i)$ the difference in the time stamps between $\mathbf{p}_t$ and $\mathbf{p}_i$. The system first calculates the two average positions (from $N-1$ frames before to $N/2$ frames before, and from $N/2-1$ frames before to current frame) (Equation 5.1). The system then calculates the vector between the two average positions and predicts the future position (Equation 5.2). Setting $N = 32$ was found to yield stable predictions.

The system then predicts the risk of a future collision based on all the predicted positions to decide whether an alarm sound should be emitted, as outlined in Figure 5.2. A collision is expected when a future pedestrian trajectory crosses the "emergency line" shown in Figure 5.2 (D). The system computes the line connecting the current and future positions of the pedestrian as a prediction of the expected future trajectory. The system then determines the intersection between this line and the emergency line. If the intersection lies within the range of the emergency line, the risk of collision is considered significant and the system emits a warning sound. Note that this calculation does not use pedestrian height information.

5.5 Designing Sound Emission Policy through an Observation Study

We studied the response of pedestrians to the audible warning signals in order to design a sound-emission policy for our system. As described in Related Work, human perception of audible emergency warnings has been studied [59, 78, 157]. There is also some understanding of how a visual stimulus can cause pedestrians to redirect their trajectory [70]. Yet, little is still known about how a pedestrian reacts to an audible signal. Such insight is important for designing an effective policy of sound emissions for our path-clearing system.

We conducted an observational study in a corridor in which the suitcase-enclosed system was made to emit different types of audible alerts (beeps). We recorded pedestrians' reactions and trajectories as shown in Figure 5.3. We designed a set of sound-emission patterns comprising various sound types and a range of timings. The system tracked pedestrians in the corridor and emitted alerts using these patterns. We analyzed the pedestrians' trajectories to determine which sound patterns were most effective at clearing the path in front of the suitcase. Based on the outcome, we then designed our sound-emission policy for evaluation in an airport.

Table 5.1: Sound-emission patterns. UL: urgency level, BF: base frequency, PD: Pulse duration, IPI: inter-pulse interval, and s: timing of the sound emission (i.e., the suitcase emits a sound alert by considering the expected position of pedestrians after s seconds).

Sound Pattern	Sound type				Timing
	UL	BF	PD	IPI	s
S1	High	1000 Hz	0.1 s	0.1 s	5.0 s
S2					2.5 s
S3	Middle	400 Hz	0.1 s	0.1 s	5.0 s
S4					2.5 s
S5	Low	400 Hz	0.5 s	0.5 s	5.0 s
S6					2.5 s
S7	Without sound emission				N/A

5.5.1 Sound-Emission Patterns

We designed 7 sound patterns (S1–S7), one of which (S7) was mute to serve as a baseline. The 6 non-baseline patterns featured combinations of 3 different sound types and 2 types of emission timings.

Beep-Alert Sound Types

We used alert sounds to represent 3 distinct levels of perceived urgency. The relationship between perceived urgency and sound parameters is well documented. We prepared 3 types of beep sounds with different urgency levels denoted High, Intermediate, and Low. Specifically, we varied the base frequency, the pulse rate, and the pitch, as given in Table 5.1. The values we used are based on recent research addressing sound urgency [192, 212].

Timing of sound emission

We also used different timings of sound emissions for each beep alerts. The system changes the timing by setting the collision detection parameter s seconds. If a system were to emit a sound alert immediately before a predicted collision with the blind user (e.g., $s < 1$)), the pedestrian in question may not be able to avoid the collision. On the other hand, a sound alert emitted too long in advance (e.g., $s = \infty$) may cause unnecessary disturbance and inconvenience and be effectively unproductive.

We selected the parameter values $s_1 = 5.0$ and $s_2 = 2.5$ seconds. The value of $s_1(5.0\ s)$ represents the time needed to travel the furthest distance in the detection range (around 10 m) when the blind user and a pedestrian are approaching at a relative speed of 4 km/h. We also used $s_2(2.5\ s)$ set to a half of s_1 to define a nearer threshold.

5.5.2 Data Collection and Analysis

This observational study considered the suitcase-enclosed system with 7 sound emission patterns placed in a straight corridor (Figure 5.3). The system tracked pedestrians and predicted their intersection with the emergency line in real time. The system also emitted sounds as specified by the adopted policy. All seven sound patterns were used in cycle.

Observations were conducted over more than four days, yielding 57 trajectories for each pattern (399 trajectories in total). Note that the system recorded a trajectory and images of the leading pedestrian who had a risk of collision when the system used a sound pattern.

Figure 5.3: The suitcase-enclosed system in a corridor, equipped to emit various types and timings of beep sounds.

We analyzed the recorded datasets to identify how trajectories were affected by the emissions, as illustrated in Figure 5.4. Some trajectories were occasionally missing or inaccurate owing to the limitations of real-time processing, as described in Pedestrian Tracking. We therefore performed a subsequent trajectory analysis using the recorded RGB and depth images to obtain more accurate pedestrian trajectories. We used OpenPose [43], a CNN-based human-body detection software, to detect parts of pedestrians' bodies from the RGB images. We then determined the central position of the detected bodies in the depth images to obtain the 3D positions of pedestrians. We conducted the analysis for all the recorded images. The laptop used analyzed images captured at a rate of 5 fps, i.e., an insufficient rate for real-time sound notifications.

5.5.3 Evaluation Measurements

We measured the *"minimal distance"* between the suitcase position and a given trajectory to investigate the relevance of different sound patterns for avoiding collisions. A longer minimal distance may be interpreted as indicating that the pedestrian has avoided the blind user by a comfortable margin. These minimal distances were determined from the 3D positions returned by OpenPose.

We considered three hypotheses for the main potential factors influencing the minimal distances: the presence or absence of a sound emission (Hypothesis 1), the sound-emission timing (Hypothesis 2); and the urgency level of the emitted sound (Hypothesis 3). We tested these hypotheses using a Kruskal–Wallis test and a Mann–Whitney U test at 5% levels of significance to discern differences within sound patterns. We also saw 95 % confidence intervals for each pattern.

5.5.4 Results

Figure 5.5 shows the minimal distance determined for each pattern. The mute baseline pattern (S7) yields the smallest average minimal distance. The Kruskal–Wallis test and the Mann–Whitney U test, done at 5% levels of significance, revealed that all the non-baseline sound patterns (S1-S6) gave longer minimal distances than S7. This result validates Hypothesis 1. Based on the statistical tests and the 95 % confidence intervals, we also observed that the sound emission patterns with 5 second timings (S1, S3, and S5) give greater minimal distances than the patterns with 2.5-second timings (S2, S4, and S6). Hypothesis 2 is thus also validated.

We then compared sound patterns with the same emission timing to assess the influence of the urgency levels. We observed no statistically significant difference among patterns

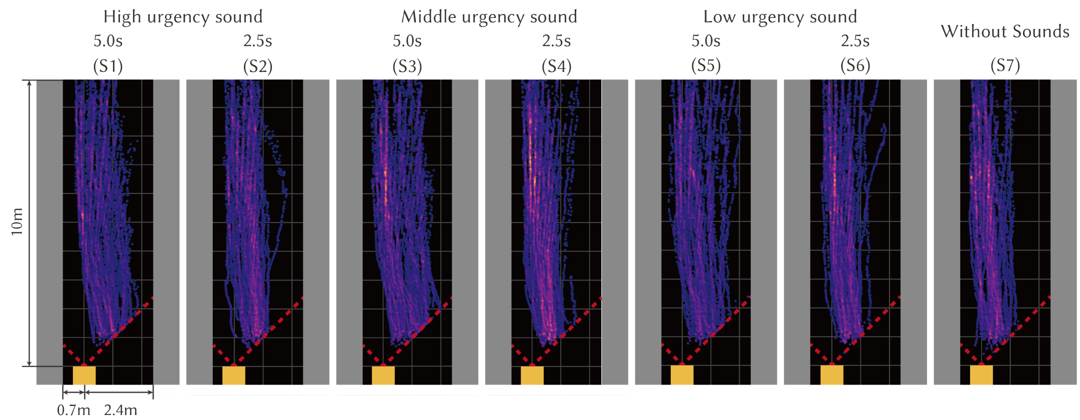


Figure 5.4: Heat maps depicting pedestrian trajectories seen from above. The suitcase (yellow rectangle) is located near the left-side wall of the corridor. The walls are shaded in gray. (Important note: the trajectory distribution appears cone-shaped owing to the ZED's limited field of view. The captured trajectories lying outside this range are consequently shorter than those along the central axis.)

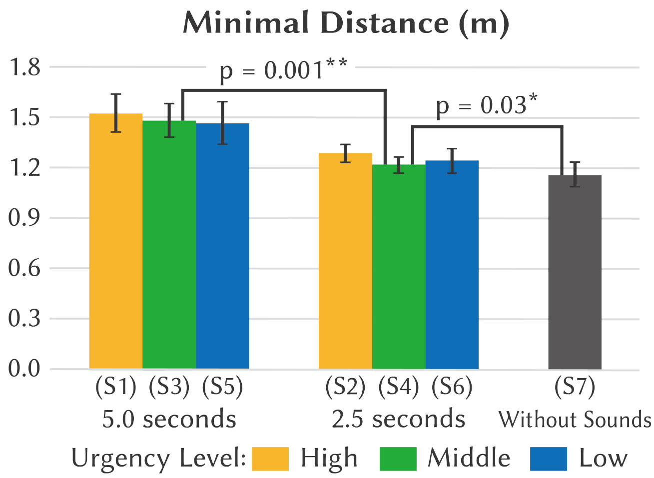


Figure 5.5: Minimal distances. The bars show the 95 % confidence intervals. p: p-value of the Mann–Whitney U test done on the minimal distance ($**$ and $*$ indicate the 0.001 and 0.03 levels of significance, respectively).

with either the 5-second (S1, S3, and S5) or the 2.5-second timings (S2, S4, and S6). We thus rejected Hypothesis 3.

We summarize our findings as follows:

- Sound warnings based on collision prediction influenced pedestrians walking toward away from the suitcase.
- Timings of the sound emissions also affected pedestrian trajectories. Patterns with a 5-second timing deflected pedestrian trajectories more effectively.
- The type of alarm sound appears not to be a significant factor affecting pedestrian trajectories.

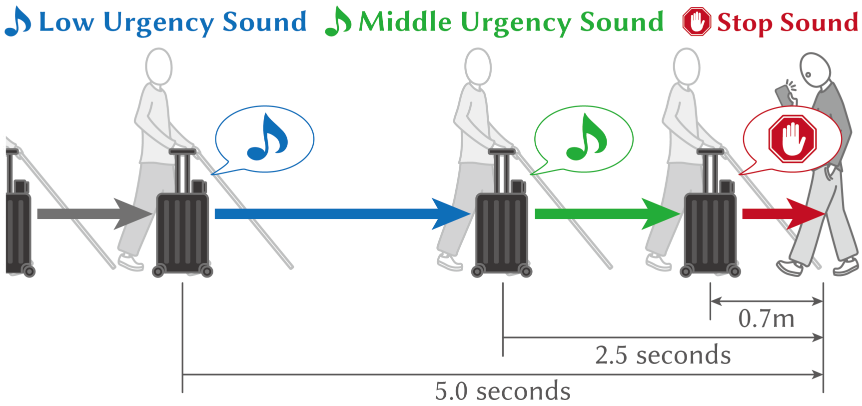


Figure 5.6: Policy of sound emissions.

5.5.5 Design of the Sound-Emission Policy

Based on the above findings, we designed a sound-emission policy for BBeep (Figure 5.6), consisting of three stages of sound emissions for preventing collisions. The system emits the following three types of alarm sounds.

1. **Low-urgency beep:**
   This sound warns of the potential risk of a collision between the blind user and pedestrians within 5 s. This sound was used in our observation study as a *low-urgency sound* with a base frequency of 400 Hz, a pulse duration of 0.5 s, and an inter-pulse interval of 0.5 s. We expect this signal to enable pedestrians to divert their path away from the blind user to avoid collision.

2. **Intermediate-urgency beep:**
   This sound indicates a potential risk of collision within 2.5 s. This sound was used in our observational study as an *intermediate-urgency sound* with base frequency 400 Hz, a pulse duration of 0.1 s, and an inter-pulse interval of 0.1 s. We also expect this signal to help avoid collision.

3. **Stop sound:**
   This sound indicates an imminent risk of collision with any obstacle (pedestrian, chair, wall, etc.) located within 70 cm. We expect this signal to prompt the blind user to come to a halt immediately.

We chose the intermediate- and low-urgency sounds for our policy. The higher the urgency level of the signal, the greater the annoyance rating of the sound alert. However, we observed that the urgency level of the sound did not affect the trajectory of oncoming pedestrians. We therefore selected two sound alerts with lower urgency and annoyance levels. By using two types of beep sounds, the blind user can know whether or not a pedestrian continues to approach. In addition, to inform the blind user of an obstacle ahead, we use a bell sound that is a completely different from the beep sounds. This bell sound is emitted whenever the system detects obstacles located within 70 cm. In our user evaluation, we recommend that the blind user stop advancing immediately upon hearing the bell sound.

5.6 Real-World User Evaluation

Our main goal was to understand the effectiveness of BBeep in clearing the path for blind travellers in crowded spaces. For that reason, we performed a real-world study where 6 blind participants (Table 5.2) navigated crowded areas at the Pittsburgh International Airport. In this study, we compared BBeep against two baseline conditions: one notifies only the blind user about collision risks, while the other does not provide any notifications.

5.6.1 Conditions

We equipped our assistive suitcase system with the capabilities to track pedestrians and predict future collisions. Based on this system, we prepared three different interfaces:

1. **Speaker interface (BBeep):**
   This interface represents our proposed system, which emits three types of sounds (low urgency, middle urgency, and stop sounds) for the blind user and other pedestrians through a speaker that is mounted on top of the suitcase.

2. **Headset interface:**
   This interface has the same behavior as BBeep, but instead of using a speaker, provides sounds only to the blind user using bone conducting headphones (to avoid blocking environmental sound).

3. **No sound interface:**
   The user also carries the suitcase, but this interface never emits sound, representing a blind user navigating by himself without additional aids.

5.6.2 Tasks

We selected several crowded gates where passengers were waiting for boarding in line or in groups. Participants were asked to walk straight along the corridor and go through the crowds until reaching a particular location, where the experimenter would ask them to stop (each task had roughly 20 meters). Participants held the suitcase handle with one hand, and used their white-cane on the other hand (Figure 5.7). Their goal was to go through the crowds effectively and avoid collisions with other pedestrians. This task enabled us to replicate a very similar setting among different participants and trials, thus enabling a fair comparison among conditions.

5.6.3 Procedure

After obtaining (IRB approved) informed consent from participants, researchers provided an overview of the study and described the three interfaces. A short training session (10 - 15 minutes) was then given to participants until they were familiar with the system alarm sounds and interfaces. Although the volume rate of the speaker interface was fixed for all participants, they were able to adjust the volume in the headset interface to make sure it was comfortable, but audible. During the training session, we explained how to hold the suitcase as it affects the accuracy of collision prediction.

Then, participants were asked to walk five similar routes using three types of interfaces (the speaker and headset interfaces twice and the no sound interface once) in a counterbalanced order. Participants were informed that a researcher would be walking behind them to guarantee their safety as well as other pedestrians' safety (Figure 5.7). They were also instructed to stop when listening the higher urgency (stop) sound to avoid colliding with pedestrians. The researcher did not intervene unless: there was an imminent risk or a deviation from the path. For example, in the latter, the researcher would tell them to slightly adjust their orientation. Also, in case the path was blocked or the participant was confused, the researcher would intervene to help the participant passing that immediate obstacle.

To observe the response of pedestrians to the system, we mounted a GoPro camera on the top of the suitcase.

Table 5.2: Demographic information of our participants.

ID	Gender	Age	Eyesight	Navigation Aid
P5.1	Female	70	Blind	Cane
P5.2	Female	70	Blind	Cane
P5.3	Male	65	Blind	Cane
P5.4	Male	46	Blind	Dog (primary) and Cane
P5.5	Male	42	Blind	Dog (primary) and Cane
P5.6	Male	58	Blind	Cane

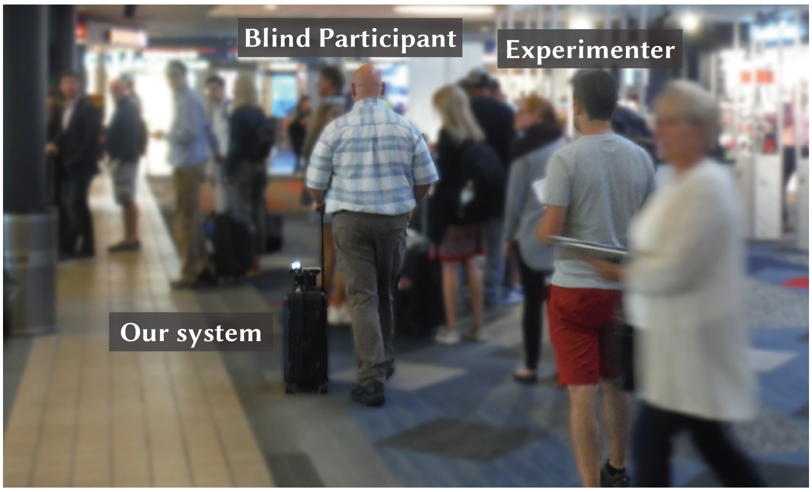


Figure 5.7: User study at an international airport. Participants walked through crowds holding the handle of an assistive suitcase-shaped system.

5.6.4 Metrics

Imminent Collision Frequency and Collision Risk Frequency

To measure how many pedestrians had an imminent risk of collision with the blind user, we defined the number of pedestrians within 70 cm as the *Imminent Collisions Frequency*. In addition, we measured the *Collision Risk Frequency* that indicates how many pedestrians had a risk of collision with the blind user within 5 s. In each task, we counted imminent collisions and risk of collision based on pedestrian detection results and our collision prediction results, respectively. We compared the three conditions quantitatively based on a 95% confidence interval (Table 5.5). In addition, we compared the two sound conditions (speaker and headset interfaces) using a Wilcoxon signed-rank test with 1% levels of significance.

Risk Continuity Ratio

This metric represents the ratio of pedestrians who had potential risks of collision, and persisted in the users' path until reaching an imminent risk of collision. To calculate the metric, we divided the Imminent Collision Frequency by the Collision Risk Frequency. Smaller values indicate that the system reduces the risk of collision between the blind user and pedestrians. We performed the same analysis described for the previous metrics. We also compared the three interfaces by using the 95% confidence interval and compared the speaker interface and the headset interface using the Wilcoxon signed-rank test with 1% levels of significance.

Table 5.3: Likert items (1: strongly disagree to 7: strongly agree) and a summary of answers. M: Median.

Questions	P5.1	P5.2	P5.3	P5.4	P5.5	P5.6	M
Q5.1							
People cleared the path when I was **using the speaker interface**.	4	6	7	7	5	4	5.5
People cleared the path when I was **not using the speaker interface.**	4	3	2	3	4	4	3.5
Q5.2							
The speaker interface helped me walk comfortably in airports.	7	4	5	6	6	5	5.5
The headset interface helped me walk comfortably in airports.	7	4	3	6	6	6	6
Q5.3							
The speaker interface is also useful in less crowded places.	5	2	6	6	6	4	5.5
The headset interface is also useful in less crowded places.	5	2	4	5	6	4	4.5

Table 5.4: The mean and standard deviation of the number of pedestrians who had potential risks of collision. 5.0 s: the number of pedestrians who had potential risk of collision within 5.0 s, 2.5 s: within 2.5 s, and 70 cm: number of pedestrians within 70cm.

Interface	5.0 s	2.5 s	70 cm
Speaker	4.75 ± 2.45	2.33 ± 1.97	0.42 ± 0.76
Headset	5.42 ± 1.93	3.08 ± 1.55	2 ± 1.35
No sound	5 ± 1.07	3.83 ± 0.99	3 ± 1.85

Post-Interview

After completing the tasks, we asked participants to rate a set of sentences using 7-point Likert Items (ranging from 1: strongly disagree, 4: neutral, to 7: strongly agree). The sentences and a summary of the answers are shown in Table 5.3. Finally, we asked open-ended questions about the advantages and challenges of each interface (speaker and headset). We also asked for suggestions to improve each interface, and in what scenarios would the participants use the Speaker Interface (BBeep).

5.6.5 Results

Quantitative Evaluation

Table 5.4 reports the number of pedestrians who had potential resks of collision. In addition, table 5.5 reports the imminent collision frequency, collision risk frequency and the risk continuity ratio. We found no significant differences between interfaces concerning the collision risk frequency ($p = 0.8$). On the other hand, our analysis revealed that the speaker interface resulted in significantly ($p = 0.005$) less pedestrians with an imminent risk of collision with the user, when compared to the headset interface. Moreover, a significant difference ($p = 0.009$) in risk continuity ratio shows that the speaker interface was more effective to reduce the number of pedestrians that had a risk of collision with the user.

Table 5.3 reports the results of six Likert scale questions. Four participants agreed that people cleared the path when they were using the speaker interface rather than the headset interface (Q5.1). On the other hand, in the other questions, we obtained similar results between two interfaces.

Table 5.5: Quantitative evaluation of each metric. It presents means and standard deviations (SD), and the lower and upper bounds of 95% confidence intervals.

Interface	Collision Risk Frequency			Imminent Collision Frequency			Risk Continuity Ratio		
	Mean and SD	Lower	Upper	Mean and SD	Lower	Upper	Mean and SD	Lower	Upper
Speaker	6.67 ± 3.75	4.55	8.79	**0.41 ± 0.76**	**0.00**	**0.85**	**0.08 ± 0.19**	**0.00**	**0.19**
Headset	5.91 ± 2.25	4.64	7.19	2.00 ± 1.35	1.23	2.76	0.37 ± 0.25	0.22	0.51
No sound	6.67 ± 2.05	5.03	8.30	3.00 ± 1.85	1.52	4.48	0.45 ± 0.21	0.28	0.63

Video Observations

Video recordings enabled us to analyze the behavior of both the blind user and sighted pedestrians, in order to complement our quantitative metrics. We observed that participants would clear the path most of the times after noticing the user. However, participants using the Headset or No Sound interfaces often collided (or had an imminent risk of collision) with pedestrians who were unaware of their presence. In most occasions, pedestrians were either talking in groups or standing in line waiting for boarding. On the other hand, when using the Speaker interface, even in the aforementioned scenarios pedestrians would hear the sound and immediately clear the path for the blind user.

Speaker interface reduced the number of with such people because the interface notified the presence of the blind user to them. There were, however, five exceptions where pedestrians approached within 70 cm radius of the participant, representing an imminent risk of collision. The reasons for them were: (1) pedestrians who were standing in front of blind participants were putting headphones on and did not hear the sound notification; (2) a blind user changed walking direction suddenly; and (3) Although pedestrians tried to clear the path for the blind person, they did not take enough space between the blind person and them (e.g., by being against the wall).

Qualitative feedback

Participants were generally aware that other pedestrians cleared the path when using BBeep, as illustrated by their comments:

> **A5.1:** *"The advantage of the speaker is [that] they [other pedestrians] cleared the path"* P5.5
>
> **A5.2:** *"People were noticing that I was approaching and people were moving away... giving me the path"* P5.4
>
> **A5.3:** *"The biggest advantage is that other people heard it [sound alert] and they would move to get out of the way"* P5.3

When using the headset interface, participants felt that being quieter was its main advantage. However, they also had the perception that pedestrians did not clear the path in the same way:

> **A5.4:** *"It [the headset interface] is more private"* P5.1
>
> **A5.5:** *"People don't notice, so I'm required to say something for them to clear the path.. in comparison to the speaker"* P5.4
>
> **A5.6:** *"[The main advantage of the headset interface is that] it's quiet. ... [The main challenge is that] it didn't get anybody's attention"* P5.2

Some participants commented that the usefulness of the speaker interface might depend on the context or environment. Places where they would use it generally include crowded public spaces and open areas:

A5.7: *"It's more useful in more crowded places"* P5.1

A5.8: *"It is useful at a grocery store, a shopping mall, and other open areas"* P5.5

In addition, it was also reported to be useful in less crowded place:

A5.9: *"The speaker interface is also useful in less crowded places, because it doesn't beep when there are no people. So, I can still take advantage when there is someone with a risk of collision"* P5.3

In contrast, using it in quieter environments was found to be inappropriate or to draw too much attention:

A5.10: *"In the airport type of settings, I would probably use the speaker settings, but if I'm in a quiet area where people are expected to be quiet, ... maybe I will not use it"* P5.4

A5.11: *"I don't agree to use the speaker interface at places supposed to be quiet like hospitals or libraries, but, in any public environment like airports, train stations, or whatever, the speaker is always gonna be appropriate."* P5.2

When asked for suggestion, two participants mentioned that the system should not only beep sounds but also provide more information about the surroundings:

A5.12: *"I'm more likely use the speaker ... but I still want to hear what's going on through the headset"* P5.3

A5.13: *"[In both interfaces] I want to understand what's happening. People are in front of me walking, coming, or standing. ... [I recommend] different output. Speaker will notify sighted people. Headset will explain what's going on."* P5.2

5.7 Discussion

5.7.1 Effectiveness of BBeep for Collision Avoidance

The real-world user study showed that BBeep was an effective tool for blind users to prevent collisions with pedestrians. While the number of pedestrians with a low risk of collision (within the next 5 seconds) was very similar among conditions, the number of pedestrians with an imminent risk of collision with the user was significantly lower for BBeep than the Headset condition. These two conditions provide the exact same sound notifications, but use different output sources (i.e., speaker or headset). This result indicates that emitting sound both to the blind user and to nearby pedestrians was effective for clearing the path for the user, and that it was more effective than notifying the user alone. Video observation of the navigation tasks at the airport corroborate these results. Participants traversed crowded areas near the gates and frequently encountered pedestrians who were unaware of them. When walking with BBeep, sighted pedestrians gained immediate awareness of the user's presence and cleared the path and, in some cases, even prompted their peers to move. Although not always aware of sighted users' behavior, participants had the perception that BBeep was more effective than the alternatives, as shown by their ratings and comments (A5.1–A5.3).

While the Headset condition was not as effective as BBeep, we found interesting techniques that intended to leverage the users' knowledge about the collision risk with pedestrians. For instance, P4 started saying *"Excuse me!"* after noticing that the collision risk persisted. Also, other participants became more effective orienting the suitcase in order to find a path without collision risks.

5.7.2 Prospective Scenarios for BBeep

We carefully designed our sound emission policy, keeping in mind that social acceptance was crucial for such an approach. In addition, conducting the experiment at the airport enabled participants to understand how would it be to use BBeep in the real-world. Participants' feedback indicates that it is acceptable to use BBeep in crowded, public spaces such as airports, train stations or shopping malls (A5.7, A5.8, A5.10, and A5.11). Indeed, users' reported comfort in using the suitcase-shaped system (Q5.2) showed very similar results between the Headset interface and BBeep. In contrast, participants' feedback regarding the use of both interfaces in less crowded places is not consensual (Q5.3). Still, some participants see advantages in using them since they do not provide notifications unless there are risks of collision (A5.9). While crowded areas seem appropriate to use BBeep, participants commented that they would not use it in very quiet places where they would attract too much attention or in places where they are supposed to be quiet, such as hospitals or libraries (A5.10 and A5.11).

5.7.3 Limitations and Future Work

Reducing the Number of Sound Emissions

The main advantage of the Headset interface was it discreetness, as it does not attract so much attention nor disturb other people (A5.4 and A5.6). However, being more private significantly impacted task performance. This relation between performance and discreetness suggests that it is important to investigate how to further reduce the number of sound emissions while maintaining its ability to clear the path for blind users. For instance, we observed that sometimes sound notifications were provided even when pedestrians had already noticed the blind user, but did not clear the path immediately. In these scenarios, possible future directions include using face tracking or gaze estimation techniques [262] to assess whether pedestrians are aware of the blind user, thus reducing the number of sound emissions.

Impressions of Pedestrians

To assess the acceptability of BBeep, it is relevant to investigate not only the impressions of blind users, but also those of sighted pedestrians. However, in this case recruiting sighted people beforehand would prevent us from evaluating BBeep's ability to help clearing the path for the blind user. We aim to further explore sighted people's impressions in the future with a different study design.

Beyond Path Clearing

In order to evaluate the impact of our approach, we focused exclusively on collision avoidance and on the ability to clear the path for the blind user. For that purpose, we used straight-line routes and did not include additional navigational challenges that could affect the results. These design decisions allowed us to run a more controlled experiment, despite being done in a real-world scenario. However, independently traversing complex environments like airports has additional significant challenges such as following a particular route, or gaining knowledge about surrounding Points of Interest (POIs).

The need to convey more informative feedback to the blind user was also mentioned by participants, who wanted to know more details about their surroundings (A5.12 and A5.13). One possible extension is to encode distance (or urgency) information continuously instead of using three pre-determined levels. A different possibility is to provide the user with additional information that is useful for orientation and mobility. In particular, P2 and P3 suggested to combine the speaker and the headset so that they provide different feedback to the user (A5.12 and A5.13). They suggested to use BBeep as is, but to describe the environment using the bone-conductive headset. Future directions may include investigating how to combine BBeep with solutions that provide turn-by-turn navigation assistance and/or convey information about relevant POIs in the vicinity of the user [6, 63, 215].

5.8 Conclusion

We proposed an assistive suitcase system, BBeep, that aims to clear the path for blind users when walking through crowded spaces, by notifying both the user and sighted pedestrians about the risks of collision. It provides sound notifications only when needed, based on pedestrian tracking and by predicting their future position in real-time. We first investigated how to convey the sound feedback effectively to sighted pedestrians and designed the sonic notification interface of BBeep. Then, we conducted a real-world user study with visually impaired people in an airport. Results showed that BBeep reduces the number of situations of imminent collision risk when compared to notifying the blind user alone. Moreover, users found BBeep acceptable and appropriate to use in crowded, public spaces such as airports, train stations or shopping malls. Yet, they were more hesitant about using it in places they are supposed to be quiet. In the future, we plan to extend our collision prediction method, by using vision-based attention analysis to reduce the number of unnecessary sound emissions when the pedestrians have already noticed the presence of the blind user.

Acknowledgement

We thank the Allegheny County Airport Authority and all study participants. We would also like to thank the anonymous reviewers for their helpful comments. This work was sponsored in part by JST CREST (JPMJCR14E1), JST AIP-PRISM (JPMJCR18ZG), NSF NRI award (1637927), NIDILRR (90DPGE0003), Uptake (CMU ML for Social Good fund) and Shimizu Corporation.

Chapter 6

Guiding System for Walking Seamless with Nearby Pedestrians[1]

"While walking alone, I always concentrate on grasping surrounding environments via auditory sensations to avoid collisions. Using the system alerted me of the risks of collisions, so I could walk more confidently."

Participant 13
IBM Japan Ltd
September 2, 2019

6.1 Introduction

Blind people face significant risks of collision with other pedestrians when walking through public spaces due to their lack of vision. According to one survey, 87.8% of blind people have collided or nearly collided with pedestrians, bicycles, and other obstacles [242]. Using a white cane is the most common method for the blind to sense obstacles and pedestrians, but it requires a user to risk their safety to physically contact the object. Therefore, blind people report that the cane is not useful in avoiding contact with walking pedestrians in crowded sidewalks and corridors or in other crowded environments [243]. Due and Lange reported that blind pedestrians rely on the collision avoidance behaviors of sighted pedestrians, such as changing trajectory or stopping [58]. That is why collision incidents happen when sighted pedestrians have difficulty noticing blind pedestrians in public spaces such as stations [2] and airports [115].

Meanwhile, sighted pedestrians continuously adapt their speed and direction using their sense of vision to make their walking seamless with nearby pedestrians [171]. We characterize such walking behaviors as two types of avoidance behavior: (1) *"on-path"* avoidance: adjusting walking speed without changing the path; and (2) *"off-path"* avoidance: changing the path and walking through free space. For example, sighted pedestrians choose the on-path avoidance when other pedestrians *will* cut across in front, but choose the off-path avoidance when people are standing still in front and talking. Our goal in this work is to enable blind people to walk seamlessly with nearby pedestrians by using the on-path and off-path avoidance, like sighted pedestrians. We argue that the on-path avoidance is more important for blind people because they have to walk along non-visually sensible landmarks. Changing their path frequently may risk them losing their way and becoming disoriented.

Research using computer vision has aimed to assist blind pedestrians to avoid obstacles or hazards [23, 38, 85, 143, 144, 197, 235, 243, 261]. These systems generate an alternative

[1]Project Page: https://wotipati.github.io/projects/IMWUT2020/IMWUT2020.html

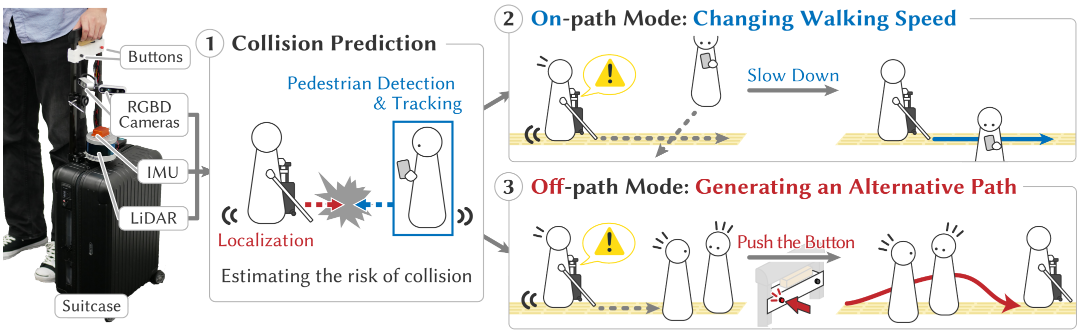


Figure 6.1: Overview of the proposed system. 1) The system predicts the potential risks of collisions using two RGBD cameras, a LiDAR sensor, and an IMU sensor. Then, when the system detects a risk of collision, it emits low-urgency alert signals. 2) Blind users who receive these alert signals can avoid a collision by changing their walking speed. In addition, 3) when pedestrians are blocking the user's path, the system emits high-urgency alert signals continuously. If the user pushes the button equipped on the suitcase, our system generates a path around the blocking pedestrians and guides the user safely around them.

path around the detected obstacles and navigate blind users. However, these systems provide only the off-path avoidance, rather than the on-path avoidance. BBeep is a sonic collision warning system to alert nearby sighted pedestrians about potential risks of a collision via beeping sounds [115]. The system assumes sighted pedestrians will give way to blind users, and a guiding system to help blind people walk seamlessly with nearby pedestrians has not been explored.

We present a guiding system to help blind people walk in public spaces by adapting their walking speed to avoid collision with approaching pedestrians (the on-path avoidance) and by enabling them to avoid standing pedestrians (the off-path avoidance). The system first predicts the risks of collisions using sensors, and then it recommends that the blind user adjust his or her walking speed to avoid a collision with a walking pedestrian or take an alternative route to avoid a collision with a standing pedestrian. For example, if a pedestrian is going to cut across in front of a blind user, the system recommends adjusting the blind user's walking speed. This is called the *"on-path mode."* If a group of people is blocking the blind user's path, the system recommends an alternative route and navigates the user. This is called the *"off-path mode."*

To realize the on-path and off-path modes, our system predicts the risks of collisions with nearby pedestrians. Our system first localizes the user's position and calculates the user's velocity using simultaneous localization and mapping (SLAM) with a light detection and ranging (LiDAR) sensor. The system uses two RGBD cameras to capture a wide field of view and detects surrounding pedestrians by applying a convolutional neural network (CNN)-based object detector from the data on the two cameras. Then, the system can accurately track and predict the motion of multiple pedestrians by compensating for camera motion using the SLAM results. By combining these sensing results, the system predicts the potential risks of collision (Figure 6.1 (1)). When the system detects a risk of collision, the on-path modes emit low-urgency alert signals for the users. By walking slowly or stopping while being alerted, the users can avoid collisions without changing their path. (Figure 6.1 (2)). When someone is blocking the blind user's path, the system continues to emit high-urgency alert signals (Figure 6.1 (3)). In such situations, the off-path mode can be initiated by the user, and the system generates an alternative path to avoid the collision with the standing pedestrian. We designed the system to be attached to everyday luggage like a rolling suitcase. We attached two cameras and a LiDAR sensor to the bar of the handle on a suitcase and asked blind users to carry it. This rolling suitcase form-factor is used as a supportive system for blind people in a recent work [115]. This suitcase-based system has several advantages such as it can capture images without significant motion-induced blur and can carry sensors and computational resources easily [115] (other advantages are described in 3.4.).

Navigation technologies for blind people commonly use an audio interface [4, 6, 32, 47, 62, 174, 190, 197, 201, 214, 215, 216, 250] or tactile interface [15, 211, 243, 256]. Each interface has its own characteristics. For example, audio interfaces can convey clear instructions, but they may block ambient sounds that blind people often use to ensure their safety [36]. Although tactile interfaces may not block these ambient sounds, they have difficulty conveying detailed information. Because of these characteristics, blind users' preferences depend on the types of tasks and environments (e.g., turn-by-turn navigation, collision avoidance, indoor/outdoor navigation, and crowded/empty spaces). In this paper, we present our implementation of tactile and audio interfaces to find out which is more suitable for our target situations with our guiding system. The audio interface alerts the risks of collisions by using beep sounds and guides the user by using text-to-speech feedback through a bone conduction headset. The tactile interface warns of the risks of collisions with a vibrating handle and navigates users with a newly developed directional lever, which shows the correct direction. We attached the two tactile devices to the handle of the suitcase.

We conducted a user study with 14 blind people in specific routes and evaluated the effectiveness of the audio and tactile interfaces and the overall guiding system. We obtained the following results.

1. Most blind participants successfully avoided the walking and standing pedestrians in both controlled and real-world environments by using both interfaces.

2. The sound-based audio interface for the on-path mode made it easier for blind participants to recognize alerts from the system than the vibration-based tactile interface. One reason was that the vibration was affected by the floor texture.

3. Participants completed tasks using the tactile interface (the directional lever) for the off-path mode significantly faster than they did using the speech-based audio interface.

Overall, participants had significantly stronger preferences for the tactile interface after the studies. The audio interface was useful in certain situations, but its tendency to block ambient sounds was a major drawback.

6.2 System Design: Guiding System for Public Spaces

Our main goal is to develop a guiding system to help blind people and to make their walking seamless with nearby pedestrians in public spaces. Such public spaces are characterized by *"restricted, impeded, and unstable flow* of pedestrians in the levels of services defined by Polus et al. [193]. In this section, we describe the design of our guiding system specifically for the following typical situations: *A blind pedestrian tries to walk through a public space such as public buildings and shopping centers. He/she is familiar with the route. He/she should be able to walk seamlessly with the surrounding pedestrians.*

- **Situation 1:**
  Other pedestrians often cut across the blind pedestrian's path at close range with average walking speed (1 – 1.5 m per second). Such pedestrians can be regarded as dynamic obstacles for him/her.

- **Situation 2:**
  A group of standing pedestrians blocks the blind pedestrian's path unintentionally. Such pedestrians can be regarded as static obstacles for him/her.

6.2.1 On-path Navigation

Blind pedestrians are usually trained to walk along specific familiar routes with non-visual landmarks given that they lack vision. We argue that a system should stick to these familiar paths as much as possible even when it navigates the blind pedestrian so as to avoid collision. We call such navigation without route-changes *"on-path"* navigation. In Situation 1, the system should recommend adjusting the blind user's walking speed without changing his or her path to avoid a collision. The technical challenge is creating a method to present such alerts in real-time. Therefore, we designed two interfaces: a sound-based audio interface and a vibration-based tactile interface based on previous work [26, 59, 87, 142, 157, 192, 196, 212]. The system first senses the walking speed and direction of the approaching pedestrian and predicts the trajectory and the risks of collisions. Then, the system alerts the user to adjust his or her walking speed to avoid a collision with a walking pedestrian (see Implementation).

6.2.2 Off-path Navigation

In Situation 2, the system needs to help a blind pedestrian avoid obstacles by having him or her move out of the path, walk through free space, and return to his or her path. Successfully navigating a user accurately along a target path by continuously presenting directions is a challenge, but previous studies showed that audio and tactile interfaces may have sufficient utility for such a situation [67, 227, 228]. Therefore, we designed and compared both interfaces. The speech-based audio navigation was designed on the basis of previous methods like Headlock [67]. A new shape-changing device to indicate the accurate direction in real-time was designed on the basis of previous research [227, 228]. We call the device the *"directional lever"* (see Section 6.3 and Figure 6.3 (2)).

6.2.3 Attachment Design

We designed the system as a set of components attachable to a standard rolling suitcase (Figure 6.1). Given the footprint of the required sensors, it will be difficult to make the system fully wearable in the near future. As an alternative, we expect such a system can be attached to daily mobile devices, such as a rolling suitcase, a shopping cart, and a wheelchair. Such a rolling suitcase with attachments can move naturally alongside a blind person, much like a guide dog or a sighted guide who walks side-by-side. Kayukawa et al.created a supportive system for blind people [115] and argued that this rolling suitcase form has four advantages: 1) for a blind user, a rolling suitcase can often act as an extended sensing mechanism for identifying changes in floor texture or as a form of protection from collisions with obstacles; 2) as a robotic sensing system, it also provides a convenient place to store and attach sensors, power, actuators, and computing resources; 3) users can walk with the system easily on flat spaces; and 4) the system can capture images without significant motion-induced blur. For these reasons, we also chose a suitcase form for our prototype system.

6.2.4 Navigation Interface Overview

The overview of the navigation process is as follows (see Figure 6.1). A blind user usually starts walking with the on-path mode. The user is instructed by the system to slow down when he or she perceives alert signals, sounds, or vibrations. The user can walk at normal speed again after the alert signal stops. This means the user's path is clear and safe.

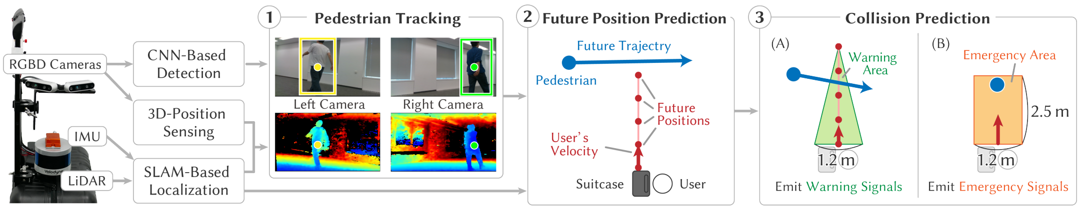


Figure 6.2: Overview of the on-path mode. 1) The system detects and tracks the pedestrians' position using two RGBD cameras and the SLAM-based localization method. 2) The system estimates the user's current position and velocity using the SLAM results and then predicts the user's future position. Next, 3) the system estimates two levels of collision risks and emits two types of alert signals (A: a warning signal and B: an emergency signal) via an audio or tactile interface.

If a pedestrian is not aware that the user is approaching and blocks the user's path, the system continues to emit alert signals. In such a situation, a user can push the start button on the handle to enable the off-path mode to avoid the standing pedestrians. The system automatically navigates the user in speech or with the directional lever. After returning to the user's usual path, the system automatically changes back to the on-path mode.

6.3 Implementation

In this section, we first describe our implementation of the system, which is characterized by its on-path and off-path modes, followed by the audio and tactile interfaces, respectively.

6.3.1 The On-path Mode

As shown in Figure 6.2, we attached a LiDAR sensor, an inertial measurement unit (IMU) sensor, and two RGBD cameras to a suitcase. The system uses these sensors to predict the risks of collisions. In what follows, we explain how the system predicts the risks of collision on a step-by-step basis.

Localization

During navigation, our system estimates the current location and direction of a user using a cartographer package[2] [93] of the robot operating system (ROS) [199]. The cartographer can localize by comparing the 3D pointcloud map previously generated and real-time scanning data from the LiDAR and IMU sensors. On the basis of the localization results, the system estimates what the user's velocity and position will be four seconds in the future (Figure 6.2 (2)).

Pedestrian Detection, Tracking, and Prediction

We use two RealSense D435 cameras[3] for tracking pedestrians to obtain a wide field of view. Each camera has an 69.4° × 42.5° field of view. By arranging the two cameras horizontally (Figure 6.2 (1)), we can obtain about a 135° × 42.5° field of view. To calibrate the relative position and orientation of each camera for LiDAR, we used an intensity-based LiDAR camera calibration tool [244]. Pedestrians are tracked using the following steps.

1. The system detects pedestrians using a YOLOv3 object detector [205]. The model is trained for detecting people using the publicly available COCO dataset [145].

[2]http://wiki.ros.org/cartographer
[3]https://www.intelrealsense.com/depth-camera-d435/

2. The system calculates the positions of detected pedestrians in camera coordinates. RGB-D images are used to calculate the 3D positions of the detected pedestrians in the camera coordinates.

3. The system calculates the positions of detected pedestrians using map coordinates. By using localization results, it compensates for camera motion and converts pedestrian positions into map coordinates from the camera coordinates.

4. The system matches detected pedestrians with tracked pedestrians. To match the detected pedestrians with the tracked pedestrians, we first use a Kalman filter [111] for each track to predict the positions in the next time step. We assume each person has a 1.0m circle size in the 2D map and calculate the intersection over union (IoU) for the detected and predicted circles. To find the best matches of tracked circles and detected circles using the IoU, we use the Hungarian algorithm [129].

5. The system estimates the velocity for each tracked pedestrian using a Kalman Filter.

These steps for detection (steps 1–3), tracking (step 4), and velocity estimation (step 5) are done using separate processes. The detection steps are done for each camera, and the detection results for each camera are merged in the tracking step. All these steps were done at about 4–5 FPS when we used a laptop computer (Intel Core i7-8750H CPU @ 2.20GHz, NVIDIA GeForce GTX 1080 Mobile GPU). On the basis of the estimated surrounding pedestrians' velocity, the system assumes that the pedestrians move at the constant velocity and predicts their positions four seconds in the future (Figure 6.2 (2)).

Collision Prediction

The system predicts the risk of a future collision on the basis of all the predicted positions of surrounding pedestrians and the blind user, then it decides whether or not to emit an alert signal. A collision is expected when a pedestrian's future trajectory crosses the "Warning Area" shown in Figure 6.2 (3) A. The system defines the area as triangular, its base length is $1.2m$, and the apex position is the user's predicted position in 4 seconds. The system can dynamically change the "Warning Area" in accordance with the user's velocity. For example, when the user is walking faster, the system predicts collisions with pedestrians in a larger area. When the user is walking slower, the system considers collisions in a smaller area. If the system detects the intersection between the pedestrian trajectory and the warning area, the system decides the user has a risk of collision and emits the low-urgency alert signals. In addition, we define the "Emergency Area" shown in Figure 6.2 (3) B. The system defines the area as a fixed-size rectangle, $1.2 \times 2.5m$. When a pedestrian is in the area, the system emits the high-urgency alert signals. In this case, we expect the blind user to stop immediately. The system estimates the risk of collisions in the "Emergency Area" by assuming that both the suitcase and the user face the same direction. Thus, in our user study, we asked blind participants to walk while keeping the suitcases in the direction they were heading.

6.3.2 The Off-path Mode

Path Planning

To navigate a user in the off-path mode, the system assumes that navigation maps include route information that is safe for blind users. In the following user studies, we assumed that the floor map had the positions of the tactile paving. Note that this study focused on navigating users on tactile paving, but other non-visual sensible landmarks, such as walls, can be used to define the blind users' path.

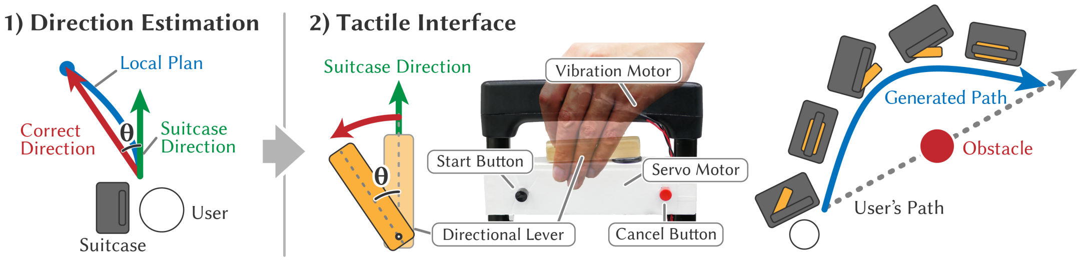


Figure 6.3: 1) In the off-path mode, our system calculates the angle (θ) between the correct direction and the suitcase direction. 2) The tactile interface has two vibration motors and a directional lever driven by a servo motor indicating the correct direction. The directional lever tells the user the correct walking direction in accordance with the generated path in the off-path mode.

To enable users to avoid standing pedestrians, the system sets the current user's position as the start position of the off-path mode, and it sets the position five meters ahead of the user on tactile paving as the goal position of the off-path mode. Then, the system plans a path to avoid the pedestrian(s) and return to the tactile paving. The path planning can be done with the navigation packages of the ROS, Navfn global path planner[4], and DWA local path planner[5] [69]. Navfn can generate a safe path avoiding obstacles such as pedestrians, walls, and static obstacles using surrounding structural information from the LiDAR sensor. When the system cannot generate a safe path, for example, in cases where no space is available for pedestrian avoidance, the system continues to emit the emergency alert to stop the blind user until a path becomes available.

Direction Estimation

Until the user reaches the defined goal position, the DWA local path planner shows a local plan with the trajectory that the user should follow at that moment. Our system uses the local plan to estimate the correct walking direction (θ in Figure 6.3 (1)). The correct direction is defined by a straight path to the end of the local plan. To guide blind users in the correct direction, the system calculates the angle between the correct direction and the suitcase direction estimated by the SLAM.

6.3.3 Interface for Navigation Instruction

Audio Interface

Our audio interface emits beeps to alert the user about the risk of collisions. We use a bone conduction headset to convey navigation information without impeding environmental sounds. Beeps have been used as a means to alert people about urgent situations, such as in aircraft [35], nuclear power plants [158], and hospital intensive care units [163]. Audio notifications can also alert drivers of an imminent risk of collision or assist in navigation [157]. The relationship between perceived urgency and sound parameters is well documented [59, 87, 157]. We prepared two types of beeps with different urgency levels denoted low and high. Specifically, we varied the pulse rate, the pitch, and the base frequency, as given in Table 6.1. The values we used are based on previous research addressing sound urgency [115, 192, 212].

In the off-path mode, the audio interface instructs the user on the correct walking direction via text-to-speech feedback such as *"Right," "Left,"* and *"Go straight."* These navigation commands were used in a prior navigation system called Headlock [67], which provides

[4] http://wiki.ros.org/navfn
[5] http://wiki.ros.org/dwa_local_planner

Table 6.1: Feedback patterns. PD: Pulse duration, IPI: inter-pulse interval, and BF: base frequency.

Interface	Urgency Level	PD	IPI	BF
Audio	Low	0.5 s	0.5 s	400 Hz
	High	0.1 s	0.1 s	1000 Hz
Tactile	Low	0.5 s	0.5 s	N/A
	High	Inf	N/A	N/A

information to navigate toward detected objects (e.g., direction and distance) via audio. We chose such simple navigation commands because a previous study revealed that blind people have difficulty adjusting their orientation slightly [216]. Our audio interface says *"Right"* or *"Left"* depending on the angle between the correct direction and the suitcase direction. When the absolute value of the angle is within 10 degrees, the system says *"Go straight."* These speech commands are emitted at one-second intervals.

Tactile Interface

To provide vibration feedback in the on-path mode, we attached two vibration motors (T.P.C., FM34F), which were connected to an Arduino Uno Rev3[6] on a suitcase handle (Figure 6.3 (2)). Studies have used vibration to alert people of emergency situations [74, 243], and the relationship between perceived urgency and vibration parameters has been shown [26, 142, 196]. Specifically, perceived urgency significantly decreases as the inter-pulse interval (IPI) increases. On the basis of previous research [26, 142, 196], we designed a low-urgency tactile signal that has an IPI and a high-urgency tactile signal that causes continuous vibrations (Table 6.1).

In the off-path mode, the system needs to convey the correct direction to avoid obstacles. The tactile interface has also been used to show directions to destinations [51, 159, 211, 224]. We designed the *"directional lever,"* which always indicates the correct direction (Figure 6.3). The directional lever is rotated by a servo motor (NEW TC, SE-A410) that is connected to an Arduino Uno Rev3. We attached the directional lever under the suitcase handle. As shown in Figure 6.3 (2), a user holds the suitcase handle with one hand and clamps the directional lever with the fingers. While the suitcase is facing the correct direction, the directional lever indicates ahead. When users should turn right (left), the lever indicates right (left) in accordance with the angle of the estimated correct walking direction (θ in Figure 6.3 (1)). For comfortable use, we clipped the angle range of the directional lever from $-30°$ to $30°$.

6.4 User Evaluations

We conducted a user study with 14 blind people. The main goals were 1) to evaluate if participants could avoid collisions using our guiding system, and 2) to evaluate the effectiveness of the tactile and audio interfaces. We asked the participants to walk a short route in a controlled environment and to walk a long route in a real-world environment.

6.4.1 Participants

As shown in Table 6.2, we recruited 14 blind participants (6m/8f) with ages ranging from 32 to 70 (Mean=50.43 and SD=10.04). Thirteen participants (P6.1–P6.13) regularly used a

[6] https://store.arduino.cc/usa/arduino-uno-rev3

Table 6.2: Demographic information on our participants and the SUS Score for each interface.

Demographic information				SUS Score (Grade)			
ID	Gender	Navigation Aid	Age	Audio		Tactile	
P6.1	Female	Cane	44	85	A+	77.5	B+
P6.2	Male	Cane	56	82.5	A	82.5	A
P6.3	Male	Cane	48	90	A+	92.5	A+
P6.4	Female	Cane	47	67.5	C	62.5	D
P6.5	Female	Cane	48	72.5	C+	72.5	C+
P6.6	Female	Cane	57	65	C	65	C
P6.7	Female	Cane	51	87.5	A+	87.5	A+
P6.8	Female	Cane	43	45	F	52.5	D
P6.9	Male	Cane	40	72.5	C+	75	B
P6.10	Male	Cane	70	57.5	C	47.5	F
P6.11	Female	Cane	32	77.5	B	70	C
P6.12	Male	Cane	55	75	B	72.5	C+
P6.13	Male	Cane	69	90	A+	95	A+
P6.14	Female	Dog (primary) and Cane	46	97.5	A+	97.5	A+
		Mean	50.43	76.1	B	75.0	B
		SD	10.04	13.7		14.6	

white cane, and one (P6.14) owned a guide dog. They considered themselves to have good orientation and mobility skills.

6.4.2 Tasks

In this study, we asked the participants to walk on tactile paving located in two types of environments: 1) controlled environments where one experimenter crossed or blocked a blind user's path, and 2) real-world environments where many different people were walking. The participants walked through these environments with either the audio or tactile interface.

Controlled Environments

To evaluate the effectiveness of our system in the same conditions across all participants, we first prepared a controlled environment that had a simple route of 16 meters with tactile paving. In that environment, one experimenter interrupted participants' walking. We prepared two conditions to evaluate the effectiveness of the on-path and off-path modes (see also Figure 6.4 (1)): A) one experimenter walked across the route at two points, and B) one experimenter blocked the tactile paving at two points. We asked the participants to walk the route four times (two interfaces × two conditions). In condition A, the experimenter started walking at the time at which they would collide with the participants. Each participant held the suitcase handle with one hand and used his or her white cane with the other. Their goal was to walk on the tactile paving while avoiding collisions with the experimenter. Participants started each task without knowing how the experimenter would behave (walk across

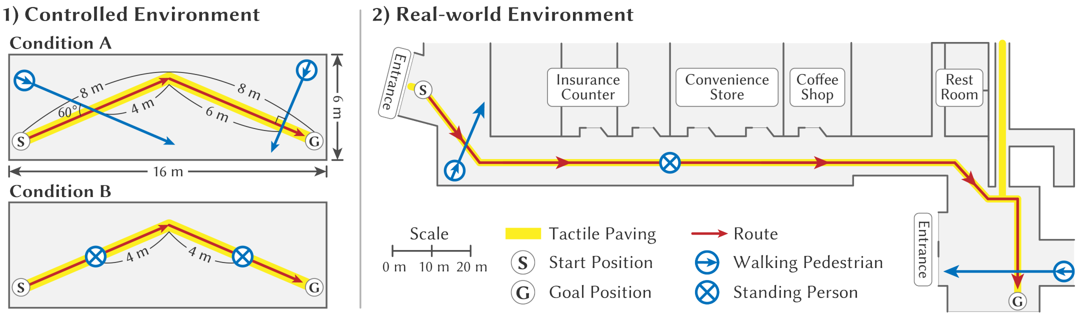


Figure 6.4: The routes used in our user study. 1) We first asked the participants to walk routes in a controlled environment. The routes included condition A, where an experimenter crossed the participants' route at two different points, and condition B, where an experimenter blocked the route at two different points. 2) Blind participants also walked a long route (around 180 meters) in a real-world environment. The route included two points where an experimenter crossed the participants' route and one point where an experimenter blocked the route.

or block their path). We asked participants to change modes on the basis of the feedback from the system.

Real-world Environments

In the study, we also asked blind participants to walk a long route (approximately 180 meters) on the ground floor of an office building (Figure 6.4 (2)). We selected this real-world environment because people constantly walk into or out of the office, restroom, convenience store, coffee shop, etc. When participants walked on the route, one experimenter crossed the participants' path at two points and blocked the path at one point (Figures 6.4 (2) and 6.5). Participants walked the long route twice using either the audio or tactile interface.

6.4.3 Procedure

We first provided an overview of the study and administered a questionnaire on demographics and navigation habits. We also surveyed the participants' opinions about using a headset to receive navigation instructions while walking alone. Next, we described the two modes for two types of interfaces to the participants and gave them a short training session (10 – 20 minutes) until they became familiar with each interface. We adjusted the volume in the audio interface to make sure it was comfortable but audible. During the training session, we explained how to hold the suitcase as it affects the accuracy of collision prediction.

For the first task, we asked the participants to walk the short route in the controlled environment while using either the audio or tactile interface. They walked the route four times while changing the interfaces and conditions (conditions A and B shown in Figure 6.4 (1)) in a counter-balanced order. For each trial, we measured the task completion time and counted the number of collisions between the participants and the experimenter. After completing the first task, participants took a post-questionnaire, which was audio recorded for further analysis. Specifically, we asked the participants to rate the following sentences using 7-point Likert items (rating from 1: strongly disagree to 7: strongly agree):

- **Q6.1:** *"The on-path mode with audio interface helped me avoid collisions while changing my walking speed.*[7] *"*

[7] All of the communications with participants were done in their native language. In this chapter, we describe any translated content in the form of "*translated content*".

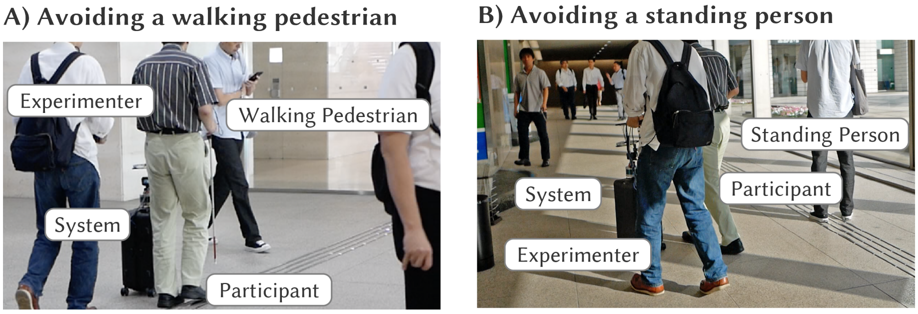


Figure 6.5: User study of a real-world environment. Participants walked on tactile paving while avoiding a walking experimenter (A) and a standing experimenter (B). One experimenter crossed the participants' path at two points and blocked the path at one point.

- **Q6.2:** *"The on-path mode with tactile interface helped me avoid collisions while changing my walking speed."*
- **Q6.3:** *"The off-path mode with audio interface helped me avoid pedestrians who blocked my path."*
- **Q6.4:** *"The off-path mode with tactile interface helped me avoid pedestrians who blocked my path."*

We also asked open-ended questions about the advantages and challenges of each interface.

The remaining tasks were performed in the real-world environment shown in Figures 6.4 (2) and 6.5. The participants had a training session to walk the route twice while using each interface to grasp the overall route. Then, they were again asked to walk the route twice while using either the audio or tactile interface. In each trial, we did not instruct the participants to use a particular mode. Instead, we asked them to change the modes by themselves on the basis of the alerts from the system. We informed the participants that a researcher would be walking behind them to guarantee their safety as well as other pedestrians' safety (Figure 6.5). The researcher did not intervene unless an imminent risk or a deviation from the path occurred. We counted the number of times imminent risks of collisions occurred for each condition. To observe the participants' movement and the response of pedestrians, we mounted a GoPro camera on the top of the suitcase during the study.

After completing all the tasks, participants took a post-questionnaire. The participants were again asked to rate the four sentences (Q6.1–Q6.4) they assessed after the first task in the controlled environment. In addition, we asked them to rate the following sentences about their preferred interface using 7-point Likert items (rating from 1: do not prefer to 7: prefer):

- **Q6.5:** *"I prefer the audio interface for navigation instructions."*
- **Q6.6:** *"I prefer the tactile interface for navigation instructions."*

They also rated the items of the system usability scale (SUS) [39]. Finally, we asked open-ended questions to gather feedback about their overall experience with each interface. The task process took around 45 minutes, while the whole experiment took approximately 90 minutes per participant.

6.5 Results

6.5.1 Audio Interface Usage

All participants commented that receiving auditory navigation commands via headsets is undesirable because these commands may interfere with ambient sounds. For that reason, 11 participants out of 14 mentioned that they never use a headset while walking:

> **A6.1:** *"When I'm walking alone, I always listen to various sounds such as footsteps and engine sounds. A headset may block these sounds, so I always walk without using one."* P6.3
>
> **A6.2:** *"When I use a headset to get navigation instructions, I always keep in mind that I need to stand in a safe zone, such as near a wall. If I receive auditory commands while walking through public spaces, I am distracted by these commands and have difficulty hearing ambient sounds."* P6.9
>
> **A6.3:** *"I sometimes use Google Maps and a headset to get instructions to the destination. While using the app, I make sure I stop walking."* P6.5
>
> **A6.4:** *"When audio commands come from a headset, I tend to concentrate on listening to them. In fact, I was once nearly hit by a car when I tried walking with a headset."* P6.10

Although seven participants (P6.5–P6.7, P6.9, P6.11–P6.13) mentioned that they sometimes use audio-based navigation systems, such as Google Maps, they also strongly agreed on the risk of using a headset while walking. In particular, three participants (P6.5, P6.7, and P6.9) mentioned that they stop walking while listening to audio commands from a navigation app.

6.5.2 Experience of Collision with Nearby Pedestrians

Thirteen participants out of 14 mentioned that they have collided or nearly collided with pedestrians in public spaces. The only exception was P6.14, who commented that her dog was very well trained at avoiding collisions. All participants except P6.2 and P6.14 also mentioned that they had experienced situations in which someone had blocked their path even on a tactile paving. P6.2 commented that he could recognize empty spaces with no standing pedestrians by listening to the ambient sounds.

6.5.3 Overall Performance

The Number of Collisions

In our controlled study, all participants reached the goal without collisions. In our real-world study, while participants did not encounter real pedestrians who blocked their path, several pedestrians crossed it. Table 6.3 reports the number of times the system emitted a warning or emergency alert. The participants were thus at risk of collision about 10 times in each trial. Blind participants could avoid such pedestrians by using the on-path mode. Most participants also successfully avoided the experimenter who crossed the participants' path at two points and blocked it at one point. However, two exceptions occurred. In these cases, participants had an imminent risk of colliding with the standing experimenter, and another experimenter had to ask them to stop. The reasons for these close calls were as follows: 1) P6.11 and P6.13 continued to walk without noticing the vibration alerts from the system; and 2) although the system generated a path to the left, the system told P6.7 to *"go straight,"* because the suitcase was facing the generated path (i.e., a difference occurred between the suitcase direction and the user's walking direction).

Table 6.3: The number of times the system emitted a warning or emergency alert.

Interface	Mean and SD	
	Warning Alert	Emergency Alert
Audio	12.6±5.7	4.1±2.2
Tactile	12.5±3.1	8.1±3.6

Table 6.4: Quantitative evaluations of the task completion time: Mean and SD: the mean and standard deviation of the task completion time; Lower and Upper: the lower and upper bounds of the 95% confidence intervals, respectively; and the p-value of the Wilcoxon signed-rank test (* indicates the significance found at the levels of 0.01).

Condition	Audio Interface			Tactile Interface			p-value
	Mean and SD	Lower	Upper	Mean and SD	Lower	Upper	
Pattern A (the on-path mode)	36.3 ± 5.85	32.8	39.8	39.4 ± 8.60	34.3	44.6	0.23
Pattern B (the off-path mode)	59.5 ± 9.97	53.5	65.5	53.6 ± 8.08	48.7	58.4	**0.015***

Task Completion Time

Table 6.4 reports the task completion time in terms of its mean and standard deviation, as well as 95% confidence intervals, obtained using each interface. The table also shows the p-value of the Wilcoxon signed-rank test. Our statistical analysis revealed that, in the off-path mode, participants who used the directional lever could avoid a standing person more quickly than those who used the speech-based audio interface. In the on-path mode, we observed no significant differences in the task completion time between the audio and tactile interfaces.

Video Observation

Video recordings enabled us to analyze the behavior of the blind user and sighted pedestrians to complement our quantitative analysis. We observed that participants could switch between the on-path mode and off-path mode on the basis of the feedback from the system. Participants mainly walked with the on-path mode and slowed down or stopped when the system emitted alert signals. When the system continued to emit alerts, they pushed the start button to enable the off-path mode. In the controlled environments, all participants could switch between the on-path and off-path modes successfully. In the real-world environment, many participants also could change the modes properly. However, as mentioned above, P6.11 and P6.13 who used the tactile interface continued to walk without noticing the vibration alerts and had an imminent risk of colliding with the standing experimenter.

6.5.4 Ratings of Our System

Figure 6.6 shows the participants' ratings: the effectiveness of the on-path mode with the audio interface (Q6.1) and tactile interface (Q6.2), the effectiveness of the off-path mode with the audio interface (Q6.3) and tactile interface (Q6.4), and the participants' preferences for the audio interface (Q6.5) and tactile interface (Q6.6). In the study, we asked Q6.1–Q6.4 after participants finished tasks in both the controlled and real-world environments. We

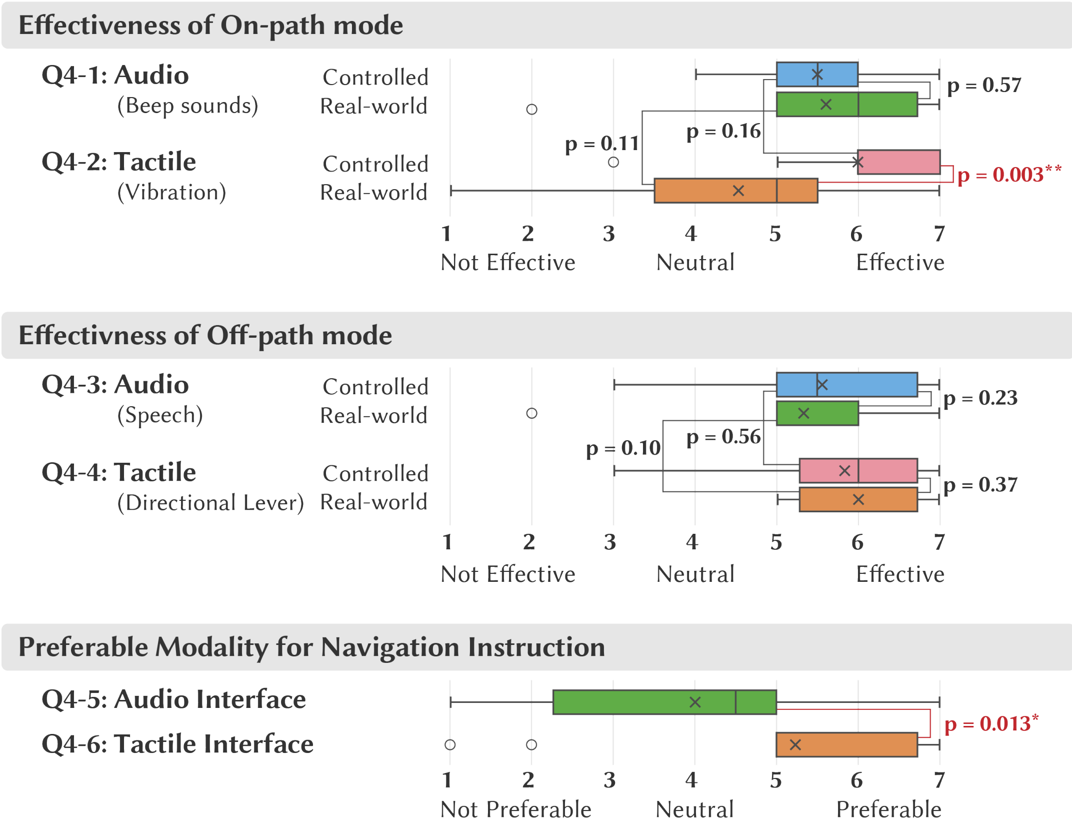


Figure 6.6: Box plots of users' ratings (Q6.1–Q6.6): Controlled and Real-world are ratings after tasks in the controlled environment and the real-world environment, respectively, and p: p-is the value of the Wilcoxon signed rank test done on each question (* and ** indicate the significance found at the levels of 0.01 and 0.001, respectively).

compared these questions using the Wilcoxon signed-rank test. The p-values of each test are shown in Figure 6.6. A comparison of answers to each question between the controlled and real-world environments showed a significant difference in Q6.2. The effectiveness of the vibration-based tactile interface significantly reduced after the participants used it in the real-world environment. In addition, we also observed a significant difference between Q6.5 and Q6.6. Our analysis found that participants had significantly stronger preferences for the tactile interface than the audio. In Q6.6, 12 participants rated the tactile interface higher than neutral on its navigation instructions (over 5 points). We observed no significant differences in the other tests.

Table 6.2 reports the scores of the system usability scale (SUS) for each participant. The mean (M) and standard deviation (SD) of the SUS score were 76.1 and 13.7 for the audio interface and 75.0 and 14.6 for the tactile interface, respectively. Participants who did not like our system mainly pointed out the difficulties with the suitcase form. We describe the feedback from them in a later section.

6.5.5 User Feedback on Our System

Overall Experiences

Participants generally agreed that the on-path mode was effective:

> **A6.5:** *"While walking alone, I always concentrate on grasping surrounding environments via auditory sensations to avoid collisions. Using the system alerted me of the risks of collisions, so I could walk more confidently"* P6.13

A6.6: *"The system emits no alert signals when my path is safe. Therefore, I could feel safe walking when no alerts were being emitted"* P6.8

A6.7: *"The system [the on-path mode] told me my path was clear and safe. So, I could walk faster while the system was not emitting alert signals."* P6.1

A6.8: *"In my workplaces [a braille library and a school for the blind], I sometimes collide with other blind pedestrians. Thus, I want to use the system in my office"* P6.11

We also got positive feedback on the off-path mode:

A6.9: *"When I encounter a group of people who block my path, I always ask them to move out of the way. By using the system, I could avoid such people by myself"* P6.12

A6.10: *"When I avoid obstacles, I sometimes lose my way and become disoriented. The system [the off-path mode] is useful because it can guide me back to my path"* P6.7

To generate an alternative path in the off-path mode, we used the navigation packages in ROS. The packages are standard for controlling robots, but P6.7 and P6.8 commented that the navigation was not smooth enough for human navigation:

A6.11: *"When I avoided pedestrians using the off-path mode, the system sometimes instructed me to alternate between going toward the right and left. Thus, I felt I lost my position and direction"* P6.7

A6.12: *"In the off-path mode, the system provided me with the same direction repeatedly, and it made me confused"* P6.8

Nine participants provided negative feedback on the suitcase form. We present our examination of this feedback in the discussion section:

A6.13: *"This suitcase-shaped system was too large and heavy for daily use"* P6.10

A6.14: *"When I walk with a cane, I want to keep my other hand free"* P6.6

A6.15: *"This system is large and I'm afraid of the additional risk of hitting it against other pedestrians, especially in a very crowded area"* P6.8

Three participants (P6.7, P6.13, and P6.14) also mentioned the need to shrink the tactile interface:

A6.16: *"If the tactile interface can be attached to my cane, I want to use it every day"* P6.13

A6.17: *"I'd be so happy if the tactile device were made small and lightweight enough for me to hold it in my hand"* P6.7

Audio Interface

All participants mentioned that the audio feedback was easy to recognize:

A6.18: *"Audio feedback is clearer than tactile feedback. Thus, I could respond to it quickly"* P6.5

A6.19: *"I could distinguish beep sounds easily because these sounds were characterized by not only the pulse rate but also the pitch"* P6.4

However, they also reported that the audio interface was disadvantageous for sensing ambient sounds:

A6.20: *"I walk while getting surrounding information using my ears, so I don't want to use a headset while walking if at all possible. In particular, when I heard auditory instructions through the headset, I had difficulty listening to footsteps"* P6.9

A6.21: *"When I was using the audio system in the quieter environment [the controlled environment], I could easily hear both ambient sounds and audio signals. However, in the noisy environment [the real-world environment] where there are many ambient sounds such as footsteps, it was necessary to use extra awareness to hear ambient sounds"* P6.5

A6.22: *"In the real-world environment, where many people were walking, there was a larger amount of information from ambient sounds than in the controlled environment. So, I had difficulty recognizing both ambient sounds and audio-based feedback"* P6.11

A6.23: *"Text-to-speech feedback distracted me more than beeping sounds"* P6.10

A6.24: *"The instructions from the system were simple. So, I could distinguish instructions even if I used the tactile system. I think it is excessive to use the audio system to convey such simple instructions"* P6.2

Tactile Interface

All participants appreciated that the tactile interface did not interrupt auditory sensations:

A6.25: *"I could easily grasp the surrounding environment and collision risks by simultaneously using my cane, my ears, and tactile signals from the system"* P6.8

A6.26: *"Audio-based feedback interfered with my auditory sensations, but I could use both the auditory sensations and tactile-based instructions. The tactile interface gave me an additional sensing modality"* P6.11

However, some participants mentioned that the vibration alerts were difficult to recognize:

A6.27: *"I had to distinguish between two vibration patterns by considering the pulse duration, so it was harder than beep sounds with changing pitch"* P6.5

In addition, eight participants commented that the effectiveness of the vibration depended on the surface of the floor:

A6.28: *"The ground of the real-world environment had a rough surface, and the handle of the suitcase also vibrated. I had difficulty distinguishing between the suitcase vibration and the vibration alerts"* P6.9

A6.29: *"[In the real-world environment,] I had to concentrate on recognizing the vibration alerts because the suitcase vibrated due to the unevenness of the floor"* P6.6

The blind users provided positive comments on the directional lever. In particular, all the participants mentioned that the directional lever was effective in both the controlled and real-world environments:

A6.30: *"The directional lever helped me adjust the walking direction because it indicated the correct direction directly"* P6.8

A6.31: *"The lever could tell me the direction more precisely than auditory commands could. If the system were to say precise directions, like 'turn right 32 degrees,' it would be time-consuming and annoying"* P6.3

A6.32: *"The directional lever was not affected by the surface of the ground and always worked effectively"* P6.7

However, three participants (P6.1, P6.2, and P6.10) mentioned that the navigation lever took time to get used to:

> **A6.33:** *"The directional lever took a while for me to get used to because it indicated the direction too precisely"* P6.2

6.6 Discussion

6.6.1 Effectiveness of The Guiding System

Both the controlled and real-world studies showed that our guiding system was effective for blind users to prevent collisions with pedestrians. Most participants successfully avoided collisions with nearby pedestrians by using the on-path and off-path modes properly. Feedback from the participants also supported the effectiveness of our system (A6.5–A6.10). They appreciated that our collision warning system enabled them to walk more confidently and with a more secure feeling than they usually do in daily life (A6.5–A6.7).

6.6.2 Audio Interface

We observed that participants listen for ambient sounds to ensure their safety while they walk through a public space. All participants commented that receiving audio feedback frequently while walking in the real-world situation was not usable because it interfered with their ability to make out useful ambient sounds such as footsteps of other pedestrians and echoes from walls (A6.1–A6.4, A6.20–A6.24). In particular, two participants commented that they did not want to use the audio system in the real-world environment because of the rich ambient sounds they needed to listen for (A6.21 and A6.22). Although the participants mentioned that they could recognize the sound alerts more clearly than the vibration alerts (A6.18 and A6.19), they had significantly stronger preferences for the tactile interface (Q6.5 and Q6.6). This result is understandable since the footsteps and other sounds from other pedestrians are faint and can be easily masked by other ambient sounds or computer-generated navigation commands. In echolocation, changes in frequency and amplitude of low-frequency sounds need to be detected to recognize changes in echoes [128]. Such recognition requires trained abilities, which vary among blind people (see 1. Introduction). From this study, audio interfaces are not a promising interaction method for navigation tasks. We believe more studies should be done to seek better interaction methods for the diverse abilities of blind people by combining audio, haptics, and other non-visual media.

6.6.3 Tactile Interface

Advantages

We observed two advantages of the tactile interface. First, the users could recognize the vibration alerts and the correct direction indicated by the lever while listening to ambient sounds (A6.25 and A6.26). Second, the directional lever enabled the participants to avoid standing pedestrians significantly faster than the speech-based audio interface did. The participants also commented that the directional lever could indicate the correct direction directly (A6.30 and A6.31), and all participants commented that the lever was not affected by the surface of the ground (e.g., A6.32). Therefore, they revealed stronger preferences for the tactile interface than for the audio interface (Q6.5 and Q6.6).

Disadvantages: Vibration Alerts

Some participants expressed some concern about recognizing the vibration alerts. In the controlled environment, all participants successfully avoided collisions using the vibration alerts. However, in the real-world environment, two participants sometimes could not recognize the vibration alerts and had imminent risks of collisions. In the controlled environment, the floor was carpeted, and the vibration alerts were clear. However, in the real-world environment, the floor was tiled with a rough surface, and the vibration alerts were mixed with the vibration of the suitcase. Eight participants commented that the effectiveness of the vibration alerts was affected by the floor texture (e.g., A6.28 and A6.29). These results indicate that the vibration alerts are affected by the type of floor surface. The participants provided significantly lower scores for the effectiveness of the vibration alerts in the real-world environments than in the controlled environments (Q6.2).

The one possible solution to overcome this limitation is to attach vibration motors to a user's body, such as on a wrist or fingers. All participants mentioned that the directional lever was always useful and effective in both environments (e.g., A6.32). This may suggest that the directional lever in a shape-changing interface was effective for the guiding system and that this interface is better than vibration signals for indicating alerts.

Participants were asked to adjust their walking speed on the basis of the pulse duration of the vibration, but it was hard for some participants (A6.27). Another possible solution is to equip brakes on the wheels of the suitcase and to control the walking speed using the physical feedback from the suitcase handle.

6.6.4 Form-factor of the System

We designed the system to be attached to standard luggage like a rolling suitcase. Rolling suitcases can be seamlessly used in public spaces, are well designed to walk with when holding the device, and can have all necessary sensors and tactile devices mounted on them, and enable images to be captured without significant motion-induced blur from the mounted camera (see also 6.2.3 Attachment Design above). Eight out of 14 participants mentioned that the suitcase-shaped system might be too heavy for daily usage (A6.13–A6.15). This suggests that the device will be accepted when they use a suitcase (or any other similar luggage) for other purposes, and then attach the system as assistive technology. At this moment, our solution is comprised of a depth camera and a laptop, which increase the weight and reduce the available space (for luggage) in the suitcase. This situation usually happens when a new assistive technology is developed [38, 191]. We expect that the size and weight of both the sensors and processors will be decreased as the device and communication technologies are systematically improved, enabling cloud-based computational power in the near future.

Another result related to the form is the possibility of mobile devices with total functionality. The directional lever was well accepted by all participants, and some commented that they would want to use the device on a daily basis if it were made the size of a mobile device (A6.9 and A6.10). We have to overcome technical challenges to enable such a form, but our results suggest a high possibility of utilizing mobile shape-changing devices.

6.6.5 Integrating a Guiding System to O&M Training Methods

Some participants commented that the experiments were their first time they were able to avoid pedestrians by following the instructions from a system (A6.9 and A6.10). Also, while all participants could learn how to use the system after a short training session (10 – 20 minutes), three of them commented that it took time to get used to the directional lever (A6.33). We asked participants to hold the suitcase handle with one hand and the white cane with the other. One participant (P6.6) commented that she want to keep her other hand free (A6.14). The current orientation and mobility (O&M) method is based on a white cane as the primary tool, and it uses all possible senses to understand the situations. Skill is required to recognize non-visual landmarks and to navigate safely in public spaces. Therefore, all blind users are strongly recommended to take O&M training [247] when they start walking independently. The use of systems for O&M is an uncharted territory for not only an individual blind user but also the entire community who supports O&M for the blind.

In the near future, we should share the results and our experience with the community and discuss how to build new O&M methodologies by fully utilizing both traditional navigation aids and new technologies, including the system we introduced in this study.

6.6.6 Autonomous Guiding Systems

The off-path mode requires a guiding system to navigate users accurately. A future solution for such navigation needs can be autonomous guiding robots [23, 24, 71, 85, 131, 230, 235]. Our non-autonomous guiding system has the advantage of allowing users to control their speed voluntarily even during an off-path situation. The directional lever successfully achieved sufficient accuracy for the off-path navigation. We believe that our guiding interface and autonomous guiding robots will complement each other to broadly satisfy blind users' needs, such as a variety of mobility skills, familiarity with a target public space, the density of crowds, and the preferences of users.

6.6.7 Other Guiding Situations in Public Spaces

We focused on two typical situations for this study: walking and standing pedestrians. No comprehensive list for such situations has been reported in previous studies to the best of our knowledge, but imagining other situations is not difficult. For example, a blind user may have difficulty in following a queue to get on a train car at a station, walking together with sighted surrounding people in the same direction, and walking through an extremely crowded public space. Such situations are beyond the scope of this study, and further research is required to cover a comprehensive set of situations.

6.6.8 Comparing with Traditional Navigation Aids

This study was not designed to compare the proposed system with traditional navigation aids such as a cane or guide dog, because traditional methods are very challenging in public spaces. Previous studies reported that blind pedestrians using only traditional methods had difficulty avoiding collisions with nearby pedestrians [2, 58, 115, 242, 243]. In fact, 13 out of 14 participants in our study mentioned that they had collided with other pedestrians while walking. On the other hand, we believe that it would be informative for researchers and developers to understand how blind pedestrians behave when using traditional methods in public spaces as the baseline. Therefore, we plan to measure, evaluate, and create a model of such behaviors and compare traditional methods with other new navigation methods, including our system.

6.7 Conclusion

We presented a guiding system equipped on a rolling suitcase to help blind people walk in public spaces seamlessly with nearby pedestrians. The system recognizes and predicts surrounding people's behavior and predicts the risks of collisions. The system then recommends the user to adjust his or her walking speed (the on-path mode) or to take an alternative path around pedestrians (the off-path mode). We implemented tactile and audio interfaces and conducted a user study with 14 blind participants. The results revealed that blind users could successfully avoid pedestrians using both interfaces; the tactile interface for the off-path mode guided blind participants significantly faster than the audio interface could; and the sound-based audio interface was easier for recognizing alerts than the vibration-based tactile interface. Overall, blind participants believed that a tactile interface could be effective because it did not block ambient sound. In future work, we will further research new tactile interfaces to eliminate the weakness of vibration alerts by focusing on shape-changing interfaces. We also plan to collaborate with orientation and mobility

(O&M) communities for building new O&M methodologies with technologies by sharing our results and experiences.

Acknowledgement

We would like to thank all participants who took part in our user study. We would also like to thank the anonymous reviewers for their helpful comments. This work was supported by JSPS KAKENHI Grant Number JP20J23018, Grant-in-Aid for Young Scientists (Early Bird, Waseda Research Institute for Science and Engineering), and JST-Mirai Program Grant Number JPMJMI19B2.

Chapter 7

LineChaser: Smartphone-Based Navigation System for Standing in Line[1]

> *"I felt that I do not need my cane anymore, as with this system I can maintain a certain amount of distance from the target. I think this system is revolutionary because I usually find a line by asking strangers, but with this system, I do not have to rely on others."*
>
> Participant 6
> Research Innovation Center, Waseda University
> August 30, 2020

7.1 Introduction

People often need to stand in waiting lines in public spaces in daily life, such as at cashier stations, bus stops, and check-in-counters at airports. This activity is challenging for blind people due to their lack of vision. People are first required to find the end of a line, a position that dynamically changes over time. It is difficult for blind people to find the end of line using information gathered by their senses (e.g., auditory cues) or tools (e.g., canes), and thus depend on the assistance of nearby people. Recent research has proposed high-accuracy indoor navigation systems to help blind people reach a destination based on static topological route maps and localization techniques [33, 44, 47, 62, 72, 85, 126, 144, 172, 175, 189, 207, 215, 259]. These systems can help users reach the *fixed* entrance of a waiting line area, but they did not focus on navigating to a changing end-of-line position.

Once a person joins a line, the next step is to follow the person in front of them as they move intermittently. Lines in public spaces are not always straight, as is the case, for example, in the long serpentine lines at airports. It is challenging for blind people to maintain a consistent distance from the person ahead of them only with their senses and tools. The larger distance currently required to comply with social distancing during the COVID-19 pandemic has increased this challenge [75]. Recent research has aimed to help blind people avoid collisions with pedestrians [85, 115]. These systems can detect positions and movements of nearby pedestrians by using computer vision technologies, but they have not been applied to line navigation.

We first developed a prototype system by focusing on the task of tracking and following a line, and recruited six blind people to test the system as a preliminary user study. The prototype system is capable of detecting and reporting the distance to the person in front of the blind user continuously using only a smartphone with an RGB camera and an infrared depth sensor. The sensing results are used to alert the users with three levels of distance information via vibration patterns to allow users to start moving forward and stop in a synchronized manner with the other people. All participants commented that they have

[1]Project Page: https://wotipati.github.io/projects/CHI2021_LineChaser/CHI2021_LineChaser.html

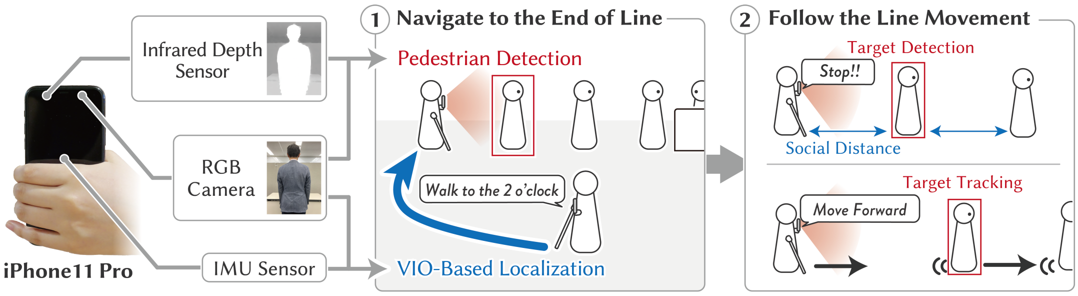


Figure 7.1: Overview of LineChaser. LineChaser helps blind people stand in waiting lines in public spaces by using only off-the-shelf smartphones. The system guides the blind user to the end of the line, and then helps them follow the line movement.

had previous experiences where they did not sense that the line was moving or, conversely, did not notice the line had stopped and bumped into the person in front of them. The prototype system enabled blind participants to detect the movement of the line and thus stand in lines with increased confidence. However, we observed several situations where the participants followed a wrong person who happened to be standing in front of the *target* (the correct person to follow).

We used feedback on the prototype system to design a smartphone-based assistive system called *LineChaser* (Figure 7.1). The system enables a blind user not only to follow line movements but also to find the end of a line. The system uses a topological route map that contains the line information, such as the place where pedestrians usually form a line. LineChaser first guides the blind user to the end of a line by using the map and a localization method with the smartphone (Figure 7.1, action (1)). To navigate, the system uses the smartphone's built-in RGB camera and infrared depth sensor, respectively, to detect nearby pedestrians and estimate their 2D positions on the map. According to the position estimation, LineChaser determines whether pedestrians are standing in a line. After guiding the user to the end of the line, LineChaser detects the last person in line (we call this person the target), and tracks the target based on the color histograms and positions of detected pedestrians (Figure 7.1, action (2)). LineChaser then uses the sensing results to instruct blind users to advance to the front of the line by moving in the right direction at the right time. We also take account of "social distancing," the distance to be maintained from other people to prevent possible infection with COVID-19. Our interview revealed that many blind people cannot maintain social distancing, as complying to it impedes blind people from having the target from aural sensing area, making it difficult to follow the target. Therefore, the system is also designed to maintain proper social distancing from the target.

To understand the usability of our system, we conducted a second study with 12 blind people. In this study, we prepared two types of lines (straight and serpentine) and asked blind participants to find the end of the line and follow the line movement with LineChaser. The results show that all participants were successfully able to both find and follow lines while maintaining social distancing. Our questionnaire results suggest that blind people face everyday difficulties when standing in lines. Also, blind participants significantly increased their confidence in standing in lines after using LineChaser, comparing to their daily experience. We also discuss future requirements to further improve LineChaser and possibly integrate it with other systems that provide day-to-day assistance for blind people.

7.2 Related Work

We specifically focus on supporting blind people to stand in line by using a smartphone. In this section, we review related work specific to this chapter that were not detailed in

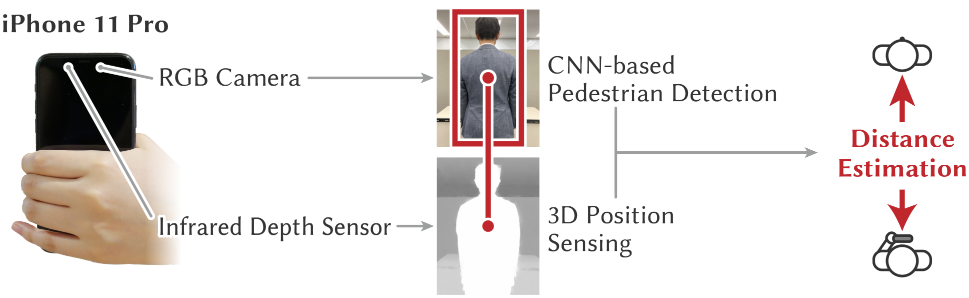


Figure 7.2: Overview of preliminary prototype system. The system uses an off-the-shelf smartphone to detect pedestrians and estimates the distance to them.

Chapter 2.

7.2.1 Smartphone-based Assistance System

Sighted assistance system such as Aira [10] and BeMyEyes [29] can be considered as one solution. While it is possible to use such services for the line standing task, it is also helpful to offer non-manual solutions to improve their independence. These services require human labor for assistance and thus may not be scalable when many users require assistance, may have limitation for service cost reduction, and may not be easy to provide round clock service for worldwide clients.

Computer vision-based systems have become popular among blind people. Seeing-AI [165], Or-Cam [182], TapTapSee [49], Aipoly [9] and Envision [42] are examples of such. These systems are able to recognize and read printed letters and even provide simple captions to pictures taken by the blind user, but they are not designed to detect surrounding people with sufficient accuracy for blind people to follow a line. It is necessary to detect the positions of surrounding pedestrians and convey suitable distance and direction information to enable line standing tasks. Accordingly, our system adopts the depth sensor of an iPhone 11 Pro and object detection system to obtain the positions of surrounding pedestrians.

7.2.2 Robots that Stand in Line

Nakauchi et al.designed a robot that finds the end of a line and follows the person in front of it [176]. Given the starting point and rough shape of the line, the robot eventually finds the end of the line by scanning each person in the line from the front of the line until no person is detected. This algorithm is effective for a robot as it is capable of changing its' orientation frequently. However, frequent instructions to change their path and orientation might disorient blind people [117]. We therefore implemented a system to support the standing-in-line task with a single smartphone and investigated the suitability of the interface and navigation method for blind people by conducting a user study.

7.3 Preliminary Study

We conducted a preliminary study with six blind people to understand: (1) their daily experience and challenges while standing in line; and (2) how to design a smartphone-based system that helps blind people stand in line. We implemented a prototype system that helps blind users to sense line movement, and asked blind participants to use the system.

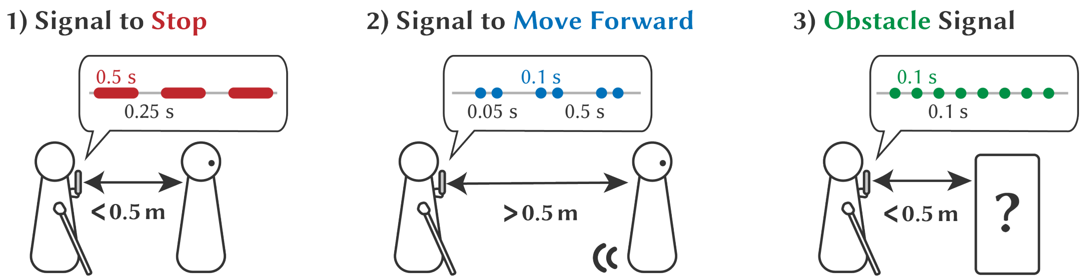


Figure 7.3: The system emits three types of vibration alerts to provide position information to users.

7.3.1 Prototype System

Pedestrian Detection and Distance Estimation

We developed a smartphone-based system that can detect surrounding people and inform about the distance to the closest person (Figure 7.2). This system intends to complement blind users' orientation and mobility skills in a social context, allowing them to stand in lines by themselves.

We used an off-the-shelf smartphone, iPhone 11 Pro[2], which is equipped with RGB image sensors and an infrared depth sensor. The system first detects pedestrians from the RGB streams by using a vision-based object recognition engine, YOLOv3-tiny [205], which recognizes "human" as an object type. Then, it automatically generates bounding boxes for all humans and selects the largest one as the target. The distance to the target is estimated from the depth data at the target's central position (Figure 7.2). All these estimation steps were done at about 2–3*fps*. We checked that the system could estimate the distance between 0.2 m and 6 m with reasonable accuracy.

Vibration Alerts

To convey distance information, we relied on vibration alerts, as audio might be less effective in noisy, crowded environments [36]. The system emits three types of vibration alerts:

1. **Signal to stop:**
   This signal indicates that another person is standing within 50 cm of the user and that the user should stop moving. We used a long vibration alert (pulse duration (PD) of 0.5 s and inter-pulse interval (IPI) of 0.25 s) (Figure 7.3 (1)).

2. **Signal to move forward:**
   This signal indicates that a person is standing in front of the user at a distance greater than 50 cm. The signal is used to prompt the user to step forward and uses a two-pulse vibration. (Figure 7.3 (2)).

3. **Obstacle signal:**
   This signal indicates an imminent risk of collision with any obstacle (pedestrian, desk, or wall) located less than 50 cm away. We used a short vibration alert (PD and IPI of 0.1 s) as the signal (Figure 7.3 (3)).

The absence of vibration indicates that the user has lost the target. In that case, users should scan the environment with their phone to find them. We set the alert distance threshold to 50 cm, because the distance of personal space while standing in line is around 40–80 cm [176] and users hold the smartphone in front of them.

[2] https://www.apple.com/iphone-11-pro/

Table 7.1: Participants' demographic information, their reaction time, and values for SUS scores.

Demographic info					Reaction Time	SUS Score	
ID	Age	Gender	Eyesight	Navigation Aid	(seconds)	(Grade)	
P7.1	22	Male	Blind	Cane	2.97 ± 0.68	77.5	B+
P7.2	33	Female	Blind	Cane	6.73 ± 5.48	37.5	F
P7.3	33	Female	Blind	Cane	4.39 ± 2.25	80	A-
P7.4	22	Male	Blind	Cane	3.50 ± 1.38	87.5	A+
P7.5	24	Male	Blind	Cane	2.43 ± 0.30	97.5	A+
P7.6	23	Male	Blind	Cane	2.18 ± 0.50	90	A+
				Average (Mean ± SD)	3.55 ± 2.66	78.3 ± 21.3	B+

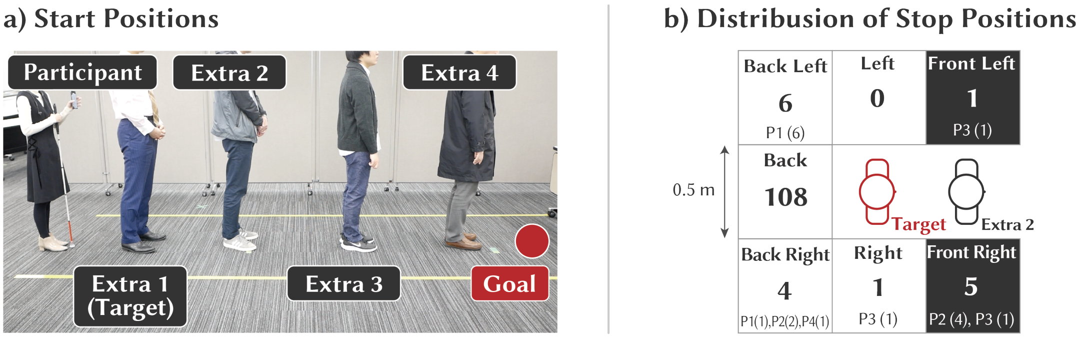


Figure 7.4: a) Starting positions of each line condition and b) Distribution of positions where blind participants stopped.

7.3.2 User Evaluation

To evaluate the effectiveness of our system, we performed a user study with 6 blind people (Table 7.1).

Tasks and Conditions

All tasks started with five people in line. The blind participants were asked to follow a line formed by four sighted people (hereafter called *extras*) in front of them (Figure 7.4 (a)). They were asked to stand in line and proceed until reaching the reception desk (the goal). A researcher signaled for the extra standing in the front position to leave the line after 30, 60, or 90 s. Waiting times were randomized for each extra and trial. Each blind participant held a smartphone with one hand and used their cane with the other hand. We stopped the task if the participant overtook the target in front of them, also referred as extra 1 in Figure 7.4 (a). We designed two types of organized straight lines. In condition C1, four extras moved one by one. In condition C2, two randomly selected extras in consecutive positions – extras 1 to 4 in Figure 7.4 (a) left the line together (i.e., one extra left the line at the same time as the person in front of the extra left the line). Condition C2 was designed to evaluate the response to irregular line movement. To simulate a crowded public space, we played ambient noise recorded at a shopping mall at 60 dB [208].

Procedure

We performed a pre-questionnaire where we asked participants about their prior experiences and challenges of standing in lines. We also asked them to rate a set of statements

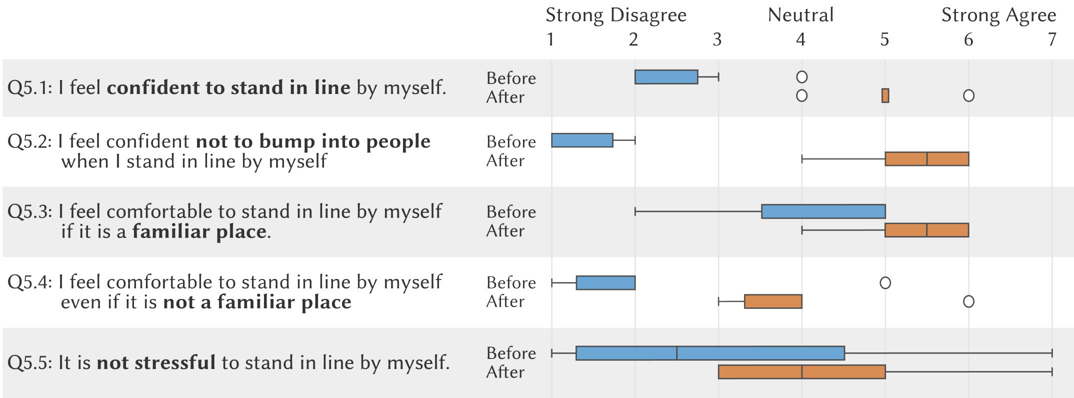


Figure 7.5: Likert scores and summary of responses before and after the preliminary experiment.

(Q7.1–Q7.5 in Figure 7.5) using 7-point Likert items (from 1: strongly disagree to 7: strongly agree). A training session of 10–15 minutes was then given to participants. During the training session, we explained how to hold the system as it affects the accuracy of the distance estimation.

Then, participants performed six trials where they stood in line using our system until reaching the reception desk. The order of the line conditions was randomized for each participant (three C1, three C2). After the trials, participants were asked a set of questions to gauge their confidence and sense of comfort (Q7.1–Q7.5), to rate the system on a system usability scale (SUS) [39], and were also asked open-ended questions to gather qualitative feedback.

Metrics

In order to better understand how the line moved, we considered both timing and position. For timing, we measured the *reaction time*, which is the duration between the movement of an extra and the movement of the following extra or blind participant. We then compared the reaction time of the blind participants and of the sighted extras. For position, we defined seven positions around a target person as the *stop positions* of a participant (Figure 7.4-b). Ideally, the participant stops right after (*Back*) the Target, but a slight deviation to either side is also acceptable (Back Left or Right).

7.3.3 Results and Discussion

Past Experiences and Opinions about Standing in Line

Participants reported that they stood in line to check out at stores and cafes (5 participants out of 6), to use a bus stop (5), to get on subway (3), to use a ticket counter at airports and stations (2), to use ticket-vending machines (1), and to use a restroom (1). Most participants reported trying to cope with standing in lines by their intuition with ambient sounds (P7.1, P7.2, and P7.6), asking people in line for help (P7.3), or touching the clothes of the person in front (P7.5). P7.4 reported that he does not stand in line by himself.

Despite their various strategies for standing in line, all participants reported occasions during which they did not realize the line was moving or bumped into the person in front of them:

> **A7.1:** *"In noisy places such as shopping centers and stations, it is hard for me to recognize when the line was moving. Even if I noticed the line movement, I can't*

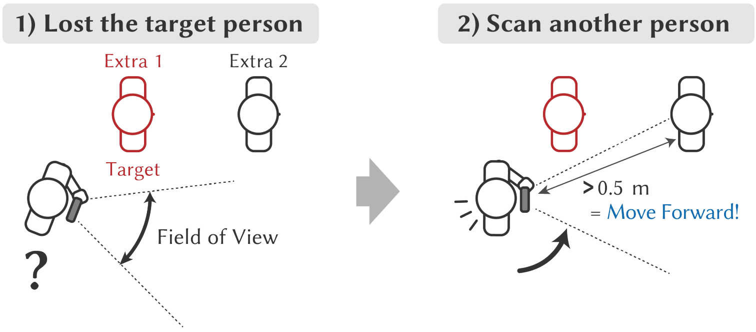


Figure 7.6: Example of a failure case where a participant overtook the target.

sense the distance to the person in front."[3] P7.6

A7.2: *"Lines are not always straight. For example, when I stand in a serpentine line, it is difficult to determine the direction I should walk in."* P7.1

Overall Performance

Figure 7.4 (b) shows the distribution of positions where participants stopped after each line movement. While the maximum trials in the study are 126 trials, we stopped the experiment as there was an occasion where P7.2 in C2 stood in front of the target person. As a result, the actual total number of trials was 125. Overall, participants successfully stopped just immediately behind the target (86.4%, 108 out of 125). We also noted that each blind participant tended to shift to a specific side during each line movement. The success rate for stopping at the correct position was 94.4% (118 out of 125) after including slight deviations to the side. The task success rate for line conditions C1 and C2 was 75% (18 out of 24), because we stopped each trial after one failure. Specifically, participants P7.3 and P7.2 overtook the target two and four times, respectively. Figure 7.6 (c) shows an example of a failure case. When the user shifted to one side and lost the target, the user scanned the area to relocate the target (Figure 7.6 (1)). However, in this case, the system detected another person and the user miss-tracked the target. (Figure 7.6 (2)). As a result, the system prompted the user to move forward, even though the actual target was standing on the user's left.

Table 7.1 reports the reaction time for each participant. The mean (M) reaction time was 3.55 seconds (SD: 2.66 sec.) for blind participants and 1.23 sec. (SD: 0.40 sec.) for the target person, significantly different by using Welch T-Test ($p < 0.001$, 95% CI: 0.95~1.49).

Subjective Ratings

Figure 7.5 shows the post-questionnaire results, in which most participants reported feeling more confident and comfortable standing in line after the experiment (with the system) than before (without the system). For questions Q7.1–Q7.4, all participants except P7.3 for Q7.3 increased their scores after the experiment. Four participants (P7.2–P7.5) also increased the score of Q7.5. Table 7.5 also reports the SUS scores [27] for each participant. The mean SUS score was 78.3 (SD: 21.3), which can be classified as "acceptable". P7.2 was the only participant with a lower score, mainly due to difficulties to hold the smartphone.

Qualitative feedback

Participants generally agreed that the system allowed users to start and stop moving forward at the right time, as illustrated by their comments:

[3] All of the communications with participants were done in their native language. In this chapter, we describe any translated content in the form of *"translated content"*.

A7.3: *"The biggest advantage of the system is that I could easily recognize the movement of a step forward from the person in front."* P7.1

A7.4: *"By using the system, I could decide when and how far I should move forward."* P7.4

A7.5: *"The system provides information on the distance, so it can reduce risks of collisions."* P7.2

Some participants provided positive feedback on our smartphone-based interface:

A7.6: *"The system is implemented on a smartphone. This is a strong advantage since it means I don't have to carry extra devices."* P7.1

A7.7: *"The system was simple and easy to use."* P7.4

In contrast, P7.2 commented that keeping the position of the system while waiting in line was difficult due to the large and heavy system:

A7.8: *"This smartphone is big and heavy, so it was difficult for me to hold the smartphone stably."* P7.2

The vibration alerts received positive feedback overall:

A7.9: *"I could distinguish vibration patterns easily. I like tactile feedback more than audio because tactile-based alerts do not block ambient sounds."* P7.2

Still, P7.3 suggested to use sound-based alerts rather than vibration:

A7.10: *"It was a little difficult to distinguish between the three types of vibration. I think that using audio cues can be a good idea."* P7.3

When asked for suggestions, two users mentioned that the system should provide more detailed distance information or the directional information of the target:

A7.11: *"When I lost track of the target, I had to relocate the target by myself while changing the direction of the system. I want to know in which direction the target is standing beforehand."* P7.4

A7.12: *"I like to know more detailed distance information. It could be a good feature to be able to change the pulse duration continuously to encode distance information."* P7.1

As another concern about our system, P7.2–P7.4 and P7.6 pointed out that they feel uncomfortable to point a smartphone to other people:

A7.13: *"My concern is that others may be wondering why I'm pointing the smartphone at people."* P7.3

A7.14: *"To turn the touch screen to other people may seem strange to surrounding people."* P7.6

7.3.4 Findings

We summarize our findings as follows:

1. Blind people face difficulties when standing in line to accomplish daily living tasks at subways, stores, cafes, and other public places, even when they use a cane and auditory senses.

2. Our prototype system allowed blind participants to follow the line movement successfully with increased confidence.

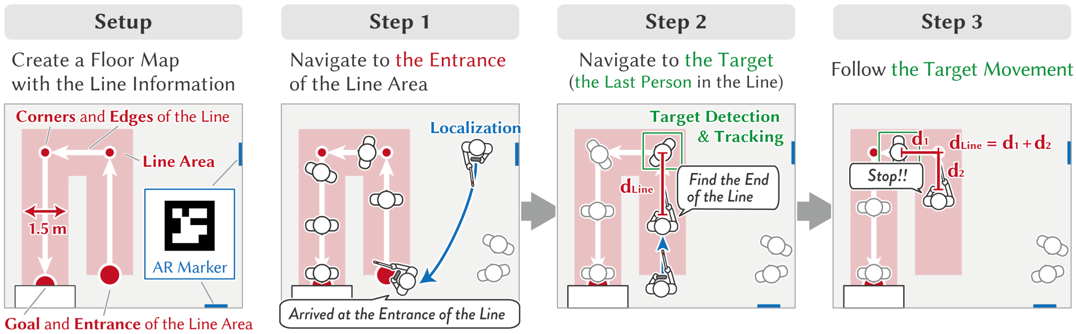


Figure 7.7: Overview of the navigation strategy. LineChaser uses a floor map that includes the line information, such as the line area, the entrance of the line area, and the corners of the line. (1) The system localizes the user's position via ARKit and AR markers placed on the wall, and guides the user to the entrance of the waiting line area. (2) Then, it guides the user to the target, the last person in the line. (3) The system tracks the target to follow the movement of the line.

3. More detailed information on the distance or direction might further improve the efficacy of the system.
4. The cause of all task failures was incorrect positioning, due to miss-tracking of the target.

On the basis of these findings, we designed and implemented LineChaser.

7.4 Implementation

This section describes the implementation of LineChaser. As shown in Figure 7.7, our system uses a floor map that includes the line information to guide the user to the end of the line. Then, the system detects and tracks the target to follow their movement.

7.4.1 Map Preparation and Localization Phase

To localize the user's current position, LineChaser uses a floor map prepared in advance and ARKit to detect augmented reality (AR) markers placed on walls. We note that there is prior work on smartphone-based indoor localization [122, 172, 189, 215], and the aim of the present work is not to advance the state-of-the-art in this respect. Instead, our contribution is the analysis and development of practical assistance for blind people to stand in lines based on such localization results.

Map Preparation

A map that consists of the information of where a line might form is created (Setup in Figure 7.7). First, we place an initial AR marker on the floor and scan it with the front RGB camera of the smartphone. Subsequently, we add the locations of the entrance, corner and line destination in the map. The orientation of the line is recognized by the directed edges of the line. Edges made by connecting each placed corner represents the center of the line. The width of the line is determined as a distance from both sides of the center of the line. Simultaneously, we place and scan additional AR markers to help users re-localize their position while they are standing in line. Finally, LineChaser creates a map that records all positions of the line relative to the AR markers.

Localization

While performing the task of standing in line, LineChaser localizes the user's current position and orientation on the prepared floor map. After the map is prepared, the blind user can now scan the initial AR marker to load the prepared map and localize their position on it. The system keeps track of the user's position using visual inertial odometry [259], which is supported by ARKit[4], whose localization errors generally range from 0.27m to 0.74 m [259] in a route about 61m long. The system tracks the user's position by combining the smartphone's motion tracking sensors and computer vision-based analysis of notable features obtained from the built-in rear RGB camera of the smartphone.

7.4.2 Line-standing Phase

Finding the End of the Line

LineChaser first guides the user to the entrance of the line and then locates the end of the line. The system conveys the direction and distance toward the entrance of the line (Figure 7.7, (Step1)). Upon reaching the entrance of the line, the system instructs the user to walk along the line area (Figure 7.7 (Step2)). While walking along the line area, the system detects nearby persons to search for the person standing at the end of the line (the person detection algorithm is described in section 7.4.2). When the system detects the person who is the last person in the line, the system recognizes this person as the target. At this point, the system acquires the initial color histogram of the target (Section 7.4.2). Finally, the system assumes that the blind user has now been guided to the end of the line and begins the line following task.

LineChaser navigates the user along a simple straight line connecting the user's current position with the entrance of the line. We note the existence of prior work on path planning for autonomous robots [69] and blind navigation [85]. These systems can generate a safe path that avoids obstacles, such as walls and static structures, using information from a LiDAR sensor to represent the structural surroundings. This study focused on developing a navigation interface, rather than a new path planning system.

Person Detection and Target Tracking with Color Histograms

LineChaser uses the front camera of the iPhone 11 Pro and YOLOv3-tiny [205] to detect pedestrians in the same way as our initial prototype system (Figure 7.2 (a)). Based on the calculated bounding boxes and the depth data from the iPhone, the system estimates the positions of detected pedestrians in the map coordinate system.

The failure case during the preliminary study occurred because of miss-tracking of the target (Section 7.3.3). Therefore, we implemented a new target tracking system. LineChaser first tracks each person based on the calculated positions for each frame (See [115] for the concrete algorithm[5]). We observed that the system can track persons at a rate of around 15 frames per second.

LineChaser uses the results of person tracking to track the target. The system distinguishes tracked people by the color of their clothes, which corresponds to the color histogram of the center area of their bounding box. When the target is initially recognized (Section 7.4.2), the system acquires his or her color histogram. Out of all detected persons, the system selects the person with the minimum value of histogram distance between the color histogram of the target and that of the detected person. For histogram distance, we adopted the Bhattacharyya distance for each a and b dimension in the Lab color space. If the histogram distance is below the threshold γ, the system recognizes the person as the

[4]https://developer.apple.com/arkit

[5]Based on our observations, we set the parameter values of the algorithm $\alpha = 0.5$ m and $\beta = 15$ frames for all of our studies.

target. We set the parameter value $\gamma = 0.40$, which is 40% of the maximum value of the color histogram distance.

Following the Target

After finding the person at the end of the line, LineChaser then instructs the user to follow the target. The system calculates d_{Line}, the distance between the user and the target along the waiting line, as shown in Figure 7.7, (Steps 2 and 3). The system instructs the user to maintain a distance of $d_0 m$ from the target. When $d_{Line} > d_0$, the system prompts the user to move forward. When $d_{Line} \leq d_0$, the system instructs the user to stop (Figure 7.7 (Step3)). We set the parameter value $d_0 = 1.7$ m, to maintain social distancing.

Considering Social Distancing

As reported in section 7.3.3, blind people usually either listen to ambient sounds or rely on others to navigate a waiting line. Since the outbreak of COVID-19, blind people, like everyone else, need to maintain a protective social distance between themselves and others. This prevents them from relying on the methods reported in section 7.3.3 to navigate the line. Therefore, we have adjusted LineChaser to maintain the social distancing.

7.4.3 Audio and Vibration Interface

LineChaser uses both audio and vibration interfaces. The audio interface conveys information about direction and distance, and the vibration interface prompts the blind user to move or stop when the user is facing the correct orientation. Participants in the preliminary study preferred to know detailed information about distance and direction to the target (A7.11 and A7.12). To take this user feedback into account, we used audio (text-to-speech) feedback for the main user interface. The audio interface can convey various types of information such as clock positions [33, 122, 148, 175, 211] or whether to move right or left [62, 67, 215, 259]. To convey a specific direction for navigation, we adopted a method based on clock position. Loomis et al.showed that instructions with clock position are capable of guiding blind people to a specific destination [148]. To explicitly convey the position of the destination to the user, for example, the system says *"Walk to the 2 o'clock, 2.1 meters ahead."* Any time the user shifts more than 30° from the expected orientation, the system repeats the instruction to the user to turn them toward the destination. In other words, the system is silent when the user is facing the expected orientation. Also, to keep the blind user in the center of the line when they are shifting aside, the system guides the blind user to slide left or right towards the center of the line.

An example of audio instructions during the line-standing phase is as follows:

1. **Starting navigation:**
   "Walk to the 2 o'clock, 2.1 meters ahead."

2. **Arriving at the entrance of the line:**
   "You arrived at the entrance of the line. You will now be walking along the line."

3. **Finding the Target:**
   "Target[6] found. Stop. The target is at the 1 o'clock, 1.5 meters ahead."

4. **Following the line movement:**
   "Walk to the 12 o'clock, 1.4 meters ahead towards the target."

5. **Ending navigation:**
   "You are now in front of the line."

[6] We used the phrase "target" because it was a concise way to express "the person in front."

Table 7.2: Participants' demographic information and their values for SUS scores.

ID	Age	Gender	Navigation Aid	SUS (Grade)	
P7.1	58	Female	Cane	60	D
P7.2	44	Female	Cane	77.5	B+
P7.3	56	Female	Cane	82.5	A
P7.4	53	Female	Cane	90	A+
P7.5	23	Male	Cane	82.5	A
P7.6	57	Female	Dog (primary) and Cane	90	A+
P7.7	49	Male	Cane	80	A-
P7.8	45	Female	Cane	90	A+
P7.9	38	Male	Cane	87.5	A+
P7.10	47	Female	Dog	82.5	A
P7.11	24	Female	Cane	100	A+
P7.12	33	Female	Cane	72.5	C+

We also integrated vibration feedback for the additional user interface, as using a double encoding of the information with both audio and vibration feedback can increase the understandability of the feedback [76, 83, 256]. We used a weak, short vibration for move signal (Each set of move signal using three-pulse-vibration: PD of 0.1 s and IPI of 0.17 s. Interval between each move signal is 0.5 s). LineChaser vibrates weakly when the blind user is facing the correct orientation and indicates to move forward. This enables the blind user to correct their orientation because they only have to face the orientation which the system vibrates weakly. As long as the user can perceive this weak vibration, he or she can walk forward until the next instruction is enunciated. Blind users are instructed to stop when they cannot perceive this vibration. Even if the user is facing the correct orientation, but the distance to the target is within 1.7m, the system will stop vibrating, indicating that the user should stop. For the emergency stop signal, we use a long and strong vibration (PD of 0.4 s and IPI of 1.0 s). This vibration is used to alert the user that they are in imminent risk of collision. If there is an obstacle in the camera field of view within 0.50m, the system alerts the user with a long and strong emergency vibration. Upon sensing this vibration, the user is expected to stop and wait for the next audible instruction.

7.5 User Evaluation

To evaluate the effectiveness of LineChaser, we performed a user study with 12 blind participants. We recruited blind people who are able to use their cane or guide dog to independently travel and who often travel independently. Also, P7.2 in the preliminary study and P7.12 in the main study were the same person. Therefore, we will report our quantitative results without the results with P7.12 in Section 7.6.2 and Section 7.6.3 because she might have a learning effect due to the participation in both studies. As shown in Table 7.2, we recruited 12 blind participants (eight females, four males) aged 23 to 58 years old (mean=43.8 and SD=12.1). All participants considered themselves to have good orientation and mobility skills. Also, all participants used smartphone in their daily lives for more than three years.

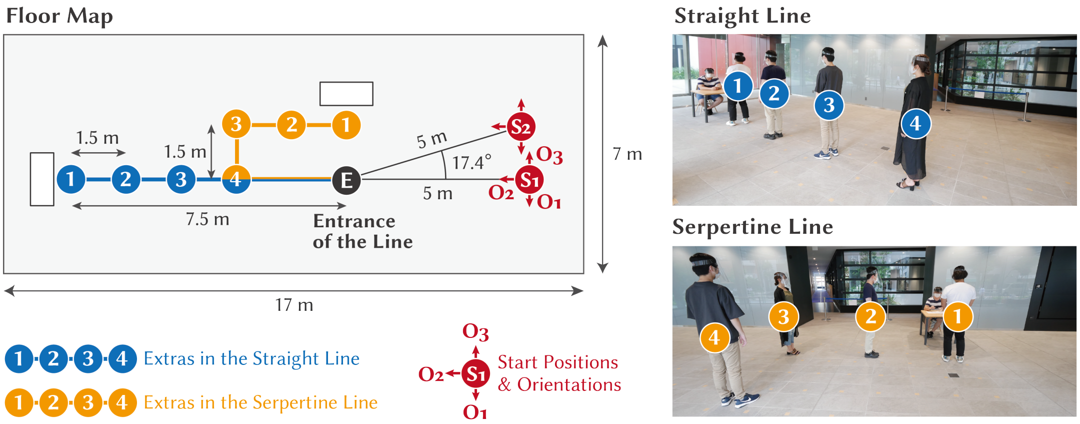


Figure 7.8: Overview of user study set up.

7.5.1 Tasks and Conditions

The blind participants were asked to find and follow a line consisting of two to four extras in front of them (Figure 7.8). To prevent the spread of COVID-19, we asked extras to cover their face with a mask and a face shield, and maintain social distancing (1.5 m) while standing in line. Similar to the preliminary study (Section 7.3.2), blind participants were asked to proceed until reaching the reception desk (the goal). All of the blind participants except P7.10 held the smartphone in their left hand and cane in their right hand. P7.10 held the smartphone with the right hand and the guide dog in the left. Blind participants were first placed 5 m away from the line in an initial position and orientation that were both randomized in each trial. There were two starting positions (**S1** and **S2** in Figure 7.8) and three initial orientations (**O1**, **O2**, and **O3** in Figure 7.8). After the blind participant successfully found a line, a researcher signaled the extra standing at the front of the line to leave the line randomly after 20, 40, or 60 s.

We designed two types of lines: **L1** was organised straight line and **L2** was organised serpentine line (Figure 7.8). Each blind participant completed six trials (three trials for each L1 and L2) of the standing-in-line task. We changed the number of extras (from two to four) in line for each trial to vary the position of the end of the line. When L2 trial has only 2 extras, the line is straight, but the route toward the end of a line is different from the L1 trial since the predefined line area is serpentine. To reach the goal of the line, participants have to make two turns. The user study was designed to simulate situations such as like a cashier line at a shopping mall or check-in-counter at an airport. To simulate a crowded public space, we played ambient noise recorded at a shopping mall at 60 dB [208].

7.5.2 Procedure and Metrics

After obtaining the IRB-approved (the Ethics Review Committee on Research with Human Subjects of Waseda University, 2020-039) informed consent from participants, similar to the preliminary study (Section 7.3.2), we performed a 15-min pre-interview during which we asked about participants' daily experiences, challenges and confidence when standing in lines. We asked participants to rate a set of statements (Q7.1–Q7.7 in Figure 7.9) using a 7 point Likert scale (from 1, strongly disagree, to 7, strongly agree). Then we gave participants around 30 min of training before the main session. We also clarified that the technical phrase "target" meant the person to follow in the training session. Based on the previous feedback, where P7.2 found it difficult to hold the smartphone (A7.8), we applied a smartphone ring to the smartphone so that the user can easily maintain a constant system position. After the training, we conducted the main session, which took around 30 minutes.

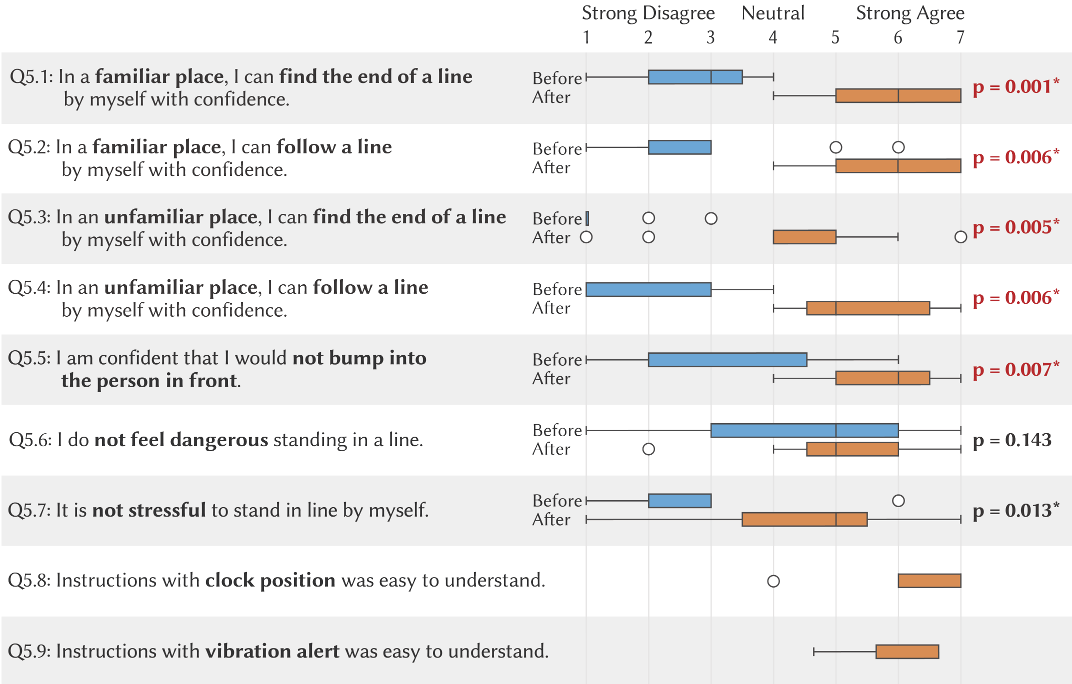


Figure 7.9: Likert scores and summary of the answers before and after the experiment. p is the value of the Wilcoxon signed-rank test done on each question (* indicates the significance found at the levels of 0.01).

To measure how accurately the blind user found the line, we measured the distribution of stop positions of the blind user when he or she found the line. As shown in Figure 7.10, we defined the area within a square of length 0.5 m as the *ideal position* when the blind user is standing in the center of a line and maintaining social distancing (1.5 m) from the person in front of them. We also defined the space around the ideal position (a total square of length 1.5 m), as *acceptable positions*. Every time a participant found the line, we added their stop position to the distribution. Every time the line moved, we added where the participant stopped to the distribution. While the experiment, we put tape on the floor and measured the actual stop positions referring to the tape.

After completing all the trials, we asked participants to answer a set of questions (the SUS [39] and Q7.1–Q7.9 in Figure 7.9) as well as open-ended questions to gather qualitative feedback. To observe how LineChaser improved the user experience when standing in line, we compared the results of the post-interview with those of the pre-interview. In total, the whole experiment took approximately 90 min per participant.

7.6 Results

7.6.1 Past Experiences about Standing in Lines

All participants reported that their main strategy for finding the end of a line was asking someone, usually a stranger or a store clerk. Four participants (P7.6, P7.7, P7.11 and P7.12) clarified that they have no other way of finding the end of a line except to ask others for help. P7.2 and P7.7 reported that they hesitate to stand in lines on their own because they think a stand-in-line task will certainly cause them trouble:

> **A7.15:** *"It it difficult for me to both find and follow any line. I do not stand in lines by myself because it is troublesome."* P7.7

Other than asking a stranger or a clerk, some mentioned that they find the end of a line by sensing positions of surrounding people by listening only in a familiar place (P7.4, P7.8, and

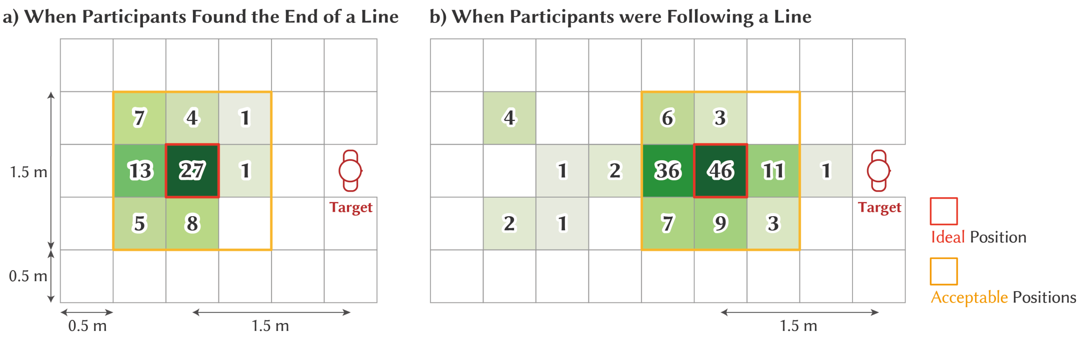


Figure 7.10: Stop positions while finding and following a line.

P7.11). P7.9 mentioned that he video calls his family to determine where the end of a line is.

Eight participants reported that they complete the task of following a line mainly by asking a stranger to notify them when a line moves or by listening to ambient sounds. Two participants reported approaching strangers apologetically as:

> **A7.16:** *"When I ask a stranger to notify me each time the line moves, I feel sorry to have them help me for a long time."* P7.5

Although one of their main skill is to use auditory senses, eight participants reported feeling that it is difficult to use auditory senses during the current COVID-19 pandemic. Not only are they required to cover their face with a mask (which dulls their auditory sense), but they must also to maintain social distancing with others, which prevents them from detecting a target from aural sensing. One participant commented:

> **A7.17:** *"I usually listen to ambient sounds or footsteps to follow a line. However, since the outbreak of COVID-19, my auditory senses has been limited by needing to wear a mask."* P7.2

7.6.2 Overall Performance

Stop Position after Finding the End of a Line

Figure 7.10 (a) shows the distribution of positions where participants stopped after finding the end of a line. All participants were able to successfully find the end of a line in all 66 trials. The success rate of finding a line at the ideal position was 40.9% (27 out of 66). In the other 59.1% (39 out of 66 trials), all participants successfully found the line within the definition of acceptable positions.

Time Took to Find the End of a Line

In L1 (straight line), the average time took to find the end of a line with two, three, four extras were 47.8s (SD: 18.0s), 49.9s (SD: 33.8s) and 41.7s (SD: 14.2s), respectively. For L2 (serpentine line), the corresponding times were 52.8s (SD: 15.6s), 55.8s (SD: 37.8), and 36.9s (SD: 13.4s). All participants tended to take more time to find the end of an L2, especially when it had only two or three people. This is mainly because they have to make 90° turn to the right once or twice to find the end of the L2 line.

Stop Positions when Following a Line

Figure 7.10 (b) shows the distribution of positions where participants stopped while following a line to the goal. All participants were able to successfully follow the line and reach the goal in all 72 trials. The success rate of following a line with the ideal positions was 34.8% (46 out of 132). Participants followed with acceptable positions 91.7% of the time (121 out of 132). For the other 11 trials, participants failed to stop within acceptable positions.

The 11 failure cases, in which participants did not stop in acceptable positions, occurred mainly due to three reasons:

- **Reason 1:** P7.1, P7.7, and P7.8 were unable to correct her orientation.
- **Reason 2:** P7.5 did not understand how to use the interface.
- **Reason 3:** P7.11 did not listen to the instructions.

Reasons 1 and 2 were the causes of 10 failures in which participants stood behind the acceptable positions (i.e., they stayed behind as the line moved forward). LineChaser vibrates to prompt forward movement only when the participant is facing the correct direction. Three participants were unable to face the correct direction and thus did not receive the vibration signal to move forward. Reason 3 was the cause of one failure in which P7.11 stood just behind the target. For the trial in which the navigation failed, P7.11 managed to stop just before colliding with the target because LineChaser issued an emergency stop signal. Also, in trials for P7.12, we observed a situation where ARKit accumulated a localization error and therefore instructed P7.12 to walk out of the line. The details of localization errors in ARKit are reported in Yoon et al. [259].

Comparisons between LineChaser and the Prototype System

Although P7.12 failed to follow the line movements with the prototype system (four trials out of six trials), when P7.12 used LineChaser, P7.12 was able to complete both tasks in all trials. P7.12 described the reason for her success as: **A7.18:***"I was able to hold the smartphone stably compared to the preliminary study."* (P7.12). She also gave a higher SUS score for LineChaser (72.5, C+) than for the prototype system (37.5, F), and gave a positive comment: **A7.19:***"The integration of the audio and vibration interfaces made me confident about facing the right direction. The new audio feedback gave me a rough image of the direction I should be facing compared to the prototype system."* (P7.12).

7.6.3 Subjective Ratings

Figure 7.9 shows the questionnaire results, which show that most participants felt more confident and comfortable standing in a line after the experiment (with the system) than before (without the system, their daily experience). We compared each question using the Wilcoxon signed-rank test with 1% levels of significance. The p-values of each test are shown in Figure 7.9. Our analysis revealed that, in Q7.1–Q7.5 except for Q7.6 and Q7.7, LineChaser received significantly ($p < 0.001$) better ratings than their daily experience. As shown by Q7.8 and Q7.9, the audio and vibration interface both received a high rating as no one rated both interfaces lower than 4 (neutral). Table 7.2 reports SUS scores [27] for each participant. The mean SUS score was 83.9 (SD: 10.1) which is an "A" rating on the SUS grade.

7.6.4 Qualitative feedback

All participants generally agreed that they were able to both find and follow a line by themselves with LineChaser:

> **A7.20:***"I am very happy to be able to find a line without needing to touch anything [rails or strangers]. Also, I was confident that I would not bump into the target as the system notifies me with an emergency signal if I am too close."* P7.3

> **A7.21:***"With this system, I can grasp my orientation and the distance to the target. Being able to grasp the distance is very important in the current situation [with regard to COVID-19]. Also, I am surprised that this system can specify the target."'* P7.4

A7.22: *"I felt that I do not need my cane anymore, as with this system I can maintain a certain amount of distance from the target. I think this system is revolutionary because I usually find a line by asking strangers, but with this system, I do not have to rely on others."* P7.6

We also received feedback from P7.10, who used LineChaser with a guide dog. She indicated that some of the instructions for LineChaser should be improved:

A7.23: *"I could both find and follow a line, while maintaining social distancing. Neither of these tasks are supported by my guide dog."* P7.10

A7.24: *"The system gave me an instruction to slide left, but guide dogs does not have such commands as they are usually taught to keep left of the owner. I felt her being confused when I slid to my left. As I am pretty confident that I will not bump into anything with a guide dog, I only need information about the direction to the end of a line and distance to the target to maintain social distancing while following them."* P7.10

Most participants (P7.2, P7.3, and P7.6–P7.12) reported feeling that the integration of audio and vibration feedback was easy to understand:

A7.25: *"The audio feedback gives me an approximate sense of the direction I should be facing. Then, I can find the exact direction with vibration feedback."* P7.9

P7.10 also commented about difficulty getting used to the interface:

A7.26: *" As I use this system I noticed that the tip for using this system is to move gently, not quick... Some people may have a hard time until they get the tip of it."* P7.9

Half of the participants (P7.4 and P7.8–P7.12) had a positive impression of the system because it was implemented on an off-the-shelf smartphone:

A7.27: *"I am happy that this system allows me to follow a line with just a single smartphone."* P7.8

However, 11 participants pointed out that the requirement to hold the iPhone so that the front camera faces others should be improved:

A7.28: *"I could easily hold the iPhone, but I hesitate to hold it by facing the camera to others."* P7.3

A7.29: *"I prefer not to hold the iPhone like this because people might think I am raising my hand"* P7.5

Also, two participants reported a physical burden of LineChaser:

A7.30: *"The way of holding the iPhone was easier compared to the prototype system, but still heavy because I had to hold my hand up."* P7.12

A7.31: *"The method of holding the iPhone by using our left hand should be improved. We blind people prefer our hand to be free."* P7.6

7.7 Discussion

7.7.1 Effectiveness of LineChaser

All participants reported that they constantly face difficulties where they do not know where the end of a line is and also when and how much the line moves while they are waiting in it (A7.15). They also reported that the social distancing norm since the outbreak of COVID-19 makes the situation more challenging than before because nearby people are farther away than the necessary distance for reliable auditory sensing (A7.17).

Although some users had various difficulties, LineChaser successfully enabled all participants to both find and follow a line. They all successfully found the end of a line at acceptable positions, indicating that LineChaser can navigate the blind user properly to the end of a line. LineChaser was also able to help blind users stop at acceptable positions 91.7% of the time while following a line (Figure 7.10 (a)). Feedback from the participants also supported the effectiveness of our system (A7.20–A7.23). LineChaser received an SUS mean score of 83.9, which is rated as "A." All participants significantly increased most of their scores for confidence and comfortableness when standing in line (Figure 7.9).

7.7.2 User Interfaces and Training

In LineChaser, we adopted both audio and vibration for the user interface as audio feedback alone may be less effective in a noisy/crowded environment [36]. As a result, participants appreciated that using a double encoding of the same information with both audio and vibration can increase the understandability of the feedback (A7.25).

While the overall results are positive, we also found opportunities to improve the user interfaces. We observed a few cases in which the system did not provide a vibration to signal to prompt the user to move forward because of an incorrect device orientation due to the sensing capability of correct orientation (Section 7.6.2 (1)). Improved real-time guidance for device orientation may reduce such failures. In one instance, P7.11 did not notice the signal to stop (Section 7.6.2 (3)). Thus, we may need to redesign the signals to function better in noisy practical environments.

Another possibility is a user interface that adapts to the skill of user's navigation aid (a cane or a guide dog). As we observed from P7.10 (A7.24), improved instructions should be designed not only for cane users but also for guide dog users. Williams et al. [249] summarized that a cane is for obstacle detection and a guide dog is for obstacle avoidance. As guide dogs can naturally prevent collisions, instructions can be reduced by eliminating collision-related information. Instead, the system can provide more information about the surrounding environment. Additionally, P7.10 reported that some instructions are not appropriate for guide dogs (e.g., slide left), and it is necessary to conduct studies that focus on users of guide dogs and carefully redesign an optimized user interface for them.

For higher usability of the system, not only a redesign of interface is required but also training method to utilize the interface should also be considered. While all participants could learn how to use the system after a short training session (30 min), P7.1, P7.5, P7.7, and P7.8 experienced ten failures because they were not accustomed to the interface of the system (Section 7.6.2 (1) and (2)). Training is an imperative part of the efforts to make these technologies available for the blind community. We plan to collaborate with orientation and mobility training communities to design new training methods by seamlessly integrating new navigation technologies into their traditional navigation tools, such as canes and guide dogs [245].

7.7.3 Integration with Navigation Systems

Our ultimate goal is to supplement practical indoor navigation systems to allow blind users to stand in line in real-world environments. Various localization methods are proposed for helping the blind people navigate, but the methods based on the radio-wave signal strength (RSS) of Wi-Fi networks or Bluetooth beacons are used by the most practical systems [72, 122, 172, 189, 215]. Such systems can achieve an accuracy of 1.5m mean error at best, but this is not sufficient to navigate a user to an end-of-line location as LineChaser did. We used the AR marker-based localization (Section 7.4.1) that had a better localization accuracy compared with RSS-based systems [172, 189]. Overall, most of the components can be integrated into an RSS-based indoor navigation system, but the function for finding the end of the line requires higher localization accuracy. Possible solutions include integrating computer vision-based end-of-line recognition, improving RSS-based localization, and integrating AR marker localization. We plan to consult development teams for indoor navigation systems to plan a road map toward integration.

7.7.4 Real-world scenarios

We conducted experiments in a controlled environment. In real-world scenarios, there are varieties of situations that we can not simulate in a lab-based study. For example, a target person may bend down while following a line. In this case, LineChaser may not have the target person in the camera field of view and then misinstruct the blind user to move forward. In a case a line has groups of people standing together and occupy the same area of the line, the system recognizes the closest person in the group as the target. The closest person in the group may dynamically change, and the system will fail target identification. Therefore, we plan to conduct real-world user studies to assess the generalizability of LineChaser in various situations in real-world settings.

7.7.5 Social Acceptance

Although we obtained positive feedback on the fact that the entire stand-in-line task could be done with an off-the-shelf smartphone (A7.27), some participants stated that they would not use this system in public spaces, mainly due to the way that the smartphone must be held (A7.29). LineChaser was implemented on iPhone 11 Pro by facing the front camera to others. When considering how the system would fit with current uses' practice, the release of iPhone 12 Pro, which is equipped with Lidar sensor on the back, may have a positive impact on the problem with the way of how the smartphone is held.

Participants were concerned about pointing a camera explicitly at others (A7.28). While cameras and sensors of smartphones are being developed to provide more opportunities to support blind people, this problem is always present and may cause *social friction*. To reduce such social friction, alternative *seamless* wearable devices (e.g., smart glasses) may play an important role as a technical solution. We also should raise the societal awareness of the computer vision-based assistive technologies for the blind.

7.7.6 Ergonomics

Given the fact that we focus on off-the-shelf smartphones, it is not easy to improve the ergonomics of the system. We improved the device's graspability by adding a small handle to the smartphone after the prototype evaluation. P7.12 rated our prototype system with a SUS score of 37.5 (F) because she had a hard time holding an iPhone 11 Pro and she reported it was too heavy to hold for several minutes. However, P7.12 rated our improved system higher because the handle helped her grasp the device comfortably (A7.18). Overall, P7.12 was able to stand in line with high confidence with LineChaser. However, P7.12 still reported a physical burden of the system (A7.30), and P7.6 preferred to free their hand while using the system (A7.31). We observed that the ergonomics aspects, such as the way of gripping the device, greatly affects the usability of the whole system, and such aspects should be taken into account in the device selection phase for assistive technologies. For example, using wearable devices which are discussed in Section 7.7.5 could be one solution to this problem as it enables users to use the system with their hands-free.

7.8 Conclusion

This work developed a smartphone-based system that helps blind people to stand in lines. We first developed a prototype system that helps blind users follow the person in front of them as they move intermittently, and performed a preliminary study with six blind people. Based on the results and feedback, we designed LineChaser. LineChaser first guides a blind user to the end of a line and then helps the blind user follow the line. LineChaser uses the RGB camera to detect nearby people and the infrared depth sensor to estimate the distance to the target. LineChaser uses an audio interface to convey detailed navigation instructions and information on the distance to the target. A vibration interface prompts the blind user to move or stop. We performed a more advanced user study with 12 blind participants and observed that LineChaser enabled all participants to complete both tasks, while maintaining appropriate social distancing. We observed that LineChaser significantly increased their confidence in standing in lines. In the future, we plan to integrate LineChaser into a navigation system with high localization accuracy and a re-designed interface that is improved based on the results of real-word user study.

Acknowledgement

We would like to thank all participants who took part in our user study. We would also thank Asuka Hirata, and Yoshiki Kubotani for their support, and the anonymous reviewers for their helpful comments. This work was supported by JST-Mirai Program (JPMJMI19B2), JSPS KAKENHI (JP20J23018), and Grant-in-Aid for Young Scientists (Early Bird, Waseda Research Institute for Science and Engineering, BD070Z003100).

Chapter 8

Blind Pilot: Robot-based Navigation System to Landmark Objects[1]

"I felt secure because walking with the robot was similar to the feeling of walking with a person."

Participant 5
IBM Japan Ltd
March 17, 2019

8.1 Introduction

Independent travel is a major challenge for blind people as they sometimes need to interact with a landmark object such as a door when entering a room, a button when using an elevator, and a chair when arriving in a lounge. Interactions with such objects require significant effort and time for blind people because of their lack of visual perception.

Several local navigation systems have been proposed in previous studies to detect landmark objects such as doors and chairs and provide information concerning the detected objects (e.g., direction and distance) via sound and speech [67] or tactile [243] feedback to blind users. These systems allow users to approach a detected object while correcting their direction; however, it is difficult for blind people to maintain a straight trajectory in open spaces [110, 248]. Therefore, such systems require blind users to quickly and repeatedly adjust their orientation for precise navigation.

To resolve this limitation, we propose a local navigation robot, *BlindPilot*, which can lead blind users directly to landmark objects. By only following the movement of BlindPilot, blind users can reach a target object faster with less effort compared to previous local navigation systems providing sound/speech feedback. In this study, we focus on a scenario in which BlindPilot detects an empty chair and guides a blind user to that chair (Figure 8.1). BlindPilot uses an RGB-D camera to detect the empty chair and LiDAR to build a 2D map of the surrounding area using simultaneous localization and mapping (SLAM). The RGB images are used to detect chairs and people using a convolutional neural network (CNN)-based generic object detector [205], whereas the depth channel is used to estimate the 2D positions of the detected chairs and people. Then, the system determines whether each chair is empty by considering the estimated 2D positions of the chairs and people. After mapping the positions of the empty chairs on a 2D map, the system generates a path to an empty chair and moves along the generated path.

To evaluate the effectiveness of BlindPilot in approaching landmark objects, we performed a user study with six blind participants. As a baseline system, we implemented a sound feedback system based on a previous study [67]. Then, we requested the blind participants to approach a chair using the proposed system or the baseline system. We observed

[1]Project Page: https://wotipati.github.io/projects/CHI2020_LBW_BlindPilot/CHI2020_LBW_BlindPilot.html

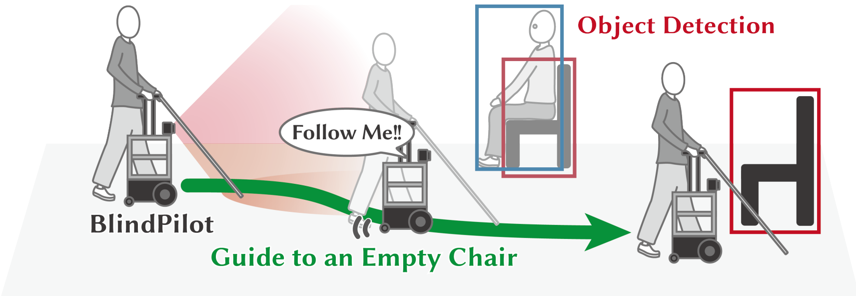


Figure 8.1: We present a local navigation robot, BlindPilot, which directly leads blind users to an empty chair.

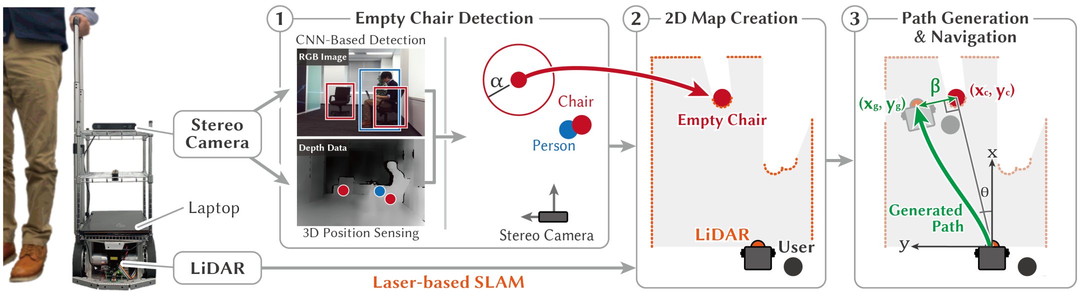


Figure 8.2: Overview of the navigation method. The system involves a stereo camera and LiDAR mounted on a mobile robot. It uses the camera to obtain RGB images and depth data. (1) The system detects chairs and people using images and obtains their 2D positions using the depth data. Then, it judges whether each chair is empty by considering the previously estimated positions. (2) It then builds a 2D map of the surrounding area using SLAM and maps the positions of the empty chairs. Finally, (3) it generates a path toward the goal position (x_g, y_g) next to the chair (x_c, y_c) and leads the blind user by moving along the selected path.

that BlindPilot allowed users to approach an empty chair significantly faster compared to the baseline system with sound feedback. Moreover, on the basis of the qualitative feedback from the blind participants, we confirmed that the robotic local navigation system could navigate blind users with a feeling of security and less effort. Based on our findings, we discuss certain future directions for research to realize a more flexible and comfortable navigation robot for blind users.

8.2 Proposed System

We developed a local navigation robot called BlindPilot, which can directly lead blind users to an empty chair. As outlined in Figure 8.2, BlindPilot navigates blind users in three steps: 1) empty chair detection, 2) 2D map creation, and 3) path generation and navigation. All these processes are performed on a laptop computer (CPU: Intel®Core™i7-8750H and GPU: NVIDIA®GeForce®GTX 1080) attached to a mobile robot.

8.2.1 Step 1: Empty Chair Detection

We implemented an empty chair detection system by combining stereo image sensing and a CNN-based object detector (YOLOv3 [205]). The system is equipped with a ZED™Stereo Camera[2] to capture RGB images and collect depth data. BlindPilot uses YOLOv3

[2]https://www.stereolabs.com/zed/

to detect chairs and people using the RGB images. We used the central areas of the detected bounding boxes to obtain the 2D positions of the detected objects in the camera coordinate system.

The system determines that a detected chair is empty if there are no existing people within a distance α from the chair position (Figure 8.2(1)). On the basis of our observations, we set the parameter value $\alpha = 1$ m for all of our studies.

8.2.2 Step 2: 2D Map Creation

The system develops a 2D map of the surrounding area via laser-based SLAM. We built this map creation system on the basis of ROS gmapping [82], which can create a floor plan and the pose trajectory of a mobile robot using laser data. In this study, we used a commercially available mobile robot (MegaRover ver. 2.0, Vstone Co., Ltd.[3]). To obtain laser data, we attached a LiDAR unit (URG-04LX-UG01, HOKUYO[4]) to the mobile robot (Figure 8.2).

After creating the 2D map, the system maps the position of the empty chair estimated in Step 1. If several empty chairs are detected in Step 1, the system selects the closest chair as the target chair.

8.2.3 Step 3: Path Generation and Navigation

The system then generates a path next to the target chair and automatically leads the blind user by following the generated route. This navigation system is built on the ROS navigation stack [156], which comprises a path planner, a localization system, and a mobile robot controller. The path planner can generate a path that avoids static obstacles.

To enable users to easily sit in chairs, the system sets the goal position to the left of the target chair. More formally, let (x_c, y_c) and θ be the position and direction of a detected empty chair, respectively, as illustrated in Figure 8.2(3), then the system computes the goal position (x_g, y_g), which the system publishes to the ROS navigation stack, as follows.

$$\begin{cases} x_g = x_c - \beta sin\theta \\ y_g = y_c + \beta cos\theta \end{cases} \tag{8.1}$$

Let β be the distance between the target chair and the goal position. Based on our observations, we set the parameter value $\beta = 0.7$m for all of our studies. When BlindPilot starts or ends the navigation, the system outputs an audio message such as *"Starting the navigation"* or *"Arrived at the destination. "*

8.3 User Evaluation

Our primary goal in this study was to understand the effectiveness of BlindPilot with respect to local navigation for blind users. Therefore, we performed a preliminary user study in which six participants (female: 2, age: 24.5 ± 4.2 (mean and SD)) approached an empty chair using the proposed system. In this study, we compared BlindPilot to an audio-based local navigation system that provides information to navigate toward a target object (e.g., direction and distance) via audio-based feedback to blind users.

8.3.1 Baseline System

As a baseline system, we implemented an audio-based local navigation system based on a previous study [67] in which Google Glass was used to estimate the position of a door and text-to-speech feedback was provided such as *"right"*, *"left"*, and *"straight X m"*

[3]https://www.vstone.co.jp/english/index.html
[4]https://www.hokuyo-aut.jp/search/single.php?serial=166

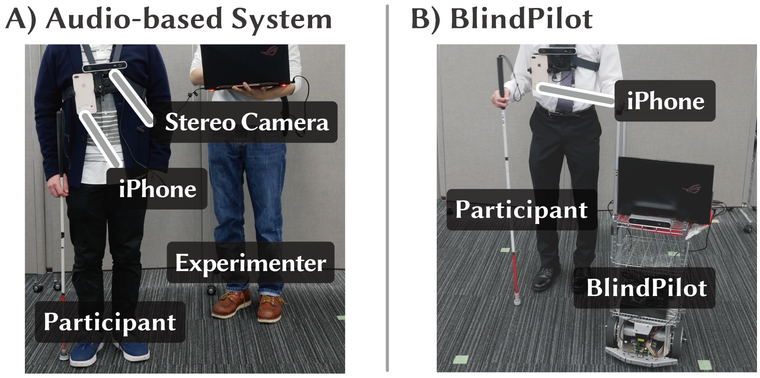


Figure 8.3: In our user study, we compared two systems: (A) an audio-based system and (B) a robot-based system (BlindPilot).

(actual algorithms can be found in [67]). In our study, we mounted a ZED camera on the participants' chests and detected the chair using the camera rather than Google Glass. While participants were approaching the chair using the baseline system, a researcher followed them while carrying a laptop (Figure 8.3 (A)).

8.3.2 Tasks

The primary focus of our study was to investigate which interface (robot-based navigation or audio-based navigation) is more effective and comfortable for blind users. Accordingly, we placed a chair in a simple square space (6 m × 11 m) and asked participants to approach the chair using the proposed system or the baseline system. We prepared three types of routes to the target chair: (1) Front (the target was straight ahead with respect to the start), (2) Right (the target was 20° to the right with respect to the start), and (3) Left (the target was 20° to the left with respect to the start). For all routes, the distance between the start position and each target was 6 m.

8.3.3 Procedure

After obtaining informed consent from the participants, the researchers provided an overview of the study and described these two systems. A short training session (10 min) was then given to the participants. In this training session, while using each system, participants approached a chair that was placed in front of them.

Then, the participants were requested to approach the chair three times using one of the systems (the proposed or baseline system) while the position of the target chair was changed (1: Front, 2: Right, and 3: Left). Then, we requested the participants to approach the chair three times using the other system. For each participant, the order of the three target positions and two systems was randomized.

To analyze the walking trajectories of the participants, we mounted an Apple®iPhone®8 Plus smartphone on the participants (Figure 8.3). We obtained the trajectories using ARKit[5], which can estimate the six degrees of freedom poses of the smartphone using the camera images. Furthermore, we recorded the behavior of the participants and counted the number of times that the participants stopped to adjust their orientation when using the baseline system.

[5]https://developer.apple.com/arkit/

Table 8.1: Likert items (1: the baseline system to 7: BlindPilot) and a summary of the answers.

No.	Question	P8.1	P8.2	P8.3	P8.4	P8.5	P8.6	Median
Q8.1	Which system was **more effective**?	4	5	7	7	7	7	7
Q8.2	Which system was **easier to use**?	2	6	7	7	7	7	6
Q8.3	Which system felt **more secure**?	5	6	6	7	7	6	6
Q8.4	Which system was **more comfortable**?	2	6	7	4	7	7	6.5

8.3.4 Metrics

Task Completion Time

We measured the times that elapsed before the participants sat on the chair (i.e., the task completion time). To compare the two systems, we hypothesized that the proposed system would result in a faster task completion time than the baseline system and validated this hypothesis based on a 95% confidence interval for the means and the Wilcoxon signed-rank test with 5% levels of significance.

Post-Interview

After completing all the tasks, we asked the participants to fill out a questionnaire[6]. We designed the questions based on the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) [54], which is a popular evaluation method for assistive devices. The total QUEST score, which is the mean of the scores, indicates the user satisfaction with the system. The statements and a summary of the answers are shown in Table 8.1. The participants answered these questions in the form of a seven-point scale ranging from 1: more inclined toward the baseline system to 4: neutral (i.e., both did or did not apply) to 7: more inclined toward BlindPilot.

Finally, to obtain qualitative feedback, we asked open-ended questions about the advantages and challenges of each system. Moreover, we asked for suggestions to improve each system and what types of landmark objects the participants would want to approach using the proposed system.

8.4 Results

8.4.1 Task Completion Time

Table 8.2 reports the task completion time in terms of its mean and standard deviation, as well as the 95% confidence intervals, obtained using BlindPilot and the baseline system. The Wilcoxon signed-rank test confirmed the statistical significance and the superiority of BlindPilot over the sound-based system for all tasks. When we compared the total task completion time, 95% confident intervals and the Wilcoxon signed-rank test confirmed the statistical significance and the superiority of BlindPilot.

Moreover, we observed that the blind participants who used the audio-based system required some additional time to adjust their orientation. Figure 8.4 shows certain examples of trajectories of participants. While the blind participants walked in a zigzag line when they were using the audio-based system (Figure 8.4 (A)), they could approach the chair smoothly using BlindPilot (Figure 8.4 (B)). Furthermore, when the participants used the sound system, they adjusted their orientation 2.9 ± 1.3 times (mean and SD) during each task.

[6] All communications with the participants were in their native language. In this chapter, we describe any translated content in the form of "*translated content*".

Table 8.2: Quantitative evaluation of the task completion time: Mean and SD: the mean and standard deviation of the task completion time; Lower and Upper: the lower and upper bounds of the 95% confidence intervals, respectively; and the p-value: the p-value of the Wilcoxon signed-rank test ($*$ and $**$ indicate the significance confirmed at the 0.1 and 0.05 levels, respectively).

Target chair position	Audio-based System			Robot-based System (BlindPilot)			p-value
	Mean and SD	Lower	Upper	Mean and SD	Lower	Upper	
Front	20.5 ± 2.75	17.3	23.7	$\mathbf{16.5 \pm 1.61}^{*}$	14.7	18.3	$\mathbf{0.04}^{**}$
Right	22.2 ± 4.30	17.2	27.1	$\mathbf{16.8 \pm 1.86}^{*}$	14.7	19.0	$\mathbf{0.06}^{*}$
Left	21.5 ± 4.57	16.2	26.8	$\mathbf{16.0 \pm 1.53}^{**}$	14.2	17.8	$\mathbf{0.09}^{*}$
Total	64.2 ± 6.87	56.3	72.1	$\mathbf{49.3 \pm 3.50}^{**}$	**45.4**	**53.6**	$\mathbf{0.04}^{**}$

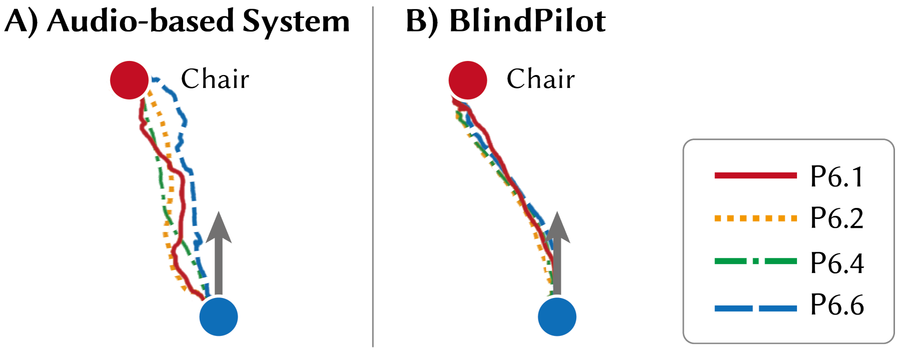


Figure 8.4: Trajectories of the blind participants approaching the chair from the left while using (A) BlindPilot or (B) the baseline system.

8.4.2 Post-Interview

Table 8.1 reports the results of our questionnaire. We confirmed that BlindPilot satisfied all but one participant (P8.1).

Positive Feedback

Five of the participants who valued BlindPilot (P8.2–P8.6) mentioned that the advantage of BlindPilot was that they could walk with a greater feeling of security:

> **A8.1:** *"Because the robot led the way, I did not have to worry about collisions with obstacles."* P8.2

> **A8.2:** *"I felt secure because walking with the robot was similar to the feeling of walking with a person."* P8.5

Although P8.1 liked the baseline system, P8.1 also acknowledged that BlindPilot provided a feeling of security:

> **A8.3:** *"I could approach a chair with a feeling of security because the robot directly guided me."* P8.1

Furthermore, five participants (P8.1–P8.4, and P8.6) reported that they could effortlessly approach the chair by following the movement of BlindPilot:

> **A8.4:** *"Because the robot automatically approached the chair, I could reach the chair easily just by following the movement of the robot."* P8.2

A8.5: *"When I used the robot system, I could approach the chair more smoothly compared to the sound system."* P8.1

A8.6: *"I could approach the chair without stress because the robot directly guided me to the chair."* P8.4

On the other hand, all participants commented that following the frequent navigational sounds of an audio-based system was difficult:

A8.7: *"It was difficult for me to walk straight. So, when I used the sound system, I had to change my orientation repeatedly."* P8.2

A8.8: *"When I used the sound system, I had to concentrate on the audio message."* P8.3

A8.9: *"The sound system required that I repeatedly adjust my orientation. This may increase the risk of collision with other pedestrians or obstacles."* P8.1

Negative Feedback

We obtained two types of negative opinions with respect to BlindPilot as follows:

A8.10: *"I want to know the position of the destination before the robot starts its navigation. I also want to know the movement of the robot. For example, the robot could turn right after saying 'Turn right.'"* P8.3

A8.11: *"The robot system does not change speed when approaching the goal."* P8.6

Other situations where blind people need local navigation

The participants said they wanted to be guided to an empty chair when they are in a train (P8.1, P8.2, P8.4, and P8.6), bus (P8.1, P8.2, and P8.4), food court (P8.2 and P8.3), or non-territorial office (P8.3 and P8.5). Furthermore, participants commented on other situations where they might need a local navigation system:

A8.12: *"I often take time to find an entrance after arriving at the front of a shop."* P8.5

A8.13: *"When I'm waiting for a train at the platform, I want to be navigated to a train door."* P8.4

8.5 Discussion

8.5.1 Advantages of BlindPilot

Our quantitative evaluation revealed that, with BlindPilot, blind participants were able to access an empty chair faster compared to when using the audio-based navigation system. As shown in Figure 8.4, while blind users could approach a chair smoothly using BlindPilot, they repeatedly adjusted their orientation and walked in a zigzag line when they used the sound-based system. Moreover, we confirmed that the majority of participants (five out of six) were satisfied with BlindPilot according to our questionnaire (Table 8.1). The feedback showed that BlindPilot could guide blind users with a feeling of security and less effort (A8.1–A8.6). Furthermore, it was observed that following the frequent navigational sounds of the audio-based system was difficult (A8.7–A8.9). These results indicate that a local navigation robot is a promising approach to enable easy local navigation for blind people with confidence (a feeling of security). In this study, BlindPilot performed better with respect to speed, a feeling of security, and the level of satisfaction. These results assume certain conditions such as correct landmark object detection and smooth movements by the robots. Furthermore, the participants reported that they wanted to use the robotic system to identify empty chairs, entrances, and doors in public spaces (A8.12–A8.13). It is necessary to evaluate multiple factors in practical public environments such as occlusion of the landmark objects by crowds and non-smooth movements that may be necessary to avoid obstacles.

8.5.2 Limitations and Possible Extensions

Explanatory Robotic System

Some participants reported the need to convey more informative feedback from the robot to the blind users. They suggested to provide information about the position of the destination and the movement of the robot in advance (A8.10). For global voice navigation systems, the importance of explanation for a current location is well known [200]. Furthermore, in autonomous robotic navigation, it is important to guide blind users while describing the next movement of the robot rather than guiding blind users in silence. Local navigation is more time critical compared to global navigation; therefore, the time available for such an explanation may be limited in practical situations. The timing, interaction methods, and amount of information should be well designed through future studies.

Flexibility of the Robotic Navigation System

Both P8.1 and P8.6 reported that a drawback of the BlindPilot experience was the inability of the user to control the walking speed. The current implementation does not allow the user to change speed when the robot is leading the user. This is one example of the well-known challenge of shared control [8], a situation when a user feels less independent when a robot does not accept any control even if the user is aware of better or more comfortable methods of completing the task. This result indicates that future versions of BlindPilot should provide shared control at least for speed control. The most popular solution is to enable a shift of control authority in real time during movement. We expect that the robot will need to provide shared control for many other aspects, such as orientation, route, and landmark object selection, to improve the user experience. The timing and interaction methods to improve control flexibility require further studies.

8.6 Conclusions

We proposed a local navigation robot, BlindPilot, which directly guides blind users toward an empty chair. The system detects the position of an empty chair using a CNN-based object detector and stereo image sensing. BlindPilot then generates a path to the chair and moves along the generated path. We evaluated the effectiveness of BlindPilot for six blind people. The results showed that BlindPilot guides blind users significantly faster than a sound-based local navigation system. Furthermore, our qualitative analysis showed that BlindPilot could guide blind users with a feeling of security and less effort as well as limitations such as explainability, and walking speed control. In future, we plan to implement and study a robotic system that can provide informative feedback and shared control toward our ultimate goal to create practical solutions for blind people to travel independently.

Acknowledgement

We would like to thank all participants who took part in our user study. We would also the anonymous reviewers for their helpful comments. This work was supported by JST ACCEL (JPMJAC1602) and JST-Mirai Program (JPMJMI19B2).

Part III

Social Acceptance of Assistive Systems

Chapter 9

Autonomous Navigation Robot's Acceptance in Public Buildings[1]

> *"I thought the robot would look more like a navigation machine, but it looks like a natural-looking suitcase and cool. I prefer looking natural and to not to be recognized as visually impaired. If the robot looked unsophisticated or conspicuous, I would not want to use it, but I love that this robot is natural."*
>
> Participant 1
> IBM Japan Ltd
> March 25, 2021

9.1 Introduction

Blind people face significant challenges when walking through large and complicated public buildings, such as shopping malls, airports, and hospitals, due to their lack of vision. Autonomous navigation robots have a great potential of transforming the daily lives of blind people by allowing them to move independently in such spaces. Recent research has proposed navigation robots that can guide blind users to their destinations and help them avoid surrounding obstacles and nearby pedestrians [85, 232, 260].

For the deployment of navigation robots in public buildings, it is imperative to obtain acceptance not only from blind users but also people in the buildings and those facility's managers. Specifically, 1) a navigation robot needs to be socially accepted by people in general, 2) a navigation robot that works in a specific building needs to be accepted by facility managers because they are responsible for keeping a building safe, reliable, and enjoyable for everyone, and 3) a navigation robot needs to be accepted by blind users so that they will use it in their daily lives. Prior works have investigated the social acceptance of assistive technologies, including wearable cameras [11, 138, 198] and computer vision-based assistance [3, 12], with visually impaired users and bystanders. However, there have been few efforts made to study acceptance and concerns of navigation robots for blind people [22, 24], and no work investigated it with facility managers.

In this paper, we conducted three studies to investigate the acceptance and concerns of three stakeholders: 1) an online survey of people in general, 2) interviews with facility managers, and 3) focus groups [81] with blind users. For these studies, we used a prototype navigation robot that is assembled into a suitcase (Figure 9.1), which can assimilate into the environment. It is capable of navigating blind people in multi-story buildings while avoiding obstacles and nearby pedestrians.

We first conducted an online survey with 300 sighted participants. The survey focused on the social acceptance of our navigation robot, which moves about with blind users. They

[1]Project Page: https://wotipati.github.io/projects/AI-Suitcase-Acceptance/index.html

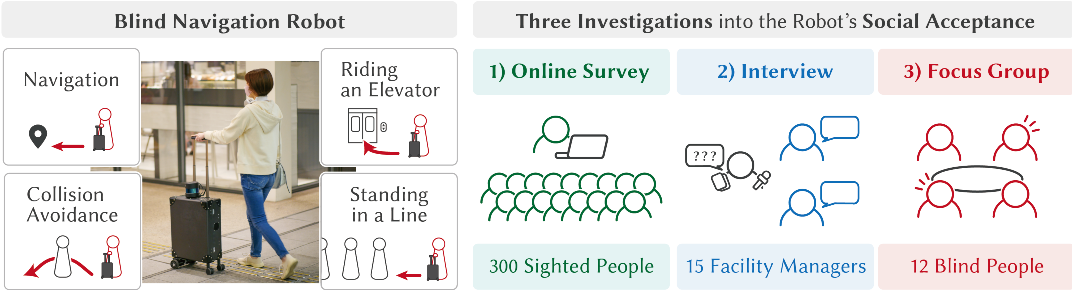


Figure 9.1: We evaluated the social acceptance and concerns regarding an autonomous robot for navigating blind people in public buildings. We conducted three studies to investigate the acceptance of and concerns with our prototype robot: 1) an online survey of 300 sighted people, 2) interviews with 15 facility managers in 6 organizations, and 3) a focus group interview including an experience session using the robot with 12 blind participants.

watched the behaviors of the robot through videos embeded in the questionnaire. Because our prototype robot is different from existing autonomous service robots such as security robots and delivery robots, we created two types of videos. One shows the robot guiding a user, and the other shows the robot moving about alone. In addition, participants were asked about their preferences. The results revealed that participants felt significantly more comfortable, less obstructed, and safer with the robot guiding a user than with the robot moving about alone. We also observed that many participants would accept being captured by a camera if the data were used for assisting blind people and would not be saved.

In the second study, we interviewed 15 facility managers from 6 entities, including 3 retail stores, 2 medical facilities, and a museum, regarding concerns that may arise when introducing robots to their buildings. We showed them the videos and observed that managers expressed concern that surrounding people might misunderstand the purpose of the robot's camera, which could cause privacy-related trouble. They also commented that the robot's movements were so natural that the user would not seem visually impaired, and thus, we may need some way to indicate that the user is visually impaired. In this paper, we define these concerns as **"privacy concerns"** and **"visibility concerns"**, respectively, and discussed these concerns with blind participants in the focus group session.

In focus group sessions, we asked 12 blind participants to use our robot in an office building and then discussed concerns that may arise when using robots in public buildings. While all participants appreciated that the robot's design looks very natural and cool, they shared various opinions regarding their concerns with the robot's appearance. Five participants commented that they prefer not to be recognized as visually impaired, but seven participants agreed with the visibility concerns and suggested informing surrounding people that they have a visual impairment. Regarding privacy, while six participants did not mind capturing people in the surroundings with a camera on the suitcase, the others were concerned that they might get into trouble if people misunderstood the purpose of the camera.

On the basis of the results and findings of the three studies, we analyzed the commonalities and differences among blind people, facility managers, and people in general. In the discussion section, we further discuss how to reduce privacy and visibility concerns in addition to safety considerations. Finally, we present future work on making autonomous navigation robots more flexible and more socially acceptable for all.

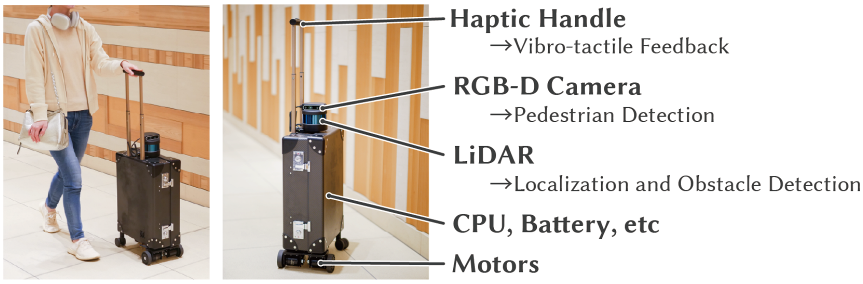


Figure 9.2: Our prototype robot uses vibro-tactile feedback on a suitcase handle; an RGB-D camera for pedestrian detection; and a LiDAR for localization and obstacle detection. Computational resources such as a CPU and battery are in the suitcase. The front wheels of the suitcase are motorized.

9.2 Related Work

We specifically focus on evaluating the social acceptance and concerns regarding an autonomous robot for navigating blind people in public buildings. In this section, we review related work specific to this chapter that were not detailed in Chapter 2.

9.2.1 Social Acceptance of Autonomous Robots

There is a large body of research investigating and discussing the social acceptance of service robots [52]. Such service robots have become part of work-life in many sectors [217]. Researches have investigated the social acceptance of shopping assistant robots [28, 151, 221], public relations robots [180], security robots [152], delivery robots [113, 184], and healthcare robots [37, 90]. Similar to the previously introduced assistive technologies, service robots involve multiple stakeholders. Niemela et al.studied the acceptance of social robots in a shopping mall with their customers, store managers, and mall managers [180]. Hebesberger et al.reported acceptance of robots in a care hospital with staff and older adults [90].

Compared with existing autonomous service robots such as security robots and delivery robots, blind navigation robots are characterized by the fact that the user always moves beside the robot. Thus, we conducted an online survey with sighted people, where we compared the social acceptance of a robot guiding a user and the robot moving about alone.

9.3 Implementation of a Robot in Public Buildings Setting

The purpose of this paper is to investigate how blind people, facility managers, and people in general perceive and accept an autonomous assistant robot that is implemented in public buildings. We designed a navigation robot with reference to prior research [24, 85, 91, 232, 260] and tested it in large multi-story buildings (a five-story shopping mall building and selected floors of a 25-stories office building). This section describes the design principles of the robot and how the designs matter to real-world scenarios.

9.3.1 Design Principle of the Blind Navigation Robot

There are three principles to the robot design.

1. **The robot considers the user who holds its handle as much as possible:**
   The navigation robot walks alongside a user, so it needs to consider not only its own body but also the user's body to compute its path and speed [85].

2. **The robot follows social norms in the building as much as possible:**
   The robot needs to consider nearby pedestrians on behalf of the user. The robot recognizes people in the surroundings so that it and the user can behave socially; avoiding people standing still, following people walking in front of the robot, social distancing, navigating through elevators, and waiting in lines [24].

3. **The robot interacts with the user through haptic and voice:**
   The robot provides its future actions in advance through haptic output on its handle and is controllable by a conversational interface on a smartphone app.

We have implemented functionalities in addition to the existing navigation robot [85, 260], including high-precision localization, social-aware navigation (e.g., , queuing and floor transitioning by using elevators), and a smartphone app to control the robot. The robot is assembled in a ready-made suitcase, which enables the robot blend in with the environment without attracting too much attention [91].

9.3.2 High-precision Localization in Multi-story buildings

First, the robot is able to know its precise location to navigate blind people to their desired destinations. Our robot utilizes radio frequency signals like Wi-Fi or Bluetooth to roughly localize its location in large multi-story buildings. Then, it uses a LiDAR sensor to localize its position and orientation in a building by matching point clouds to a pre-built map to get a precise location to usually within several inches. The current implementation uses the Cartographer ROS package[2] for point cloud registration. The system requires collecting radio frequency signals and LiDAR point clouds data in the building to build the map in advance.

9.3.3 Social-aware Navigation

In addition to localization, the robot can detect and track people in its surroundings with its RGB-D (RGB image + depth) camera and estimate the status (position, speed, walking direction) of each person. It recognizes people in an image by using Yolo V4 [34] at rate of 10 frames per second, and it tries to behave socially on the basis of the recognition results. It keeps social distance as much as possible, avoids people standing in its path, and waits in lines. To get in a line, the current implementation uses metadata for possible queuing positions. This means that the current robot can deal with pre-defined static line location like with footprint markers, which have become popular in the current pandemic situation. While in a line, the robot continuously measures the distance from the person in front of it and proceeds towards the head of the line. The robot is also able to navigate through elevators with little help from the user. It asks users to push the elevator buttons by indicating in detail the location of the buttons to call the elevator. How robots communicate with an elevator management system can often be a problem. It is usually solved by using robot hands or using a network API. In our scenario, the robot can sometimes rely on the user.

We also implemented a smartphone app as an interface for the robot in addition to the custom tactile handle of the suitcase [85]. Similar to some navigation apps [215, 250], the smartphone app utilizes a speech-to-text engine to help the user easily input their destination. The app sends the destination to the robot by using a custom Bluetooth profile. It also has function for speaking messages received from the robot with a text-to-speech engine. The user can choose a means of audio output that connects to their smartphone. We recommend that users use the smartphone speaker or an open ear headset such as a bone conduction headset.

[2] https://opensource.google/projects/cartographer

9.4 Online Survey of People in General

We first explored the social acceptance of our navigation robot with online sighted participants. The main research questions of this online survey were **"how will people in general accept our robot moving about in public buildings"** and **"how does social acceptance change between the robot guiding blind users and the robot moving about alone"**. We recruited more than 300 participants via a crowdsourcing marketplace (CrowdWorks[3]) and asked them to answer our questions after watching two types of videos that presented the features of our robot. These videos presented the same features but one video showed the robot guiding a user, while the other showed the robot moving about alone. The videos and the list of all questions are available in our supplementary material.

9.4.1 Video Stimuli

As shown in Figure 9.3, participants watched two videos: video A, which showed the robot guiding a user with a blindfold, and video B shows the robot moving about alone. Video A and B have the same content and presented the robot's four features as follows:

- **Feature 1: Navigation**
  The robot can navigate users to their destinations while avoiding obstacles.
- **Feature 2: Pedestrian Avoidance**
  The robot can avoid collisions with nearby pedestrians by stopping if a pedestrian is going to cut across in front of the robot or by moving through free space if a group of people is blocking the robot's path.
- **Feature 3: Riding an Elevator**
  The robot can detect the opening/closing of an elevator door and get on/off the elevator at the desired floor.
- **Feature 4: Standing in a Line**
  The robot can navigate the user to the end of a line and follow the line movement.

The two videos presented these features from the same place, same camera angle, and with the same movements. The only difference between the two is whether a user is next to the robot or not.

9.4.2 Procedures

On the instructions page, we first mentioned that our online survey contained secret questions for checking participation and determining compensation to encourage participants to read and answer all the questions thoroughly; it was adapted from a performance-based payment approach [96]. After that, the online survey opened with demographic questions and several questions about experiences with interacting with visual impairments and robots in public spaces. Then, participants watched either video A (robot and user) or B (robot only) and answered questions about the social acceptance of the robot presented in the video. After completing the questions, they watched the other video and also answered the same questions. The order of videos to be watched was randomized for each participant.

There were three secret questions (one in video A and two in video B) that we used to check whether online participants watched the videos: how many people were standing in lines in video A for Feature 4 (Standing in a Line), how many people blocked the robot's path in video B for Feature 2 (Pedestrian Avoidance), and how many people were in the elevator in video B for Feature 3 (Riding an Elevator). We marked a response as invalid

[3]https://crowdworks.co.jp/en/

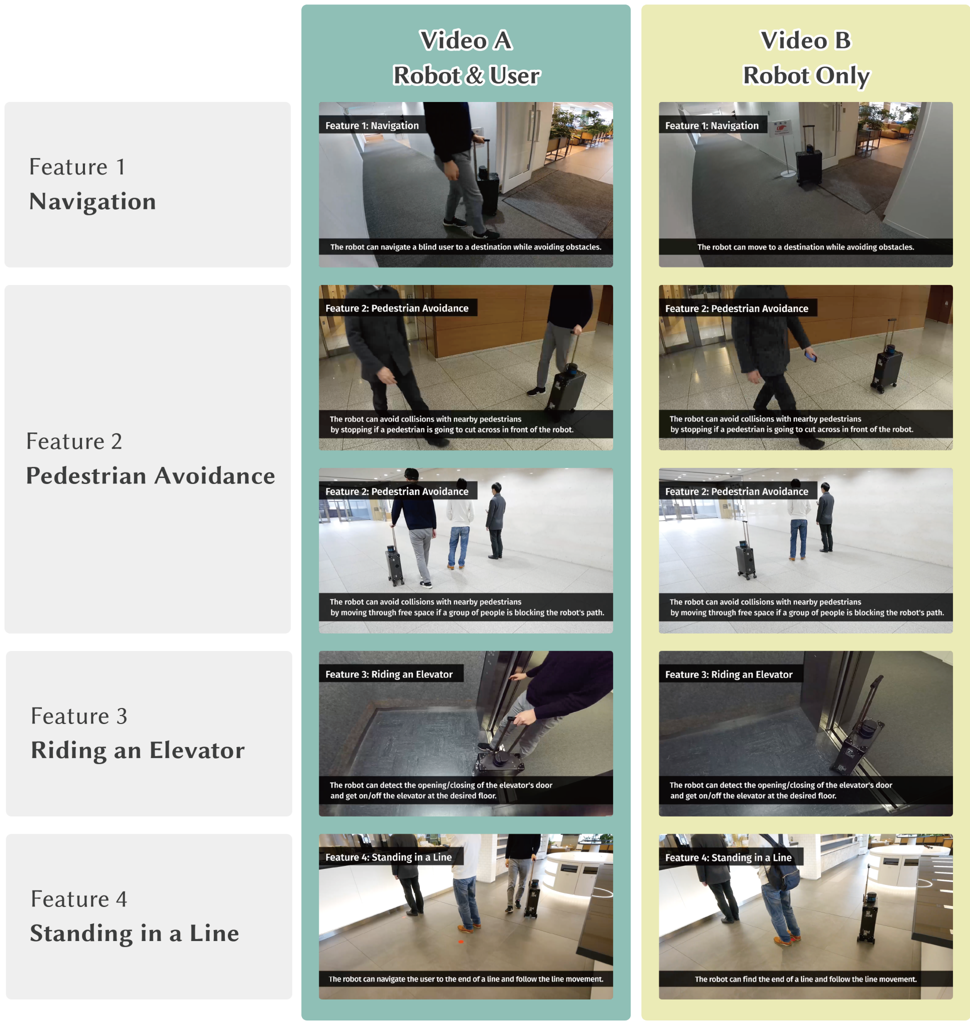


Figure 9.3: Video stimuli used in our online survey. While the two videos presented the robot's features from the same place, same camera angle, and with the same movements, video A shows the robot assisting a blind person, and video B shows the robot moving about alone.

and did not use it for our analysis if a participant had a wrong answer to any of these questions. Participants whose responses were marked as valid were compensated 1$ for their participation.

9.4.3 Questionnaire about Social Acceptance of Robot

To evaluate the overall social acceptance toward the robot, we asked participants to rate the following sentences using 7-point Likert items (rating from 1: strongly disagree to 7: strongly agree):

- *"I would feel uncomfortable if the robot were moving about 'alone' (or 'with a blind person') in public buildings."*[4]
- *"I would feel that the robot is disruptive if it were moving about 'alone' (or 'with a blind person') in public buildings."*

[4]All communication with the participants was in their native language. In this chapter, we describe any translated content in the form of *"translated content."*

- *"I would feel unsafe if the robot were moving about 'alone' (or 'with a blind person') in public buildings."*

One of our research questions is "social acceptance change between the robot guiding blind users and the robot moving about alone." Thus, we asked participants to rate how they accepted the robot after watching each video.

The other research question is "how will people in general accept our robot moving about in public buildings." To investigate the robot's camera acceptance, we asked participants to answer whether they are okay with the robot's camera capturing them in each of four conditions (C1 – C4) using 7-point Likert items. As shown in Figure 9.6, these conditions are characterized by whether the captured data is used for only blind people assistance [C1 and C2] or not (including other purposes) [C3 and C4] and whether the captured data is saved [C2 and C4] or not (one-time detection only) [C1 and C3]. Also, participants rated whether they were okay with the robot's camera capturing them if it were to be used for the robot's specific features (Features 1–4 in 9.4.1). We also asked participants to rate these questions about camera acceptance for the cases of the robot moving about alone or moving about with a blind person after watching each of the videos. Finally, participants answered what kind of personal information they do not mind being captured by the robot's camera, such as age, height, clothes, action, and so on.

9.4.4 Results

We summarize the remarkable results in this section. The list of all questions and the summary of the answers are shown in our supplementary material (Appendix A).

About Participants

While we got 315 responses in total, 15 participants incorrectly answered our secret questions and were excluded from our analysis. As a result, we acquired answers from 300 individuals (Male: 155, Female: 142, and Decline to State: 3). Participants ranged from 18 to 67 in age (Mean = 37.94 and SD = 9.58); only adults, age 18 or older, were allowed to participate in our online survey.

Sixty-three participants out of 300 had a personal, volunteer, or work experience interacting with people who had visual impairments, and 53 had helped strangers with visual impairment in public spaces. One hundred and twenty three participants out of 300 had seen robots moving about in public spaces, and only one participant had been involved in the development, promotion, marketing, or sale of robots.

Robot Only vs. Robot and User

Figure 9.4 shows the questionnaire results regarding the overall social acceptance in terms of uncomfortable, obstructed, and unsafe feelings toward the robot guiding a user or moving about alone. We compared the acceptance for each by using a Wilcoxon signed-rank test with 5% levels of significance. The p-values of each test are shown in Figure 9.4. Our analysis revealed that, for all questions, the robot guiding a user received significantly higher social acceptance than the robot moving about alone.

Figure 9.5 shows the questionnaire results regarding the acceptance of the camera for each feature (Feature 1–4). We also compared the robot guiding a user and moving about alone for each feature using a Wilcoxon signed-rank test, and the p-values of each test are shown in Figure 9.5. Our statistical analysis revealed that the robot guiding a user received significantly (at the levels of 5% for Features 2 and 3 and 0.1% for Features 1 and 4) higher camera acceptance for all features than the case of the robot moving about alone.

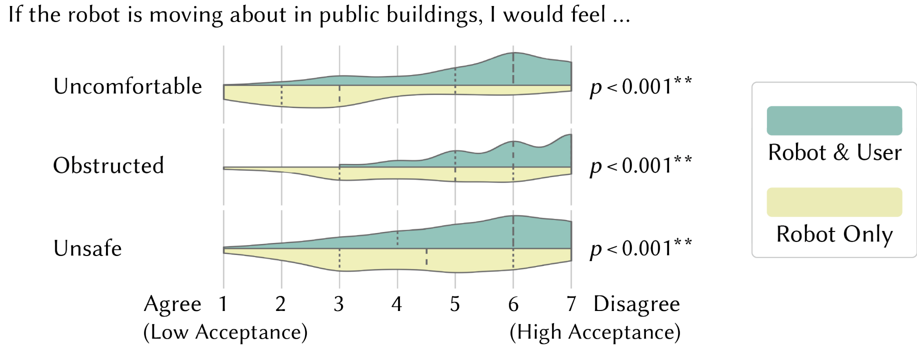


Figure 9.4: Overall social acceptance for cases in which the robot is guiding a user (Robot & User) and moving about alone (Robot Only). p: p-value of the Wilcoxon signed-rank test done for each question (* and ** indicate the significance found at the levels of 0.05 and 0.001, respectively).

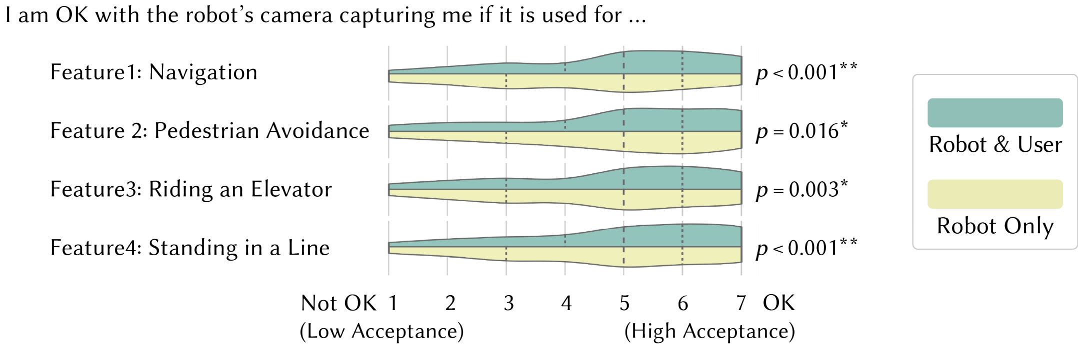


Figure 9.5: Camera acceptance for cases in which the robot is guiding a user (Robot & User) and moving about alone (Robot Only). p: p-value of the Wilcoxon signed-rank test done for each question (* and ** indicate the significance found at the levels of 0.05 and 0.001, respectively).

Camera Acceptance

Figure 9.6 shows the results of the camera acceptance of the robot guiding a user when we compared the acceptance in four conditions (C1–C4). Figure 9.6 also shows the p-values of a Wilcoxon signed-rank test with 0.1% levels of significance. Our analysis revealed that the camera acceptance was higher when the captured data was used for assisting blind people [C1 and C2] rather than including other purposes [C3 and C4]. In addition, we observed that the camera acceptance was higher when the captured data was used for one-time detection only [C1 and C3] rather than when the data was saved [C2 and C4]. As a result, the highest acceptance was for C1 (using the captured data for assisting blind people and one-time detection only), followed in order by C2, C3, and C4. When the captured data was used for assisting blind people only and was not saved at all [C1], 75.3% of participants answered that they were OK with the robot's camera capturing them (5–7 points). However, 19% of participants answered that they did not want to be captured by the robot's camera.

9.5 Interview with Facility Managers

Our ultimate goal is to implement our robot in public buildings (e.g., , shopping malls, hospitals, and museums) and to allow blind people to walk independently and freely in such buildings by using it. To explore the concerns that may arise when introducing robots to public buildings, we conducted semi-structured interviews with 15 facility managers in 6

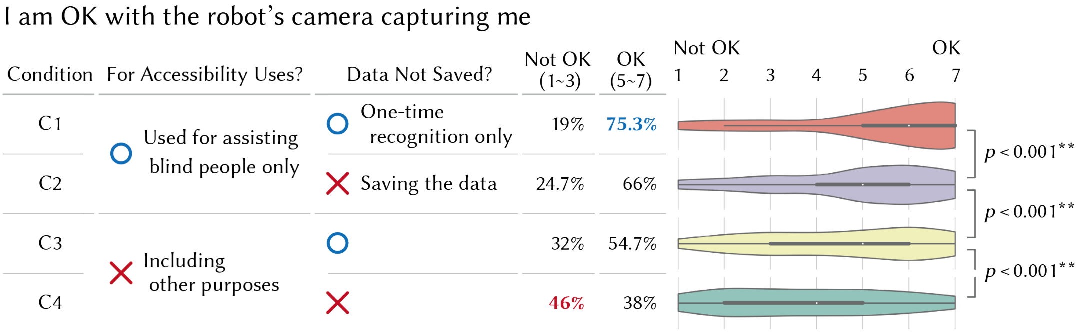


Figure 9.6: Overall camera acceptance for each condition. p: p-value of the Wilcoxon signed-rank test done on each condition (** indicates the significance found at the levels of 0.001).

Table 9.1: Interview Participants.

ID	Facility Type	Position (Number of People)
F9.1	Shopping Mall	Tenant Management (1)
F9.2	Rehabilitation Center	Vice Director and O&M Specialist (1) Technical Advisor (1)
F9.3	Polyclinic	Management Improving Team (1)
F9.4	Real Estate Development Co.	Corporate Social Responsibility Promoter (1) Customer Support (1)
F9.5	Science Museum	Facility Management (1) Visitor Service (2) Strategy Management (2)
F9.6	Discount Store	Corporate Officer (1) Business Consultant (2) Store Designer (2)

organizations (Table 9.1). Out of the six organizations we interviewed, two had contacted the authors' research group in the past, and four were contacted for the first time.

9.5.1 Procedure

Because of the COVID-19 outbreak, all interviews were conducted over videoconferencing. The sixth author was the primary interviewer in all interviews. The first and second authors attended most interviews as secondary interviewers who asked follow-up questions. Interviews started by asking participants how they assist blind people who visit their facilities:

- *"Do you have a manual or guidelines for blind visitors?"*
- *"Have you had any troublesome situations that occurred when a blind person visited your facilities?"*

Then, participants watched a video that showed the robot guiding a user and presented the robot's four features. The video was used in our online survey as video A. The details of the video contents were described in Section 9.4.1. After watching the video, we asked participants questions about the concerns that may arise when introducing the robot to their facilities:

- *"Do you have any safety concerns about the robot moving about in your facilities?"*
- *"What criteria and processes must be needed so that your organization judges the robot to present no safety problems?"*
- *"Do you have any privacy concerns about the robot's camera?"*
- *"Do you have any suggestions for improving the robot?"*

The interviews lasted approximately 60 minutes each. The sessions were audio recorded and transcribed for further analysis.

9.5.2 Findings

Policies for Blind Visitors

Three facilities (a rehabilitation center, polyclinic, and science museum) had a manual or guidelines for assisting blind visitors:

> **A8.1:** *"As part of staff training, we take a course to experience a simulation of visual impairment by wearing goggles."*
>
> F9.2: Rehabilitation Center, Technical Advisor

> **A8.2:** *"We have instruction manuals for assisting people with disabilities and manuals that mention visually impaired people. Although our facility does not have a system for directly guiding visually impaired people to exhibits, we have a supplementary tool to help them understand exhibit contents through voice."*
>
> F9.5: Science Museum, Visitor Service

> **A8.3:** *"We have instruction manuals for taking care of people with disabilities, including visually impaired people."*
>
> F9.3: Polyclinic, Management Improving Team

Three other organizations (a shopping mall, real estate development corporation, and discount store) did not have such manuals for blind visitors:

> **A8.4:** *"In our facility, there are no clear rules or documents on how to assist visually impaired people."*
>
> F9.1: Polyclinic, Shopping Mall, Tenant Management

> **A8.5:** *"Our company does not have a department that handles accessibility. Our company is not well prepared for assisting people with disabilities."*
>
> F9.6: Discount Store, Business Consultant

> **A8.6:** *"We do not have guidelines or training courses so far. I cannot say our facility actively takes care of visually impaired people."*
>
> F9.4: Real Estate Development Co., Customer Support

Positive Comments

F9.1, F9.3, and F9.5 commented that our suitcase-shaped robot's appearance and movements were natural and intelligent:

> **A8.7:** *"I was surprised that the robot was in the shape of a suitcase. The robot's movements were so intelligent that surrounding people might not notice that the user was visually impaired."* F9.3: Polyclinic, Magagement Improving Team

A8.8: *"This robot was moving about naturally, and I did not feel any safety issues with it. ... The size of the robot is an advantage of this robot. Because this robot is about the size of a suitcase, stores can respond to blind users like they do to customers carrying a suitcase and do not need to take any special measures."*
F9.1: Shopping Mall, Tenant Management

F9.3 and F9.4 mentioned that this navigation robot would be especially useful for people visiting the facility for the first time:

A8.9: *"I think the best use case for this robot is to guide our patients. Since the hospital is large and complex and it is difficult to know where to go, especially for first-time patients, this robot would be helpful for such people."*
F9.3: Polyclinic, Magagement Improving Team

A8.10: *"I think this robot would be useful not only for people with visual impairments but also for other people. It would be a good idea to introduce this robot to facilities where many people, such as inbound tourists, are visiting for the first time."* F9.4: Real Estate Development Co., Customer Support

Privacy Concerns

All organizations answered that they would be OK with the robot's sensors capturing the surrounding information if the captured data is used for assisting blind visitors:

A8.11: *"If the robot is used for assisting blind people and does not record surrounding information, we would like to ask stores in our shopping malls to accept the robot the same as guide dogs. I think it would be essential to inform the stores of the robot's purpose and the handling of the captured data in advance."*
F9.1: Shopping Mall, Tenant Management

However, all organizations except for F9.6 were concerned that surrounding people might get suspicious of the cameras attached to the suitcase. They suggested announcing the use of the cameras and handling of the data to visitors of the facility. Specifically, they suggested that a sticker should be attached to these suitcases to inform people who see the system of the purpose of the robot and that posters at the entrance or in-house announcements should inform visitors that an assistance robot for blind people is working in the building:

A8.12: *"There is a concern that customers may misunderstand the purpose of the robot's camera, which could cause some trouble. When we installed a temperature monitoring system during the COVID-19 outbreak, we received an inquiry about how the face recognition data is managed. If the robot lets the surrounding visitors know that it is used for supporting blind users, they will accept the robot's camera."*
F9.3: Polyclinic, Magagement Improving Team

A8.13: *Since service robots are still not common, and some visitors may be suspicious of the robot's camera, it would be good to announce that the robot is lent by the facility. For example, we can put stickers on robots and stores, broadcasts inside the building, and use digital signage."* F9.1: Shopping Mall, Tenant Management

A8.14: *"We should inform visitors that an assistive robot is moving throughout the museum and how captured data is handled by putting a poster at the entrance. Since we agree that blind users do not want to attract people's attention, the robot itself may not have to be overly conspicuous."*
F9.5: Science Museum, Facility Management

Safety Concerns and Visibility Concerns

While the robot's movements were described as so natural that the user would not seem visually impaired (A8.7 and A8.8), four organizations (F9.1 and F9.3–F9.5) commented that, for safety reasons, the robot and user should inform people that the user is visually impaired:

A8.15: *"If surrounding people notice users of guide dogs, they will move out of the way. This robot may be perceived as a suitcase for travel. I think that surrounding people would not notice that the user is visually impaired or would not avoid them."*
F9.1: Shopping Mall, Tenant Management

A8.16: *"If surrounding people notice people who are visually impaired, they will avoid collision. In the case of hospitals, surrounding staff members will help. If the user does not mind, the system should inform surrounding people that the user is visually impaired."* F9.3: Polyclinic, Management Improving Team

A8.17: *"I think that the robot will assimilate into surrounding environments and that people will not avoid it. Because the crowdedness of the facility changes depending on the season, I'm worried about whether the robot can avoid collision in crowded situations. If the robot informs surrounding people that the user is visually impaired, people could avoid them, reducing the risk of collision."*
F9.4: Real Estate Development Co., Corporate Social Responsibility Promoter

In this paper, we defined these concerns about whether blind people should notify their presence to surrounding people as a **"visibility concern"**, and we discussed it with blind people in a focus group session (Section 9.6).

In relation to the safety concerns that appear when introducing autonomous robots to public facilities, the need for criteria that can be used to verify the safety of robots was mentioned by participants:

A8.18: *"Because our facility does not have any past cases of safety verification, the introduction of robots will face high barriers in terms of safety at this stage. If public guidelines issued by the government, standards, and safety tests could be used for objective evaluation, I think they could be criteria for introducing robots."*
F9.4: Real Estate Development Co., Customer Support

A8.19: *"Our facility does not have any standards for evaluating safety. The ISO standards have three types[5] of standards for service robots [ISO 13482]. This suitcase would not apply to any of them, so we need to define a new standard."*
F9.2: Rehabilitation Center, Vice Director and O&M Specialist

A8.20: *"As for guidelines regarding the safety of robots, our stores are unusual environments (corridors are very narrow). Even if there are no legal problems, we do not know if our stores would not have any problems, so I think we need to verify this in our stores."* F9.6: Discount Store, Corporate Officer

As other safety concerns, we got the following comments from participants:

A8.21: *"Many of our visitors are children. I'm worried about the risk of children making unexpected movements and colliding with the robot."*
F9.5: Science Museum, Visitor Service

A8.22: *"In our stores, the corridors between shelves are narrow [1 m–1.6 m]. There is the risk of shoppers or clerks tripping."* F9.6: Discount Store, Store Designer

[5] Mobile servant robot, physical assistant robot, and person carrier robot.

Table 9.2: Demographic descriptions of focus group participants.

ID	Age	Gender	Eyesight	Primary Aid
P9.1	26	Male	Blind since age 10	White cane
P9.2	33	Female	Blind since age 14	White cane
P9.3	23	Male	Blind since age 4	White cane
P9.4	38	Male	Blind since age 30	White cane
P9.5	56	Male	Blind since birth	White cane
P9.6	43	Female	Blind since age 3	White cane
P9.7	64	Male	Blind since age 42	White cane
P9.8	31	Male	Blind since age 26	Guide dog
P9.9	29	Male	Blind since age 5	White cane
P9.10	61	Female	Blind since age 45	White cane
P9.11	25	Male	Blind since age 15	White cane
P9.12	52	Male	Blind since age 8	White cane

9.6 Focus Group with Blind People

In the sessions, we discussed safety, privacy, and visibility concerns with the robot mentioned in the online survey of people in general and the interviews with facility managers.

9.6.1 Participants

We recruited 12 blind participants (9m/3f), with ages ranging from 23 to 64 years old (Table 9.2). For the recruiting, we used an e-newsletter for people with visual impairments and the participants were compensated $90 for their time. We recruited blind people who satisfy the following conditions: (1) consider themselves to have good orientation and mobility skills; (2) often travel independently by using their cane or guide dog; and (3) familiar with using their smartphone.

9.6.2 Procedure

We had three sessions with four participants each (P9.1–P4, P9.5–P9.8, and P9.9–P9.12). This study was organized into two sections. Participants first tried to use our robot in an office building. Then, participants and researchers conducted a focus group [81] that discussed how to widely adopt autonomous navigation robots in the real world. The focus group session took around 90 minutes, while the whole study took approximately 2.5 hours, and blind participants were compensated 90 $ for their time.

Trial Session of Our Robot

Figure 9.7 shows the routes we used in the study. The routes included a long route (approximately 160 m) on the first floor and a short route using an elevator to move to another floor. After obtaining (IRB approved) informed consent from participants, researchers gave an overview of the study and described the interface of the robot. We especially explained how to place their left hand on the handle of the robot, how to change the robot's speed, and the vibration signals provided from the suitcase handle.

Then, four participants were asked to walk on the route on the first floor (approximately 160 m) using the robot. As shown in Figure 9.7, when participants walked on the route, 1)

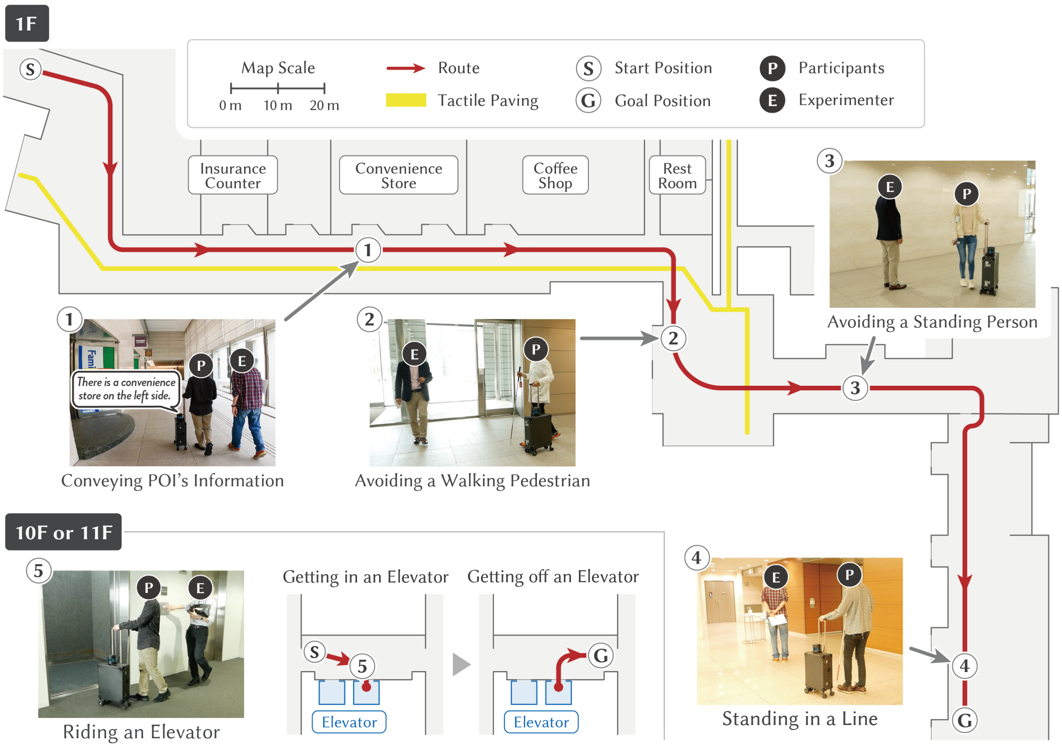


Figure 9.7: The route used in the focus group with blind people. **1F:** We first asked participants to walk along the route (approximately 160 m) on the first floor. The route included points where (1) the robot conveyed information on POIs, (2) an experimenter crossed the participants' route, (3) an experimenter blocked the route, and (4) participants stood in a line. **10F or 11F:** We then asked participants to move from a floor to another floor by riding an elevator with the robot (5).

the robot gave information on the surroundings (e.g., *"There is a convenience store on the left side."*), 2) one experimenter crossed the participants' path, and the robot stopped to avoid collision, 3) one experimenter blocked the path, and the robot avoided them by moving through free space, and 4) the robot navigated the user to the end of the line and they followed the line movement. On the first floor, people constantly walked into or out of the office, restroom, coffee shop, convenience store, etc. While participants were walking with the robot, a researcher was walking behind them to explain the features of the robot and guarantee their safety as well as other pedestrians' safety.

After reaching the goal position on the first floor, participants moved from one floor to another by riding an elevator with the robot (Figure 9.7–5). Two participants moved from the tenth floor to the eleventh floor, and two other participants moved from the eleventh floor to the tenth floor.

The trial session took around 30 minutes per participant. We conducted the session using two robots and finished all trial sessions with four participants in 60 minutes. During the tasks, we recorded how participants used the system with a video camera for further analysis.

Focus Group Session

After finishing all routes, researchers performed a focus group session with participants. The session was semi-structured to focus on the safety, privacy, and visibility concerns that were pointed out in our online survey and interviews. We also asked the participants in what scenarios would they use our robot. The session was audio recorded and transcribed for further analysis.

9.6.3 Findings

Visibility Concerns

All participants appreciated that the robot's design looked natural and cool. When we asked them about the visibility concerns pointed out by the facility managers (i.e., , blind people should inform others that they are visually impaired for safety), five participants (P9.1–P9.3, P9.6, and P9.10) commented that they would not want to emphasize that they are visually impaired:

A8.23: *"It is great that the robot looks stylish. I thought the robot would look more like a navigation machine, but it looks like a natural-looking suitcase and cool. I prefer looking natural and to not to be recognized as visually impaired. If the robot looked unsophisticated or conspicuous, I would not want to use it, but I love that this robot is natural."* P9.1

A8.24: *"I do not want to emphasize that I am visually impaired. I'm trying to live so that I do not look visually impaired. It is great that the design of the robot is based on a suitcase and is natural and modest."* P9.2

A8.25: *"It is good that this suitcase-shaped robot may not make me look like a visually impaired person, unlike when walking with a guide dog, which may make it obvious."* P9.6

A8.26: *"I do not use a white cane as a symbol of visual impairment. I use it as a necessary aid for walking. If I can feel safe with the suitcase, I would trust it and walk without the cane."* P9.3

Seven participants (P9.4, P9.5, P7–P9.9, P9.11, and P9.12) agreed with the visibility concerns and suggested showing that users have a visual impairment. P9.1, P9.2 and P9.11 also commented that, when they bump into someone, showing their white cane can reduce the possibility of being in trouble:

A8.27: *"I usually walk with a white cane. I'm using a white cane for a walking aid and making people aware I am visually impaired. I think that the suitcase-shaped robot looks cool, but I'm a little worried about the surrounding people not noticing me. I want to show that I have a visual impairment in some way. For example, designing a new symbol for suitcase robots and putting it as a sticker on the robot and promoting the symbol such as through SNS."* P9.4

A8.28: *"While I understand that some blind people don't want to emphasize that they are visually impaired, I also realize the advantage of it. For example, when I am walking with my guide dog, surrounding people sometimes ask me, 'do you want some assistance?' "* P9.8

A8.29: *"When I get into an accident such as a collision with someone, if they are aware that I am visually impaired, it can reduce the possibility of me being in trouble."* P9.11

A8.30: *"When I bump into someone, it is important to show that I am visually impaired. If I have a white cane, the person may apologize, so I can feel safe."* P9.12

Safety Concerns

Participants shared various impressions regarding their safety concerns with the robot. Although only one participant (P9.2) tried the robot without holding cane, six participants

(P9.1, P9.2, P9.6, P9.7, P9.9, and P9.10) commented that they would not use a white cane while walking with the robot to keep another hand free:

> **A8.31:** *"I think there is less risk of collision in a place like today's place (office environment). However, I wondered how the robot would behave in crowded situations such as stations during rush hours."* P9.12

> **A8.32:** *"Because the suitcase slowed down its speed when it moved near walls and people, I was not concerned about endangering surrounding people."* P9.4

> **A8.33:** *"I folded the white cane away and walked with the robot. If I were also using the white cane, I would have concerns that both hands are busy. So, I would not use a white cane while walking with the robot."* P9.2

> **A8.34:** *"I think I would not use the cane while walking with the robot. I want to move about by putting one hand on the robot and keeping the other hand free because I don't want both of my hands busy."* P9.10

> **A8.35:** *"I hope I can fold away my white cane because I do not want both to be hands busy. If the robot can be trusted to also recognize gaps, I want to walk without a cane."* P9.6

Privacy Concerns

Our study revealed that, while some facility managers were concerned that surrounding people may misunderstand the purpose of the robot's camera, which could cause trouble, people in general would tend to accept the robot's camera if it were used for assisting blind people only and the captured data were not saved at all. When we described these results regarding privacy concerns, while six participants (P9.1 and P9.2, P9.6, and P9.9–P9.11) did not mind surrounding people being captured with the camera on the suitcase, the other participants (P9.3–P9.5, P9.7, P9.8 and P9.12) were concerned that they might get into trouble if people misunderstood the usage of the camera:

> **A8.36:** *"Surveillance cameras are widely used and accepted because the usages of these cameras are understood by society. Similarly, the suitcase's camera will be accepted if surrounding people understand that the camera is used for assisting visually impaired people."* P9.7

> **A8.37:** *"If surrounding people do not understand that the camera on the suitcase is necessary for supporting visually impaired people, there is a concern that people will misunderstand the usage of the camera and the user might get into trouble. I hope the system will be widely used, and surrounding people will understand the usage of the camera."* P9.3

> **A8.38:** *"If surrounding people will be concerned about privacy and so on, I think it might be better to clarify the usage of the camera on the suitcase."* P9.5

Prospective Scenarios for Autonomous Navigation Robots

All participants appreciated that they could walk alone by following the robot's movement only. When asked about in what scenarios the participants would use our robot, the participants commented that they would want to use the robot to walk alone while feeling the surrounding information and atmosphere, which is difficult for them to do with a white cane:

> **A8.39:** *"I hope I can walk around the neighborhood by using the robot. Visually impaired people walk to clearly defined destinations by preparing in advance well or utilizing some tool. I want to use the robot for taking a walk without a specific destination. The robot has a great advantage in that I can walk alone without having to worry about anyone else."* P9.3

A8.40: *"I want to do window shopping while listening to surrounding information at a shopping mall. I can do some shopping if someone helps me, but I feel sorry asking when I do not have specific purpose. By using the suitcase, I want to walk freely in a shopping mall by myself and do window shopping."* P9.2

A8.41: *"When I usually move about, I think only about 'which intersection to turn at' or 'which traffic light to turn at.' With the guidance of the robot, I want to walk while sensing the surrounding atmosphere and information."* P9.4

9.7 Discussion

9.7.1 Social Acceptance of Suitcase-shaped Navigation Robot

In this study, we observed that social acceptance would be higher for navigation robots assisting blind people than for robots operating alone in the study with people in general. The robot guiding a user received significantly higher social acceptance than the robot moving about alone, as indicated in all questions (Section 9.4.4). One of the major concerns, in general, is the use of cameras, but acceptance was higher if the captured data is to be used for assisting blind people for one-time detection without storing the data (Section 9.4.4). This effect is similar to the AT-effects of HMD usage [198], in other words, people tend to accept technology if it is used for assistive purposes Still, this finding is unique since autonomous robots usually move around alone, and there are fewer use cases of them being accompanied by a human. Our findings suggest that this characteristic will lower the barrier to deploying robots in public buildings. While we have yet to further investigate this, these findings may suggest that other use cases of robots accompanying people with disabilities may have similar advantages, such as autonomic wheelchairs or navigational shopping carts for the elderly. The results suggest that providing navigation to people with disabilities can open the door to the deployment of service robots in public buildings in the near future.

9.7.2 Visibility Concerns

The robot was designed to appear as a standard suitcase to make it possible for the robot to assimilate into the environment in public buildings. The design principle was successfully adapted, and facility managers commented that the use cases were not distinguishable from those in which a person is walking with a standard suitcase. Throughout the study, we received divergent opinions on this aspect.

All blind participants rated the design highly or expressed that they were comfortable because the robot was able to assimilate into the surrounding environment (A8.23–A8.26). However, facility managers from four out of six organizations mentioned concerns with such assimilation. They thought that it should be clear to others that a person is visually impaired so that they are safer. The managers thought that a sighted visitor can proactively avoid a collision if they notices a blind person is approaching them (Section 9.5.2).

Five out of 12 blind participants maintained their preference toward the seamless-look in public buildings even after understanding the existence of the visibility concerns (A8.23–A8.26). One notable comment was, *"I prefer to look natural and not be recognized as visually impaired"* (A8.23). On the other hand, seven blind participants suggested ways to make visual impairment clear to others such as through the use of a new symbol for blind navigations robot that can be attached to the robots (A8.27 and A8.28). Three blind participants agreed that there are situations in which they need to make their blindness visible, such as in a crowded environment (A8.29 and A8.30). In such environments, they prefer to keep white canes in hand. Facility managers also suggested solutions such as putting a signboard on top of robots, publicly announcing the existence of navigation robots, or displaying posters to explain the purpose and features of the robots (A8.15–A8.17).

Faucett et al. also reported similar conflicts. They described that externally imposed disability identities may stigmatize those with disabilities as being incapable, or even stereotypically unkempt, and unfashionable. In contrast, they also mentioned that the visibility of the assistive technology, white cane as an example, affords credibility, allowing the user to communicate their needs without words [65].

As for legislation topic, in some countries such as Japan, Germany, and Austria, traffic regulations presume that visually impaired people will carry a white cane or walk with a guide dog to be "visible" in order to use public roads. In other countries such as the US and UK, there are no rules expecting the visually impaired to have a mobility aid making them visible, but drivers are asked to yield to visually impaired pedestrians using mobility aids. When navigation robots for the blind become mature and ready for use in public places, amendments to regulations will be required in some countries.

The challenge regarding the visibility concern can be one of the essential challenges we face for broader deployment. Visibility may have the effect of educating the general public about the necessity for such technologies and, consequently, increase social acceptance. Visibility is mandated in some countries as a part of traffic safety. In addition, blind users tend to feel comfortable assimilating into public buildings. We realized that we need studies to seek a balance between visibility and assimilation along with broadening the usage of such technologies and then propose amendments to regulations to reflect the latest technologies.

9.7.3 Privacy Concerns

Privacy concerns were major aspects we tried to investigate on the basis of previous research [3, 11, 12, 138, 198]. An RGB-D (RGB image + depth) camera was attached on top of the robot, and we expected that the feature might invoke concerns for privacy. We found that even the general public relatively accepts the usage of a camera device for accessibility purposes (Section 9.4.4); 75% of people accepted the usage of a camera as long as it were only for blind users and data was not saved. However, only 38% of people accepted the use of a camera if the device were used not only for the blind and data was saved. Facility managers from five out of six organizations expressed concerns about capturing images inside their facilities (A8.12–A8.14). One thing that they were in unanimous agreement about was that visitors might accept the camera if the purpose of assisting the blind were evident and intuitive. Thus, the idea of increasing visibility, as discussed in the previous section, is one solution to this privacy concern (A8.12–A8.14). Six blind participants expressed concern that there could possibly be misunderstandings regarding the non-consent nature of using the camera (A8.38). They agreed to make the purpose of the robot visible.

Privacy concerns are usually considered to be the most serious challenge toward practical deployment. We found that it may not be a deal breaker for social acceptance and practical deployment among stakeholders if it is clear that the purpose is to assist visually impaired people. We still need to explore ways of balancing visibility and comfort for users, but at least, the problem space we need to solve can be narrowed down to the methods and levels of visibility.

9.7.4 Safety Concerns

Safety is one obvious concern for practical deployment. As for the subjective sense of safety for the blind people, six participants reported that they felt safe and would not use a white cane (A8.31–A8.35). In contrast, there is no known objective safety measure for such autonomous navigational robots, and this fact made facility managers uncomfortable (A8.18–A8.20).

There is one set of standards for services robots, ISO 13482 [103], that define three types of robots: 1) mobile servant robots, 2) physical assistant robots, and 3) person carrier robots. There is no definition for autonomous navigation robot for visually impaired people in the standards, but it could be categorized as a physical assistant robot. Complying with such standards could be a promising way of lowering the barrier to implementation for facility managers and also increasing safety for users. It is necessary to provide our information to standards organizations to clarify the definition of the robot and requirements.

9.7.5 Limitations

This study was conducted in Japan. Social acceptance may vary by country, so we hope researchers in other countries will be able to conduct comparative studies. Such comparison may reveal new ways of approaching the public to improve acceptance.

This study is based on data collected in a short period of time, about two months, and this data can be considered as a first impression data of this technology. It is expected that social acceptance may change over time in accordance with exposure to use cases or related information like global trends. We believe that a longitudinal study should be conducted to understand the long-term social acceptance trends among all stakeholders.

In the online survey and the interview sessions with facility managers, participants did not see the real robot moving in the wild with an actual user in person. It was difficult to show the real robot in person due to the current pandemic situation. We designed videos to be understandable and also tried to cover major scenarios. However, demonstrating a real robot at the facilities of these managers may make it easier for them to imagine practical deployment. We will keep interviewing facility managers on how we can expand pilots and deployments.

9.8 Conclusion

We investigated acceptance and concerns regarding autonomous navigation robots for blind people in public buildings by conducting three studies: an online survey of people in general, interviews with facility managers, and a focus group interview with blind participants. We found that acceptance was higher when the blind navigation robot accompanied a blind person than when the robot moved about alone from online survey participants. However, facility managers expressed concern that their customers might misunderstand the purpose of the robot's camera, which could cause trouble because it would not be clear that the camera was being used to aid those with visual impairments. We then discussed privacy and visibility concerns with blind participants. They appreciated that the robot had the potential to assimilate into the surrounding environment, and five of them commented that they prefer not to be recognized as visually impaired. We then discussed how we could fill the gap between facility managers and blind users. One possible solution is to increase awareness by putting a sticker on the suitcase and/or putting up posters in buildings. Further investigation is needed to seek a balance between visibility and assimilation while broadening the usage of such robots. As future work, we would like to make our robot smarter to offer new mobility experiences such as those that the participants shared during the focus group (A8.39–A8.41). For example, window shopping, finding a favorite restaurant, and walking around the neighborhood while feeling the surrounding atmosphere.

Acknowledgement

We would like to thank all participants who took part in our user study. We would also like to thank the anonymous reviewers for their helpful comments. This work was supported by JST-Mirai Program (JPMJMI19B2), JSPS KAKENHI (JP20J23018), and Consortium for Advanced Assistive Mobility Platform.

Part IV

Discussion and Conclusion

Chapter 10

Discussion

> *"I have a risk of collisions and some other accidents while walking with a white cane only. Why do they [facility managers] worry too much about the safety issues of assistive systems? I think it's much safer to use a combination of a system and a white cane than to walk with a white cane only. "*
>
> Participant 1
> IBM Japan Ltd
> March 25, 2021

10.1 From the Laboratory to the Real World

As shown in Figure 10.1, our systems were evaluated in various situations, from the controlled environments of buildings to dynamic real-world environments. Our four projects were evaluated in the following controlled environments:

- **One-shot wayfinding system:**
  Wayfinding decision-making situations reproduced in public buildings (Section 4.1).
- **BlindPilot:**
  A laboratory space containing an empty chair (Section 8.3).
- **Corridor-Walker:**
  Controlled corridors in which people other than the participant were barred from entering (Section 4.2).
- **LineChaser:**
  Queues formed by two–four sighted people whose movements were controlled by the experimenters (Section 7.5).

Our systems were also evaluated in the following real-world environments:

- **Open space in an office building:**
  To evaluate our suitcase-shaped guiding system, we asked blind participants to walk along a set route on the ground floor of an office building. During the evaluation, other people constantly walked into or out of the office, restroom, convenience store, coffee shop, and other places (Section 6.4).
- **Crowded area in an International Airport:**
  To evaluate BBeep, we asked blind participants to walk through a crowded area of Pittsburgh International Airport. We selected several crowded gates where passengers were waiting to board, either in a line or in groups (Section 5.6).

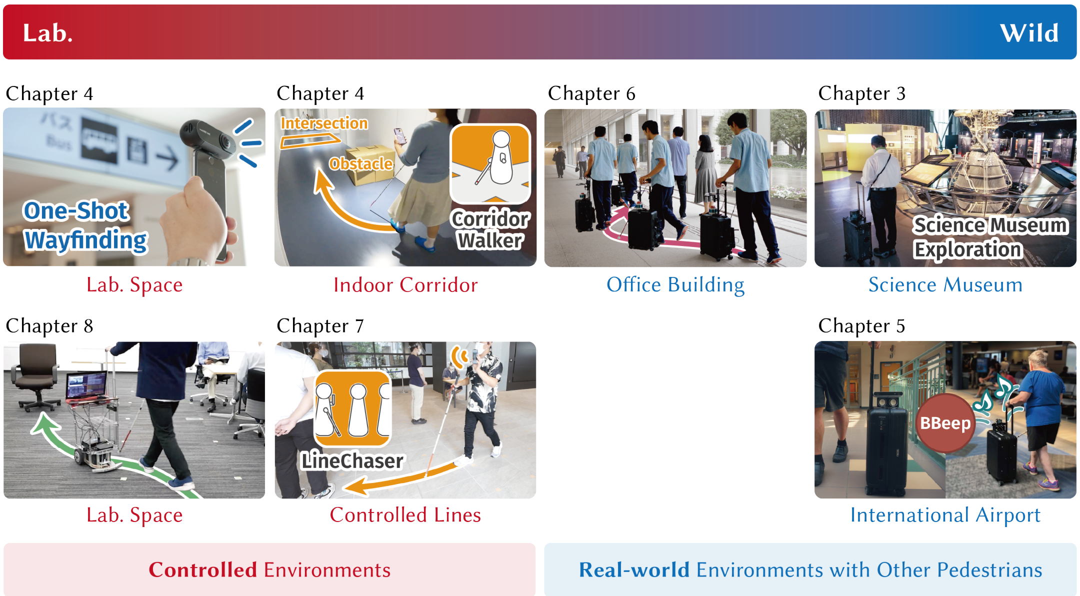


Figure 10.1: Summary of the environments in which our user studies were performed: laboratory, indoor corridor, office building, science museum, and an international airport.

- **A science museum during opening hours:**
  To evaluate our museum exploration system, we asked blind participants to freely explore and experience one floor of the museum for 90 minutes using our system. During the study, we also asked sighted visitors to complete a short questionnaire about the robot's social acceptance (Section 3.4).

Based on these experiences, this thesis suggests that **future accessibility research be conducted not in the laboratory, but in the dynamic environments of the real world**.

10.1.1 Workable Systems in Real Environments

Our assistive systems based on current computer vision and robotics technologies were safely implemented in real-world environments. Blind participants using BBeep could walk through the crowded area of an international airport without colliding with nearby pedestrians (Chapter 5). Blind participants could avoid walking and standing pedestrians in a real-world environment by continuously adapting their speed and direction, mimicking the behaviors of sighted pedestrians (Chapter 6). Blind participants could safely and freely explore the science museum during its opening hours (Chapter 3). These results suggest that current technologies are reaching a level that can assist real users in the real world. Future accessibility research should be extended from the controlled environs of the laboratory to real-world environments. In complex real-world environments, the evaluation will encounter many unexpected events and provide various findings.

10.1.2 High Social Acceptance of Assistive Technologies

We found that accessibility research is amenable to testing in real-world studies and social implementation In Chapter 9, we observed that the social acceptance level of the general public is higher for navigation robots assisting blind people than for robots operating alone. In particular, the robot guiding a user received significantly higher social acceptance than the robot moving about alone. The use of cameras roused concern but the acceptance was higher if the data were captured once for assisting blind people and later discarded rather

Figure 10.2: Summary of the environments in which our user evaluations were conducted, ranging from the laboratory to an international airport.

than stored. People tended to accept technology designed for assistive purposes. This characteristic of human perception is expected to lower the barrier of deploying robots and other assistive technologies in public buildings. Although further investigation is required, these findings suggest similar advantages in other use cases of robots accompanying people with disabilities, such as autonomic wheelchairs and navigational shopping carts for the elderly. Therefore, providing navigation to people with disabilities can open the door to deploying service robots in public buildings in the near future. Accessibility research can potentially lead social implementations of the latest technologies, including autonomous robots.

10.1.3 Various Form Factors for Various Users and Situations

Assistive systems intended for ubiquitous use by blind people require a properly designed hardware system. This thesis employed various types of devices, as described below and shown in Figure 10.2.

- **Autonomous Robots:**
  Our navigation robots are equipped with an intuitive handle that blind users can grasp while being guided (Chapters 3 and 8). These robots have a LiDAR sensor that localizes the robot's position, an RGBD camera that recognizes surrounding objects (e.g., pedestrians and chairs), a haptic handle that provides vibro-tactile feedback, a mini PC with a GPU, a battery, and motors.

- **Suitcase Shaped Devices:**
  Our collision-avoidance systems are implemented by attaching devices to a standard rolling suitcase (Chapters 5 and 6). These systems have a LiDAR sensor that detects surrounding obstacles, RGBD cameras that detect the positions of surrounding pedestrians, and a laptop that processes the data. In addition, BBeep has a speaker that alerts nearby pedestrians to a potential collision risk (Chapter 5). In Chapter 6, we designed a tactile interface that warns of collision risks with a vibrating handle and navigates users with a newly developed directional lever that shows the correct direction. Both tactile devices are attached to the handle of the suitcase.

- **360-degree smartphone camera:**
  Our wayfinding system uses a 360° camera attached to a smartphone (Section 4.1).

- **Smartphone only:**
  Our two assistive systems require only a smartphone for use (Section 4.2 and Chapter 7). Using a LiDAR sensor equipped to an iPhone 12 Pro [17], Corridor-Walker constructs a 2D occupancy grid map of the surrounding environment (Section 4.2). LineChaser uses the smartphone's built-in RGB camera and an infrared depth sensor, respectively, to detect nearby pedestrians and estimate their 2D positions on the map (Chapter 7).

Of course, autonomous robots with many sensors and devices can provide highly intelligent assistance. However, smartphone-based systems will likely become popularized in the blind community because most blind people already own a smartphone.

In future work, we hope to distribute our smartphone assistance to any user with a LiDAR-equipped smartphone. Blind people using the assistance devices in real-world environments can potentially discover new use cases and needs for a mobile-based system. For example, in the user evaluation of Corridor-Walker (section 4.2), the participants raised many situations in which the systems could be used (e.g., hospitals, shopping malls, metro transfers, and restaurants), and six participants suggested that Corridor-Walker could facilitate the construction of mental maps of unfamiliar environments. We expect that after distributing the system, various unintended use cases will be discovered and will extend the system's application range.

10.2 Future Directions: Toward more Inclusive Public Spaces

What challenges face the adoption of assistive systems in the real world? One challenge is rising the social acceptance level of assistive systems. As pointed out in Chapter 9, assistive technology for blind people raises visibility concerns, which should be addressed in future research. For assimilation into the environments of public buildings, the robot has the outward appearance of a standard suitcase. The design principle was successfully adopted. Facility managers commented that the use cases were indistinguishable from those of walking with a standard suitcase. Throughout our study, we received divergent opinions on this aspect.

All blind participants appreciated the robot's design because the robot assimilated into the surrounding environment. However, facility managers mentioned safety concerns with such assimilation because the visual impairment of the user might not be obvious to others. The managers considered that sighted visitors can proactively avoid a collision if they notice an approach by a blind person. In addition, our online survey found that acceptance of a technology was raised if the technology is used for assistive purposes.

Five blind participants preferred the seamless assimilation into public buildings even after understanding the visibility concerns of facility members. Meanwhile, seven blind participants suggested ways to clarify their visual impairment to others; for example, by attaching a distinctive symbol to the blind navigation. Three blind participants agreed that their blindness should be visible in certain situations, such as crowded environments. In such environments, they prefer to hold their white canes.

Faucett et al. [65] reported similar conflicts. They mentioned that externally imposed disability identities may stigmatize those with disabilities as incapable, stereotypically unkempt, and unfashionable. In contrast, they mentioned that a visible assistive technology such as a white cane affords credibility, allowing the user to communicate their needs without words.

The two pedestrian-avoidance systems presented in Chapters 5 and 6 represent a trade-off between the visibility of a blind user and the safety performance of the system. BBeep clears the path for blind users walking through crowded spaces. In this process, it notifies both the user and sighted pedestrians of a collision risk. In other words, the BBeep system enhances the visibility of blind users. The results showed that BBeep allows blind users to walk safely without collisions, even through a crowded international airport. Participants accepted and appropriately used BBeep in crowded public spaces, but hesitated to use the system in quiet spaces. Meanwhile, the guiding system presented in Chapter 6 aims for seamless walking of blind people in public spaces shared by nearby pedestrians. Blind people using this system avoided collisions with approaching pedestrians by adapting their walking speed. They also avoided standing pedestrians by changing their path and walking through free space, mimicking the behavior of sighted pedestrians. In other words, the guiding system reduces the visibility of blind users. Although this system supported blind users in the open space of an office building, its utility might be reduced in extremely crowded public spaces, such as stations and airports during rush hour. In such situations, a high-visibility system such as BBeep might be more useful than the guiding system.

The visibility concern presents a major barrier to broader deployment. Visibility can educate the general public on the necessity for such technologies, thus raising their social acceptance level. Visibility is mandated as a part of traffic safety in some countries. In addition, blind users tend to feel comfortable when assimilating into public buildings. We realized that real-world studies are essential for balancing the visibility–assimilation tradeoff and for broadening the usage of such technologies. Regulations can then be amended to reflect the latest technologies.

Chapter 11

Conclusion

This thesis aimed to improve the independence of blind people moving through public spaces such as shopping malls, airports, and museums, which have complex structures and various facilities. Moreover, these spaces are dynamic environments shared by other people. Blind people are deprived of independent travel through public spaces because they must rely on assistance from others. This thesis explored how computational technologies can increase the independence of blind people in public spaces and the challenges faced by blind people using assistive technologies to advance into the public domain.

To achieve its goals, this thesis presented and evaluated assistive technologies in the context of human–computer interaction. The main contributions are as follows:

First, we developed technologies that will assist blind visitors in exploring public spaces independently, similarly to sighted visitors who enjoy window shopping or museum visiting. We presented three assistive systems for exploration in public spaces: 1) a navigation robot that enables blind people to safely and independently explore a science museum and increase their autonomy in socially inclusive ways; 2) a wayfinding system by which blind people can recognize surrounding signage in public buildings; and 3) smartphone-based walking assistance by which blind people can avoid obstacles and recognize intersections in indoor corridors (Part I).

Second, we considered the social behaviors of blind users in the presence of surrounding people in public spaces. We proposed four systems that assist the social behaviors of blind people in public spaces: 1) an assistive suitcase system *BBeep*, which supports blind people walking through crowded environments; 2) a guiding system that helps blind people to walk seamlessly with nearby pedestrians in public spaces; 3) a smartphone-based assistive system named *LineChaser*, which navigates a blind user to the end of a line and continuously reports the distance and direction to the last person in the line, enabling users to join a queue and follow its movement; and 4) an assistive robot *BlindPilot*, which guides blind users to an object (such as an empty chair) using an intuitive handle (Part II).

Third, this thesis investigated and discussed the acceptance and concerns regarding autonomous navigation robots used by blind people in public buildings. For this purpose, we conducted an online survey of the general public, interviews with facility managers, and a focus group interview with blind participants. We identified safety, privacy, and visibility concerns related to blind visitors using an assistive robot in public spaces (Part III).

Based on these evaluations and investigations, we discussed the challenges and opportunities of providing blind people with assistive systems that improve their independence in public spaces. In conclusion, the next accessibility research will not be confined to controlled laboratory spaces but will extend to real-world environments in the quest for flexible and socially acceptable assistive systems.

Appendix A

Questionnaire and Results of the Online Survey in Chapter 9

A.1 Questionnaire

A.1.1 Question Types

There were four different types of questions used in our survey:

- **7-scale:** 7-scale Agreement/Disagreement
  Strongly agree/disagree, Moderately agree/disagree, Slightly agree/disagree, and Neutral

- **Y/N:** Yes and No

- **Open:** Open-ended

- **Check:** Check all that apply

A.1.2 Pre-Questions

Q9.1: [Open] *"Please indicate your gender."*

Q9.2: [Open] *"Please indicate your age."*

Q9.3: [Y/N] *"Have you had a personal, volunteer, or work experience interacting with visually impaired people?"*

Q9.4: [Y/N] *"Have you helped strangers with visual impairments in public spaces?"*

Q9.5: [Y/N] *"Have you seen robots moving about in public spaces?"*

Q9.6: [Y/N] *"Have you been involved in the development, promotion, marketing, or sale of robots?"*

A.1.3 After Watching Video A (Robot & User)

Q9.7A: [7-scale] *"I would feel uncomfortable if the robot moves about with a blind person in public buildings."*

Q9.8A: [7-scale] *"I would feel obstructed if the robot moves about with a blind person in public buildings."*

Q9.9A: [7-scale] *"I would feel unsafe if the robot moves about with a blind person in public buildings."*

Q9.10A: [Check] *"What kind of information about yourself would you not mind being detected by the camera attached to the robot assisting blind people?"*

- Body outline
- Distance
- Face
- Face orientation
- Facial expression
- Age
- Gender
- Height
- Hair color
- Objects being carried
- Ethnicity
- Action
- (If you know the user) Your name
- (If you do not know the user) Your name
- All information about me
- None of the above

Q9.11A: [7-scale] *"Feature 1 (Navigation) should be implemented in the robot assisting a blind person."*

Q9.12A: [7-scale] *"I am OK with the robot's camera capturing me if it is used for Feature 1 (Navigation) of the robot assisting a blind person."*

Q9.13A: [7-scale] *"The robot navigating a blind person to a destination looked natural."*

Q9.14A: [7-scale] *"Feature 1 (Navigation) could malfunction."*

Q9.15A: [7-scale] *"Feature 2 (Pedestrian Avoidance) should be implemented in the robot assisting a blind person."*

Q9.16A: [7-scale] *"I am OK with the robot's camera capturing me if it is used for Feature 2 (Pedestrian Avoidance) of the robot assisting a blind person."*

Q9.17A: [7-scale] *"The robot and a blind person avoiding pedestrians looked natural."*

Q9.18A: [7-scale] *"Feature 2 (Pedestrian Avoidance) could malfunction."*

Q9.19A: [7-scale] *"Feature 3 (Riding an Elevator) should be implemented in the robot assisting a blind person."*

Q9.20A: [7-scale] *"I am OK with the robot's camera capturing me if it is used for Feature 3 (Riding an Elevator) of the robot assisting a blind person."*

Q9.21A: [7-scale] *"The robot and blind person riding an elevator looked natural."*

Q9.22A: [7-scale] *"Feature 3 (Riding an Elevator) could malfunction."*

Q9.23A: [7-scale] *"Feature 4 (Standing in a Line) should be implemented in the robot assisting a blind person."*

Q9.24A: [7-scale] *"I am OK with the robot's camera capturing me if it is used for Feature 4 (Standing in a Line) of the robot assisting a blind person."*

Q9.25A: [7-scale] *"The robot and a blind person standing in a line looked natural."*

Q9.26A: [7-scale] *"Feature 4 (Standing in a Line) could malfunction."*

Q9.27A: [7-scale] *"I am OK with the camera attached to the robot assisting a blind person capturing me, if it is used for assisting blind people only and the captured data is used for one-time detection only and not saved."*

Q9.28A: [7-scale] *"I am OK with the camera attached to the robot assisting a blind person capturing me, if it is used for assisting blind people only and the captured data is saved."*

Q9.29A: [7-scale] *"I am OK with the camera attached to the robot assisting a blind person capturing me, if it is used for not only assisting blind people and the captured data is used for one-time detection only and not saved."*

Q9.30A: [7-scale] *"I am OK with the camera attached to the robot assisting a blind person capturing me, if it is used for not only assisting blind people and the captured data is saved."*

Q9.31: [Check] *How many people were standing in the line when video A was introducing Function 4 (Standing in a Line)?*

- One person
- Two people
- Four people

A.1.4 After Watching Video B (Robot Only)

Q9.7B: [7-scale] *"I would feel uncomfortable if the robot moves about in public buildings."*

Q9.8B: [7-scale] *"I would feel obstructed if the robot moves about in public buildings."*

Q9.9B: [7-scale] *"I would feel unsafe if the robot moves about in public buildings."*

Q9.10B: [Check] *"What kind of information about yourself would you not mind being detected by the robot's camera?"*

- Body outline
- Distance
- Face
- Face orientation
- Facial expression
- Age
- Gender
- Height
- Hair color
- Objects being carried
- Ethnicity
- Action
- (If you know the user) Your name
- (If you do not know the user) Your name
- All information about me
- None of the above

Q9.11B: [7-scale] *"Feature 1 (Navigation) should be implemented in the robot."*

Q9.12B: [7-scale] *"I am OK with the robot's camera capturing me if it is used for Feature 1 (Navigation) of the robot."*

Q9.13B: [7-scale] *"The robot moving to a destination looked natural."*

Q9.14B: [7-scale] *"Feature 1 (Navigation) could malfunction."*

Q9.15B: [7-scale] *"Feature 2 (Pedestrian Avoidance) should be implemented in the robot."*

Q9.16B: [7-scale] *"I am OK with the robot's camera capturing me if it is used for Feature 2 (Pedestrian Avoidance) of the robot."*

Q9.17B: [7-scale] *"The robot avoiding pedestrians looked natural."*

Q9.18B: [7-scale] *"Feature 2 (Pedestrian Avoidance) could malfunction."*

Q9.19B: [7-scale] *"Feature 3 (Riding an Elevator) should be implemented in the robot."*

Q9.20B: [7-scale] *"I am OK with the robot's camera capturing me, if it is used for Feature 3 (Riding an Elevator) of the robot."*

Q9.21B: [7-scale] *"The robot riding an elevator looked natural."*

Q9.22B: [7-scale] *"Feature 3 (Riding an Elevator) could malfunction."*

Q9.23B: [7-scale] *"Feature 4 (Standing in a Line) should be implemented in the robot."*

Q9.24B: [7-scale] *"I am OK with the robot's camera capturing me, if it is used for Feature 4 (Standing in a Line) of the robot."*

Q9.25B: [7-scale] *"The robot standing in a line looked natural."*

Q9.26B: [7-scale] *"Feature 4 (Standing in a Line) could malfunction."*

Q9.27B: [7-scale] *"I am OK with the robot's camera capturing me, if it is used for assisting blind people only and the captured data is used for one-time detection only and not saved."*

Q9.28B: [7-scale] *"I am OK with the robot's camera capturing me, if it is used for assisting blind people only and the captured data is saved."*

Q9.29B: [7-scale] *"I am OK with the robot's camera capturing me, if it is used for not only assisting blind people and the captured data is used for one-time detection only and not saved."*

Q9.30B: [7-scale] *"I am OK with the robot's camera capturing me, if it is used for not only assisting blind people and the captured data is saved."*

Q9.32: [Check] *"How many standing people did the robot avoid when video B was introducing Function 2 (Pedestrian Avoidance)?"*

- One person
- Two people
- Three people

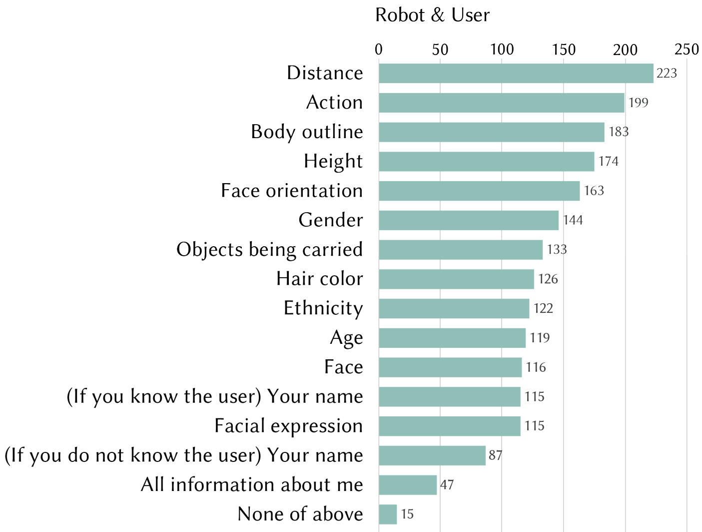
Q9.33: [Check] *"How many people were riding in the elevator when video B was introducing Function 3 (Riding an Elevator)?"*

- Zero
- One person
- Two people

A.2 Results

Figure A.1: Results of Q9.1–Q9.9. p: p-value of the Wilcoxon signed-rank test done on each question (** indicates the significance found at the levels of 0.01).

Q10: What kind of information about yourself would you not mind being detected by the camera attached to the robot [assisting blind people/moving about alone]?



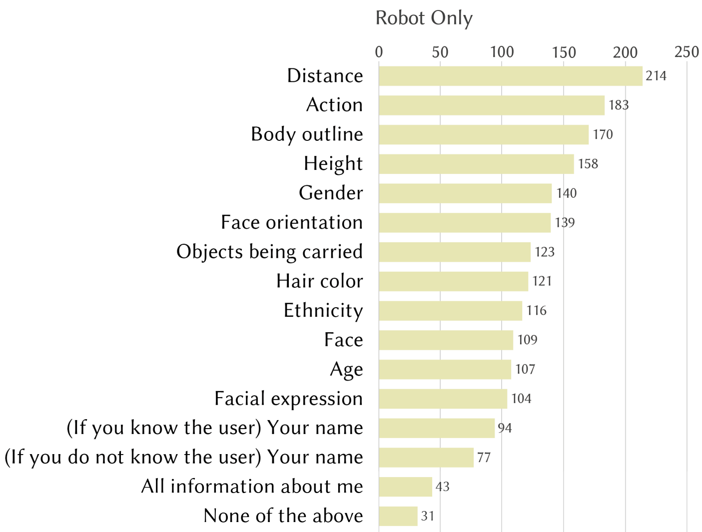


Figure A.2: Results of Q9.10.

Figure A.3: Results of Q9.11–Q9.25. p: p-value of the Wilcoxon signed-rank test done on each question p: p-value of the Wilcoxon signed-rank test done on each question (* and ** indicate the significance found at the levels of 0.05 and 0.01, respectively).

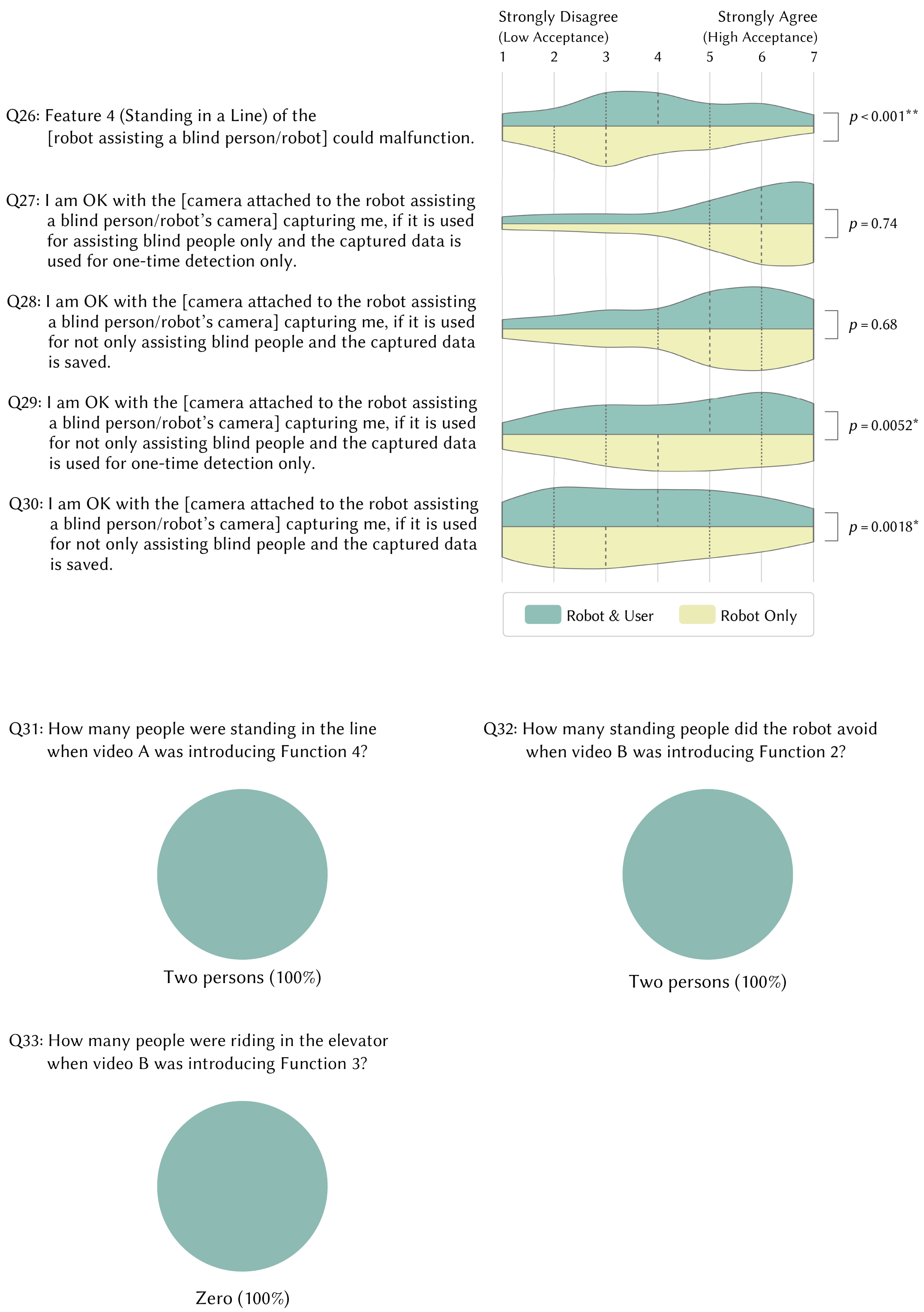


Figure A.4: Results of Q9.26–Q9.33. p: p-value of the Wilcoxon signed-rank test done on each question (* and ** indicate the significance found at the levels of 0.05 and 0.01, respectively).

Bibliography

[1] Dustin Adams, Lourdes Morales, and Sri Kurniawan. "A Qualitative Study to Support a Blind Photography Mobile Application". In: *Proceedings of the ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '13)*. New York, NY, USA: ACM, 2013, pp. 1–8. URL: https://doi.org/10.1145/2504335.2504360.

[2] Shikha Agarwal, Abir Mullick, and GG Ray. "Wayfinding Challenges and Strategies in Mumbai Suburban Railway". In: *Proceedings of the International Conference on Research into Design (ICoRD '17)*. New York, NY, USA: Springer, 2017, pp. 161–172. URL: https://doi.org/10.1007/978-981-10-3518-0_14.

[3] Tousif Ahmed, Apu Kapadia, Venkatesh Potluri, and Manohar Swaminathan. "Up to a Limit? Privacy Concerns of Bystanders and Their Willingness to Share Additional Information with Visually Impaired Users of Assistive Technologies". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.3 (Sept. 2018). URL: https://doi.org/10.1145/3264899.

[4] Dragan Ahmetovic, Federico Avanzini, Adriano Baratè, Cristian Bernareggi, Gabriele Galimberti, Luca A Ludovico, Sergio Mascetti, and Giorgio Presti. "Sonification of rotation instructions to support navigation of people with visual impairment". In: *Proceedings of the 2019 IEEE International Conference on Pervasive Computing and Communications (PerCom '19)*. Los Alamitos, CA, USA: IEEE, 2019, pp. 1–10. URL: https://doi.org/10.1109/PERCOM.2019.8767407.

[5] Dragan Ahmetovic, Cristian Bernareggi, Kristian Keller, and Sergio Mascetti. "MusA: Artwork Accessibility through Augmented Reality for People with Low Vision". In: *Proceedings of the 18th International Web for All Conference (W4A '21)*. New York, NY, USA: ACM, 2021. ISBN: 9781450382120. URL: https://doi.org/10.1145/3430263.3452441.

[6] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. "NavCog: a navigational cognitive assistant for the blind". In: *Proceedings of the ACM Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM. New York, NY, USA, 2016, pp. 90–99. URL: https://doi.org/10.1145/2935334.2935361.

[7] Dragan Ahmetovic, Nahyun Kwon, Uran Oh, Cristian Bernareggi, and Sergio Mascetti. "Touch Screen Exploration of Visual Artwork for Blind People". In: *Proceedings of the Web Conference 2021 (WWW '21)*. New York, NY, USA: ACM, 2021, pp. 2781–2791. URL: https://doi.org/10.1145/3442381.3449871.

[8] Peter Aigner and Brenan McCarragher. "Shared control framework applied to a robotic aid for the blind". In: *IEEE Control Systems Magazine* 19.2 (1999), pp. 40–46. URL: https://doi.org/10.1109/37.753934.

[9] Aipoly. *Aipoly Vision. Sight for Blind and Visually Impaired.* Retrieved in December 27, 2020 from http://aipoly.com. 2020.

[10] Aira. *Aira.* Retrieved in December 27, 2020 from https://aira.io/. 2020.

[11] Taslima Akter, Tousif Ahmed, Apu Kapadia, and Swami Manohar Swaminathan. "Privacy Considerations of the Visually Impaired with Camera Based Assistive Technologies: Misrepresentation, Impropriety, and Fairness". In: *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*. New York, NY, USA: ACM, 2020, pp. 1–14. URL: https://doi.org/10.1145/3373625.3417003.

[12] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan C. Semaan. ""I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications". In: Berkeley, CA, USA: Publisher USENIX Association, 2020.

[13] Abdulrhman Alkhanifer and S. Ludi. "Disorientation Factors that Affect the Situation Awareness of the Visually Impaired Individuals in Unfamiliar Indoor Environments". In: *HCI*. New York, NY, USA: Springer, 2015, pp. 89–100. URL: https://doi.org/10.1007/978-3-319-20687-5_9.

[14] Dario Allegra, Francesco Alessandro, Corrado Santoro, and Filippo Stanco. "Experiences in Using the Pepper Robotic Platform for Museum Assistance Applications". In: *Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP '18)*. New York, NY, USA: IEEE, 2018, pp. 1033–1037. URL: https://doi.org/10.1109/ICIP.2018.8451777.

[15] Tomohiro Amemiya, Jun Yamashita, Koichi Hirota, and Michitaka Hirose. "Virtual leading blocks for the deaf-blind: A real-time way-finder by verbal-nonverbal hybrid interface and high-density RFID tag space". In: *Proceedings of the IEEE Conference on Virtual Reality (VR '04)*. Los Alamitos, CA, USA: IEEE, 2004, pp. 165–287. URL: https://doi.org/10.1109/VR.2004.1310070.

[16] Giorgos Anagnostakis, Michalis Antoniou, Elena Kardamitsi, Thodoris Sachinidis, Panayiotis Koutsabasis, Modestos Stavrakis, Spyros Vosinakis, and Dimitris Zissis. "Accessible Museum Collections for the Visually Impaired: Combining Tactile Exploration, Audio Descriptions and Mobile Gestures". In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '16)*. New York, NY, USA: ACM, 2016, pp. 1021–1025. URL: https://doi.org/10.1145/2957265.2963118.

[17] Apple. *iPhone 12 Pro - Technical Specifications*. Retrieved in January 17, 2022 from https://support.apple.com/kb/SP831?locale=en_US. 2021.

[18] Vassilios S Argyropoulos and Charikleia Kanari. "Re-imagining the museum through "touch": Reflections of individuals with visual disability on their experience of museum-visiting in Greece". In: *Alter* 9.2 (2015), pp. 130–143. URL: https://doi.org/10.1016/j.alter.2014.12.005.

[19] The Metropolitan Museum of Art. *For Visitors Who Are Blind or Partially Sighted*. 2018. URL: https://www.metmuseum.org/events/programs/access/visitors-who-are-blind-or-partially-sighted.

[20] Saki Asakawa, João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. "The Present and Future of Museum Accessibility for People with Visual Impairments". In: *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18)*. New York, NY, USA: ACM, 2018, 382–384. ISBN: 9781450356503. URL: https://doi.org/10.1145/3234695.3240997.

[21] Saki Asakawa, João Guerreiro, Daisuke Sato, Hironobu Takagi, Dragan Ahmetovic, Desi Gonzalez, Kris M. Kitani, and Chieko Asakawa. "An Independent and Interactive Museum Experience for Blind People". In: *Proceedings of the 16th International Web for All Conference (W4A '19)*. New York, NY, USA: ACM, 2019. ISBN: 9781450367165. URL: https://doi.org/10.1145/3315002.3317557.

[22] Mauro Avila Soto and Markus Funk. "Look, a Guidance Drone! Assessing the Social Acceptability of Companion Drones for Blind Travelers in Public Spaces". In: *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '18)*. New York, NY, USA: ACM, 2018, 417–419. ISBN: 9781450356503. URL: https://doi.org/10.1145/3234695.3241019.

[23] Mauro Avila Soto, Markus Funk, Matthias Hoppe, Robin Boldt, Katrin Wolf, and Niels Henze. "DroneNavigator: Using Leashed and Free-Floating Quadcopters to Navigate Visually Impaired Travelers". In: *Proceedings of the ACM SIGACCESS Conference on Computers and accessibility (ASSETS '17)*. New York, NY, USA: ACM, 2017, pp. 300–304. URL: https://doi.org/10.1145/3132525.3132556.

[24] Shiri Azenkot, Catherine Feng, and Maya Cakmak. "Enabling building service robots to guide blind people: A participatory design approach". In: *Proceedings of the ACM/IEEE International Conference on Human Robot Interaction (HRI '16)*. Los Alamitos, CA, USA: IEEE Press, 2016, pp. 3–10. URL: https://doi.org/10.1109/HRI.2016.7451727.

[25] Shiri Azenkot, Richard E Ladner, and Jacob O Wobbrock. "Smartphone haptic feedback for nonvisual wayfinding". In: *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '11)*. New York, NY, USA: ACM, 2011, pp. 281–282. URL: https://doi.org/10.1145/2049536.2049607.

[26] Carryl L Baldwin, Jesse L Eisert, Andre Garcia, Bridget Lewis, Stephanie M Pratt, and Christian Gonzalez. "Multimodal urgency coding: auditory, visual, and tactile parameters and their impact on perceived urgency". In: *Work* 41.Supplement 1 (2012), pp. 3586–3591. URL: https://doi.org/10.3233/WOR-2012-0669-3586.

[27] Aaron Bangor, Philip Kortum, and James Miller. "Determining what individual SUS scores mean: Adding an adjective rating scale". In: *Journal of usability studies* 4.3 (2009), pp. 114–123. URL: https://doi.org/10.5555/2835587.2835589.

[28] Willy Barnett, Adrienne Foos, Thorsten Gruber, Debbie Keeling, Kathleen Keeling, and Linda Nasr. "Consumer perceptions of Interactive Service Robots: A Value-Dominant Logic perspective". In: *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN '14)*. Piscataway, NJ, USA: IEEE Press, 2014, pp. 1134–1139. URL: https://doi.org/10.1109/ROMAN.2014.6926404.

[29] BeMyEyes. *BeMyEyes*. Retrieved in December 27, 2020 from https://www.bemyeyes.com/. 2020.

[30] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and Tom Yeh. "VizWiz: Nearly Real-Time Answers to Visual Questions". In: *Proceedings of the ACM Annual Symposium on User Interface Software and Technology (UIST '10)*. New York, NY, USA: ACM, 2010, 333–342. URL: https://doi.org/10.1145/1866029.1866080.

[31] BB Blasch, SJ LaGrow, and WR De l'Aune. "Three aspects of coverage provided by the long cane: Object, surface, and foot-placement preview". In: *Journal of Visual Impairment and Blindness* 90 (1996), pp. 295–301. URL: https://doi.org/10.1177/0145482X9609000404.

[32] Simon Blessenohl, Cecily Morrison, Antonio Criminisi, and Jamie Shotton. "Improving indoor mobility of the visually impaired with depth-based spatial sound". In: *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW '15)*. Los Alamitos, CA, USA: IEEE, 2015, pp. 26–34. URL: https://doi.org/10.1109/ICCVW.2015.62.

[33] Jeffrey R Blum, Mathieu Bouchard, and Jeremy R Cooperstock. "What's around me? Spatialized audio augmented reality for blind users with a smartphone". In: *Proceedings of the International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services (MobiQuitous '11)*. New York, NY, USA: Springer, 2011, pp. 49–62. URL: https://doi.org/10.1007/978-3-642-30973-1_5.

[34] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. *YOLOv4: Optimal Speed and Accuracy of Object Detection*. 2020. arXiv: 2004.10934 [cs.CV].

[35] GP Boucek Jr, James E Veitengruber, and Wayne D Smith. *Aircraft alerting systems criteria study. Volume II. Human factors guidelines for aircraft alerting systems*. Tech. rep. BOEING COMMERCIAL AIRPLANE CO SEATTLE WA, 1977.

[36] Nicholas A Bradley and Mark D Dunlop. "Investigating context-aware clues to assist navigation for visually impaired people". In: *Proceedings of Workshop on Building Bridges: Interdisciplinary Context-Sensitive Computing, University of Glasgow*. GBR: Strathprints, 2002, pp. 5–10.

[37] Elizabeth Broadbent, Rebecca Stafford, and Bruce MacDonald. "Acceptance of Healthcare Robots for the Older Population: Review and Future Directions". In: *International journal of social robotics* 1.4 (2009), p. 319. URL: https://doi.org/10.1007/s12369-009-0030-6.

[38] Michael Brock and Per Ola Kristensson. "Supporting blind navigation using depth sensing and sonification". In: *Proceedings of the ACM International Joint Conference on Pervasive and ubiquitous computing adjunct publication (UbiComp '13)*. ACM. New York, NY, USA, 2013, pp. 255–258. URL: https://doi.org/10.1145/2494091.2494173.

[39] John Brooke et al. "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7. URL: https://doi.org/.

[40] Wolfram Burgard, Armin B Cremers, Dieter Fox, Dirk Hähnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. "The Interactive Museum Tour-Guide Robot". In: *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence*. AAAI '98/IAAI '98. Palo Alto, CA, USA: AAAI, 1998, pp. 11–18. URL: https://www.aaai.org/Library/AAAI/1998/aaai98-002.php.

[41] Serap Buyurgan. "The Expectations of the Visually Impaired University Students from Museums." In: *Educational Sciences: Theory and Practice* 9.3 (2009), pp. 1191–1204.

[42] Envision Technologies B.V. *Envision AI. Enabling vision for the blind*. Retrieved in December 27, 2020 from https://www.letsenvision.com. 2020.

[43] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. "Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields". In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE. Los Alamitos, CA, USA, 2017, pp. 1302–1310. URL: https://doi.org/10.1109/CVPR.2017.143.

[44] Hsuan-Eng Chen, Yi-Ying Lin, Chien-Hsing Chen, and I-Fang Wang. "BlindNavi: A navigation app for the visually impaired smartphone user". In: *Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems*. New York, NY, USA: ACM, 2015, pp. 19–24. URL: https://doi.org/10.1145/2702613.2726953.

[45] Ping-Ling Chen, Wafaa Saleh, and Chih-Wei Pai. "Texting and walking: a controlled field study of crossing behaviours and inattentional blindness in Taiwan". In: *Behaviour & Information Technology* 36.4 (2017), pp. 435–445. URL: https://doi.org/10.1080/0144929X.2016.1240234.

[46] Tzu-Kuan Chuang, Ni-Ching Lin, Jih-Shi Chen, Chen-Hao Hung, Yi-Wei Huang, Chunchih Tengl, Haikun Huang, Lap-Fai Yu, Laura Giarré, and Hsueh-Cheng Wang. "Deep Trail-Following Robotic Guide Dog in Pedestrian Environments for People who are Blind and Visually Impaired - Learning from Virtual and Real Worlds". In: *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA '18)*. Piscataway, NJ, USA: IEEE Press, 2018, pp. 1–7. URL: https://doi.org/10.1109/ICRA.2018.8460994.

[47] Sakmongkon Chumkamon, Peranitti Tuvaphanthaphiphat, and Phongsak Keeratiwintakorn. "A blind navigation system using RFID for indoor environments". In: *2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON '08)*. Vol. 2. Los Alamitos, CA, USA: IEEE, 2008, pp. 765–768. URL: https://doi.org/10.1109/ECTICON.2008.4600543..

[48] Giovanni Ciaffoni. *Ariadne GPS - Mobility and ,ap exploration for all*. Retrieved in December 27, 2020 from https://www.ariadnegps.eu/. 2020.

[49] Cloudsight. *TapTapSee. Mobile camera application designed specifically for the blind and visually impaired iOS users*. Retrieved in December 27, 2020 from http://www.taptapseeapp.com. 2020.

[50] Angela Constantinescu, Karin Müller, Monica Haurilet, Vanessa Petrausch, and Rainer Stiefelhagen. "Bring the Environment to Life: A Sonification Module for People with Visual Impairments to Improve Situation Awareness". In: *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20)*. New York, NY, USA: ACM, 2020, pp. 50–59. URL: https://doi.org/10.1145/3382507.3418874.

[51] Akansel Cosgun, E Akin Sisbot, and Henrik I Christensen. "Guidance for human navigation using a vibro-tactile belt interface and robot-like motion planning". In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '14)*. Los Alamitos, CA, USA: IEEE, 2014, pp. 6350–6355. URL: https://doi.org/10.1109/ICRA.2014.6907796.

[52] Michael Decker. "Service Robots in the Mirror of Reflective Research". In: *Poiesis & Praxis* 9.3 (2012), pp. 181–200. URL: https://doi.org/10.1007/s10202-012-0111-8.

[53] Katherine Deibel. "A Convenient Heuristic Model for Understanding Assistive Technology Adoption". In: *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. New York, NY, USA: ACM, 2013. ISBN: 9781450324052. URL: https://doi.org/10.1145/2513383.2513427.

[54] Louise Demers, Rhoda Weiss-Lambrou, and Bernadette Ska. "The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress". In: *Technology and Disability* 14.3 (2002), pp. 101–105. URL: https://doi.org/10.13072/midss.298.

[55] Andrés A Díaz-Toro, Sixto E Campaña-Bastidas, and Eduardo F Caicedo-Bravo. "Vision-Based System for Assisting Blind People to Wander Unknown Environments in a Safe Way". In: *Journal of Sensors* 2021 (2021). URL: https://doi.org/10.1155/2021/6685686.

[56] Mengdi Ding. *Arts and Technology: How one Museum's App is Connecting Visitors and Driving Institutional Change*. Americans for the Arts. 2016. URL: https://namp.americansforthearts.org/2019/05/15/arts-and-technology-how-one-museumâĂŹs-app-is-connecting-visitors-and-driving-institutional-change.

[57] Karen Duarte, José Cecílio, and Pedro Furtado. "Easily guiding of blind: Providing information and navigation-smartnav". In: *Wireless Internet*. New York, NY, USA: Springer, 2014, pp. 129–134. URL: https://doi.org/10.1007/978-3-319-18802-7_18.

[58] Brian Due and Simon Bierring Lange. "The Moses effect: The spatial hierarchy and joint accomplishment of a blind person navigating". In: *Space and culture* 21.2 (2018), pp. 129–144. URL: https://doi.org/10.1177/1206331217734541.

[59] Judy Edworthy, Sarah Loxley, and Ian Dennis. "Improving auditory warning design: Relationship between warning sound parameters and perceived urgency". In: *Human factors* 33.2 (1991), pp. 205–231. URL: https://doi.org/10.1177/001872089103300206.

[60] Christin Engel, Karin Müller, Angela Constantinescu, Claudia Loitsch, Vanessa Petrausch, Gerhard Weber, and Rainer Stiefelhagen. "Travelling More Independently: A Requirements Analysis for Accessible Journeys to Unknown Buildings for People with Visual Impairments". In: *Proceedings of the 22th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*. New York, NY, USA: ACM, 2020, pp. 1–11. URL: https://10.1145/3373625.3417022.

[61] Agebson Rocha Façanha, Ticianne Darin, Windson Viana, and Jaime Sánchez. "O&M Indoor Virtual Environments for People Who Are Blind: A Systematic Literature Review". In: *ACM Transactions on Accessible Computing (TACCESS)* 13.2 (2020), pp. 1–42. URL: https://doi.org/10.1145/3395769.

[62] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. "The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. New York, NY, USA: ACM, 2012, pp. 425–432. URL: https://doi.org/10.1145/2207676.2207735.

[63] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. "Indoor human navigation systems: A survey". In: *Interacting with Computers* 25.1 (2013), pp. 21–33. URL: https://doi.org/10.1093/iwc/iws010.

[64] José Faria, Sérgio Lopes, Hugo Fernandes, Paulo Martins, and João Barroso. "Electronic white cane for blind people navigation assistance". In: *Proceedings of the IEEE Conference on World Automation Congress (WAC '10)*. Los Alamitos, CA, USA: IEEE, 2010, pp. 1–7.

[65] Heather A. Faucett, Kate E. Ringland, Amanda L. L. Cullen, and Gillian R. Hayes. "(In)Visibility in Disability and Assistive Technology". In: *ACM Trans. Access. Comput.* 10.4 (Oct. 2017). ISSN: 1936-7228. URL: https://doi.org/10.1145/3132040.

[66] Rodney Fewings. "Wayfinding and airport terminal design". In: *The journal of navigation* 54.2 (2001), pp. 177–184. URL: https://doi.org/10.1017/S0373463301001369.

[67] Alexander Fiannaca, Ilias Apostolopoulous, and Eelke Folmer. "Headlock: a wearable navigation aid that helps blind cane users traverse large open spaces". In: *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility (ASSETS '14)*. New York, NY, USA: ACM, 2014, pp. 19–26. URL: https://doi.org/10.1145/2661334.2661453.

[68] Vítor Filipe, Filipe Fernandes, Hugo Fernandes, António Sousa, Hugo Paredes, and João Barroso. "Blind navigation support system based on Microsoft Kinect". In: *Procedia Computer Science* 14 (2012), pp. 94–101. URL: https://doi.org/10.1016/j.procs.2012.10.011.

[69] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. "The dynamic window approach to collision avoidance". In: *IEEE Robotics & Automation Magazine* 4.1 (1997), pp. 23–33. URL: https://doi.org/10.1109/100.580977.

[70] Masahiro Furukawa, Hiromi Yoshikawa, Taku Hachisu, Shogo Fukushima, and Hiroyuki Kajimoto. ""Vection field" for pedestrian traffic control". In: *Proceedings of the ACM International Conference on Augmented Human (AH '11)*. ACM. New York, NY, USA, 2011, p. 19. URL: https://doi.org/10.1145/1959826.1959845.

[71] Georgios Galatas, Christopher McMurrough, Gian Luca Mariottini, and Fillia Makedon. "eyeDog: an assistive-guide robot for the visually impaired". In: *Proceedings of the ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '11)*. New York, NY, USA: ACM, 2011, p. 58. URL: https://doi.org/10.1145/2141622.2141691.

[72] Aura Ganz, Siddhesh Rajan Gandhi, Carole Wilson, and Gary Mullett. "INSIGHT: RFID and Bluetooth enabled automated space for the blind and visually impaired". In: *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. Los Alamitos, CA, USA: IEEE, 2010, pp. 331–334. URL: https://doi.org/10.1109/IEMBS.2010.5627670.

[73] Giuseppe Ghiani, Barbara Leporini, and Fabio Paternò. "Supporting Orientation for Blind People Using Museum Guides". In: *CHI'08 extended abstracts on Human factors in computing systems (CHI EA '08)*. New York, NY, USA: ACM, 2008, pp. 3417–3422. URL: https://doi.org/10.1145/1358628.1358867.

[74] Giuseppe Ghiani, Barbara Leporini, and Fabio Paternò. "Vibrotactile feedback to aid blind users of mobile guides". In: *Journal of Visual Languages & Computing* 20.5 (2009), pp. 305–317. URL: https://doi.org/10.1016/j.jvlc.2009.07.004.

[75] Nicholas A. Giudice. *COVID-19 and blindness: Why the new touchless, physically-distant world sucks for people with visual impairment.* Retrieved in December 27, 2020 from https://medium.com/@nicholas.giudice/covid-19-and-blindness-why-the-new-touchless-physically-distant-world-sucks-for-people-with-2c8dbd21de63. 2020.

[76] Nicholas A Giudice, Benjamin A Guenther, Nicholas A Jensen, and Kaitlyn N Haase. "Cognitive mapping without vision: Comparing wayfinding performance after learning from digital touchscreen-based multimodal maps vs. embossed tactile overlays". In: *Frontiers in Human Neuroscience* 14 (2020), p. 87. URL: https://doi.org/10.3389/fnhum.2020.00087.

[77] Nicholas A Giudice, William E Whalen, Timothy H Riehle, Shane M Anderson, and Stacy A Doore. "Evaluation of an accessible, real-time, and infrastructure-free indoor navigation system by users who are blind in the mall of america". In: *Journal of Visual Impairment & Blindness* 113.2 (2019), pp. 140–155. URL: https://doi.org/10.1177/0145482X19840918.

[78] Christian Gonzalez, Bridget A Lewis, Daniel M Roberts, Stephanie M Pratt, and Carryl L Baldwin. "Perceived urgency and annoyance of auditory alerts in a driving context". In: 56.1 (2012), pp. 1684–1687. URL: https://doi.org/10.1177/1071181312561337.

[79] Desi Gonzalez. *A path with choice: What we learned from designing an inclusive audio guide*. Museums and the Web (MW '17). 2017. URL: https://mw17.mwconf.org/paper/a-path-with-choice-what-we-learned-from-designing-an-inclusive-audio-guide/.

[80] Google. *Google Maps*. Retrieved in December 27, 2020 from https://maps.google.com. 2020.

[81] Thomas L Greenbaum. *The Fandbook for Focus Group Research*. Los Angeles, CA, USA: Sage, 1998. URL: https://doi.org/10.4135/9781412986151.

[82] Giorgio Grisetti, Cyrill Stachniss, Wolfram Burgard, et al. "Improved techniques for grid mapping with rao-blackwellized particle filters". In: *IEEE transactions on Robotics* 23.1 (2007), p. 34. URL: https://doi.org/10.1109/TRO.2006.889486.

[83] William Grussenmeyer, Jesel Garcia, and Fang Jiang. "Feasibility of using haptic directions through maps with a tablet and smart watch for people who are blind and visually impaired". In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. New York, NY, USA: ACM, 2016, pp. 83–89. URL: https://doi.org/10.1145/2935334.2935367.

[84] João Guerreiro, Eshed Ohn-Bar, Dragan Ahmetovic, Kris Kitani, and Chieko Asakawa. "How Context and User Behavior Affect Indoor Navigation Assistance for Blind People". In: *Proceedings of the 15th International Web for All Conference (W4A '18)*. New York, NY, USA: ACM, 2018. URL: https://doi.org/10.1145/3192714.3192829.

[85] João Guerreiro, Daisuke Sato, Saki Asakawa, Huixu Dong, Kris M Kitani, and Chieko Asakawa. "CaBot: Designing and Evaluating an Autonomous Navigation Robot for Blind People". In: *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. New York, NY, USA: ACM, 2019, pp. 68–82. URL: https://doi.org/10.1145/3308561.3353771.

[86] International Health Facility Guidelines. "Wayfinding Guidelines International Health Facility Guidelines". In: (2016). URL: http://www.healthfacilityguidelines.com/GuidelineIndex/Index/Wayfinding-Guidelines.

[87] Ellen C Haas and Judy Edworthy. "Designing urgency into auditory warnings using pitch, speed and loudness". In: *Computing & Control Engineering Journal* 7.4 (1996), pp. 193–198. URL: https://doi.org/10.1049/cce:19960407.

[88] Kozue Handa, Hitoshi Dairoku, and Yoshiko Toriyama. "Investigation of priority needs in terms of museum service accessibility for visually impaired visitors". In: *British journal of visual impairment* 28.3 (2010), pp. 221–234. URL: https://doi.org/10.1177/0264619610374680.

[89] Peter E Hart, Nils J Nilsson, and Bertram Raphael. "A formal basis for the heuristic determination of minimum cost paths". In: *IEEE transactions on Systems Science and Cybernetics* 4.2 (1968), pp. 100–107. URL: https://doi.org/10.1109/TSSC.1968.300136.

[90] Denise Hebesberger, Tobias Koertner, Christoph Gisinger, and Jürgen Pripfl. "A long-term autonomous robot at a care hospital: A mixed methods study on social acceptance and experiences of staff and older adults". In: *International Journal of Social Robotics* 9.3 (2017), pp. 417–429. URL: https://doi.org/10.1007/s12369-016-0391-6.

[91] Marion A. Hersh and Michael A. Johnson. "A Robotic Guide for Blind People. Part 1. A Multi-national Survey of the Attitudes, Requirements and Preferences of Potential End-users". In: *Appl. Bionics Biomechanics* 7.4 (Oct. 2010), pp. 277–288. ISSN: 1176-2322. URL: http://doi.org/10.1080/11762322.2010.523626.

[92] Marion A. Hersh and Michael A. Johnson. "A Robotic Guide for Blind People Part 2: Gender and National Analysis of a Multi-national Survey and the Application of the Survey Results and the CAT Model to Framing Robot Design Specifications". In: *Appl. Bionics Biomechanics* 9.1 (Jan. 2012), pp. 29–43. ISSN: 1176-2322. URL: http://doi.org/10.3233/ABB-2011-0034.

[93] Wolfgang Hess, Damon Kohler, Holger Rapp, and Daniel Andor. "Real-time loop closure in 2D LIDAR SLAM". In: *Proceedings of the IEEE Conference on Robotics and Automation (ICRA '16)*. Los Alamitos, CA, USA: IEEE, 2016, pp. 1271–1278. URL: https://doi.org/10.1109/ICRA.2016.7487258.

[94] Phoebe Hillemann. *Art Museum Tours For Visitors Who Are Blind*. 2016. URL: https://americanart.si.edu/blog/eye-level/2016/09/317/art-museum-tours-visitors-who-are-blind.

[95] Juan David Hincapié-Ramos and Pourang Irani. "CrashAlert: enhancing peripheral alertness for eyes-busy mobile interaction while walking". In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM. New York, NY, USA, 2013, pp. 3385–3388. URL: https://doi.org/10.1145/2470654.2466463.

[96] Chien-Ju Ho, Aleksandrs Slivkins, Siddharth Suri, and Jennifer Wortman Vaughan. "Incentivizing High Quality Crowdwork". In: *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. Geneva, Switzerland: International World Wide Web Conference Committee, 2015, 419–429. URL: https://doi.org/10.1145/2736277.2741102.

[97] Leona Holloway, Kim Marriott, and Matthew Butler. "Accessible maps for the blind: Comparing 3D printed models with tactile graphics". In: *Proceedings of the 2018 CHI conference on human factors in computing systems (CHI '18)*. New York, NY, USA: ACM, 2018, pp. 1–13. URL: https://doi.org/10.1145/3173574.3173772.

[98] Hsieh-Chang Huang, Ching-Tang Hsieh, and Cheng-Hsiang Yeh. "An indoor obstacle detection system using depth information and region growth". In: *Sensors* 15.10 (2015), pp. 27116–27141. URL: https://doi.org/10.3390/s151027116.

[99] Andreas Hub, Joachim Diepstraten, and Thomas Ertl. "Design and Development of an Indoor Navigation and Object Identification System for the Blind". In: *SIGACCESS Access. Comput.* 77–78 (Sept. 2003), 147–152. ISSN: 1558-2337. URL: https://doi.org/10.1145/1029014.1028657.

[100] Andreas Hub, Joachim Diepstraten, and Thomas Ertl. "Design and development of an indoor navigation and object identification system for the blind". In: *Proceedings of the ACM SIGACCESS Conference on Computers and accessibility (ASSETS '04)*. 77-78. New York, NY, USA: ACM, 2004, pp. 147–152. URL: https://doi.org/10.1145/1029014.1028657.

[101] Andreas Hub, Tim Hartter, and Thomas Ertl. "Interactive Tracking of Movable Objects for the Blind on the Basis of Environment Models and Perception-Oriented Object Recognition Methods". In: *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06)*. New York, NY, USA: ACM, 2006, 111–118. URL: https://doi.org/10.1145/1168987.1169007.

[102] Kiyohide Ito, Makoto Okamoto, Junichi Akita, Tetsuo Ono, Ikuko Gyobu, Tomohito Takagi, Takahiro Hoshi, and Yu Mishima. "CyARM: an alternative aid device for blind persons". In: *Proceedings of the ACM CHI Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM. New York, NY, USA, 2005, pp. 1483–1488. URL: https://doi.org/10.1145/1056808.1056947.

[103] Theo Jacobs and Gurvinder Singh Virk. "ISO 13482-The new safety standard for personal care robots". In: *Proceedings of the ISR/Robotik 2014; 41st International Symposium on Robotics*. Berlin, Germany: VDE, 2014, pp. 1–6.

[104] Rabia Jafri, Rodrigo Louzada Campos, Syed Abid Ali, and Hamid R Arabnia. "Visual and infrared sensor data-based obstacle detection for the visually impaired using the Google project tango tablet development kit and the unity engine". In: *IEEE Access* 6 (2017), pp. 443–454. URL: https://doi.org/10.1109/ACCESS.2017.2766579.

[105] Dhruv Jain. "Pilot Evaluation of a Path-Guided Indoor Navigation System for Visually Impaired in a Public Museum". In: *Proceedings of the 16th international ACM SIGACCESS conference on Computers & Accessibility (ASSETS '14)*. New York, NY, USA: ACM, 2014, pp. 273–274. URL: https://doi.org/10.1145/2661334.2661405.

[106] Anna Jankowska, Agnieszka Szarkowska, Krzysztof Krejtz, Anita Fidyka, Jaroslaw Kowalski, and Marcin Wichrowski. "Smartphone app as a museum guide. Testing the Open Art application with blind, deaf, and sighted users". In: *Rivista internazionale di tecnica della traduzione* 19 (2017), pp. 113–130. URL: https://doi.org/10.13137/2421-6763-17354.

[107] Chandrika Jayant, Hanjie Ji, Samuel White, and Jeffrey P. Bigham. "Supporting Blind Photography". In: *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '11)*. New York, NY, USA: ACM, 2011, 203–210. URL: https://doi.org/10.1145/2049536.2049573.

[108] Watthanasak Jeamwatthanachai, M. Wald, and G. Wills. "Indoor navigation by blind people: Behaviors and challenges in unfamiliar spaces and buildings". In: *The British Journal of Visual Impairment* 37 (2019), pp. 140 –153. URL: https://doi.org/10.1177/0264619619833723.

[109] U.S. Department of Justice. *Expanding Your Market: Maintaining Accesibility in Museums*. 2009. URL: https://www.ada.gov/business/museum_access.htm.

[110] Hernisa Kacorri, Eshed Ohn-Bar, Kris M Kitani, and Chieko Asakawa. "Environmental Factors in Indoor Navigation Based on Real-World Trajectories of Blind Users". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM. New York, NY, USA, 2018, p. 56. URL: https://doi.org/10.1145/3173574.3173630.

[111] Rudolph Emil Kalman. "A new approach to linear filtering and prediction problems". In: *Journal of basic Engineering* 82.1 (1960), pp. 35–45. URL: https://doi.org/10.1115/1.3662552.

[112] Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. "Freedom to Roam: A Study of Mobile Device Adoption and Accessibility for People with Visual and Motor Disabilities". In: *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '09)*. New York, NY, USA: ACM, 2009, 115–122. ISBN: 9781605585581. URL: https://doi.org/10.1145/1639642.1639663.

[113] Sebastian Kapser and Mahmoud Abdelrahman. "Acceptance of autonomous delivery vehicles for last-mile delivery in Germany – Extending UTAUT2 with risk perceptions". In: *Transportation Research Part C: Emerging Technologies* 111 (2020), pp. 210–225. ISSN: 0968-090X. URL: https://doi.org/10.1016/j.trc.2019.12.016.

[114] Robert K Katzschmann, Brandon Araki, and Daniela Rus. "Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26.3 (2018), pp. 583–593. URL: https://doi.org/10.1109/TNSRE.2018.2800665.

[115] Seita Kayukawa, Keita Higuchi, João Guerreiro, Shigeo Morishima, Yoichi Sato, Kris Kitani, and Chieko Asakawa. "BBeep: A sonic collision avoidance system for blind travellers and nearby pedestrians". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. New York, NY, USA: ACM, 2019, pp. 1–12. URL: https://doi.org/10.1145/3290605.3300282.

[116] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. "BlindPilot: A Robotic Local Navigation System That Leads Blind People to a Landmark Object". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. New York, NY, USA: ACM, 2020, pp. 1–9. URL: https://doi.org/10.1145/3334480.3382925.

[117] Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. "Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4.3 (2020), pp. 1–22. URL: https://doi.org/10.1145/3411825.

[118] Seita Kayukawa, Daisuke Sato, Masayuki Murata, Tatsuya Ishihara, Akihiro Kosugi, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. "How Users, Facility Managers, and Bystanders Perceive and Accept a Navigation Robot for Visually Impaired People in Public Buildings". In: *Proceedings of the 31st IEEE International Conference on Robot & Human Interactive Communication*. IEEE RO-MAN 2022. Piscataway, NJ, USA: IEEE Press, 2022.

[119] Seita Kayukawa, Hironobu Takagi, João Guerreiro, Shigeo Morishima, and Chieko Asakawa. "Smartphone-Based Assistance for Blind People to Stand in Lines". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. New York, NY, USA: ACM, 2020, pp. 1–8. URL: https://doi.org/10.1145/3334480.3382954.

[120] Sulaiman Khan, Shah Nazir, and Habib Ullah Khan. "Analysis of Navigation Assistants for Blind and Visually Impaired People: A Systematic Review". In: *IEEE Access* 9 (2021), pp. 26712–26734. URL: https://doi.org/10.1109/ACCESS.2021.3052415.

[121] Beomjoon Kim and Joelle Pineau. "Socially Adaptive Path Planning in Human Environments Using Inverse Reinforcement Learning". In: *International Journal of Social Robotics* 8.1 (2016), pp. 51–66. ISSN: 1875-4805. URL: https://doi.org/10.1007/s12369-015-0310-2.

[122] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. "Navigating Visually Impaired Travelers in a Large Train Station Using Smartphone and Bluetooth Low Energy". In: *Proceedings of the 31st Annual ACM Symposium on Applied Computing (SAC '16)*. New York, NY, USA: ACM, 2016, 604–611. URL: https://doi.org/10.1145/2851613.2851716.

[123] Seongho Kim, Taeyeon Kim, Choong Sun Kim, Hyeongdo Choi, Yong Jun Kim, Gyu Soup Lee, Ockkyun Oh, and Byung Jin Cho. "Two-Dimensional Thermal Haptic Module Based on a Flexible Thermoelectric Device". In: *Soft Robotics* 7.6 (2020), pp. 736–742. URL: https://doi.org/10.1089/soro.2019.0158.

[124] Yeongmi Kim, Arturo Moncada-Torres, Jonas Furrer, Markus Riesch, and Roger Gassert. "Quantification of long cane usage characteristics with the constant contact technique". In: *Applied ergonomics* 55 (2016), pp. 216–225. URL: https://doi.org/10.1016/j.apergo.2016.02.011.

[125] Kris M. Kitani, Brian D. Ziebart, James Andrew Bagnell, and Martial Hebert. "Activity Forecasting". In: *Proceedings of the IEEE European Conference on Computer Vision (ECCV '12)*. Ed. by Andrew Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid. Heidelberg, Zweigniederlassung: Springer, 2012, pp. 201–214. URL: https://doi.org/10.1016/B978-0-12-809276-7.00014-X.

[126] Eunjeong Ko, Jin Sun Ju, and Eun Yi Kim. "Situation-based indoor wayfinding system for the visually impaired". In: *Proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '11)*. New York, NY, USA: ACM, 2011, pp. 35–42. URL: https://doi.org/10.1145/2049536.2049545.

[127] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. "Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. New York, NY, USA: ACM, 2020, 1–19. ISBN: 9781450367080. URL: https://doi.org/10.1145/3313831.3376162.

[128] Andrew J Kolarik, Silvia Cirstea, Shahina Pardhan, and Brian CJ Moore. "A summary of research investigating echolocation abilities of blind and sighted humans". In: *Hearing research* 310 (2014), pp. 60–68. URL: https://doi.org/10.1016/j.heares.2014.01.010.

[129] Harold W Kuhn. "The Hungarian method for the assignment problem". In: *Naval research logistics quarterly* 2.1-2 (1955), pp. 83–97. URL: https://doi.org/10.1007/978-3-540-68279-0_2.

[130] Vladimir Kulyukin, Chaitanya Gharpure, John Nicholson, and Grayson Osborne. "Robot-assisted wayfinding for the visually impaired in structured indoor environments". In: *Autonomous Robots* 21.1 (2006), pp. 29–41. URL: https://doi.org/10.1007/s10514-006-7223-8.

[131] Vladimir Kulyukin, Chaitanya Gharpure, John Nicholson, and Sachin Pavithran. "RFID in robot-assisted indoor navigation for the visually impaired". In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '04)*. Vol. 2. Los Alamitos, CA, USA: IEEE, 2004, pp. 1979–1984. URL: https://doi.org/10.1109/IROS.2004.1389688.

[132] Bineeth Kuriakose, Raju Shrestha, and Frode Eika Sandnes. "Tools and Technologies for Blind and Visually Impaired Navigation Support: A Review". In: *IETE Technical Review* 0.0 (2020), pp. 1–16. URL: https://doi.org/10.1080/02564602.2020.1819893.

[133] Masaki Kuribayashi, Seita Kayukawa, Daisuke Sato, Chieko Asakawa, Hironobu Takagi, and Shigeo Morishima. "Designing a Smartphone-Based Assistance System for Blind People to Recognize Intersections and Obstacles in Indoor Corridors". In: *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '21)*. New York, NY, USA: Springer, 2021, pp. 1–3. URL: https://doi.org/10.1007/978-3-030-94822-1.

[134] Masaki Kuribayashi, Seita Kayukawa, Daisuke Sato, Chieko Asakawa, Hironobu Takagi, and Shigeo Morishima. "Corridor-Walker: Mobile Indoor Walking Assistance for Blind People to Avoid Obstacles and Recognize Intersections". In: *Proceedings of the 24th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '22)*. New York, NY, USA: ACM, 2022, pp. 1–22. URL: https://doi.org/10.1145/3546714.

[135] Masaki Kuribayashi, Seita Kayukawa, Hironobu Takagi, Chieko Asakawa, and Shigeo Morishima. "LineChaser: A Smartphone-Based Navigation System for Blind People to Stand in Line". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. New York, NY, USA: ACM, 2021. URL: https://doi.org/10.1145/3411764.3445451.

[136] Steven Landau, William Wiener, Koorosh Naghshineh, and Ellen Giusti. "Creating Accessible Science Museums With User-Activated Environmental Audio Beacons (Ping!)" In: *Assistive Technology* 17.2 (2005), pp. 133–143. URL: https://doi.org/10.1080/10400435.2005.10132103.

[137] Walter S. Lasecki, Phyo Thiha, Yu Zhong, Erin Brady, and Jeffrey P. Bigham. "Answering Visual Questions with Conversational Crowd Assistants". In: *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. New York, NY, USA: ACM, 2013. URL: https://doi.org/10.1145/2513383.2517033.

[138] Kyungjun Lee, Daisuke Sato, Saki Asakawa, Hernisa Kacorri, and Chieko Asakawa. "Pedestrian Detection with Wearable Cameras for the Blind: A Two-Way Perspective". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. New York, NY, USA: ACM, 2020, 1–12. ISBN: 9781450367080. URL: https://doi.org/10.1145/3313831.3376398.

[139] Linda Lee, J Lee, Serge Egelman, and David Wagner. "Information Disclosure Concerns in the Age of Wearable Computing". In: *NDSS Workshop on Usable Security (USEC '16)*. Vol. 1. Reston, VA, USA: Internet Society, 2016, pp. 1–10. URL: https://doi.org/10.14722/usec.2016.23006.

[140] Young Hoon Lee and Gérard Medioni. "RGB-D camera based navigation for the visually impaired". In: *Proceedings of the RSS*. Vol. 2. Pennsylvania State University, University Park, USA: Citeseer, 2011. URL: https://doi.org/10.1016/j.cviu.2016.03.019.

[141] Young Hoon Lee and Gerard Medioni. "Wearable RGBD indoor navigation system for the blind". In: *Proceedings of the European Conference on Computer Vision (ECCV '14)*. New York, NY, USA: Springer, 2014, pp. 493–508. URL: https://doi.org/10.1007/978-3-319-16199-0_35.

[142] Bridget A Lewis, Jesse L Eisert, and Carryl L Baldwin. "Effect of tactile location, pulse duration, and interpulse interval on perceived urgency". In: *Transportation research record* 2423.1 (2014), pp. 10–14. URL: https://doi.org/10.3141/2423-02.

[143] B Li, JP Muñoz, X Rong, Q Chen, J Xiao, Y Tian, A Arditi, and M Yousuf. "Vision-based Mobile Indoor Assistive Navigation Aid for Blind People." In: *IEEE transactions on mobile computing* 18.3 (2019), pp. 702–714. URL: https://doi.org/10.1109/TMC.2018.2842751.

[144] Bing Li, J Pablo Munoz, Xuejian Rong, Jizhong Xiao, Yingli Tian, and Aries Arditi. "ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind". In: *Proceedings of the European Conference on Computer Vision (ECCV '16)*. New York, NY, USA: Springer, 2016, pp. 448–462. URL: https://doi.org/10.1007/978-3-319-48881-3_31.

[145] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. "Microsoft coco: Common objects in context". In: *Proceedings of the European Conference on Computer Vision (ECCV '14)*. New York, NY, USA: Springer, 2014, pp. 740–755. URL: https://doi.org/10.1007/978-3-319-10602-1_48.

[146] Jacobus C Lock, Iain D Gilchrist, Iain D Gilchrist, Grzegorz Cielniak, and Nicola Bellotto. "Experimental Analysis of a Spatialised Audio Interface for People with Visual Impairments". In: *ACM Transactions on Accessible Computing (TACCESS)* 13.4 (2020), pp. 1–21. URL: https://doi.org/10.1145/3412325.

[147] Jack M Loomis, Reginald G Golledge, and Roberta L Klatzky. "Navigation system for the blind: Auditory display modes and guidance". In: *Presence* 7.2 (1998), pp. 193–203. URL: https://doi.org/10.1162/105474698565677.

[148] Jack M Loomis, Yvonne Lippa, Roberta L Klatzky, and Reginald G Golledge. "Spatial updating of locations specified by 3-D sound and spatial language." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28.2 (2002), p. 335. URL: https://doi.org/10.1037/0278-7393.28.2.335.

[149] BeSpecular LTD. *BeSpecular. Let blind people see through your eyes.* Retrieved in December 27, 2020 from https://www.bespecular.com. 2016.

[150] Chen-Lung Lu, Zi-Yan Liu, Jui-Te Huang, Ching-I Huang, Bo-Hui Wang, Yi Chen, Nien-Hsin Wu, Hsueh-Cheng Wang, Laura Giarré, and Pei-Yi Kuo. "Assistive Navigation Using Deep Reinforcement Learning Guiding Robot With UWB/Voice Beacons and Semantic Feedbacks for Blind and Visually Impaired People". In: *Frontiers in Robotics and AI* 8 (2021). URL: https://doi.org/10.3389/frobt.2021.654132.

[151] Yvonne Ludewig, Nicola Döring, and Nadine Exner. "Design and evaluation of the personality trait extraversion of a shopping robot". In: *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN '12)*. Washington, DC, USA: IEEE, 2012, pp. 372–379. URL: https://doi.org/10.1109/ROMAN.2012.6343781.

[152] Joseph B. Lyons, Thy Vo, Kevin T. Wynne, Sean Mahoney, Chang S. Nam, and Darci Gallimore. "Trusting Autonomous Security Robots: The Role of Reliability and Stated Social Intent". In: *Human Factors* 2020 (2020). URL: https://doi.org/10.1177/0018720820901629.

[153] Shachar Maidenbaum, Shlomi Hanassy, Sami Abboud, Galit Buchs, Daniel-Robert Chebat, Shelly Levy-Tzedek, and Amir Amedi. "The "EyeCane", a new electronic travel aid for the blind: Technology, behavior & swift learning". In: *Restorative neurology and neuroscience* 32.6 (2014), pp. 813–824.

[154] Roberto Manduchi and James M. Coughlan. "The Last Meter: Blind Visual Guidance to a Target". In: *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. New York, NY, USA: ACM, 2014, 3113–3122. URL: https://doi.org/10.1145/2556288.2557328.

[155] Kanak Manjari, Madhushi Verma, and Gaurav Singal. "A survey on Assistive Technology for visually impaired". In: *Internet of Things* 11 (2020), p. 100188. ISSN: 2542-6605. URL: https://doi.org/10.1016/j.iot.2020.100188.

[156] Eitan Marder-Eppstein, Eric Berger, Tully Foote, Brian Gerkey, and Kurt Konolige. "The office marathon: Robust navigation in an indoor office environment". In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '10)*. IEEE. Los Alamitos, CA, USA, 2010, pp. 300–307. URL: https://doi.org/10.1109/ROBOT.2010.5509725.

[157] Dawn C Marshall, John D Lee, and P Albert Austria. “Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness”. In: *Human factors* 49.1 (2007), pp. 145–157. URL: https://doi.org/10.1518/001872007779598145.

[158] Edward Marshall and Sue Baker. “Alarms in nuclear power plant control rooms: current approaches and future design”. In: *Human factors in alarm design*. Taylor & Francis, Inc. Milton Park, UK, 1995, pp. 183–191.

[159] Manuel Martinez, Angela Constantinescu, Boris Schauerte, Daniel Koester, and Rainer Stiefelhagen. “Cognitive evaluation of haptic and audio feedback in short range navigation tasks”. In: *Proceedings of the International Conference on Computers for Handicapped Persons (ICCHP ’14)*. New York, NY, USA: Springer, 2014, pp. 128–135. URL: https://doi.org/10.1007/978-3-319-08599-9_20.

[160] Natalina Martiniello, Werner Eisenbarth, Christine Lehane, Aaron Johnson, and Walter Wittich. “Exploring the use of smartphones and tablets among people with visual impairments: Are mainstream devices replacing the use of traditional visual aids?” In: *Assistive Technology* (2019), pp. 1–12. URL: https://doi.org/10.1080/10400435.2019.1682084.

[161] Rajesh Kannan Megalingam, Souraj Vishnu, Vishnu Sasikumar, and Sajikumar Sreekumar. “Autonomous Path Guiding Robot for Visually Impaired People”. In: *Cognitive Informatics and Soft Computing*. Ed. by Pradeep Kumar Mallick, Valentina Emilia Balas, Akash Kumar Bhoi, and Ahmed F. Zobaa. Singapore: Springer Singapore, 2019, pp. 257–266. URL: https://doi.org/10.1007/978-981-13-0617-4_25.

[162] Apostolos Meliones and Demetrios Sampson. “Blind MuseumTourer: A System for Self-Guided Tours in Museums and Blind Indoor Navigation”. In: *Technologies* 6.1 (2018), p. 4. URL: https://doi.org/10.3390/technologies6010004.

[163] Christina Meredith and Judy Edworthy. “Are there too many alarms in the intensive care unit? An overview of the problems”. In: *Journal of advanced nursing* 21.1 (1995), pp. 15–20. URL: https://doi.org/10.1177/001872089303500408.

[164] Susana Mesquita and Maria João Carneiro. “Accessibility of European museums to visitors with visual impairments”. In: *Disability & Society* 31.3 (2016), pp. 373–388. URL: https://doi.org/10.1080/09687599.2016.1167671.

[165] Microsoft. *Seeing AI. A free app that narrates the world around you.* Retrieved in December 27, 2020 from https://www.microsoft.com/en-us/seeing-ai. 2020.

[166] Microsoft. *SoundScape.* Retrieved in December 27, 2020 from https://www.microsoft.com/en-us/research/product/soundscape/. 2020.

[167] Sewon Min, Jordan Boyd-Graber, Chris Alberti, Danqi Chen, Eunsol Choi, Michael Collins, Kelvin Guu, Hannaneh Hajishirzi, Kenton Lee, Jennimaria Palomaki, et al. “NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned”. In: *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*. Ed. by Hugo Jair Escalante and Katja Hofmann. Vol. 133. Proceedings of Machine Learning Research. PMLR, 2021, pp. 86–111. URL: https://proceedings.mlr.press/v133/min21a.html.

[168] MIPsoft. *Blindsquare.* 2020.

[169] Hideo Mori and Shinji Kotani. “Robotic travel aid for the blind: HARUNOBU-6”. In: *Proceedings of the European Conference on Disability, Virtual Reality, and Assistive Technology (ECDVRAT ’98)*. Reading, UK: The University of Reading, 1998. URL: https://doi.org/10.1007/978-1-4471-1580-9_22.

[170] John Morris and James Mueller. "Blind and deaf consumer preferences for android and iOS smartphones". In: *Inclusive designing*. New York, NY, USA: Springer, 2014, pp. 69–79. URL: https://doi.org/10.1007/978-3-319-05095-9_7.

[171] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. "How simple rules determine pedestrian behavior and crowd disasters". In: *Proceedings of the of the National Academy of Sciences (PNAS '11)* 108.17 (2011), pp. 6884–6888. URL: https://doi.org/10.1073/pnas.1016507108.

[172] Masayuki Murata, Dragan Ahmetovic, Daisuke Sato, Hironobu Takagi, Kris M Kitani, and Chieko Asakawa. "Smartphone-based indoor localization for blind navigation across building complexes". In: *Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom ('18))*. Los Alamitos, CA, USA: IEEE, 2018, pp. 1–10. URL: https://doi.org/10.1109/PERCOM.2018.8444593.

[173] The Andy Warhol Museum. *Accessibility-The Andy Warhol Museum*. 2018. URL: https://www.warhol.org/accessibility-accommodations/.

[174] Madoka Nakajima and Shinichiro Haruyama. "Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing". In: *Proceedings of the IEEE International Conference on Communications in China (ICCC '12)*. Los Alamitos, CA, USA: IEEE, 2012, pp. 524–529. URL: https://doi.org/10.1109/ICCChina.2012.6356940.

[175] Madoka Nakajima and Shinichiro Haruyama. "New indoor navigation system for visually impaired people using visible light communication". In: *EURASIP Journal on Wireless Communications and Networking* 2013.1 (2013), p. 37. URL: https://doi.org/10.1186/1687-1499-2013-37.

[176] Yasushi Nakauchi and Reid Simmons. "A social robot that stands in line". In: *Autonomous Robots* 12.3 (2002), pp. 313–324. URL: https://doi.org/10.1023/A:1015273816637.

[177] Amal Nanavati, Xiang Zhi Tan, and Aaron Steinfeld. "Coupled Indoor Navigation for People Who Are Blind". In: *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. New York, NY, USA: ACM, 2018, pp. 201–202. ISBN: 978-1-4503-5615-2. URL: https://doi.org/10.1145/3173386.3176976.

[178] Arshad Nasser, Kai-Ning Keng, and Kening Zhu. "Thermalcane: Exploring thermotactile directional cues on cane-grip for non-visual navigation". In: *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*. New York, NY, USA: ACM, 2020, pp. 1–12. URL: https://doi.org/10.1145/3373625.3417004.

[179] United Nations. *Convention on the Rights of Persons with Disabilities (CRPD)*. 2006. URL: https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html.

[180] Marketta Niemelä, Päivi Heikkilä, Hanna Lammi, and Virpi Oksman. "A Social Robot in a Shopping Mall: Studies on Acceptance and Stakeholder Expectations". In: *Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction*. Cham: Springer International Publishing, 2019, pp. 119–144. ISBN: 978-3-030-17107-0. URL: https://doi.org/10.1007/978-3-030-17107-0_7.

[181] Illah R Nourbakhsh, Clayton Kunz, and Thomas Willeke. "The mobot museum robot installations: A five year experiment". In: *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '03)*. Vol. 4. IEEE. New York, NY, USA: IEEE, 2003, pp. 3636–3641. URL: https://doi.org/10.1109/IROS.2003.1249720.

[182] Orcam. *Help People who are Blind or Partially Sighted - OrCam*. Retrieved in December 27, 2020 from https://www.orcam.com/en/. 2020.

[183] Joyojeet Pal, Anandhi Viswanathan, and Ji-Hye Song. "Smartphone Adoption Drivers and Challenges in Urban Living: Cases from Seoul and Bangalore". In: *Proceedings of the ACM Indian Conference on Human Computer Interaction (IHCI '16)*. New York, NY, USA: ACM, 2016, 24–34. URL: https://doi.org/10.1145/3014362.3014364.

[184] Agnivesh Pani, Sabya Mishra, Mihalis Golias, and Miguel Figliozzi. "Evaluating public acceptance of autonomous delivery robots during COVID-19 pandemic". In: *Transportation research part D: transport and environment* 89 (2020), p. 102600. URL: https://doi.org/10.1016/j.trd.2020.102600.

[185] Suraj R. Pardeshi, Vikul J. Pawar, Kailas D. Kharat, and Sachin Chavan. "Assistive Technologies for Visually Impaired Persons Using Image Processing Techniques – A Survey". In: *Recent Trends in Image Processing and Pattern Recognition*. Ed. by K. C. Santosh and Bharti Gawali. Singapore: Springer Singapore, 2021, pp. 95–110. ISBN: 978-981-16-0507-9. URL: https://doi.org/10.1007/978-981-16-0507-9_9.

[186] Phil Parette and Marcia Scherer. "Assistive Technology Use and Stigma". In: *Education and Training in Developmental Disabilities* 39.3 (2004), pp. 217–226. ISSN: 15470350. URL: http://www.jstor.org/stable/23880164.

[187] Jagannadh Pariti, Vinita Tibdewal, and Tae Oh. "Intelligent Mobility Cane-Lessons Learned from Evaluation of Obstacle Notification System using a Haptic Approach". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. New York, NY, USA: ACM, 2020, pp. 1–8. URL: https://doi.org/10.1145/3334480.3375217.

[188] Huang Peng, Guangming Song, Jian You, Ying Zhang, and Jie Lian. "An Indoor Navigation Service Robot System Based on Vibration Tactile Feedback". In: *International Journal of Social Robotics* 9.3 (2017), pp. 331–341. ISSN: 1875-4805. URL: https://doi.org/10.1007/s12369-017-0403-1.

[189] Manoj Penmetcha, Arabinda Samantaray, and Byung-Cheol Min. "Smartresponse: Emergency and non-emergency response for smartphone based indoor localization applications". In: *Proceedings of the International Conference on Human-Computer Interaction (HCII '17)*. New York, NY, USA: Springer, 2017, pp. 398–404. URL: https://doi.org/10.1007/978-3-319-58753-0_57.

[190] Helen Petrie, Valerie Johnson, Thomas Strothotte, Andreas Raab, Steffi Fritz, and Rainer Michel. "MoBIC: Designing a travel aid for blind and elderly people". In: *The Journal of Navigation* 49.1 (1996), pp. 45–52. URL: https://doi.org/10.1145/191028.191051.

[191] Huy-Hieu Pham, Thi-Lan Le, and Nicolas Vuillerme. "Real-time obstacle detection system in indoor environment for the visually impaired using microsoft kinect sensor". In: *Journal of Sensors* 2016 (2016). URL: https://doi.org/10.1155/2016/3754918.

[192] Ioannis Politis, Stephen A Brewster, and Frank Pollick. "Evaluating multimodal driver displays under varying situational urgency". In: *Proceeding of the ACM CHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM. New York, NY, USA, 2014, pp. 4067–4076. URL: https://doi.org/10.1145/2556288.2556988.

[193] Abishai Polus, Joseph L Schofer, and Ariela Ushpiz. "Pedestrian flow and level of service". In: *Journal of transportation engineering* 109.1 (1983), pp. 46–56. URL: https://doi.org/10.1061/(ASCE)0733-947X(1983)109:1(46).

[194] Benjamin Poppinga, Charlotte Magnusson, Martin Pielot, and Kirsten Rassmus-Gröhn. "TouchOver map: audio-tactile exploration of interactive maps". In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. New York, NY, USA: ACM, 2011, pp. 545–550. URL: https://doi.org/10.1145/2037373.2037458.

[195] Vivek Pradeep, Gerard Medioni, and James Weiland. "Robot vision for the visually impaired". In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. San Francisco, CA, USA: IEEE, 2010, pp. 15–22. URL: https://doi.org/10.1109/CVPRW.2010.5543579.

[196] Stephanie M Pratt, Bridget A Lewis, BN Penaranda, Daniel M Roberts, Christian Gonzalez, and Carryl L Baldwin. "Perceived urgency scaling in tactile alerts". In: *Proceedings of the of the Human Factors and Ergonomics Society Annual Meeting (HFES '12)*. Los Angeles, CA: SAGE Publications Sage CA, 2012, pp. 1303–1306. URL: https://doi.org/10.1177/1071181312561378.

[197] Giorgio Presti, Dragan Ahmetovic, Mattia Ducci, Cristian Bernareggi, Luca Ludovico, Adriano Baratè, Federico Avanzini, and Sergio Mascetti. "WatchOut: Obstacle sonification for people with visual impairment or blindness". In: *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. New York, NY, USA: ACM, 2019, pp. 402–413. URL: https://doi.org/10.1145/3308561.3353779.

[198] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K. Kane. "The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. New York, NY, USA: ACM, 2016, 4884–4895. ISBN: 9781450333627. URL: https://doi.org/10.1145/2858036.2858130.

[199] Morgan Quigley, Ken Conley, Brian P. Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. "ROS: an open-source Robot Operating System". In: *ICRA Workshop on Open Source Software*. Los Alamitos, CA, USA: IEEE, 2009, p. 5.

[200] Pablo-Alejandro Quinones, Tammy Greene, Rayoung Yang, and Mark Newman. "Supporting visually impaired navigation: a needs-finding study". In: *Extended Abstracts on ACM CHI Conference on Human Factors in Computing Systems (CHI EA '11)*. ACM. New York, NY, USA, 2011, pp. 1645–1650. URL: https://doi.org/10.1145/1979742.1979822.

[201] Lisa Ran, Sumi Helal, and Steve Moore. "Drishti: an integrated indoor/outdoor blind navigation system and service". In: *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications (PerCom '04)*. Los Alamitos, CA, USA: IEEE, 2004, pp. 23–30. URL: https://doi.org/10.1109/PERCOM.2004.1276842.

[202] Kyle Rector, Rachel Bartlett, and Sean Mullan. "Exploring Aural and Haptic Feedback for Visually Impaired People on a Track: A Wizard of Oz Study". In: *Proceedings of the ACM SIGACCESS Conference on Computers and accessibility (ASSETS '18)*. New York, NY, USA: ACM, 2018, pp. 295–306. URL: https://doi.org/10.1145/3234695.3236345.

[203] Kyle Rector, Keith Salmon, Dan Thornton, Neel Joshi, and Meredith Ringel Morris. "Eyes-Free Art: Exploring Proxemic Audio Interfaces For Blind and Low Vision Art Engagement". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.3 (2017), pp. 1–21. URL: https://doi.org/10.1145/3130958.

[204] Joseph Redmon and Ali Farhadi. "YOLO9000: Better, Faster, Stronger". In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR '17)*. IEEE. Los Alamitos, CA, USA, 2017, pp. 6517–6525. URL: https://doi.org/10.1109/CVPR.2017.690.

[205] Joseph Redmon and Ali Farhadi. *YOLOv3: An Incremental Improvement*. 2018. arXiv: 1804.02767 [cs.CV].

[206] Jillian M Rickly, Nigel Halpern, Marcus Hansen, and John Welsman. "Travelling with a Guide Dog: Experiences of People with Vision Impairment". In: *Sustainability* 13.5 (2021), p. 2840. URL: https://www.mdpi.com/2071-1050/13/5/2840.

[207] Timothy H Riehle, Shane M Anderson, Patrick A Lichter, Nicholas A Giudice, Suneel I Sheikh, Robert J Knuesel, Daniel T Kollmann, and Daniel S Hedin. "Indoor magnetic navigation for the blind". In: *Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '12)*. Los Alamitos, CA, USA: IEEE, 2012, pp. 1972–1975. URL: https://doi.org/10.1109/EMBC.2012.6346342.

[208] Jean-Paul Rodrigue, Claude Comtois, and Brian Slack. *The geography of transport systems*. Abingdon, Oxfordshire, UK: Routledge, 2016.

[209] André Rodrigues, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. "Getting Smartphones to Talkback: Understanding the Smartphone Adoption Process of Blind Users". In: *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '15)*. New York, NY, USA: ACM, 2015, 23–32. URL: https://doi.org/10.1145/2700648.2809842.

[210] Alberto Rodríguez, J Javier Yebes, Pablo F Alcantarilla, Luis M Bergasa, Javier Almazán, and Andrés Cela. "Assisting the visually impaired: obstacle detection and warning system by acoustic feedback". In: *Sensors* 12.12 (2012), pp. 17476–17496. URL: https://doi.org/10.3390/s121217476.

[211] David A Ross and Bruce B Blasch. "Wearable interfaces for orientation and wayfinding". In: *Proceedings of the fourth international ACM conference on Assistive technologies*. New York, NY, USA: ACM, 2000, pp. 193–200. URL: https://doi.org/10.1145/354324.354380.

[212] Shadan Sadeghian Borojeni, Susanne CJ Boll, Wilko Heuten, Heinrich H Bülthoff, and Lewis Chuang. "Feel the Movement: Real Motion Influences Responses to Take-over Requests in Highly Automated Vehicles". In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM. New York, NY, USA, 2018, p. 246. URL: https://doi.org/10.1145/3173574.3173820.

[213] Shozo Saegusa, Yuya Yasuda, Yoshitaka Uratani, Eiichirou Tanaka, Toshiaki Makino, and Jen-Yuan Chang. "Development of a guide-dog robot: human–robot interface considering walking conditions for a visually handicapped person". In: *Microsystem Technologies* 17.5 (2011), pp. 1169–1174. ISSN: 1432-1858. URL: https://doi.org/10.1007/s00542-010-1219-1.

[214] Manaswi Saha, Alexander J Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. "Closing the Gap: Designing for the Last-Few-Meters Wayfinding Problem for People with Visual Impairments". In: *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. New York, NY, USA: ACM, 2019, pp. 222–235. URL: https://doi.org/10.1145/3308561.3353776.

[215] Daisuke Sato, Uran Oh, João Guerreiro, Dragan Ahmetovic, Kakuya Naito, Hironobu Takagi, Kris M Kitani, and Chieko Asakawa. "NavCog3 in the wild: Large-scale blind indoor navigation assistant with semantic features". In: *ACM Transactions on Accessible Computing (TACCESS)* 12.3 (2019), p. 14. URL: https://doi.org/10.1145/3340319.

[216] Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. “Navcog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment”. In: *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS ’17).* New York, NY, USA: ACM, 2017, pp. 270–279. URL: https://doi.org/10.1145/3132525.3132535.

[217] Nina Savela, Tuuli Turja, and Atte Oksanen. “Social Acceptance of Robots in Different Occupational Fields: A Aystematic Literature Review”. In: *International Journal of Social Robotics* 10.4 (2018), pp. 493–502. URL: https://doi.org/10.1007/s12369-017-0452-5.

[218] Marcia Joslyn Scherer. *Living in the state of stuck: How assistive technology impacts the lives of people with disabilities.* Brookline, MA, USA: Brookline Books, 2005.

[219] Victor R Schinazi, Tyler Thrash, and Daniel-Robert Chebat. “Spatial navigation by congenitally blind individuals”. In: *Wiley Interdisciplinary Reviews: Cognitive Science* 7.1 (2016), pp. 37–58. URL: https://doi.org/10.1002/wcs.1375.

[220] Huiying Shen and James M. Coughlan. “Towards a Real-Time System for Finding and Reading Signs for Visually Impaired Users”. In: *Proceedings of the International Conference on Computers Helping People with Special Needs (ICCHP’12).* New York, NY, USA: Springer, 2012, 41–47. URL: https://doi.org/10.1007/978-3-642-31534-3_7.

[221] Chao Shi, Satoru Satake, Takayuki Kanda, and Hiroshi Ishiguro. “How Would Store Managers Employ Social Robots?” In: *Proceedings of the Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI ’16).* Piscataway, NJ, USA: IEEE Press, 2016, 519–520. ISBN: 9781467383707. URL: https://doi.org/10.1109/HRI.2016.7451835.

[222] Kristen Shinohara and Jacob O. Wobbrock. “In the Shadow of Misperception: Assistive Technology Use and Social Interactions”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’11).* New York, NY, USA: ACM, 2011, 705–714. ISBN: 9781450302289. URL: https://doi.org/10.1145/1978942.1979044.

[223] Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. “Interactive humanoid robots for a science museum”. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot Interaction (HRI ’06).* New York, NY, USA: ACM, 2006, pp. 305–312. URL: https://doi.org/10.1145/1121241.1121293.

[224] Shraga Shoval, Johann Borenstein, and Yoram Koren. “The Navbelt-A computerized travel aid for the blind based on mobile robotics technology”. In: *IEEE Transactions on Biomedical Engineering* 45.11 (1998), pp. 1376–1386. URL: https://doi.org/10.1109/10.725334.

[225] Jennie Small, Simon Darcy, and Tanya Packer. “The embodied tourist experiences of people with vision impairment: Management implications beyond the visual gaze”. In: *Tourism Management* 33.4 (2012), pp. 941–950. ISSN: 0261-5177. URL: https://doi.org/10.1016/j.tourman.2011.09.015.

[226] John P Snyder. “Flattening the earth: two thousand years of map projections”. In: (1997), pp. 5–8.

[227] Adam J Spiers and Aaron M Dollar. “Design and evaluation of shape-changing haptic interfaces for pedestrian navigation assistance”. In: *IEEE transactions on haptics* 10.1 (2016), pp. 17–28. URL: https://doi.org/10.1109/TOH.2016.2582481.

[228] Adam J Spiers, Janet Van Der Linden, Sarah Wiseman, and Maria Oshodi. “Testing a Shape-Changing Haptic Navigation Device With Vision-Impaired and Sighted Audiences in an Immersive Theater Setting”. In: *IEEE Transactions on Human-Machine Systems* 48.6 (2018), pp. 614–625. URL: https://doi.org/10.1109/THMS.2018.2868466.

[229] Julian Striegl, Claudia Lotisch, Jan Schmalfuss-Schwarz, and G. Weber. "Analysis of Indoor Maps Accounting the Needs of People with Impairments". In: *ICCHP*. Vol. 12377. New York, NY, USA: Springer, 2020. URL: https://doi.org/10.1007/978-3-030-58805-2_36.

[230] Susumu Tachi, Kazuo Tanie, Kiyoshi Komoriya, and Minoru Abe. "Electrocutaneous communication in a guide dog robot (MELDOG)". In: *IEEE Transactions on Biomedical Engineering* BME-32.7 (1985), pp. 461–469. URL: https://doi.org/10.1109/TBME.1985.325561.

[231] Sebastian Thrun, Maren Bennewitz, Wolfram Burgard, Armin B Cremers, Frank Dellaert, Dieter Fox, Dirk Hahnel, Charles Rosenberg, Nicholas Roy, Jamieson Schulte, et al. "MINERVA: A second-generation museum tour-guide robot". In: *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*. Vol. 3. IEEE. New York, NY, USA: IEEE, 1999, pp. 1999–2005. URL: https://doi.org/10.1109/ROBOT.1999.770401.

[232] Kazuteru Tobita, Katsuyuki Sagayama, and Hironori Ogawa. "Examination of a Guidance Robot for Visually Impaired People". In: *Journal of Robotics and Mechatronics* 29.4 (2017), pp. 720–727. URL: https://doi.org/10.20965/jrm.2017.p0720.

[233] Nelson Daniel Troncoso Aldas, Sooyeon Lee, Chonghan Lee, Mary Beth Rosson, John M Carroll, and Vijaykrishnan Narayanan. "AIGuide: An Augmented Reality Hand Guidance Application for People with Visual Impairments". In: *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20)*. New York, NY, USA: ACM, 2020, pp. 1–13. URL: https://doi.org/10.1145/3373625.3417028.

[234] Yu-Chih Tung and Kang G Shin. "Use of Phone Sensors to Enhance Distracted Pedestrians' Safety". In: *IEEE Transactions on Mobile Computing* 17.6 (2018), pp. 1469–1482. URL: https://doi.org/10.1109/TMC.2017.2764909.

[235] Iwan Ulrich and Johann Borenstein. "The GuideCane-applying mobile robot technologies to assist the visually impaired". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 31.2 (2001), pp. 131–136. URL: https://doi.org/10.1109/3468.911370.

[236] Mohammad Moeen Valipoor and Angélica de Antonio. "Recent trends in computer vision-driven scene understanding for VI/blind users: a systematic mapping". In: *Universal Access in the Information Society* (2022), pp. 1–23. URL: https://doi.org/10.1007/s10209-022-00868-w.

[237] Savvas Varitimiadis, Konstantinos Kotis, Andreas Skamagis, Alexandros Tzortzakakis, George Tsekouras, and Dimitris Spiliotopoulos. "Towards implementing an AI chatbot platform for museums". In: *Proceedings of the International Conference on Cultural Informatics, Communication & Media Studies (CICMS '20)*. Mytilene, Island: The Department of Cultural Technology and Communication (DCTC), 2020. URL: https://doi.org/10.12681/cicms.2732.

[238] Roberto Vaz, Diamantino Freitas, and António Coelho. "Blind and Visually Impaired Visitors' Experiences in Museums: Increasing Accessibility through Assistive Technologies". In: *International Journal of the Inclusive Museum* 13.2 (2020). URL: https://doi.org/10.18848/1835-2014/CGP/v13i02/57-80.

[239] Roberto Vaz, Diamantino Freitas, and António Coelho. "Perspectives of Visually Impaired Visitors on Museums: Towards an Integrative and Multisensory Framework to Enhance the Museum Experience". In: *9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (DSAI '20)*. New York, NY, USA: ACM, 2020, pp. 17–21. URL: https://doi.org/10.1145/3439231.3439272.

[240] Roberto Vaz, Diamantino Freitas, and António Coelho. "Visiting Museums from the Perspective of Visually Impaired Visitors: Experiences and Accessibility Resources in Portuguese Museums". In: *International Journal of the Inclusive Museum* 14.1 (2021). URL: https://doi.org/10.18848/1835-2014/CGP/v14i01/71-93.

[241] N Veeranjaneyulu. "A Meta-Analysis on Obstacle Detection for Visually Impaired People". In: *i-manager's Journal on Pattern Recognition* 6.1 (2019), pp. 40–62. URL: https://doi.org/10.26634/jpr.6.1.15523.

[242] Japan Federation of the Visually Impaired. *Research Project Report for Developing Assistive Systems that Ensure the Safety of the Visually Impaired People in Out-door Spaces (in Japanese)*. 2013. URL: http://nichimou.org/wp-content/uploads/2014/02/9aff1483649e8747de10bfbd5c8bca55.pdf.

[243] Hsueh-Cheng Wang, Robert K Katzschmann, Santani Teng, Brandon Araki, Laura Giarré, and Daniela Rus. "Enabling independent navigation for visually impaired people through a wearable vision-based feedback system". In: *2017 IEEE international conference on robotics and automation (ICRA '17)*. Los Alamitos, CA, USA: IEEE, 2017, pp. 6533–6540. URL: https://doi.org/10.1109/ICRA.2017.7989772.

[244] Weimin Wang, Ken Sakurada, and Nobuo Kawaguchi. "Reflectance Intensity Assisted Automatic and Accurate Extrinsic Calibration of 3D LiDAR and Panoramic Camera Using a Printed Chessboard". In: *Remote Sensing* 9.8 (2017), p. 851. ISSN: 2072-4292. URL: https://doi.org/10.3390/rs9080851.

[245] Richard Welsh. *Foundations of orientation and mobility*. Tech. rep. American Foundation for the Blind, 1981.

[246] Lorraine Whitmarsh. "The benefits of guide dog ownership". In: *Visual impairment research* 7.1 (2005), pp. 27–42. URL: https://doi.org/110.1080/13882350590956439.

[247] William R Wiener, Richard L Welsh, and Bruce B Blasch. *Foundations of orientation and mobility*. Vol. 1. New York, NY, USA: American Foundation for the Blind, 2010. ISBN: 0891284486.

[248] Michele A Williams, Caroline Galbraith, Shaun K Kane, and Amy Hurst. "Just let the cane hit it: how the blind and sighted see navigation differently". In: *Proceedings of the 16th International ACM SIGACCESS Conference on Computers and accessibility (ASSETS '14)*. ACM. New York, NY, USA, 2014, pp. 217–224. URL: https://doi.org/10.1145/2661334.2661380.

[249] Michele A Williams, Amy Hurst, and Shaun K Kane. "" Pray before you step out" describing personal and situational blind navigation behaviors". In: *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. New York, NY, USA: ACM, 2013, pp. 1–8. URL: https://doi.org/10.1145/2513383.2513449.

[250] Jeff Wilson, Bruce N Walker, Jeffrey Lindsay, Craig Cambias, and Frank Dellaert. "Swan: System for wearable audio navigation". In: *Proceedings of the IEEE international symposium on wearable computers (ISWC '07)*. Los Alamitos, CA, USA: IEEE, 2007, pp. 91–98. URL: https://doi.org/10.1109/ISWC.2007.4373786.

[251] Paul F Wilson, Janet Stott, Jason M Warnett, Alex Attridge, M Paul Smith, and Mark A Williams. "Evaluation of touchable 3D-printed replicas in museums". In: *Curator: The Museum Journal* 60.4 (2017), pp. 445–465. URL: https://doi.org/10.1111/cura.12244.

[252] Anxing Xiao, Wenzhe Tong, Lizhi Yang, Jun Zeng, Zhongyu Li, and Koushil Sreenath. *Robotic Guide Dog: Leading a Human with Leash-Guided Hybrid Physical Interaction.* 2021. arXiv: 2103.14300 [cs.RO].

[253] Takuma Yagi, Karttikeya Mangalam, Ryo Yonetani, and Yoichi Sato. "Future person localization in first-person videos". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '18).* Los Alamitos, CA, USA: IEEE, 2018, pp. 7593–7602. URL: https://doi.org/10.1109/CVPR.2018.00792.

[254] Yutaro Yamanaka, Seita Kayukawa, Hironobu Takagi, Yuichi Nagaoka, Yoshimune Hiratsuka, and Satoshi Kurihara. "One-Shot Wayfinding Method for Blind People via OCR and Arrow Analysis with a 360-degree Smartphone Camera". In: *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '21).* New York, NY, USA: Springer, 2021. URL: https://doi.org/10.1007/978-3-030-94822-1_9.

[255] Keiichi Yamazaki, Akiko Yamazaki, Mai Okada, Yoshinori Kuno, Yoshinori Kobayashi, Yosuke Hoshi, Karola Pitsch, Paul Luff, Dirk Vom Lehn, and Christian Heath. "Revealing Gauguin: engaging visitors in robot guide's explanation in an art museum". In: *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '09).* New York, NY, USA: ACM, 2009, pp. 1437–1446. URL: https://doi.org/10.1145/1518701.1518919.

[256] Koji Yatani, Nikola Banovic, and Khai Truong. "SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12).* New York, NY, USA: ACM, 2012, pp. 415–424. URL: https://doi.org/10.1145/2207676.2207734.

[257] Cang Ye, Soonhac Hong, Xiangfei Qian, and Wei Wu. "Co-robotic cane: A new robotic navigation aid for the visually impaired". In: *IEEE Systems, Man, and Cybernetics Magazine* 2.2 (2016), pp. 33–42. URL: https://doi.org/10.1109/MSMC.2015.2501167.

[258] Kumar Yelamarthi, Daniel Haas, Daniel Nielsen, and Shawn Mothersell. "RFID and GPS integrated navigation system for the visually impaired". In: *Proceedings of the 53rd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS '10).* Piscataway, NJ, USA: IEEE Press, 2010, pp. 1149–1152. URL: https://doi.org/10.1109/MWSCAS.2010.5548863.

[259] Chris Yoon, Ryan Louie, Jeremy Ryan, MinhKhang Vu, Hyegi Bang, William Derksen, and Paul Ruvolo. "Leveraging Augmented Reality to Create Apps for People with Visual Disabilities: A Case Study in Indoor Navigation". In: *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19).* New York, NY, USA: ACM, 2019, pp. 210–221. URL: https://doi.org/10.1145/3308561.3353788.

[260] Limin Zeng, Björn Einert, Alexander Pitkin, and Gerhard Weber. "HapticRein: Design and Development of an Interactive Haptic Rein for a Guidance Robot". In: *Proceedings of the Computers Helping People with Special Needs (ICCHP '18).* Ed. by Klaus Miesenberger and Georgios Kouroupetroglou. Cham: Springer International Publishing, 2018, pp. 94–101. ISBN: 978-3-319-94274-2.

[261] Limin Zeng, Markus Simros, and Gerhard Weber. "Camera-based mobile electronic travel aids support for cognitive mapping of unknown spaces". In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. New York, NY, USA: ACM, 2017, p. 8. URL: https://doi.org/10.1145/3098279.3098563.

[262] Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. "Mpiigaze: Real-world dataset and deep appearance-based gaze estimation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017). URL: https://doi.org/10.1109/TPAMI.2017.2778103.

[263] Yuhang Zhao, Shaomei Wu, Lindsay Reynolds, and Shiri Azenkot. "A Face Recognition Application for People with Visual Impairments: Understanding Use Beyond the Lab". In: *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '18)*. New York, NY, USA: ACM, 2018, 1–14. URL: https://doi.org/10.1145/3173574.3173789.

[264] Yu Zhong, Pierre J Garrigues, and Jeffrey P Bigham. "Real time object scanning using a mobile phone and cloud-based visual search engine". In: *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. New York, NY, USA: ACM, 2013, pp. 1–8. URL: https://doi.org/10.1145/2513383.2513443.

[265] Yu Zhong, Walter S. Lasecki, Erin Brady, and Jeffrey P. Bigham. "RegionSpeak: Quick Comprehensive Spatial Descriptions of Complex Images for Blind Users". In: *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. New York, NY, USA: ACM, 2015, 2353–2362. URL: https://doi.org/10.1145/2702123.2702437.

[266] Annuska Zolyomi, Anushree Shukla, and Jaime Snyder. "Technology-Mediated Sight: A Case Study of Early Adopters of a Low Vision Assistive Technology". In: *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17)*. New York, NY, USA: ACM, 2017, 220–229. ISBN: 9781450349260. URL: https://doi.org/10.1145/3132525.3132552.

Works of Lists

Journal Papers and Conference Full Papers

1. Seita Kayukawa, Keita Higuchi, João Guerreiro, Shigeo Morishima, Yoichi Sato, Kris Kitani, and Chieko Asakawa. 2019. "BBeep: A Sonic Collision Avoidance System for Blind Travellers and Nearby Pedestrians". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19).*
DOI: https://doi.org/10.1145/3290605.3300282

2. Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2020. "Guiding Blind Pedestrians in Public Spaces by Understanding Walking Behavior of Nearby Pedestrians". In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT).* 4, 3, Article 85, 22 pages.
DOI: https://doi.org/10.1145/3411825

3. Masaki Kuribayashi*, Seita Kayukawa*, Hironobu Takagi, Chieko Asakawa, and Shigeo Morishima (* - equal contribution). 2021. "LineChaser: A Smartphone-Based Navigation System for Blind People to Stand in Line". In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21).*
DOI: https://doi.org/10.1145/3411764.3445451

4. Yutaro Yamanaka, Seita Kayukawa, Hironobu Takagi, Yuichi Nagaoka, Yoshimune Hiratsuka, and Satoshi Kurihara. 2021. "One-Shot Wayfinding Method for Blind People via OCR and Arrow Analysis with a 360-degree Smartphone Camera". In *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '21).*
DOI: https://doi.org/10.1007/978-3-030-94822-1_9

5. Seita Kayukawa, Daisuke Sato, Masayuki Murata, Tatsuya Ishihara, Akihiro Kosugi, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2022. How Users, Facility Managers, and Bystanders Perceive and Accept a Navigation Robot for Visually Impaired People in Public Buildings. In *Proceedings of the 31st IEEE International Conference on Robot & Human Interactive Communication (IEEE RO-MAN '22).*

6. Masaki Kuribayashi, Seita Kayukawa, Jayakorn Vongkulbhisal, Daisuke Sato, Chieko Asakawa, Hironobu Takagi, and Shigeo Morishima. 2022. "Corridor-Walker: Mobile Indoor Walking Assistance for Blind People to Avoid Obstacles and Recognize Intersections". In *Proceedings of the 24th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '22).*
DOI: https://doi.org/10.1145/3546714

7. Wang Xiyue, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2022. "Bento-Museum: 3D and Layered Interactive Museum Map for Blind Visitors ". In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22).*
DOI: https://doi.org/10.1145/3517428.3544811

Conference Short Papers, Demonstrations, and Posters

8. Seita Kayukawa, Hironobu Takagi, João Guerreiro, Shigeo Morishima, and Chieko Asakawa. 2020. “Smartphone-Based Assistance for Blind People to Stand in Lines”. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI ‘20 LBW).*
   DOI: https://doi.org/10.1145/3334480.3382954

9. Seita Kayukawa, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. 2020. “BlindPilot: A Robotic Local Navigation System that Leads Blind People to a Landmark Object”. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI ‘20 LBW).*
   DOI: https://doi.org/10.1145/3170427.3189085

10. Masaki Kuribayashi, Seita Kayukawa, Jayakorn Vongkulbhisal, Daisuke Sato, Chieko Asakawa, Hironobu Takagi, and Shigeo Morishima. 2021. “Designing a Smartphone-Based Assistance System for Blind People to Recognize Intersections and Obstacles in Indoor Corridors”. In *Proceedings of the 18th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous ’21 Poster).*
    DOI: https://doi.org/10.1007/978-3-030-94822-1

11. Seita Kayukawa, Keita Higuchi, Ryo Yonetani, Masanori Nakamura, Yoichi Sato, and Shigeo Morishima. 2018. “Dynamic Object Scanning: Object-Based Elastic Timeline for Quickly Browsing First-Person Videos”. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI ‘18 LBW and DEMO).*
    DOI: https://doi.org/10.1145/3170427.3189085

12. Ryo Shimamura, Seita Kayukawa, Takayuki Nakatsuka, Shoki Miyagawa, and Shigeo Morishima. 2019. “A Study on the Sense of Burden and Body Ownership on Virtual Slope”. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR ’19 Poster).*
    DOI: https://doi.org/10.1109/VR.2019.8797960

Other publications inluding conference papers published in a domesting conference are shown in my web page[1].

Awards

13. Best Paper Award. IPSJ Interaction 2019 (a Domestic Conference in Japan). March, 2019.

14. Azusa Ono Memorial Award. Waseda University. March, 2019.

15. IPSJ Yamashita SIG Research Award. Information Processing Society of Japan (IPSJ). March, 2020.

16. Best Paper Award. JSSST WISS 2020 (a Domestic Conference in Japan). December, 2020.

17. Outstanding Student Paper Award. MobiQuitous 2021. (co authored paper, 1st author: Yutaro Yamanaka). November, 2021

[1] https://wotipati.github.io/