博士論文

Smartphone-based Mobility Aid System Architecture for the Visually Impaired

（スマートフォンを用いた視覚障碍者向け移動支援システムアーキテクチャに関する研究）

金 智恩

ABSTRACT

People who are blind and visually impaired face many challenges in achieving independent mobility and travel. Despite considerable research efforts into building effective mobility aids, no single solution has yet been widely accepted by the blind community, mainly because the existing systems cannot satisfy all of blind people's needs; they usually require additional dedicated hardware that is expensive and cumbersome, are difficult to learn and use, or are not aesthetically appealing. Meanwhile, the advent and rapid adoption of smartphones that are equipped with a rich set of sensors such as GPS, accelerometers, and digital compasses has opened up new opportunities for improving blind people's travel experiences in an unobtrusive, inexpensive, practical, and effective way.

This dissertation proposes a mobility aid system architecture for the blind that can be used to create easy-to-use and readily usable assistive technology solutions that are delivered via mainstream mobile devices such as smartphones without requiring that users carry additional special hardware. The architecture supports a set of major tasks involved in blind people's travel activities such as *navigation & orientation* both outdoors and indoors and *environmental access* to points-of-interest (POI), public transit, buildings, signage, and electronic kiosks. This architecture has been derived based on the experiences of iterative prototyping with blind users for smartphone-based applications that aid targeted mobility tasks. The proposed architecture consists of three layers: The user interface, functionality, and infrastructure. In particular, the user interface layer includes a set of interface design techniques that are intuitively usable by blind users for touch-based mobile devices.

To verify the effectiveness of the proposed architecture, I developed three prototypes: SaSYS, TalkingTransit, and StaNavi, which provide mobility aids for POI search, public transit use, and indoor navigation, respectively. Experimental evaluations involving a number of blind people show that each system was generally well accepted by its intended users. User studies for both SaSYS and TalkingTran-

sit were conducted in the laboratory, each with 11 blind participants, where the prototype was compared to an existing approach. A field test was conducted for StaNavi, in which eight blind participants inside Tokyo Station were asked to independently navigate to a given destination. This was the first thoroughly conducted indoor navigation study with blind people under real-world conditions.

This architecture's applicability was then demonstrated by reflecting on the three prototypes' development processes and system designs. In addition, the potential of this architecture to create further mobility aids was also demonstrated by conducting a rapid prototyping of TalkingBuilding that aids building use and by discussing the development of applications that aid outdoor navigation and provide access to signage and electronic kiosks. This shows that the proposed architecture can be applicable to all target mobility scenarios in this dissertation.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor, Prof. Ken Sakamura, for his guidance, support, and patience during the years I have been pursuing my Ph.D. degree. I consider myself extremely fortunate to have had the privilege to work under his supervision for all my years at the University of Tokyo. I also would like to thank my co-advisor Prof. Noboru Koshizuka for offering insightful comments and constructive suggestions which helped me in all the time of my research. This dissertation would not have been possible without their great support and encouragement.

I also would like to express my sincere gratitude to my other committee members: Prof. Jun Rekimoto, Prof. Akihiro Nakao, and Prof. Toru Ishikawa, for their valuable comments, suggestions, and discussions throughout my research.

I am greatly indebted to my mentor, Dr. Bessho Masahiro, for helping shepherd me through the daunting journey to completion of my Ph.D. He has always been available to discuss my research questions and ideas. Words cannot express my heartfelt thanks for his consistent encouragement, understanding, support, and advice that helped me overcome many (many!) difficult situations.

I also would like to thank all the past and present members of the Sakamura-Koshizuka lab, particularly Yuichi Bando, Manami Matsuura, Reza Aryaditya, Arto Nabito, Kounakis Giannis, Syafril Bandara, Hangli Ge, Xiaohui Peng, and Satoshi Asano for their assistance and support, as well as for their friendship in and out of the lab. I am especially grateful to Dr. Fahim Khan for his help in improving my writing skill by offering valuable feedback on my research papers.

I would like to extend my sincere gratitude to my blind and visually impaired research participants who generously gave their time, effort, and insights. Without their passionate participation and feedback, this accomplishment would not have been possible. A special acknowledgement goes to Akiyoshi Takamura who generously has shared his wide knowledge of assistive technologies for the blind and experiences of working as a teacher of visually impaired students. He has always provided me with invaluable advice and comments throughout this research.

I gratefully acknowledge the financial support of the Ministry of Education, Culture, Sports, Science and Technology of Japan over six years (2010–2016) through the MEXT program, without which I would not have had the opportunity to pursue higher studies in Japan. I am also grateful to the research funding received from the Global Creative Leader (GCL) program of the University of Tokyo.

And finally, I reserve my deepest gratitude to my parents and my brother, who have stood by my side at every stage of my life with their unconditional love and support. Thank you for believing in me even when I didn't believe in myself. I will always remain indebted to them.

This dissertation is dedicated to my parents.

Table of Contents

Chapter 1 Introduction **1**

1.1 Background 1

1.2 A Brief History of Assistive Technology for the Blind 3

1.3 The Travel Activities of the Blind 5

1.4 Research Objective and Overview 6

1.5 Contributions 8

1.6 Dissertation Outline 10

Chapter 2 Related Work **12**

2.1 Touch-based Mobile Device Accessibility 12

2.2 Aiding for POI Search 14

2.2.1 Map Exploration Systems 15

2.2.2 Location-aware Systems Supporting POI Discovery 17

2.3 Aiding for Public Transit Use 19

2.4 Aiding for Indoor Navigation 24

2.4.1 Indoor Positioning Techniques 24

2.4.2 Indoor Guidance Systems 25

2.5 Summary 30

Chapter 3 Smartphone-based Mobility Aid System Architecture for the Blind **31**

3.1 Requirements for Achieving Wide User Acceptance 32

3.2 Approaches 33

3.2.1 Use Mainstream Devices 33
3.2.2 Conduct Iterative Co-design with Blind users 34
3.2.3 Use Recent Technologies such as IoT and Open Data 35
3.3 The Architecture Overview 38
3.4 The User Interface Layer 40
3.4.1 Easy Navigation of Information Hierarchy 41
3.4.2 Areas-of-Interest Selection 45
3.5 The Functionality Layer 47
3.6 The Infrastructure Layer 49
3.7 Summary 50

Chapter 4 SaSYS: A Location-aware System for Aiding POI Search 51
4.1 Motivation 51
4.2 SaSYS System 53
4.2.1 Design Process 54
4.2.2 Scanning Mode 55
4.2.3 Contents Listening Mode 57
4.2.4 Implementation 59
4.3 Evaluation 62
4.3.1 Participants 62
4.3.2 Procedure 63
4.3.3 Results 66
4.4 Discussion 71
4.5 Summary 73

Chapter 5 TalkingTransit: A Location-aware System for Aiding Public Transit Use 74
5.1 Motivation 74
5.2 TalkingTransit System 76
5.2.1 System Overview 76

5.2.2 Requirement Analysis 78
5.2.3 User Interface 79
5.2.4 In-station Information 83
5.3 Evaluation 84
5.3.1 Participants 86
5.3.2 Procedure 86
5.3.3 Results 87
5.4 Discussion 91
5.5 Summary 92

Chapter 6 StaNavi: A Location-aware System for Aiding Indoor Navigation 93

6.1 Motivation 93
6.2 StaNavi System 95
6.2.1 Design Process 95
6.2.2 System Overview 96
6.2.3 Information Provided for Easy Navigation 97
6.2.4 User Interface 103
6.2.5 Implementation 107
6.3 Evaluation 109
6.3.1 Participants 109
6.3.2 Procedure 109
6.3.3 Results 113
6.4 Discussion 116
6.5 Summary 118

Chapter 7 Discussion 119

7.1 Reflection on the Implemented Scenarios 119
7.1.1 Aiding for POI Search 120
7.1.2 Aiding for Public Transit Use 123

7.1.3 Aiding for Indoor Navigation 126
7.2 Rapid Prototyping Scenario: Access within Buildings 129
7.2.1 System Overview 130
7.2.2 User Interface 132
7.2.3 Reflection on the Effectiveness of the Architecture 136
7.3 Other Scenarios 136
7.3.1 Outdoor Navigation 136
7.3.2 Access to Street Signage 137
7.3.3 Access to Electronic Kiosks 138
7.4 Summary 139

Chapter 8 Conclusion 140

8.1 Reflections and Insights 141
8.2 Future Work 143

List of Figures

1.1 Tasks within the travel activity by the blind. 5

2.1 User interacting with a smartphone using a Braille display 13

3.1 The iterative design process used in the development of mobility aid systems. 34

3.2 The architecture of a smartphone-based system that assists the mobility of blind people. 39

3.3 An example of fast navigation for a contacts list. 43

3.4 The basic idea of swipe-to-search. 47

4.1 Blind person who enjoys traveling by exploring surrounding POIs. . 52

4.2 Participant searching for POIs using SaSYS 54

4.3 Examples of swipe-to-search 56

4.4 Participants performing TTS controls 58

4.5 The gesture set used for TTS controls. 59

4.6 A screenshot after performing swipe-to-search 60

4.7 Examples of contextual cues 61

4.8 An example of performing Pointer Test 63

4.9 Participant conducting a given task 64

4.10 Participants' preference of the number of areas they would like the search space to be divided into. 68

5.1 TalkingTransit system architecture. 77

5.2 Screen transition flow and performed gestures of TalkingTransit . . 80
5.3 Users interacting with TalkingTransit 81
5.4 Participant receiving information from a BLE marker. 84
5.5 The mobile website developed for the user study 85
5.6 Average task completion time in seconds for both systems 88

6.1 StaNavi system architecture. 96
6.2 Areas defined in StaNavi and some points-of-reference. 100
6.3 Examples of navigation cues provided by StaNavi. 101
6.4 An example of route overview information 102
6.5 Screen transition flow and performed gestures of StaNavi. 104
6.6 Participant interacting with StaNavi: Determining directions 106
6.7 Flowchart of StaNavi's navigation process. 108
6.8 Maps showing given destinations and the routes 110
6.9 A neck strap and speaker used in the experiment. 111
6.10 Participants interacting with StaNavi: Listening to route instructions 112

7.1 The iterative design process of the SaSYS system. 121
7.2 The SaSYS system from the proposed architecture perspective . . . 122
7.3 The iterative design process of the TalkingTransit system. 124
7.4 The TalkingTransit system from the proposed architecture perspective 125
7.5 The iterative design process of StaNavi system. 127
7.6 The StaNavi system from the proposed architecture perspective . . 128
7.7 The TalkingBuilding system from the proposed architecture perspective 131
7.8 A mobile application using Smart Building API. 133
7.9 Screen transition flow and the performed gestures of TalkingBuilding. 134

List of Tables

2.1 The user interfaces adopted by existing public transit information systems for the blind. 24

3.1 The functionalities provided by the implemented prototype systems. 48

4.1 The ratings on Ease-of-Use and Usefulness for each interaction method 66

4.2 Results of questionnaire on usefulness of TTS control functionality . 69

5.1 Gestures and corresponding auditory feedback. 83

5.2 NASA-TLX WWL scores for each task 89

5.3 Questionnaire results on the usefulness of both systems 89

5.4 User's preference of both systems for each task 90

6.1 Overview of the information provided by StaNavi. 99

6.2 Data collected for each route. 113

6.3 Questionnaire results on overall system effectiveness. 115

6.4 The ratings on Usefulness and Ease-of-Use for each functionality . . 116

Chapter 1

Introduction

1.1 Background

Mobility, which can be defined as "the ability to move independently from one point to another [1]," is an integral part of everyday life—a key component of engagement in major life activities such as caring for oneself, eating, learning, and working[1]. The ability to get around is something that most people take for granted. However, people with disabilities have difficulty traveling independently, resulting in limited access to healthcare, education, employment, shopping, recreation, and public services. This may lead to them suffering from mobility-related social exclusion, defined by Kenyon et al. [2] as:

> "*The process by which people are prevented from participating in the economic, political and social life of the community because of reduced accessibility to opportunities, services and social networks, due in whole or in part to insufficient mobility in a society and environment built around the assumption of high mobility*" (2002, pp. 210–211).

[1]Americans with Disabilities Act of 1990, U.S. Code, Title 42, Chapter 126, Section 12102, Section (2)(A).

Visual impairment and blindness are disabilities that significantly affect personal mobility. Vision, which constitutes approximately 70% of the body's sensory receptors [3], is the primary sensory system with which humans navigate given environments [4]. It allows people on the move to gather information such as direction, distance, speed, and the general layout of spaces (placement of landmarks and obstacles). Furthermore, pre-trip planning activities may also require vision to access websites or print materials such as maps, newspapers, brochures, and guidebooks to obtain information on routes, weather conditions, and public transit such as vehicle schedules, fares, and real-time service status. Indeed, people with visual impairment, who have the same right to "personal mobility" as people without disabilities[2], are often at a severe disadvantage regarding the achievement of free movement. Thus, visually impaired individuals require mobility aids to lead autonomous and independent lives and to participate in society.

According to the World Health Organization, there are an estimated 285 million visually impaired people worldwide as of 2016, of whom 39 million are blind and 246 million have moderate to severe visual impairment; the number of people with vision loss will increase as the population ages in many countries [5]. For example, the prevalence of vision loss in the United States (U.S.) is expected to double by 2030 due to the rapid growth of the elderly proportion [6]. Vision loss and blindness not only negatively impact the quality of life of this population, but also impose a considerable economic burden on society as a whole [7]. For example, the economic cost of visual impairment was an estimated $15.8 billion in Canada in 2007 [8]; $72.8 billion in Japan in 2007 [9]; and $139 billion in the U.S. in 2013 [10]. This emphasizes the importance of facilitating access to and use of affordable and cost-effective mobility aid services by people with visual impairments to help alleviate the burden on individuals and society.

[2] Article 20 of the UN Convention on the Rights of Persons with Disabilities

1.2 A Brief History of Assistive Technology for the Blind

Assistive technology (AT) has been an important means of helping blind people[3] perform daily life activities such as shopping, working, paying bills, and taking medicine. It may range from relatively low-tech tools such as eyeglasses and white canes to high-tech tools such as electronic devices. The widespread use of personal computers in the 1980s brought advancements in computer-based AT. For example, screen readers converted the contents on a computer screen into a form that was accessible to blind people such as synthesized speech. In addition to screen readers, AT including screen magnifiers, optical character recognition software, Braille displays, and money readers have enhanced blind people's ability to perform daily life tasks by allowing them to use computers and access printed material.

In particular, considerable research effort has been devoted to improving the accessibility of the Web, as it has become an increasingly popular means of information delivery platform. Text-based Web pages were made accessible to blind people using a screen reading technology in the late 1990s [11]. Web browsers that help blind users access multimedia content have been developed in the late 2000s [12], and more recently, systems such as Sasayaki [13] have been proposed to enhance their Web browsing experience in an unobtrusive and intelligent manner.

Meanwhile, with regard to mobility support, the traditional white cane is the blind community's most commonly adopted tool; it is simple, inexpensive, yet highly effective. In the late 1990s and early 2000s, the widespread adoption of Global Positioning System (GPS) technology, coupled with advances in wireless technology and handheld device hardware such as mobile phones, tablet PCs, and personal digital assistants (PDAs), has expanded the possibility of developing mobility aids for blind people.

Many research prototypes using GPS-based positioning have been proposed to guide blind travelers in outdoor environments [14, 15, 16]. Several commercial navi-

[3]For the sake of simplicity, hereafter I use the expression "*blind people*" to refer to both the blind and those with a visual impairment, unless otherwise stated.

gation products have also been developed, such as the BrailleNote GPS and Trekker from Humanware[4] and the LoadStone GPS[5]. Some other solutions have been developed to assist with indoor navigation; these are typically based on technologies such as RFID, infrared, and computer vision (a discussion of previous systems and their effectiveness can be found in Chapter 2). Despite many research efforts in this area, no single solution has yet been widely accepted by the blind community, mainly because the existing systems cannot satisfy all of blind people's needs; they usually require additional dedicated hardware that is expensive and cumbersome, are hard to learn and use, or are not aesthetically appealing [17, 18]. Consequently, the white cane remains the primary mobility aid for blind travelers [19].

Nevertheless, the advent and rapid adoption of smartphones[6] in the early 2010s, combined with fast and reliable network connectivity, has opened up new opportunities for improving blind people's travel experiences in unobtrusive, inexpensive, practical, and effective ways. Smartphones are becoming increasingly popular among the blind community [20] thanks to built-in accessibility features including screen readers such as Apple's VoiceOver [21] and Google's TalkBack [22]. Moreover, important emerging trends in the field of information and communication technology (ICT) such as the Internet of Things (IoT) and open data [23] (as discussed further in Chapter 3) also have great potential to create useful and cost-effective applications for the blind [24]. Thus, taking advantage of these emerging technologies, this dissertation focuses on facilitating the use of smartphones as readily accessible, affordable, and readily usable travel aids for blind people across all possible mobility challenges.

[4] `http://www.humanware.com/`

[5] `http://www.loadstone-gps.com/`

[6] The Oxford English Dictionary defines a smartphone as "a mobile phone capable of running general-purpose computer applications, now typically with a touchscreen interface, camera, and Internet access." It generally has advanced computing capability and is equipped with a rich set of sensors such as GPS, Wi-Fi, Bluetooth, accelerometers, and digital compasses.

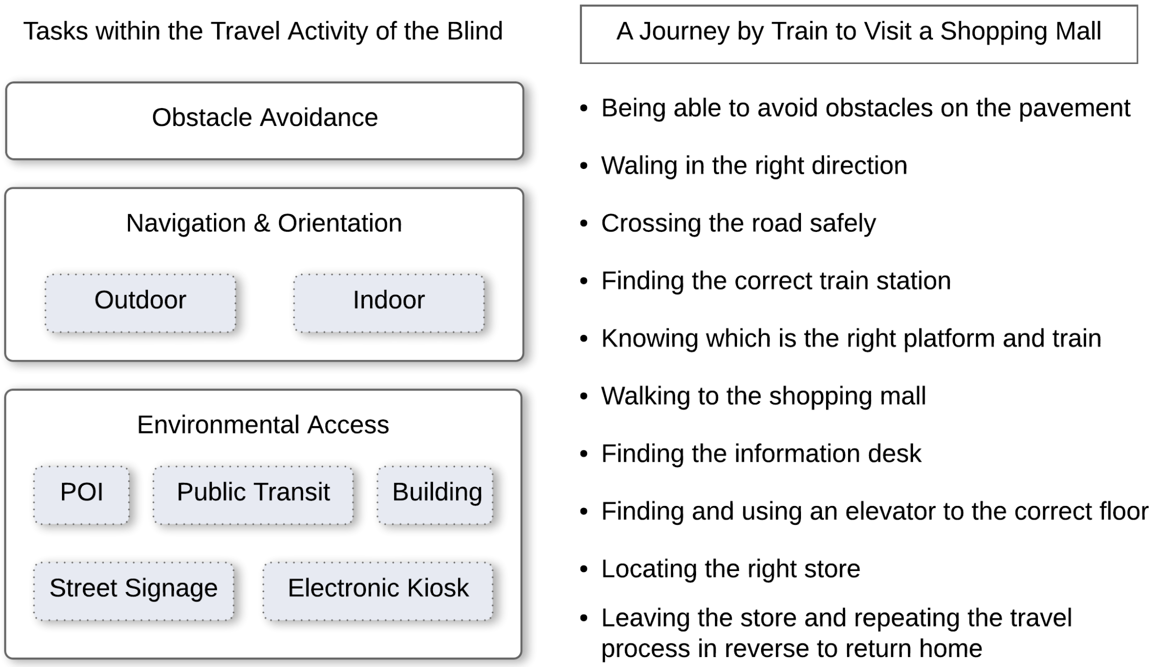


Figure 1.1: Tasks within the travel activity by the blind.

1.3 The Travel Activities of the Blind

Before describing the scope of this dissertation, it seems worthwhile to note the different types of activity involved in achieving independent mobility. Mobility as defined by Foulke [25] is "the ability to travel safely, comfortably, gracefully, and independently." In Brambring's model [26], the mobility of the blind involves two types of activity: the perception of objects and the process of orientation. Mobility can also be interpreted as activity that includes a set of capabilities: space perception, orientation, wayfinding, navigation, and obstacle avoidance [19]. Meanwhile, travel activities can be interpreted as a wider concept that encompasses both mobility and environmental access that involves issues related to the ability to obtain necessary information while traveling; for example, access to street signs, information about location, public transport information, etc. [17]. Although there were slight differences between classification models, they are basically similar.

Thus, I classified the tasks involved in travel activity into three categories based on existing models [17, 19, 27]: Obstacle avoidance, navigation & orientation, and environmental access. Navigation & orientation can be divided into outdoor and indoor tasks. Tasks related to environmental access include access to points-of-interest (POI), public transit, buildings, street signs, and electronic kiosks. Figure 1.1 shows this classification with an example based on Hersh and Johnson (2010, p. 168) describing the tasks involved in a train journey across a city to visit a shopping mall. As the example shows, blind travelers may need to perform different types of task to make even a relatively simple trip, indicating the importance of supporting their ability not only to orient themselves and navigate to a waypoint but also to access surrounding environmental information. I refer to the tasks shown in Figure 1.1 as the range of tasks that mobility aid can support in this dissertation, and use the terms "mobility aid" and "travel aid" interchangeably.

1.4 Research Objective and Overview

The goal of this dissertation is to propose a mobility aid system architecture for blind people that supports a range of travel activities related to *navigation & orientation* and *environmental access*, using readily available mobile devices such as smartphones without necessitating any specialized hardware. The architecture aims to provide a practical working solution in the context of actual use in everyday life.

The research adopts an iterative co-design approach to achieve this. Specifically, I have conducted iterative prototyping with blind users for smartphone-based applications that aid target travel activities, coupled with participant observation and interviews to identify user needs and establish requirements for facilitating travel aids that have a high possibility of gaining wide user acceptance; to identify possible approaches to respond to established requirements; and to provide a useful basis for building the architecture.

To verify the effectiveness of the proposed architecture, I have developed three travel aid systems, particularly for helping blind people perform the following tasks:

- Point-of-interest search (Chapter 4).
- Public transit use (Chapter 5).
- Indoor navigation (Chapter 6).

I tackled the above mobility challenges for the following reasons. Regarding the tasks in *environmental access*, I first focused on aiding POI search, recognizing the strong need for exploring and learning about one's environment that has not been fully satisfied by existing guidance systems. I then focused on improving public transit accessibility because public transit systems play a vital role in transportation within urban areas. Regarding the tasks in *navigation* & *orientation*, outdoor navigation has been well explored and GPS-based commercial smartphone applications such as BlindSquare [28] and Ariadne GPS [29] have been deployed for blind users. Meanwhile, indoor navigation remains largely unsolved [30] due to the unavailability of GPS signals in indoor environments; therefore, I addressed this challenge.

These prototypes were evaluated with user studies. The user studies for the POI search and public transit information systems were conducted in the laboratory, each with 11 blind participants, where the prototype was compared to an existing approach. A field test was conducted for the indoor navigation system with eight blind participants inside Tokyo Station, a large and complex space densely crowded with people, where the participants were asked to independently navigate to a given destination four times (the total shortest path length was approximately 600 m) under daily routine conditions. This was, to the best of my knowledge, the first thoroughly conducted user test that presented clear evidence from real-world situations that the effectiveness of using smartphones and readily available positioning techniques such as Bluetooth Low Energy (BLE) technology for facilitating independent indoor navigation for blind people.

Then, I demonstrated the applicability of the proposed architecture by reflecting on the development processes and system designs of the three prototypes that were developed and tested with blind users. Moreover, I conducted rapid prototyping to demonstrate how this architecture can be applied in developing applications for other travel aid scenarios. Specifically, I developed a smart building application that considered the possibilities of enhancing *environmental access within buildings* for blind people. In addition, how the proposed architecture can be used to support other travel activities such as *accessing street signage* and *electronic kiosks* and *navigating outdoors* was discussed.

1.5 Contributions

The main contribution of this dissertation is to propose a new technological solution, a travel aid system architecture that helps blind people perform a set of mobility tasks, which is readily usable via a modern mobile device such as a smartphone without the need for users to carry additional special hardware. This research shows how an effective yet unobtrusive mobility aid system that is well-accepted by blind individuals can be developed, and can contribute to the field of assistive technology in which considerable efforts have been made to enhance the independent mobility of blind people but no single solution has achieved wide adoption. The specific contributions are summarized below.

Prototype Systems

Low-cost and real-time travel aid systems that can be used from commercial off-the-shelf mobile devices:

- SaSYS, a location-aware system that enables users to search for surrounding POIs and detailed information about them (Chapter 4).
- TalkingTransit, a location-aware system that enables users to obtain real-time service status and timetables of public transit in Tokyo (Chapter 5).
- StaNavi, a navigation system that provides turn-by-turn voice directions inside a large train station using BLE technology (Chapter 6).

- TalkingBuilding, a location-aware system that helps blind people access within buildings and use their facilities (Chapter 7).

Interface Design Techniques

The unique techniques in designing user interfaces for touch-based mobile devices that are intuitively usable by and specialized for blind users:

- Swipe-to-Search, a swipe gesture-based interaction that allows users to retrieve relevant information on a desired direction by performing a swipe gesture on the touchscreen in that direction.
- Fast Navigation, a multi-touch interaction that allows users to rapidly locate a desired item from a long list of items.
- Shake-to-Main, a shake gesture-based interaction that allows users to directly return to the main screen or another important navigation point of applications.
- Hierarchical page structure, which allows users to easily make menu selections and navigate applications.

User Study Results

- Comparison evaluation of Swipe-to-Search, a swipe gesture-based interaction, and Pointing-to-Search, a pointing gesture-based interaction, for areas-of-interest selection in POI retrieval (Chapter 4).
- Comparison evaluation of TalkingTransit and searching mobile websites using VoiceOver, a gesture-based screen reader, to obtain public transit information (Chapter 5).
- Performance data of an indoor navigation task using StaNavi inside Tokyo Station, which, to the best of my knowledge, is the first thoroughly conducted field test under real-world conditions in the literature (Chapter 6).

1.6 Dissertation Outline

The remainder of this dissertation is organized as follows:

Related Work Chapter 2	Discusses related work in the domain of mobile assistive technology to support the independent travel for the blind, including accessible interaction techniques on touchscreen-based mobile devices and prior research efforts related to the mobility tasks: POI search, public transit use, and indoor navigation that this dissertation mainly tackles.
Architecture for Mobility Aids Chapter 3	Proposes a smartphone-based mobility aid system architecture for the blind that supports a set of mobility tasks related to navigation & orientation and environmental access.
Prototypes Chapters 4, 5, and 6	Describes systems that I have designed and developed to verify the effectiveness of the proposed architecture. Chapter 4 introduces SaSYS, a location-aware system for aiding POI searches. Chapter 5 introduces TalkingTransit, a location-aware system for aiding public transit use. Chapter 6 introduces StaNavi, a location-aware system for aiding indoor navigation.

Discussion Chapter 7	Demonstrates the applicability of the proposed architecture by reflecting on the development processes and system designs of the three prototypes, by conducting a rapid prototyping of TalkingBuilding that aids access within buildings, and by discussing the development of applications that aid other mobility tasks such as navigating outdoors and gaining access to street signage and electronic kiosks.
Conclusion Chapter 8	Draws conclusions based on the findings of this work. This chapter also discusses the reflections and insights that I gained through this work, and future research directions.

Chapter 2

Related Work

This chapter summarizes previous work in the domain of mobile assistive technology to support the independent travel for individuals with visual impairment. As this research aims to facilitate the use of mainstream mobile devices as useful travel aids, Section 2.1 initially describes currently possible ways and strategies for blind people to use mobile devices such as smartphones and then reviews accessible interaction techniques on touchscreen-based mobile devices. Sections 2.2, 2.3, and 2.4 respectively discuss prior research efforts that mainly focus on target mobility tasks of this dissertation: POI search, public transit use, and indoor navigation.

2.1 Touch-based Mobile Device Accessibility

As handheld devices such as smartphones have gained widespread popularity, the number of blind people who use smartphones in daily activities has been increasingly growing [20]. Although the use of modern mobile devices can be challenging to blind people due to the prevalence of touchscreens in such devices, they can use smartphones with built-in screen reading technologies such as Apple's VoiceOver [21] and Android's TalkBack [22]. In particular, VoiceOver is one of the most popular screen reading technologies among the blind community [31]. It helps blind users navigate touchscreens without seeing it. For example, by touching or swiping, VoiceOver will speak what is on the screen without activating what has been

Figure 2.1: User interacting with a smartphone using a Braille display[7].

touched; by double tapping in anywhere on the screen, whatever VoiceOver last said will be activated. Another option for blind people to access touchscreen-based mobile devices is to use external equipment such as a refreshable Braille display and a Bluetooth keyboard. As shown in Figure 2.1, the Braille display converts information on the screen into the form of raised dots that represent Braille characters, allowing the blind user to read what is on the screen.

Meanwhile, many research efforts have been investigating interaction techniques to make touch-based mobile devices accessible to people with visual impairment. For example, Slide Rule [32] uses a set of multi-touch gestures for browsing menu and selecting an item to use several applications. earPod [33] allows eyes-free menu selection by utilizing slide gestures on a circular touchpad and auditory feedback. Verma et al. [34] proposed "Quick Scrolling" that enables blind users to quickly find a desired item from a long list with multiple fields. It vertically splits the

[7] Author's screenshot provided under the Creative Commons Attribution (CC BY) license. Source: Christophe Tacquet (2016). "Braillant BI 32 Humanware Canada. Plage braille." Retrieved from `https://www.youtube.com/watch?v=vPe7GW9mvUI` (Accessed 20 Aug 2016).

screen, and scrolling a specific part of the screen only reads aloud a specific field of the items; for example, a user can listen to timestamps only from a list of e-mail items that consist of fields including a timestamp, a sender name, and a subject, by scrolling the leftmost part of the screen. McGookin et al. [35] used simple swipe and tap gestures for a touchscreen-based MP3 player. Sánchez and Aguayo [36] divided a touchscreen into nine zones and used each zone as a virtual button for their mobile messenger application. Although many approaches have been made using touch gestures, some leverage built-in sensors in smartphones such as accelerometers and gyroscopes. One notable example is Virtual Shelves [37] that allows blind users to select an application shortcut by pointing the device in mid-air as if the items were located in front of them.

Text-based contents are getting increasingly accessible to the user, but visual-based contents (e.g., maps) still remain inaccessible [38]. While existing interaction techniques cover several useful applications such as a mobile messenger, address book, e-mail, and mobile player [39], little attention has been given to gesture-based interaction techniques for map-based applications that are often used in traveling (e.g., Google Maps, Foursquare, etc.). The swipe-to-search method that I designed in Chapter 4 has potential to be applied to these map-based applications for visually impaired users.

2.2 Aiding for POI Search

Since the key to creating a successful POI search system for blind people is to provide an interaction method for interacting with geo-located data in a simple and intuitive manner [40], the survey in this section largely focuses on IT-assisted interaction techniques that support exploratory traveling for blind travelers. The section discusses two major areas of work that give an opportunity for blind users to explore and learn about their surroundings: (a) map exploration systems that enable users to learn spatial and geographic information, and (b) location-aware systems that provide POI information in situ while on the move.

2.2.1 Map Exploration Systems

Although the literature review basically considers mobile device-based assistive technologies, some desktop-based systems are also investigated in terms of the way they deliver information about environments.

The Blind Audio Tactile Mapping System (BATS) [41]

In BATS, a user can browse a map with a tactile input device such as a mouse, gamepad, or joystick. As the user moves a cursor on a map, relevant geographic information is provided through the two types of spatial sound: Auditory icons and TTS that respectively deliver region types (e.g., city, forest, and river) and names. The sense of distance to a POI is represented by the volume level of the sounds. It also utilizes a directional force feedback generated by an input device to alert POIs.

Auditory Map [42]

Pielot et al. [42] developed Auditory Map that delivers POI locations on a map using a tangible user interface and spatial sonification. It leverages computer vision technology to trace a tangible object regarded as the representation of the listener on a virtual map. When a user gets closer to the POIs by moving and rotating his "virtual listener," relevant spatial sounds are played (e.g., playing church bells near the church) so that he can localize the POIs on the map.

Timbremap [43]

Some map exploration systems allow blind users to grasp layout of streets. Timbremap uses the different types of sonification to guide a user's finger to trace routes on mobile touchscreens. The Timbremap interface has two modes to convey indoor layouts. First, the line hinting mode that helps the user's touch follow along path segments. If the user's finger deviates from a route, auditory feedback is provided to return the finger to the route; for example, when the finger deviates to the left, a beeping sound is played in the right ear to guide the user to the path. Second, area hinting mode that empty spaces adjacent to the paths offer different auditory icons so that the user can understand the outline of a geometry. Similarly, Poppinga et al. [44] examined the use of vibration and speech feedback on handheld touchscreens to make a road network accessible to blind people.

Ariadne GPS [29]

Ariadne GPS is an iOS application that supports map exploration and location-aware POI search with VoiceOver. In the map mode, when a user's finger touch a map element such as a street and a city, its name is read aloud. In addition, a variety of different auditory icons such as the sound of water (rivers, lakes, or oceans), birds (parks), and streets (steps, bikes, or cars) are provided to help users easily follow the map elements.

SpaceSense [45]

While map exploration systems listed above provide spatial information in detail, SpaceSense offers high-level spatial relationships between two locations. It uses custom spatial tactile feedback hardware consisting of nine vibration motors on a mobile device. The motors generate vibrations to a user's palm indicating locations of POIs. For example, if the user's destination is on the north-east side from the current position, vibrotactile feedback is given at the top right-hand corner of the device, showing the approximate direction to the destination. It also conveys the sense of distance by adjusting the vibration intensity at four different levels.

However, these techniques used in map exploration systems described in 2.2.1 require users to concentrate on sound and/or vibrotactile feedback, which makes them difficult to use while on the move. Furthermore, many of these systems need haptic input devices or dedicated hardware (costly and bothersome to carry). Thus, approaches in map exploration systems may not be useful for exploring urban environments.

2.2.2 Location-aware Systems Supporting POI Discovery

Bellotti et al. [46]

Bellotti et al. developed a tour guide system on PocketPC mobile devices using RFID technology for a large-scale flower exhibition. They deployed RFID tags to every POI in the exhibition so that a user with the handheld device equipped with a RFID reader can receive automatic notification of nearby POIs in the vicinity. The POI contents were particularly designed for blind visitors; the description included visual, olfactory, and tactile features of the POIs. The system was tested in the real-world context of use during the exhibition with more than 100 blind people.

The Chatty Environment [47]

Similar to the guide system described above, the Chatty Environment is also based on RFID technology to inform users about surrounding objects they are passing by. It utilizes different types of RFID tag according to objects. For example, for large and sparsely placed objects (e.g., classrooms and other facilities in a university building), it uses long-range tags to allow a user to detect the objects from a distance; however, for small and densely placed objects (e.g., items in a supermarket), it uses short-range tags to allow the user to only detect a desired item without having to listen to a very long list of items.

In Situ Audio Services (ISAS) [48]

ISAS is a GPS-based mobile application running on a smartphone platform that offers surrounding POI information via speech and spatial sounds. When walking down a street, information about the POIs in front of a user is automatically delivered in the form of spatialized audio. Although the user studies conducted in the real-world showed that ISAS was generally preferred by blind individuals, concentrating on spatial sounds to localize POIs while walking might affect safe mobility.

This type of tag-based push approach may be appropriate for providing POI information in usage scenarios such as museum and zoo tours or shoppings. Although systems like ISAS support POI search in urban environments with only smartphones, passively receiving information might be thought to be somewhat overwhelming and distracting [48]. Thus, it would be more appropriate to allow blind users to search for areas-of-interest at their own convenience.

Talking Points 3 (TP3) [49]

TP3 is a smartphone-based system that addresses this issue by providing both push- and pull-based information retrieval. The system provides three interaction modes for a better interaction with surrounding POIs. First is Automatic Notification; this automatically notifies a user of nearby POIs within 10 feet. Second is Nearby Locations; this allows the user to explicitly retrieve POIs within 30 feet. Third is Directional Finder; this allows the user to seek for distant POIs by pointing a handheld device. In this manner, TP3 enhances the chance to discover POIs and encourages the user to learn the surroundings in the process of wayfinding.

BlindSquare [28]

BlindSquare, currently one of the most popular iOS applications among the blind community, is a GPS-based POI search application using the data available on FourSquare and Open Street Map. It provides both push- and pull-based POI retrieval interactions; it automatically announces POIs within a radius that is adjustable by a user, and in the Look Around mode, it allows the user to retrieve POIs in a desired direction by pointing his smartphone.

The pointing-based mobile interaction has been recognized as an effective means of location selection [50, 51], and adopted by several exploration and navigation systems for sighted people as well [40, 52]. However, "pointing-to-search" technique may have problems especially for visually impaired users; for example, the user might lose his sense of direction and spatial orientation because he must make an

180 degree turn while searching backward. Furthermore, pointing allows the user to show only desired directions without any mechanism for specifying distance. Allowing the user to adjust the scope of search space could provide a better user experience when the user wants to search further or closer areas.

2.3 Aiding for Public Transit Use

This section investigates assistive systems for allowing blind people to access public transit information. First, the section describes each system in terms of its functionalities and technologies to obtain transit information. Recognizing the challenge of browsing public transit information that typically generates long lists of items (e.g., train timetables), this section then discusses the interaction techniques adopted by existing systems for blind users to support easy interaction.

HapticTransit [53]

Many systems have focused on getting the current position of vehicles. HapticTransit is a smartphone-based system that monitors the real-time location of a bus currently riding in and notify blind users when to get off via vibrotactile and visual feedback. The system uses a smartphone's GPS to determine the bus's current position and sends it to a server that calculates real-time bus arrivals. Regarding haptic feedback, it first generates a weak vibration at one stop before a desired stop, and then the vibration gets stronger as the bus reaches near the destination. In addition, the users can receive additional information about surrounding POIs, landmarks, and tourist attractions.

Ariadna [54]

Ariadna puts a device on a bus, which continuously transmits its GPS position to the central server using General Packet Radio Service protocol. The server then sends bus arrival notifications to mobile phones owned by visually impaired users waiting for the bus using the same protocol. The notifications are delivered via voice messages after a user input the code number of a bus stop using the mobile phone's keypad. Moreover, the system allows bus drivers to be informed when there is an awaiting blind rider at a bus stop so that proper support can be provided to such riders.

Mobi+ [55]

Zhou et al. [55] installed Mobi+ cards, a dedicated embedded system that consists of GPS, a RFID reader, a speaker, etc. both on buses and at bus stops. When a blind rider with a RFID tag indicating the types of impairment comes to the bus stop, the Mobi+ card recognizes him and provides a sound alarm of upcoming buses. The frequency of alarming sound gradually increases as the bus approaches. It also notifies the driver of the presence of riders with special needs such as wheelchair users and riders with baby strollers waiting at the next stop so that he can park the bus carefully and install the pallet properly.

The RAMPE Project [56]

The RAMPE is a PDA-based system that offers information to assist the use of buses and tramways. The RAMPE system consists of three components: Wi-Fi-enabled PDA, fixed base stations deployed at bus stops communicating with the user device using Wi-Fi, and a central server providing real-time public transit information to the base stations. When arriving at a certain bus stop, blind users can obtain information such as the line numbers and bus schedules via TTS. In particular, the system provides real-time notifications about changes in the service status (e.g., delays) and then requires the users to confirm the messages by pressing a button. In such a case, the keyboard turns to a "dynamic mode" that makes all buttons the "confirmation buttons" to allow the users to easily acknowledge the emergency notifications.

Although these systems are helpful while waiting or riding a bus, it may not be useful when the user wants to know vehicle schedule or real-time service status in advance for a pre-trip plan. Furthermore, they need special equipment or cause battery draining by continuous use of smartphone's sensors. I believe that it would be clearly beneficial for blind individuals to access rich transit information anytime and anywhere without any dedicated hardware.

GoBraille [57]

Some research efforts have been investigating the use of crowdsourcing in collecting contextual information around bus stops. GoBraille is a smartphone-based application that allows deaf-blind people to obtain real-time bus arrival information and crowdsourced information about landmarks of bus stops such as whether the stop has a shelter or a bench, and the description of how to get a given bus stop from the nearest intersection that is entered by deaf-blind riders waiting for a bus. The system makes use of smartphone's built-in GPS and compass to determine a user's current location. The user interacts with GoBraille using a Braille display connected to smartphones via the Wi-Fi connection. For example, once at a stop, the user first determines his stop by pointing the smartphone towards the street and pressing a button, inputs a route number using the Braille display, and then relevant information is delivered through his Braille display.

Bus Stop CSI (Crowdsourcing Streetview Inspections) [58]

Recognizing the importance of providing contextual cues to help blind travelers identify their bus stops, Hara et al. [58] explored the possibility of leveraging crowdsourcing to collect such data. They proposed a custom tool, called Bus Stop CSI, to gather landmarks around bus stops such as bus stop signs, shelters, benches, trash cans, newspaper boxes, and traffic poles. It allows online crowd workers (Amazon's Mechanical Turk3) to find target bus stops and label landmarks using Google Street View.

Although the concerns about using crowdsourcing such as getting incorrect and outdated information have been reported [57], it could be a feasible approach to collect information that needed by a special user group. Meanwhile, some systems have focused on providing not only useful functionalities but also user interfaces that can help blind people more easily interact with their systems.

Stopman [59]

Stopman is a mobile phone-based system that allows blind users to search for timetables of buses in the Tampere city area. A user basically interacts with Stopman by using voice commands with speech recognition that is performed on a server. When starting Stopman, the user can obtain a list of bus lines by telling a certain stop name to spoken dialogue prompts. In the smartphone-based version of the system, text messages are provided on the screen as well as the voice messages, and a built-in keypad or a joystick can be used as an additional input method.

Landau et al. [60]

Landau et al. developed an audio-haptic transit map presenting train routes of the New York City subway for blind riders. Raised-lines and small depressions indicate train lines and stations respectively. A blind rider can get relevant audio information by dropping a computerized pen into each depression. The pen tip is equipped with a camera so that it can read a certain pattern on the map's surface to find the location the pen pointing at.

Brunet et al. [61]

Brunet et al. designed vibrotactile patterns for a custom-made tactile wrist bracelet to assist public transit riders by informing events such as the arrival of vehicles, accidents, and discovery of POIs. The bracelet equipped with eight vibration motors generates a variety of different tactile patterns by adjusting the intensity and order of vibration. Although not explicitly designed for blind people, this haptic-based approach can be an effective method for such travelers as well because it supports eyes-free and ears-free interaction.

ABLE (Accessible Bussing through Location Estimation) Transit [62]

ABLE Transit is a mobile application running on smartphones that provides bus timetables of the nearest bus stop. It determines a user's current location using built-in GPS and then retrieves bus schedule data made publicly available by a local transport agency in the General Transit Format Specification (GTFS) format. When launching the application, users can obtain a list of bus arrival times by pressing the "refresh" button on the screen and access the list by using an accessibility tool such as a screen reader, a screen magnifier, and a Braille display.

Korbel et al. [63]

Korbel et al. developed Public Transport Explorer, a smartphone-based system that assists blind riders by providing real-time timetable information and a vehicle's current position retrieved from a local bus operator's system. The system adopts a dedicated hardware that is equipped with a Bluetooth sensor and a keyboard to allow blind users to easily interact with it while walking with a white cane.

Table 2.1 shows input and output methods used in systems described in this section. Regarding the "touch" input for HapticTransit [53], please note that there was no explicit description about the input method; however, it is running on Android mobile devices so that the input type was assumed to be touch. Although speech input can be beneficial for blind users, it may give a bad user experience in a noisy place and some users might feel uncomfortable with performing speech input in front of public. Furthermore, haptic- and Braille display-based interaction techniques require additional equipment that can be costly and cumbersome. TalkingTransit described in Chapter 5 provides simple yet effective gesture-based user interfaces co-designed with visually impaired people and does not require any dedicated hardware.

Table 2.1: The user interfaces adopted by existing public transit information systems for the blind.

Research	**User Interface**	
	Input	**Output**
HapticTransit [53]	Touch	Vibrotactile and Visual
Ariadna [54]	Keypad	Voice message
Mobi+ [55]	N/A	Voice message
RAMPE [56]	Keypad	TTS
GoBraille [57]	Braille display	
Stopman [59]	Speech	Voice message
Landau et al. [60]	Computerized pen	TTS
Brunet et al. [61]	N/A	Vibrotactile
ABLE Transit [62]	Screen magnifier, Screen reader, Braille display	
Korbel et al. [63]	Bluetooth keyboard	TTS

2.4 Aiding for Indoor Navigation

This section investigates positioning techniques that have been used for indoor environments. Then, the section explores existing indoor guidance systems designed for blind travelers.

2.4.1 Indoor Positioning Techniques

Many localization techniques have been developed for indoor spaces where GPS is ineffective, and can be roughly categorized into two categories: infrastructure-free and infrastructure-based.

Infrastructure-free approaches

A typical positioning technique that does not rely on infrastructure is dead-reckoning (DR). It determines a user's current position by calculating displacement from a previously known position utilizing sensors such as accelerometers and gyroscopes that generally come with smartphones. The main limitation of DR is the accumulation of errors as time passes. Another approach is vision-based localization using cameras, where the images are processed to match the pre-collected features of the environment or to recognize a specific visual pattern such as a QR code. However, this approach can be subject to lighting conditions, occlusions, etc., and its computational complexity may impose a burden on a mobile device with limited energy and computing power.

Infrastructure-based approaches

Infrastructure-based approaches deploy beacons in a physical space. The beacons typically emit wireless signals with unique identifiers that are associated with location information. Such beacon-based technologies include RFID, infrared (IR), Wi-Fi, and BLE. IR and RFID technologies can achieve high accuracy [64], but they require dedicated hardware to read tags. Wi-Fi is one of the most popular technologies for indoor localization using smartphones because Wi-Fi access points are already widely deployed in most buildings. However, a fingerprinting method that is often used in Wi-Fi positioning needs the construction of signal strength database, which is laborious and time-consuming. In this research, I employ BLE technology for my indoor guidance system, the most accurate localization method that is currently available on modern smartphones with a low cost ($5 per unit) and low power consumption (runs for years on a coin cell) [65].

2.4.2 Indoor Guidance Systems

Existing systems have employed one or more positioning techniques to assist blind travelers in indoor spaces.

Tsirmpas et al. [66]

Tsirmpas et al. installed passive RFID tags on a floor and developed a wearable module integrating a reader and ultrasonic finder to detect both tags and obstacles. Specifically, the RFID reader is placed on a user's leg. The wearable module communicates with a server that calculates the shortest path dynamically via Wi-Fi connection and provides voice instructions to a destination.

Alghamdi et al. [67]

Alghamdi et al. proposed a smartphone-based indoor navigation system using active RFID technology. Tags are deployed both at corners and known locations such as room doors, elevators, and near the entrances inside a building that act as reference points during a navigation process. A RFID reader is attached to a user's waist belt and communicates with a mobile software running on a smartphone via Bluetooth connection.

Faria et al. [68]

Some systems augment a standard white cane with sensors instead of requiring blind users to wear devices. Faria et al. attached a RFID reader to the cane so that a user could follow a pathway equipped with tags. The cane is also equipped with a vibration motor around the handle, which allows the user to be informed every time it detects a tag. When the user detects a tag that represents a POI, relevant audio information is provided. Moreover, they deployed a group of tags (a cluster of tags that is regarded as one tag) at each waypoint to provide the best possible localization accuracy.

Guerrero et al. [69]

Guerrero et al. used an IR LED mounted cane and Wii Remotes (Wiimotes) to localize and guide a user. A number of LEDs attached to the cane emit IR radially, and Wiimotes installed in an indoor environment detect the LEDs for determining the user's current position. A server that receives information about the cane position from the Wiimotes performs proper calculation and then sends navigational information to the user's smartphone.

Jain [70]

Jain exploited both IR and DR technologies to direct blind users inside a building aiming to use as few tags as possible. He developed a custom waist-worn module integrating an IR receiver and accelerometer sensor that respectively detect IR sensors installed on the wall and count the number of steps. By utilizing the accelerometer (DR), the distance between tags can be shortened (approximately from 8–10 m to 3 m) so that the number of needed tags can be reduced. The module then sends its location to a mobile phone that delivers verbal route instructions and vibrotactile feedback.

Headlock [71]

Vision-based approaches have recently attracted much attention [71, 72]. Headlock uses an optical head-mounted display to help blind users lock onto a landmark (e.g. a door) across large open spaces. An interaction with Headlock mainly consists of two phases. First, the discovery phase; a user can scan the surrounding by moving one's head to find a landmark. When detecting a landmark, Headlock switches to the guidance phase; the user can receive continuous feedback to veer correctly and reach the landmark. For the both phases, relevant information is delivered via auditory icons and TTS.

Joseph et al. [72]

Joseph et al. integrated a head-mounted camera and waist-worn Kinect to obtain a user environment. They also developed a custom haptic belt equipped with six vibration motors to generate directional information to help the user orient himself and head in the right direction. When the user detects POIs in a building, location-relevant information is also offered via voice messages, which guides him to a desired destination with vibrotactile feedback.

While these systems can achieve accurate navigation, they require dedicated hardware that could be costly and cumbersome to carry. My system uses smartphones without requiring any additional hardware for users. Meanwhile, some research efforts have investigated leveraging computing and the sensing capabilities of smartphones without additional external sensors.

Spindler et al. [73]

Spindler et al. developed a smartphone app that offers spatial information and directions in a large airport using Wi-Fi fingerprinting. Although Wi-Fi fingerprinting is one of the widely adopted technologies in the field of indoor positioning, in their field test, it sometimes provided insufficient accuracy to guide blind passengers. To help blind users compensate for the lack of precision positioning, it allows them to obtain not only route directions but also spatial descriptions of the surrounding environment that can be used as some cues to navigate the airport. However, the field test was conducted via the Wizard of Oz technique where an experimenter follows a user and updates contents according to the user's location.

Flores and Farcy [74]

Flores and Farcy used DR to infer a user's location through step counting. They utilized a gyroscope sensor to determine the user's (relative) orientation instead of using a compass sensor because it tends to be subject to magnetic inference in a place such as a train station. A barometer sensor is also used to detect floor changes. The user receives turn-by-turn voice directions to a destination. The application was tested in both a commercial center and a subway station. However, it suffered from errors in the estimated user's walking distance caused by different step lengths due to hesitation in walking.

SIMO (Simplified Information for Mobility and Orientation) [75]

SIMO fuses various sensors such as a barometer, Wi-Fi, an accelerometer, and a digital compass on a smartphone to locate a user. It provides route instructions in the form of letters and graphics with large-size and high-contrast for partially sighted users. Any textual instructions on the screen can be read aloud with accessibility options for blind users. However, the application asks the user to keep pointing the smartphone in the same direction as the user's body, which is impractical in some situations.

Seeing Eye Phone [76]

Seeing Eye Phone uses smartphone cameras to periodically capture images of an environment. It then sends the captured images to a server that holds a floor plan of a building and image features with their 3D locations. The server searches for matching images to determine a user's location and orientation, and calculates a route to a desired destination. The user can interact with the application by issuing predetermined voice commands, and the navigational instructions are delivered via TTS.

Navatar [77]

Navatar employs DR to locate users. Notably, the system enlists the help of users to correct DR errors; users are required to confirm and input typical landmarks in an indoor environment such as hallway intersections, doors, and floor transitions before receiving the next instruction. In addition, the system makes use of landmarks in the route directions (e.g., "*go straight until you find a hallway intersection*") rather than metrics.

Although these approaches can work well in relatively uncrowded and confined indoor areas, they may not be useful in large and complex spaces crowded with people. Furthermore, most systems in the literature have been tested in specific routes or places such as university campuses and have not been thoroughly conducted with intended users, lacking in clear evidence from real-life situations [48]. Consequently, no single solution has gained widespread acceptance [19]. StaNavi described in Chapter 6 provides several features that can be helpful when performing navigation under realistic conditions.

2.5 Summary

This chapter discussed prior work related to mobility aids for blind people in POI search, public transit use, and indoor navigation. Section 2.1 briefly investigated assistive technologies and research efforts that make touch-based mobile devices accessible to blind people. It highlights that textual contents can be made accessible, but non-textual contents (e.g., marker icons showing POIs on a map) may be difficult to access. Section 2.2 described electronic systems that aid POI search, mainly in terms of interaction methods to deliver surrounding information. A pointing-based interaction can be considered as an effective smartphone-only method to search POIs in a desired direction, but it may not be useful in some situations. Section 2.3 described assistive systems that aid public transit use. Although many systems provide transit information while waiting or riding vehicles, there are relatively few solutions made available for blind people to access real-time transit information anywhere, anytime. Furthermore, most previous systems have incorporated dedicated hardware or expensive equipment, mainly for the vehicle tracking and/or easy interaction. Section 2.4 described indoor guidance systems. While a number of systems allow blind people to navigate indoor environments, they either require special hardware, suffer from unsatisfactory positioning accuracy, or may not be useful in certain situations such as in crowded places and large open spaces.

To summarize, despite much research effort has been made in this area, no single solution has yet been widely accepted by the blind community, mainly because the existing systems cannot satisfy all of blind people's needs; they usually require additional dedicated hardware that is expensive and cumbersome, are difficult to learn and use, or are not aesthetically appealing. This motivated me to facilitate a smartphone-only solution that supports a wide range of mobility challenges of blind people while traveling. The following chapters present how this can be achieved.

Chapter 3

Smartphone-based Mobility Aid System Architecture for the Blind

This chapter describes the architecture for a mobility aid system for blind people that can be used from commercial off-the-shelf mobile devices such as smartphones without requiring additional hardware that aims to provide practical working solutions in the context of actual everyday use. The system supports a wide range of mobility challenges that blind individuals face while on the go. This has basically been derived based on the experiences of designing and developing smartphone-based applications in collaboration with blind users—the POI search system, public transit information system, and indoor navigation system, which are introduced in Chapters 4, 5, and 6 respectively. This chapter starts by identifying the requirements for facilitating mobility aids that have a high possibility of widespread user acceptance and large-scale deployment, followed by a description of possible approaches to respond to the requirements. The chapter then describes the mobility aid system architecture, specifically including a set of techniques for user interface design that are specialized for blind users' unique needs.

3.1 Requirements for Achieving Wide User Acceptance

Towards mobility aid systems that are actually used, the development of the three different prototypes described later in this dissertation has largely focused on how blind people experience the proposed systems alongside their technical aspects. Closely and continuously collaborating with a totally blind teacher at a special school for the visually impaired, the design process for all three prototypes has directly involved blind users. In particular, since the teacher is familiar with AT for the blind and has worked with visually impaired students for a long time, his comments and advice were very helpful in gaining a sense of how applications can be made readily usable even to blind people who have little or no experience with mobile devices and understanding the important factors for the successful integration of AT into their daily activities. Through these development experiences with the target users, the findings from the user studies of each prototype, and the literature review, I have identified the main requirements critical to facilitating the creation of a mobility aid system that is sufficiently usable in practice for real-world use and hence can be well received by the blind community. These requirements are:

- *A system should be socially acceptable.* Personal preferences in social contexts—mainly related to the concerns about how others perceive assistive devices and the person using them—have been recognized as a critical factor that influences the decision whether or not to use AT in daily activities [78, 79].
- *A system should provide interfaces and functionalities that are sufficiently tailored to blind people's needs.* Since their needs are unique and often hidden (they themselves may not even be aware of what they want unless they have first-hand experience of an application), if not properly addressed, accomplishing tasks may cause them considerable difficulty.
- *A system should be cost-effective and affordable.* More importantly, the system should ensure value for money from the users' perspective for the con-

tinued use of AT; i.e., "the advantage a device offers a user must outweigh the costs of using it or the device will most likely be discontinued [80]."

3.2 Approaches

The following approaches can basically respond to the requirements described in Section 3.1:

- Use mainstream devices and do not require any dedicated hardware.
- Conduct iterative co-design with blind users.
- Use recent technologies such as IoT and Open Data.

Each of these approaches is discussed below in detail.

3.2.1 Use Mainstream Devices and Do Not Require Any Dedicated Hardware

Employing a modern handheld device such as a smartphone, tablet, Apple iPod Touch without requiring any dedicated hardware or additional device is absolutely essential in terms of achieving social acceptability and cost-effectiveness. In fact, existing systems for blind people often involve dedicated hardware that is somewhat bulky, obtrusive, and can be socially awkward to carry [17], which may both entail additional costs and cause them to feel stigmatized and isolated, resulting in AT abandonment [81]. The result of empirical investigation with people with visual impairments showed that the use of mainstream devices would give them a feeling of belonging to their sighted peers; make them look more capable of doing things like everyone else; and look aesthetically acceptable [78], emphasizing the importance of using mainstream technologies for large-scale adoption. Moreover, exploiting mainstream mobile devices can also be beneficial from a cost-effectiveness perspective, because most such devices incorporate multiple wireless technologies (cellular networks, Wi-Fi, Bluetooth, NFC, etc.) and be equipped with a rich sensor set (GPS, camera, compass, gyroscope, accelerometer, etc.). Hence, it enables

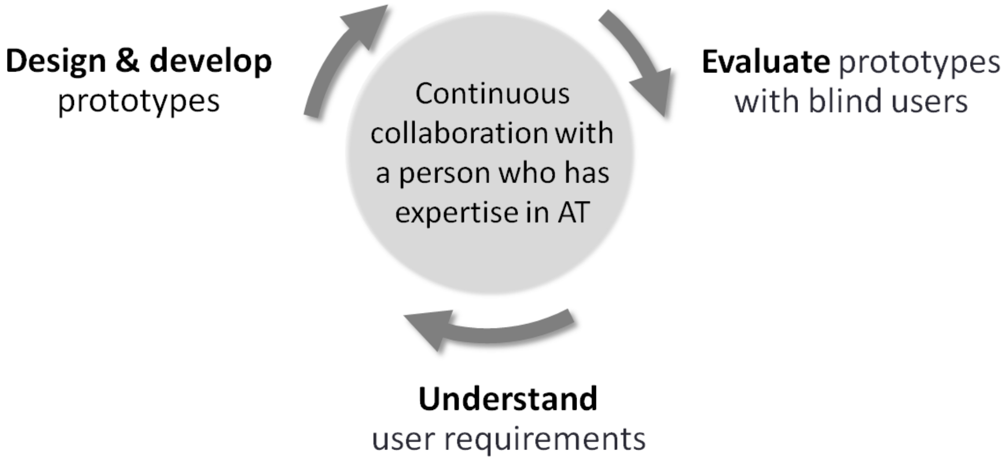


Figure 3.1: The iterative design process used in the development of mobility aid systems. Each iteration is developed via close collaboration with an expert on AT for the blind, and consists of three phases: 1) design and develop a prototype, 2) evaluate the prototype with blind users, and 3) understand and specify user requirements.

the development of a wide range of location-aware applications that could allow blind users to deal with various travel activities with a single device rather than necessitating carrying multiple gadgets.

3.2.2 Conduct Iterative Co-design with Blind users

When considering the use of mainstream mobile devices that are often equipped with touchscreens without physical buttons as end-user terminals, one challenging design issue that emerged is that interactions with such devices, which are visually demanding due to the nature of touchscreen manipulation, should be supported in a non-visual manner that might be unintuitive for sighted people. Given the situation, the direct involvement of real users during the development process and iterative design are the key to achieving highly usable systems [82, 83]. Conducting prototyping tests with blind users can resolve inherent uncertainties about whether blind users can interact with systems and how they perceive the provided user interfaces, and can identify potential needs and usability problems. Iterative prototyping and obtaining such user feedback can then progressively facilitate the development of interfaces and functionalities that are well suited to blind users. Figure 3.1 shows the iterative design process that I applied to the development of the prototypes introduced in Chapters 4–6. This design process generally complies with ISO standard 9241-210 that specifies the following four activities required for the human-centered design of any interactive system:

1. Understanding and specifying the context of use.
2. Specifying user requirements.
3. Producing design solutions.
4. Evaluating the design.

Meanwhile, it can be very helpful to involve people who have broad knowledge of and experience with AT for the blind as an advisor throughout the entire development process because their insights into AT may reduce the trial-and-error time of the design process and lead to better designs.

3.2.3 Use Recent Technologies such as IoT and Open Data

Exploiting contextual information has been recognized as a critical aspect in facilitating effective mobility aids that respond to the blind travelers' needs [15]. Context can be defined as "*any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves* [84]." For example, the context related to a blind person on the move may include the identification of the user and devices (e.g., nearby elevators in a building), location (e.g., GPS coordinates, heading direction, proximity to POI), date, time, environmental conditions (e.g., the current temperature, sounds, smell, lightning, surface textures), in addition to relevant information that can be derived from known contexts (e.g., a list of nearby restaurants and their detailed information such as menus, opening hours, and user reviews that can be retrieved based on the current location). Unsurprisingly, when users' mobile devices are allowed to access richer context data, blind users can gain a better understanding of their environments and more support and opportunities for independent mobility [85]. For example, blind travelers at a train station can receive automatic notification regarding the surrounding facilities, and when near a platform they can obtain information about the real-time arrivals and timetable for that platform.

The current trend towards the Internet-of-Things has great potential to generate rich context data, where the concept of IoT envisions that sensors, actuators, and network connectivity are embedded into physical things such as doors, windows, lights, home appliances, and other items, and then everyday things, both physical and virtual (e.g., multimedia contents and application software) things, can be connected to each other via a global network such as the Internet [86]. This connectivity allows large amounts of diverse context data to be gathered to the cloud from physical objects.

Moreover, the amount of data resources (virtual things) that have become publicly available and can be used to provide useful information to blind travelers has grown at an exponential rate over the last few years, mainly from two different sources. First, *user-generated content* from social media sites; the huge rise in popularity of social media such as social networks (e.g., Facebook) and microblogging (e.g., Twitter) sites has yielded an explosion of social data—for example, Twitter has attracted over 200 million monthly active users as of 2013, who generate over 500 million tweets per day [87]. In particular, social data such as local business reviews on Yelp, Foursquare, etc., and geo-tagged tweets seem to be helpful in keeping blind users informed of their surroundings. Second, *Open Data* refers to the release of proprietary data owned by major organizations in machine-readable formats under an open license that permits anyone to freely use, reuse, and redistribution for any purpose[8]. In particular, there has been a growing movement towards

[8] `http://opendefinition.org/guide/`

open data from governments and public sector organizations [88] that has opened up a vast amount of datasets on a wide range of categories that may include education, health, geospatial, and transportation through portals such as Data.gov in the US, Data.gov.uk in the UK, and Data.go.jp in Japan, which currently provide 184,055, 17,105, and 37,436 datasets respectively. According to DataPortals.org, there are 519 open data portals across all levels of government (national, state, and local) around the world as of July 2016, which promotes third parties to create innovative applications and services for their citizens. One important public sector data asset from a mobility support perspective is real-time public transit information that includes vehicle positions, service status, and timetables that are already being offered by a number of major cities around the world such as London[9], New York City[10], and Sydney[11].

More importantly, software interfaces that access, process, and retrieve the data resources mentioned above are also increasingly being made open to the public, often in the form of Web APIs that are available over the Internet [89]. For example, the ProgrammableWeb site, which is one of the most popular directories of publicly available Web APIs, lists more than 15,000 APIs as of July 2016, and covers a wide range of categories including social (e.g., Facebook, Twitter), mapping (e.g., Google Maps, Foursquare), localization (e.g., Yahoo Local Search, Yelp), and government. These open APIs are generally accessible using HTTP from various clients written in any language and running on any platform (e.g, mobile, desktop, and the Web) [90].

I believe that the use and integration of IoT resources such as sensor data, social data, open data, and location data that can represent user context, all of which are generated from "things," alongside Web APIs that efficiently access and control "things" can greatly expand the possibilities for the easy, fast, and cost-effective development of value-added IT services that benefit people with disabilities without

[9] `https://tfl.gov.uk/info-for/open-data-users/`

[10] `http://datamine.mta.info/`

[11] `http://www.transport.nsw.gov.au/opendata`

requiring any dedicated infrastructure [91].

3.3 The Architecture Overview

This section describes the architecture of a mobility aid system that can be accessed via off-the-shelf handheld devices such as smartphones without requiring additional dedicated hardware or expensive equipment such as a Braille keyboard. As illustrated in Figure 3.2, it consists of three layers:

- *The user interface layer* provides the eye-free capability for users to control the system's functionalities. It also includes some new interfaces that are specially designed for blind people and allow easy and effective interaction with the system.
- *The functionality layer* provides functionalities that can support a wide range of travel activities including obtaining knowledge of the surrounding POIs, using public transit systems, and navigation.
- *The infrastructure layer* allows the system to obtain the users' current location, retrieve context-relevant information that supports the needs of blind people on the move, and control surrounding devices if necessary.

The user interface and functionality layers run in a client application on mobile devices that are typically equipped with various positioning sensors (GPS, compass, etc.). Each of these layers is described in the next sections.

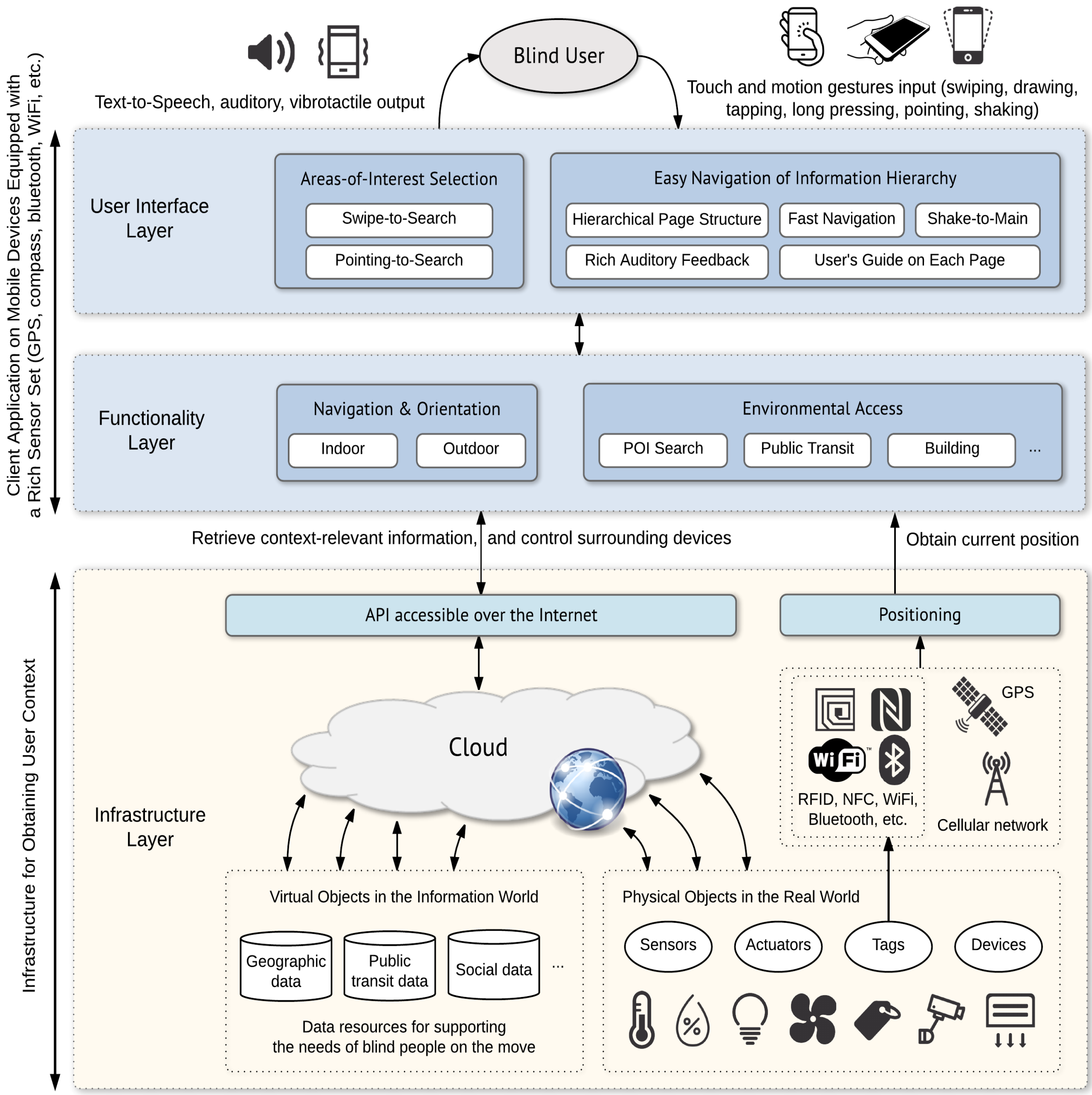


Figure 3.2: The architecture of a smartphone-based system that assists the mobility of blind people.

3.4 The User Interface Layer

User interface design plays a crucial role in developing easy-to-use and practical working mobility aid systems for blind people on mainstream devices. Based on the experiences of three case studies (described in the following chapters) in which smartphone applications are developed with blind people, the architecture provides practical techniques that can be utilized in user interface design to create applications on touchscreen-based mobile devices accessible to and usable by blind users without requiring specialized hardware. Taking advantage of modern mobile devices' multi-touch capabilities, built-in TTS engine, and built-in sensors that can be used to capture device movement (e.g., accelerometers and gyroscopes), the architecture uses the following as input and output in the mobile devices' user interfaces:

Input types. One important user requirement that has been identified through this work for the readily usable and accessible mobile applications by blind people is that inputs should be made with confidence and certainty. Recognizing this need, the architecture makes use of simple touch gestures such as swiping, tapping, and long presses. Users can perform gestures anywhere on the device's touchscreen, thus allowing them to use an application easily without having to understand the screen's spatial (visual) layout. Two types of motion gesture are also used for certain tasks: Pointing the device towards a direction of interest and shaking the device. Although speech recognition could also be an effective form of input, the architecture does not employ this method because it has some usability issues—users might feel embarrassed about talking to their device in public and recognition accuracy may decrease in noisy environments such as train stations, crowded restaurants, and in moving vehicles.

Output types. Using voice messages generated by a TTS synthesizer as the main form of output, the architecture also utilizes additional non-speech auditory and vibrotactile feedback cues to emphasize the information being delivered or to improve user interaction with applications.

3.4.1 Interface Design Techniques That Are Useful in the Easy Navigation of Applications' Information Hierarchy

The following techniques can be used to design user interfaces that not only support effective navigation of a large amount of information for blind users but also help them feel in control and confident to browse a navigation hierarchy.

Hierarchical Page Structure

Blind users navigate applications strictly by the audio channel. Unlike visual output (e.g., a list of menus) that can be scanned quickly and consistently, speech output can be slow due to its necessarily serial nature [92]. Furthermore, audio information is transitory (temporal), so that users have to remember spoken items to make menu selections [93]. This can increase users' cognitive load due to their limited short-term memory capacity, which can hold only approximately seven (plus or minus two) items for a brief period of time [94]. Thus, the amount of information presented to users per screen should not be overwhelming, and the navigation path to a desired content page should be as simple and predictable as possible. A hierarchical structure seems well suited for providing such a navigation style, because it can minimize the number of choices that must be considered per screen [95] and can facilitate intuitive navigational flow by dividing information into meaningful categories and levels and then arranging them into a logical structure.

One important aspect to consider when designing a hierarchy of applications (defining hierarchical relationships between screens) is the balance between breadth and depth [94]. Although hierarchical menu design increases the navigational depth for applications with relatively simple structures, it can basically make it easier for blind users to make menu selections by presenting clear and concise alternatives on each screen. For example, the destination input of a navigation system for shopping malls can be performed more efficiently with a two-level list—categories (e.g., exits, restaurants, facilities, etc.) then a specific destination (e.g., women's toilet)—rather than with a single-level list of all destinations. However, the type of applications that require that users browse massive contents inherently gets quite deep. For example, when designing a train timetable application for Tokyo, which has more than 1,200 stations, searching for a station name would require that users make several iterative choices to narrow down the alternatives; this may generate a deep application structure that can increase the chance that users feel lost in the hierarchy [95]. The following four techniques can be particularly useful to help blind people use this kind of application.

Fast Navigation

This is a multi-touch interaction technique that can let users rapidly locate a desired item from a long list, such as contacts, song titles, and subway station names in a big city. It uses one- and two-finger swipe up and down gestures to navigate the list of items of the screen: The one-finger swipes pass items one by one, while the two-finger swipes move among specific items that indicate meaningful positions in the list. For example, when navigating a contacts list in Figure 3.3, users can scan the names in serial order using a one-finger swipe up and down, or can jump to the first item in the list that begins with the previous or next letter in the alphabet; for instance, if the user performs a two-finger swipe down while scanning the names starting with J, then "Kevin," which is the first item starting with K, will be read aloud. This allows users to quickly move to a part of the list that includes the item for which they are searching. Moreover, because this technique allows the efficient navigation of a long list of items on one screen, complex and deep application structures can be simplified by reducing the depth of the menu hierarchy. Without "fast navigation," a contacts list may require a two-level menu selection; for example, the first letter of a name in the alphabet list, then a desired name starting with the selected letter.

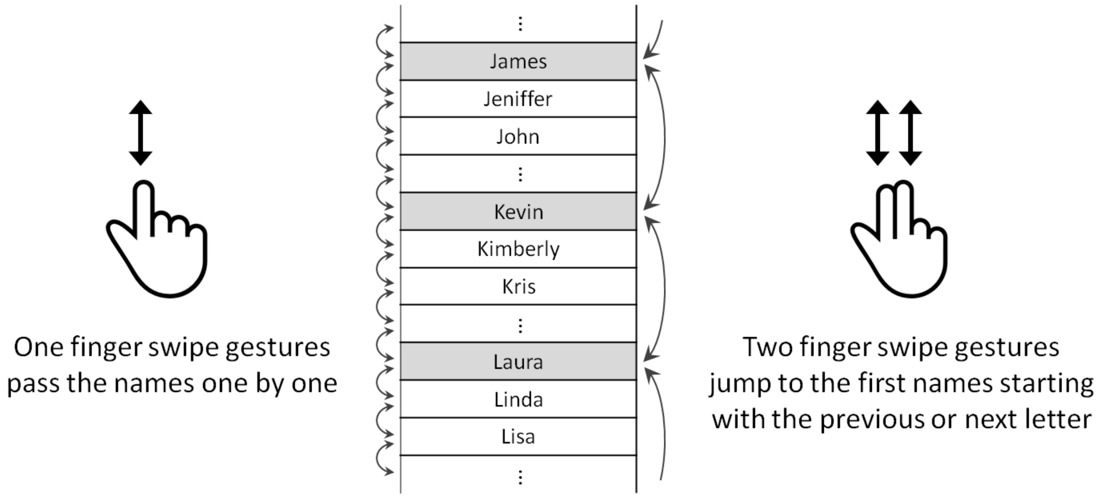


Figure 3.3: An example of fast navigation for a contacts list sorted in alphabetical order by first name.

Shake-to-Main

An interaction technique using a shake gesture can allow users to directly return to the main screen or applications' important navigation points. It provides a quick exit in the case of emergency situations such as when users mistakenly chose incorrect items or became confused about where they are within the navigation hierarchy. Users can also get easy access to the main page when they want to start over without having to go back through a series of screens. This technique can help users stay in control and feel confident and comfortable when using applications. Shaking gestures can be particularly useful for blind people in terms of allowing one-handed navigation. For example, users may navigate applications with a single hand by performing the following simple gestures: swiping up and down to browse menus, double-tapping to select one, and "shake-to-main." In this manner, it can respond to blind people's requirement for the one-handed use of applications when they are carrying a white cane such as navigation that is often used while walking.

Rich Auditory Feedback

Since touchscreens typically lack tactile landmarks such as physical buttons, blind users, especially those who are unfamiliar with touchscreen devices, may have difficulty performing touch inputs and feel uncertain about whether their input was successfully recognized. Thus, it is crucial to provide rich auditory feedback (non-speech sounds) for all touch and motion gestures to let users know that the applications have recognized their inputs as intended. An application should play different sound effects for each gesture so that it can help users intuitively identify which gesture they performed. Regarding the assignment of sound effects, using sounds that are somewhat associated with either gestures (e.g., a bounce sound for tapping, a whoosh sound for swiping, etc.) or actions (e.g., a page flip sound for navigating items, a trash tossing sound for deleting items, etc.) can help users learn and remember the necessary operations and actions.

User's Guide on Each Screen

When considering the situation of a blind person launching an application for the first time, they might have no idea how or where to start. Sighted users may intuitively guess how to use an application as soon as they see the screen because the visual interfaces for mobile applications often involve typical components such as buttons, sliders, toggle switches, etc. that are obvious in their use. However, blind users hardly get any initial context of what is on a screen such as an approximation for the number of items in a list, possible actions that can be performed, appropriate gestures to trigger actions, etc. Even if they received some training on the use of the application, it may be difficult to memorize all the key operations and actions in a short period of time. They can also forget how to use an interface if they have not used an application in a long time. Thus, blind users should be able to obtain information on the current screen and instructions on how to interact with their applications at any time. The following shows essential information to include in a user guide for each of an application's screens:

- Whether a screen transition occurred.
- Current location in an application.
- What users can or should do on the current screen.
- Summary of the screen's content.
- Key operations and actions that can be performed on this screen.

For example, a train time search application can provide the following user guidance when a user triggers a screen transition to the screen that presents a list of train stations: "You have just entered the nearby train stations list screen. There are three stations. Please navigate the list by swiping up and down, and make a selection by double tapping. You can also add a station to the favorites using a two-finger long press." This allows blind users, especially those who are not familiar with touchscreen mobile devices, to handle a complex navigation hierarchy, and helps them stay in the context of what they are doing.

3.4.2 Interaction Techniques for Areas-of-Interest Selection

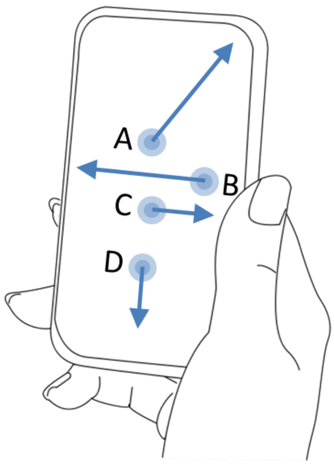
Location-based services (LBS) generally offer information associated with a geographical location such as latitude and longitude coordinates. Many applications typically use a map to allow users to obtain this kind of information, such as placing pushpins at the locations of nearby local businesses, drawing route lines, and displaying a bearing arrow, but map-based interfaces are visually demanding and hence largely inaccessible to people with visual impairments. Thus, it is important to provide blind users with techniques for interacting with geo-located information spaces so that they can easily and effectively perform specific tasks that are often required in LBS such as searching for surrounding POIs or determining which direction the user is facing. The following two techniques can be used for such tasks by allowing blind users to intuitively indicate areas-of-interest.

Pointing-to-Search

This is a pointing-based interaction that enables users to retrieve relevant information on a desired direction by pointing their mobile device in that direction. It utilizes the mobile device's built-in compass sensor to determine the direction of the users' device is facing. Since pointing gestures are intuitive and easy to perform, this method has been identified as a useful interaction technique for letting blind people indicate their areas-of-interest [96] and has been adopted in commercial apps [28] and research prototypes [49]. Nevertheless, application scenarios should be sufficiently taken into account before adopting this interaction technique. If an application requires users to perform a search while mainly facing forward (between 0 and 180 degrees), pointing-to-search would be appropriate; however, if users have to search backwards, this method would not be suitable because turning their entire body to point backwards may cause the loss of sense of direction.

Swipe-to-Search

Swipe gesture-based interactions allow users to retrieve relevant information on a desired direction by performing a swipe gesture on the touchscreen in that direction. For example, users can obtain information on their left by simply performing a swipe left gesture (Gesture B in Figure 3.4), and users can adjust the range of search space in accordance with the swiped distance. As shown in Figure 3.4, long and short swipes can indicate distant areas (Gestures A and B) and nearby areas (Gestures C and D) respectively. This is a new interaction technique I designed to allow blind people to intuitively and easily select areas-of-interest. It has been adopted by my POI search application as described in Chapter 4, and has been shown to be easy to perform and useful for blind users through a user study [96]. The main advantages of this method lie in the fact that it allows users to easily search all directions without turning and to simply adjust the search range. Moreover, it would be beneficial compared to pointing-to-search, in that swipe-to-search can be performed in a non-obtrusive manner (by just sliding one's finger on the



Performed Gesture	Indicated Area-of-Interest
Gesture A	Distant area on the front-right
Gesture B	Distant area on the left
Gesture C	Nearby area on the right
Gesture D	Nearby area on the back

Figure 3.4: The basic idea of swipe-to-search. Swipe gestures can be used to roughly specify the direction (relative to the direction the user is facing) and distance of areas-of-interest.

touchscreen), considering that blind people may not want to use a device or system that draws much attention [17].

3.5 The Functionality Layer

This layer provides a set of functionalities that allows blind people to accomplish a wide range of tasks involved in travel activities which can be mainly classified into two categories: Navigation & orientation and environmental access. In particular, recognizing the strong need of blind travelers to explore and learn about their environments, the functionality layer focuses on not only guiding them to a desired destination but also enhancing their ability to obtain the information they require while traveling. For example, Table 3.1 shows the details of what information users can obtain and what they can do using the functionalities that provided by the prototype systems that I have developed in this dissertation. All these functionalities are operable via the user interfaces described in the previous section.

Table 3.1: The functionalities provided by the implemented prototype systems.

Category		Provided Functionalities in Detail
Navigation & Orientation		– Obtain turn-by-turn instructions to a predetermined destination. – Obtain the directions that users are facing. – Obtain a route overview that provides a picture of the entire journey in advance.
Environmental Access	POI Search	– Obtain information about nearby local businesses and facilities such as contacts, locations, real-time events, promotions, and user reviews.
	Public Transit Use	– Obtain the arrival times for vehicles and a full timetable. – Check the real-time service status of public transit such as real-time updates on delays, service changes, and suspensions. – Obtain automatic notification of in-station information.
	Access within Buildings	– Monitor a room status such as lighting, temperature, and humidity. – Control in-building machines such as air conditioners and elevators. – Obtain automatic notification of in-building information.

3.6 The Infrastructure Layer

The infrastructure layer is responsible for obtaining user context (e.g., location, date, time, and any relevant information that can be used to support the needs of blind travelers) and providing it to the client application. Based on the notion of IoT described in Section 3.2, the infrastructure collects contextual information from both virtual and physical objects to the cloud, and offers APIs to access and control "objects" via the Internet (See Figure 3.2). From the client application's perspective, the infrastructure basically allows:

- *Positioning.* Tags that are densely deployed in urban spaces—which are generally based on wireless technologies such as RFID, NFC, Wi-Fi, and BLE and transmit a reference (a unique identifier) to the information associated with their specific location—can achieve seamless positioning across indoor and outdoor environments alongside GPS and cellular networks.
- *Retrieving context-relevant information.* Large and diverse data resources are gathered in the cloud from difference sources; for example, environmental sensors (e.g., temperature, humidity, light, and electricity), public sector (e.g., real-time schedules for public transit), private sector (e.g., geographical data), and end users of online services (e.g., restaurant reviews). Since such data can be accessed over the Internet, modern mobile devices that are generally Internet-enabled can easily retrieve information that can be helpful for the independent travel of blind people using a known context (e.g., location, date, and time) as a key.
- *Controlling surrounding devices.* Various types of devices including home appliances and in-building machines (e.g., refrigerators, air conditioners, elevators, surveillance cameras, and lighting) can send their status and receive commands over the Internet. Such capabilities allow client applications to remotely control surrounding devices according to user needs. For example, a blind person may want to check the current floor of the elevator and call it

when stood in front of it without having to find the call button panel on the wall.

Please note that this infrastructure is not specifically designed as a mobility aid system for blind people but rather for any applications and services that could benefit anyone's daily life.

3.7 Summary

This chapter described how a successful mobility aid system that supports a wide range of blind people's travel activities using readily available mobile devices without necessitating any specialized hardware can be developed by providing user requirements, approaches, and conceptual architecture, based on the experiences and findings of the studies described in Chapters 4–6. With approaches based on the use of mainstream devices, iterative co-design with blind users, and IoT technology, the proposed architecture of a mobility aid system consists of three layers: The user interface, functionality, and infrastructure. In particular, the user interface layer includes a set of interface design techniques that are intuitively usable by blind users for touch-based mobile devices such as smartphones. I believe that this chapter can serve as a comprehensive reference for a smartphone-based mobility aid system that is easy-to-use and practical for blind individuals.

Chapter 4

SaSYS: A Location-aware System for Aiding POI Search

The preceding chapter introduced the architecture of a smartphone-based mobility aid system derived from several prototypes that assist a variety of challenges faced by blind people on the move. This chapter[12] present SaSYS, the first prototype that addresses one of the most important needs in environmental access: Acquisition of POI information. It enables blind people to search for surrounding local businesses and facilities using simple swipe gestures on touchscreen.

4.1 Motivation

Traveling to unfamiliar places still remains one of the major challenges faced by people with visual impairments. Although many guidance systems providing turn-by-turn directions to predetermined destinations have been developed, such systems are inadequate in that they fail to address one of the most important navigational needs: exploring and learning one's surroundings [49, 97]. Blind people often get knowledge about what is around them by walking with sighted people. However, it is important to enhance their ability to explore surrounding points-of-interest for making their travels more independent and enjoyable as shown in Figure 4.1. This

[12]Portions of this chapter were previously published in [96]

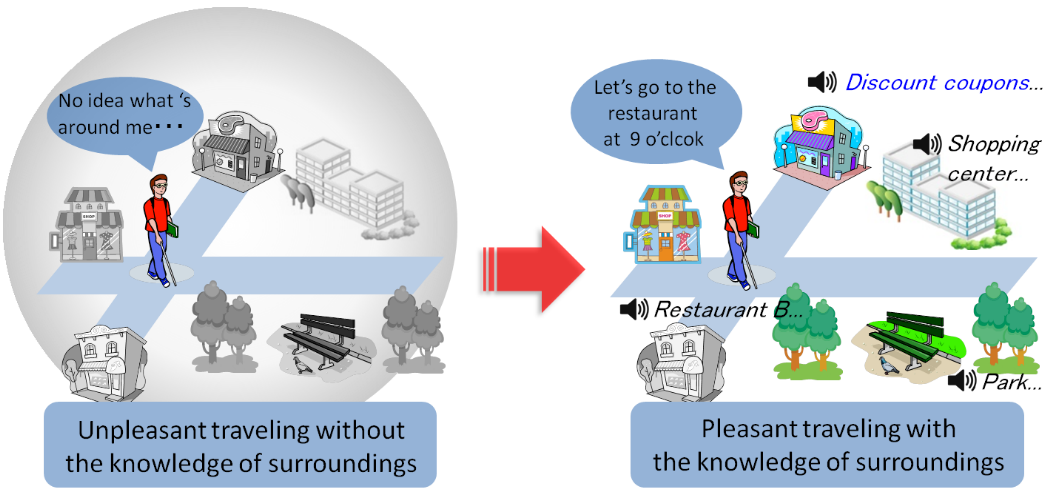


Figure 4.1: Blind person who enjoys traveling by exploring surrounding POIs.

need is also recognized by large-scale field tests in research projects, carried over a decade, using ubiquitous computing technologies for enhancing mobility experience of the blind [98, 99, 100].

Map exploration systems [43, 45] are one of the effective options for blind people to learn an environment. They often adopt tactile or auditory feedback to interact with information on maps, and this may require dedicated hardware. These interaction techniques may be useful for learning an environment in detail, but are hard to apply to discover POIs while walking. Automatically notifying immediate surroundings could be another option [46, 47], but may not be useful when users want to retrieve POIs at their discretion. Thus, how to allow users to interact with the surrounding geo-located information space is the key to effectively support exploratory traveling [40]. To this end, I introduce a novel and simple touch-based interaction method, *swipe-to-search* that allows blind users to select areas-of-interests. Users intuitively specify the direction (relative to the direction they are facing) and distance of areas-of-interest by swipe gestures on touchscreen. For example, users can find nearby POIs in front of them by performing short swipe-up gestures, and distant POIs on their right by performing long swipe-right gestures. This approach can be easily implemented on handheld devices without requiring any dedicated hardware and having users to constantly focus on tactile or auditory feedback. In this chapter, it is shown that swipe-to-search is usable and acceptable to blind users. Pointing a smartphone towards areas-of-interest has been used for retrieving POIs [49], but this method could have usability issues as discussed in Chapter 2.

Based on swipe-to-search, I developed SaSYS (Swipe and Scan Your Surroundings) on a smartphone, which enables users to search for POIs such as restaurants and stores. SaSYS provides interaction methods that can be used to control audio information by allowing the user to draw shapes on touchscreen. For example, while listening to user reviews about some restaurants, the user can skip to the next review by drawing a triangle. SaSYS also offers auditory and olfactory features of POI such as “There is a café on the right, near which coffee aroma can be experienced.”

4.2 SaSYS System

SaSYS is a location-aware system running on touch-based handheld devices that allows blind users to discover POIs and get detailed information about them. To provide useful and easy-to-use interaction methods for searching and obtaining POI information, I designed SaSYS to achieve the following:

- SaSYS enables users to search for POIs in all directions, and allows them to adjust the scope of search areas at will.
- SaSYS enables users to obtain information effectively from auditory information, as needed, by allowing them to choose what to listen to.

Furthermore, I explored what kind of POI information could be useful for blind people [17, 101]. SaSYS can provide olfactory and auditory features of POI (currently manually collected in a specific area).

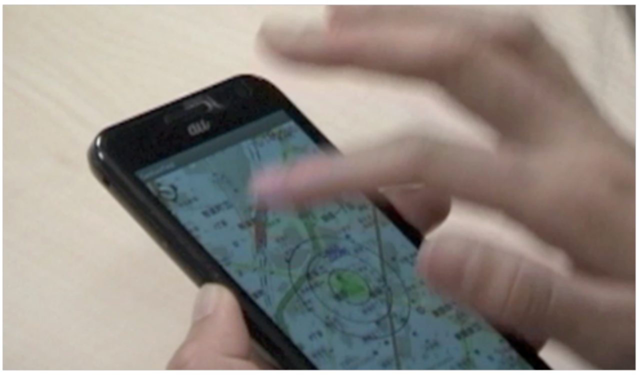


Figure 4.2: Participant searching for POIs using SaSYS by performing swipe gestures.

4.2.1 Design Process

Figure 4.2 shows the current SaSYS prototype, which was developed through three iterations of design, implementation, testing and refinement. First, I conducted initial prototype test with the blind teacher to assess the feasibility of the swipe-to-search approach, and discover potential usability problems. He provided the following feedback for ensuring better user experience:

- *Only intentional touches should be recognized; the prototype should provide some kind of a "touch trigger" to avoid accidental activation.* Therefore, I chose the volume keys on the side to act as a "touch trigger," effectively ignoring touch inputs when one of the volume keys is not pressed.

- *Let users know what appears onscreen, and provide instructions in order not to get lost while navigating the application.* Thus, I gave instructions for non-triggered touches as well. For example, if the POIs list shows up on screen, the following voice instruction is given: "POIs list. Swipe up and down to navigate POIs and double-tap to select one."

- *Minimize the number of drawings for TTS controls to remember them easily.* Therefore, I made the basic shapes to mean different interactions depending on whether or not they are drawn while the volume key is pressed. For example, while listening to audio information, drawing a greater-than sign with the volume-up key pressed would cause *speech rate up*, whereas doing the same with the volume-down key pressed would skip to the next sentence (Figure 4.5).

After refining the prototype as described above, I again consulted the blind teacher to ensure usability improvements before conducting a user study. Moreover, I demonstrated the refined prototype to 3 blind individuals at the 2013 TRON Enableware Symposium[13] to examine the swipe-based system's usability. One of them was iPhone user and the others had never experienced using smartphones. All of them could understand, and perform swipe-to-search with interest, indicating the potential of my interaction technique.

SaSYS interactions consist of two modes: Scanning mode to retrieve POIs by swipe gestures, and content listening mode to listen details of selected POI information. The following subsections present two modes illustrating how I achieve the design goals and describe implementation of SaSYS in detail, highlighting its novel interactive features.

4.2.2 Scanning Mode

Scanning mode allows the user to retrieve POIs and select one from the list of POIs. SaSYS's swipe-to-search interaction is derived from very simple and intuitive ideas. Suppose that you were asked to present a desired direction on the paper. Drawing an arrow might be one of the easiest ways to show the direction. Additionally what if you were asked to indicate the relative distances you want to know? You can intuitively express the sense of distance by drawing different lengths of arrows. I have implemented these ideas on a touchscreen smartphone. The user first touches any point on touchscreen, which is regarded as the current

[13] `http://www.tron-enableware.org/en/`

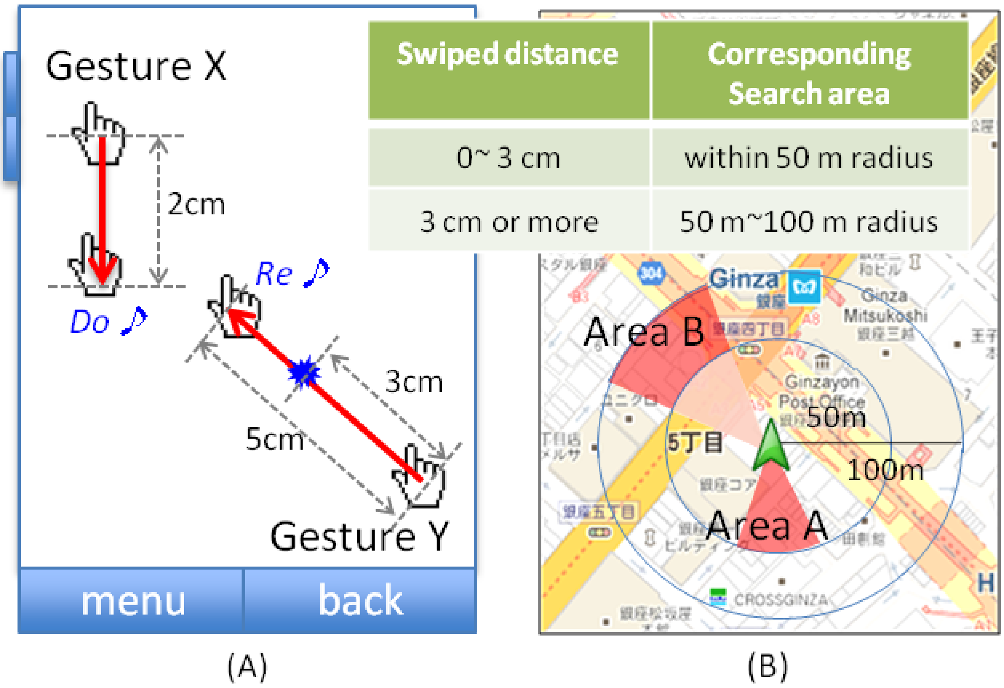


Figure 4.3: Examples of swipe-to-search. (A) Gesture X retrieves POIs within Area A, and (B) Gesture Y retrieves POIs within Area B.

location, then draw a line towards a desired direction. This simple gesture gives the user a POIs list within the area-of-interest.

An example of swipe-to-search is illustrated in Figure 4.3. In the example, swiped distances between 0 to 3 cm retrieve POIs within 50 m radius, and 3 cm or more retrieve POIs within from 50 to 100 m radius in the 8 cardinal directions. The user can customize the swiped distances and their corresponding search radii. When the user performs a short swipe-down gesture ("Gesture X" in Figure 4.3A), SaSYS offers a POIs list within 50m behind the user (Area A in Figure 4.3B). Also, the user can retrieve POIs between 50m and 100m front-left by performing a long swipe gesture diagonally ("Gesture Y" in Figure 4.3A and Area B in Figure 4.3B). In this manner, the user can intuitively explore nearby places by short swipes and distant places by long swipes in all directions.

POI retrieval process includes three steps: (a) measure a direction and distance of a swipe gesture, (b) convert swiped distance to predetermined search ranges and calculate the latitude and longitude of the center of a target area, and (c) get POIs information within the area and read the search results via TTS.

In addition to explicit speech feedback about swipe gestures performed by the user, SaSYS provides vibrotactile and auditory feedback for assisting the user to perform swipe-to-search. For example, while performing Gesture Y as depicted in Figure 4.3A, the user will receive vibrotactile feedback as he advances from Area A to Area B, crossing the 50 m radius of Area A (3 cm gesture equivalent). The duration of the vibration will be in effect corresponding to $(3 \pm \alpha)$ cm, where α (currently 20 mm) is a configurable parameter. It helps the user perform swipe gestures as he intended by allowing him to know how long he slid his finger. Auditory feedback, one of 7 notes in the major scale, is given according to the radius of search areas (e.g., Figure 4.3B has two areas, A and B) when the user's finger lifts off the screen after finishing the swipe gesture. Swipe gestures specifying further areas get higher notes; for example, if search space is divided into 3 areas, then swipe gestures searching for the third area play "Mi."

The user can perform swipe-up and swipe-down gestures to navigate the list of retrieved POIs. SaSYS reads aloud brief description of the POI that includes the name and distance from user's current location. The user can select POI for detailed information by double-tap. A hardware back button (standard for most Android phones) is used to go back to the swipe-to-search mode from navigating the POI list.

4.2.3 Contents Listening Mode

After the user selects POI on the list, SaSYS provides more detailed information with several controls that make TTS repeat, forward, skip, etc. When sighted people search for POI using mobile applications (e.g., nearby restaurants for lunch), they might look over or into detailed information to filter information they want

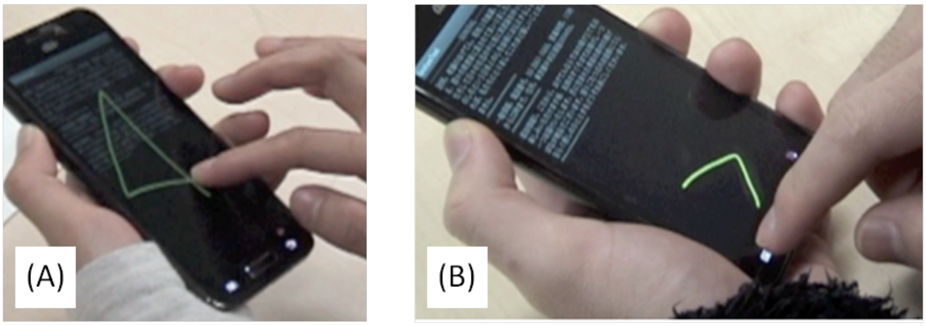


Figure 4.4: Participants performing TTS controls: (A) skip by drawing a triangle and (B) forward by drawing a greater-than sign.

to get (e.g., the latest review). However, it is difficult for blind users to get desired information efficiently through unilateral auditory information. The user may have to listen again from the beginning for missed information that sighted people may not need even a second to read again. Although the need for TTS controls has been reported [102], few systems provide this feature. Therefore, it is necessary to provide TTS controls feature for blind users to obtain desired information effectively from TTS by allowing them to choose what to listen to.

I use gesture-based interactions that allow the users to control TTS by letting them draw shapes on touchscreen (Figure 4.4). TTS controls include play, pause, repeat, go forward and backward by the sentence, skip to the next and previous by the item (e.g., each customer review), and speech rate up and down. By providing various control features, SaSYS helps the user "look over" or "look into" detailed information. I initially assigned gestures to each TTS control in accordance with design guidelines for accessible touchscreens [35, 39]. For denoting gestures, Figure 4.5 shows a set of shapes provided by SaSYS. However, in the future, it would be desirable to enable the user to customize the gesture-shape mapping.

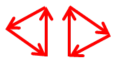
Functionality	User Interface
Play and Pause	Double tap with pressing the volume down key
Forward	Drawing  with pressing the volume down key
Backward	Drawing  with pressing the volume down key
Next Content	Drawing  with pressing the volume down key
Previous Content	Drawing  with pressing the volume down key
Change the Information Category	Drawing  with pressing the volume up key
Repeat	Drawing  with pressing the volume down key
Speech Rate Up	Drawing  with pressing the volume up key
Speech Rate Down	Drawing  with pressing the volume up key

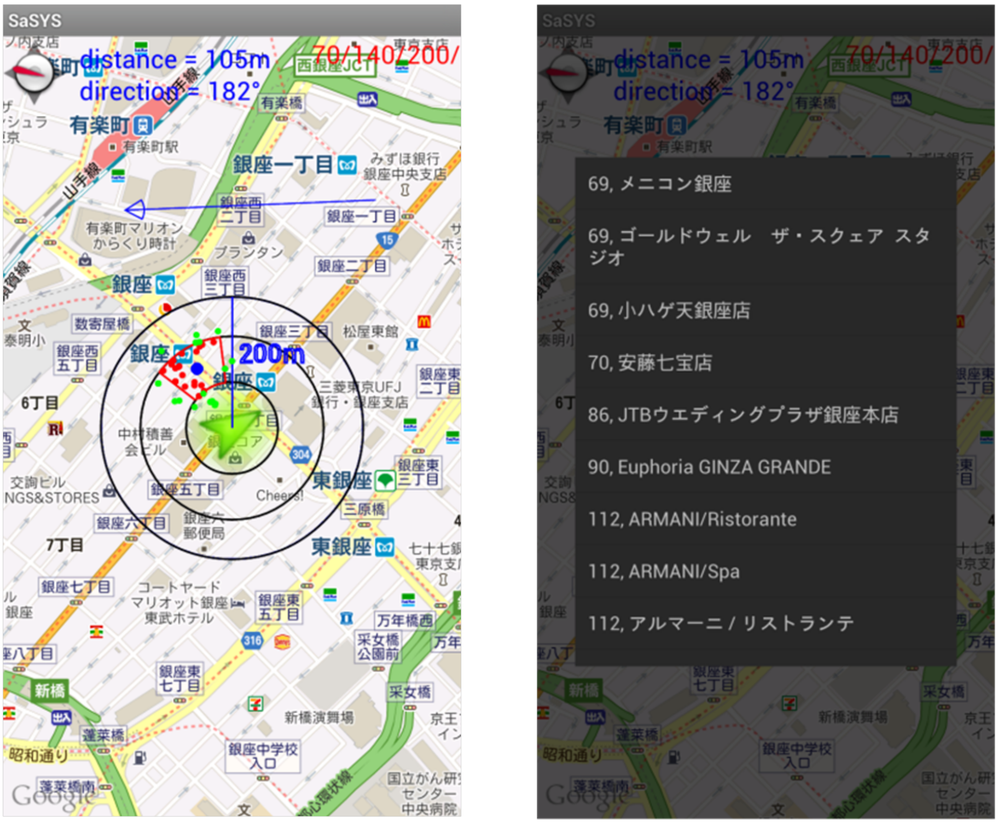
Figure 4.5: The gesture set used for TTS controls.

4.2.4 Implementation

SaSYS is implemented on Android platform, and currently running on Samsung Galaxy S2 and Galaxy Nexus. SaSYS makes use of smartphone's GPS-based positioning and a built-in compass to identify the user's current position and orientation. I used commercially available SOVOX Classic TTS engine[14] with Japanese language pack to provide POI information and interact with the user. I also utilized Gesture Builder which is an application included in Android 1.6 and higher SDK platforms to create custom gestures used for TTS control interactions. Gesture Builder allows developers to create a library of pre-defined gestures that can be used in their applications.

When the user performs a swipe gesture, SaSYS retrieves POIs on the map (Figure 4.6A), and then immediately transitions to the screen that shows the list

[14] `http://www.nuance.com/products/SVOX/index.htm`



(A) (B)

Figure 4.6: A screenshot after performing swipe-to-search. (A) The blue arrow shows the gesture performed by the user; the green and red dots present retrieved POIs but only red dots will be introduced to the user. (B) The list of retrieved POIs.

Selected POI	Description provided by SaSYS
Bank	There is a bakery on the right, near which the smell of bread can be experienced
Bakery	It smells of bread
Post Office	There is a bakery on the left, near which the smell of bread can be experienced

Figure 4.7: Examples of contextual cues as auditory and olfactory features of POI.

of search results (Figure 4.6B). After the user selects POI, SaSYS moves to the screen that shows detailed information (Figure 4.4).

SaSYS offers rich information about points-of-interest. The current implementation of SaSYS supports a use case of practical importance of going around in Ginza, a famous shopping and entertainment district in Tokyo. Kokosil[15] API was used to provide local businesses information in Ginza. To help the user find desired information easily, I categorized POI information into four types: general information, user reviews, advertisements (e.g., discount coupons, events, etc.), and contextual cues. The user can move between not only items but also categories so as to save time by listening to only information of interest. General information includes a short description about POI, business hours, address, regular holidays, and additional features such as history of POI. User reviews and advertisements are retrieved from Kokosil and Twitter. Advertisements often contain URL for further information, but it can be tiresome for the user listen to the meaningless characters via TTS. Thus, SaSYS examines posts and removes URLs if any. Contextual cues

[15] http://home.ginza.kokosil.net/en/

are information about auditory and olfactory features of POI. For example, aroma of coffee near a café or sound of outdoor air conditioner units. This contextual information can be useful to visually impaired people when they get near to a desired POI. To the best of my knowledge, there is no content provider that offers auditory or olfactory features of POI. Therefore, I manually collected contextual information of POIs in Ginza. This information is created as a CSV file. When the user selects POI, SaSYS searches this file whether the POI has any contextual features. Figure 4.7 presents contextual cues provided by SaSYS.

4.3 Evaluation

To verify the usability of interaction techniques provided by SaSYS, I conducted experimental evaluation with 11 visually impaired participants in the laboratory. The user study consists of three parts: A) verifying the ease-of-use and usefulness of swipe-to-search approach by comparing to pointing-to-search approach, B) verifying the effectiveness of TTS controls feature, and C) obtaining user feedback on the overall system. The whole process of the experiment was conducted in Japanese, and was video taped for later analysis.

4.3.1 Participants

Eleven visually impaired participants (four male and seven female; P1-P11) were recruited for the study. Their degree of visual impairments varied: Nine of them were blind and two had low vision. The average age of the participants was 23.09 (SD=8.1). Two participants were iPhone users. The others are not smartphone users but three of them had experience with touchscreen mobile devices. All participants use white canes as mobility aids and six participants had experience with GPS navigation systems. The entire study took on average 2 hours.

(A) (B)

Figure 4.8: An example of performing Pointer Test: (A) searching POIs on front-left by pointing the smartphone to that direction, and (B) a screenshot after performing (A).

4.3.2 Procedure

Part A: Swipe-Based Interaction

For comparison with pointing-based interaction, I developed an Android application called “Pointer Test” that provides the same functionalities as SaSYS, but using pointing-to-search. Pointer Test retrieves POIs by pointing the smartphone in a desired direction and double tapping on the screen (Figure 4.8).

To allow the users to develop a clear impression of both approaches, I asked them to perform the following task using both applications:

> *“Imagine that you stand at an exit of Ginza subway station. Before traveling, you want to know what is around you. Explore your surroundings with the given application. Please try to find four specific*

Figure 4.9: (A) Participant putting blocks on a LEGO board to present spatial relationships among the location of stores. (B) Spatial relationship of POIs as perceived by a participant.

> *stores in Ginza and remember their locations. After 10 minutes, represent each location on a LEGO board with blocks, and tell us the names of each store."*

Both SaSYS and Pointer Test used the same POIs set which consists of 66 POIs in Ginza within a 200 m radius centered on the participant's simulated location. I divided SaSYS's search space into three areas by swiped distances: searches from 0 to 70m radius (corresponds to from 0 to 2 cm swipes), from 70 to 140 m radius (corresponds to from 2 to 4 cm swipes), and from 140 to 200 m radius (corresponds to over 4 cm swipes). I limited the number of POIs for the experimental purpose because navigating too many search results was cumbersome to the participants. The participants were asked to remember the direction and distance of the given POIs when they discover them and present their location on the LEGO board (the center of the board was regarded as the current position of the user) (Figure 4.9). I used LEGO to enable the participants to simply express the relative spatial relationships among stores. To avoid making the task into a test of how well the participants were able to remember the given locations, the users were allowed to ask the name of given stores while performing the task.

To minimize the influence of area familiarity from the former session, I set different current locations for each session (exit A3 at first and then exit B4 of Ginza subway station, regardless of which application would be used first). I randomly chose six participants to use swipe-to-search first and the others to use pointing-to-search first. At the beginning of each task, I conducted a training session for all participants until they became familiar with the application. After finishing the task for both applications, the participants were asked to answer the questionnaire about ease-of-use and usefulness of each approach.

Part B: TTS Controls

The participants were asked to answer three questions while listening to POI information about a restaurant. To avoid making the task too simple, I added more user reviews and advertisements manually, collected from social network services. The questions include *(Q1) What time does the restaurant close?*, *(Q2) how much the cheapest menu of the restaurant in the user reviews?*, and *(Q3) what services does the coupon offer during lunch?*

After the participants finished the task, I asked them to answer the questionnaire about the effectiveness of TTS controls provided by SaSYS.

Part C: User Feedback on the Overall System Design

After the participants had completed the tasks I conducted an interview about the overall experience with SaSYS. I asked the participants to provide any comments on SaSYS including additional functionalities that they would like to have, potential use cases, gesture-based interactions on the touchscreen, etc.

I also investigated whether the auditory and olfactory features of POI provided by SaSYS could be useful for visually impaired people. Furthermore, I asked the participants to mention other helpful contextual information that they would like to acquire through SaSYS.

4.3.3 Results

Part A: Swipe-Based Interaction Results

The participants generally were able to complete the task using both applications. All participants could discover the given stores, but some participants had difficulties in remembering the location and name of the stores when they were asked to put the blocks on the LEGO board. 11 participants presented the location of four given stores using the two interaction methods, 88 items in total. They were able to properly put the blocks with the store names 71 times, only the locations without the names 4 times, and completely failed to put the blocks 13 times (8 times with SaSYS and 5 times with Pointer Test). After finishing both sessions, the participants rated each interaction method in terms of usefulness and ease-of-use using a five-point Likert scale (1=Disagree strongly, 5=Agree strongly). The results are shown in Table 4.1.

Table 4.1: The ratings on Ease-of-Use and Usefulness for each interaction method (Mean, SD).

Interaction Method	Ease-of-Use	Usefulness
Swipe-to-search	3.91 (1.04)	4.36 (0.81)
Pointing-to-search	3.18 (1.33)	3.72 (1.19)

Although a Wilcoxon test did not find a significant difference regarding both ease-of-use (z=1.26, p=.24) and usefulness (z=1.51, p=.14), 9 of 11 participants preferred swiped-based interaction to pointing-based interaction (P9 chose swipe-to-search on the condition that search space is not divided by swiped distances) when they were asked to indicate their favorite of the two interaction methods.

The participants were also asked to comment on the reasons why they thought the method they chose was better than the other. The participants who liked *swipe-to-search* commented as follows:

- Easy to search backward while walking (P1, P6, P7, P8, P9, and P11).
- Pointing a device is not realistic in a crowded place, whereas swiping on the screen allows him to intuitively imagine his surroundings (P2).
- Pointing a device on the street would be embarrassing (P4).
- Adjusting search areas by swiped distances would be useful in different use cases (P10).

On the other hand, the participants who liked *pointing-to-search* commented as follows:

- Retrieving POIs by double tapping was easy to use and simple (P3 and P5).

The participants explicitly mentioned the concerns about losing their sense of direction and the awareness of their body position in relation to the surroundings when they turn around for searching backward with Pointer Test. P8 said, *"Compared to performing swipes, it is quite demanding to turn around with holding the device [for searching]. I found [Pointer Test] tiresome for having to change my body position."* P7 also pointed out, *"Turning the direction [of the device] or changing one's own body position would end up losing the sense of direction."*

Although most participants preferred SaSYS's swipe-based interaction method, some participants were either neutral or positive about Pointer Test's pointing-based interaction method. P2 noted, *"I think it would be useful to go to the [discovered] place by moving the device [with Pointer Test] after understanding the surroundings by moving the finger [with SaSYS]...So it is hard to say which one is better"* and added *"but surely, the one with moving the finger [swipe-based interaction method] is necessary as a means of imaging the surroundings."* Similarly, P6 commented, *"While I generally prefer swipe [based interaction] mode, it would be good to know [information about] a building right in front of me by pointing the device and double-tapping."* Integrating SaSYS with wayfinding systems such as TP3 [49] could respond to the needs.

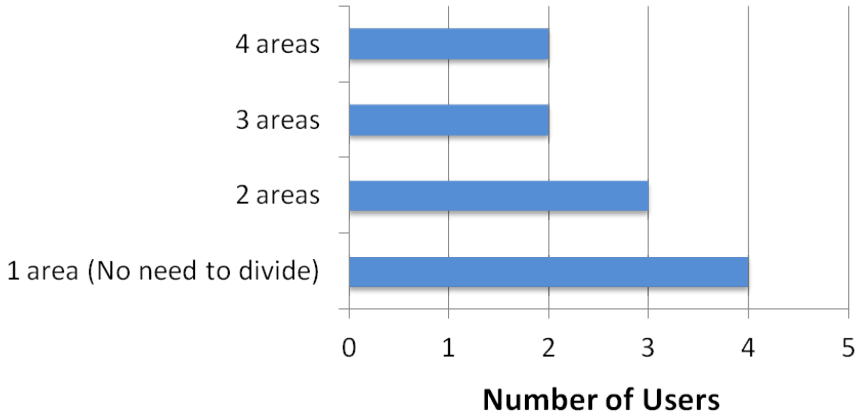


Figure 4.10: Participants' preference of the number of areas they would like the search space to be divided into.

Some participants found it difficult to control the distance by sliding their fingers on the touchscreen. P11 thought that dividing search space by swiped distances was not a desirable feature for him because he has a slight finger-motion impairment. However, P10 liked swipe-to-search because she could control search areas by swiped distances. I asked the participants about the desirable number of areas for this feature. The preference was varied as shown in Figure 4.10. The result indicates that interaction designs for blind users should take individual characteristics into account. Thus, the number of areas should be customized according to each user's preference.

Part B: TTS Controls Results

Overall, participants could easily remember the given drawings for each control and perform the task. 8 of 11 participants answered correctly to all the questions, but 3 participants gave a wrong answer only to one question each (both P1 and P6 answered incorrectly to Q(3), and P9 answered incorrectly to Q(2)). SaSYS takes 248 seconds to read the prepared POI information with normal speech rate. The participants used TTS controls on average for 504 seconds (SD=119) to find answers for the given questions. 10 of 11 participants found that they needed TTS controls when they were asked to indicate whether the TTS control feature is necessary.

The participants were also asked to rate the usefulness of each TTS control provided by SaSYS using a five-point Likert scale (1=Disagree strongly, 5=Agree strongly). Overall, the participants found the given functionalities were useful. In particular, the "speech rate" control was thought of as the most useful functionality whereas "repeat" was thought of as the least useful among the given functionalities (Table 4.2).

Table 4.2: Results of questionnaire on usefulness of each TTS control functionality.

Functionality	**Average Rating (SD)**
Repeat	3.72 (1.19)
Forward	4.18 (0.75)
Backward	4.27 (0.79)
Speech Rate	4.28 (0.4)
Skip to Other Items	4.27 (0.64)
Skip to Other Categories	4.36 (0.67)

Most participants liked the drawings I assigned and the recognition rates of the drawings were acceptable, but some participants had difficulty drawing swirls to repeat. P3 mentioned, *"It would be better to assign symbols less similar [to triangles to the repeat control] because swirls are recognized as triangles."* As noted earlier, it is important to allow the user to customize interactions as much as possible.

Overall, the participants liked gesture-based interaction to control auditory information. When I asked the participants about their preference between button-based interface and gesture-based interface, if the drawing recognition was not a problem, surprisingly 8 of 10 participants (who thought the TTS control feature was necessary) chose the gesture-based interface. This shows that symbol-based gestures have a potential to be used as an effective interface with touchscreen for blind people.

Part C: User Feedback on the Overall System Design Results

The participants explicitly commented that SaSYS would encourage them to explore their surroundings and to participate in shopping and leisure activities. P6 said, *"I would like to use [systems like SaSYS] all the time. I would like to search stores right now."*

P11 who is currently using an iPhone compared SaSYS with existing navigation systems providing turn-by-turn instructions. P11 pointed out, *"Navigation systems require predetermined destinations...I think they are useless in situations like going around Asakusa (one of the famous districts in Tokyo)."* He also commented, *"I think I can use [SaSYS] like equipment for leisure time amusement. For example, when I want to go shopping or look for some restaurants in Ginza."*

I investigated additional features for improving the user experiences with SaSYS. Several participants thought that it would be better to share their findings (e.g., a bakery near the station) with other users. Some participants suggested that it is important to know the current direction they are facing.

I found that all participants thought contextual information (olfactory and auditory feature of POI) would be helpful when they get near to a desired POI (or destination). I also asked the participants to whether SaSYS requires any additional features. Several participants noted that information about the following would be helpful: Shapes and types of entrances (e.g., whether it is a manual door or an automatic door), the position of the entrance (whether in the middle or the end of a building), whether there are steps or slopes near the entrances and the texture of the floor of the entrance hall (e.g., carpets, concrete, wood). Furthermore, some participants mentioned that the characteristics such as the colors of signboards or the shapes of buildings could be used as clues when they ask the way to sighted individuals. P6 and P11 suggested that the order of buildings in a city block with street names would be helpful. For example, "Your are on Harumi Avenue. You can find buildings A, B, and C in order." Although the current prototype of SaSYS does not provide turn-by-turn navigation to a POI, I believe that POI description should contain useful information for blind people that sighted people may take for granted.

A number of participants enjoyed the experiences with SaSYS and the user study. P10 said, *"Overall, I enjoyed [the user study]. It would be very useful once I get used to smartphones."*

4.4 Discussion

Regarding part A of the user study, it was encouraging that the participants intuitively understood the concept of swipe-to-search and were able to find POIs. Most participants preferred the swipe-to-search method, but two participants thought the pointing-to-search method was better for them because of its simplicity. I observed that they seemed to avoid touch input as much as possible, which made them feel the pointing-based interaction was simpler than swipe-based interaction. Several users also were positive to the pointing-to-search method. Further extension of SaSYS could add the pointing-based interaction method to allow users to select a proper mode according to their preferences or use cases.

Although several participants thought dividing search space by swiped distances could be useful, some participants found that it was difficult or tiresome to adjust swiping distances in order to select desired search areas. I found two possible explanations for why this feature was not well accepted. First, individual ability and attributes (e.g., spatial ability, tactile sensibility, age) could affect the performance [103]; and to perform the task, users had to interact much more with SaSYS. To scan all areas, SaSYS required users to search 24 times (8 directions with 3 search areas), which made the system seem more tiresome compared to Pointer Test (which only takes 8 times to scan non-divided search space). Note that I extremely limited the number of POIs (66 stores) for the task. I believe that adjusting search areas is necessary when dealing with dozens of POIs within a specific area in real-world scenarios. This issue will be improved by allowing users to customize the number of search areas.

During part B of the user study, I observed that several participants spent much more time on finding the answer to *Q(3)* (find a coupon available at lunch time) compared to *Q(1)* and *Q(2)*. They seemed to have difficulty figuring out the starting point of each item (a coupon), whereas they easily distinguished items (user reviews) from each other because each review starts with the date of posting. I believe that numbering items or giving users feedback (e.g., vibration or auditory icons) at the end of each item could provide a better user experience in case of auditory information.

It was interesting that most participants used and rated the speech rate control the highest. Overall, users preferred speech rate up and down functionalities to quickly "look though" the contents rather than fast-forwarding or skipping functionalities. The participants seemed to be concerned about the possibility of missing out on important information. One participant used only speech rate controls to perform the task. This suggests that providing just the speech rate controls could allow users to obtain information effectively. On the other hand, the repeat control was thought relatively less useful because the recognition rates of swirl shapes were poor, and the backward control could substitute for the repeat functionality. I believe that assigning and registering custom drawings by the user could improve this issue.

This study also confirmed that providing contextual information could benefit blind users. The current implementation of SaSYS provides only manually collected olfactory and auditory features of POIs in Ginza, but these features can be affected by certain situations (e.g., rains can affect olfactory senses). Furthermore, it would be desirable to provide richer contextual information anywhere and at any time suggested by the participants. Future work should investigate into how to collect and manage contextual information for blind users.

4.5 Summary

This chapter presented a novel and intuitive interaction technique, swipe-to-search, to allow blind people to search for the surrounding geo-located information space. Based on swipe-to-search, I developed SaSYS using off-the-shelf smartphones. SaSYS also provides TTS control feature, which gives blind users liberty to what to listen to.

The results of this study revealed valuable insights into the design requirements of an effective location-based service like SaSYS. It showed that swipe-based interaction method was accepted well by users as a means of location selecting. Users also preferred swipe-to-search method to existing pointing-to-search method. Moreover, it showed that TTS controls are highly useful to visually impaired users. Users thought the speech rate control was the most useful functionality. The user study also showed that olfactory and auditory features of POI would be helpful to visually impaired people. In short, the participants found that SaSYS would be useful for exploring unfamiliar places and discovering surrounding POIs.

Chapter 5

TalkingTransit: A Location-aware System for Aiding Public Transit Use

This chapter[16] focuses on enhancing public transit accessibility for blind people. Specifically, taking advantage of open data from transit authorities and operators, the chapter presents TalkingTransit, a location-aware system that enables blind people to obtain real-time service status and timetables of public transit in Tokyo that has more than 1200 stations and 1500 bus stops. The key challenge in designing TalkingTransit was to allow blind users to deal with such a vast amount of information easily. To address this problem, in this chapter, a set of user interfaces is designed by utilizing multi-touch gestures and smartphone's built-in sensors.

5.1 Motivation

Modern cities are equipped with public transit systems including underground and surface train, and bus networks. People of all walks of life—commuters, students, senior citizens, and people with special needs—greatly depend on public transit for their mobility needs. Particularly, people with visual impairments, most of whom are unable to drive, highly rely on buses or trains for getting to work, school, shopping and other facilities. However, they face many challenges

[16] Portions of this chapter were previously published in [24]

when using public transit due to poor access to information such as timetables, line numbers of arriving vehicles, location of facilities at station – which are mostly provided in visual formats (signs, prints, digital displays, etc.) [54, 63]. Unexpected situations involving real-time changes like delays and suspension of services can cause additional difficulties. Taking into account the fact that there are more than 285 million visually impaired people and rate of vision loss will double along with aging population by 2030 [6], providing accessible transit services is crucial to inclusive cities.

The recent trend of making proprietary data increasingly publicly available by major organizations, coupled with the governmental commitment to open data at national and international levels including the G8 forum[17], has greatly expanded the possibilities for creating value-added IT services. Recently, the Japanese government has been promoting building Open Data infrastructures [104], and as a part of this initiative, *OpenData API* has been developed for efficiently retrieving massive transit information in the Tokyo metropolitan area.

By using the Open Data infrastructure and IoT technology, this chapter presents TalkingTransit (TT) that enables users to obtain (1) real-time service status and schedules using smartphones and (2) detailed in-station information via BLE technology. TT offers a set of intuitive interfaces iteratively co-designed with blind users, which allows them to access mass transit information in Tokyo, one of the most complex public transit systems in the world. Existing push-based systems [53, 54, 55] notifying users of arriving vehicles often use smartphone's GPS to trace a bus, and/or custom hardware (e.g., tactile bracelet), but may not be useful for checking timetables or service status in advance. Some pull-based systems [57, 62, 63] allowing users to retrieve transit information adopt a Braille display or keyboard for easy interaction, but these are expensive and cumbersome as discussed in Chapter 2. Unlike the existing systems, TT doesn't require any dedicated hardware and can access real-time information anytime, anywhere by using open data

[17] `https://www.gov.uk/government/publications/open-data-charter`

infrastructure. Using OpenData API, TT can access massive information made open by 11 railway operators including East Japan Railway and private railways; 2 subway operators (75 lines and 1270 stations in total) and Toei Bus (132 routes and 1585 stops). Information retrieved by OpenData API include real-time service status, current positions of vehicles, timetables, and in-station information from in-situ BLE markers and sensors (temperature, humidity, pollen, etc.).

5.2 TalkingTransit System

TalkingTransit is a location-aware system running on smartphones that allows blind users to obtain transit information including vehicle schedule, and service status which gives real-time updates on delays, service changes and suspensions of trains in Tokyo.

5.2.1 System Overview

Through iterative design process with blind users, which is described in the next subsection, the current version of TT includes four main functionalities:

- **Timetable search:** Users can get the arrival times for the next vehicle, as well as a full timetable by using one of the two search modes according to their usage scenarios. *GPS-based search* allows users to search for the timetables of trains/buses within 500 m of their current position, which can be useful when they are going to nearby stations or already at bus stops. On the other hand, *alphabetical search* allows users to find the timetables of trains/buses by a particular station name, which can be helpful when they make a journey plan in advance.

- **Check the real-time status of railway services:** Using this function, users can avoid any possible inconvenience due to service changes.

- **Bookmarks:** Users can get easy access to timetables for vehicles that they

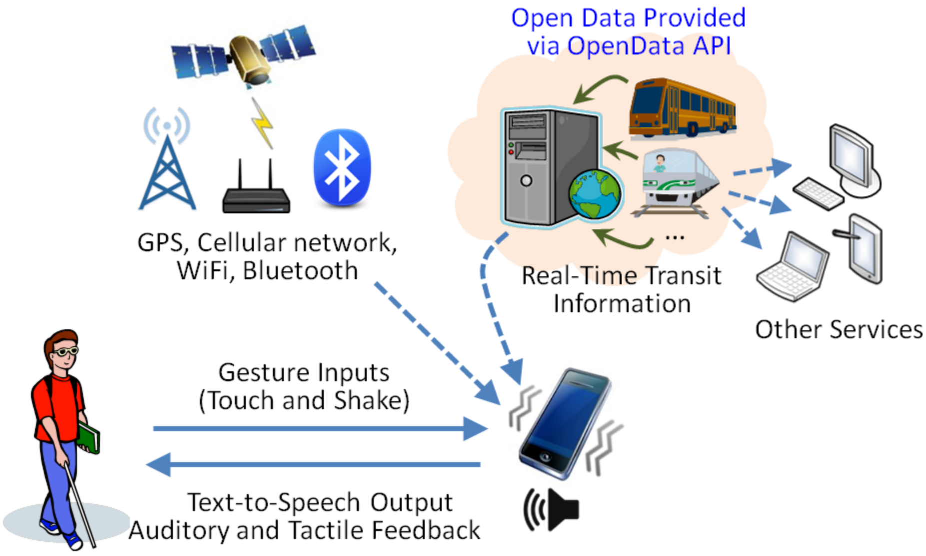


Figure 5.1: TalkingTransit system architecture.

frequently use by bookmarking pairs of stations and train lines or stations and bus routes.

- **Automatic notification of in-station information:** TT automatically notifies users of points-of-reference for a nearby platform and facilities (e.g., stairs, escalators, toilets) within 2 m of their current position. The notification consists of a sound alert and a short description of the spot. If the user gets notification of a nearby platform, he can directly skip to the page containing timetables of trains in that platform.

As illustrated in Figure 5.1, TT system architecture consists of user smartphone, positioning infrastructure, and open data infrastructure. Positioning infrastructure including GPS, Wi-Fi, and Cellular network allows the user to get his current location, and search nearby stops and stations. BLE is used in the railway station, which has been recognized as an appropriate technology for indoor positioning because of its range (approximately 1∼2 meters), low power consumption, and cost-effectiveness. The Open Data infrastructure provides real-time transit information via OpenData API not only TT but also many third-party services. TT is running on a touch-based smartphone equipped with sensors for positioning. The user interacts with TT using gestural input and TTS output with auditory and tactile feedback.

TT is implemented on iOS 7 platform, and currently running on Apple iPod touch and iPhone 5S which have GPS and Bluetooth. TT makes use of the device's built-in accelerometer and gyroscope sensors to detect shake-motion events. I also used AVSpeechSynthesizer API included in iOS 7 and higher platforms to offer TTS output without using VoiceOver.

5.2.2 Requirement Analysis

The current system was developed through an iterative design process in consultation with the blind teacher. First, I developed a prototype system running on iOS, and conducted a pilot study to understand potential requirements for a better system. The first design for the user interfaces was based on VoiceOver which is used by more than 100,000 blind people [31]. The main interaction with the prototype was performed by three-finger swipes following VoiceOver's gesture for scrolling between pages. Users can navigate the list with swipe up and down, select an item with swipe right, and go back to the previous page with swipe left. The prototype was tested with 10 visually impaired users (six male and four female), each of whom used the system for about an hour. This testing identified the following requirements for ensuring a better user experience:

- One- or two-finger swipe and double tap would be easy to perform. Swiping left and right with three fingers is tiring.
- The demand for filtering huge text data to find desired information in a list should be reduced.
- The demand for following which page the user is in while navigating the application should be reduced.

- Explicit feedback for what happened after performing touch inputs should be provided.
- Direct access to timetables of frequently used vehicles should be provided.
- Location-based information in the subway stations for identifying a right platform to ride should be provided.

5.2.3 User Interface

Users interact with TT using a set of touch gestures (one- and two-finger swipe up, down, left, single-tap, double-tap, and three-finger long press) and a shake gesture. As shown in Figure 5.2, swipe up and down to navigate the list, double-tap to select, two-finger swipe left to go previous page, single-tap to get description of the current page, three-finger long press to bookmark, and a shake to go back to the main page. When starting the application, users can access transit information by selecting one of four items in the main menu: Favorites, service status, bus, and train (Figure 5.2A). From the favorites menu (Figure 5.2B), users have quick access to timetables of their registered vehicles by selecting one of two types of public transit and then a specific vehicle (Figure 5.2C). In the service status menu (Figure 5.2D), users can check the real-time status of 11 railway and subway transit operators whether there is a notification or not. If there are any messages from an operator, users can see the details by selecting the operator (Figure 5.2E). In the bus and train schedule menu (Figure 5.2F), users can find vehicle timetables by two ways: Search by current position and by station names (ordered by Japanese alphabet called Hiragana). For example, with location-based search in the bus schedule menu, users can get the list of bus stops within 500 m radius. Then they can select a desired stop (Figure 5.2H) and route (Figure 5.2I) to access the timetable (Figure 5.2K). Interacting with the three functionalities above would give a list of 10 items or less, which makes it very easy to navigate using swipe gestures.

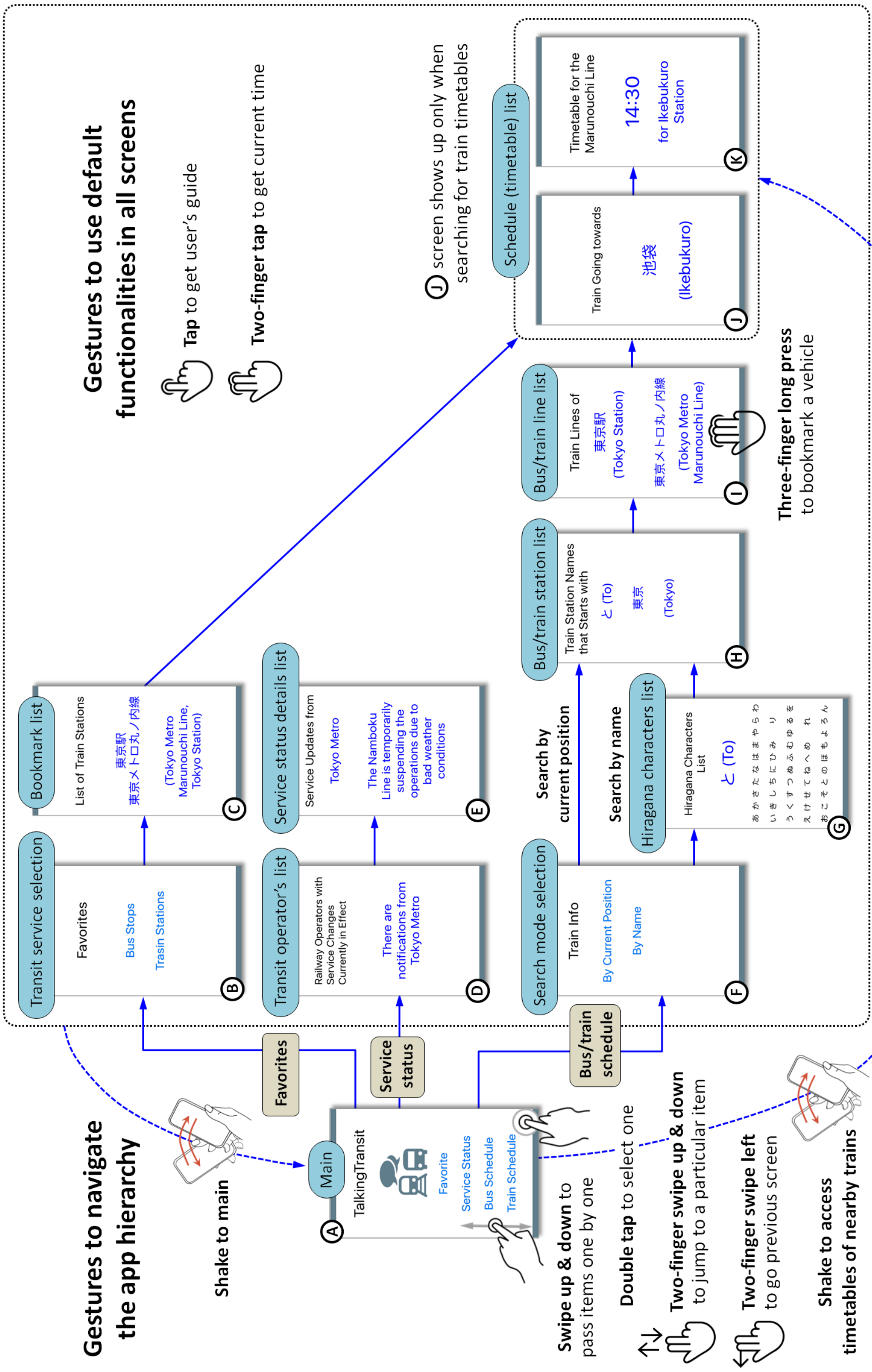


Figure 5.2: Screen transition flow and performed gestures of TalkingTransit. Screens F to K show an example of searching for a train schedule by name.

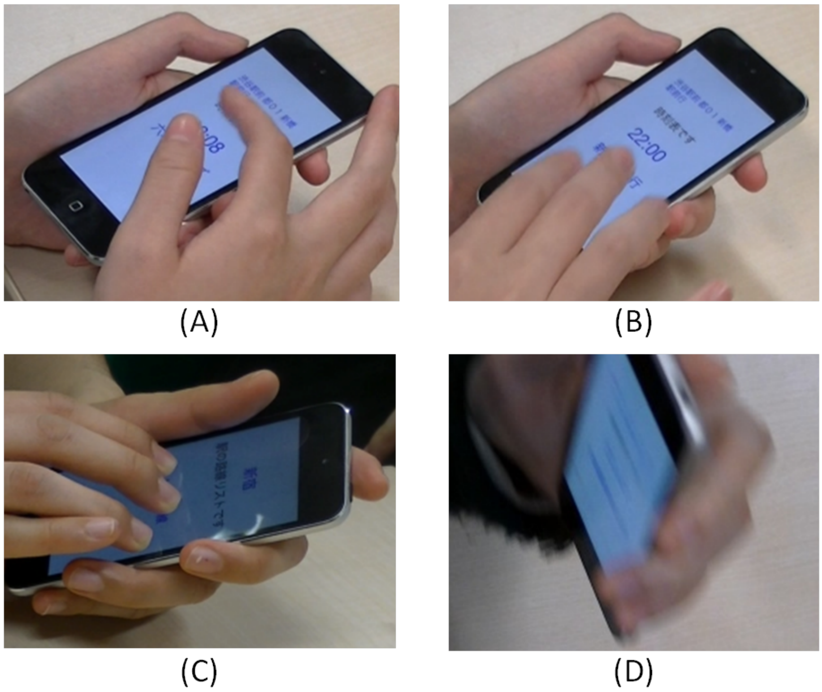


Figure 5.3: Users interacting with TT: (A) Navigating a list by passing items one by one, (B) jumping to a certain position, (C) bookmarking a vehicle, and (D) going back to the main page.

Fast Navigation

Searching by station name can be a challenge because there are more than 1500 bus stops and 1200 stations in Tokyo. To reduce search time to some extent, I categorized stops and stations by the first character of their name. Then, users are required to do the following steps in order: (1) selecting the first character of a bus stop (one of 46 Hiragana characters as shown in Figure 5.2G), (2) selecting the stop in the list of all stops that start with that character (more than 35 items), (3) selecting a line number of the bus, and (4) finding a certain time in a timetable. Step 4 can be overwhelming as well even though the timetable list starts with the next arrival time. Thus, in order to reduce the demand for navigating a long list including steps 1, 2, and 4, I support fast navigation that can be activated by two-finger swipe up and down, while one-finger swipes navigate the items one by one (Figure 5.3A).

For faster navigation, I divided the long list into groups in a way that users can intuitively know how the items are grouped and tagged the first item of each group so that users can jump between the groups by performing two-finger swipes. The grouping strategies are as follows. For step 1, dividing Hiragana syllables into 5 consecutive characters; for step 2, dividing the bus stop list with the same first character into groups where the first two characters (syllables) are the same; and for step 4, dividing a timetable into every hour so that users can jump to the time of the first vehicle per hour (Figure 5.3B).

Shake-to-Main

To reduce the demand for following which page the user is in while navigating the application, I provide "shake-to-main" so that users can immediately go back to the main page from anywhere in TT (Figure 5.3D). I also added bookmark functionality. Users can add a vehicle to the bookmark list in the bus/train line list page by performing three-finger long press (auditory feedback is provided when the gesture is recognized) and remove one from the bookmark list page by performing the same gesture (Figure 5.3C).

Rich Feedback

To let users ensure TT has recognized their inputs as they intended, I provide feedback for every gesture input. When users launch TT, audio information about the existence of any notification from transit operators (e.g., "There are service changes currently in effect, so please check the transit status menu") is offered with vibrotactile feedback. Auditory feedback is provided for every gesture input as shown in Table 5.1. The different sound effects were used, which are somewhat relevant to its gesture (e.g., whooshing sound for swiping gestures) so that it allows users to easily identify gestures they performed.

Table 5.1: Gestures and corresponding auditory feedback.

Gesture	Description	Sound
One- and two-finger swipe up/down	Navigate a list	Page flip
	If last item	Page flip and ding
Two-finger swipe left	Go previous	Whooshing
Three-finger long press	Add a favorite	Short bell ring
	Delete a favorite	Tossing trash
Shake	Go main	Chimes
Double-tap	Select an item	Glass bell hit
Single-tap	Get description	Bounce
None	BLE detection	Message notification

5.2.4 In-station Information

At the same time, I provide in-station information based on user's current position. It may be noted that extracting information through BLE markers, which have a greater range than NFC or RFID, is particularly useful for blind users, as it enables them to simply walk along a concourse or aisle and receive information seamlessly without having to get near any marker. When a smartphone detects a BLE marker's signal, TT receives the marker's unique ID. Then TT retrieves relevant information about the marker – a short description of the place, the floor number, latitude and longitude coordinates of the marker's position, etc. – via OpenData API. Among this information, only the place description is read aloud to the users. If users detect a marker installed near a platform (stairs, elevators, and escalators), they can quickly access schedules of trains on that platform by shaking their smartphones from the main menu (Figure 5.2A). This gesture directly transitions to the screen for selecting a train going towards their destinations (Figure 5.2J). This feature would be very helpful for blind riders to find a right platform

Figure 5.4: Participant receiving information from a BLE marker.

or exit, especially in large stations like Shinjuku station that has 30 platforms and 66 exits. As a part of the open data infrastructure mentioned in Section 5.1, 42 BLE markers were deployed in Shinjuku station near toilets, lockers, and exits. 10 BLE markers were installed at stairs near 8 platforms, 1 toilet, and 1 exit of New South Exit concourse of Shinjuku station where I conducted a field test (Figure 5.4). After refining the prototype as described above, I consulted the blind teacher to ensure usability improvements before conducting a user study. Generally, he was satisfied with the refined design of TT but he commented that users should be able to check the current time when navigating timetables. Thus, I added a feature for getting the current time by two-finger single tap from any page of the application.

5.3 Evaluation

To verify the ease-of-use and usefulness of TT, experimental evaluation was conducted with 11 visually impaired participants in the laboratory. Reading websites using mobile devices with a screen reader like VoiceOver is generally well accepted

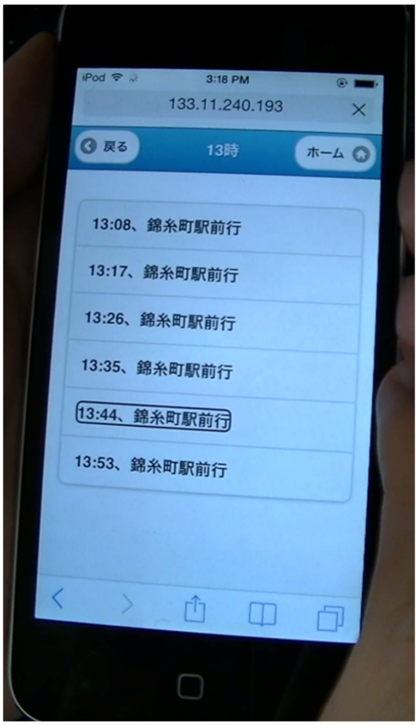


Figure 5.5: The mobile website developed for the user study showing bus schedules.

by visually impaired users. Thus searching a website of timetables using VoiceOver can be a practical solution although it is not created for blind people. Therefore, TT was compared to a mobile website that I developed providing the same information using VoiceOver (Figure 5.5). For a fair comparison as much as possible, the site's structure was designed almost the same as TT's screen flow. In addition, when users access a timetable, the page scroll bar moves near the item (schedule) closest to the current time so that they can quickly find the next vehicle schedule, which gives a similar navigation experience as with TT. The client side was developed using javaScript and jQuery Mobile. PHP was used for the server side. Users can see the site with Safari on iOS mobile devices using VoiceOver.

5.3.1 Participants

11 visually impaired participants (three male and eight female; P1-P11) were recruited for the study. Their degree of visual impairments varied: Eight of them were blind and three had low vision (one of them had only light perception). The average age of the participants was 19.9 (SD = 1.5). To obtain transit information, all smartphone users were using a native application, and the others were using a feature phone's service. Five participants said they also often ask information to train station staffs or other riders. The entire study took approximately 2 hours. The whole process of the experiment was conducted in Japanese, and was video tapped for later analysis.

5.3.2 Procedure

Considering the possible usage scenarios of TT, I asked the participants to perform the following four tasks: (1) Check the service status of the given transit operator, (2) search the nearest bus stop or subway station where you can ride the given bus route or train line, and find the next vehicle schedule 30 minutes from now, (3) bookmark the given vehicle, and (4) find the last departure time of bookmarked vehicle in step 3.

Each participant performed three trials of each task for both systems. The participants used Safari on iPod touch with VoiceOver when browsing the mobile website. Given operator, station, and vehicle for each task was different in every trial. I also set different current position for each session to avoid learning effect. I set the current time as 13:15 for both sessions, but set different locations at my laboratory and Tokyo Dome (about 1.5 km away from my laboratory) respectively. At the beginning of each session, a training session was conducted for all participants until they felt familiar with both systems. The order in which systems were used was counter-balanced across the participants.

At the end of each trial, the participants were instructed to return to the main page. Each trial began when the participant performed a touch input in the main page, and ended when the participant answered the given question. When the participants had difficulty in performing the task, the experimenter assisted them to perform the proper gesture to continue the task. After finishing each task, the participants assessed their subjective mental workload involved in performing the task using the NASA Task Load Index (TLX) [105]. The participants were also asked to answer a questionnaire about each system at the end of each session.

5.3.3 Results

To evaluate ease-of-use, I measured task completion time as a quantitative indicator and used NASA-TLX for assessing mental workload as a qualitative indicator. For evaluating usefulness of TT, a questionnaire survey was conducted. Each participant performed 3 trials of the 4 tasks for each of the two systems, making a total of 264 trials. They completed all trials successfully except for the 5 trials in task 2: for 3 trials, they mistakenly thought that they should find the next vehicle after 13:30 (the participants were supposed to find the next vehicle after 13:45 which was 30 minutes from "now"); for the other 2 trials, they selected right stop (station) for the given vehicle line but not the closest one from the current location.

Task Completion Time

The mean time for 3 trials of a task was considered as task completion time per task. However, regarding the completion time of task 2, I used only the records of successful trials to calculate the mean time (three participants who made the mistakes in task 2 were able to complete at least one or two trials). The average completion time for each task is shown in Figure 5.6. Paired t-test revealed that TT was significantly faster than browsing the website with VoiceOver in all tasks: service status ($p<0.013$), nearby stops ($p<0.001$), bookmark ($p<0.001$), and last departure time ($p<0.035$).

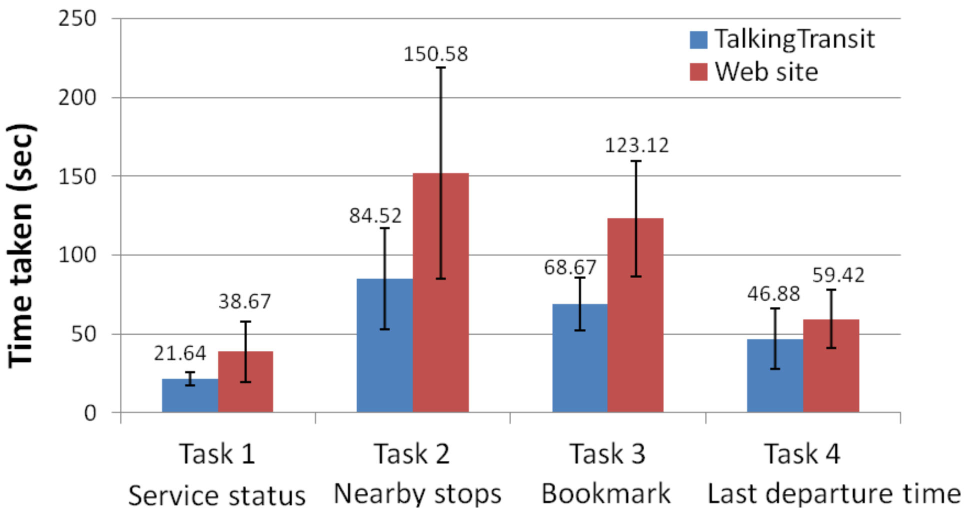


Figure 5.6: Average task completion time in seconds for both systems. Shown with standard deviations.

Subjective Ratings

The participants were asked to rate on six subscales of the NASA-TLX questionnaire for each task: Mental Demands, Physical Demands, Temporal Demands, Effort, Own Performance, and Frustration. Each of these dimensions was rated within a 100-pioints range with 5-point increments. Then I calculated and used an overall workload score (WWL) based on a weighted average of ratings on the six subscales [105]. Table 5.2 shows the average WWL score for each task. Paired t-test found that the participants felt significantly less workload with TT when performing the tasks for service status ($p<0.017$), nearby stops ($p<0.049$), and bookmark ($p<0.027$), showing that TT system could provide simple and intuitive user interfaces.

After each session, the participants were asked to complete a brief questionnaire to investigate the usefulness of the system used in that session. The participants rated five statements using a 5-Likert scale (1=Disagree strongly, 5=Agree strongly). The list of statements, mean values, and significant differences are shown in Table 5.3. A Wilcoxon test found that the following items were significant:

Table 5.2: NASA-TLX WWL scores (0–100) for each task (Mean, SD). Lower means less workload.

Tasks	TalkingTransit	Website
1. Service status	22.85 (15.16)	35.64 (19.94)
2. Nearby stops	34.25 (18.95)	41.81 (22.51)
3. Bookmark	30.76 (19.51)	40.19 (18.00)
4. Last departure time	29.08 (18.19)	34.96 (16.68)

Table 5.3: Questionnaire results (Mean, SD). Starred items were rated significantly higher for TalkingTransit.

Statement	TalkingTransit	Website
Useful*	4.82 (0.40)	4.09 (0.83)
Easy to use*	4.45 (0.69)	3.55 (0.93)
Easy to learn	4.82 (0.40)	4.27 (1.01)
Felt in control	4.45 (0.69)	3.82 (0.98)
Would use on smartphone*	4.64 (0.67)	3.64 (1.12)

Useful ($p<0.03$), *Easy to use* ($p<0.03$) and *Would use on smartphone* ($p<0.01$). There were no significant differences regarding both *Easy to lean* and *Felt in control* showing that the participants could get used to both interaction techniques without much difficulty. These results indicate that the participants felt more comfortable and were able to get familiar with TT easily. In addition, I asked the participants to choose the easier to use system for each task. Generally, VoiceOver users were either positive or neutral about the website. However, when they asked to indicate their favorite of the two systems, all participants preferred TT to the website as shown in Table 5.4.

Table 5.4: User's preference of both systems for each task. The number of users choosing each option is displayed.

Tasks	TalkingTransit	Website	Neutral
1. Service status	7	1	3
2. Nearby stops	9	1	1
3. Bookmark	9	2	0
4. Last departure time	5	4	2
Entirely	11	0	0

User Feedback

Overall, the participants found TT was easy-to-use because of its simple and intuitive user interfaces. Even participants who have not experienced touchscreens were able to easily use TT. P11 stated, *"I was reluctant to own a smartphone, but after using TT, I found it easy."* The participants commented positively about TT's ability to navigate fast and provide rich auditory feedback. P7 said, *"It was quite good to get a sound notification upon reaching the end of a list."* 3 Participants particularly liked "shake-to-main." A number of users said they anticipated public release of TT so that they could use it in their daily lives.

I also investigated additional features for improving the user experiences with TT. P3 and P11 suggested that jumping to the first and last items with three-finger swipes would be useful. Some had difficulty with long-press gestures. The average number of performed gestures to add a bookmark, which requires a long-press gesture, was 1.5. P1 and P3 commented that it would be better to change the bookmark gesture from three-finger long-press to double-tap. Moreover, the participants mentioned that providing additional information such as a list of all bus stops on a route, surrounding point-of-interests, and fare information would be helpful.

5.4 Discussion

The user study results indicate that TT allows users to access desired transit information quickly and easily. It was encouraging that users could intuitively understand the fast navigation feature. All users used this feature every time they encountered a long list. I observed that users seemed to feel a sense of control and freedom because of the shake-to-main feature that provided them an "emergency exit" through which they could escape from an unexpected situation and could start over like pressing a reset button. Regarding feedback, users preferred implicit manner (e.g. tossing trash sound for deleting a bookmark) rather than explicit manner (e.g., "This is the first item of a list"). Some users seemed to be annoyed at the verbose feedback of VoiceOver. Three users said, *"VoiceOver talks too much!"* Although VoiceOver's gestures allow users to use various applications, it may be difficult to provide user interfaces specialized for each application [106]. This study demonstrated that utilizing smartphone's built-in sensors and customized touch gestures, mobile applications can be made highly accessible to blind users. Although all users thought TT was better than the site, some iPhone users liked the mobile website design as well. Thus, it would be desirable to let them choose the system according to their preference. Providing more choice can bring higher satisfaction for each visually impaired person because their individual ability and attributes are varied (e.g., touchscreen experiences, age, degree of visual impairments). Furthermore, TT can benefit elderly citizens who might have difficulty reading text in the small screen space.

In addition, to identify the feasibility of TT in the field, I conducted a user test at Shinjuku station with three blind participants (Figure 5.4). The test included a 15-minute training, 20-minute test, and 15-minute interview. They walked about 50 m with an experimenter, passing the toilet, exit, and two platforms in order along around New South Exit. Users could generally get information from BLE markers. All users were particularly pleased with getting schedules by shaking the device near a platform. They also commented that it would be more helpful with additional information about which train comes from which side of the platform. Because this information was not available in the current Open Data infrastructure, I gave feedback about information needs to the open data project community for improving the breadth and quality of data. One participant mentioned that he wanted BLE markers to be deployed anywhere in the street so that he can go shopping and enjoy traveling.

5.5 Summary

In this chapter, I have proposed a public transit information system, Talking-Transit—a smartphone-based location-aware system for visually impaired passengers utilizing open data and IoT technology. Unlike the existing systems, TT does not require any additional hardware and provides access to information anytime anywhere, which can offer the improved accessibility of public transit for blind travelers by carrying a smartphone only. The user study demonstrated that the interaction design including fast navigation, shake-to-main, and rich auditory feedback could give better user experience than existing approaches that mainly use VoiceOver, allowing blind users to obtain mass transit information in Tokyo quickly and easily. This study also revealed the potential of open data that foster cost-effective application developments for blind individuals.

Chapter 6

StaNavi: A Location-aware System for Aiding Indoor Navigation

This chapter[18] addresses the challenge of navigating indoor environments, especially in public places where blind people often face difficult situations such as in large crowds, convoluted paths, or wide-open spaces. The chapter introduces StaNavi, a navigation system that provides turn-by-turn voice directions inside a large train station using Bluetooth Low Energy technology. To facilitate indoor navigation of blind travelers under such real-world conditions, in this chapter, a set of features that mainly includes rich navigational cues and interfaces for one-handed use is provided.

6.1 Motivation

Individuals who are blind and visually impaired face challenges when traveling to unfamiliar places. For outdoor navigation, GPS-based commercial smartphone apps, such as BlindSquare [28] and Ariadne GPS [29], have been deployed and gained popularity among blind users thanks to built-in accessibility features on mobile devices (e.g., Apple's VoiceOver). However, navigating indoors where GPS is ineffective remains largely unsolved, especially for large and complex public spaces

[18]Portions of this chapter were previously published in [107]

(shopping malls, train stations, airports, etc.) that are a part of our everyday lives. In particular, transportation settings are a critical domain that must be addressed, because public transit, such as subways and trains, plays a key role in the lives of people with visual impairments by offering access to employment, education, shopping, etc.

The goal of this study is to propose a practical working solution, especially the design and implementation of a client application, which enables blind people to move around independently inside large and complex public buildings, such as train stations and airports, using a readily available positioning technique and modern smartphones without the need for users to carry additional devices. The leveraging proximity detection capability of BLE technology and a smartphone's built-in compass, StaNavi allows blind travelers to determine their current location and obtain turn-by-turn instructions to a destination. I implemented and deployed it at Tokyo Station, one of the busiest train stations in the world, which has more than 400,000 passengers and 3,700 trains per day. Indeed, facilitating a useful and reliable navigation system for blind travelers inside such a large and complex space involved several practical problems, such as fluctuations in proximity and compass readings and factors that reduce their sense of orientation commonly found in a train station – high noise levels, bumping into people, wide-open spaces (e.g., lobbies), etc. [108].

I iteratively designed StaNavi in collaboration with blind users in order to provide features that can enable them to cope with such challenging conditions and reach their destinations. These features mainly include: (1) Simple and one-handed user interfaces that can be performed easily by users holding a cane, (2) a route overview that summarizes the total route in terms of the main areas and turns on the route, and (3) navigation cues that describe convoluted paths and open spaces to help blind people orient themselves. For example, StaNavi provides spatial relations among objects in a large space around ticket gates – facing south, there are ticket gates in front, stores on the right side, a ticket office on the left side, and a corridor leading to platforms at the back.

6.2 StaNavi System

StaNavi is a BLE-based guidance system that runs on smartphones that allows blind users to obtain their current location, points-of-interest, and turn-by-turn directions to a destination in Tokyo Station. In the remainder of this section, I present an iterative design process and describe the system and its implementation in detail, highlighting the novel features specialized for the intended users.

6.2.1 Design Process

The current version of StaNavi has been developed via four iterations of implementation, testing, and refinement in collaboration with the blind teacher. The first prototype only provided automatic notification of the user's current location and some relevant information when detecting a beacon in order to identify how blind users experience the BLE-based information service and investigate the feasibility of navigation functionality. Two blind people (one female and one male) – both of whom are not smartphone users and use Tokyo Station five to 10 times a year – tested the prototype in the field for about an hour. They could easily understand the concept of our system and provided comments about the representation of surroundings and importance of knowing the directions they are facing.

Based on my observations and feedback from the initial prototype test, I developed the navigation functionality and consulted the blind teacher. He was generally positive about the functionalities of StaNavi and gave advice on user interfaces and route descriptions for ensuring effective navigation. After refining the prototype, I again asked the blind teacher to thoroughly test the prototype system in the field by thinking aloud while navigating to several destinations. This test allowed me to identify and fix potential usability problems before conducting a user study. Through the design process described above, I have extracted the following user requirements that can realize independent navigation for blind travelers in a large train station:

- One-handed operation should be allowed because the other hand is often

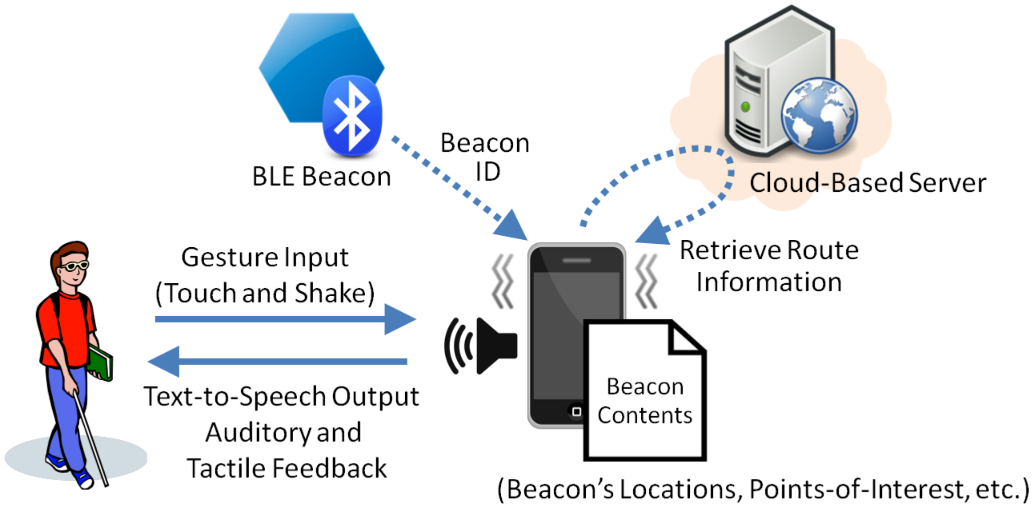


Figure 6.1: StaNavi system architecture.

occupied with holding a cane or guide dog.

- Destination input should be performed with confidence and certainty. Speech input may not work well in noisy places, and text input may be hard to perform for users unfamiliar with smartphone use.
- Information on the current location and direction faced should be accessed directly and easily anywhere on the app.
- Overall understanding of the space should be supported.
- Route descriptions should be as detailed as possible, particularly for convoluted paths or open spaces.
- Warnings of poor signals and deviations from a route should be provided; in case of signal loss, the last position should be repeated.

6.2.2 System Overview

As illustrated in Figure 6.1, the StaNavi system architecture consists of a user smartphone, BLE infrastructure, and a cloud-based server that calculates the route from origin to destination. The BLE infrastructure allows users to determine their approximate locations in close proximity to points-of-reference (POR) in a station (e.g., platforms, exits, toilets, elevators). When a smartphone detects a BLE beacon's signal, StaNavi receives the beacon's unique ID, which can be used as a key to access relevant information (a location name, nearby POI, directions to adjacent beacons, etc.) from contents stored by the app. StaNavi takes advantage of BLE beacons installed on signs and the ceiling that have been experimentally deployed by East Japan Railway. StaNavi runs on a BLE-enabled smartphone. The user interacts with StaNavi using gestural inputs and TTS output with auditory and tactile feedback. Based on user requirements described in the previous subsection, the current version of StaNavi includes three main functionalities.

- **"Where am I":** Users can obtain direct access to their current location and four directions that are faced – the Marunouchi side, Yaesu side (both are famous districts in Tokyo), South, and North.
- **"Free Roam":** Users can obtain automatic notification of their current location when they are within range of PORs. The user also can get detailed information if available. This functionality allows users to explore and learn about the surrounding facilities of the train station.
- **Navigation:** Users can obtain turn-by-turn directions from their current location to a desired destination. To guide the user to a destination, StaNavi introduces PORs (beacons) on the route in order. Every time the user's presence is detected in the vicinity of a POR, StaNavi provides instructions on how to get from the current POR to the next POR. The user also can obtain navigation cues if they are available.

6.2.3 Information Provided for Easy Navigation

Recognizing the importance of meeting information needs for independent traveling in a large and complex environment, StaNavi provides rich information that helps blind people understand a space and navigate to and from places with ease. Table 6.1 shows an overview of the information provided by StaNavi.

Current Location

Spatial representation of a large environment is generally organized into a hierarchical structure in human memory [109] (e.g., the location of a desk in a room on the first floor of a building). To let blind people intuitively and easily determine where they are in such a large station, StaNavi presents their current location in two levels of a hierarchy (e.g., "at the women's toilet in the Marunouchi area"):

- *Area*: Meaningful chunks of a train station, such as passages, lobbies, and around gates. Figure 6.2 shows the areas defined in StaNavi.
- *Point-of-Reference*: Specific locations that can be important reference points when navigating the station and can be identified by beacons (platforms, gates, escalators, elevators, staircases, intersections, the entrances of facilities such as toilets and stores, etc.). A total of 55 PORs were used in StaNavi (Figure 6.2 shows some PORs).

Detailed Information

In "Free Roam" mode, StaNavi offers detailed information about surroundings that can raise a blind person's spatial awareness. This information includes the width of the passage, the number of staircase steps to the platforms, the presence of facilities such as elevators and escalators in the passages, two relative directions (left and right) when standing against the staircase to the platforms, and the floor level difference in the lobby.

Table 6.1: Overview of the information provided by StaNavi.

Functionality	Information Type		Example Message
Free Roam	Current Location		“You are near the Tokaido line platform in the Central Passage area. Detailed information exists.” “The passage width is approximately 17 m. 34 staircase steps to the platform.”
	Detailed Information about Surroundings		
Navigation	Route Instruction	Current Location	“You are near the Shinkansen South Transfer Gate in the Central Passage area. Turn right near the 7th platform and go 25 m to the south. 50 m left to the destination. Navigation cue exits.” “Access to the Shinkansen South Transfer area is between the 7th and the 9th platform.”
		Direction to Next POR	
		Distance to Next POR	
		Distance Left to Destination	
		Navigation Cue	
	Route Overview		See Figure 6.4.
	Warning of Poor Signals and Deviations		”You may be off the route. You can reroute from the Navi Main menu.”

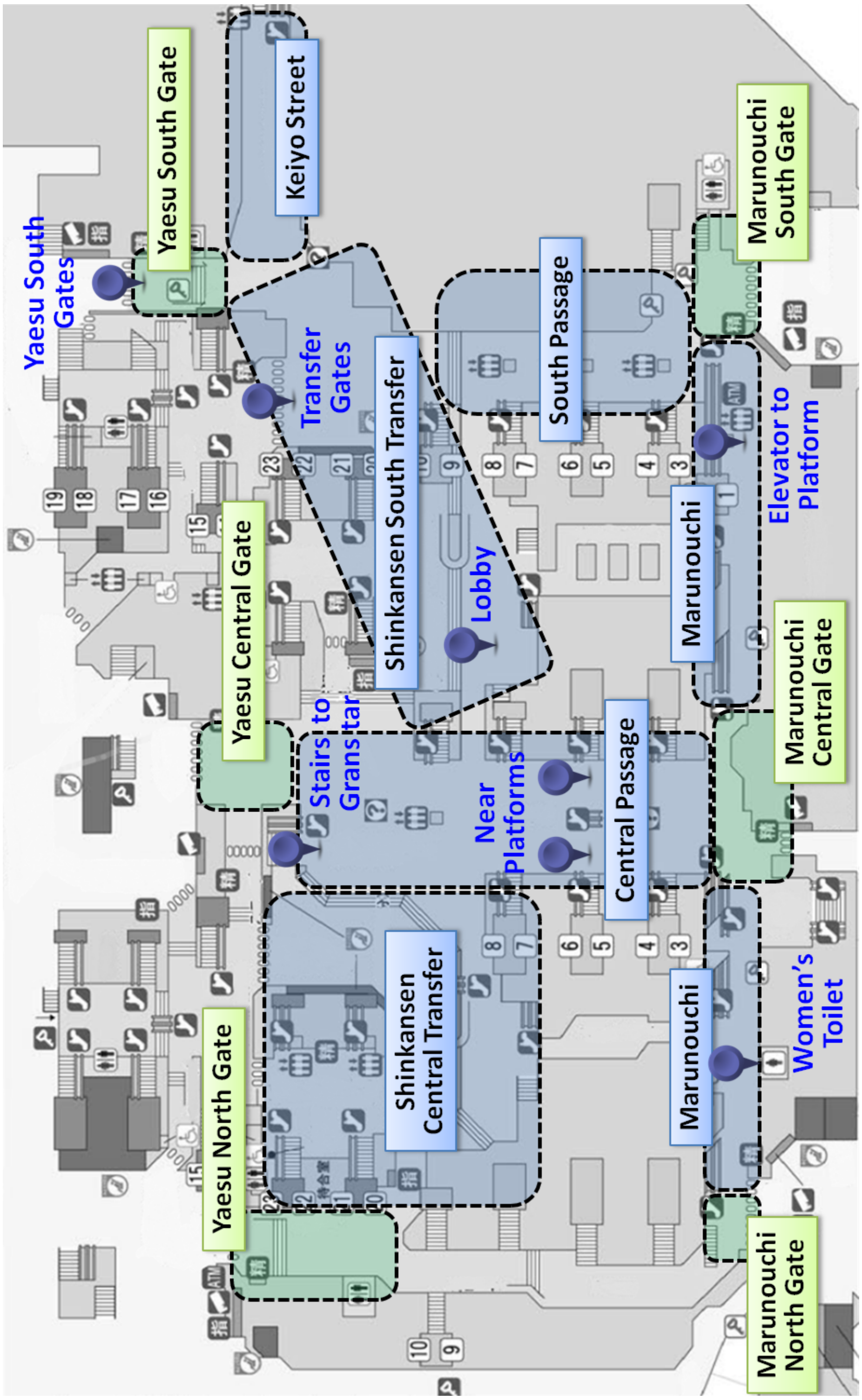


Figure 6.2: Areas defined in StaNavi and some points-of-reference.

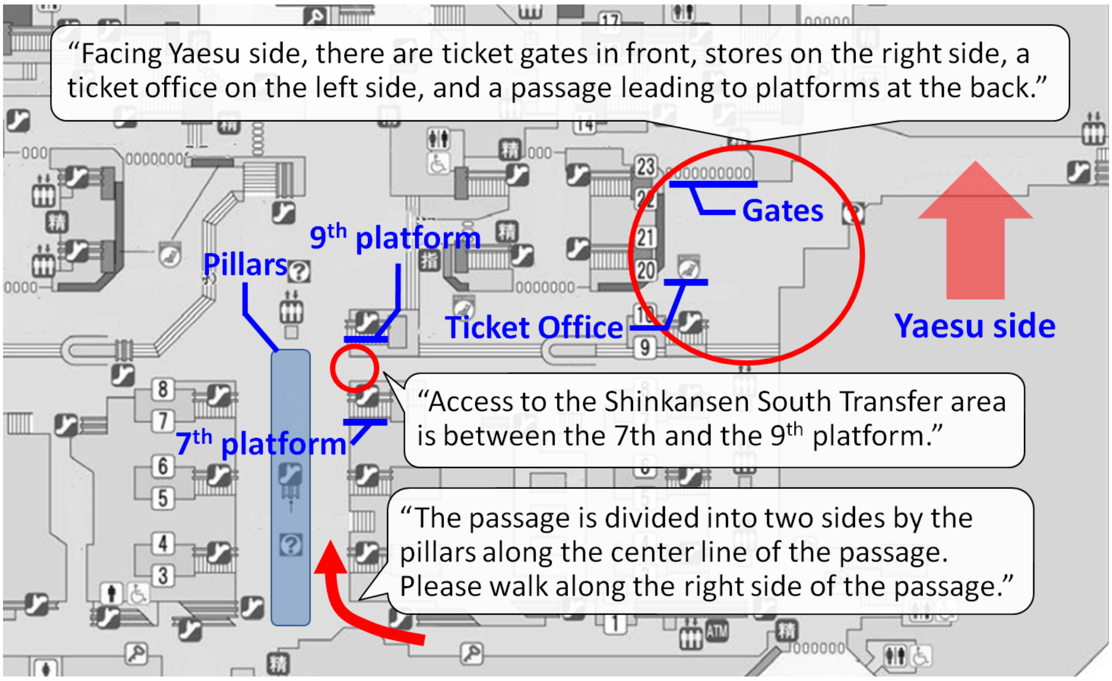


Figure 6.3: Examples of navigation cues provided by StaNavi.

Navigation Cues

A default instruction message (Table 6.1) notifies the user when detecting a POR on the route. Regarding the description of direction changes, I used four absolute directions (Marunouchi side, Yaesu side, South, and North). For the diagonal directions, a clock-positioning system was used (e.g., "facing south, change direction to 2 o'clock"). In addition to the default instruction, StaNavi offers navigation cues to help blind travelers orient themselves in convoluted paths and open spaces. This information mainly includes the exact position of a byway, the description of a path that goes through wide passages, and the description of spatial relations to surroundings in a large space as shown in Figure 6.3.

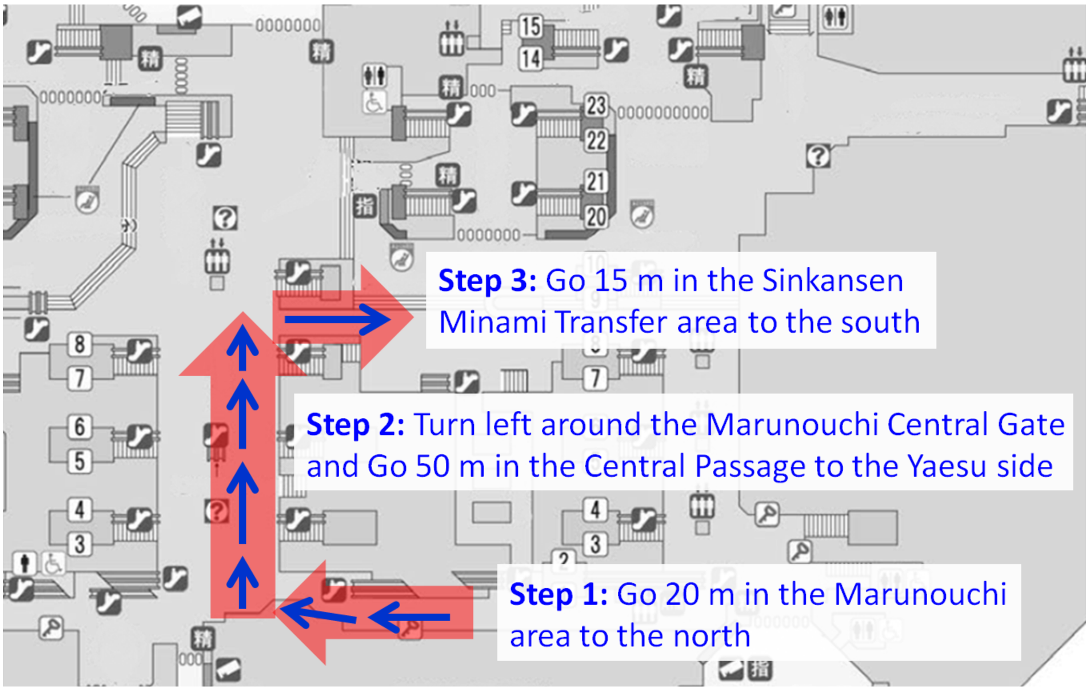


Figure 6.4: An example of route overview information. Red arrows indicate area-level steps that summarize the total route, and blue arrows show turn-by-turn directions.

Overview of the Route

The importance of route preview information has been recognized for helping blind people establish an overall cognitive map of an area [110]. In the design process, I have also identified that gaining a route overview in advance may enhance user confidence in the navigation process. StaNavi provides a route overview in a way that divides the total route into area-level steps, as shown in Figure 6.4. Audio information on the route overview begins with the total distance and the number of areal-level steps to a destination, and is then followed by area-level steps in order.

Warnings of Poor Signals and Deviations

If the user is outside of all beacon regions, StaNavi provides a warning and reads aloud the last known position. For example, "currently, you are outside of the beacon region. Your last known location is the women's toilet in the Marunouchi area." If the user deviates from the route (detecting a beacon that is not included on the route), StaNavi provides a message shown in Table 6.1.

6.2.4 User Interface

To enable users to control StaNavi with one hand, a set of simple touch gestures was used—one-finger swipe up and down, two-finger swipe left, one- and two-finger tap, double-tap, and long press—and a shake gesture, as shown in Figure 6.5. Users can perform the touch gestures anywhere on the screen without being restricted by a fixed layout so that it allows one-handed thumb use of StaNavi. Note that the command assigned to the two-finger swipe left gesture can be substituted by a shake gesture.

I took advantage of interfaces that have been designed for TalkingTransit [24] that have been shown to be useful for blind users in navigating the app's information hierarchy. When starting StaNavi, users can choose one of the two modes according to their usage scenarios: "Free Roam" mode to explore the station and "Navigation" mode to reach their destinations (Figure 6.5A). In "Free Roam," StaNavi automatically reads aloud the current location with both auditory and tactile feedback when detecting a beacon (Figure 6.5B). Users can obtain detailed information by swiping down. In "Navigation," users can input a desired destination by selecting one of three categories (Figure 6.5C), then a specific location (Figure 6.5D). This hierarchical selection allows users to easily input their destinations with confidence and certainty. After selecting a destination, StaNavi transitions to the Navi Main screen (Figure 6.5E) and sends a request to the server for route information. When the download is finished, users will receive a voice notification that says, "The route information has been downloaded." In the "Navi

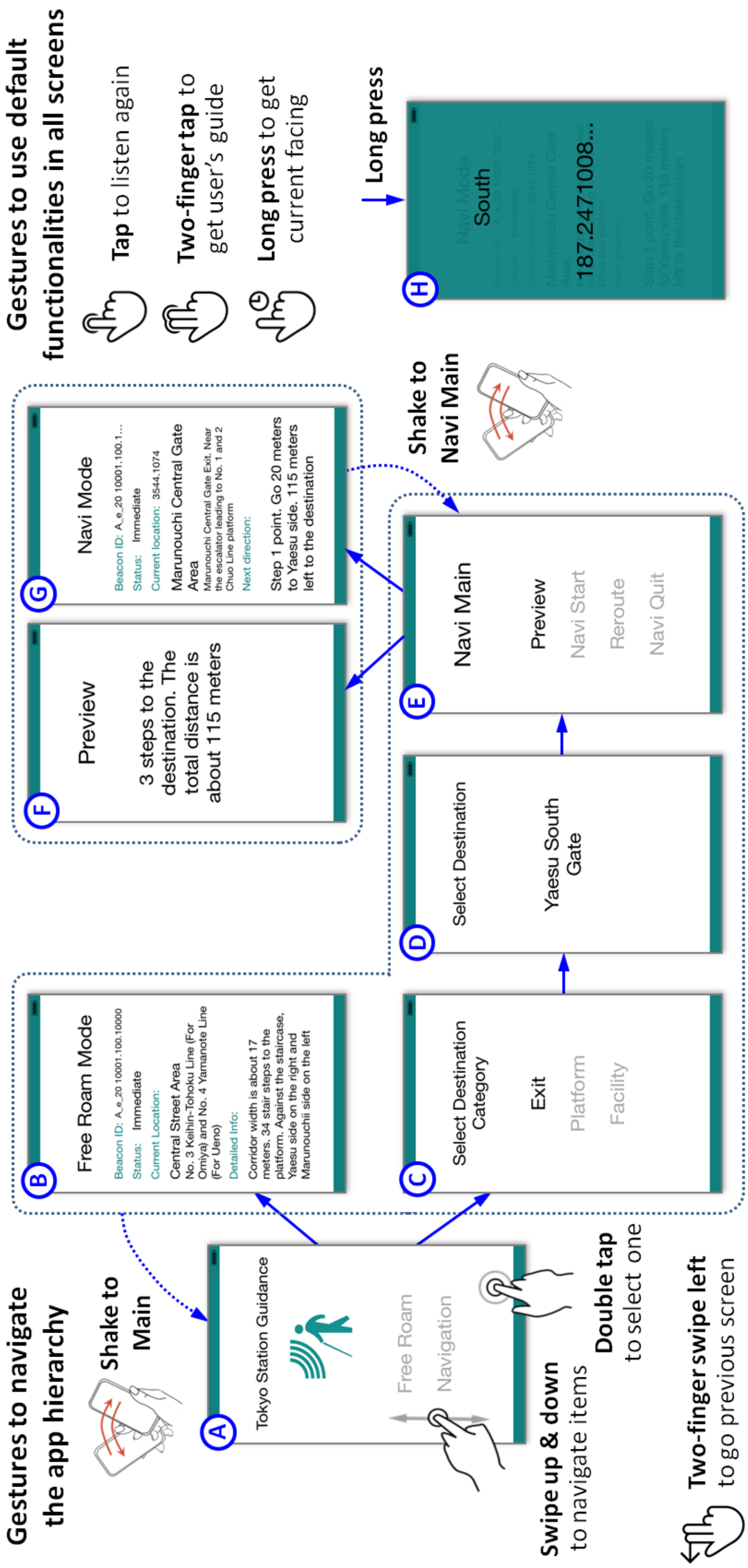


Figure 6.5: Screen transition flow and performed gestures of StaNavi.

Main" menu, users can select one of four items: Preview, Navi Start, Reroute, and Navi Quit (Figure 6.5E). In "Preview," users can access the route overview information, as described in the previous subsection, by performing swipe up and down gestures that read aloud each area-level step (Figure 6.5F). Selecting "Reroute" will re-request route information from the current location to the selected destination without a screen transition. When the download is finished, a voice notification will be produced. Users can go back to the top screen by selecting "Navi Quit." Selecting "Navi Start" transitions to the "Navi Mode" screen (Figure 6.5G). In "Navi Mode," StaNavi reads aloud turn-by-turn instructions with both auditory and tactile feedback whenever it detects a beacon. Users can access navigation cues by swiping down.

To allow users to determine the direction they are facing, I adopted a pointing-based interaction that enables users to retrieve relevant information on a desired direction by pointing smartphones in that direction. This method has been identified as a useful interaction technique for letting blind users indicate their areas-of-interest [96] and has been adopted by commercial apps, such as BlindSquare [28]. In all screens, tapping and holding (long press) with one finger brings up the screen (Figure 6.5H), which reads aloud four absolute directions of Tokyo Station according to the direction that the smartphone is pointing toward. Figure 6.6 shows a participant determining his directions. Users can access their current location in most screens (Figure 6.5C, D, E, and F) by one-finger tap; in "Free Roam" and "Navi Mode," a one-finger tap repeats the last description. A two-finger tap anywhere in the app provides guidance on the use of each screen. For example, "Destination category selection screen. Swipe up and down to navigate categories and double-tap to select." Additionally, users can go back to the top screen or "Navi Main" directly by shaking their smartphones, which can be helpful for not only one-handed operation, but also for escaping from an unexpected situation and starting over, like pressing a reset button.

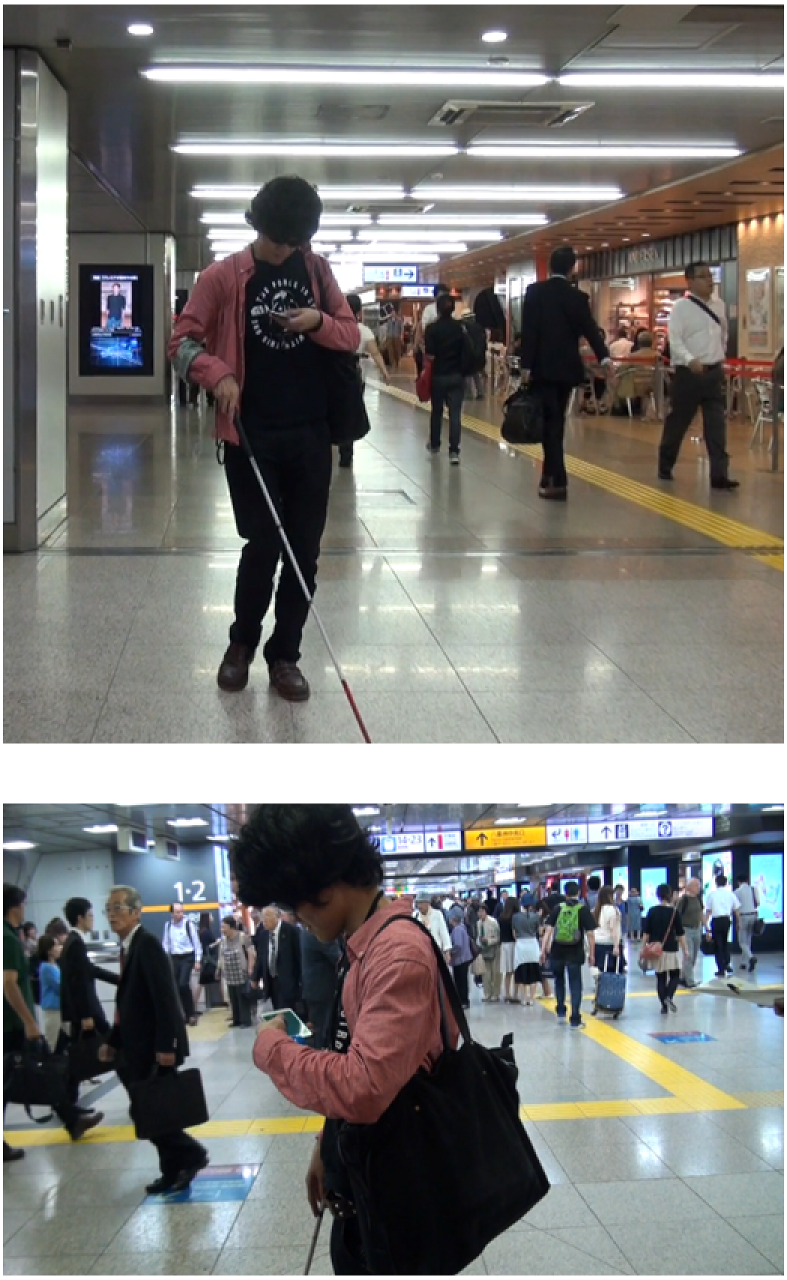


Figure 6.6: Participant interacting with StaNavi: Determining directions by pointing the device.

6.2.5 Implementation

StaNavi is implemented on iOS 8 platform and currently runs on BLE-enabled Apple iPhone 5S and 6. StaNavi makes use of the device's built-in accelerometer and gyroscope sensors to detect shake-motion events. I also used AVSpeechSynthesizer API included in iOS7 and higher platforms to offer TTS.

In the early development stage, after testing the prototype several times in the field, I found that proximity estimations were unstable and fluctuated greatly in some areas where beacons were installed densely, which could cause blind people to feel anxious and confused. Therefore, to improve beacon detection stability, I empirically adjusted which beacons are used for StaNavi. In particular, I grouped adjacent beacons to represent the same POR and ignored some beacon signals.

The overall navigation process of StaNavi is shown in Figure 6.7. When a user selects a destination, StaNavi communicates with the server via HTTP. The server calculates the shortest path using Dijkstra's algorithm and returns relevant information in XML format. StaNavi then parses it and makes route instructions for reaching the destination. The system scans for beacons every 2 seconds, updates the current location, and provides proper instructions until the user detects the last beacon on the route. If StaNavi detects a beacon not on the route, it waits for 30 seconds before sending an off-route message in order to ensure user deviation.

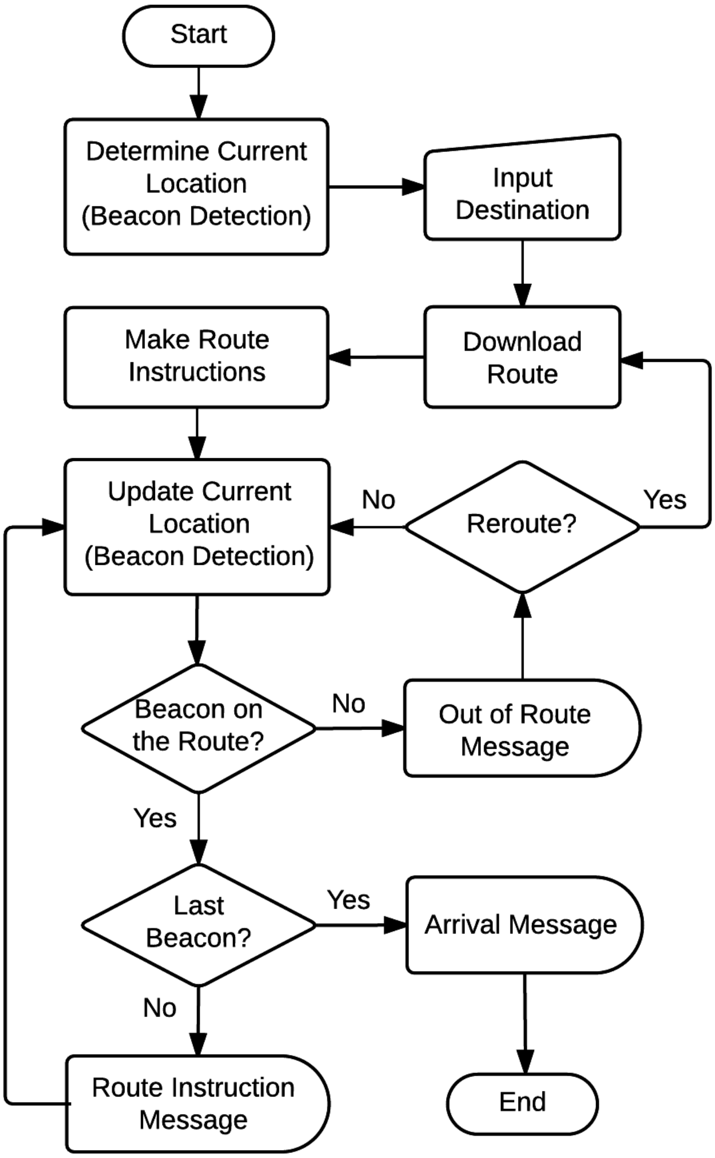


Figure 6.7: Flowchart of StaNavi's navigation process.

6.3 Evaluation

Experimental evaluation of StaNavi was conducted in the field to verify (1) the effectiveness of enabling people with visual impairments to independently navigate a complex train station and (2) the usefulness of features for supporting easy navigation. StaNavi was tested in the context of actual use in day-to-day life, rather than in a carefully controlled environment. The field was on the first floor of Tokyo Station where a total of 100 beacons were deployed. The experiment was conducted over four days, and eight visually impaired people participated in total. User interactions with StaNavi—detected beacon identifiers, use of functionalities such as direction determination and rerouting, and screen transitions—were recorded with timestamps in a text (.txt) log file.

6.3.1 Participants

I recruited eight visually impaired participants (three male and five female; P1-P8) with an average age of 24.4 (SD=7.3). All of them were totally blind except one who had low vision. Six participants were current iOS users. Seven participants go out and use public transit every day, and all of them use a white cane. Additionally, I asked about the frequency of Tokyo Station use; four participants use it at least once a year; two participants use it at least once a month; one participant uses it at least once a week; and one participant rarely uses this station. The entire study took approximately 2 hours. The whole process of the experiment was conducted in Japanese and was videotaped for later analysis.

6.3.2 Procedure

Considering key locations inside Tokyo Station, the participants were asked to perform the following four navigation tasks as shown in Figure 6.8: (1) Route 1: From Marunouchi Central Gate to Shinkansen South Transfer Gate, (2) Route 2: From Shinkansen South Transfer Gate to Keiyo Street, (3) Route 3: From Keiyo

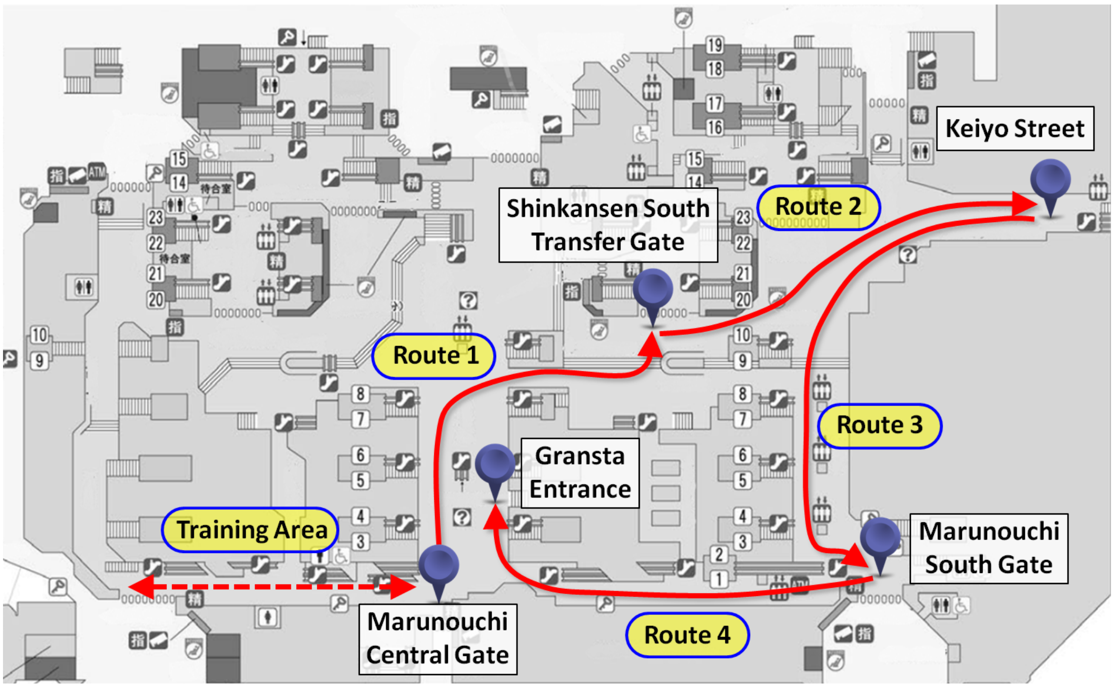


Figure 6.8: Maps showing given destinations and the routes that the participants took. The total distance is about 600 meters.

Street to Marunouchi South Gate, and (4) Route 4: From Marunouchi South Gate to Gransta Entrance.

To help the participants develop a clear impression of the usefulness of the route overview information, I asked them to use the preview functionality before performing the first two tasks or last two tasks. Half of the users followed Routes 1 and 2 with route previews and followed Routes 3 and 4 without route previews; the other half did the opposite. At the beginning of the tasks, I, along with colleagues, provided all participants with training and practice until they felt familiar with StaNavi in the training area shown in Figure 6.8. For the users' convenience, a neck strap for smartphones was provided. A portable speaker was also offered, which directly plugged into the earphone jack of the smartphone to amplify sound, because Tokyo Station is often very noisy as shown in Figure 6.9.

For each task, the participants were led to the start location where the exper-

Figure 6.9: A neck strap and speaker used in the experiment.

imenter would check the detection of the first beacon on the route and hand the device to them. Each task began when the participant selected "Navi Start," and ended when the participant said he or she had reached the given destination. I especially asked the participants to navigate to the given destinations as independently as possible utilizing the reroute functionality if needed. Two or three experimenters monitored the participants' safety at all times, but provided assistance only in dangerous situations. The experimenter intervened in the case of a user's request for help; outside of the experiment area (e.g., when entering stores, platforms); and outside of the range of the beacon's signal. In those cases, the experimenter led the users to the next POR location on the route so that they could continue the task. After finishing all tasks, the participants were asked to answer a questionnaire to collect qualitative experiences. Figure 6.10 shows the participants navigating the station using StaNavi.

Figure 6.10: Participants interacting with StaNavi: Listening to route instructions while walking.

6.3.3 Results

To evaluate the effectiveness of enabling independent navigation, I analyzed data on task completion rate, task completion time, deviation, and help-seeking situations, as shown in Table 6.2. For evaluating the usefulness of features provided by StaNavi, a questionnaire survey was conducted.

Task Completion Rate and Time

The task was considered a success when the users detected the last beacon on the route, which means that they were able to navigate within 2 meters of the target location. Each participant performed four tasks, amounting to a total of 32 trials. They completed all trials without failure, taking alternative paths (rerouting) appropriately. The average task completion time is shown in Table 6.2. Considering the time taken to reroute, which particularly occurred more than one time per trial in Routes 1 and 2, the time taken for navigating to the given destinations would be acceptable.

Table 6.2: Data collected for each route.

	Route 1	**Route 2**	**Route 3**	**Route 4**
Success Rate	1.00	1.00	1.00	1.00
Route Length (m)	153	114	171	138
Average Time (s)	550	466	403	301
Time Stdev	285	145	79	58
Number of Deviations	17	10	5	6
Average Number of Deviations	2.1	1.3	0.6	0.8
Number of Interventions	4	5	1	4

The Number of Deviations

The number of times that rerouting occurred—the participants regarded themselves as off the route and tried to take an alternative route—was counted as the number of the user's deviations. Particularly, the participants often deviated from Routes 1 and 2 as shown in Table 6.2, which include some challenging points where people with visual impairments may have difficulty orienting themselves. A total of 38 deviations occurred during 32 trials, requiring an average 1.2 times of rerouting per trial. However, this would be acceptable considering the use of StaNavi under realistic conditions, such as magnetic inference on the compass readings and fluctuations of proximity readings.

The Number of Interventions

During the whole experiment, a total of 14 interventions were provided (Table 6.2) for the following reasons: Nine times to secure participants' safety when they entered stores and platforms; four times to resume tasks when they were outside of beacon range; and one time to help a user who accidentally launched another app. Note that there were no explicit requests for help from the participants. Furthermore, the critical situation of users going out of all beacon regions and not being able to return to the navigation process occurred only four times, indicating the participants could generally reach their destinations independently.

Subjective Ratings

After finishing all tasks, the participants were asked to complete a brief questionnaire to investigate the effectiveness of the system. They rated five statements using a five-point Likert scale (1=Disagree strongly, 5=Agree strongly). The list of statements, mean values, and standard deviations are shown in Table 6.3. The results confirm that StaNavi was generally well accepted by the participants. In particular, all participants found the system very useful and would strongly like to use it on smartphone, indicating StaNavi has great potential for large-scale adoption, given that six out of eight participants were smartphone users. The participants were also asked to rate each functionality in terms of usefulness and ease-of-use using the same five-point Likert scale. Overall, the participants found the given functionalities very useful as shown in Table 6.4. On the other hand, perceived ease-of-use was somewhat varied by functionality, although generally higher. In particular, some users had difficulty using the direction determination functionality because they were required to keep pressing the screen while pointing the device in desired directions.

Table 6.3: Questionnaire results on overall system effectiveness.

Statement	**Average rating**
Useful	4.88 (0.35)
Easy to use	4.00 (0.93)
Easy to learn	4.38 (0.92)
Felt in control	4.25 (0.71)
Would use on smartphone	4.75 (0.46)

User Feedback

Overall, a number of participants liked the system. P6 said, *"I appreciate that I could reach my destination just as I wanted, without anxiety."* P5 stated, *"It would be helpful to understand a train station I have never been."* Some participants commented positively about features provided by StaNavi. Three users were particularly fond of the route overview information. P5 said, *"It was useful in that I could predict the route in advance. This made me feel comfortable even if I might forget it during navigation."* P2 stated, *"Current location was easy to understand because of its hierarchical representation."* P3 commented that the navigation cues were helpful.

Table 6.4: The ratings on Usefulness and Ease-of-Use for each functionality (Mean, SD).

Functionality	**Usefulness**	**Ease-of-Use**
Free Roam	4.88 (0.35)	4.63 (0.52)
Navigation	4.75 (0.46)	4.00 (0.76)
Destination Selection	4.88 (0.35)	4.88 (0.35)
Current Location Check	4.88 (0.35)	4.75 (0.46)
Direction Determination	4.88 (0.35)	3.13 (1.25)
Preview	4.50 (0.93)	4.88 (0.35)
Reroute	4.75 (0.46)	4.25 (1.04)

Meanwhile, five participants found the clock-positioning system for heading in diagonal directions difficult to understand. Two participants suggested that although it increases the number of turns they should make, it would be better to use only three types of navigational instructions: Turn left, turn right, and go forward. Additionally, some participants provided recommendations to refine the interface for determining directions. For example, P3 commented on the need for a user's guide for using the direction determination functionality, such as "Rotate the device until the compass reading stabilizes." In addition, several participants mentioned the need for additional information on surrounding shops and facilities. Moreover, three participants suggested that route guidance based on tactile paving would be helpful.

6.4 Discussion

It was encouraging that all participants achieved independent navigation to given destinations successfully. They actively used navigation cues and rerouted themselves by appropriately considering the concept of proximity sensing—it only let them know they were within close range of a specific location, rather than a precise position—and unstable sensor readings that might have sometimes provided incorrect information on their current location or direction faced. I observed a number of deviations of users in Routes 1 and 2. The beacon installed at the turning point of Route 1, the byway located between two platforms, was often detected after users passed that point because it was difficult to pinpoint such a byway with a relatively narrow width in a timely manner, although BLE technology generally provided sufficient accuracy in localization. Route 2, where users were instructed to proceed in a diagonal direction, was tricky to navigate for most participants. One participant stressed the need for continuous tactile or auditory feedback to orient herself and keep walking in the right direction in the case of moving diagonally. Further extension of StaNavi could integrate other positioning methods, such as DR, to improve position accuracy to reduce the incidence of deviation.

When users received the off-route message, they tended to become close to a corner, wall, or pillar to interact with StaNavi and avoid other passengers; however, those spots sometimes receive no signal, which caused users to disappear from all beacon regions. This issue will be partially addressed by putting beacons in those spots or providing an additional instruction to move away from those spots. I also observed that users effectively used navigation cues to orient themselves. In particular, all participants took advantage of information on spatial relationships in surroundings in the open space of Route 2. I believe that the key to navigation tasks for blind travelers in a large and complex space is effective support for their understanding of situations such as poor signals, spatial relations to surrounding objects, and environmental patterns (e.g., tactile paving). Additional environmental cues, such as auditory (e.g., a beeping sound from a card reader at the entrance) and olfactory (e.g., the coffee aroma from a café) information can be helpful for identifying their location. The next prototype will include environmental cues to allow blind travelers to gain a better understanding of surroundings so that more independent and confident navigation can be achieved.

6.5 Summary

This chapter presented an indoor navigation system, StaNavi—a smartphone-based navigation system for blind travelers in a large train station utilizing BLE technology. Unlike the existing systems, StaNavi can be used from off-the-shelf smartphones without requiring additional hardware, and provides a set of features for ensuring easy and effective navigation. The user study conducted under realistic conditions demonstrated that the interaction design for one-handed operation and information on navigation cues and route overviews could encourage independence and confidence when navigating large and complex indoor environments. The user study presented "real-world" evidence of the effectiveness of using smartphones and BLE technology for facilitating independent navigation for people with visual impairments.

Chapter 7

Discussion

The previously presented projects, SaSYS (Chapter 4), TalkingTransit (Chapter 5), and StaNavi (Chapter 6), demonstrated how the proposed architecture can be used to develop applications that are really usable and well accepted by blind people in aiding POI search, public transit use, and indoor navigation. This chapter verifies the general applicability of the proposed architecture across every mobility scenario that this research covers: *Navigation* & *orientation* both outdoors and indoors, and *environmental access* to POI, public transit, buildings, street signage, and electronic kiosks. Specifically, the chapter shows how the architecture can be applied to develop applications that support scenarios by reflecting on the three implemented systems, conducting a rapid prototyping of a system that aids access within buildings, and discussing the development of applications that aid in other mobility scenarios.

7.1 Reflection on the Implemented Scenarios

The prototype systems were implemented for scenarios of aiding POI search, public transit use, and indoor navigation, and their effectiveness was verified through user studies. Many blind research participants who used smartphones mentioned that the prototypes were practical enough to integrate into their everyday lives. This section demonstrates how the proposed architecture has been applied to each successful application. I first present each prototype's design process, which consists of iterative design & development, evaluation, and user requirement identification, which shows how an iterative co-design approach can contribute to incrementally improving the quality and functionality of a mobility aid system for blind users. Then, I describe the system design of each prototype from the proposed architecture perspective to ensure that such a mobility aid system can be made based on this architecture.

7.1.1 Aiding for POI Search

The iterative design process of SaSYS

As shown in Figure 7.1, the current version of SaSYS, the POI search system introduced in Chapter 4, has been developed by iterating through four times of prototyping and testing. The early iterations of the design process allowed me to clarify the feasibility of using swipe gestures on touchscreens for blind people to indicate areas-of-interest, which helped me gain confidence in my unique "swipe-to-search" interaction method. Moreover, because SaSYS is the first of several mobility aid systems that I co-designed with blind people in this dissertation, its design process allowed me to identify the basic design requirements for creating applications in touch-based mobile devices for blind people. These include that blind users should be able to know what appears onscreen and obtain instructions to use their applications at any time; and should be able to perform touch inputs with confidence and certainty that their inputs will be recognized as intended.

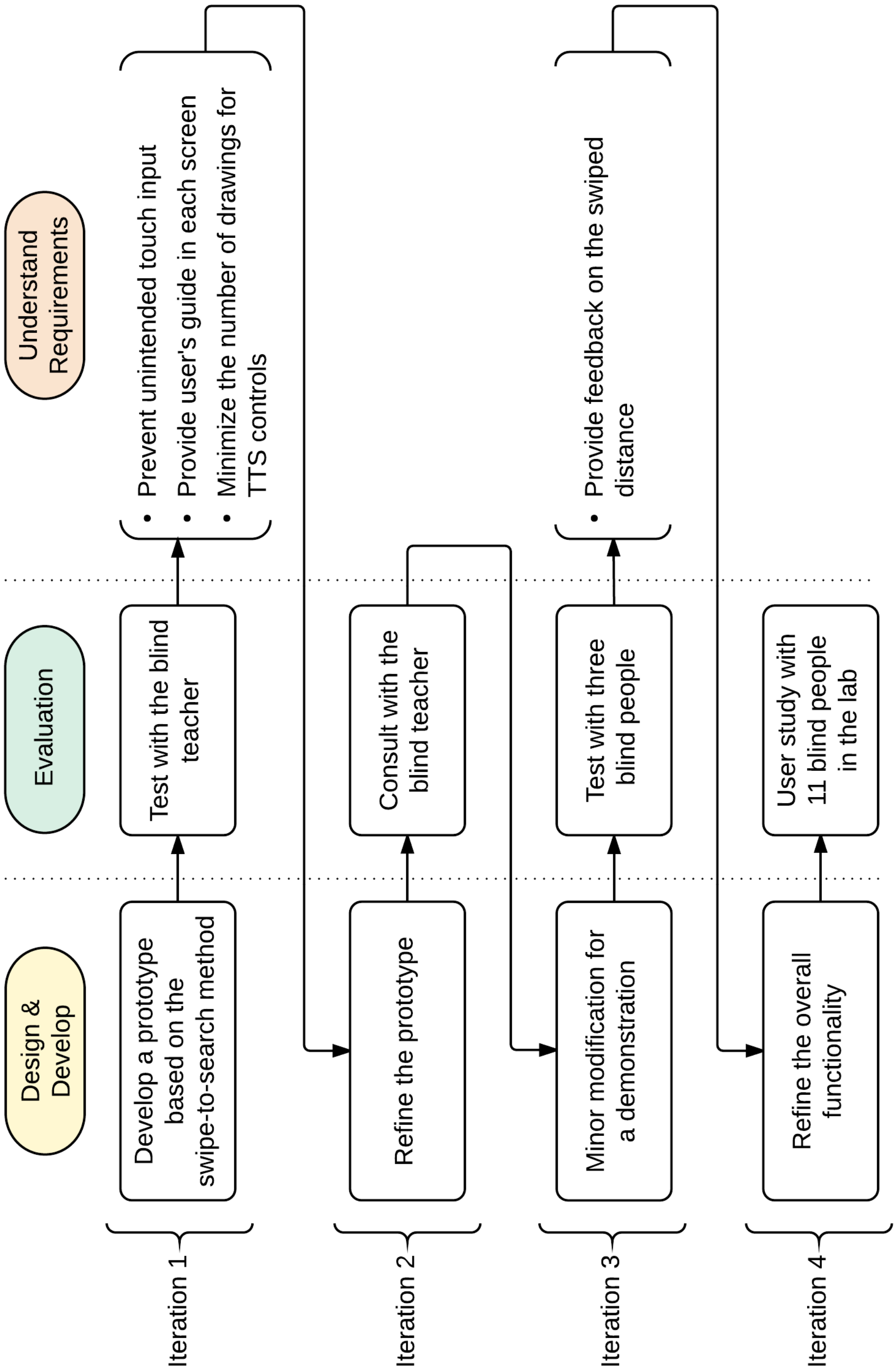


Figure 7.1: The design process of the SaSYS system, involving four iterations of implementation, testing, and refinement.

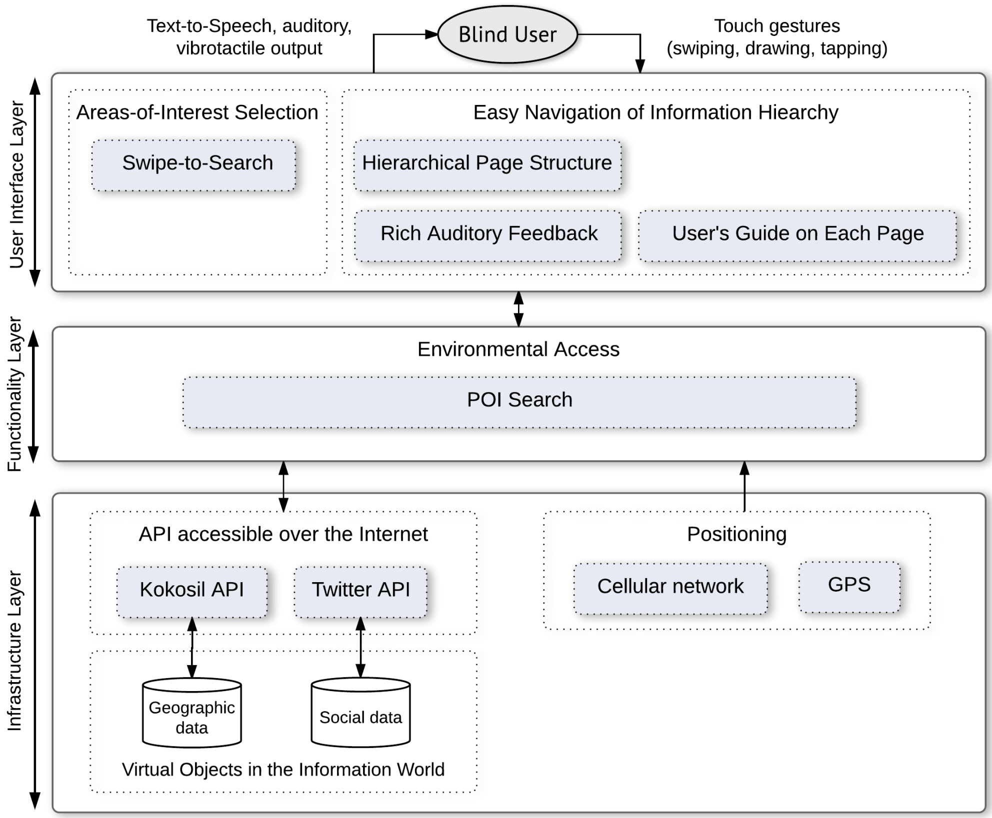


Figure 7.2: A representation of the SaSYS system from the proposed architecture perspective.

The architectural design of SaSYS

Figure 7.2 describes the structure of the SaSYS system from the proposed architecture perspective. The SaSYS system is made up of the following components in each layer of the architecture:

- *The infrastructure layer.* To provide POI information in areas-of-interest, SaSYS leverages geographic data and social data sourced respectively from Kokosil and Twitter. Each data resource is accessed using Web APIs: The Kokosil API for the geographic data; and the Twitter API for the social data. SaSYS makes use of GPS and cellular networks to obtain a user's current position.

- *The functionality layer.* It provides the functionality to support the mobility tasks in environmental access: POI search.
- *The user interface layer.* Users interact with SaSYS using touch gesture inputs and TTS outputs with auditory and vibrotactile feedback. It employs a swipe-to-search technique for the selection of areas-of-interest, and it employs the following techniques for easy navigation of the information hierarchy. First, it uses a hierarchical page structure that consists of three levels: The top page where users perform swipe-to-search at the first level; a list of retrieved POIs at the second; and detailed POI information divided into four categories at the third. Second, it provides auditory feedback for each touch input. In particular, auditory feedback on swipe-to-search, where a single note is played after performing a swipe gesture and the pitch changes according to the swiped distance. In addition, vibrotactile feedback is given when performing swipe-to-search. Lastly, it provides user's guide on each page.

7.1.2 Aiding for Public Transit Use

The iterative design process of TalkingTransit

As shown in Figure 7.3, the current version of TalkingTransit, the public transit information system introduced in Chapter 5 has been developed by iterating through three times of prototyping and testing. The first iteration, where I developed a prototype system that provides user interfaces based on VoiceOver gestures and conducted a pilot study, provided the opportunity to collect sufficient user feedback from 10 blind participants. This process helped me understand the user requirements and generate ideas such as “fast navigation” and “shake-to-main” to satisfy the specified requirements. In the following iteration, I then investigated whether the initial requirements could be satisfied, and captured another essential requirement for a better system. The prototype was refined again before conducting the user study.

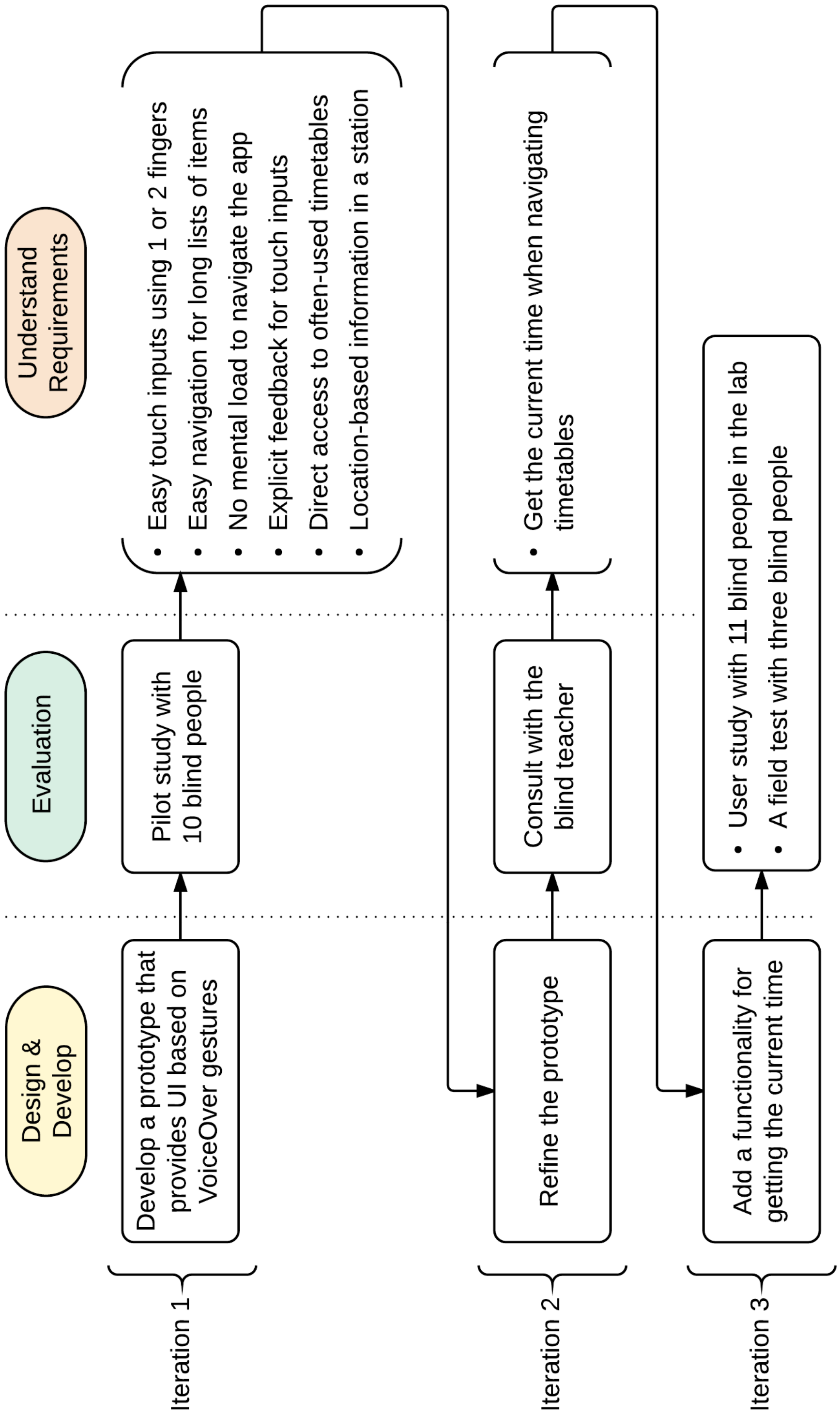


Figure 7.3: The design process of the TalkingTransit system involving three iterations of implementation, testing, and refinement.

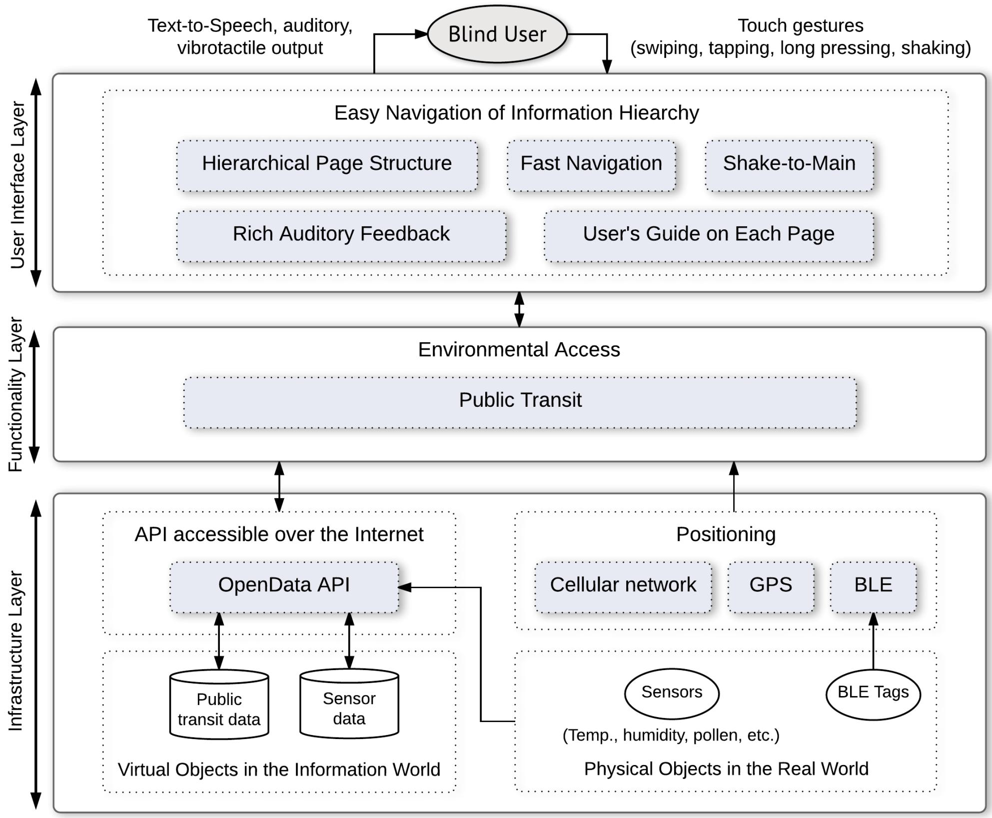


Figure 7.4: A representation of the TalkingTransit system from the proposed architecture perspective.

The architectural design of TalkingTransit

Figure 7.4 describes the structure of the TalkingTransit system from the proposed architecture perspective. The TalkingTransit system is made up of the following components in each layer of the architecture:

- *The infrastructure layer.* To provide real-time public transit information anytime, anywhere, TT leverages data including timetables, service status, and environmental conditions inside train stations that are made open by 11 railway operators; 2 subway operators; and Toei Bus in Japan. The data resources are accessed using a Web API: The OpenData API. TT makes use of GPS, cellular networks for outdoors, and utilizes BLE technology inside train stations to obtain a user's current position.

- *The functionality layer.* It provides the functionality to support the mobility tasks in environmental access: Public transit use.
- *The user interface layer.* Users interact with TT using touch and motion gestures inputs and TTS outputs with auditory and vibrotactile feedback. It employs the following techniques for easy navigation of the information hierarchy. First, it uses a hierarchical page structure as shown in Figure 5.2. Second, it supports fast navigation in browsing the Hiragana characters, station names, and timetable lists. Third, it provides shake-to-main to immediately return to the top page. Fourth, it provides auditory feedback for each input. Different sound effects are used to easily distinguish between performed gestures. Vibrotactile feedback is given when notifying users of any service changes. Lastly, it offers user's guide on each page.

7.1.3 Aiding for Indoor Navigation

The iterative design process of StaNavi

As shown in Figure 7.5, the current version of StaNavi, the indoor navigation system inside large train stations in Chapter 6 has been developed by iterating through four times of prototyping and testing. The first iteration investigated the feasibility of using BLE technology as an indoor navigation aid for blind people in which it pinpoints points-of-reference on a route rather than continuously locates and tracks a user. This initial field test allowed me to assess the potential of BLE technology and understand user needs by observing and interviewing the participants. Conducting extra field tests in the two following iterations helped me gain insights into functionality and information that could help realize the independent navigation of blind people, which led to the development of the refined, fully functional prototype.

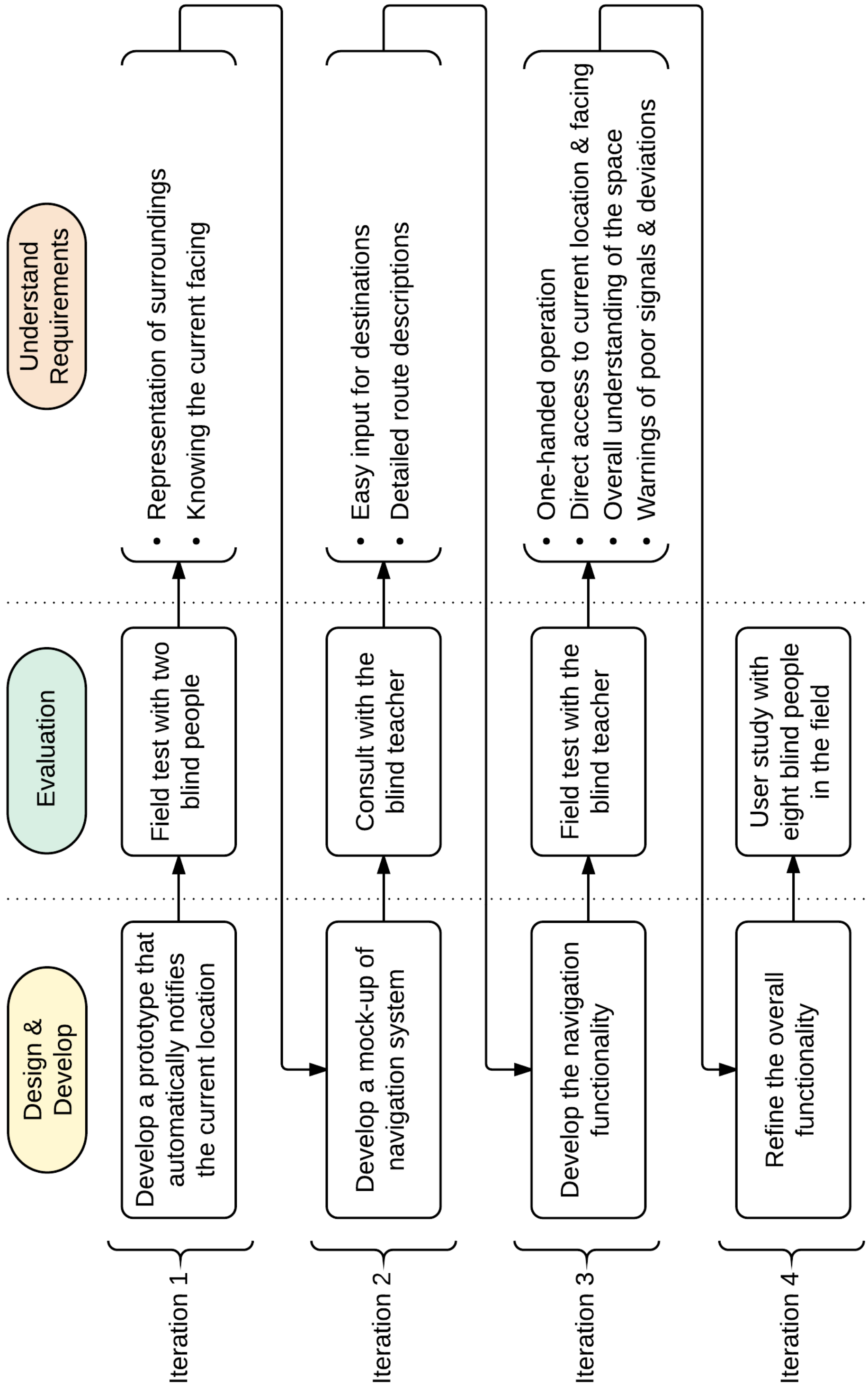


Figure 7.5: The design process of the StaNavi system involving four iterations of implementation, testing, and refinement.

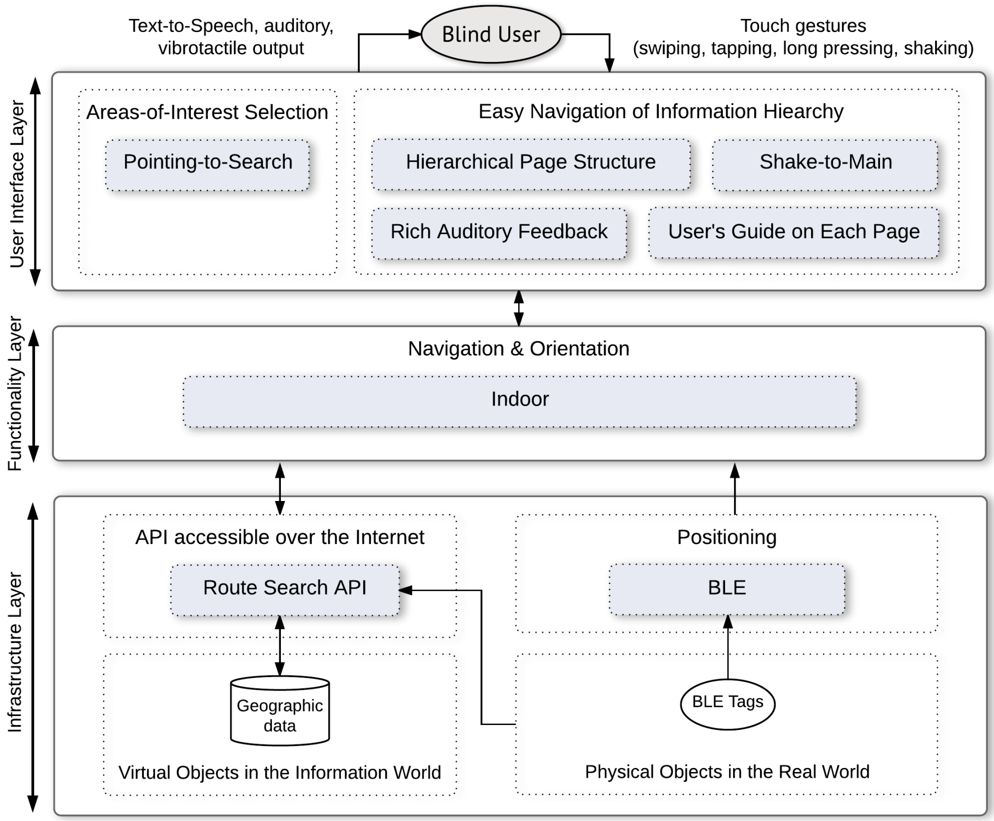


Figure 7.6: A representation of the StaNavi system from the proposed architecture perspective.

The architectural design of StaNavi

Figure 7.6 describes the structure of the StaNavi system from the proposed architecture perspective. The StaNavi system is made up of the following components in each layer of the architecture:

- *The infrastructure layer.* To provide directional instructions, StaNavi takes advantage of the geographic data resource of Tokyo Station that is accessed using a Web API: The Route Search API. StaNavi makes use of BLE beacons that have been experimentally installed by East Japan Railway to obtain a user's current position.

- *The functionality layer.* It provides the functionality to support navigation & orientation in indoor environments.
- *The user interface layer.* Users interact with StaNavi using touch and motion gestures inputs and TTS outputs with auditory and vibrotactile feedback. It employs a pointing-to-search technique for the selection of areas-of-interest, and it employs the following techniques for easy navigation of the information hierarchy. First, it uses a hierarchical page structure as shown in Figure 6.5. Second, it provides shake-to-main to directly return to the top page or "Navi Main" page (Figure 6.5E). Third, it provides auditory feedback for each input. Both auditory and vibrotactile feedback are given when detecting a BLE beacon. Lastly, it offers user's guide on each page.

7.2 Rapid Prototyping Scenario: Access within Buildings

This section aims to demonstrate how the proposed architecture can be applied to developing applications for other travel aid scenarios by conducting a rapid prototyping of a system that aids access within buildings and use their facilities. I prototyped a smart building application, TalkingBuilding (TB), which is a location-aware system that allows blind people to obtain their current location, monitor environmental conditions such as temperature and humidity, and control devices such as air conditioners (A/C) and elevators at the Daiwa Ubiquitous Computing Research (DUCR) Building, a state-of-the-art smart building at the University of Tokyo.

According to the U.S. Environmental Protection Agency, people spend an approximate average of 87% of their time in indoor environments such as at work, school, and home [111]. Thus, providing assistance to blind people not only to move between places but also to accomplish daily life tasks inside the buildings where they work or live can improve their quality of life. The following scenario illustrates one possible use of TB:

Mary, a blind student with a guide dog, arrives at her laboratory. Upon entering the room, she launches TB and checks the status of the lights. TB says: "A304, the student room. All lights are off." Although Mary is totally blind, she can remotely turn the light near her seat on by performing a simple touch gesture to indicate her presence and to make it easier for her dog to see. She also navigates to the A/C menu in TB and sets it to the desired temperature.

After a while, Mary attends a meeting for a group project; the meeting room for which is on the top floor. Once at the elevator hall, she starts TB again and checks the elevator's current floor. TB says: "The elevator is now on the second floor." She then remotely calls the elevator using TB without having to find the control panel.

Mary is unfamiliar with the layout of the top floor, and so she walks along the corridor with TB that reads aloud her current location when she passes within range of POIs such as doors and toilets. Upon nearing the meeting room, TB says: "You are in front of A604, the meeting room." Mary finds the right place and enters.

TB helps blind people achieve this type of task by providing improved access to building environments so that it can enhance their ability to become more independent and exercise control over their lives. The following subsections describe the system and its implementation in detail.

7.2.1 System Overview

Figure 7.7 illustrates the structure of the TalkingBuilding system from the proposed architecture. The TB system consists of the following components in each layer of the architecture:

- *The infrastructure layer.* It provides data including real-time environmental conditions and POIs within buildings. The Smart Building API is used to

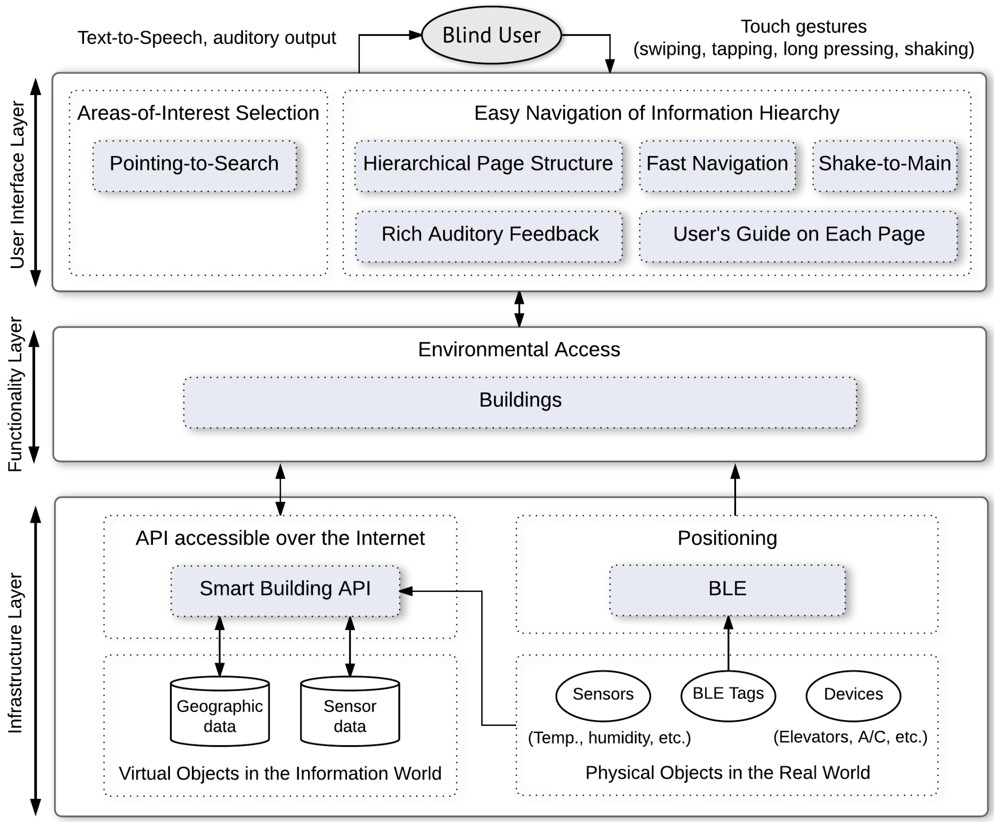


Figure 7.7: A representation of the TalkingBuilding system from the proposed architecture perspective.

access the data resources and control in-building machines over the Internet. TB utilizes BLE technology inside buildings to obtain a user's current position.

- *The functionality layer.* It provides the functionality to support the mobility tasks in environmental access: Access within buildings.
- *The user interface layer.* Users interact with TB using touch and motion gestures inputs and TTS outputs with auditory feedback. For areas-of-interest selection, it employs the pointing-to-search technique. For easy navigation of information hierarchy, it employs the hierarchical page structure as shown in Figure 7.9, fast navigation, and shake-to-main. It also provides auditory

feedback for every input and user's guide on each page. The user interface is described in detail in the following subsection.

Considering potential usage scenarios, the current version of TB includes three main functionalities:

- *Room status monitoring.* Users can check the current status of a room: The lighting (on/off), temperature, and humidity. Users can also obtain a list of remotely-operable devices in the room.
- *In-building machine control.* Users can remotely call an elevator; turn on and off lights in a room; and control a room air conditioner (turning on/off, set the temperature, fan speed, and airflow direction).
- *Automatic notification of in-building information.* Users can obtain automatic notification of their current location when they are within range of the BLE tags that are installed in front of each door and near the facilities such as a toilet. In addition, users can determine four directions they are facing.

TB is implemented on iOS 8 platform, and currently runs on BLE-enabled Apple iPhone 5S and 6. TB utilizes the device's built-in accelerometer and gyroscope sensors to detect shake-motion events. AVSpeechSynthesizer API included in iOS7 and higher platforms was used to offer TTS. I took advantage of Smart Building API designed for applications in the DUCR Building to provide the TB functionalities. Using Smart Building API, in-building machines such as elevators and A/C can be operated and managed based on real-time environmental contexts, such as the current user's location, temperature, humidity, and electricity usage that are collected from various environment sensors, BLE tags, and cameras deployed throughout the building as shown in Figure 7.8.

7.2.2 User Interface

Users interact with TB using a set of touch gestures (one- and two-finger swipe up and down, two-finger swipe left, one- and two-finger tap, and double-tap) and a

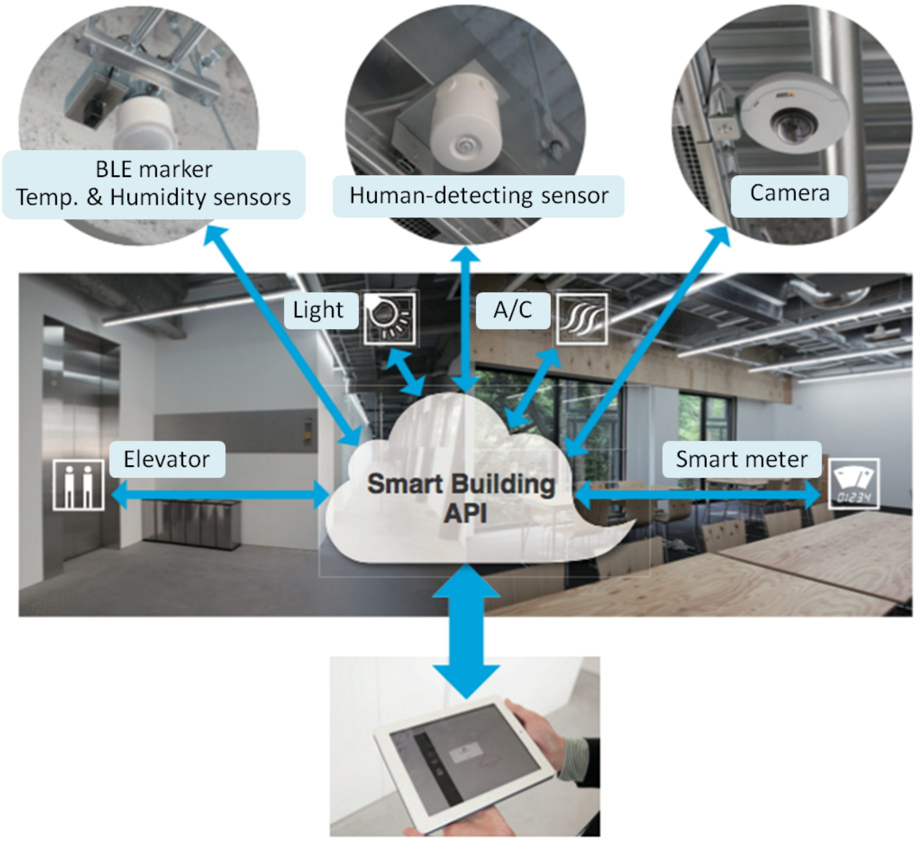


Figure 7.8: A mobile application using Smart Building API.

shake gesture, as shown in Figure 7.9. When starting TB, users can select a desired room where they want to monitor from a room list in the DUCR Building (Figure 7.9A). Users can browse the room list one by one by performing one-finger swipe up and down gestures, while two-finger swipes jump to the first item (room) of one floor up or down.

After selecting a room, TB transitions to the "Room monitoring" screen (Figure 7.9B) and sends a request to the server for retrieving the room status that provides two separate lists: First, the list of environmental conditions including light, temperature, and humidity; and second, the list of in-room machines including the name of devices that can be controlled remotely. Users can navigate each list by performing one-finger swipe up and down gestures, while two-finger swipes

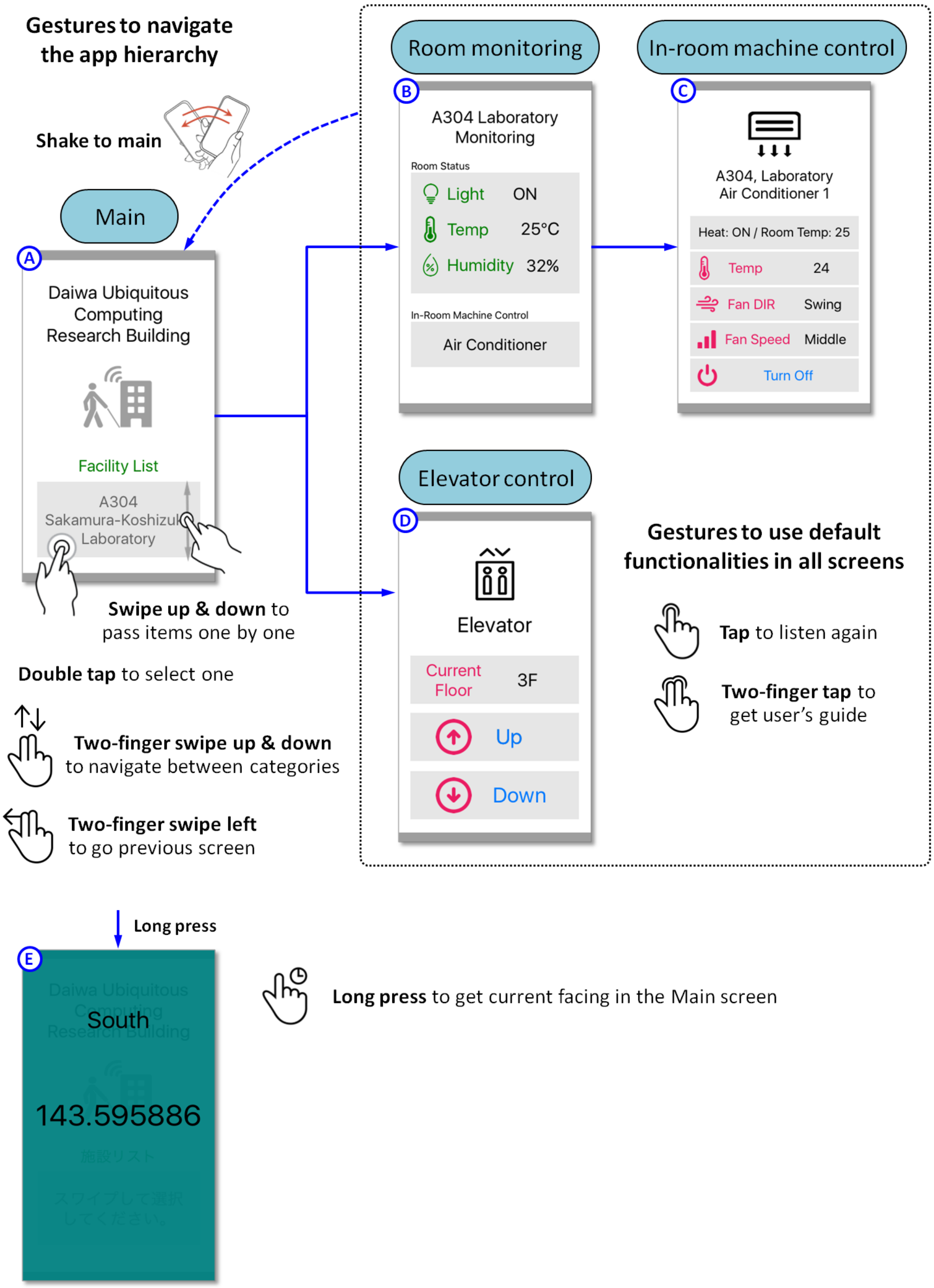


Figure 7.9: Screen transition flow and the performed gestures of TalkingBuilding.

switch between the two lists. When the download is finished, users will receive a voice notification that says, "The room status has been downloaded," followed by a full description of the room; for example, "A304, the laboratory room monitoring screen. Currently, the lights are on, the temperature is 25°C, and the humidity is 32%. There are two remotely-controllable devices in this room."

Selecting a device such as lights and A/C transitions to the "In-room machine control" screen (Figure 7.9C) that allows users to control it. Basically, users can navigate the control menu of the selected device by performing one-finger swipe up and down gestures, adjust the value of each menu item, if possible, by performing one-finger swipe left and right gestures (e.g., for the A/C control, swipe left or right may increase or decrease the temperature setting by one degree), and then confirm the change or their selection by double-tap.

Users can control an elevator by selecting an elevator hall from the room list in the "Main" screen (Figure 7.9A). Upon the transition to the "Elevator control" screen (Figure 7.9D), TB automatically specifies the elevator closest to a user based on the current location. In this screen, the user can select one of three items: the Current Floor, and the Up and Down buttons. Every time the user selects the "Current Floor", TB reads aloud the current floor of the elevator. Selecting one button and confirming it by double-tap will call the elevator.

Users can obtain their location information in the "Main" screen (Figure 7.9A). In this screen, TB automatically reads aloud the current location with auditory feedback when detecting a BLE tag. A pointing-based interaction is used to determine directions that are faced. Tapping and holding (long press) with one-finger brings up the screen (Figure 7.9E), which reads aloud the four cardinal directions according to the direction that the user's device is pointing towards. A two-finger tap anywhere in TB offers guidance on the use of each screen. For example, "Main screen. Swipe up and down to navigate the room list and double-tap to select." In addition, users can return to the top screen directly by shaking their devices.

7.2.3 Reflection on the Effectiveness of the Proposed Architecture

In this subsection, I discuss how the architecture helped create more useful travel aid systems inside building environments by describing its benefits in TB design. In general, although the current version of TB has not yet been tested by blind people, they could conceivably use TB because its design is based on the user interfaces that were well accepted by the blind users who participated in this research. In particular, the fast navigation technique was helpful for providing a way for target users to quickly scan a room list in the "Main" screen (Figure 7.9A). This technique was also of great use in the "Room monitoring" screen (Figure 7.9B), in that it allowed easy navigation for distinct groups of items in one screen without having to create deeper page structures.

Moreover, the IoT infrastructure in the DUCR Building facilitated not only obtaining rich information on the real-time environmental conditions but also controlling the surrounding devices remotely with just the device's Internet access capability. I believe that this infrastructure has the potential to provide blind individuals with further assistance to accomplish a wide range of daily tasks.

7.3 Other Scenarios

This section discusses how the proposed architecture can be used to develop applications that aid in other mobility scenarios such as outdoor navigation and environmental access to street signage and electronic kiosks.

7.3.1 Outdoor Navigation

Outdoor navigation systems can basically be realized based on the same idea as the StaNavi system in terms of user interface design, as shown in Figure 7.6. The fast navigation technique may also be used if a navigation application should display a long list of items such as nearby places of interest or popular destinations. Positioning infrastructure such as GPS and cellular network can be used to obtain a user's current location in an outdoor environment, coupled with other technologies such as BLE that can pinpoint the locations of navigational waypoints (e.g., landmarks, crossroads, and street corners) more accurately. The navigation systems can leverage publicly available geospatial data resources that provide useful information for blind people's navigation; for example, the OpenStreetMap for the blind[19] includes information such as tactile paving and audible traffic signals.

7.3.2 Access to Street Signage

It is possible to realize environmental access to street signage and signboards if the concept of IoT is applied to such "things." A radio frequency (RF) tag-based approach can be used to identify them with smartphones, but deploying tags on every sign seems too costly and therefore may not be practical. Nevertheless, I believe that tag-based identification remains a feasible approach to providing blind people with the functionality to access information about nearby signage and signboards. The following describes a potential application scenario:

> Mary holds her mobile device in her hand while walking around town. When she passes in front of a building, the device vibrates to tell her that there is some signage content and reads the building name aloud. Selecting it, she begins to listen to the content on that building's signboards. Arriving at an intersection, the device vibrates and reads the intersection name aloud. Selecting it and then pointing her device in her desired direction, she begins to listen to the content of the street signage in that direction. After finding the correct direction, she continues on her way.

Achieving such application requires that the positioning infrastructure, where RF tags (such as BLE tags that can be read by smartphones) are installed on the front of buildings or waypoints in urban environments, can be used to allow applications to determine blind people's locations in close proximity to such specific points.

[19] `http://wiki.openstreetmap.org/wiki/OSM_for_the_blind`

Simultaneously, if geospatial data resources including information about signage are made available, applications could provide relevant information when detecting installed tags. The user interface design may follow that of SaSYS as shown in Figure 7.2, except that instead of using "swipe-to-search," it employs "pointing-to-search" to select areas-of-interest.

Although the proposed architecture could support access to information about signage to a certain extent, the most common approach in the literature is based on computer vision technologies that detect and recognize text and signs in urban scenes [112]. Recently, a crowdsourcing-based approach has also been proposed to address this type of mobility challenge. For example, blind people can ask questions about street signage by taking pictures and receiving answers promptly from sighted people on the Web via a mobile application such as VizWiz [113] and VizLens [114]. Integrating these approaches with the RF tag-based method could provide better access to street signage.

7.3.3 Access to Electronic Kiosks

Consider the following potential application scenario that aids access to electronic kiosks such as vending machines and ATMs, which can be realized by the proposed architecture:

> Mary is navigating a train station. When she passes a beverage vending machine, the device notifies her of the machine's presence via vibrotactile and auditory feedback. Feeling thirsty, she selects it and begins to listen to the goods available. After choosing her drink, she easily buys it over the Internet by performing simple touch gestures without having to find the slot and insert the correct coins or bills.

An example of infrastructure for such applications would be electronic kiosks installing BLE tags so that smartphones can detect such machines in the vicinity. Meanwhile, if Web APIs for online payments processing and the real-time retrieval of information about electronic kiosks (e.g., inventory data from vending machines) are made available, applications could remotely monitor and control them. The user interface design may follow one similar to that of TalkingTransit as shown in Figure 7.4. The hierarchical page structure may consist of: A top page at the first level, a list of menu items about an electronic kiosk at the next few levels, and a payment page on the last level. A fast navigation technique can be used if the electronic kiosk has a long list of items to display. This approach can be of benefit to everyone. For example, people in wheelchairs who may have difficulty pushing out-of-reach buttons on vending machines can also buy goods more easily.

7.4 Summary

This chapter presented how the proposed architecture can be used to create applications for supporting a range of mobility scenarios in this research. Three prototype systems, SaSYS, TalkingTransit, and StaNavi, have been developed to support scenarios such as aiding POI search, public transit use, and indoor navigation, and these systems have been well received by blind users. Section 7.1 demonstrated that each prototype system can be explained by the proposed architecture. Section 7.2 conducted a rapid prototyping of TalkingBuilding for the scenario of assisting access within buildings, and demonstrated that the proposed architecture can be applied to develop systems that facilitate the use of buildings for the blind. Section 7.3 briefly described examples of using the proposed architecture to realize systems that aid outdoor navigation and provide access to signage and electronic kiosks.

This chapter has clearly demonstrated that the proposed architecture can cover all target mobility scenarios using mobile devices such as smartphones without necessitating that blind users acquire any additional hardware, showing that this architecture is adequate to achieve the objectives of this dissertation.

Chapter 8

Conclusion

This dissertation has proposed a mobility aid system architecture for blind people that can create easy-to-use and readily usable assistive technology solutions that are delivered via mainstream mobile devices such as smartphones without requiring that users carry additional special hardware. The architecture covers a set of mobility challenges that blind people face while on the move: *Navigation & orientation* both outdoors and indoors, and *environmental access* to POIs, public transit, buildings, signage, and electronic kiosks—which are considered major tasks during the travel activities of such individuals. This architecture has been derived based on the experiences of conducting iterative prototyping with blind users for smartphone-based applications that aid target mobility tasks. Its effectiveness was demonstrated via the implementation and evaluation of SaSYS, TalkingTransit, and StaNavi, which respectively provide mobility aids for POI search, public transit use, and indoor navigation. In addition to verifying the correct composition of the three implemented systems from the proposed architecture perspective, the potential of this architecture to create further mobility aids was also demonstrated by prototyping TalkingBuilding to aid the use of buildings and by discussing the development of applications that aid outdoor navigation and provide access to signage and electronic kiosks. I can conclude from this that the proposed architecture can be applicable across all the target mobility tasks in this dissertation.

8.1 Reflections and Insights

The work presented here reinforces the importance of using mainstream technologies to support blind people's mobility needs in terms of gaining widespread adoption. I have observed a rise in the number of blind research participants who own a smartphone during my dissertation research. The number of smartphone users who participated in the user study for SaSYS in 2012 was two of 11 participants, five of 11 participatns for TalkingTransit in 2014 owned one, as did six of the eight participants for StaNavi in 2015. The blind teacher who closely collaborated in this research once talked about the growing popularity of smartphones among his students.

Moreover, according to recent statistics, approximately 100,000 blind people use an Apple iPhone [31]. A more recent study investigating the use of Facebook by 50,000 visually impaired iPhone users shows that they engage in the Facebook activities as often as sighted users [20]. This statistic, the study results, and my research experiences clearly suggest that smartphones are becoming "mainstream" among the blind community. Thus, I believe that providing mobility aids via modern mobile devices can strongly support the integration of such technology into daily life.

The iterative co-design approach adopted throughout this research undoubtedly facilitated the development of easy-to-use and readily usable smartphone-only solutions that help blind people travel independently without requiring additional dedicated hardware. Generally, a system's overall quality can be improved as the number design cycle repetitions increases; however, design iterations may lead to increased development time and costs. One important question that may be raised by this approach, though outside of the scope of this dissertation, is "what is the appropriate number of design iterations to guarantee some degree of system usability for blind people?"

During the development of SaSYS, TalkingTransit, and StaNavi, the respective number of iterations performed until each prototype reached an acceptable level of usability were four, three, and four. This empirical evidence suggests that the notion offered by Nielsen (1993, p.32) that "iterating through at least three versions can substantially improve usability" can be also applied to the development of mobile applications designed for people with visual impairments. Thus, it is desirable to conduct at least three iterations of implementation, user testing, and refinement to provide acceptable-quality systems that assist blind people's mobility.

All four prototypes, SaSYS, TalkingTransit, StaNavi, and TalkingBuilding, have benefited from existing IoT infrastructure where virtual resources (sensor data, social data, open data, etc.) and physical resources (sensors, BLE tags, elevators, etc.) can be accessed over the Internet. However, not all the resources in the IoT infrastructure were freely available. Differences existed in the level of access to resources; for example, access to the public transit data used by TalkingTransit was publicly available, whereas access to sensor data, BLE tags, and the devices used by TalkingBuilding were exclusive to the DUCR Building residents. The important point here is that this IoT infrastructure is shared across other services and basically was not designed or built with blind people in mind.

By taking advantage of this shared infrastructure, my research projects have mainly focused on the design and implementation of client applications, which helped reduce the development time and costs for the prototype systems. The four prototypes presented in this dissertation show clear evidence that movements towards the open APIs that IT companies and public agencies make available to third-party developers and open data can facilitate easy, rapid, and cost-effective prototyping and can therefore provide opportunities for the creation of innovative value-added IT services that benefit groups of people within niche markets such as the elderly and people with disabilities.

Despite exponential growth in the number of smartphone applications over the last few years, there remains a lack of applications related to visual impairment. As of 2013, the Apple iTunes store offers only 30 low vision-related applications [115] out of 475,000 iPhone and iPod applications [116], and the Google Play store offers only 33 low vision-related applications [115] out of more than 800,000 Android applications [117]. Possible reasons for this may include a lack of awareness among developers about how many blind users own smartphones and use them as often as the general population [20], and a lack of references or best practice applications for developers. Thus, I hope that the proposed architecture in this dissertation, which includes a set of techniques for designing user interfaces for touch-based mobile devices that are intuitively usable and specialized for blind people, will play an important role in encouraging the development of accessible mobile applications for blind smartphone users.

8.2 Future Work

This section provides suggestions for future research directions.

The Adoption and Integration of Wearable Technology

While the presented mobility aid system architecture only uses smartphones as user terminal devices without necessitating any additional specialized hardware, given the growing popularity of general-purpose wearable devices in mainstream consumer market [118] such as smart watches (e.g., Apple Watc[20], Samsung Galaxy Gear2[21]) and smart glasses (e.g., Sony SmartEyeGlass[22], Vuzix M100[23], Epson Moverio BT-200[24]), the adoption of such devices as alternative input/output methods could enrich and extend the proposed architecture. Indeed, two wearable technologies—smart watches and smart glasses—have recently been recognized by the blind community for their potential to positively impact blind people's access to information while on the move [119].

Modern smart watches that typically have vibration capabilities and are equipped with a rich set of sensors such as GPS, accelerometers, and gyroscopes could offer

[20] `http://www.apple.com/watch/`
[21] `http://www.samsung.com/uk/gear-s2/`
[22] `http://developer.sony.com/devices/mobile-accessories/smarteyeglass/`
[23] `http://www.vuzix.com/Products/m100-smart-glasses`
[24] `http://epson.com/moverio-augmented-reality-smart-glasses`

the navigation functionality in a more simple and unobtrusive manner by eliminating the need for blind people to hold their smartphone while walking. For example, a blind user could perform simple touch gestures (e.g. swipes and taps) on a smart watch or non-touchscreen gestures (e.g. shaking, tilting) for menu selection and navigation [120], which would be spoken through his earphones by the screen reader of the smartphone in his pocket or bag. Then, route instructions could be delivered to the user via auditory or haptic feedback using different vibration patterns [121], which would be less disruptive than spoken directions.

Smart glasses equipped with cameras could enhance the functionality that access to information about street signage. For example, smart glasses could capture images in a blind user's line of sight and then automatically detect and recognize text and objects in a surrounding environment, which would be read aloud to the user via the screen reader on his smartphone. This capability could also be used to extend the proposed architecture to cover obstacle avoidance, which is an important navigational task in blind people's travel activities but outside of the scope of this dissertation. Moreover, the navigation functionality benefits from smart glasses with landmark identification software such as Headlock [71].

Towards Large-scale Deployment in the Real World

This dissertation has demonstrated that the three prototypes (as described in Chapters 4, 5, and 6) developed based on the proposed architecture are well-accepted by blind people. Generally, my blind research participants were pleased with my prototypes, and many of them who own a smartphone have expressed their desire to install and use the prototypes in their daily lives. However, the large-scale deployment of such applications in the real world would require expanding the coverage of the IoT infrastructure, which is often accompanied by several challenges such as network connectivity, security, privacy, and standardization [122]. The need for common standards for IoT data interoperability was particularly apparent during the processes of developing my prototype systems.

One standardization area that has been recognized by this research is open government data. Because open data come from several different sources with heterogeneous structures, integrating such datasets may require significant effort [123]. Although TalkingTransit largely benefited from the Open Data API that allows developers to easily retrieve the real-time information released by 14 public transit operators in Tokyo via a single point of access, there were some minor integration problems. For example, the dataset of train stations in the Tokyo metropolitan area was downloaded and locally saved in the application; it contained related attributes including the full names of railway operators. However, the Open Data API provided train timetables by railway operator where one operator name was abbreviated, which caused a mismatch when extracting information about that operator. Given that there are many stakeholders and players in open data (for example Japan has more than 200 railway companies), data standards should be established to maximize the effective reuse of their data, which "allows applications created in one locality to be easily used by another locality or readily integrated into private applications [124]." Identifying metadata needs for describing the characteristics of datasets (e.g., data structures, types, access rights, authorship, etc.) and developing mechanisms to ensure the quality of such metadata would help establish open data standards.

Another standardization area is a spatial data model for representing indoor environments. Many outdoor LBS have been deployed thanks to the large amount of map data and contents widely available on the Internet (e.g., Google Maps, OpenStreetMap) with standardized addressing schemes (e.g., postal codes, street names). In contrast to outdoor spatial information, map data and relevant content for indoor LBS tend to be created using proprietary formats without uniformity in indoor addressing schemes, which makes the LBS developed for a specific building difficult to reuse for other buildings. The importance of spatial data standards in driving the global deployment of indoor LBS has recently been emphasized by IoT stakeholders [125], and several efforts have been made to that end. IndoorGML[25], the recent standard adopted by the Open Geospatial Consortium that defines a data model and XML schema for describing indoor spaces, is one notable example. Although IndoorGML seems a promising model for indoor navigation systems, it currently lacks the ability to represent detailed information about landmarks and clues that can help blind people determine their current location and orient themselves such as tactile, (e.g., wood floors, carpet), auditory (e.g., a beeping sound from a card reader at a train station entrance gate), or olfactory (e.g., the coffee aroma from a café) information. Thus, future work should extend IndoorGML to incorporate such contextual information to support the development of indoor LBS for people with disabilities.

[25] `http://www.opengeospatial.org/standards/indoorgml`

References

[1] A. E. Patla and A. Shumway-Cook. Dimensions of Mobility: Defining the Complexity and Difficulty Associated with Community Mobility. *Journal of Aging and Physical Activity*, vol. 7, no. 1, pp. 7–19, 1999.

[2] S. Kenyon, G. Lyons, and J. Rafferty. Transport and social exclusion: investigating the possibility of promoting inclusion through virtual mobility. *Journal of Transport Geography*, vol. 10, no. 3, pp. 207–219, 2002.

[3] P. Wolfe. *Brain Matters: Translating Research into Classroom Practice (2nd Edition)*. Association for Supervision and Curriculum Development, 2010.

[4] J. J. Gibson and L. E. Crooks. A theoretical field-analysis of automobile-driving. *The American Journal of Psychology*, vol. 51, no. 3, pp. 453–471, 1938.

[5] World Health Organization. Visual impairment and blindness. `http://www.who.int/mediacentre/factsheets/fs282/en/`, 2011. Accessed 13 July 2016.

[6] America, P. B. Vision Problems in the US. `http://www.preventblindness.net/site/DocServer/VPUS_2008_update.pdf`, 2008. Accessed 8 July 2014.

[7] J. Köberlein, K. Beifus, C. Schaffert, and R. P. Finger. The economic burden of visual impairment and blindness: a systematic review. *BMJ Open*, vol. 3, no. 11, p. e003471, 2013.

[8] A. F. Cruess, K. D. Gordon, L. Bellan, S. Mitchell, and M. L. Pezzullo. The cost of vision loss in Canada. 2. Results. *Canadian Journal of Ophthalmology*, vol. 46, no. 4, pp. 315–318, 2011.

[9] C. B. Roberts, Y. Hiratsuka, M. Yamada, M. L. Pezzullo, K. Yates, S. Takano, K. Miyake, and H. R. Taylor. Economic cost of visual impairment in japan. *Archives of Ophthalmology*, vol. 128, no. 6, pp. 766–771, 2010.

[10] J. S. Wittenborn and D. B. Rein. Cost of Vision Problems: The Economic Burden of Vision Loss and Eye Disorders in the United States. NORC at the University of Chicago. Prepared for Prevent Blindness America, Chicago, IL, 2013. Available at: http://costofvision.preventblindness.org.

[11] C. Asakawa and T. Itoh. User Interface of a Home Page Reader. In *Proceedings of the Third International ACM Conference on Assistive Technologies*, pp. 149–156. ACM, 1998.

[12] H. Miyashita, D. Sato, H. Takagi, and C. Asakawa. Aibrowser for Multimedia: Introducing Multimedia Content Accessibility for Visually Impaired Users. In *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 91–98. ACM, 2007.

[13] D. Sato, S. Zhu, M. Kobayashi, H. Takagi, and C. Asakawa. Sasayaki: Augmented Voice Web Browsing Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2769–2778. ACM, 2011.

[14] J. M. Loomis, J. R. Marston, R. G. Golledge, and R. L. Klatzky. Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance. *Journal of Visual Impairment & Blindness*, vol. 99, no. 4, pp. 219–232, 2005.

[15] A. Helal, S. E. Moore, and B. Ramachandran. Drishti: an integrated navigation system for visually impaired and disabled. In *Proceedings Fifth International Symposium on Wearable Computers*, pp. 149–156. IEEE, 2001.

[16] H. Petrie, V. Johnson, T. Strothotte, A. Raab, S. Fritz, and R. Michel. MOBIC: Designing a Travel Aid for Blind and Elderly People. *Journal of Navigation*, vol. 49, no. 1, pp. 45–52, 1996.

[17] M. Hersh and M. A. Johnson (Eds.). *Assistive Technology for Visually Impaired and Blind People*. Springer-Verlag London, 2008.

[18] R. Manduchi and S. Kurniawan (Eds.). *Assistive Technology for Blindness and Low Vision*. CRC Press, 2012.

[19] P. Strumillo. Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation. In *3rd International Conference on Human System Interaction*, pp. 17–24. IEEE, 2010.

[20] S. Wu and L. A. Adamic. Visually Impaired Users on an Online Social

Network. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3133–3142. ACM, 2014.

[21] VoiceOver. http://www.apple.com/accessibility/voiceover/. Accessed 13 July 2016.

[22] TalkBack. https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback&hl=en. Accessed 13 July 2016.

[23] J. Rivera and R. Meulen. Gartner's 2015 Hype Cycle for Emerging Technologies Identifies the Computing Innovations That Organizations Should Monitor. Gartner Newsroom, 2015. Available at: http://www.gartner.com/newsroom/id/3114217.

[24] J. Kim, M. Bessho, N. Koshizuka, and K. Sakamura. Enhancing Public Transit Accessibility for the Visually Impaired Using IoT and Open Data Infrastructures. In *Proceedings of the First International Conference on IoT in Urban Space*, pp. 80–86. ICST, 2014.

[25] E. Foulke. The perceptual basis for mobility. *Research Bulletin*, no. 23, pp. 1–8, 1971.

[26] M. Brambring. Mobility and orientation processes of the blind. In D. H. Warren and E. R. Strelow (Eds.), *Electronic Spatial Sensing for the Blind*, pp. 493–508. Springer, 1985.

[27] S. Harper and P. Green. A Travel Flow and Mobility Framework for Visually Impaired Travellers. In *International Conference on Computers Helping People with Special Needs*, pp. 289–296. OCG Press, 2000.

[28] BlindSquare. http://blindsquare.com/. Accessed 13 July 2016.

[29] Ariadne GPS. http://www.ariadnegps.eu/. Accessed 3 November 2016.

[30] M. B. Dias, A. Steinfeld, and M. B. Dias. Future Directions in Indoor Navigation Technology for Blind Travelers. In H. A. Karimi (Ed.), *Indoor Wayfinding and Navigation*, pp. 203–226. 2015.

[31] U.S. Bureau of Engraving and Printing. Eyenote. http://moneyfactory.gov/images/EyeNote_Press_Release_4_4-19_2_4.pdf, 2011. Accessed 5 June 2014.

[32] S. K. Kane, J. P. Bigham, and J. O. Wobbrock. Slide Rule: Making Mobile Touch Screens Accessible to Blind People Using Multi-touch Interaction

Techniques. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 73–80. ACM, 2008.

[33] S. Zhao, P. Dragicevic, M. Chignell, R. Balakrishnan, and P. Baudisch. Earpod: Eyes-free Menu Selection Using Touch Input and Reactive Audio Feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1395–1404. ACM, 2007.

[34] N. Verma, P. Shetye, D. Gangopadhyay, M. Bisht, and I. Pavlyshak. Find It: Information at Hand. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp. 467–470. ACM, 2010.

[35] D. McGookin, S. Brewster, and W. Jiang. Investigating Touchscreen Accessibility for People with Visual Impairments. In *Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges*, pp. 298–307. ACM, 2008.

[36] J. Sánchez and F. Aguayo. Mobile Messenger for the Blind. In *Universal Access in Ambient Intelligence Environments*, pp. 369–385. Springer, 2007.

[37] F. C. Y. Li, D. Dearman, and K. N. Truong. Leveraging Proprioception to Make Mobile Phones More Accessible to Users with Visual Impairments. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 187–194. ACM, 2010.

[38] M. C. Buzzi, M. Buzzi, B. Leporini, and L. Martusciello. Making Visual Maps Accessible to the Blind. In *Proceedings of the 6th International Conference on Universal Access in Human-computer Interaction: Users Diversity - Volume Part II*, pp. 271–280. Springer, 2011.

[39] S. K. Kane, J. O. Wobbrock, and R. E. Ladner. Usable Gestures for Blind People: Understanding Preference and Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 413–422. ACM, 2011.

[40] S. Robinson, P. Eslambolchilar, and M. Jones. Sweep-Shake: Finding Digital Resources in Physical Environments. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, no. 12, pp. 1–10. ACM, 2009.

[41] P. Parente and G. Bishop. BATS: The Blind Audio Tactile Mapping System.

In *Proceedings of the ACM Southeast Regional Conference*, pp. 132–137, 2003.

[42] M. Pielot, N. Henze, W. Heuten, and S. Boll. Tangible User Interface for the Exploration of Auditory City Maps. In *Proceedings of the 2nd International Conference on Haptic and Audio Interaction Design*, pp. 86–97. Springer, 2007.

[43] J. Su, A. Rosenzweig, A. Goel, E. de Lara, and K. N. Truong. Timbremap: Enabling the Visually-impaired to Use Maps on Touch-enabled Devices. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp. 17–26. ACM, 2010.

[44] B. Poppinga, C. Magnusson, M. Pielot, and K. 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*, pp. 545–550. ACM, 2011.

[45] K. Yatani, N. Banovic, and K. 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*, pp. 415–424. ACM, 2012.

[46] F. Bellotti, R. Berta, A. De Gloria, and M. Margarone. Guiding visually impaired people in the exhibition. *Mobile Guide*, vol. 6, pp. 1–6, 2006.

[47] V. Coroama. Experiences from the Design of a Ubiquitous Computing System for the Blind. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, pp. 664–669. ACM, 2006.

[48] S. A. Panëels, A. Olmos, J. R. Blum, and J. R. Cooperstock. Listen to It Yourself!: Evaluating Usability of What's Around Me? For the Blind. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2107–2116. ACM, 2013.

[49] R. Yang, S. Park, S. R. Mishra, Z. Hong, C. Newsom, H. Joo, E. Hofer, and M. W. Newman. Supporting Spatial Awareness and Independent Wayfinding for Pedestrians with Visual Impairments. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 27–34. ACM, 2011.

[50] P. Fröhlich, R. Simon, L. Baillie, and H. Anegg. Comparing Conceptual Designs for Mobile Access to Geo-spatial Information. In *Proceedings of*

the 8th Conference on Human-computer Interaction with Mobile Devices and Services, pp. 109–112. ACM, 2006.

[51] E. Rukzio, K. Leichtenstern, V. Callaghan, P. Holleis, A. Schmidt, and J. Chin. An Experimental Comparison of Physical Mobile Interaction Techniques: Touching, Pointing and Scanning. In *Proceedings of the 8th International Conference on Ubiquitous Computing*, pp. 87–104. Springer, 2006.

[52] R. Wasinger, C. Stahl, and A. Krüger. M3I in a Pedestrian Navigation & Exploration System. In *Proceedings of the 5th International Symposium on Human-Computer Interaction with Mobile Devices*, pp. 481–485. Springer, 2003.

[53] R. Jacob, P. Mooney, B. Shalaik, and A. Winstanley. HapticTransit: Tactile feedback to notify public transport users. *International Journal of Digital Information and Wireless Communications*, vol. 1, no. 1, pp. 204–218, 2011.

[54] M. Markiewicz and M. Skomorowski. Public Transport Information System for Visually Impaired and Blind People. In J. Mikulski (Ed.), *Transport Systems Telematics*, pp. 271–277. Springer, 2010.

[55] H. Zhou, K.-M. Hou, D. Zuo, and J. Li. Intelligent Urban Public Transportation for Accessibility Dedicated to People with Disabilities. *Sensors*, vol. 12, no. 8, pp. 10678–10692, 2012.

[56] G. Baudoin, O. Venard, G. Uzan, A. Rousseau, Y. Benabou, A. Paumier, and J. Cesbron. The RAMPE Project: Interactive, Auditive Information System for the Mobility of Blind People in Public Transports. In *Proceedings of the 5th International Conference on Intelligent Transportation Systems Telecommunications*, 2005.

[57] S. Azenkot, S. Prasain, A. Borning, E. Fortuna, R. E. Ladner, and J. O. Wobbrock. Enhancing Independence and Safety for Blind and Deaf-blind Public Transit Riders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3247–3256. ACM, 2011.

[58] K. Hara, S. Azenkot, M. Campbell, C. L. Bennett, V. Le, S. Pannella, R. Moore, K. Minckler, R. H. Ng, and J. E. Froehlich. Improving Public Transit Accessibility for Blind Riders by Crowdsourcing Bus Stop Landmark Locations with Google Street View. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, no. 16, pp. 1–8. ACM, 2013.

[59] M. Turunen, T. Hurtig, J. Hakulinen, A. Virtanen, and S. Koskinen. Mobile speech-based and multimodal public transport information services. In *Proceedings of MobileHCI 2006 Workshop on Speech in Mobile and Pervasive Environments*, 2006.

[60] S. Landau, G. Bourquin, J. Miele, and A. Van Schaack. Demonstration of a universally accessible audio-haptic transit map built on a digital pen-based platform. In *Proceedings of the 3rd International Workshop on Haptic and Audio Interaction Design*, pp. 23–24, 2008.

[61] L. Brunet, C. Megard, S. Panëels, G. Changeon, J. Lozada, M. P. Daniel, and F. Darses. "Invitation to the voyage": The design of tactile metaphors to fulfill occasional travelers' needs in transportation networks. In *2013 World Haptics Conference (WHC)*, pp. 259–264. IEEE, 2013.

[62] N. Harrington, L. Antuna, and Y. Coady. ABLE Transit: A Mobile Application for Visually Impaired Users to Navigate Public Transit. In *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, pp. 402–407. IEEE, 2012.

[63] P. Korbel, P. Skulimowski, P. Wasilewski, and P. Wawrzyniak. Mobile applications aiding the visually impaired in travelling with public transport. In *2013 Federated Conference on Computer Science and Information Systems*, pp. 825–828. IEEE, 2013.

[64] H. Liu, H. Darabi, P. Banerjee, and J. Liu. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1067–1080, 2007.

[65] X. Zhao, Z. Xiao, A. Markham, N. Trigoni, and Y. Ren. Does BTLE measure up against WiFi? A comparison of indoor location performance. In *European Wireless 2014; 20th European Wireless Conference*, pp. 1–6. VDE, 2014.

[66] C. Tsirmpas, A. Rompas, O. Fokou, and D. Koutsouris. An indoor navigation system for visually impaired and elderly people based on Radio Frequency Identification (RFID). *Information Sciences*, vol. 320, pp. 288–305, 2015.

[67] S. Alghamdi, R. Van Schyndel, and A. Alahmadi. Indoor navigational aid using active RFID and QR-code for sighted and blind people. In *2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 18–22. IEEE, 2013.

[68] J. Faria, S. Lopes, H. Fernandes, P. Martins, and J. Barroso. Electronic white cane for blind people navigation assistance. In *2010 World Automation Congress*, pp. 1–7. IEEE, 2010.

[69] L. A. Guerrero, F. Vasquez, and S. F. Ochoa. An indoor navigation system for the visually impaired. *Sensors*, vol. 12, no. 6, pp. 8236–8258, 2012.

[70] D. Jain. Path-guided indoor navigation for the visually impaired using minimal building retrofitting. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*, pp. 225–232. ACM, 2014.

[71] A. Fiannaca, I. Apostolopoulous, and E. 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*, pp. 19–26. ACM, 2014.

[72] S. L. Joseph, X. Zhang, I. Dryanovski, J. Xiao, C. Yi, and Y. Tian. Semantic Indoor Navigation with a Blind-User Oriented Augmented Reality. In *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3585–3591. IEEE, 2013.

[73] M. Spindler, M. Weber, D. Prescher, M. Miao, G. Weber, and G. Ioannidis. Translating Floor Plans into Directions. In *Proceedings of the 13th International Conference on Computers Helping People with Special Needs*, pp. 59–66. Springer, 2012.

[74] J. Z. Flores and R. Farcy. Indoor Navigation System for the Visually Impaired Using One Inertial Measurement Unit (IMU) and Barometer to Guide in the Subway Stations and Commercial Centers. In *Proceedings of the 14th International Conference on Computers Helping People with Special Needs*, pp. 411–418. Springer, 2014.

[75] T. Gallagher, E. Wise, H. C. Yam, B. Li, E. Ramsey-Stewart, A. G. Dempster, and C. Rizos. Indoor navigation for people who are blind or vision impaired: where are we and where are we going? *Journal of Location Based Services*, vol. 8, no. 1, pp. 54–73, 2014.

[76] D. Zhang, D.-J. Lee, and B. Taylor. Seeing Eye Phone: A Smart Phone-based Indoor Localization and Guidance System for the Visually Impaired. *Machine Vision and Applications*, vol. 25, no. 3, pp. 811–822, 2014.

[77] N. Fallah, I. Apostolopoulos, K. Bekris, and E. 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*, pp. 425–432. ACM, 2012.

[78] K. Shinohara and J. 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*, pp. 705–714. ACM, 2011.

[79] T. L.-B. Pape, J. Kim, and B. Weiner. The shaping of individual meanings assigned to assistive technology: a review of personal factors. *Disability and Rehabilitation*, vol. 24, no. 1–3, pp. 5–20, 2002.

[80] M. L. Riemer-Reiss and R. R. Wacker. Factors Associated with Assistive Technology Discontinuance Among Individuals with Disabilities. *Journal of Rehabilitation*, vol. 66, no. 3, pp. 44–50, 2000.

[81] P. Parette and M. Scherer. Assistive technology use and stigma. *Education and Training in Developmental Disabilities*, pp. 217–226, 2004.

[82] J. Nielsen. Iterative user-interface design. *Computer*, vol. 26, no. 11, pp. 32–41, 1993.

[83] J. Karat. Evolving the Scope of User-centered Design. *Communications of the ACM*, vol. 40, no. 7, pp. 33–38, 1997.

[84] A. K. Dey. Understanding and Using Context. *Personal and Ubiquitous Computing*, vol. 5, no. 1, pp. 4–7, 2001.

[85] N. Bradley and M. Dunlop. Navigation AT: Context-aware Computing. In M. A. Hersh and M. A. Johnson (Eds.), *Assistive Technology for Visually Impaired and Blind People*, pp. 231–260. Springer London, 2008.

[86] Recommendation Y.2060. Overview of Internet of Things. ITU-T, 2012.

[87] S. Kim. Twitter's IPO Filing Shows 215 Million Monthly Active Users. *ABC News*, 3 October 2013. `http://abcnews.go.com/Business/twitter-ipo-filing-reveals-500-million-tweets-day/story?id=20460493`. Accessed 5 November 2016.

[88] S. Moore. Gartner Highlights Top 10 Strategic Technology Trends for Government. Gartner Newsroom, 2015. Available at: `http://www.gartner.com/newsroom/id/3360317`.

[89] A. DuVander. 9,000 APIs: Mobile Gets Serious. `http://www.programmableweb.com/news/9000-apis-mobile-gets-serious/2013/04/30`, 2013. Accessed 5 November 2016.

[90] A. Guruge. *Web Services: Theory and Practice.* Digital Press, 2004.

[91] J. Kim, M. Bessho, N. Koshizuka, and K. Sakamura. Mobile applications for assisting mobility for the visually impaired using iot infrastructure. In *Proceedings of 2014 TRON Symposium (TRONSHOW)*, pp. 1–6. IEEE, 2014.

[92] R. Halstead-Nussloch. The Design of Phone-based Interfaces for Consumers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, vol. 20, pp. 347–352. ACM, 1989.

[93] M. Marics and G. Engelbeck. Designing voice menu applications for telephones. In M. Helander, T. Landauer, and P. Prabhu (Eds.), *Handbook of Human-Computer Interaction*, pp. 1085–1102. Elsevier, 1997.

[94] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, vol. 63, no. 2, pp. 81–97, 1956.

[95] K. L. Norman. *The Psychology of Menu Selection: Designing Cognitive Control at the Human/Computer Interfacee.* Greenwood Publishing Group Inc, 1991.

[96] J. Kim, M. Bessho, N. Koshizuka, and K. Sakamura. SaSYS: A Swipe Gesture-Based System for Exploring Urban Environments for the Visually Impaired. In G. Memmi and U. Blanke (Eds.), *Mobile Computing, Applications, and Services*, pp. 54–71. Springer International Publishing, 2014.

[97] S. K. Kane, C. Jayant, J. O. Wobbrock, and R. 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*, pp. 115–122. ACM, 2009.

[98] K. Sakamura and C. Ishikawa. Internet of Things—From Ubiquitous Computing to Ubiquitous Intelligence Applications. *Internet of Things-Global Technological and Societal Trends From Smart Environments and Spaces to Green ICT*, pp. 115–141, 2011.

[99] K. Sakamura and G. Santucci. Japan-Europe Cooperation on ucode tech-

nologies. In I. G. Smith (Ed.), *The Internet of Things 2012: New Horizons*, pp. 340–351. CASAGRAS2, 2012.

[100] K. Sakamura. Ubiquitous ID Technologies 2011. `http://www.t-engine.org/ja/wp-content/themes/wp.vicuna/pdf/en_US/UID910-W001-110324.pdf`, 2011. Accessed 13 July 2016.

[101] R. Golledge, R. Klatzky, J. Loomis, and J. Marston. Stated preferences for components of a personal guidance system for nonvisual navigation. *Journal of Visual Impairment & Blindness*, vol. 98, no. 3, pp. 135–147, 2004.

[102] R. Guy and K. Truong. CrossingGuard: Exploring Information Content in Navigation Aids for Visually Impaired Pedestrians. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 405–414. ACM, 2012.

[103] J. Oliveira, T. Guerreiro, H. Nicolau, J. Jorge, and D. Gonçalves. Blind People and Mobile Touch-based Text-entry: Acknowledging the Need for Different Flavors. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 179–186. ACM, 2011.

[104] N. Koshizuka. IoT, Ubiquitous Computing, and Open Data for Smart Environments. *ITU-T Workshop on "Internet of Things Trend and Challenges in Standardization"*, 2014.

[105] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock and N. Meshkati (Eds.), *Advances in Psychology*, vol. 52, pp. 139–183. North-Holland, 1988.

[106] S. K. Kane, M. R. Morris, A. Z. Perkins, D. Wigdor, R. E. Ladner, and J. O. Wobbrock. Access Overlays: Improving Non-visual Access to Large Touch Screens for Blind Users. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pp. 273–282. ACM, 2011.

[107] J. Kim, M. Bessho, S. Kobayashi, N. Koshizuka, and K. 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*, pp. 604–611. ACM, 2016.

[108] A. A. Alkhanifer and S. Ludi. Visually Impaired Orientation Techniques in Unfamiliar Indoor Environments: A User Study. In *Proceedings of the 16th*

International ACM SIGACCESS Conference on Computers & Accessibility, pp. 283–284. ACM, 2014.

[109] T. P. McNamara. Mental representations of spatial relations. *Cognitive Psychology*, vol. 18, no. 1, pp. 87–121, 1986.

[110] D. A. Ross and G. W. Kelly. Filling the gaps for indoor wayfinding. *Journal of Visual Impairment & Blindness*, vol. 103, no. 4, p. 229, 2009.

[111] N. E. Klepeis, W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, W. H. Engelmann, et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *Journal of Exposure Analysis and Environmental Epidemiology*, vol. 11, no. 3, pp. 231–252, 2001.

[112] X. Chen and A. L. Yuille. Detecting and reading text in natural scenes. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 2, pp. 366–373. IEEE, 2004.

[113] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, et al. VizWiz: Nearly Real-time Answers to Visual Questions. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, pp. 333–342. ACM, 2010.

[114] A. Guo, X. A. Chen, H. Qi, S. White, S. Ghosh, C. Asakawa, and J. P. Bigham. VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 651–664. ACM, 2016.

[115] B. Martínez-Pérez, I. De La Torre-Díez, and M. López-Coronado. Mobile Health Applications for the Most Prevalent Conditions by the World Health Organization: Review and Analysis. *Journal of Medical Internet Research*, vol. 15, no. 6, p. e120, 2013.

[116] Apple Press Info. App Store Tops 40 Billion Downloads with Almost Half in 2012. `http://www.apple.com/pr/library/2013/01/07App-Store-Tops-40-Billion-Downloads-with-Almost-Half-in-2012.html`, 2013. Accessed 13 July 2016.

[117] M. Ehrenkranz. iOS App Sales vs. Android App Sales: Apple's iTunes Store Earns More Revenue, Google Play Gaining Ground. `http://www.`

idigitaltimes.com/ios-app-sales-vs-android-app-sales-apples-itunes-store-earns-more-revenue-google-play-gaining-ground, 2013. Accessed 1 October 2016.

[118] R. T. Llamas, E. Taylor, F. Jeronimo, K. Liew, K. Shikita, and P. Soriano. IDC FutureScape: Worldwide Wearables 2016 Predictions. International Data Corporation, 2016.

[119] H. Ye, M. Malu, U. Oh, and L. Findlater. Current and Future Mobile and Wearable Device Use by People with Visual Impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3123–3132. ACM, 2014.

[120] A. Guo and T. Paek. Exploring Tilt for No-touch, Wrist-only Interactions on Smartwatches. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 17–28. ACM, 2016.

[121] J. R. Marston, J. M. Loomis, R. L. Klatzky, and R. G. Golledge. Nonvisual route following with guidance from a simple haptic or auditory display. *Journal of Visual Impairment & Blindness*, vol. 101, no. 4, pp. 203–211, 2007.

[122] D. Bandyopadhyay and J. Sen. Internet of Things: Applications and Challenges in Technology and Standardization. *Wireless Personal Communications*, vol. 58, no. 1, pp. 49–69, 2011.

[123] J. Hendler. Data Integration for Heterogenous Datasets. *Big Data*, vol. 2, no. 4, pp. 205–215, 2014.

[124] J. Bonaguro. Open Data in San Francisco: Institutionalizing an Initiative. July 14, 2014.

[125] ITU-T. Location matters: Spatial standards for the Internet of Things. *ITU-T Technology Watch Report*, September 2013.