Improving Auditory Navigation in Public Buildings for Blind and Visually impaired People

Xiaowei Wang

TORY

audigator

NAVI

Collaborating with Stichting Geluid in Zicht
Supervisor Team:
Dr. R. van Egmond (TU Delft)
ir. R. J. Jansen (TU Delft)
Hannes Wallrafen (Geluid in Zicht)
Henke Baars (Geluid in Zicht)

PREFACE

Having great empathy for blind and visually impaired people, I started this project as my graduation project eight months ago focusing on the improvement of accessibility to public buildings for them. As a master student of Design for Interaction, I approached the goal from the perspective of an interaction designer. By exploring the possibility of sonification, tryouts were made throughout the project to delivery effective and simple navigational information of the interior of public buildings to blind and visually impaired people.

This thesis is the final outcome of this eight-month graduation project. It documents the process from problem explorations to final concept design and evaluation. The assignment was initially explored into broad directions (including sound design, user study of blind and visually impaired people, navigation methods and programming) and thereafter narrowed down to a concrete concept through diverse iterative research by design cycles. Interesting results and remarkable findings throughout the project were also included.

The project was developed on behalf and in cooperation with Geluid in Zicht, a Dutch organization specializing in sound design for blind and visually impaired people. Their involvement provided first-hand insights on inclusive design as well as offering blind and visually impaired participants. ID-StudioLab of Delft University of Technology offered theoretical and technical support on sound design and user-centered design. All these supports greatly helped me in pursuing my goal.

With joy and sorrow, this project has come to an end. However, this is definitely not the end for my exploration in the non-visual world. I have no doubt that I will continue designing to benefit people of all ages and abilities.

Have fun reading!

Xiaowei Wang

January 2012

ACKNOWLEDGMENT/致谢

This thesis would not have come into existence without the help of a lot of people. Firstly, I owe my deepest gratitude to my supervisor team: René van Egmond, Reinier Jansen, Hannes Wallrafen, and Henke Baars. Your supervision and support truly helped the progression and smoothness of this project. All studies during the project would be nothing without the enthusiasm and imagination from all of you. I would specially emphasize my gratitude to my daily supervisor Reinier Jansen, the person I faced at least once a week. I am not able to count how many meetings I had with you, and how much time you spent on this project, even your personal time. I am now able to switch from “I will have a try” to “I will do it”. Thank you Reinier.

It was a great honor to work with Geluid in Zicht and its partners. I would also like to thank Ans Withagen, Robert de Kloe and Marten van Doorn, who helped me organize user studies and field explorations.

My grateful thanks also go to all people from ID-StudioLab for your support, your inspiration, and your shared enthusiasm. Special thanks are given to Aadjan ven der Helm and Rob Luxen for your help throughout this project.

Great appreciations go to the contribution of Muyun Xiao. It was a pleasure sharing thoughts with you, getting inspired and encouraged. I am not able to name all my friends supporting me during this project and all the participants. Thank you all.

Finally, I would like to deeply thank my parents for their support. Going back home for Spring Festival with you is always a force driving me forward. A paragraph in Chinese is specially given to them:

由衷地感谢我的父母 , 谢谢你们对我的支持。回家过年一直是驱使我尽快完成这个项目的动力。

ABSTRACT

Compared with people with normal vision, blind and visually impaired people (B&VI) suffer from not only limited accessibility to public buildings but also poor experience of these buildings. However, the keen senses of sound and touch of B&VI, which are honed through daily practice and compensated by the loss of vision in brain, could be utilized to make up the inexistence of sight in navigation functionally and experientially. By exploring the possibilities of sonification, this project was intended to design an auditory navigation system of the interior of a public building for B&VI users. The aim of this system is to help B&VI to navigate in a building independently without any difficulties and to satisfy their curiosity to explore the building as well.

The main approach of this project is research by design. The research includes field explorations of several public buildings and a literature study. According to the findings of literature study, B&VI have the same ability to perceive, process and understand spatial concepts as people with normal vision do, but that they do so more slowly and in different ways. Visual information can provide direct spatial relations of objects, while the auditory or tactile information that can be used for recognizing spatial relations have to be stored in memory and processed one by one. Previous studies [19] showed that the main difficulty in indoor navigation and orientation is the missing of known landmarks. Based on the results mentioned above, a scale model for indoor navigation with auditory landmarks of known sound could help B&VI to establish sense of spatial relations. Designing such a scale model needs two steps: selecting the representative auditory landmarks and integrating the auditory landmarks to the scale model.

User interviews were conducted to obtain insights of general navigation strategies of B&VI, which contributed to the designing of the frame of the model. Basic design requirements were generated and eight common landmarks, including landmarks for toilets, receptions, canteens, offices, elevators, staircases, doors, and smoking rooms, were selected for further development. In addition, a sound test with blind participants was conducted to select representative sounds.

In order to better analyze how people perceive space through auditory information and how to make the best use of auditory landmarks for indoor navigation, two user studies were conducted with blindfolded participants. These studies showed that the most effective auditory information tends to be clear signal sounds rather than environment sounds recorded in reality.

Through iterative building-and-testing cycles, the final concept Audigator was designed. Audigator is an interactive sound system, which could be applied to interior scale models, for navigation inside public buildings of blind and visually impaired people. With the help of a finger sleeve in blue color, Audigator could track the hand movement of users in the scale model by color detection with a webcam. Sound will be processed and played according to the relative position of finger and the preset positions of sound sources in the scale model. Quadraphonic sound setting of four speakers is applied to provide surround sound. Combined with the tangible objects representing sound sources, which are installed in the scale model, the sound play can help users to recognize the main functional objects or areas in the scale model and provide navigational information to the users.

Table of contents

1. Introduction

Compared with people with normal vision, blind and visually impaired (B&VI) people suffer from not only limited accessibility to public buildings but also poor experience of these buildings. When designing a public building and its facilities, usually insufficient attention has been paid to the demands of B&VI people and the ways they can experience the exterior and interior of a public building. "Geluid in Zicht" (Sound in Vision in Dutch, GiZ) is an organization that focuses on improving the accessibility and enriching the experience of public building for the B&VI. This graduation project is one part of the program and fully supported by GiZ and its partners.

Previous research [1] stated that people who have been blind from birth use visual parts of their brain to hone their sense of sound and touch. These keen senses could be used to help the B&VI better navigate their world including the public buildings. By exploring the possibilities of sonification, this project was intended to design an auditory navigation system for the interior of a public building for the B&VI. At the same time, people with normal sight can also use the new navigation system to enrich their experience of the building.

1.1 Problem Definition

Despite of the visual limitations, blind and visually impaired people are also the target users of most public buildings, such as hospitals, theatres and city halls. However, the special needs of the B&VI were frequently neglected when designing a public building. As a result, the accessibility of B&VI people to such a building is not sufficient enough to prevent them from potential hazards. Even for those buildings designed friendly to B&VI users, usually, the aspects of independence, dignity and experience of the B&VI were compromised for safety and functional purposes. Furthermore, overlooking the B&VI's exceptional auditory sensations could be seen as a waste of experiential resource.

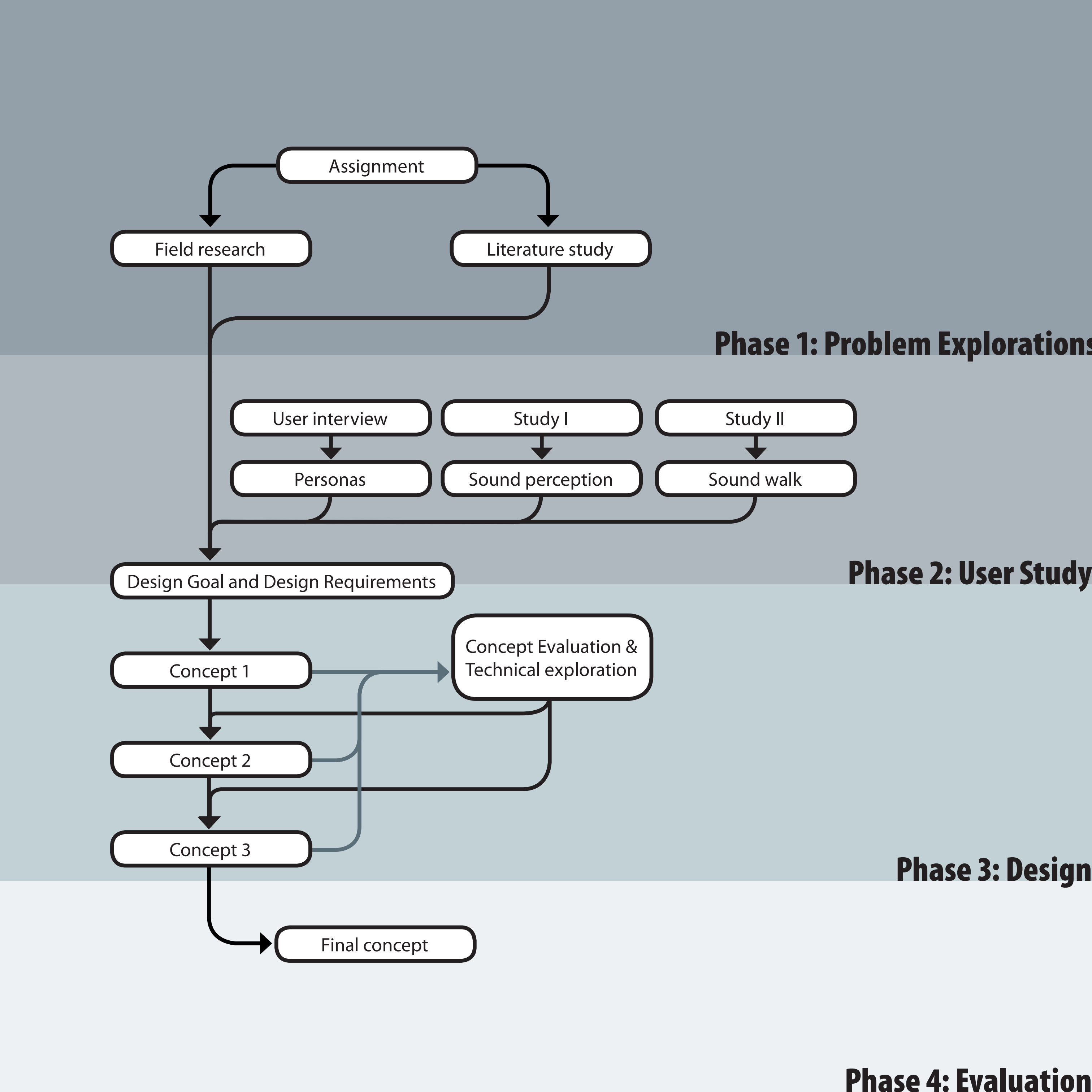
1.2 Assignment

Firstly, several questions regarding the main purpose of the project should be answered, such as "How to scale the acoustical environment of a building, and how does it influence one's experience" through research by design iterations. Based on the results, an initial interactive navigation system, such as a scale model of the interior of a public building would be developed. Afterwards, user tests would be conducted with the prototypes. The results were used to refine the prototypes. The design goal and design requirements would be identified according to users' needs. Although most acoustic features are closely related to a specific building, instead of focusing merely on the characteristics of the specific building, universal design criteria of public buildings for the B&VI drawn from this project should also be documented. It could be applied to design for other public buildings and contributed to the GiZ project. Since audition was selected as the main modality to explore, the concept should preferably base on sound. If possible, other sensations with experimental or functional values could also be used to achieve the project goal. Aiming at the B&VI meant the principles of safety, dignity, and independence should be taken into consideration throughout the project. The interactive design concept should be tested and evaluated by the B&VI, and the final results should be able to contribute to the GiZ project for further development. As the design concept would not only be used by the B&VI, attractions to sighted people should also be emphasized.

1.3 Approach

Figure 1.1 >>
Structure of this project

The project is a research through design project (Figure 1.1), including iterative research by design cycles. The first phase of problem explorations (Chapter 2) consists of a literature study and field research. Literatures relating to the B&VI, sound


Assignment
Field research
Literature study
Phase 1: Problem Explorations
User interview
Study I
Study II
Personas
Sound perception
Sound walk
Design Goal and Design Requirements
Phase 2: User Study
Concept 1
Concept Evaluation &
Technical exploration
Concept 2
Concept 3
Phase 3: Design
Final concept
Phase 4: Evaluation

design, public design, and scale models of buildings were all involved to build a solid theoretical background of this project. Meanwhile, a field research was conducted in public buildings to get general insights of how the B&VI navigate in public buildings. The design goal was thus confirmed. Design requirements, which set the boundaries for whole project, concluded this phase. The second phase is user studies. Within this phase, interviews and observations (Chapter 3) on the B&VI were carried out. The results from the user studies were used as guidelines for this project. Moreover, personas were made based on the user analysis to specify the target group. The design phase (Chapter 5), including several iterative sub cycles, were based on the findings of two previous phases. Since it is an iterative research by design project, within each iterative cycle, user studies and evaluation were involved. Selected public buildings were also examined within this phase. By the end of this phase, a final concept was generated for the final develop phase. ID-Studiolab of Delft University of Technology (TU Delft) supported the design phase.

1.4 Partners

Rijksgebouwendienst ensured the access to public buildings, Koninklijke Visio and Bartiméus provided B&VI participants, and educational partners, including Delft University of Technology, Radboud University Nijmegen, Utrecht University and Utrecht School of the Arts, offered the theoretical and technical support.

2. Problem Explorations

In this chapter, the problems in this project will be further explored in order to formulate a design goal. The study was started with field explorations of several public buildings and then the literature study was conducted. The field research involving six public buildings and four of which will be considered for further research. The acoustic, functional, and experiential aspects of the public buildings were analyzed, and recommendations for future research and design were given based on the exploration. The main findings of the literature study will be concluded in three directions: spatial perception of B&VI people, sound localization, and indoor navigation. The design goal identified from the exploration and detailed design requirements drawn based on the findings will conclude this chapter.

Figure 2.1
Faculty of Architecture, TU Delft

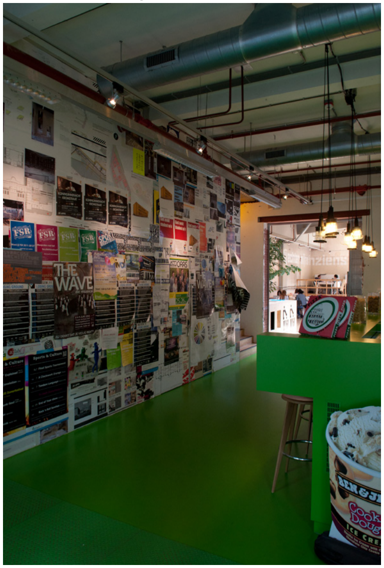
2.1 Initial Exploration of Public Buildings

The initial explorations were planned while the literature search were conducted, to get an initial inspiration on how B&VI people explore a public building. On the one hand, the findings of literature study could be verified through explorations. On the other hand, new insights from explorations could provide more perspectives for literature search.

Faculty of Architecture, TU Delft

The Faculty of Architecture at TU Delft is the largest faculty of the university and one of the largest architecture schools in the world. After the old building was completely burned down in 2008, the faculty has moved to the new building (Figure 2.1) which was initially designed as a church. The huge building consists of three main areas: Oost, West, and Zuid. In total, there are two huge lecture theatres, seven big lecture halls, twelve medium and small size classrooms, and also two restaurants. All of the spaces are accessible to disabled people. Since it is redecorated from a former church, the heights of ceiling varies with the size and function of the space. Moreover, numerous varieties of materials are used for indoor decoration. Therefore, the building is ideal for the initial exploration for this project.

Figure 2.2
Corridor leading to cafeteria



The researchers were equipped with a recorder to document the whole process of exploration* and to enable active listening. Three routes were planned to cover the three main areas of the building. Each route was initially explored with eyes open and then blindfolded for comparison. Through the exploration, several interesting findings were obtained.

The difference in ceiling height and size of the rooms can be clearly characterized by sound. For example, there is a corridor (Figure 2.2) with much lower ceiling and in narrower width leading to the main cafeteria. When stepping into the cafeteria from the corridor, its spaciousness could be noticed instantly. Therefore, echo or reverberation characteristics of a space is potentially useful to represent a space in the design phase.

When blindfolded, close-range sound sources caught researchers' attention more easily than far-range sources. For example, during the exploration with eyes open, the sound sources were treated equally and were categorized according to visual positions into groups. However, when blindfolded, the researcher tended to be more sensitive and cautious to close-range sound sources such as people walking by.

When multiple sound sources existed, the researcher tended to follow one and to

* all mentioned files are included in digital appendix (Chapter 2/ exploration in BK/)

ignore other sound sources. For example, in an area for prototyping, there were multiple groups of students working. When eyes were open, the focus of listening kept on switching from group to group. On the contrary, when blindfolded, the attention was focused on a certain group for much longer time before switching to another group. Several groups were not even noticed.

When walking blindfolded, the change of materials of the floor can be identified clearly by sound. Firstly, it was assumed that the notion of material changes could be attributed to the tactile feeling of friction between the sole and the floor. However, when replaying the records, it was discovered that the material change can be noticed only by the sound of friction.

Figure 2.3
Building of Koninklijke Visio Onderwijs (photo by Hanneke Hollander)

As mentioned in the literature study [2, 3], when blindfolded, the researcher felt insecure when walking in large open space, such as in the lecture hall, without any assistance.

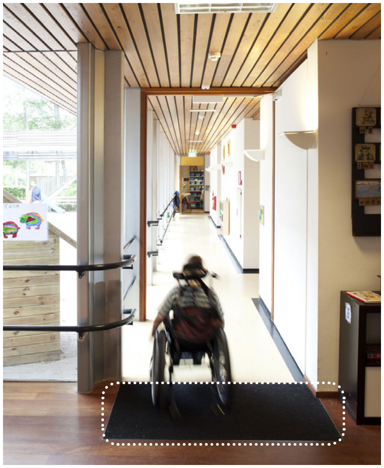
The findings of this initial exploration led to research questions in the user research phase and are elaborated in Chapter 3.

Building of Koninklijke Visio Onderwijs

Koninklijke Visio is a Dutch organization that provides visually and intellectually challenged people with residential care and education, and Koninklijke Visio Onderwijs (Figure 2.3) is a school for visually impaired children. The new building of Koninklijke Visio Onderwijs is a good example of well-designed public buildings for B&VI users. To obtain some inspirations, the new building was explored and Ans Withagen, a specialist of blind children education, was interviewed.

Koninklijke Visio Onderwijs is a two-floor building* with a swimming pool on the roof. The main structure of the building is a double-'H' shape without large open space. Withagen mentioned that the double-'H' structure can provide B&VI people more assistance in order to prevent them from getting lost. Large open space is avoided for the same reason. The ground floor of the building is used for classrooms, while the first floor is mainly for offices and other special purposes, such as painting room and indoor playground. All the accessible areas for blind and visually impaired students are equipped with handrails along the wall, which could provide secure feelings for the B&VI. Withagen mentioned that the B&VI, especially congenitally blind people, have more difficulty perceiving curves. For example, blind people will be likely to

Figure 2.4
A corridor in the study area in Koninklijke Visio Onderwijs. The special carpet (white dot line) is placed at each turning point in the building. The B&VI children can make use of it as reference (photo by Hanneke Hollander)



* included in digital appendix (Chapter 2/ Building of Koninklijke Visio Onderwijs/)

get lost if they are following a curve route. Therefore, curves and all other curved shapes are avoided in the building. The planned routes for the B&VI are divided into straight segments which are connected by 90 degree angle. A different material of the floor, carpet, is used for the turning joint (see Figure 2.4). When talking about the spatial perception of the B&VI, Withagen confirmed that they have the potential to comprehend spatial relations. Through training and practice, congenitally blind people are able to make maps of their familiar buildings. Furthermore, to establish a more pleasant environment, sound absorbing materials, mostly wood, were installed on the ceilings. According to Withagen, wood is also the favorite material of the B&VI because of its warm temperature and soft touching.

Huge open spaces are frequently used in public buildings. Therefore, when designing for B&VI users, clear routes with reference points or landmarks or lines should be provided to prevent them from getting lost. Withagen also verified the design requirements that curved routes should be avoided and short straight segments should be used.

Although the accessibility of normal public buildings is not comparable to the building of Koninklijke Visio Onderwijs, universal rules of the B&VI's navigation applied in this building could be used in the design phase.

2.2 Research of Public Buildings

A research* was conducted in four public buildings: Den Haag city hall, Binnenhof, Castle Muiderslot and Muziekgebouw aan't IJ (Figure 2.5). The buildings were chosen

Figure 2.5
Four public buildings of the research: (from left) Den Haag city hall, Binnenhof, Muziekgebouw aan't IJ and Castle Muiderslot (photos from Internet)

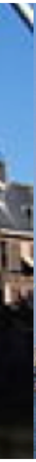
* more pictures and recordings are included in digital appendix (Chapter 2/ Public buildings/)

by GiZ since they are representative for public buildings with accessibility to the B&VI. The purpose of this research was to gather common features of public buildings for the design phase and select certain buildings as cases for the user research phase. Moreover, in this research, features of the buildings that could contribute to the design phase would be generated. Diverse aspects involving both functionality and experiential perception were analyzed, with extra attention to acoustic aspect. The existed accessibility issues of the buildings would also be examined.

During the research, pictures were taken for later analysis and recorders were used to document sound features. The buildings were examined by following routes designed for B&VI users. Each research was conducted with local guides (management officers) and tutors. Furthermore, Hannes Wallrafen, a tutor from GiZ, is visually impaired and he offered a lot of insights as a target user during the exploration.

Starting points

According to the design requirements generated from literature study, a clear starting point is essential to a navigation tool, especially for the B&VI. Among these buildings, Castle Muiderslot has a clear starting point – only one entrance to the castle. Although there are multiple entrances, Den Haag city hall also has a clear starting point since only the main entrance has accessibility to people with disabilities, and the reception is also close to the main entrance. However, the starting points of Binnenhof and Muziekgebouw aan't IJ are not as clear as the other two buildings. The layout of Binnenhof is too complicated to draw a starting point. Muziekgebouw aan't IJ has three entrances connecting respectively to pedestrian, parking lot, and public transportation.



Destinations

Den Haag city hall and Binnenhof are clearly function-oriented. Although there are also tourists, these two buildings have clear destinations based on functionality, such as wedding room, counters, and a conference room for Den Haag city hall, and meeting rooms for Binnenhof. Castle Muiderslot is a place of interests that attracts mostly tourists. Typically, the experience-oriented buildings do not have obvious destinations. However, Castle Muiderslot provides two routes for tourists: a knight route for children and a tower route for other visitors. Such routes could be seen as destinations of Castle Muiderslot. Muziekgebouw aan't IJ could be a good example representing the importance of both functional and experiential aspects. Tourists could wander freely in the modern building while the various concert halls are clear destinations for audiences.

Acoustic characteristics

Considering the different functions and structures of the four public buildings, the acoustic characteristics vary apparently. Due to the atrium and low sound absorbing materials (marble, metal and glass) applied, Den Haag city hall provides a non acoustic friendly environment with complex reverberation. To make it even worse, a restaurant, two semi-open elevator halls, a waiting area and several stores, which provide disturbing sounds, are also located in or connected to the main hall. As a consequence, additional ceilings or separations were built for the working areas (reception, counters) in the main hall. By contrast, Binnenhof offers a better-designed indoor acoustic environment. Thanks to the jagged glass ceilings and louvered shutters used on every window, echo is efficiently mitigated. Double layer windows and other soundproof materials are used for the conference rooms and there is no open lounge or coffee corner. Therefore, noise inside Binnenhof is controlled at a significantly lower level compared to Den Haag city hall. Castle Muiderslot is located in an open area near a river. The ancient wood windows all around the castle are not soundproof. Therefore, wind influences much on the acoustic environment of rooms with windows inside the castle. On the one hand, the sound of wind could be seen as noise, which may deteriorate the localization of the B&VI. On the other hand, the experiential aspect of wind could be attractive to B&VI visitors. The interior of the castle, even floor, is commonly made of wood that sound from rooms nearby and upstairs could be heard clearly. Moreover, several interactive games for children with sound effects are placed separately in different areas inside the castle. It is relatively difficult to locate a sound source in the rooms. The acoustic performance of Muziekgebouw aan't IJ is perfectly calculated and designed inside the music hall. Even the chairs for audience are specially made to reduce direct reflection of sound. Since double layer structure is applied in order to isolate sound, the acoustic environment outside is totally different from the inside of the music hall. For the outside, wooden materials are used for stairs and floor while concrete is applied to the outer layer of the sound isolation wall. The approximately 20-meter high ceiling is made of glass with steel supporting structures. The feeling of large open space when talking is quite obvious.

Locations for navigation tools

The ideal location for a navigation tool, which is preferably close to the starting point, should be accessible by B&VI users. Sufficient space should be provided for the free exploration of the navigation tool as well as for the sound play. Therefore, crowded areas inside the building should be avoided. Since Den Haag city hall and Castle Muiderslot have clear starting points, the navigation tools could be placed near the starting points. In spite of multiple entrances, Muziekgebouw aan't IJ offers an ideal location for the navigation tool – by the side of the ticket box. However, the complex layout of Binnenhof raises a baffling question to place the navigation tool. The compromising solution is to place the navigation tool in the conjoining point of four branches.

Conclusion

After examining the four public buildings based on aspects of starting point, destination, acoustics, and ideal locations for navigation tools, Den Haag city hall was selected as the ideal building for further analysis for the following reasons:

1. There is a clear starting point for B&VI visitors - the main entrance.
2. An ideal location for a navigation tool is provided near the starting point.
3. The building is function oriented and the destinations are clear. Routes could be planned according to destinations.
4. Public area and security area of the building are clearly separated.
5. The building is accessible for research.

2.3 Literature Explorations

A literature study was conducted on three main topics: blind and visually impaired people, indoor navigation and sound localization. The main findings are presented below.

2.3.1 Spatial Perception of the B&VI

For most people, vision plays an important role in establishing precise and real-time perception of spatial information. When it comes to B&VI people, especially congenital blind people, it would therefore seem reasonable to assume that they will have difficulty conceptualizing spatial relations. Indeed, many studies have focused on such areas to explore the performance of people who are blind or with visual impairment. These studies have produced contrasting findings, and three broad approaches to research into the area of how individuals who are blind or partially sighted perceive space and comprehend spatial relations can be identified: the 'deficiency theory', the 'inefficiency theory' and the 'difference theory' [4, 5, 6, 7, 8].

'Deficiency theory' states that the lack of visual experience inevitably results in a lack of spatial understanding. According to this theory, without visual assistance, the spatial information collected by the blind exists in the state of segments, which could not be connected or formed to a spatial understanding such as an overview. However, research evidence has showed that this is an extreme view [5, 6, 7, 9, 10, 11, 12]. Recent studies show that even the congenital blind are able to reestablish the spatial relation in their mind based on their other sensations [13].

'Inefficiency theory' states that blind people can understand and mentally manipulate spatial concepts but are not as effective as sighted individuals in their perception of space, largely because their information is derived from auditory and haptic cues which are regarded as less rich sources than vision [14, 15, 16]. Rieser's study [15] indicates that visual experience plays an important role in the development of sensitivity to changes in perspective structure. For the congenital blind, since they do not have visual experience, the perspective structure of spatial information is mainly depended on analogical thinking and deduction of former experience. Special training is needed in order to develop these skills. For example, three objects A, B, and C exists in one room, B is above A, and C is to the left of A. According to this theory, a congenitally blind person is able formulate the spatial relations of the three objects in their mental picture by linking B and C relatively to A. Without a holistic view, if A is removed from the room, then the relation between B and C does not exist any longer in the mental picture. Holistic spatial perception of A, B, and C can only be established through special education.

'Difference theory' states that B&VI people have the same ability to process and understand spatial concepts as the sighted, but that they do so more slowly and by different means [17]. Visual information can provide spatial relation of objects directly, while the auditory or tactile information related to spatial relation should be stored in memory and processed one by one. Consequently, this process of B&VI people will take more time to finish.

In this project, the 'difference theory' would be applied. The B&VI are able to comprehend spatial relations and rebuild the information by different means. However, considering the specialty of B&VI users, the spatial information should be provided according to their memory and cognitive ability in order to ensure sufficient time to process the reformulation. For example, the detail level of the final concept should be controlled according to the expected time of exploration. Otherwise, too much information provided will overload the short term memory of the users, which may result in the failure to establish a holistic view of the building.

When trying to estimate the size of different objects, B&VI people frequently use self-reference coding strategies [4]. For example, the B&VI estimate the height of ceiling compared with his or her own height. Millar [10] found that external frameworks are harder to access for blind people, and especially congenitally blind people tend to use an adaptive strategy that draws on an egocentric frame of reference.

The perception of surroundings of the B&VI is mostly drawn from previous experience of direct sensory contact with an object. For example, a blind person could estimate the height of his desk or the length of his bed using tactile and kinaesthetic cues. Analogical thinking and deduction from previous experience that involves no direct contact could be another source of spatial perception. For instance, in Andreou's study [4], a congenital blind subject was given tactile cards with different animals in braille (goldfish, blackbird, cat, dog, horse and elephant) and he was asked to put the cards in a row from the smallest to the biggest according to the size of the animals in real life. He put the cards of the animals in the correct order saying that he tried to 'picture' some of the animals that he was familiar with. He assumed that blackbirds are quite small animals from the way they sound. He stated he could hear the blackbirds singing in the morning and they got a high-pitched sound. He assumed that the pitch of the sound an animal makes can be reliably used to determine its size (i.e. a cow produces a low sound because it is big). He apparently establishes a link of sound pitch and the size of animal based on deduction of previous experience. The perception could also be drawn from other auditory information. It is reported that echo plays an important part in enabling the B&VI to comprehend the characteristics of the surrounding environment. For example, the B&VI could work out the size of a room by its sounds and echoes [4]. These auditory perceptions are valuable sources of information for 'sensing' the presence of objects [18]. Moreover, another source could be indirect knowledge which is not tied to obvious personal experiences but is conveyed through the descriptions of other people and secondary sources such as books or media [4].

In conclusion, spatial perception of the B&VI is tightly related to themselves and their former experience. It is necessary to provide realistic and reliable information and feedback in order to facilitate spatial relations for the B&VI.

2.3.2 Indoor Navigation of the B&VI

For indoor navigation, the B&VI have to be navigated through unfamiliar spaces without human guidance. Nordin [19] claimed that there are two essential points to navigate in such environments. Firstly, adequate information about the travel path is needed, so the blind person can walk with confidence and safety. Secondly, recognition of objects through the travel path which is followed is needed [19]. In order to provide adequate information about the travel path, two main basic methods are applied to navigation devices developed: 1) Navigation based on information from sensors, which determine the position of the blind (piloting methods) and 2) find the current position of the blind based on information for the previous position and an estimate of velocity and direction of movements (path integration methods or dead reckoning) [20, 21].

For B&VI people, navigation in unfamiliar buildings is even more difficult than outdoors because of the poor reception of GPS (Global Positioning System, which is widely used for outdoor navigation) signal. According to Ivanov [20], the main difficulties in indoor navigation and orientation are that the known landmarks are lost, overcoming obstacles can be risky, not all the blind can read Braille tags, and the price of existing systems for indoor navigation is beyond the purchasing ability of the people with visual disabilities [20]. Moreover, there are more obstacles for the B&VI in unfamiliar buildings that may cause serious injuries, such as stairs, all kinds of steps, elevators, revolving doors, and automatic doors [2].

The needs of an indoor navigation system for the B&VI in unfamiliar buildings are identified. This project is formulated to explore possibilities based on such needs. Existing products are mostly focused on functional aspects of indoor navigation, but the experiential parts have been hardly taken into consideration. However, in the assignment of this project, experiential aspects of public buildings should also be emphasized and communicated to the B&VI. Therefore, the navigation experience of users is also one of the key aspects for this project.

2.3.3 Sound Localization

Despite the long lasting debate of compensatory plasticity of the loss of vision, sound is essential to establish spatial perception for the B&VI. Sound localization relies on the neural processing of implicit acoustic cues [22, 23]. Humans estimate the location of a source by taking cues derived from one ear (monaural cues), and by comparing cues received at both ears (difference cues or binaural cues). Interaural differences in phase (IPDs) and sound level (ILDs) are both employed by the human auditory system to extract the horizontal coordinate of the sound with respect to the head (sound source azimuth). Complex spectral shape cues, which arise from the diffraction of acoustic waves at the pinna cavities, enable the system to determine the position of the sound source in the median plane (sound elevation) and to disambiguate frontal from rear locations. These differences are known as the head-related transfer functions (HRTFs) [22].

According to Zwiers's study [24], signal-to-noise (S/N, noise indicates the semantic noise, such as people talking) ratio of sound sources influences the performance of sound localization of the B&VI more obvious than sighted people. The data showed that for high S/N ratios, localization by the B&VI and sighted subjects is similar for both azimuth and elevation. At decreasing S/N ratios, the accuracy of the elevation response components deteriorated earlier than the accuracy of the azimuth component for both sighted and the blind. However, elevation accuracy of the blind deteriorated much earlier compared with the sighted. The blind are less able to extract the elevation-related spectral cues in the more complex acoustic environment, a condition more reminiscent of the natural situation. For them, it is easier to locate sounds in the horizontal plane than in the vertical plane.

Another study by Fazzi [18] showed that sound cues facilitate localization of one object, but only when they do not 'conflict' with tactile information. In the presence of conflicting tactile and auditory cues, the blind first respond to the tactile information, subsequently reducing their ability to use the auditory information.

In enclosed rooms not only the direct sound of a sound source is arriving at the listener's ears, but also sound which has been reflected at the walls. The auditory system only analyzes the direct sound, which arrives first, for sound localization, but not the reflected sound, which arrives later. So sound localization remains possible even in an echoic environment. This echo cancellation occurs in the Dorsal Nucleus of the Lateral Lemniscus (DNLL) [25].

Human binaural hearing perception makes use of differences in time of arrival (IPDs), overall level differences (ILDs), and spectral differences at the two ears induced by acoustic propagation effects around the torso, head, and pinnae. These differences are categorized by the Head Related Transfer Function (HRTF) [26]. With this information, the human brain is capable of localizing sounds. Using acquired HRTF data, is it possible to render the correct acoustic information for auditory spatial perception, thus allowing the ability to create 3D virtual auditory environments though binaural synthesis over simple headphones. For the virtual display of close range by sound, the range of perceived auditory distance is compressed about two folds relative to the physical range [27, 28]. For instance, in Speigle's study [29], participants were asked to estimate the distance of virtual sound source. The result indicates a clear tendency that there is over estimation of distance for near target (below 400 meters) and under estimation for far targets (above 400 meters). And for target within 100 meters, the indicated distance is approximately twice the physical distance. Therefore, one should be cautious when implementing sound as the only cue for real distance in the design phase.

Dramas [26] verified that the accuracy of localization was directly dependent on the stimulus, azimuth, and distance. It is possible to optimize sound localization, even within the near space, by adjusting the sound source. That means it should be possible to use synthesized sounds to provide information about the location of particular objects. A navigation system could be developed based on these findings.

2.3.4 Results

The ability of B&VI people to perceive space and to comprehend spatial relations has been examined by studies. In order to design for the B&VI, independence, safety and confidence [19] are the key values that should be taken into consideration when designing. Research in this area [2, 3] indicates that the blind people prefer to walk along walls rather than the middle of the room, because normally there is no orientation line for their cane in the middle, and because they try to avoid crowds of people as well. Thus, navigation paths should be aligned along walls. The number of turnings should be minimized. The route should be constructed by short straight segments with 90° angle between them. The blind users can easily identify doors, walls, and stairs with white cane, so such objects could be used as reference in the navigation tool.

Orientation inside a building is important because it provides a frame of reference throughout description. In most cases compass orientation is applied while in some cases a clock face is chosen. Blind people are familiar with both of these forms of orientation and use them regularly in other aspects of their lives [30].

The B&VI prefer to implement navigation with verbal commands to non-speech information. However, according to the research by Loomis [3], virtual spatial sound display is also welcomed by the B&VI, despite the concern of blocking environment sounds by headphones, which could lead to potential danger such as failure to react to emergence. One should be cautious when implementing sound as the only cue for real distance in the navigation tool, considering the potential compression in perception of virtually displayed distance [27, 28].

2.4 Design Goal

According to the findings of literature study, the B&VI have the same ability to process and understand spatial concepts as the sighted, but that they do so more slowly and by different means (§2.3.1). Visual information can provide spatial relations of objects directly, while auditory or tactile information relating to spatial relation would be stored in memory and processed one by one. According to research, the main difficulty in indoor navigation is the missing of known landmarks. Based on the results mentioned above, a scale model with auditory landmarks, which could help to make use of known sound sources as landmarks and establish spatial relations, could fulfill the need of indoor navigation for the B&VI. The task of designing such a scale model was split into two main parts: to select representative auditory landmarks and to integrate auditory landmarks in the scale model.

Based on the findings of literature study and field exploration, the design goal of the design phase of this project is:

To design an interactive interior scale model of a public building with sound navigation system for blind and visually impaired users and to fulfill their curiosity to explore the building.

2.5 Design Requirements Part 1

The scale model should provide a holistic view of the building, orientation inside the building, starting point, main destinations and reference points (landmarks). The landmarks should be represented by or linked to known sources and be identifiable for B&VI users. The overview of the building involves describing the shape of the building and the key subsections of it. The accessible public areas and security areas that are mostly not accessible for public should be separated clearly in the design.

The detailed level of the scale model should be controlled considering the functionality of scale model as well as memory capacity of B&VI users. The scale model should neither provide inadequate information of the routes nor overload users' short term memory. The information provided by sound should not conflict with information from other sources, such as tactile and linguistic. Since the distance perception of sound is not accurate, extra attention should be paid if sound is used as the only clue for distance. Most public buildings are function-oriented, which means the purpose of visiting such buildings is clear before the visit. Therefore, certain routes connecting the starting points to the functional destinations could be predesigned.

3. User Study

In this chapter, findings of a user interview session will be presented. User interviews were conducted to gain insights of general navigation strategies of B&VI, which contributed to the design of scale model, and to drawing frequently used landmarks from public buildings. Design requirements were generated and common landmarks were selected for further development. Two user studies with blindfolded participants were conducted in order to better understand how people perceive space through auditory information, and how to make use of auditory landmarks for indoor navigation. The procedure and findings will also be included in this chapter.

3.1 User Interview

Since the target user group was fixed from the beginning, phases of this project were all carried out as a user-centered design, which tried to optimize the product around how users can, want, or need to use the product, rather than forcing the users to change their behavior to accommodate the product. Therefore, a verbal interview session with B&VI participants is essential to get general insights, to highlight their needs and to set up following researches.

3.1.1 Goal

By interviewing blind and visually impaired people, general insights into their orientating strategy, experience on spaciousness and perception of sound would be obtained. Concomitant findings, which might contribute to the conceptualization and design phase, would be documented and analyzed.

3.1.2 Setup

Semi-structured interviews were conducted with five B&VI participants, covering congenital blind, acquired blind and visually impaired. Considering the influence of navigating tools (such as guide dog, white cane, tracker etc.), participants with diverse travelling habits and tools were selected.

Interview questions were categorized into topics including strategies of orientation, experiences on spaciousness, indoor and outdoor orientation and how they experience scale models. The ratio and number of closed-ended questions and open-ended questions were controlled so that the interview could be finished within 45 minutes. The interview questions can be found in Appendix I.

Figure 3.1
Picture of one user interview. The blue maps on the table were tactile maps provided.

The interviews were conducted in groups, which enabled participants to exchange ideas and inspire each other.

3.1.3 Results and Discussion

A summary of the interview results is presented below. Detail results of each interview can be found in Appendix II.

Strategies of orientation

Landmarks (reference points) play an indispensable role in memorizing a route for the B&VI. Instead of measuring distance by counting steps, all participants confirmed that their navigation relied on remembering landmarks. Criteria for choosing landmarks differ from person to person. However, recognition of landmarks mostly depends on sound, which could be produced by landmarks or made by a white cane. For indoor situation, frequently mentioned landmarks include sideway, staircase, doors, corridors, elevators, changing of materials on floor, and height of ceilings.

Mental pictures were often mentioned by participants during the interviews. The mental pictures were created with useful information collected to memorize a route or to understand the relation between landmarks. However, the mental pictures described by different participants differ apparently. People with more visual experience tend to have a more complete and holistic mental picture consisting of more imagery information. They are able to connect objects in the picture. Therefore, the relations between objects exist in a net structure. By contrast, people who are congenitally blind or with less visual experience seem to have mental pictures including more tactual or auditory information. The information is occasionally fragmentary (each segment consists of the information from a certain area of the whole environment). Therefore, there is a certain linear sequence or direction of the information in their mental pictures. Missing a key segment or putting the segments in random order will consequently result in getting lost. Furthermore, the interviews confirmed that a clear starting point would be helpful for building all kinds of mental pictures.

Experience on spaciousness

The participants stated that they could be aware of the height and size of the space by echo changes when entering a space. Sound emitted by footstep or white cane and the echo could be used as a cue. According to the findings of literature study [4,10], the measurement is different between the B&VI and the sighted. Participants confirmed that they have a notion of distance based on time estimation and echo, rather than counting steps. However, distance is not a key factor in their spatial perception.

Outdoor and indoor navigation

The light sense of some participants still functioned to a degree that they could perceive high contrast color and light. They could make use of this for outdoor and indoor navigation, for example by using the sun to orientate. The auditory landmarks are essentially helpful for both indoor and outdoor navigation. For indoor navigation, the height of ceiling perceived by echo can be used as an additional clue. Inconsistent with the literature, the participants did not tend to walk along the wall because they are able to make use of echoes to keep walking straight. Additionally, participants with a guide dog apply a different strategy compared to participants with only a white cane, because the guide dog can navigate directly to destination and prevent them from potential hazards.

Scale model

Participants suggested that the amount of information provided by the scale model should be controlled: neither too much to overload their memory capacity nor too little not be able to complete the navigation. The legend provided with scale model is necessary. They prefer symbolic information than linguistic because it would take much time to read Braille. The scale of the model is not important according to the participants, but mixing two scales in one model would confuse users.

3.1.4 Conclusion

The information obtained in the interviews was evaluated and some noteworthy findings were listed as follow:

About the blind and visually impaired

In varying degrees, some blind and visually impaired people are still able to perceive light. Thus, it implies that light or high contrast colors with transparent materials could be used as reference or guidance in the scale model to facilitate indoor navigation. For the B&VI, the safety issue always comes first and the functional aspects of a building are more important than the experiential ones.

Orientation strategy

The friction sound between floor and feet are frequently used as a navigation cue, which could also be included in the scale model design. The spatial perception of a room (size, height, distance from wall) is also used by blind users. If the spatial feeling could be created and presented virtually, it can be applied to the scale model as a navigation cue.

Scale model

If a tool (for instance, a tracker, a map or a scale model) is provided, blind and visually impaired users tend to rely more on the tool than their navigation ability. In this case, every effort should be made to include all recognizable landmarks in the scale model. Otherwise, when a landmark, which is not provided by the tool, is encountered in reality, the user will feel lost. Candidate landmarks are office, staircase, door, canteen, smoking room, reception, elevator, and toilet. Since the total amount of information should be controlled considering the memory capacity of users and to help them better focusing on the functional information as well. The non-essential information should be eliminated from the scale model, such as decorations of the building and landmarks from unreachable areas.

Although blind and visually impaired people have diverse choices of starting point selection, the main entrance is generally accepted by most of them. Therefore, in order to provide a clear starting point and prevent confusion, the main entrance or the location of scale model (generally close to main entrance) should be appointed as the starting point. When it comes to the selection of landmarks, the rule of starting point can no longer be applied, since the users may not properly perceive the designed landmarks. Therefore, more sufficient landmarks should be provided in both quantity and variation as well, enabling the users to make their own selection. Moreover, a legend of the scale model would help to describe landmarks.

The scale of the model and distance are not important according to the interviews, which allows the exaggeration of important information. However, since the mix of multiple scales in one model may lead to confusion, the relative distance between landmarks should be based on reality. Additionally, while the users with less visual experience may not be able to establish a holistic mental picture of the scale model, the maximum distance between objects should be controlled to prevent isolated landmarks (far from other objects in the scale model).

Considering the diverse approaches of the B&VI with white cane and with guide dog, the both strategies should be taken into account when designing the scale model.

3.2 Persona

Based on the interviews, B&VI people could be categorized into two user groups according to their travelling habits and personalities.

On group of blind and visually impaired people tends to be more dependent. They rely more on external support from others and they are less explorative in daily lives. Mostly, if possible, they are more willing to travel with partners or in groups. Therefore, when problems are encountered, they would directly turn to help from others.

By contrast to the group introduced above, the other group of people is more independent and adventurous. They feel proud of themselves since they could travel alone without other’s help and they feel like to actively explore the unknown things.

The two groups are better explained and visualized in the personas next page.

Anna is 38 years old. She lost sight since she was eight years old so she learnt to use Braille. She likes travelling and makes every effort to be independent. Therefore, she usually travels alone to all kinds of destinations, such as bank, home-doctor clinic. Her white cane is her only tool for travelling. She has learnt to make use of the tick sound of her cane and its echo to detect the circumstance and locate herself. She goes to work every morning alone and does her groceries mostly in the afternoon in a small supermarket close to her home. For indoor navigation, she prefers something that could enable her to travel and explore independently without exposing her to everyone around as she is blind and she needs help.

BIOGRAPHY, ATTITUDES AND NEEDS

Erik is 44 years old. He became blind 20 years ago because of an accident. He didn't learn to use Braille because he felt it is not necessary. His guide dog, Liam, takes him to his destinations. Thanks to Liam, he is able to do his daily things independently. However he is not a real fan of the sunshine, he would rather stay at home than travel on noisy and crowded street. When he has to travel to unfamiliar places, he will normally go with his wife, and she will describe the surroundings to him. If he has to travel alone, he will bring a tracker, which helps him to locate and navigate. For indoor navigation, he prefers effective stuffs which will enable him to achieve his destination more quickly.

Anna is planning to go to Muziekgebouw aan 't IJ in Amsterdam for a concert. She decided to take public transport. Since it is the first time for her to attend a concert there, she searches online for some information for the concert, the music hall and transportation as well. For the travel, Anna prepares the timetable and transfer plan of train and tram and she also notes that the ticket office which is her first destination is close to the main entrance. The cloakroom is also nearby while the concert hall is on the first floor.

On the way to railway station, **Anna** is very familiar with the surroundings. She has many landmarks with sound, smell and tactile information. She could make use of the slope of ground to locate herself. The tick sound of the cane is used as the indication of ground material, and the echo of it could illustrate the surrounding environment. Tactile sideway is also very helpful. After entering Muziekgebouw aan 't IJ, she prefers to walk along walls to avoid crowd and keep her walking straight. The echo of cane tick is more obvious indoors so she can perceive the space with it. The cashier sound and service talk directly leads her to the ticket office where she finds much information about the concert and the music hall. The lady in the office also shows her the direction which she can follow to the concert hall and the toilet. When she is confused or lost, she always tries to trace back to the previous landmark so that she could start over. She never count steps but she could calculate distance by her walking speed and time. She enjoys a wonderful music night.

TRAVEL PLAN AND PREPARATION

NAVIGATION AND TRAVEL STRATEGY

Erik is attending a wedding of his cousin in Den Haag city hall. Since his wife is fully scheduled, he has to go to the wedding hall alone. Instead of searching online, asking is always more efficient and effective for him. The first information source appears in his mind is his wife. He calls special transportation service to the building. As a backup, he also added the address to his tracker. The only information he needs is the location of the reception where he can ask for help.

Liam is a well trained guide dog that he could keep **Erik** from crowd and potential hazards. Liam can also remember the daily routes of Erik. Thanks to Liam, Erik could walk more directly and faster to his destination. After arriving at the city hall by special transportation, Erik finds the main entrance as his starting point. Normally he will ask for someone to direct him to the reception. However, there is a tactile sideway indoors in Den Haag city hall connecting the main entrance to the reception. So, he is able to find the front desk by himself. The lady of reception shows him the way. Liam keeps him walking straight towards the wedding room. When he gets lost or confused, Liam is able to bring him back to the starting point such as the main entrance and he could find somebody for help.

As demonstrated in the personas, Anna could be seen as an active and independent explorer while Erik is more passive and dependent. Since the passive users as Erik will generally ask for help from others, they rarely use scale model for the navigation tool. Therefore, when designing the functional aspect of the scale model, the focus should be placed on the need of active users. However, the experiential aspect of the scale model could be beneficial and attractive to both active users and passive users.

3.3 Study I: Sound Perception

The design goal of this project, "to design an interactive interior scale model of a public building with sound which provides blind and visually impaired people with information of navigation and fulfills their curiosity to explore the building", can be reached by two main approaches: experiential approach - to fulfill the curiosity, and a functional approach: to navigate in public buildings. This research was conducted for the experiential approach: to explore the possibility of sound presentation and sound perception. Through this research, the research question of "how people perceive, understand and describe sound and sound space?" would be explored. Furthermore, through this research, the most attractive (easy to be noticed) aspect or element of sound in public buildings could be identified and used in the scale model. Additionally, the recordings made in fixed positions would be compared with the ones recorded while moving to figure out the functional way to make the recordings for the development of the scale model. The results of this study would be contributed to the development of a scale model to fulfill curiosity.

3.3.1 Method

Participants

Eleven tests were conducted including one pilot test. Participants selected for this test were all students or recently graduated students from the faculty of Industrial Design Engineering (IDE), Delft University of Technology. They had been studying or working in the faculty building for at least one year, so they were familiar with the environment of the building.

Stimuli

There were six sound clips played to each participant. All the recordings* (.wav format, 24-bit/96kHz linear) were taken by the same recorder (Roland Edirol R-09HR) with two built-in omnidirectional microphones. Four of the recordings were taken in IDE, which were used as sound of familiar space for the participants. By contrast, the other two were recorded in Den Haag city hall, which could be seen as an unfamiliar space.

* all mentioned files are included in digital appendix (Chapter 3/ study I/)

Among these sound clips, three were recorded while moving and the others were in fixed positions in order to test the perception of relative movement between sound sources and recorder.

Recording No.1 was 31 seconds long and was taken on the ground floor of IDE, pointing at two elevators (see Figure 3.2 left). In this recording, the room where the recorder was placed is approximately 30 square meters and four meters high with glass walls. There is an automatic door to the right of the recorder. There were background sound sources including the traffic outside the glass wall on two sides, people walking outside the building and sound of machinery in workshops nearby. Meanwhile, the foreground sound sources of automatic door and elevator were relatively apparent. Since the room has a closed environment, the background sounds from distance were relatively low in volume in this recording. Therefore, compared to other recordings, the foreground sound sources in this clip are more clear. It was selected as the introducing test to the participants.

Recording No.2 (40 seconds) was recorded while walking from the coffee corner to the cashier in the canteen of IDE. The canteen connects the main hall, which is very large (Figure 3.2 middle) and more than 10 meters high. The clip was recorded during lunchtime. There was a clear sound source in the clip of the POS (point-of-sale) device from the cashier. The beep sound could be seen as a representative sound of the environment. Other sound sources included clatter of cutleries, people talking while selecting and purchasing food and sound of preparing food from the kitchen. Since the recording was taken during lunchtime, the canteen was noisy and the foreground and background sound sources were mixed and difficult to tell apart. This clip was recorded while moving. It was placed after Recording 1 in order to compare the difference of spatial perception of moving and fixed position recorders.

Figure 3.2
From left: the elevator hall in IDE with an automatic door on the right; canteen of IDE; main entrance of Den Haag city hall

Recording No.3 (38 seconds) was recorded in Den Haag city hall while walking from the main entrance to the reception. The hall is twelve-floor high with metal wall, glass windows and ceiling (Figure 3.2 right). Besides a revolving door from distance, there was no clear foreground sound source in this recording. The other sound source was mainly from a waiting area nearby. This recording was taken in a more spacious environment than the previous ones and the environment was unfamiliar to all participants. It was selected to represent an unfamiliar environment with no easy-recognizable foreground sound source.

Recording No.4 (33 seconds) was taken while walking along the Balustrade (a long table for computers) on the first floor of IDE. In this clip, students of IDE were using computers on the right side (Figure 3.3 left). As in Recording 3, there was no clear foreground sound source either. The background sound sources included clatter of cutleries from distance, computers and people moving. The recording was made while moving. It was selected to represent a familiar environment without easy-recognizable foreground sound source.

Recording No.5 (57 seconds) was recorded in the elevator hall of Den Haag City Hall. There were four elevators, two on the left and two on the right of the recorder. The small room (approximately 8 square meters) is semi-open to the main hall of the city hall (Figure 3.3 middle). Background sound in the environment included walking sound of people and chat from a waiting area. There were clear foreground sound sources, dings of elevators that indicate the arrivals. The sound of elevator is a very common sound source and is assumed to be a very familiar sound to frequent visitors. This clip was selected as an unfamiliar environment with clear sound sources. With

Figure 3.3
From left: Balustrade of IDE; elevator hall of Den Haag city hall; coffee corner of IDE

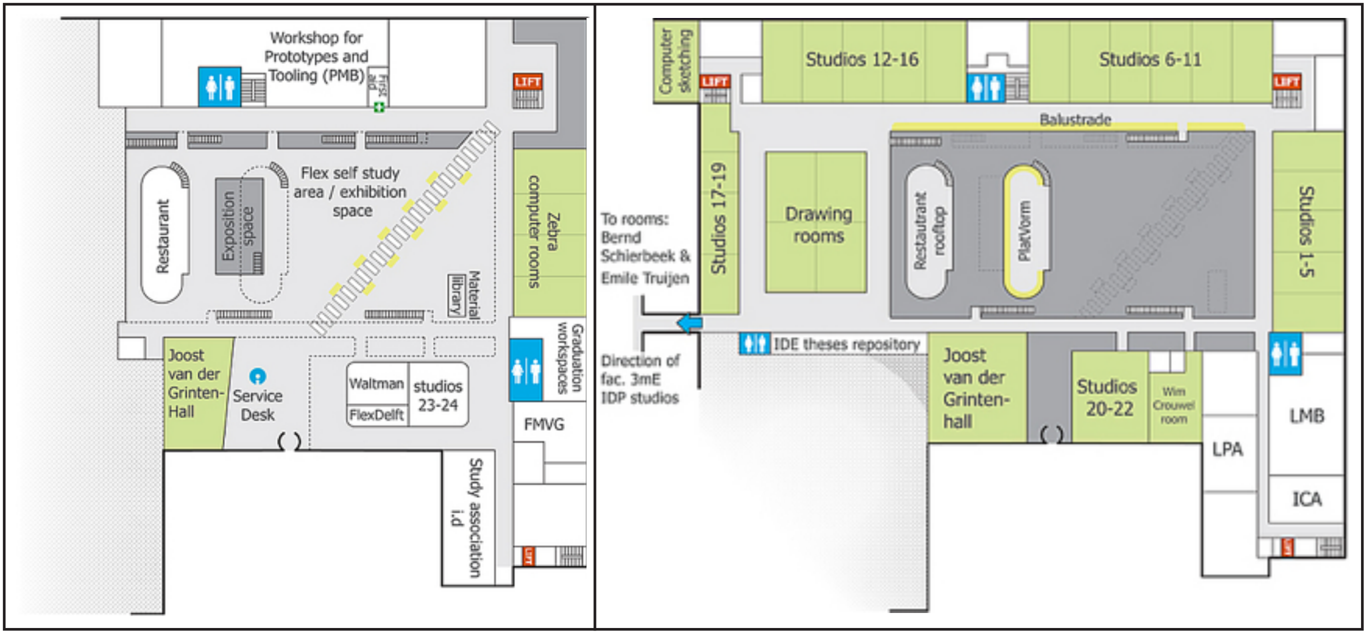


Figure 3.4
The map of IDE provided to participants (left: ground floor; right: first floor)

this clip, the participants would be tested whether they could tell apart the familiar foreground sound from an unfamiliar background environment.

Recording No.6 was taken during lunchtime pointing at the coffee corner of IDE (Figure 3.3 right). The recorder was closely surrounded by people. The talking sound, which is normally seen as the background sound, was quite loud and close in this clip, even clearer than the foreground sound source, a service bell. With this clip, the ability to perceive a space with strong distractions would be tested.

3.3.2 Apparatus

As illustrated in Figure 3.4, two maps (ground floor and first floor) of IDE were provided to the participants so that they could mark the perceived location of the recorder on the maps. The tests were carried out in a quiet studio. The sound clips were played with the application of Sound Studio of an OSX laptop. An open-back headphone of Sennheiser HD-515 was used.

3.3.3 Procedure

Participants performed the test separately. They listened to the recordings in sequence and blindfolded. After finishing each clip, five questions were asked. They answered the questions verbally while the researcher took notes and recordings. They were allowed to listen back again when they felt necessary. The questions are listed below:

- Can you describe the space where this sound clip was recorded?

This question was designed to figure out how participants perceive a space through listening.

- Do you think the clip was recorded in a fixed location or while moving?

In order to tell apart if a sound clip is recorded while moving or in a fixed position correctly, the participant should be able to select fixed ones from the sound sources. This question was designed to demonstrate the ability of the participants to distinguish sound sources.

- Can you name the sound sources that you can identify from the sound clip?

Through this question, the most identifiable sound sources could be found, which could be used in the design phase as triangulation indicator for auditory landmarks described in §3.1.3.

- Is there anything special that you noticed of the environment?

The question was designed to illustrate the attention paid to acoustic feature of the building and other noticeable auditory factors.

- Are you familiar with this space, could you mark it on the map? (A building map of faculty IDE was provided so that they could mark the position.)

The question was expected to test how the participants would link sound sources to reality, based on which the strategy of localization of sound sources could be verified.

3.3.4 Results

All results were documented in forms, which could be found in Appendix III. Comparison between tasks and results are made below:

Firstly, participants had difficulty telling the size of space from recordings. In line with the literature (§ 2.3.3), for the small size and middle size space (smaller than 100m^2), the reported size tended to be large than reality. By contrast, for the large open space (larger than 500m^2), the reported size tended to be smaller. For the height of ceiling, there was no clear mistake pattern.

When asked to tell whether one recording was made in fixed position or while moving, the accuracy rates (the percentage of correct answers out of total answers) of recordings made in fixed position were much higher than the moving recordings (Figure 3.5, the

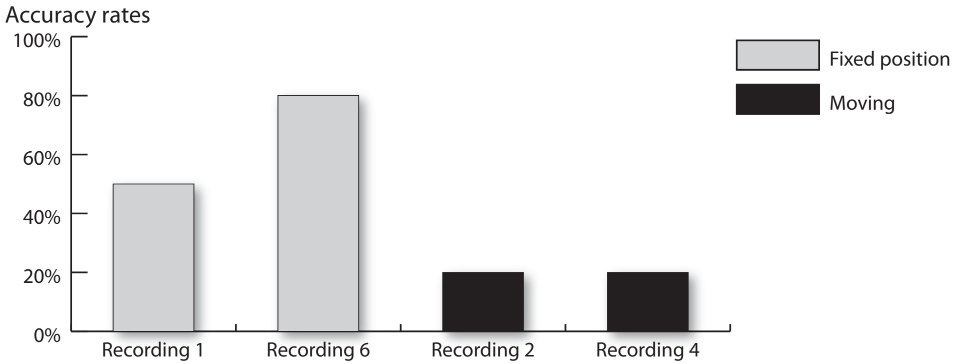


Figure 3.5
Accuracy rate of judging whether the recording was recorded in fixed position or while moving (for the four recordings made in IDE: Recording 1 and 6 were recorded in fixed position; Recording 2 and 4 were made while moving)

left two columns indicate the accuracy rate of recordings in fixed position, the right two while moving), which means participants had tendency to believe a recording was made in fixed position unless clear clues existed indicating the movement of recorder.

Talking about sound sources, the most frequently mentioned sound sources were people walking and talking. Human sound could be a clue when making judgement, but also led to confusion occasionally. For example, people taking from distance could indicate the large size of space, however, one participant assumed it as people talking next door. Consequently, he misjudged the space as a studio. Rather than human sound, permanent sound sources were also frequently mentioned during the tests, such as elevators and doors. The participants regularly used these sounds sources as sound landmarks to recognize a space. Apart from the permanent sound sources, several representing temporary sound sources were also helpful, such as the service bell of coffee corner and cutlery sound of canteen.

The background sound or environment sound was seldom used by the participants to recognize a space. However, they made frequent use of the changes in acoustic features, such as reverberation, to tell the recording was made while moving.

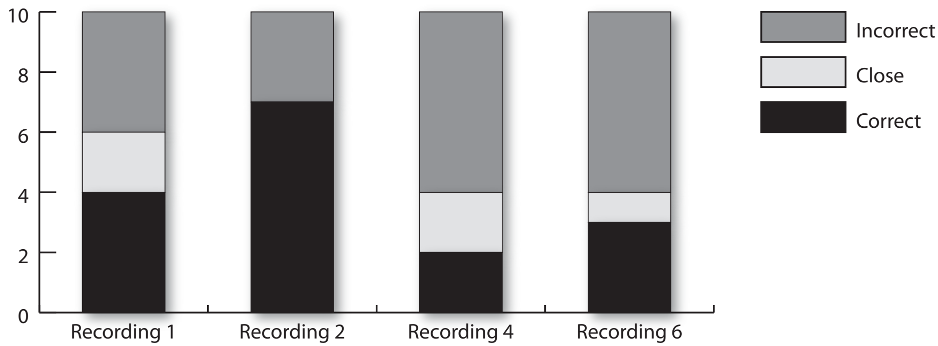


Figure 3.6
The accuracy rates of linking sound sources to reality of the four recordings made in IDE

Participants performed better in recognizing a space by sound when representing sound sources exist. As illustrated in Figure 3.6, the accuracy rate of Recording 2 was apparently higher than others thanks to the existence of the cutlery sound and POS machine sound. The recordings made while moving confused participants if they were not able to figure out the movement of recorder. For example, Recording 4 was recorded while moving from one end of the Balustrade (close to canteen) to the other (close to an automatic door). One participant recognized both the cutlery from canteen and the automatic door. However, he was confused by the presence of both these sound sources and misjudged the location.

3.3.5 Discussion

During the tests, when asked about the perceived space, participants could describe how large it was immediately. However, they often needed to listen back to the sound clip again to tell the height of the space. Also the perception of height was less accurate than of size. The fact indicates that either people pay more attention to the azimuth plane when perceiving a space by sound or the height of space is relatively more difficult to perceive by sound.

The most frequently mentioned element of the environment perception is the material of the floor. Participants noticed the material changes of floor. Some participants could describe physical features of the material, like hard or smooth, and some even could link the sound to a certain material, such as marble. This suggests that the changes of floor materials could be perceived through sound, and therefore could be used as a cue when designing a navigation system with sound.

The perception of the actual size of a space differed by participants. However, when participants referred one space to the previous one (such as “much higher than previous one”, “narrower than previous”), the relative size tends to be more correct and accurate. Lacking the reference of space could be the reason leading to such results. Considering that visual experience plays a significant role in space perceiving, the situation could be even worse for B&VI users, especially the congenital blind, who are more sensitive to spatial sound. The ability of blind and visually impaired people to link a spatial sound to reality should be tested by further study.

Both Sound Recording 3 and Recording 5 were recorded in Den Haag city hall, which was seen as an unfamiliar location for all the participants. In Recording 5 there were clear sound signals of the frequent dings of elevators, while there was no easy recognizable sound signal from Recording 3. By comparing the results of these two tasks, it is noticeable that the perceived locations of Recording 5 are more accurate than of Recording 3. Through the test, there was also more information provided to

Recording 5 than Recording 3 by the participants. Comparing the results of Recording 2 and Recording 4 also drew the similar conclusions. This suggests that sound signals are the key factor to link a sound environment to reality.

In the results of Sound Recording 2 and Recording 4, most participants recognized clatter of cutlery. However, in Recording 2 the cutlery sound was in a close location while it was relatively far from the recorder in Recording 4. When asked about the familiarity, the participants were able to figure out Recording 2 was recorded in canteen and Recording 4 was in a position with a certain distance from the dining area. By comparing the results, the strategy of locating a sound source of the participants was revealed by the researcher, which was also mentioned by one participant. They would first recognize a sound source and then make a virtual circle based on the distance perceived from the sound source. By cross comparing multiple circles of sound sources, they were therefore able to locate the recorder.

According to the strategy mentioned above, when a sound clip is recorded while moving, the virtual circles are no longer accurate enough to locate the recorder due to the change of distance. This deduction was also confirmed by the tests. The accuracy of location judgment of sound clips recorded by fixed recorders is relatively higher than moving recorders.

The recorder was closely surrounded by people in Recording 6. Several participants mentioned during the test that the close distance noise made by crowd distracted and influenced them while perceiving the space. Two participants were even not able to describe the size of the space because of the noise. When designing a navigation tool, such close noise should be avoided.

3.3.6 Conclusion

Compared to the background sound sources or acoustic feature of a space, the foreground sound sources, especially the representing sound source such as service bell, are more helpful to distinguish a space. The background sound sources should be controlled since sometimes they would distract the users or influence the space perception. For the virtual sound display, pure sound source could be tested for the following study.

Although indoor spaces of public buildings vary in height, users focus more on size than height when perceiving a space by mere sound. When representing different spaces by sound, one should be careful when using merely the cue of height differences. Additionally, the change of floor materials could be used as a cue to distinguish spaces. Since participants can notice the change of sound space or relative size more easily than the absolute size of a space, it could be helpful to provide a reference sound space for the scale model.

3.4 Study II: Sound Navigation

The user interviews demonstrated the navigation strategy and spatial perception of B&VI people. The study of sound perception explored how people perceive, understand and describe sound and sound space. Based on the results of these studies, a study of how to connect sound and sound space to the navigation for blind and visually impaired people was necessary.

3.4.1 Goal

Through this research, the question of how to apply the findings of sound space into indoor navigation design would be explored and answered. Different means of sound display would be tested with users, and the most effective way would be selected for further development in the scale model design.

3.4.2 Methods

Since the previous tests on sound perception were finished in the faculty of Industrial Design Engineering (IDE), TU Delft, in order to make better use of the results, this study was also conducted in the same building. Contrary to the previous tests, the participants of this study were unfamiliar with the environment where the recordings were made. Thus, students or staffs from IDE were avoided during the selection of participants.

Figure 3.7
Study II: sound navigation

Participants

Six sighted participants, who had never entered the building of IDE, were selected for this study.

Stimuli

Nine sets of recordings* of three routes (three for each route) were made in IDE. The three routes shared the same starting point – the main entrance of the faculty. The three destinations of routes were (see Figure 4.8): PMB store in the basement, elevator hall on the ground floor and IDP studios on the first floor. For each route, three sets of recordings were made with a dummy head microphone (model: Sennheiser MKE-2002; recorder: Roland Edirol R-09HR). Dummy head recording is a method used to make binaural recordings, that allow a listener wearing headphones to perceive the directionality and the room acoustics of single or multiple sources.

* all mentioned files are included in digital appendix (Chapter 3/ study II/)

For Set 1, the recordings were made while walking continuously from the starting point to destinations. For Set 2, several sound clips were recorded in fixed positions along the route, preferably close to sound landmarks. The distance between recording points were smaller than 20 meters. And when recording, the dummy head was always facing next turning point. For Set 3, recordings were made in fixed positions whenever there was a turning point, door or stair to take. The dummy head faced next turning point while recording. And if there was a turning point, the dummy head turned as well.

3.4.3 Procedure

Participants	1&4	2&5	3&6
Test 1	Route 1 Set 1	Route 1 Set 2	Route 1 Set 3
Test 2	Route 2 Set 2	Route 2 Set 3	Route 2 Set 1
Test 3	Route 3 Set 3	Route 3 Set 1	Route 3 Set 2

Table 3.1 Tasks for participants

The scale model would be placed in a fixed location in the public building. To better simulate the using situation, a map of IDE, which represents the scale model to B&VI users, was shown to the participants at the starting point. The participants were allowed to listen for unlimited times to the recording, which simulated the exploration of the B&VI with the scale model. Afterwards, the participants were asked to follow the route in the recordings and navigate to the destination. The map was taken back after they started navigating. The participants were allowed to take the recordings with them and listen back while navigating. As in Table 3.1, each participant finished three tests of three routes and sets (for example, “Route 1 Set 2”, “Route 2 Set 3”, “Route 3 Set 1”). During the tests, participants were observed and the time of finishing each route was noted. The participants were asked to think out loud during navigating. After finishing three tasks, a short interview session was conducted.

3.4.4 Apparatus

Figure 3.8 shows the map given to the participants (without the routes). All recordings were played by Roland R-09HR with an open-back earphone (Sennheiser OMX-95).

3.4.5 Results and Discussion

In general, the participants only finished four tasks out of eighteen successfully, among which three were finished by one participant. During the process, participants

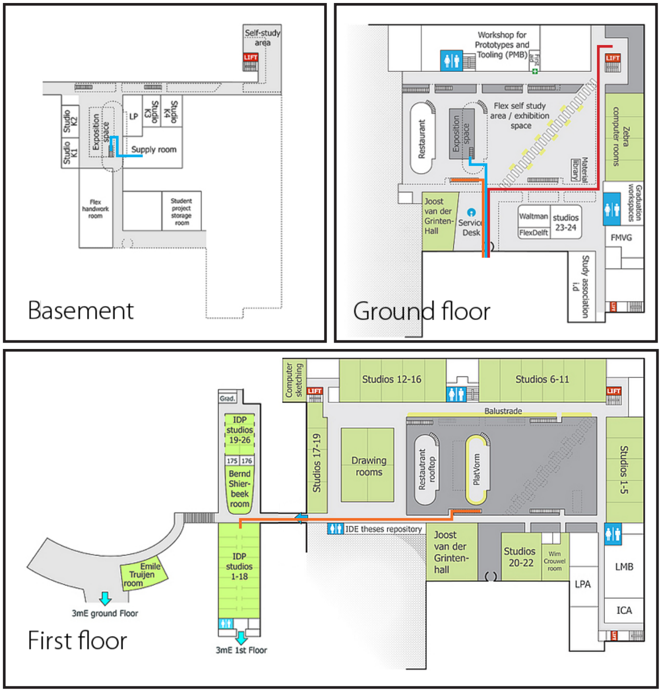


Figure 3.8
The map of IDE with the three planned routes. Red - Route 1; Blue - Route 2; Orange - Route 3

performed better with Set 2, recordings taken in fixed positions close to sound landmarks.

"Can you tell me the moving speed in this recording? There can be a door but where it is?" – one participant asked during navigation with sound set 1.

For sound set 1, the continuous recording, the most frequent encountered difficulty was to match the moving speed in the recording to reality. For example, the talking sound in the recording was captured in a small room of basement, but when listening, the participant was standing on ground floor. Therefore, there was no chance to find the proper room in reality. The only participant successfully finished this task was by comparing the sound sources in this recording to reality for several times. She manually repeated the sound clip by sound sources. For example, when she recognized a printer from the recording, she will pause the recorder and search for the nearest printer in the building.

"I didn't hear anything in this one… did you give me the right recordings?" – one participant mentioned during navigation with sound set 2

Two of the six participants successfully found the destination of set 2. They performed better because the recordings were mostly made closely to sound landmarks. By repeatedly listening to one clip, they were able to recognize several sound clips and link them to reality. However, the dilemma of this set was the choice of distance between two recording positions: in reality, sometimes the distance between two recognizable sound sources is much larger than 20 meters. If the distance is too long, the participants will get lost between two recordings, and if an extra recording is made between two sound sources, the recording itself will be confusing and misleading.

> *"I know you turned left, because I can hear the crowded noise was behind me, and now it is on the left. But I don't have any clue where to make the turning. And one more thing, you didn't mark the height of ceiling and floor material on the map." – during the interview, one participant mentioned*

Even worse than set 2, the distance between two clips was not given for sound set 3, which could be up to tens of meters. The performance with set 3 was therefore the worst among the three sets of sounds. Even they can figure out the direction of turning, it was rather difficult for the participants to find the proper turning points with the vague clue from the clips.

3.4.6 Conclusion

The navigation based on merely recordings is difficult to achieve. Sound navigation is influenced and limited by the quality of recording and the quality of player as well. Therefore, in normal cases, recordings can only be used as additional cue, rather than a standalone tool in navigation. For example, people are able to perceive the change in ceiling height, but much more effect should be made if they want to figure out the exact height. People are able to tell the direction of sound sources and the relative distance difference between them (eg. A is farther than B), but it is rather difficult for them to tell the actual distance sound source (eg. A is x meter away). Consequently, when applying sound to scale model, in order to make better use of relative perception of sound sources, it would help to provide reference to sound sources, or use multiple sound sources together as reference to each other.

3.5 Design Requirements Part 2

In addition to the design requirements mentioned §2.5, design requirements drawn from user studies are listed below.

Principally, the scale model should be safe, which means anything that might lead to potential hazards to blind users should be avoided. Independence should be the second key value after safety. The blind users should be able to explore the design freely and spontaneously without any assistance from others. Consequently, the design should also be easy to get started and the interaction should be intuitive even for B&VI users without any experience with scale model. Even if the scale model will probably be fixed near the entrance of building, inviting sound could be helpful to attract users. If linguistic instructions or descriptions had to be provided, non-Braille user should be taken into consideration. The scale model should inform users through fixed auditory landmarks instead of moving ones. The users should be able to explore the model to get sonic feedback regarding these landmarks.

For the functional aspect, the main functional areas of the building should be clearly illustrated by the scale model. The virtual display should be easy to link to reality. The current location of users should be noticeably marked on the scale model. Rather than explaining merely the main destinations of visitors, experiential display of the interior environment should also be taken into account.

Additionally, the scale model should fit into the interior of the building and the sound play should be controlled to an acceptable level, which will not influence the environment of the building. It should also be appealing and provide useful information to sighted users.

4. Research Summary

In this chapter, the design goal identified from problem explorations (Chapter 2) will be further explained based on users' needs, since this is a user-centered design. The design requirements generated from problem explorations and user studies (Chapter 2) will be summarized, interpreted, and then categorized into general requirements, design requirements of model base and design requirements of sound.

4.1 Design Goal

The design goal identified from the problem explorations is to design an interactive interior scale model of a public building with sound navigation system for blind and visually impaired users and to fulfill their curiosity to explore the building. The design goal will be expatiated below:

Firstly, to design **an interactive interior scale model of a public building** indicates that the final design concept should be an interior rather than exterior scale model of a certain public building with extra focus on human-product interaction and related interactive techniques.

Secondly, although user with normal vision should also be considered all through the design process, **blind and visually impaired users** remain the main target group to focus. Since a user-centered design process will be employed, the concept design phase will be largely focused on satisfying the needs and desires of B&VI users.

"**With sound**" implies sound will be utilized as the important but not the only method to communicate and realize the function. Through iterative research by design cycles, auditory representation of interior of public buildings will be developed and tested with B&VI users.

The functional objective of the concept is **to provide B&VI users with navigation information** while the experiential objective is **to help them to explore the buildings**. From the functional perspective, the main criteria assessing the concept should be the effectiveness and efficiency of navigation. From the experiential perspective, the concept is supposed to compensate the loss of vision in establishing spatial perception of public buildings for B&VI users.

4.2 Design Requirements

General requirements

The scale model should be safe, so anything that might lead to potential hazards to blind users should be avoided. It should allow B&VI users to explore freely and spontaneously without assistance from others. The design should also be easy for B&VI users to get started and the interaction should be intuitive even for blind users without any experience with scale model.

The scale model should provide a holistic view of the building, orientation inside the building, a starting point, main destinations and reference points (landmarks). The detailed level should be controlled considering the functionality of scale model and memory capacity of B&VI users as well, neither to provide inadequate information about the route nor to overload the short-term memory of users. Any misleading or confusing information should be eliminated. If any scale or distance is employed in the scale model, extra attention should be paid considering the different navigation strategies of B&VI users.

Model base design requirements

The base of the scale model should be able to provide a floor plan, the spatial layout, key subsections, and distinguishable separation between public and security areas of the public building and the current position of users. If linguistic instructions or descriptions are provided, non-Braille users should be taken into consideration. Since the lighting condition may vary, the concept should be adjustable according to ambient light. The concept should also be attractive to and provide useful information for sighted users.

Sound design requirements

All landmarks should be represented by sound. The selected sound should be universal and recognizable for B&VI users and easy to related to reality. The information provided by sound should not conflict with tactile information. Sound play should be controlled to an acceptable level, which will not influence the overall environment of the building.

5. Concept Development

Based on the design criteria, several concepts were developed to test. Due to the low accessibility of target users (the B&VI), the initial concepts were tested with blindfolded users.

The concepts developed in this phase were concentrated mostly on form giving, user experience and functionality. Other aspects that should be taken into consideration, such as starting point, material, lighting, are shelved for two main reasons: firstly, to which building the final scale model would be applied had not been decided during the design phase, so the final environment to place the scale model is unknown. Secondly, the frame of the final concept is not fixed and requirements of such features may change according to the frame. Thus, ruling out such features would enable the researcher to concentrate more on the development of the more contributing aspects (form, experience, functionality).

One additional sound test with B&VI participants to select representing sounds would also be included. Guidelines for final concept development will be concluded in this chapter.

5.1 Concept 1: Virtual Reality of Sound

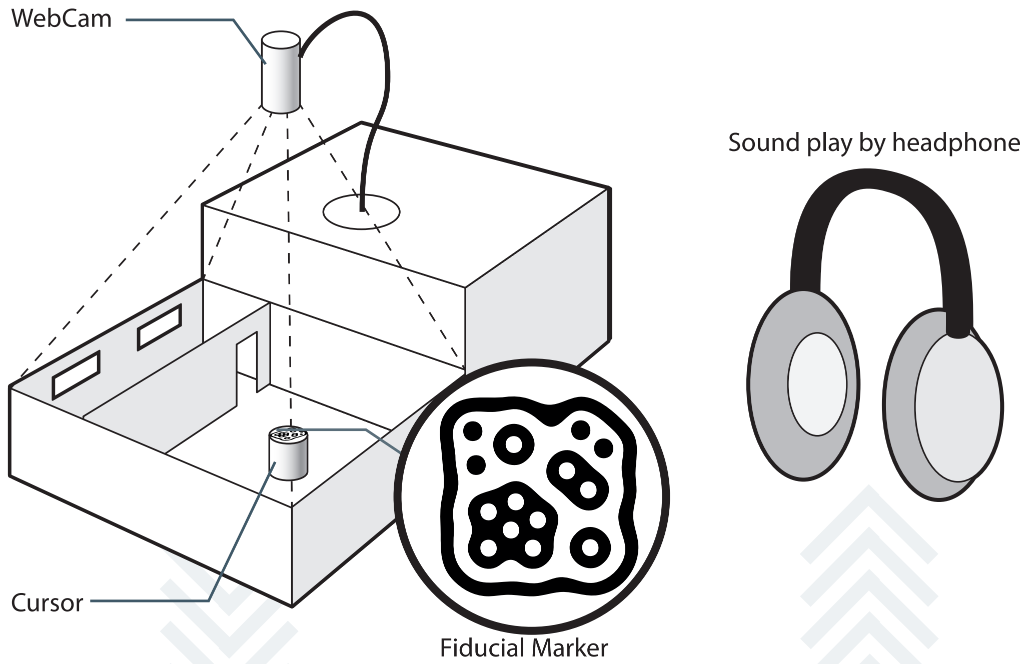
This concept provides an interactive scale model with built-in sound. A tangible cursor (see Figure 5.1) was developed to indicate user in the scale model. When the cursor is moved by users, sound will be played according to the current position, speed and angle. In order to explore representation of multiple building floors, a multiple drawer structure of two layers was applied to the scale model. When user pulls out a drawer, a webcam will detect the identification number of drawer (supported by reacTIVision, which will be explained in §5.1.3) and activate the related patch in the program. Positions (X/Y coordinates in program) of sound sources are preset in the program: a sound of a male walking upstairs for the staircase [31], a dishwasher sound [32] for the kitchen and a typewriter [33] sound for the office of the scale model. When the cursor is added, the program will detect its current position and direction. The relative distance and angle between cursor and sound sources will be calculated in real-time by the program. A binaural toolkit (will be explained in §5.1.3) will process sound based on relative distance and angle. The processed sound will be played spontaneously by headphone (stereo sound). With the sound, users are able to explore the scale model by moving the cursor. User can switch to the other layer of the scale model by pulling out the other drawer any time they want. The removal of cursor will stop sound play.

5.1.1 Representations of Reality

Dryer [34] stated that people apprehend reality only through representations of reality, such as texts, discourses, images and sounds: there is no such thing as direst or unmediated access to reality. For B&VI people, due to the lack of imagery experience, sound is vital to the representations of reality in scale model. The aim of this concept is to create a virtual sound environment as close to reality as possible inside the scale model.

Figure 5.1
Tangible cursor used in Concept 1

In this concept, the representation of sound is divided into sub-tasks: time, distance, direction (azimuth and elevation), speed and reverberation. The sub-tasks are solved separately using programming and prototyping. Firstly, a real-time video and sound processing engine is applied, which means the sound rendered and played by the system represents what is happening at the moment when users interact with the model. Secondly, for distance and direction representations of the sound sources, positions (X/Y coordinates) of the cursor and movable sound sources are captured. The relative distance and direction between the cursor and movable/fixed sound sources are calculated and then scaled up from the scale model to reality. In order to achieve the effects of movement, previous coordinates and angles are stored temporarily in the program. Comparison of the current and stored coordinates and angles will provide speed and direction of the movement. The reverberation of sound



Hardware

Software

reacTIVision host
Process video signal, track fiducial markers in the video

HRTF patch in Max/MSP
Sound processing

reacTIVision client in Max/MSP
Position and angle processing of the cursor

Max/MSP
Calculate relative distance and angle between cursor and sound sources

Figure 5.2
Main components of Concept 1 and the working process

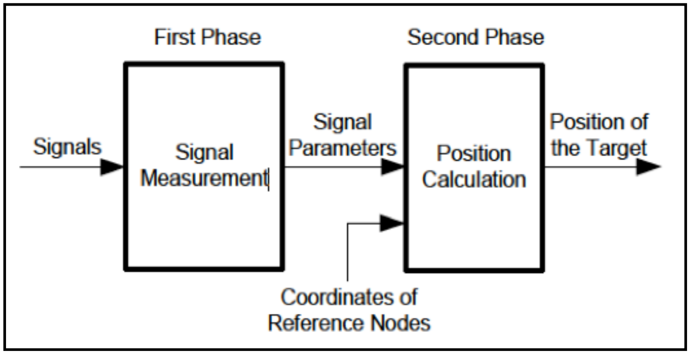


Figure 5.3
Two phases of indoor localization

(will be further explained in §5.2.2) is achieved by using real sound recordings.

5.1.2 Indoor Localization

For navigation, indoor localization is essential. An indoor localization system, as Dempsey [35] defined, is a system that can determine the position of something or someone in a physical space such as in a hospital, a gymnasium, or a school, continuously and in real-time. As illustrated in Figure 5.2, the process of localization involves two main phases [36]: signal measurement and position calculation. In the first phase, position-related signals, including arrival time, signal strength and direction, are collected by the receivers. In the second phase, the physical position of the target object will be calculated based on the signal parameters extracted from the first phase. The most common used technique for calculation is based on ranging, whereby distance and angle estimations are obtained [37]. Geometric approaches will also be applied to calculate the position of the target object as the intersection of position lines obtained from the position-related parameters at reference nodes.

5.1.3 Components

As illustrated in Figure 5.2, the system consists of a cursor with a fiducial marker, a scale model of a fictitious floor plan, a webcam with stand, a computer with Max/MSP and a headphone. The cursor in the scale model indicates user in the public building. The scale model provides spatial and structural information of the public building. The webcam captures real-time video for Max/MSP and headphone is used for sound play.

Max/MSP

Max is a visual programming language for music and multimedia developed by Cycling '74 [38]. It has been widely used by composers, performers, software designers, researchers, and artists for creating innovative recordings, performances, and installations.

MSP (short for Max Signal Processing) is an audio extension for Max, which allows users to manipulate digital audio signals in real-time and create their own synthesizers and effects processors.

In this concept Max/MSP is used to translate information from the physical cursor and the scale model to sound output.

Disclaimer: All the origin sound recordings used in these concepts are legally downloaded from http://www.freesound.org/. All the sounds are licensed under the Creative Commons 0 License (with No Copyright).

ReacTIVision

ReacTIVision is an open source computer vision framework developed by Martin Kaltenbrunner and Ross Bencina for object tracking and multi-touch finger tracking [39]. ReacTIVision makes use of a set of black-and-white visual patterns: fiducial markers (as in Figure 5.3). There are in total 216 fiducial markers with its own serial number, which can be used by pasting on surface of physical objects. A standalone application detects X/Y coordinates, angle, speed, and acceleration of objects from a camera input, and transmits event-messages to an UDP port. A dedicated Max/MSP object will listen to this UDP port, and allows for further processing of tag information.

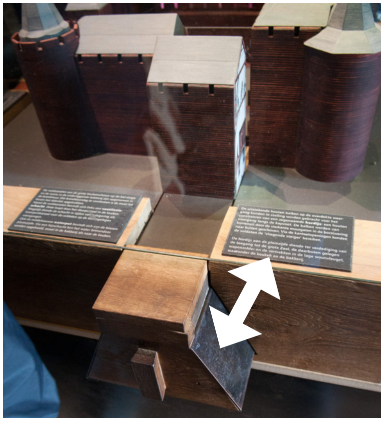


Figure 5.4
The scale model of castle Muiderslot. It is placed in one of the towers in the castle. The white arrow indicates the direction, following which the users could slide open the model and see the inside.

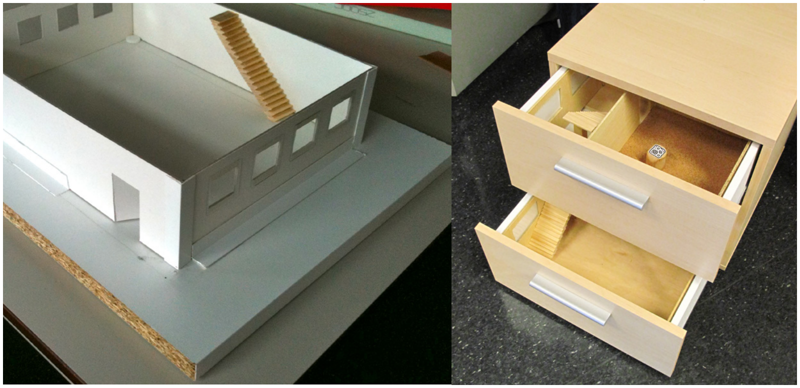
Binaural tools for Max/MSP

A binaural toolkit developed by UC Davis College of Engineering is used in this concept. It is a patch using Head Related Transfer Function (HRTF) measurements in Max/MSP [40]. With the help of this patch, any sound source can be altered to sound as if it is coming from a certain direction external to the body rather than flat on the side of one's head. Both azimuth and elevation angle can be adjusted in this patch. Therefore, it is the ideal toolkit for multi-layer building scale model. The binaural sound effect only works with headphone.

Multi-layer building structure

Several card board prototypes were made to explore representations of building structures (Figure 5.5 left). Inspired by the scale model of castle Muiderslot, which could be slid open (see Figure 5.4), a drawer structure was finally selected to the scale model to represent the multi-layer building. The testable prototype compromised a two-drawer cabinet (see Figure 5.5 right) in which a scale model of a two-layer building was installed. Each layer of the scale model was marked with a fiducial marker that the camera could recognize which layer was in use and switch to the programs of the layer accordingly*.

Figure 5.5
Left: one of the card board prototypes to explore structures. Right: the cabinet structure of the scale model and the two layers



5.1.4 Building

The webcam (Apple iSight), linked to a computer with firewire, was mounted to a lamp-stand on top of the cabinet. Since a two-layer structure was selected, when switching from one layer to the other, the webcam took time to autofocus. After experiments, the best compromising solution was to provide sufficient lighting and

* all mentioned files are included in digital appendix (Chapter 5/ concept 1/)

a clear enough object (the fiducial marker) in the field as focus. In order to indicate direction of the blind users, the cursor was formed into a water drop shape (Figure 5.1), and the triangular arrow indicated the face direction. However, in order to avoid misunderstanding, there was only one cursor provided for two layers, user should also bring the cursor to new layer.

Figure 5.6
One participant in the test. He was wearing a headphone and blindfolded. The webcam above was capturing the image of the model and cursor. The model was switched to the ground floor.

5.1.5 User Tests and Evaluation

Setup

Three master students from Department of Industrial Design Engineering, TU Delft, participated in the user test. All participants were blindfolded before exposed to the scale model (Figure 5.6). They were told the object is a scale model of an unfamiliar public building and asked to explore the prototype without further instruction provided. After exploration, they described the building based on their perception of the scale model and gave comments on the experience. The concept would be explained and four questions were asked about sound and scale model:

1. During the exploration, is there any confusion about the sound or the model?
2. Can you relate the sound to reality or not, why?
3. How do you evaluate the cursor? Were you able to link the cursor to yourself or not?
4. Is the structure of scale model clear? How to improve?

Results and discussion

Figure 5.7
During the test, fiducial marker on the cursor was covered by hand.

In general, all participants were able to retell the main structure of the building and each functional area correctly. However, several issues were encountered as below.

> *"I don't know what I did. But the sound stopped. I am totally helpless when blindfolded." - a participant said when the program stopped because she blocked the marker by her hand*

During the process, one apparent drawback of reacTIVision was noticed. Since the fiducial marker represents the cursor in reacTIVision, it should be pasted on top of the cursor. However, when provided with the cursor, all three participants spontaneously hold it on top. As a consequence, their hands blocked the fiducial marker, which resulted in the detection of cursor removal in the program. Even after being verbally instructed about the working principle, they kept on covering or partially covering the fiducial marker unconsciously (see Figure 5.8).

Although extra light was provided, when participants put hands in the webcam field, the webcam started auto-focusing. The auto-focusing process took much time and would lead to unstable signal output. The same situation occurred when switching from one drawer to the other. For further development, it is better to select either a manual focus webcam or a much faster auto-focusing one.

The recordings failed to provide reverberation effects of the sound sources. After applying the HRTF direction, reverberation of the recordings was no longer recognizable. When asked about the reverberation, participants mentioned that they got most information of the room size from the scale model. The reverberation led to confusion. The wet recording (reality sound with effects) was unable to recreate the reverberation corresponding to reality.

> *"Maybe because of your headphone, or maybe because of the program, when I turned around, there is also a staircase in front of me. But I found only one from the scale model" - a participant confused*

Due to the limitation of the binaural toolkit, the azimuth direction is constricted to -80 to +80 degree. A maximum value of 80 degree was set in the program, as a result, any relative angle larger than 80 degree was treated as 80. During the process, such constrict led to frequent confusion.

> *"I know you want me to use this thing (cursor) as a person in the building. But until you told me, I didn't know I can rotate it" - one participant mentioned when asked about the cursor*

All participants understood that sound were changing according to the position of the cursor, but the direction of cursor was vague that one participant failed to notice the rotating function of the cursor.

> *"I don't think blind people can figure out the right and left channels..."*

Considering the target user group of blind people, blindfolded participants complained they had difficulty getting started. Without help from researcher, two blindfolded participants mixed the left and right channels of the headphone, which led to confusion. The need of B&VI people should be taken into consideration when choosing components.

Figure 5.8
The scale model is made of wood, including the stair case in the middle.

Although the material of the scale model (see Figure 5.7) is not part of this test, one participant argued that the material of stairs in the scale model did not match the sound. Because he felt the staircase as soft and wooden but the footstep sound for the staircase was much harder than he expected. He suggested to use either real material or totally unrelated material such as plastic or paper.

The two-layer structure was positively evaluated based on their feedback. However, the participants mentioned, for sighted people, the cabinet structure might be not attracting because people would not think it as a scale model. It could be improved by adding lights and transparent material to it.

One participant suggested that there should not be blind spot (an area on the model with no sound or tactile information) in the scale model, otherwise users, especially B&VI users, could probably get lost when reaching the blind spot. Because of the unconscious covering of the fiducial marker, the program reset several times during the tests. One participant suggested adding a consistent sound as an indication that the program was working properly.

5.2 Concept 2: Sound Illusion

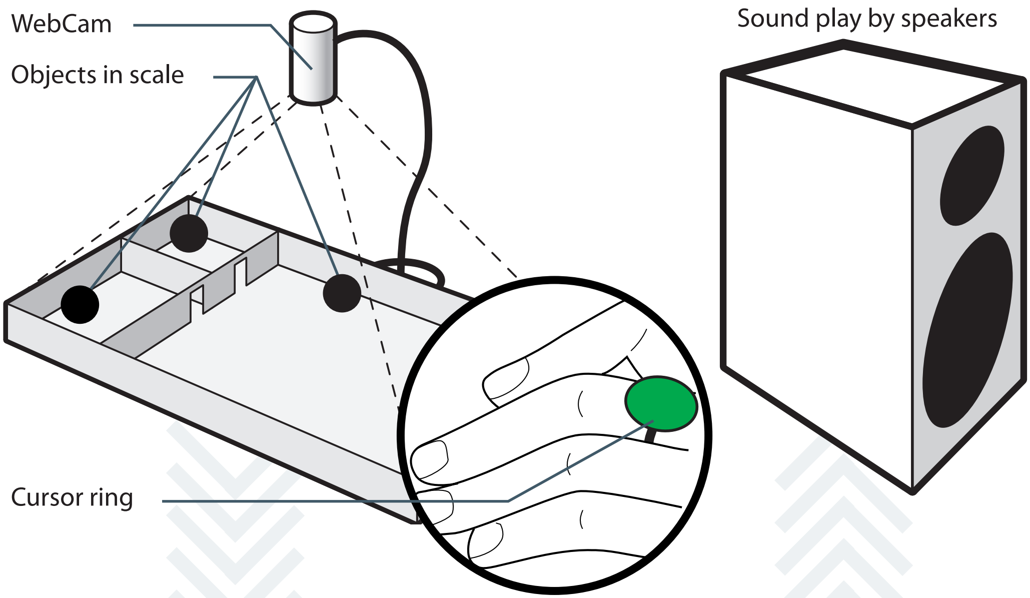
Based on evaluating the previous concept, a new concept was developed. Instead of focusing merely on the representation of reality, the idea of creating a sound illusion had been brought in. The illusionary representation allows more freedom to explore sound and form. Moreover, the issues encountered in Concept 1 could also be solved or improved in this concept.

A new one-layer scale model (Figure 5.10) was built with landmarks of public buildings, which were drawn from the user study (§3.1), such as toilet, reception and staircase. Sound of this concept is also processed based on video capturing of the position of a cursor, which is manipulated by users. However, instead of using a tangible cursor, users are able to use their hands as a cursor by this concept, which frees their hands to touch the scale model. Based on this feature, tangible objects are developed, which means users are able to not only hear the sound sources, but touch them as well. Moreover, the sound player was changed from headphone to speakers. Therefore, as stated in the design requirements (§4.2), this concept has the potential to be used by multiple users.

5.2.1 Auditory Illusion

Auditory illusion does not mean to create sound out of nothing. It is based on reality but not restrained to reality. It is the redesign of sound with the purpose of simplification and symbolization. In this concept, to create the illusion, sound benefiting the navigation or experience of users will be boosted, while confusing or distracting sound will be removed or dimmed. Conforming to the findings of literature study (§ 2.3.2), the higher signal-to-noise ratio of illusionary sound will also help the B&VI users to localize sound sources.

In this concept, instead of using the wet sound recording (sound that has been



Hardware

Software

Color detection patch in Max/MSP
Process video input, detect preset color of cursor ring, and output the position of cursor ring

Max/MSP
Calculate the speed of cursor ring, and its relative distance to preset sound sources; sound process

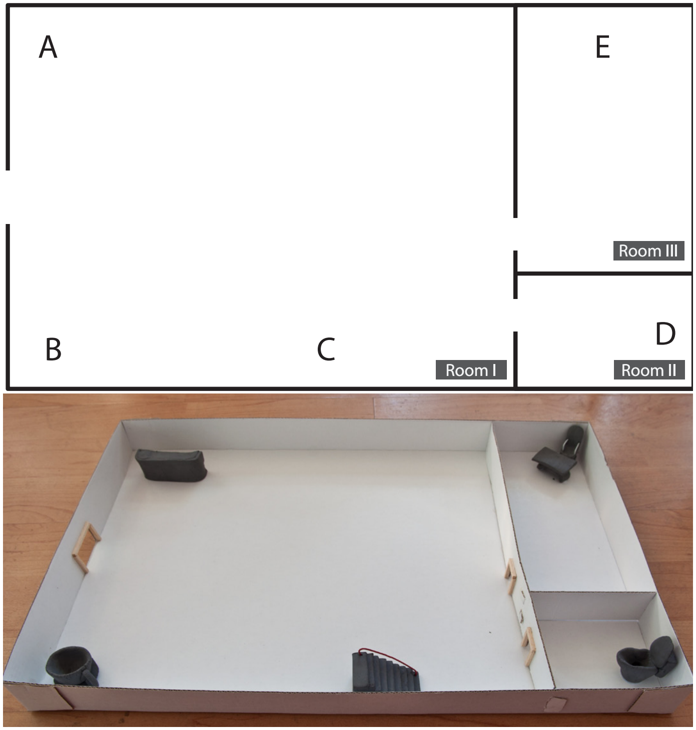
Figure 5.9
Main components of Concept 2 and the working process

processed through an effect or effects), dry sound (a sound signal without any effects) clips, which could be synthesized or recorded in sound booth, are employed. The direction of sound is temporarily removed from this concept so that users could better concentrate on the distance, which also solves the issues raised by the limitation of azimuth angle simulation of the HRTF extension. Consequently, speakers could be introduced to replace headphones for the sound reproduction, which prevents the potential confusion of left and right channels for B&VI users.

5.2.2 Components

This concept consists of a cursor ring in a solid color, a scale model of a fictitious floor plan with tangible landmarks, a webcam with stand, a computer with Max/MSP and speakers (see Figure 5.9). The cursor ring and the scale model represent user and the indoor environment of a public building. Exploration with the cursor ring in the scale model indicates the navigation process of users in a public building. The scale model is not based on a real building but a combination of features from multiple buildings. The programming language is also Max/MSP in this concept.

Figure 5.10
Layout of the scale model (three rooms and five sound sources); picture of the scale model



To free the fingertip

The previous concept made use of a physical cursor. As a result, users had to keep their hands on the cursor during exploration, which restricted the freedom of touching and feeling the scale model. Since more tactile information would be employed for further development, it would be beneficial to free the fingertip of user. Consequently, the cursor was reformed into a cursor ring. In form of a ring, the cursor could be worn on user's index finger, without sacrificing tactile functions of the finger.

Color detection and color tracking

The cursor ring is no longer large enough for a fiducial marker of reacTIVision. Since the number of objects to be tracked by camera will not be large and the directional tracking was removed, the color detection toolkit for Max/MSP can be used. With a predefined threshold of the RGBA (Red, Green, Blue, and Alpha) value of a color, the patch can track the position of a certain color in the field. With different RGBA settings, the patch can track multiple objects in different colors at the same time. Thus, the color detection patch also has the potential to be used by multiple users as reacTIVision. A firewire webcam required by reacTIVision, is no longer a necessary. Therefore, the issue of auto-focusing, which was encountered in the previous concept, was solved by introducing a USB webcam with manual focus.

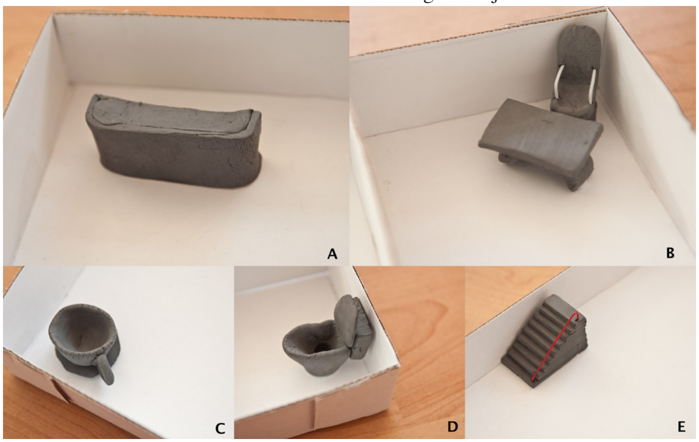
Sound setup

As shown in Figure 5.10, the model is divided into three rooms of different sizes. Room I represents a big hall; Room II is a small size toilet; Room III is a middle size office. Five sound sources were added (marked as A-E in Figure 5.10). A represents the reception of the building; B represents the coffee corner; C is an open staircase; D is a toilet; E is an office desk.

In line with the findings of literature study (§ 2.3.2), dry sounds* with high siginal-to-noise ratio were used to represent the sound sources. In Room I, a service bell sound [41] was chosen for the reception; cutlery sound [42] was selected for the coffee corner; and a male footstep sound [43] walking upstairs was chosen for the staircase. In Room II, four sounds [44] were selected and will be played in a random order loop: a flushing sound, another flushing sound, a sound of running water and a sound of washing hands. Three sounds [45] in a random order loop were applied to Room III: a printer sound, a telephone sound, and a keyboard sound. Since sound source A, B, and C are in the same room (Room I), the sound clips used for these three sound sources were edited in order to create a more harmonious sound environment. The high frequency (above 2000Hz) of A, the low frequency (below 500Hz) of C and the middle frequency (500-2000Hz) of B were boosted and other frequency areas were dimmed.

As to indicate of the presence of cursor, a signal sound of male footstep is added. The delay between each step changes according to the speed of cursor movement. A random delay was added to each step to make the footstep sounds real. And three different sets of step sounds [46] are played randomly for the same purpose.

Figure 5.11
Tangible objects in the scale model



Tangible sound sources

Instead of merely positioning sounds in the scale model, tangible objects were added. Figure 5.11 shows the five tangible objects added corresponding to the five sound sources: A – reception, B – coffee corner, C – staircase, D – toilet, and E – Office. The tangible objects would help users to formulate the connection between sounds and sound sources. Combining sound with tangible objects would prevent misunderstanding in perceiving sound.

* all mentioned files are included in digital appendix (Chapter 5/ concept 2/)

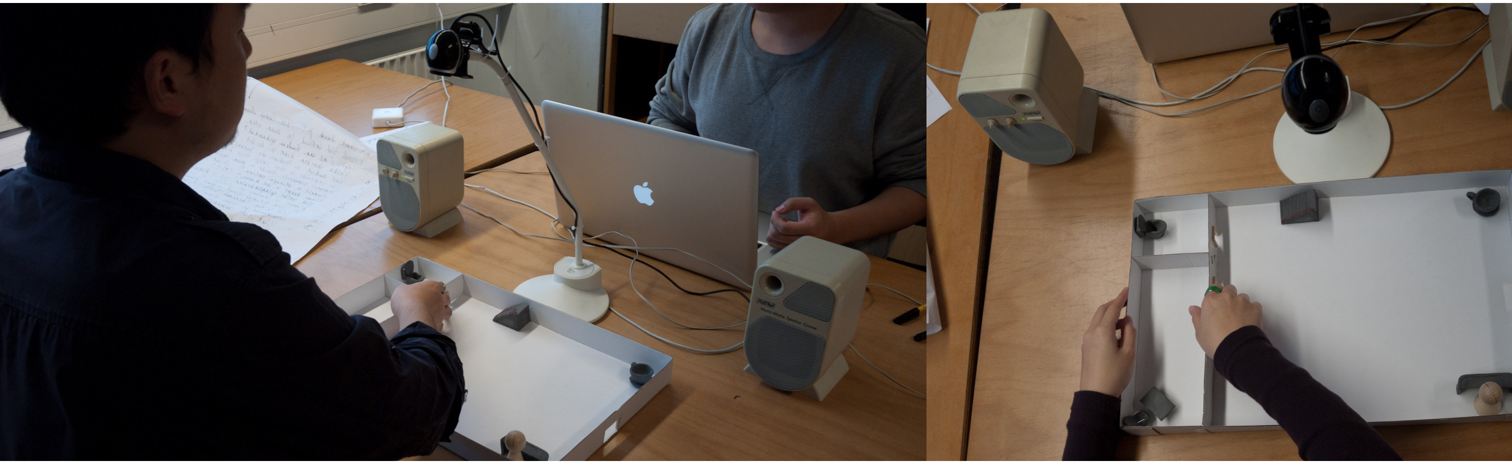


Figure 5.12
Setup of the user test and the test process: from left: the setup of scale model, speakers, webcam and computer; one participant exploring the scale model with webcam above; one participant was drawing map of the scale model after test (the scale model on top is covered by paper); one finished map

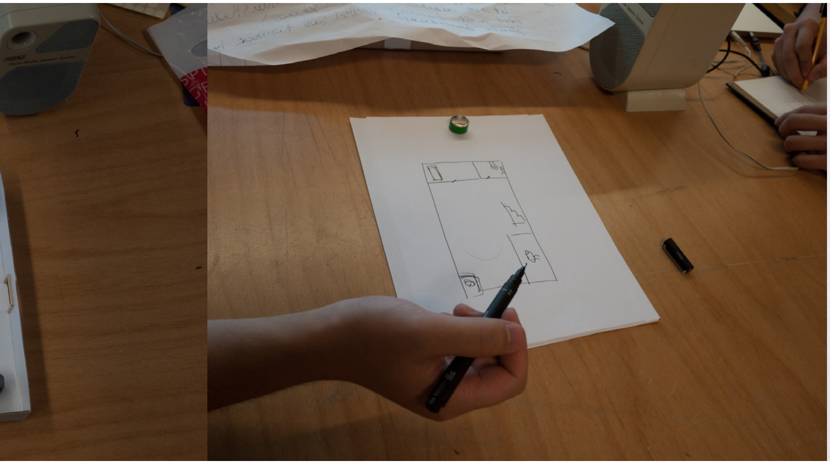
Reverberation

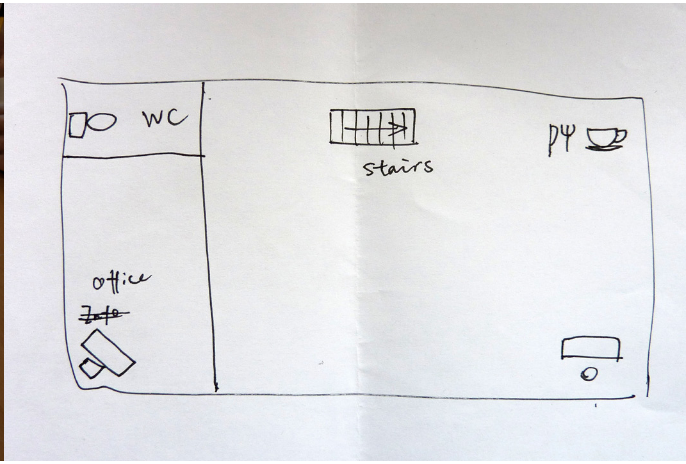
Reverberation is the persistence of sound in a particular space after the original sound is removed [47]. A reverberation, or reverb, is created when a sound is produced in an enclosed space causing a large number of echoes to build up and then slowly decay as the sound is absorbed by the walls and air [48]. In this concept, a reverberation toolkit developed by Randy Jones [49] was applied to the program. In this patch, four diameters (size of the room, decay time, high frequency damping, and diffusion) could be adjusted for the sound output. Any dry sound input could be rendered real-time into wet sound with reverberation effects. In this concept, three different settings are defined for the three rooms. Therefore, the sound reverberation of each room could reflect the acoustic characteristics of the room, which better demonstrates the acoustic environment.

5.2.3 Building and Use Process

Setup of the new webcam remains the same as the previous one in Concept 1. Several cursor rings were made to fit different finger size. In order to prevent interference from the color of human skin and clothes, the distinctive green color was chosen for the cursor ring. Speakers are connected for sound reproduction.

In the captured video, the scale model will be processed as a 320*240 matrix. The three rooms are preset as three different areas in X/Y in the matrix. Positions of sound sources are also predefined in the program. When the cursor ring is added, the program will detect its position and switch on the sound sources in the corresponding room spontaneously. When the cursor ring moves from one room to another, the





program will turn off the sound sources of the previous room and switch on the new ones accordingly. When switched on, the sound sources have a minimum volume of 10%. And when being approached within the preset distance threshold of 50 units, volume of the sound source will increase linearly with the decline of distance and other sound sources in the same room will be muted temporarily. A update log of program is included in Appendix IV.

5.2.4 User Tests and Evaluation

Setup

Seven master students from TU Delft participated in the test (including one pilot test). All the participants were blindfolded and given unlimited time to explore the scale model (see figure 5.12). The only instruction provided before exploration was that they were going to explore a scale model of a public building and they had to wear the cursor ring to get started. After exploration, the researcher covered the scale model and they were asked to draw a map (Appendix VII) of the scale model with all information collected from exploration. Thereafter, the concept was explained to participants with the scale model uncovered. Based on the experience, three questions were asked about the sound and the scale model:

1. During the exploration, was there any confusion about sound or the model?
2. Can you link the sound to reality or not, why?
3. Which objects did you recognize by sound and which by tangible objects?

Results and discussion

> *"I heard a microwave, but I don't know why there is a microwave in the hall. It seems weird but there are a canteen and a cafe in one room..." - one participant said after she misunderstood the reception as canteen*

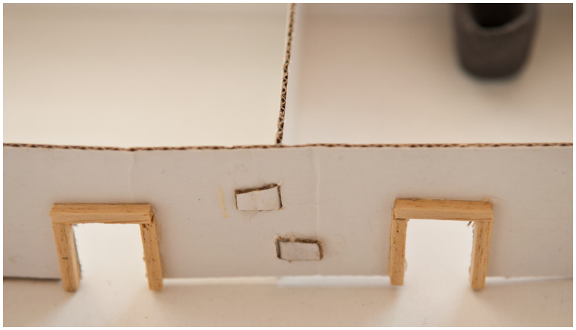
In general, all participants did well in making the map of the model. Although there were more sound sources in this concept, participants finished exploring significantly faster than Concept 1. The service bell sound of reception is confusing that one participant perceived it as sound of microwave oven, and another participant considered it as sound of elevator bell. The sound clip of reception should be improved or changed to a more representative sound. The sound of staircase is too close to the footstep sound that several participants got confused. For further development, the staircase sound should be more distinguishing. Doors (see Figure 5.13) were often overlooked during the tests, it is necessary to add sound to doors if the doors are used as landmarks.

During the process, participants kept on changing gestures of hands to touch the tangible objects and the scale model. Meanwhile, the direction of cursor ring, kept on changing as their hands, which were not always in the designed gesture. The changing direction resulted in unstable cursor detection. Moreover, the glossy material with high light reflection also influenced the detection. For the following concepts, it is better to select matte and soft material of larger size, such as a finger glove.

Although the issue of how to get started was shelved for this concept, participants questioned it. They believed that B&VI users would have difficulty finding the ring and manipulating it correctly. Therefore, for future development, a more intuitive way to approach the cursor should be developed.

Comparing to Concept 1, the added tangible objects were very effective to link the sound to reality. The positions of the sound sources on the map were more accurate than the reported positions in Concept 1.

Figure 5.13
The doors were often overlooked during the user test.



> *"I heard someone was typing and then when I touched something the telephone rang. I've no clue what I did..." - one participant said while exploring the office*

The random order of sound clips led to confusion. For the office, four different sound clips were applied, and the participants kept on trying to figure out what led to the change of sound. For further development, it would be beneficial to split the sound into two sets: signal sound and background sound. The background sound, which will be played in low volume whenever cursor enters the room, indicates the presence of sound sources of the room. And the signal sound, which signifies the sound source, will be played when the objects are being approached.

5.3 Scale Model of Den Haag City Hall

The previous concepts were all based on virtual buildings and tested with blindfolded users. In order to apply the findings to target users – B&VI people, a new concept based on a real building should be tested with target users. Den Haag city hall was selected as the ideal building for further exploration, during the field research of four public buildings. Therefore, a new model was developed based on Den Haag city hall. The scale model was designed to test whether one would be able to find a target location based on information provided by the auditory scale model.

5.3.1 Sound as Legend

Since this project is part of the longer Geluid in Zicht project, the possibility of applying this sound system into other existing scale models was tested. Two scale models developed by Muyun Xiao from a parallel project, focusing on tactile information were selected. The scale models include one 3D model (Figure 5.14 top) and one 2.5D (Figure 5.14 bottom, 2D map with embossed information) tactile map. Since the tactile information would also be tested, in order to rule out influence from auditory information added, the sound system was used as additional information source– legend to both scale models. By including specific sounds to explain the tangible objects, the possibility of integrating tactile and auditory was investigated in this research.

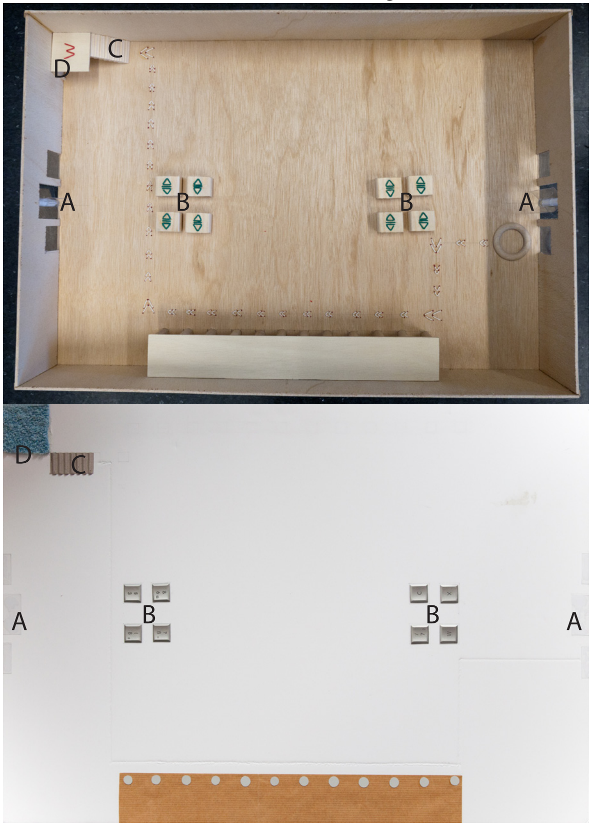


Figure 5.14
The 3D (top) and 2.5D scale model (bottom): A. revolving door; B. elevator; C. staircase; D. wedding room

5.3.2 Setup

The main structure and working process of the sound system remained the same as Concept 2. A webcam detected the position and movement of the cursor ring by color detection, and sound was processed by Max/MSP and played by two speakers. There were six clear sound sources including two revolving doors (A in Figure 5.14), two elevator halls (B in Figure 5.14), a staircase (C in Figure 5.14) and a wedding room (D in Figure 5.14). A sound of an elevator bell was selected as the legend of the elevators, while the staircase sound from concept 2 was used to signify the stairs. However, the sound of revolving doors is too subtle to be noticed and the mechanism sound of the engine is not representative or even confusing. After experiments (providing multiple sounds to participants and asking them to select a most representative sound), a mechanism sound of door lock was generated to indicate revolving doors. Although reality sounds were selected for other sound sources, wedding music was the first option to represent a wedding room. Since there was a route planned on each scale model, the reverberation effects of sound sources were adjusted along the route. In order to emphasize the reverberation effects, footstep sounds were introduced as indication. Considering the large open space in the main hall (§2.2), extreme reverberation settings were applied to the sound sources. For the area with additional ceiling and pillars in the scale model, the reverberation settings (room size, decay time, high frequency damping, and diffusion) was adjusted lower*.

5.3.3 User Tests and Evaluation

Participants

Six blind or visually impaired people participated in the tests. The participants covered different genders (three males and three females), visual experience (three visual impaired and three blind), and travel tools (one with guide dog and five with white canes). Considering they joined this research in an unfamiliar public building voluntarily, they can be seen as the active explorers of the target group (§3.2 Persona).

Procedure

Six participants were divided into two groups: three with the 3D scale model and three with the 2.5D model. B&VI participants were evenly distributed for each scale model. The two groups conducted their tests individually in quiet rooms in Den Haag city hall.

Part I Exploring the scale model

Before the test, a short introduction (see Appendix V) was provided, including the project background, the purpose of this test, test setup and the planned route presented in the scale model. The assignment "to find the wedding room" was also given verbally. After the introduction, participants were allowed to explore the scale model. They were requested to inform the researcher after memorizing all information needed. The time of exploration was measured by researcher. Methods of observational research were used to figure out how they used the prototype and how to apply the collected information for navigation. Three questions were asked for the model exploration:

- Could you find where the starting point is? How?
- Could you find where the wedding room is? How?
- Do you use any landmarks to memorize the route? I you do, could you please indicated what they are?

* all mentioned files are included in digital appendix (Chapter 5/ concept city hall/)

Part II Field exploration

The participants were asked to perform the memorized route in the building, with the researcher following. They were requested to use their familiar navigating skills and navigation tools. During the way finding process, they were encouraged to think aloud and the researcher documented the process by taking notes. The time taken to find the destination was also measured for analysis.

Part III interview

A semi-structured interview was conducted after each test. The interview questions covered three main directions: scale model, sound, and strategies in general. The detailed interview questions could be found in Appendix V. More open-ended questions than closed-ended questions were included to inspire them to provide more information.

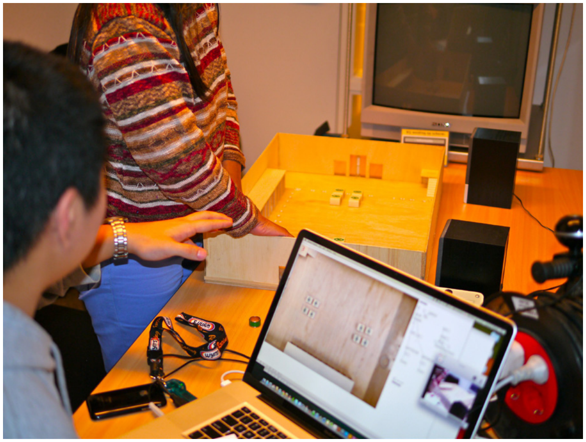


Figure 5.15
Exploration of the scale model (Photo by Henke Baars)

5.3.4 Results

Five participants succeeded in finding the wedding room while one participant using the 3D scale model failed. The results are listed and explained below.

> *"Absolutely!" – one participant answered when asked about whether the sound legend was helpful during exploration of the model*

Overall, participants reacted positively to the sound arrangement in the scale model. Since a similar research was previously conducted without sound, according to the researcher, the time of model exploration reduced significantly with sound information. The decreased time consumption should be attributed to the sound legend. Even with tactile icons, the tangible objects were too symbolic to be recognized directly. Typically, blind users need to firstly understand the meaning behind the icons and then associate the icons to the corresponding object. For example, a tactile icon of up-and-down arrows were placed on the surface the elevator. B&VI users should firstly translate the perceived icon to an up-and-down button, and identify the tangible object of elevator based on the translation. However, with sound information, the translation process was not necessarily needed or much faster, which was confirmed by this test.

Figure 5.16
Navigating in Den Haag city hall (Photo by Henke Baars)

> *"Why is a staircase here? I didn't expect this because it was not on the model." One participant said during the exploration, when he encountered a staircase near elevator hall, which is not*

included in the scale model.

B&VI users are the sophisticated users of sound cues in reality. During the exploration, they identified more sound sources than provided in the scale model. According to the participants, they tended to trust the scale model rather than their navigation ability, and the unexpected landmarks consequently led to confusion.

The only failure in navigation could be ascribed to the misjudgment of the scale. Although the scale of tactile maps and scale models were not frequently used by the B&VI in navigation, according to the findings of user study (§ 3.1), it did actually help them to predict the size of an unfamiliar space. For example, since the tangible object on the scale model, which represents the starting point, was over exaggerated, the participant, who failed the test, misjudged the building to be smaller than it actually is. Therefore, during the navigation, she got lost in the middle of the big hall when she reached a wall, because she assumed that it was the end of the building.

Contradictions of the current sound (a sound clip of door lock) assigned to the revolving doors led to a dilemma: either to choose a real revolving door which may be too subtle to be recognized, or to choose a sound of normal door which might not be linked directly to a revolving door. Although confusion about door types were met during the tests, feedback from users indicated that sound of a lock or normal door would work functionally. Similar issues were confronted when designing sound for pillars, which were used in the model. A balance point between functionality and experience should be worked out.

An interesting thinking pattern of route selection was presented by one participant. When designing the scale model, several clear and intuitive landmarks were included and made physical, while other confusing or potentially dangerous landmarks were excluded. The planned route was to follow the landmarks on the scale model. However, when touching the planned landmarks, one participant intuitively interpreted them as obstacles. Therefore, instead of taking the planned route by following the landmarks, the participants switched to the empty side of the model, which turned out to be full of obstacles in reality. Reflecting on the design of the model, the unbalanced distribution of the landmarks could be one reason behind this situation. Therefore, based on reality, even the landmarks, which are not directly contributing to the planned route, should be included to prevent such situation.

One participant used a guide dog, while the others travelled with white canes. They used totally different strategies when planning the route and memorizing it. Instead of memorizing all landmarks and turning points, the participant with a guide dog just focused on direction and the important landmarks near destinations, such as stairs and doors. The other landmarks were not necessary since the guide dog could prevent him from getting lost. During navigation in reality, he just informed the guide dog to follow a certain direction and to find a stair. However, the strategy turned out to be not functional due to the existence of multiple entrances and stairs in the main hall. The guide dog led him back and forth to different entrances in the hall even when being told to find a staircase.

5.3.5 Conclusion

Based on the findings from B&VI users, several design decisions were made in this part. Furthermore, some questions were still open to test in further concepts and studies.

Auditory and tactile

The potentials of combining tactile and auditory information were verified by facts. The combined representations were evaluated positively by participants and did not cause any conflict or confusion. The auditory information of the wedding room and elevators worked more effectively than tactile information, while the tactile icons of stair and revolving doors were clearer than the sound. Complementary advantages of combining tactile and auditory information could contribute to the scale model design.

Frame of scale model

More sufficient landmarks should be provided in the scale model. Even the irrelevant landmarks for a certain destination could be used for positioning in the building. Moreover, adding all recognizable landmarks to the scale model based on reality would prevent confusion led by encountering unexpected landmarks. Therefore, considering the potential dense distribution of landmarks, the radius of each sound source should be controlled to avoid information overload in certain areas.

Although scale and distance, which were underappreciated by B&VI participants during user interview (§3.1), failed to perform a leading role in the scale model design, it did actually help the B&VI participants to estimate the distance and the size of the space according to the findings from this test. Scale of the model should be added to present the size of the space. The demonstration of scale could make better use of the distance from entrance to where the scale model locates in reality. To achieve this, the location of scale model should be clearly marked on the model base. Therefore, B&VI users could estimate the size of the space by comparing the distance from entrance to the scale model with the demonstration on the model.

If the location of the scale model is presented as mentioned above, the starting point will not necessarily be provided. Users will use the location of scale model as starting point spontaneously. Additional starting points rather than the location of scale model will lead to confusion according to the test. Moreover, thanks to the direct representation of sound, auditory input could be used to emphasize the location of scale model.

Open questions

Although participants reacted positively to the sound selection, the best solution of representing some sound sources, such as the revolving doors discussed above, should be tested out by further studies. Furthermore, a balanced point between abstracted sound and reality sound should be figured out. Thus, a more uniformed set of sound could be developed.

All tests were conducted in quiet rooms, while in reality the scale model would be placed in open public areas. The placing environment of the scale model was not included in the scope of this test, which could be seen as a limitation of this test.

5.4 User Test of Sound

During the user test of the scale model of Den Haag city hall, blind participants encountered functional issues regarding the sound clips added to the scale model. The best way to solve the conundrum was to have blind users make the decision. A user test was set up to figure out the most representative sounds for the common landmarks, which would be used in the scale model design.

5.4.1 Goal

Through this test, the most representative sounds for the selected landmarks would be picked out from a list of sound candidates. By analyzing how B&VI people responded to certain sounds and how they link auditory information to reality, the general criteria of sound assessment for the scale model design could subsequently be drawn.

5.4.2 Methods

Participants

Eleven students (12-20 years old, marked as P1-P11) from Bartiméus participated in the study. The participants were selected covering both blind and visual impaired (7 blind and 4 visual impaired), and both congenital blind and acquired blind (5 congenital blind, 6 acquired blind). They were educated from early age and without mental disorder. Furthermore, a congenitally blind employee (59 years old, marked as P12) of Bartiméus also participated in the test in the end.

Stimuli*

Eight common landmarks or functional places were identified from field research and could be applied to the scale model: toilet, staircase, smoking room, reception, office, elevator, door, and canteen. In total 25 sound clips representing the eight landmarks were made for the test, which are described below respectively. Considering the potential use in scale model, all the selected recordings were mixed down to mono with similar volume level.

Toilet
1. A nine-second dry sound of a toilet flush.
2. A nine-second dry sound of running water. Water is being poured into a container with water. The flow of water changes while pouring.
3. A nine-second recording made in a small room of washing hands with running water.

Staircase
1. A 24-second recording of a female better in high heel walking downstairs in a small stairwell. The walking speed is low and the floor material is hard.
2. A nine-second recording of a male in leather shoes running downstairs in a relatively larger staircase with hard floor material.

Figure 5.17
A congenitally blind employee (59 years old) of Bartiméus also participated in the test for comparison.

Smoking room
1. A twelve-second recording of a male igniting a cigarette, smoking and coughing.
2. A three-second dry sound of a metal lighter, including opening the cap, igniting and closing the cap.

Reception
1. A synthesized female voice of 3 seconds saying “may I help you?”
2. A six-second synthesized dry sound of a service bell.

Office
1. A dry sound of a telephone ring for 8 seconds. It is a typical ring tone for digital telephone.
2. An eight-second recording of a printer while printing

* all mentioned files are included in digital appendix (Chapter 5/ sound test/)

continuously.
3. A recording of typing with a mechanical keyboard of seven seconds.
4. A ten-second recording of an office, including telephone ringing, people talking from distance with clear reverberation.

Elevator
1. A four-second recording of one beep from an elevator bell. The mechanism sound of elevator door closing could be heard in background.
2. A five-sound recording of an elevator door sliding open.
3. A recording of an elevator motor for twenty seconds. The motor starts, works for seconds, and stops in the end of the recording.

Door
1. A seven-second recording of closing an automatic door.
2. A 17-second recording of people walking through a revolving door into a building while talking (obvious reverberation).
3. A five-second recording of door lock while closing and locking the door.
4. A recording of opening and closing a creaking small size door of six seconds.
5. A recording of closing a large size metal door of nine seconds.

Canteen
1. A ten-second recording of cutlery of dishes while washing them, with slight water sound in background.
2. An 18-second recording of a steam coffee machine while making an espresso.
3. A recording of environment sound in a restaurant of 20 seconds, including people talking, cutleries of tableware with obvious space feeling.
4. A 21-second recording made near a cashier, with beeping from scanning barcode and background noise.

5.4.3 Apparatus

All the recordings were all in .wav format (16-bit) and played by iTunes on a laptop. Two-channel speakers were used for the sound reproduction. The speakers were placed in fixed position and the volume was set to a static level throughout the whole process.

5.4.4 Procedure

The test was conducted in a small office in Bartiméus. Participants were invited one by one. The test started with a brief introduction to both the project and the research. Before listening to the sound clips, participants were asked to report the locations represented by the sounds immediately after recognition. The definition of public building was emphasized with examples provided. Each sound clip was played in a single loop (repeating the single clip) till the participants responded. The order of the playlist was randomized before each test. The researcher measured time of the correct answers and took notes of the incorrect answers. In case the participants were not able to communicate in English, a Dutch-English interpreter was involved.

5.4.5 Results

The main results of each landmark are separately described here. Detailed test results can be found in Appendix VI.

Toilet

> *"It must be a fountain…" – even being repeatedly reminded that the sound clips represented certain areas in public buildings, a participant insisted after listening to the sound of running water*

In general, participants performed satisfactorily to all three sound clips of toilet. Three participants failed to recognize the sound of running water in toilet, since the sound shared certain similarity with the sound of drinking water. Although the other two sound clips were both perfectly recognized and linked to the correct landmark of toilet by all participants, the average responding time (5.5 seconds) to the sound clip of flushing toilet was slightly shorter than the sound of washing hands (6.9 seconds).

Staircase

Nearly all participants could describe the sound recordings representing staircases as people walking or running. However, they often failed to figure out whether it was walking on flat ground or stairs. Only three participants per sound clip correctly identified. Comparing the two clips, the performance of the male sound exceeded the female sound because of the changing tempo of walking, which was considered as encountering the turnings of the stairs.

Figure 5.18
The set up of user test

Smoking room

Seven out of twelve participants successfully established the link between the sound of smoking and smoking room while only four correctly responded to the sound of lighter. The misleading cue included in the smoking sound was the breathing sound, which was often related to sleeping according to the participants. Moreover, two participants regarded the smoking or lighter sound to smoking areas outside public buildings, rather than a smoking room.

Reception

The accuracy rate of answers to the female voice sound slightly surpassed the service bell. However, respectively, only five and two participants successfully figured out the two sound clips of reception. Since the female was offering help in the clip, the most frequent misidentification was a shop (mentioned four times during the test). The bell sound was not distinguishable enough to be connected to a service bell. Therefore, various areas of public buildings were mentioned by participants, such as cashier, elevator, or room with doorbell.

Office

> *"A telephone is ringing but I don't know where it is, it could be the reception or an office or any other rooms" – said by a participant after listening to the telephone sound*

The four sound clips selected were all recognizable according to the test. Except for one thirteen years old boy who failed at all four clips, only three mistakes were made while telling the sound sources. However, even when the sound sources were identifiable, issues were encountered connecting the sound sources to a certain area of public buildings. Only three participants of the telephone sound, four of the printer sound and three of the environment sound related the sound to an office. The performance with the typing sound worked best: ten out of twelve participants could tell the correct landmark of office within an average time of 9.5 seconds.

Elevator

The recognition rate of the elevator sound was surprisingly low. Only twelve out of thirty-six total responses (33%) were correct. The accuracy rate of motor sound (5/12) was slightly higher than the other two sounds of an elevator bell (3/12) and the opening of an elevator door (4/12). The mechanical sound of an elevator door and motor sound were occasionally confused with more accessible machines such as a washing machine.

Door

All participants succeeded in recognizing the door represented by the sound of opening and closing of a creaking door within an average time of 4.6 seconds. By contrast, the sound of an automatic door and a revolving door failed in that no participant came up with the correct respond. The sound of a door lock (10/12, average time 5.7 seconds) and the sound of a large metal door (7/12, average time 7.4 seconds) could be seen as potential alternatives. As mentioned above, the mechanical sound could mislead participants to other more familiar objects.

Canteen

> *"I heard dishes and water, it could be a kitchen or a café, and I guess café if it is a public building" – the thinking process of a participant*

Once again, the mechanical sound of a coffee machine failed to represent the canteen or coffee corner of a public building. Half (6/12) of the participants were confused by the machinery sound. Nevertheless, participants performed significantly better in identifying the other three sound clips (cutlery sound, environment sound, and cashier sound), they correctly responded to all thirty-six sound sources with only eight errors made to connect them to landmarks. Among those three sound clips, the sound of cutlery worked best with the accuracy rate of 11/12 and the average time of 6 seconds. The environment sound of canteen could be used as an alternative with the same accuracy rate and slightly longer responding time (9 seconds). Because the cashier sound could be confusing and misconnected to other areas such as shops and supermarkets, six participants failed to figure out it was representing canteen areas in public buildings.

5.4.6 Discussion

In this part, the contribution of research results to the scale model design are discussed.

Experience

The results of several sound clips failed to meet the expectations, especially the sound of elevator and printer. After finishing each test, additional questions were asked to participants about their experience with elevators and printers. The answers confirmed the assumption that their sound perception was limited by their daily experience. For example, one interpreter of the test, who was also an employee of Bartiméus, mentioned that elevators in Bartiméus were only used for multiple disabled students and emergencies. The participants who were living in houses without an elevator had little experience of elevators, which could be the reason behind their failures. In comparison, the relatively older participant (P10, 20 years old) performed better in both recognizing sound sources and telling the represented landmarks. To better verify this assumption, a 59 years congenitally blind employee of Bartiméus (P12) was invited to the test. Predictably, she responded faster (average time of 4 seconds) and with higher accuracy rate (17/25) than any teenage participants. Referring to the scale model design, daily experience of the B&VI should be taken into account. Additionally, more common sound sources should be chosen and inaccessible ones should be avoided.

“Chorus” vs. “solo”

> *“The more you can identify from sound, the bigger chance you can recognize the area.” – one senior participant mentioned when being asked how to represent an office*

The strategy of “a single sound to represent single sound source” was employed at the beginning of this project, which could be seen as a stereotype. Although multiple sounds were applied to one sound source (see §5.2.2, three sound clips for toilet, four sound clips for office) in Concept 2, the sounds were played exclusively one after another. The possibility of mixing multiple sounds to represent one area in scale model had not been tested so far. Since the four sound clips selected for office were not representative enough, a follow-up question was asked to participants “which sound do you think can best represent an office?” Several participants suggested combining the sound of telephone with the sound of typing. In the end of the test, a mixed sound of telephone, printer and typing was played to two non-participant blind students, and one of them successfully figured out the represented area as an office. Although failing to recognize the sound of a printer, one participant suggested that it could be a room like the one where the research took place, which was a small office. The potential of implementing multiple sounds for one landmark should be explored for the scale model design.

Sound loop

An issue arose when the sound clips were playing in loops. Effects of fade-in and fade-out were added to the beginnings and endings of all sound clips in order to create a hint of starting and ending. However, while being played in a loop, time of the gap between attack and previous release was too short for several sound clips, producing a strange unfinished feeling. Since sound clips could probably be played in loops in the scale model design, a certain length of silence should be added to the end of each sound clip or provide a better recording, or edit appropriate loop points in order to achieve a better loop effect.

Language barrier

For the representation of reception, a female voice of “may I help you” was selected. However, during the test, a non-English speaking participant mentioned that he felt the voice was asking for something but he did not understand the meaning. When applying linguistic information to the scale model, the language should be selected based on the need of majority of users, even the simple words, such as greetings. The same issue could be encountered when designing the instruction of scale model: non-Braille users should be taken into consideration if Braille are applied.

Congenital blind vs. acquired blind

The debate on a difference in sound perception between congenitally blind people and acquired blind people was ongoing for a long time. The distribution of congenital blind and acquired blind was considered while selecting participants. However, the results of congenitally blind participants and acquired blind participants had no apparent difference. Therefore, based on this research, no extra attention should be paid on the visual experience when designing sound for the scale model.

5.4.7 Conclusion

In this part, the decisions of sound selection are made. The criteria of sound selection drawn from the test are elaborated on. Additionally, remarkable findings that can contribute to the design of scale model are discussed.

Selection

Selection was made from the sound candidates based on the test and would be applied to the scale model. For toilets, the flushing sound was selected and the sound of washing hands could be used as alternative. The sound of opening and closing of a creaking door should be selected to represent doors in the scale model, and the lock sound and closing large size metal door could also be used as backups. Cutlery sound could best represent canteen and the second option could be the environment sound of restaurant. The sound of smoking could be selected for the smoking room. The sound for staircase should be edited to better illustrate the feeling of walking on steps. For reception, elevator, and office, multiple sounds could be applied and evaluated, since a single sound clip was not representative enough.

Criteria of sound selection

More auditory landmarks than the selected eight may be needed for the scale model design. Therefore, criteria to make the selection are needed and are discussed below.

Firstly, too general sound sources could not be used alone to represent a landmark, for example, the telephone sound from the test was far from enough to represent an office. If no representative sound source can be found, then mixing multiple general sound sources could be one solution.

Secondly, sound should be representative. For example, the service bell sound worked well to represent reception, however, it was frequently confused with elevator bell and doorbell during the test. Another example could be machinery sound like elevator engine and automatic door. It seemed obvious when an idea already existed in mind. However, without the idea, the participants found no clue to identify the sound source. In order to prevent such confusion, more representative sound sources should be selected.

Thirdly, if possible, dry sounds work better than for real recordings. The spatial feeling embedded in the recording could mislead users when perceiving space. For example, the recording of a revolving door used in the test embedded a feeling of huge space. Even when it is applied to a revolving door in smaller space in the scale model, the spatial feeling embedded could not be eliminated.

Scale model design

Based on the discussion above, multiple sound sources would be added to the scale model to represent one landmark if one sound source is not enough. Moreover, when adding any linguistic information to the scale model, the language (including Braille) of users should be taken into consideration.

5.5 Design Guidelines

Design guidelines for the final concept as well as other auditory navigation tools can be drawn based on the outcome of literature search, findings of user studies and the feedback from user evaluation of prototypes. Each guideline is described respectively below.

Proper size of scale model

According to literatures, B&VI people tend to use the body-centered strategy for exploring the scale model. Considering installation of webcam and speakers and multi-layer structure of scale model, it is convenient and beneficial to provide a fixed working direction of the scale model. This implies that users should be able to reach any corner of the scale model by hand and manipulate sound reproduction without changing positions. To achieve this, the size of scale model should be determined based on relevant ergonomics data of target users.

Incorporating a clear starting point

B&VI people make use of scale and distance of a scale model to estimate the real size of the represented public building, according to the findings of the concept evaluation. Therefore, incorporating a clear scale is beneficial from the perspective of functionality.

Precisely positioning a starting point, where the scale model would be placed in the building, could be an effective solution for providing a scale. B&VI users could estimate the size of the space by comparing the distance from entrance to the scale model with the demonstration on the model.

Providing large size cursors of matte and soft material

The glossy material with high light reflection, which was applied to the cursors of several previous concepts, led to unstable output of color detection. For the final concept, if color detection and tacking will be used, it is favorable to select matte and soft material of larger size, such as a finger sleeve. Moreover, choosing a distinctive color for the cursor could prevent the influence from the color of human body and clothes.

Including more sufficient landmarks

According to the user test with B&VI participants, the irrelevant landmarks for a certain destination could be used for positioning in the building. Moreover, adding all recognizable landmarks to the scale model would prevent confusion led by encountering unexpected landmarks in reality.

Adding sound to each landmark

The doors without sound were treated as irrelevant landmarks or even ignored by participants during the evaluation of Concept 1 (§5.1). Therefore, sound should be added to all landmarks no matter whether it is relevant to a destination of not. Considering the potential dense distribution of landmarks, the radius of each sound source should be controlled to avoid information overload in certain areas.

Selecting proper sound for landmarks

According to the sound test with B&VI participants, distinguishable sounds should be selected to represent landmarks. Too general sound, which could lead to confusion with other sound sources, should not be excluded. Dry sound with higher S/N ratio works better than reverberant environment sound.

Multiple sounds for one landmark

If a proper mix is applied, multiple sounds could be used to represent one landmark, which was confirmed by the test of B&VI users. Therefore, rather than searching for the most representative sound for one landmark, combining multiple sounds could be an effective solution.

6. Final Concept - Audigator

In this chapter, the final design concept, Audigator, will be introduced. The main components and workflow of Audigator will be explained in detail. A user test session was conducted with twenty-two blindfolded participants (including two pilots). The main findings will be described and discussed in this chapter.

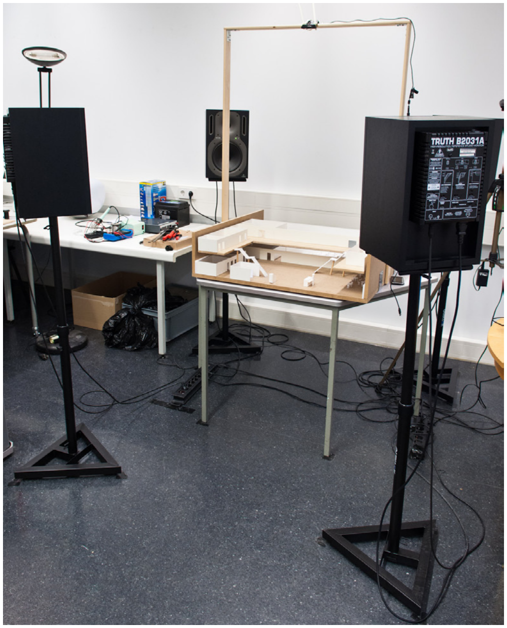
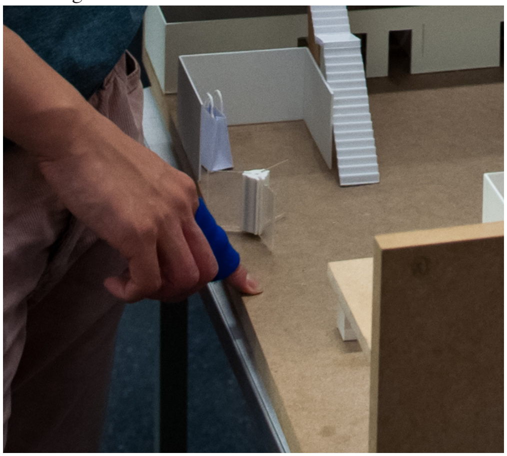


Figure 6.1
Main components of Audigator

Figure 6.2
Blue finger sleeve as a cursor



Combining the two words of "auditory navigator", the final concept is named as "Audigator". Audigator is an audio system for B&VI users and it could be adjusted and applied to any interior scale models of public buildings. Since the programming platform of Max/MSP was proved effective by the previous concepts, Audigator was also developed based on it. The input method of video color detection and tracking was implemented considering the positive evaluation. Thanks to the new introduced four-speaker setting of sound reproduction, panning of sound can be brought in again to enrich experience and help users to better locate the sound sources. To present the system, a two-layer scale model of a factitious congress center was built with the drawer structure tested in Concept 1 (§5.1.3).

6.1 Concept Description

Figure 6.1 shows the main components of Audigator: four speakers, a webcam, a blue finger sleeve, a computer with Max/MSP installed (not included in Figure 6.1), a 4-channel external sound card (not included in Figure 6.1), a distance sensor with a Phidget interface kit (not included in Figure 6.1), and a two-layer moveable scale model with tangible objects. The blue finger sleeve (see Figure 6.2) is used as the cursor (represents user in the scale model). The webcam mounted on top of the scale model is used for the video input to track the cursor. The distance sensor and Phidget interface kit are connected to the computer to detect the position of the upper layer of scale model. The 4-channel external sound card is also connected to the computer to drive the four speakers, which are positioned for quadraphonic sound output.

6.2 How It Works*

As illustrated in Figure 6.3, Audigator starts from cursor inputs (including auto detection of cursor and cursor simulation) and outputs quadraphonic sound (four-channel surround sound) with four speakers. The workflow of Audigator consists of two main phases: cursor tracking phase and sound processing phase. Cursor tracking phase translates cursor movements on the scale model into signals to trigger certain sound output. Sound processing phase processes sounds based on signals obtained from cursor tracking phase and preset attributes in the program. The processed sound will thereafter be sent to the four speakers for sound reproduction.

* all mentioned files are included in digital appendix (Chapter 6/)

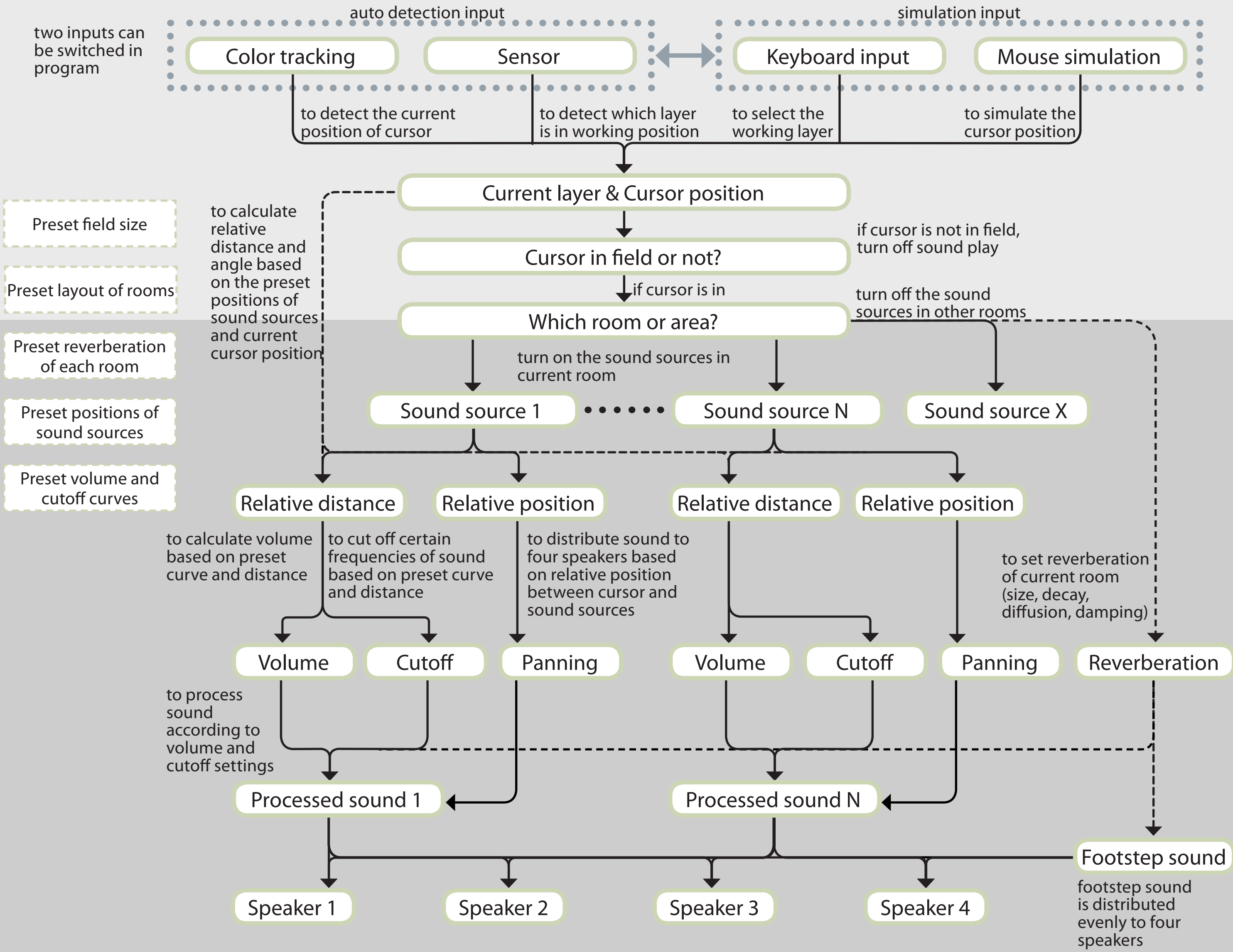


Figure 6.3
Workflow of Audigator

6.2.1 Cursor Tracking Phase

Audigator provides two options for cursor tracking: auto detection of cursor and cursor simulation. The auto detection of cursor makes use of the above mentioned blue finger sleeve and webcam. The webcam captures video and sends to the connected computer for real-time process. The color detection toolkit (see §5.2.2) detects the preset color in capture video and outputs the position of that color. Cursor simulation is developed and included mainly for the purpose of parameter and sound adjustment. When the input method is switched to cursor simulation in the program, users are able to simulate cursor movement with a mouse and select working layer by keyboard. The outputs of both tracking options are X/Y coordinates of cursor with a number (0 or 1 for this two layers scale model) indicating the active layer of scale model.

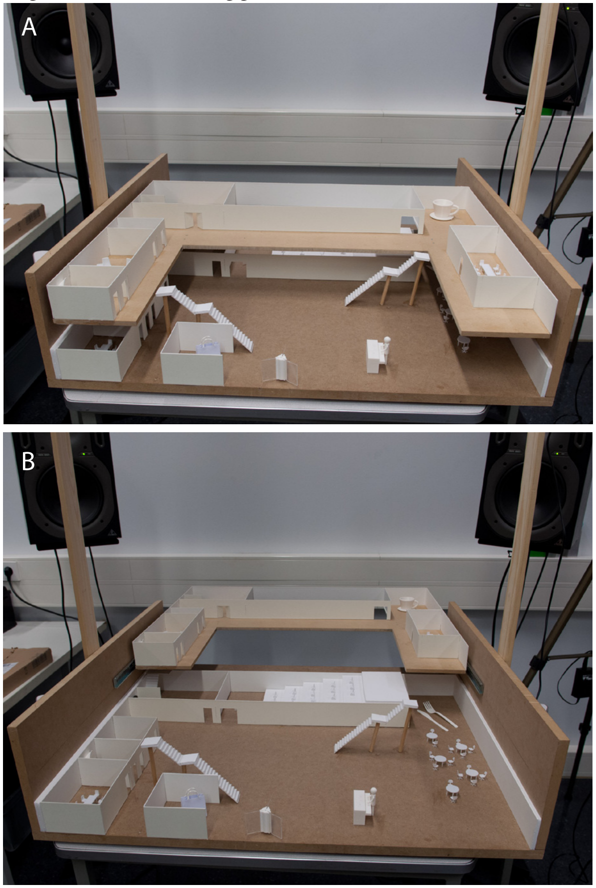


Figure 6.4
The two working positions of the scale model (A: first floor in working position; B: ground floor in working position)

After receiving the number of active layer, the program will consequently switch on the corresponding patch of the layer and compare the X/Y coordinates of cursor with preset model boundaries to detect whether the cursor is in model field or not. If so, the sound play will be switched on, and if not, the sound signal will be cut. When a cursor is detected in model field, the coordinates of cursor will be compared to the preset room layout in the program to figure out in which room the cursor is. The room name will be packed and sent to next phase with the current layer number and X/Y coordinates of cursor.

6.2.2 Sound Process Phase

Initiated by the packed message from cursor tracking phase, sound process phase starts from sounds selected from previous sound study with B&VI participants (§5.4). The sound sources in the room or area, where the cursor exists, will be switched on. Consequently, sound sources in other rooms or areas will be muted. For each sound source switched on, comparison between position of the sound source and current cursor position will be made to adjust volume based on preset decay curves. Moreover, preset cutoff curves are applied to cut off high frequencies of the sounds (will be explained in §6.3.7) based on the relative distance between sound source and cursor. The processed sound will be distributed to four speakers according to the panning setting generated from the relative position between cursor and sound source in the scale model. In a parallel patch, footstep sound is generated corresponding to the moving speed of cursor and it will be distributed evenly to four speakers.

6.3 Under the Hood

The main components and features of Audigator mentioned above are further explained below.

6.3.1 Color Tracking

Several improvements were made to achieve more accurate and stable color tracking output. On the one hand, as mentioned in the design guidelines (§5.5), a larger finger sleeve made of blue elastic fiber material is used to replace the former cursor ring. The finger sleeve is designed to cover both joints of right index finger (left when left-handed). Therefore, it will significantly reduce the chance of losing tracking of cursor when users exploring the scale model and changing hand gestures. On the other hand, a wide-angle high-definition webcam (Logitech C615) was introduced. Benefited from the wide-angle webcam, the maximum size of the scale model could be extended to fit larger buildings with more details. The high-resolution video will increase the accuracy of color detection.

6.3.2 Drawer Structure

A new drawer structure of the scale model is developed based on the cabinet model of Concept 1 (§5.1.3). As shown in Figure 6.4, the main frame of the scale model is made into a U-shaped supporting structure with front, top and back walls removed. When the two layers are fully folded, users are able to explore the upper layer and the atrium area of the lower layer. When the first floor is slid aside, the whole ground floor will be exposed to the webcam. In this expended position, users are able to explore the objects on the ground floor.

6.3.3 Sensor and Phidgets

Phidgets are a set of “plug and play” building blocks for USB sensing and control from computers. For Audigator, a Phidget interface kit is used to connect and control a distance sensor, which is installed in the scale model to detect the two working positions (the expanded position and the folded position). When the upper layer is in working position (Figure 6.4 A), the distance sensor on the board will be exposed with a larger distance sensed. By contrast, when the upper layer is slid aside, it will cover the sensor and the sensor will return a relatively lower distance value. With a preset threshold value of the distance, it is able to detect the current position of the upper layer. Thus, the corresponding patch of each layer will be activated.

6.3.4 Factitious Congress Center

The scale model was built based on a factitious two-layer congress center. The factitious building consists of eleven rooms (Figure 6.5): three meeting rooms, two toilets, two elevator halls, one office, one smoking room, one cross-layer auditorium and one semi-open souvenir shop. There is a large atrium in the middle of the building containing two stairs, a canteen, a coffee corner, a reception and a main entrance. Since the atrium is divided into small areas by rooms in the middle with different ceiling heights, based on the reverberation features, the atrium is divided into six sub-areas on the ground floor as well as six sub-areas on the first floor. Most rooms connect to the atrium by a door, with the exceptions of the two toilets and the cross-layer auditorium, which have two doors.

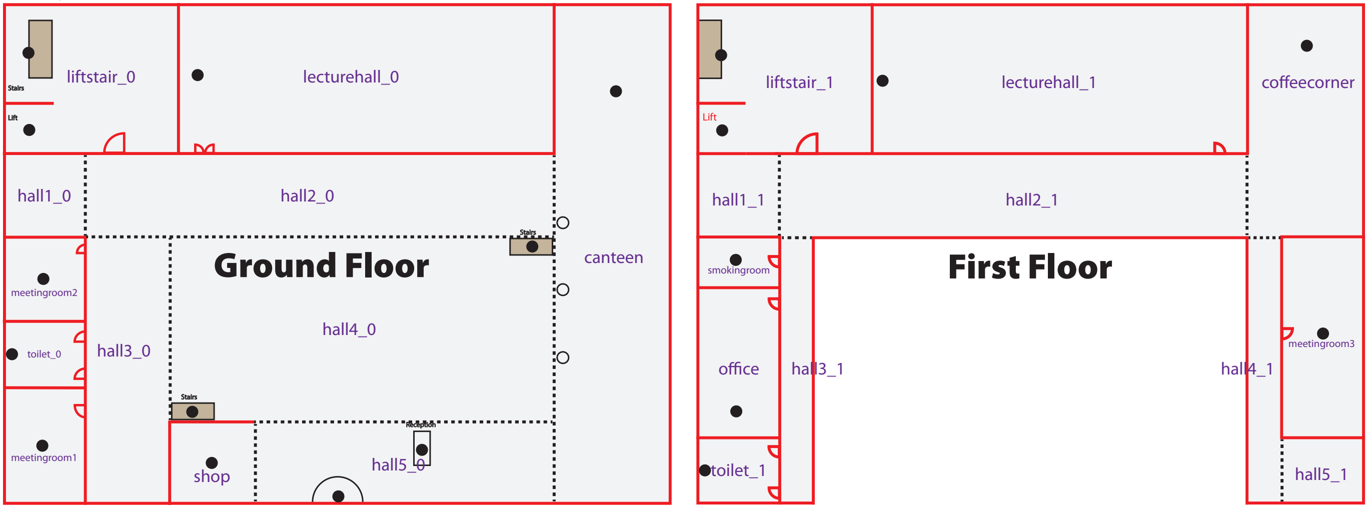
6.3.5 Reverberation

For each area or room, four parameters (size, decay time, high frequency damping, diffusion) of reverberation settings are predefined in the program. Size and decay time are set mainly based on the size of room/area and height of ceiling. High frequency damping and diffusion are set according to function of the room and wall/ceiling/floor materials. All sound sources in the room as well as the footstep sound are played with the reverberation setting of current room.

6.3.6 Sound Sources

According to the results of the sound test with B&VI participants (§5.4), the flushing

Figure 6.5
The layout of both floors



sound was selected for toilets; the sound of a female walking on stairs was used to represent staircases; the sound of locking a door and the sound of a large door were chosen for the doors; the cutlery sound and the canteen sound was selected to represent the coffee corner and the canteen area in the scale model. Several adjustments were made for some sounds based on the conclusions of the sound test: the sound of a lighter and the smoking sound (kept only the parts of smoking and coughing) were mixed to represent smoking room; a sound mix of the typing sound and the telephone sound was used for the office. A new sound recorded near the reception of Den Haag city hall was applied to represent the reception in the scale model. And a new mixed sound of elevator door sliding and engine was made for the elevators in the scale model. Since, in the previous sound test, the cashier sound was mostly recognized as a supermarket, it was used in Audigator to represent the souvenir store. For the new introduced auditorium and meeting rooms, no match could be found from the tested sound clips. Therefore, an applauding sound and three meeting sounds were applied respectively to the auditorium and meeting rooms.

Since most of the functional rooms are blocked from the atrium by doors, when exploring the atrium, users are not able to hear the sound sources inside the rooms but only the sound of doors. Therefore, when the cursor staying in the atrium, the function-representing sounds inside rooms and the respective sound of doors will be mixed, which allows users to tell the function of a room when passing by it.

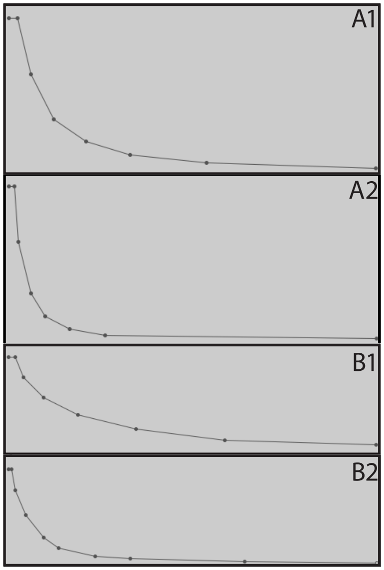
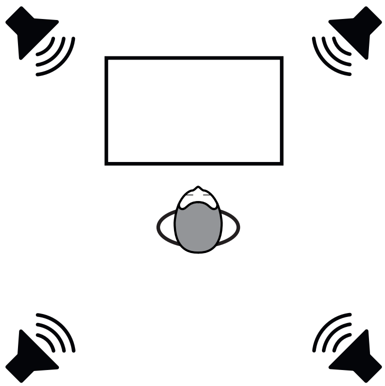


Figure 6.6
Curves applied to volume control and high frequencies cutoff
A1: cutoff curve for normal sound sources
A2: cutoff curve for doors
B1: volume curve for normal sound sources
B2: volume curve for doors

6.3.7 Sound Setup

For the previous concepts, the decay of sound with distance was set linearly in the programs. However, in reality, sound decays with distance following an ‘inverse square law’: sound level will decline by 6dB for each doubling of distance [50]. Therefore, for Audigator, instead of linear change, several curves (Figure 6.6, B1 and B2) were set for the decay in volume with distance. As shown in Figure 6.6, curve B1 was applied to normal sound sources, while curve B2 was applied to doors in order to prevent chaos in the atrium, because the doors in the atrium were close to each other. In addition, it is always necessary to take attenuation into account due to the absorption of sound by the air, which may be substantial at higher frequencies. In general, low frequency waves travel further than high frequency waves because there is less energy transferred to the medium. Similarly to volume, curves (Figure 6.6, A1 and A2) were applied to cut off high frequencies as distance increasing.

Figure 6.7
Positioning of four speakers with user and scale model



6.3.8 Quadraphonic Sound

Figure 6.7 shows positioning of the four speakers of Audigator. The user is located in the geometry center of four speakers. By spreading sound signals to four speakers, it

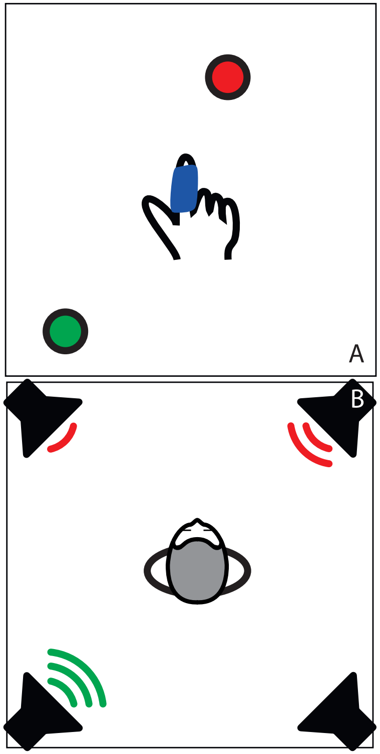


Figure 6.8
A: blue - cursor in the scale model; red - sound source 1; green - sound source 2
B: user and four-speaker setting: red - sound from sound source 1; green - sound from sound source 2

>> Figure 6.9
Tangible objects used in Audigator

is possible to simulate surrounding sound in azimuth plane. The sound panning was processed to represent the relative position between cursor and sound source in the scale model by creating sound in the same position to the user while using (Figure 6.8).

6.3.9 Tangible Sound Sources

Using tangible objects to represent sound sources was evaluated positively by both B&VI and blindfolded participants in previous concept evaluations (§5.2, §5.3). Therefore, a new series of tangible objects were developed for Audigator. As shown in Table 6.1, for each of the twenty sound sources (except for doors), a tangible object was developed. To prevent interfering color detection toolkit, all tangible objects were made in white or black color.

Sound source	Tangible object (number in Figure 6.9)
4 staircases	4 plastic stairs (1)
2 elevators	Two sets of 3D printed indoor buttons of elevators (8)
3 meeting rooms	Long meeting tables with chairs (2)
1 revolving door	A 3D printed axis with transparent plastic doors (1)
1 cross layer auditorium	Stepped floor with chairs arrays (4)
1 souvenir shop	A paper shopping bag (1)
2 toilets	Two 3D printed toilets (6)
1 coffee corner	A set of 3D printed coffee cup with saucer
1 smoking room	A 3D printed pipe (7)
1 office	A plastic office desk with a PC, a printer and a telephone on it, and three chairs (5)
1 reception	3D printed reception table with Braille and a dummy behind it (9)
1 canteen	A pair of 3D printed knife and fork and dining tables and chairs (3)

Table 6.1 Tangible objects applied in Audigator


1
2
3
4
5
6
7
8
9

6.4 User Test and Evaluation

A user test session was carried out with 22 blindfolded participants (including two pilot tests). They were divided randomly into two groups and asked to plan a route in the scale model. Half of them were tested as control group, providing the scale model and tangible sound sources. The other eleven participants were provided Audigator as experimental group. The accuracy rate and time consumption were measured and compared.

6.4.1 Participants

Twenty-two master students from Department of Industrial Design Engineering, TU Delft participated in the test. They were selected randomly covering both genders (9 males and 13 females) and three nationalities (Dutch, Chinese, and Korean). All the participants were well educated without language barrier in communication.

6.4.2 Setup

The test was set up in a quiet room in TU Delft. As shown in Figure 6.10, the scale model was fixed to a one-meter high desk surrounded by four speakers. A laptop was connected for Max/MSP. A video camera was installed behind the scale model to record the whole test.

6.4.3 Apparatus

All sound clips were monophonic in .wav format (16-bit). The four speakers used were all Behringer Truth B2031A. The external sound used was Gamesurround Muse XL Pocket LT3. The front and rear stereo outputs of the sound card were split into two monophonic channels for two speakers.

Figure 6.10
Setup of user test

6.4.4 Procedure

All participants were blindfolded before entering the test room. The researcher led them to the scale model and put their hands on the left and right boundaries of the scale model. Both groups were provided with the same verbal introduction:

> *"The project is to design a indoor navigation tool based on sound for blind and visually impaired people. Between your left and right hands, there is a scale model for a factitious two-layer congress center. Keep in mind the main function of this building is for meetings. I will first give you three minutes to explore freely on the scale model. Don't get into details, just focus on the structure and layout.*

You will come back to the model later. I will inform you the time every minute. There are a lot of tangible objects in the model. If you recognize them and what it represents for in a congress center, please tell me directly.”

For the experimental group, the cursor sleeve was put on their index fingers. They were informed that the sound would be played according to the position of their index finger.

After the three minutes exploration, a task was given to the participants to make a route plan. They were asked to start from the revolving door to reach a series of destinations in the scale model. They were informed the tasks were timed and asked to report immediately after reaching the destination, and the following destination would be provided afterwards. They would not be informed whether the point they reached was correct or not. It was allowed to skip any destinations after exploration and the destination would be marked as failed by researcher. The destinations provided are listed below:

- You just entered the congress room through this revolving door.
- You want to go to the reception to check timetable for a lecture.
- After checking the timetable, you are already late for the lecture. Go to the auditorium, and enter from the back door.
- During the coffee break, you want to grab a cup of coffee from the coffee corner.
- Then you are going to the restroom.
- Smoke a cigarette could be enjoyable. Find the smoking room.
- Go back to the auditorium to finish the lecture.
- After finishing the lecture, you want to meet a friend in an office.
- Take the elevator downstairs and have lunch in the canteen.
- Before leaving, check the store for some souvenirs.
- Leave the building through the revolving door.

After finishing all the tasks, several questions were asked to the participants.

- How do you evaluate the scale model in general?
- Did you encounter any difficulty in planning the route? Why?
- How do you evaluate the drawer structure?
- *How do you evaluate the sound?
- *Did you experience any confusion led by sound?
- *How do you evaluate the tangible objects? Did you figure out the connection between the objects and sound?
- How do you evaluate the appearance?

(*) questions were only asked to the experimental group

6.5 Results

In general, Audigator was proved effective, because both accuracy rate and average time consumption of the experimental group with Audigator exceeded the control group. Detailed results are described below. Raw data could be found in Appendix VIII.

Accuracy

Correct recognition of represented objects or areas and effective route plan was the fundamental function of a navigational tool. Therefore, before jumping into time measurement, the accuracy rate of each tasks were firstly compared and analyzed. A task would be marked as correct only when the correct object or area was found and reported, and an effective route was planned connecting the previous and current destination. Moreover, if a door was clearly mentioned for the destination, for example the back door of auditorium, entering from other doors to the correct destination would also be marked as incorrect. All the skipped tasks were marked as incorrect.

As shown in Figure 6.11, except for two tasks (from main entrance to reception – 90%/80%; from store to entrance – 100%/100%), the accuracy rates of the experimental group with Audigator of the other nine tasks are all higher than the control group. Especially for the two tasks with relatively higher reported difficulty – from the reception to the back door of auditorium and from the office to the elevator, the accuracy rates of participants with Auditory are significantly higher than control group (control/experimental – 20%/80%; 20%/70%). With Audigator, participants reached no lower than 80% accuracy rate for eight out of ten tasks, and finished all tasks with

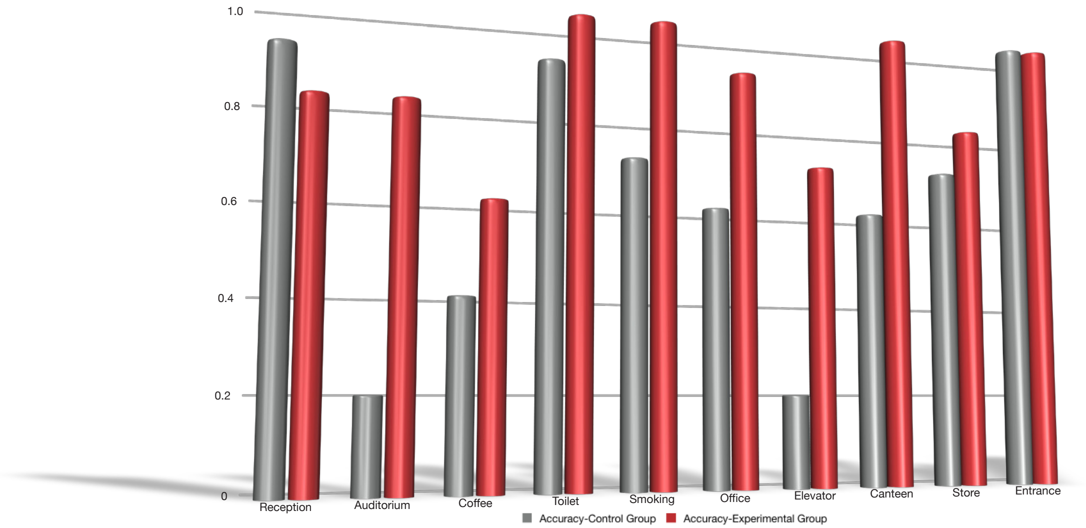


Figure 6.11
Accuracy rate of control and experimental group

higher than 60% accurate.

For the toilet, the smoking room, the canteen, and the main entrance, all participants of the experimental group correctly recognized them and planned routes. By contrast, for the same four tasks, respectively, only 90%, 70%, 60%, and 100% of the participants of the control group made correct route plans to the right destinations.

Time consumption

Time consumption of each task was measured from the task was provided till the participants reported finding the destination. Participants were not allowed to change their previous answers after the following tasks were given. When analyzing time consumption per task, the incorrect answers and skipped tasks were excluded. To make the time consumption more continuous and reliable, the returning from smoking room to the auditorium, which was excluded from the accuracy rate analysis, was taken into account for this analysis.

The blue line of Figure 6.12 indicates the average total time consumption of the control group from the start to each task. The green line represents the same for experimental group with Audigator. The average time to finish all tasks was 411 seconds for the control group and 243 seconds for the experimental group. The time consumption of participants with Audigator was noticeably less than without.

The blue columns show the time consumption of correct answers for each task of the control group, while the green columns represent experimental group with Audigator. Participants of the control group spent more time on seven out of eleven tasks, while they finished faster for the other four. Compared with the experimental group, for the

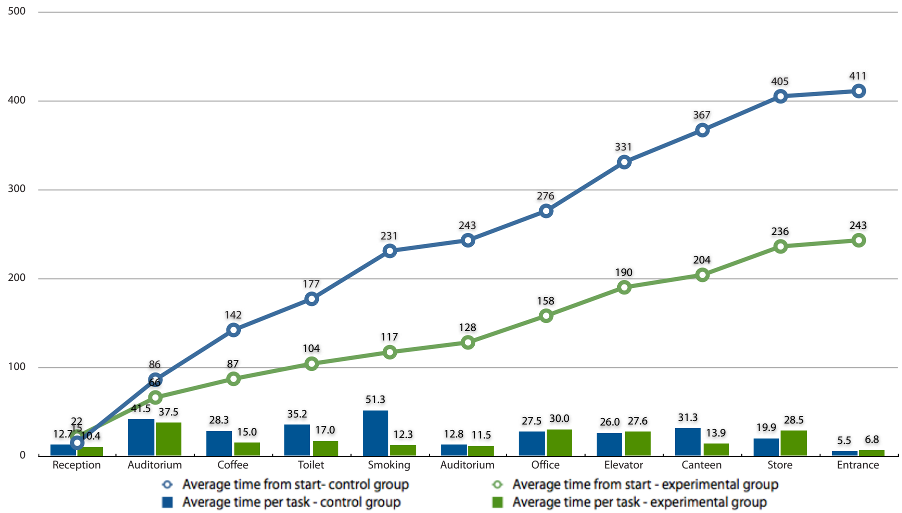


Figure 6.12
Time consumption of control and experimental group

four tasks on which they spent less time, the time difference was not obvious with the largest difference of 43%. However, the control group spent twice to three times the time as the experimental group did in four out of the rest seven tasks.

6.6 Discussion

Remarkable findings from the user test and interviews are described below.

Time consumption

Participants of both groups did not encounter much difficulty in making route plans. They spent most time recognizing the objects and figuring out the represented areas.

As shown in Figure 6.12, the time spent on the first five tasks differed substantially between the control and experimental groups, while the time for last six tasks were roughly the same. It could be assumed that the more time spent on exploring the first five tasks by the control group was sufficient to cover the whole scale model. The same process took less time for the experimental group with Audigator.

The difference of average total time consumptions (including incorrect answers and skipped tasks) between the two groups was more obvious compared with the average total time consumption of correct answers. The reason behind such situation could be that rather than reporting an incorrect answer assertively, participants of control group spent much time exploring or got lost before coming to an incorrect answer. The less cues provided by the scale model without Audigator should be blamed.

Drawer structure

> *"I like the drawer structure, but I didn't use it much, because I can touch the ground floor without sliding the first floor away" – one participant mentioned when asked about the drawer structure*

In general, the drawer structure was positively evaluated. However, the lack of usecue caused difficulty in intuitively sliding the moveable floor. Although the participants were instructed that the first floor could be slid aside, they still tended to touch the ground floor with first floor above it, which could cause the lost tracking of cursor of Audigator. Moreover, two working positions were developed for the drawer structure: extended position and folded position. However, during the test, the moveable layer was often left in between of the two designed positions, which consequently led to malfunction of the sensor and program. Several participants also reported during the test that they were not able to find the upper layer or they did not know which layer they were touching. A more smooth sliding mechanism with clear working positions marked could solve such problem.

The strategies of exploration

In the beginning of the user tests, three minutes were given to explore the scale model without providing any task. From the free exploration of the scale model, four exploring strategies were identified: diverging (explore from areas close to user body to farther parts), converging (start from border to user body), wall following (explore following a wall), and human simulation (use finger as a visitor).

> *"It looks much larger than I felt." – after exposed to the scale model, one participant, who applied converging strategy, mentioned*

More participants applied diverging strategy than converging. According to the interviews, when seeing the scale model after finishing all the tasks, participants applied diverging strategy tended to feel the perceived model was larger than they saw, while participants using converging strategy felt the perceived model smaller than it was in reality. It was relatively more difficult for the participants with diverging strategy to find objects in the far corners, while the participants with converging strategy tended to overlook objects close to their body. Since a cursor was provided to represent user in the scale model, participants of the experimental group with Audigator tended to explore with two hands together or even with only one hand, rather than scanning with two hands separately. Therefore, all participants of the experimental group applied diverging and human simulation strategy. For the improvement of Audigator, frequently used diverging strategy should be taken into consideration by emphasizing objects in far corners, such as enlarging the radius of sound.

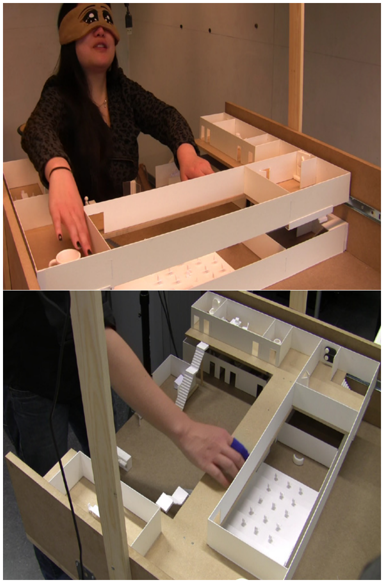


Figure 6.13
A participant of control group uses both hands exploring different areas; a participant exploring with one hand

> *"You should definitely add something on the wall to tell me there is something nearby, otherwise how can I know where to find it" – one participant of control group recommended during interview*

Figure 6.14
A participant started exploration from the far end of the scale model

Participants who explored following walls reported difficulty in finding the reception, which was located in the middle of the atrium. The tangible objects located in the middle of a room were also overlooked. They strongly recommended adding Braille or touchable icons on the wall as a cue to indicate the function of a certain room or area.

> *"You changed the model? I didn't notice these things just now." – a participant of experimental group expressed confusion, since during the three minutes free exploration she did not even find the tangible objects*

Most participants of experimental group applied human simulation strategy because of the existence of finger sleeve. One most reported difficulty they encountered was the design of door. Although the doors were already cut through, the lintel of a door was still an obstacle for those participants, because they were not able to enter the room through the door. Therefore, they often stuck their finger with cursor in the door, and listened to the sound. As a consequence, two participants did not notice the tangible objects inside rooms during the three minutes of free exploration. Removing lintels of doors could be a solution for such situation.

Tangible objects

> *"The tables and chairs are in real scale, so I didn't expect a pipe of this size. I thought it was a toilet or basin." – when the scale model was revealed, a participant recognized the pipe and she mentioned*

The tangible objects was made in uniform sizes rather than uniform scales. The mix of scales led to confusion during the test. There is always a dilemma when making such decisions: either to unify sizes of the icons, which may lead to misperception, or to unify the scale, which means too small objects should be avoided, such as the coffee cup or pipe in this case.

Common mistakes

For the control group without sound cues, the back door of auditorium, the coffee corner and the elevator were the most difficult tasks for blindfolded participants. The mistakes made when searching for the back door of the auditorium were mainly due to the misperception of the cross-layer structure of the auditorium and the two-layer presentation of the scale model. Several participants skipped this task because they failed to find the back door on the ground floor, while actually it was located on the first floor. As mentioned above, the location (in the far corner) of the coffee cup and elevator buttons, which represents the coffee corner and the elevator in the scale model, caused mistakes of participants who applied diverging exploration strategy.

6.7 Limitation

The test was conducted with blindfolded participants. The exploration strategy of B&VI users could be different. The understanding of certain sounds and tangible objects could also be different from blindfolded participants. For example, blindfolded participants could perceive the buttons of elevators wrongly, since they did not have experience with Braille. Therefore, the results could not be directly applied to B&VI users.

Although the matte and soft material of the new applied finger sleeve could reduce direct reflection, which was believed to be the main reason of unstable color detection, occasionally, the output of color detection was still not stable enough during the test. Therefore, simulation of cursor was used as the input method throughout the test. Since the simulation of cursor was manipulated by researcher, delay and errors were unavoidable.

7. Conclusion

In this chapter, conclusions of the whole project will be given. Evaluation of Audigator will be made based on the design goal and design requirements. Recommendations on both research and future improvement of Audigator will also be included.

7.1 Audigator

Audigator is an interactive sound system, which could be applied to interior scale models, for navigation inside public buildings of blind and visually impaired people. With the help of a finger sleeve in blue color, Audigator could track the hand movement of users in the scale model by color detection with a webcam. Sound will be processed and played according to the relative position of finger and preset positions of sound sources in the scale model. Quadraphonic sound setting of four speakers is applied to provide surround sound. Combined with the tangible objects representing sound sources, which were installed in the scale model, the sound play can help users to recognize the main functional objects or areas in the scale model and provide navigational information to the users.

During the development of Audigator, tryouts of 3D interior scale models were also made. A drawer structure was developed for the video detection of multiple layer buildings without compromising the spatial relations between the layers. The upper layers could be slid aside by users, if they want explore the lower layers.

Audigator could be applied universally to interior scale models of diverse public buildings. A set of tangible objects and sounds of common functional areas or objects in public buildings have been proved representative by blind and visually impaired users. The combination of tangible objects with sounds and the program were made modular. By simply placing the objects in scale model and setting the sounds in program, Audigator could be applied to diverse buildings.

7.2 Evaluation

Does Audigator fulfill the design goal?

The design goal of this project is to design an interactive interior scale model of a public building with sound navigation system for blind and visually impaired users and to fulfill their curiosity to explore the building.

Interactive?

Yes, Audigator is absolutely an interactive model, with clear user input – hand movement, with clear product feedback – sound play, and with clear connection between input and feedback.

Interior scale model of public buildings?

Yes, Audigator is developed for interior scale models without ceiling. The sounds and objects are designed for common areas and objects in public buildings.

For blind and visually impaired people?

Yes, sounds using and objects selection are all based on the test with blind and visually impaired people. The needs identified from interviews were taken into consideration throughout this project.

Providing navigation information?

Yes, with the help of combining of sound and tangible sound sources, users are able to recognize the functional areas and objects in the building. By using the scale model with sound, users can formulate spatial relations of the objects and make route plan.

Fulfill curiosity to explore the building?

Yes, users are able to open the building, touch the building, and hear the building as thorough as they want.

Does Audigator fit the design requirements?

Audigator is safe: the main physical parts are made of paper, plastic, and MDF without any sharp borders or corners. The mechanism of drawer track was hiding between the upper layer and the frame of scale model, where hands cannot reach.

The interaction of Audigator is easy and clear: by simply putting hands on the model, the sound feedback will be triggered. However, since the main scope of this project was to focus on the interaction and functional parts, how to get started with the model has not yet been well designed.

Audigator provides a holistic view of the building, orientation inside the building, a clear starting point (in this case a revolving door), main destinations, and reference points. The detail level of Audigator is controlled: all irrelevant decorating elements are excluded, and all helpful landmarks are added. The abstraction of each tangible objects added is also based on the consideration of detail level. Audigator was proved functional through tests with blindfolded participants, and no misleading or confusing information was reported. For the navigation strategy of B&VI users, it has not yet been tested with end users.

For the design of model base, clear floor plans and layouts are provided. The separation between public and security areas has not yet been included and tested, but it is feasible by adding ceilings to the security areas, which will prevent users to explore the unreachable areas. Braille is used for the elevator buttons. However, it is used for the representation of reality, since the elevator buttons in reality also have Braille on it. Through the test, it did not confuse the blindfolded participants without any experience of Braille. No extra linguistic instruction is given.

The color detection toolkit is highly adjustable in the program, which allows it to be fit to various lighting condition. The potential of multiple-user manipulation is kept, although it has not yet been added to current version. A question about the appearance was asked to the blindfolded participants after exposed to the scale model, all of them evaluated it positively. The whole setup looks interesting and attractive to the sighted users.

All sounds used in Audigator have been tested with blind and visually impaired participants. The most representative and effective sounds were selected from the test. In order to make it universal, no linguistic information, such as verbal commands, is involved in Audigator. During the test with blindfolded participants, no confliction between sound and tactile information was encountered or reported. The whole setup is built based on a fictitious building that it is not able to test the influence from the sound play of Audigator to the environment.

7.3 Recommendations

As the starting point of a two-year project (GiZ), the scope of this graduation project is how to help B&VI users independently navigate in public buildings by using the sound system developed. From the perspective of future development, this section provides recommendations on future research and design directions.

7.3.1 Recommendations on Research

Extra attentions should be paid on the following topics for future development.

In the user study phase (§3.1), interviews with the B&VI were conducted to obtain insights of their navigation strategy, and spatial perception. The interviews were carried out in a studio or their working space. The following user studies and initial concepts were all based on the findings of the interviews. However, as mentioned in §5.3, when the B&VI participants were involved to finish tasks in a real public building, their strategy and perception differed as they mentioned in the interview. Even for one blind participant who was involved in both two sessions, conflicts existed between her reported strategy during interview and the applied strategy in reality. Therefore, for future research, if possible, it would be better to bring participants to the real using situation to inspire them and to gather more reliable insights.

The two user studies (§3.3 & §3.4) were conducted and the first two concepts (§5.1 & §5.2) were tested with blindfolded participants due to the difficulty to gather enough B&VI participants. Extra efforts were paid to verify the findings with B&VI participants and conflictions were actually found through the verifying phase. However, the verifying with the B&VI could not cover all the aspects and findings. Therefore, it is not avoidable to apply unverified findings or insights to design. For future development, it is ideal to involve end users (not necessarily to be many, only to verify the insights and to get feedback) during the research and design phase.

The tests of concepts were all carried out in quiet rooms that the influence from environment was not involved. Future tests in real using environment inside public buildings should be conducted.

The tasks for the test of Audigator were provided in a fixed sequence. The difference in time consumption between the control and experimental group was more obvious for the first tasks than last ones. In order to rule out the influence from task sequence, it could be better to involve another group with different sequence of tasks as comparison. Also for this test, participants did not perform the planned route in reality due to the time limitation of the project. The other reason behind this was that the scale model was built based on a factitious building in order to involve all the functional areas or objects. If time permits, it is more reliable to make the scale model based on a real public building, and ask participants to perform the planned route in reality.

The dilemma between unifying the sizes and unifying the scales of the tangible objects should also explored for future development.

7.3.2 Recommendation on Design

Audigator could be improved for the following directions.

For a public design, Audigator is ideal to be used by multiple users at the same time. Although the potential of multiple manipulations was considered and kept during the designing of Audigator, it has not been developed yet. For future improvement, multiple-user construction should be involved.

The drawer structure of Audigator was evaluated positively during the user test. However, more structures could be tested to solve the existing issues of the drawer structure, such as the program errors caused by position of upper layer (in between of two working positions). If the drawer structure will be kept, a more intuitive usecue should be provided for B&VI users.

The unstable output of the current applied color detection toolkit leads to program errors. The improvements made in redesigning the cursor and adjusting settings in program have not solved this problem. Tryouts with other object tracking methods could be made for future improvement.

The material choices of current model base have not yet been tested with B&VI users. According to the interviews, the cardboard material used for the model base seemed fragile while touching. Instead of actually touching it, several participants kept their hands floating in the air just to hear the sound feedback. More materials should be tested out with B&VI users for future development.

According to the feedback from blindfolded participants, it could be helpful to add instructions or certain objects on the wall indicating important functional areas or objects nearby. For future development, tests could be conducted with different setups of this improvement.

8. Reflection

In this chapter, process of the whole project will be evaluated and reflected from the perspective of the author.

Inclusive design, which benefits people of all ages and abilities, is always a domain of design that I have special interested in. To design products and environments to be usable by all people to the greatest possible extent, without the need for adaptation or specialized design, is also the pursuit of my design career. In addition, ever since I was introduced to the world of Hi-Fi during high school, I have been enjoying so much playing with headphones and amplifiers. This project is a perfect combination of my ambition and interests. Consequently, when it is started, I could not help myself diving into the non-visual world.

As things went on, this project drove me into four directions: study of blind and visually impaired people, indoor navigation, sound design, and programming. Correspondingly, I was split into four: the researcher me, the technician me, the designer me, and the programmer me. I was kind of lost switching my role during the exploration phase. Meanwhile, a full list of never-known terminologies drove me crazy and frustrating. It was the interest and ambition mentioned above kept me fighting and finally got survived in the end. When looking back from the end of this project, I do think this is one of the main reasons why I enjoyed this project so much.

The most important thing I learned from this project is how to manage a project as my project rather than a project I am participating in. There is no planned route to follow and no boundaries to keep me on track. I had to make arrangements and decisions. It was pretty tough in the beginning since I had no experience conducting such a project individually and in such a scientific manner. When I got used to the managing role, I started getting into my stride dealing with problems and eventually cracking them all throughout the project.

It was enjoyable to work with my supervisor team, all stakeholders of this project, blind and visually impaired participants, and all my friends I may have more than once blindfolded. From all kinds of supervision and collaboration, I learned how to explain and sell my ideas to both designers and non-designers, how to communicate with both technicians and officials, how to request help from both friends and strangers, and how to get myself inspired and keep myself motived from both colleagues and other designers.

For the process of this project, one thing that I was not satisfied is the final concept Audigator was not tested with end users – blind and visually impaired people. Through the process, B&VI participants should have been involved more. It would have saved much time and effort without the test-and-verify phases of blindfolded tests.

Graduation project is a milestone for my master program. However, I would not regard it as the end of my study. I would continue studying by applying iterative study by design cycles to my future design career as well as to my life, just as the iterative research by design cycles applied to this project.

9. References

[1] Renier, L. A., Anurova, I., De Volder, A. G., Carlson, S., VanMeter, J., & Rauschecker, J. P. (2010). Preserved Functional Specialization for Spatial Processing in the Middle Occipital Gyrus of the Early Blind. Neuron, 68(1), 138–148. Elsevier Inc. doi:10.1016/j.neuron.2010.09.021.

[2] Hub, A., Diepstraten, J., & Ertl, T. (2004). Design and development of an indoor navigation and object identification system for the blind, 852. University of Stuttgart: Collaborative Research Center SFB 627 (Nexus: World Models for Mobile Context-Based Systems) New York: ACM Press.

[3] Loomis J. et al. (2005), Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance, Journal of Visual Impairment and Blindness, Vol.99, pp.219-232, 2005.

[4] Andreou, Y. (2010). Using the voice of the child who is blind as a tool for exploring spatial perception. British Journal of Visual Impairment.

[5] Fletcher, J. F. (1980) Spatial Representation in Blind Children 1: Development Compared to Sighted Children. Journal of Visual Impairment and Blindness 74(12): 381–5.

[6] Fletcher, J. F. (1981) Spatial Representation in Blind Children 2: Effects of Task Variations. Journal of Visual Impairment and Blindness 75(1): 1–3.

[7] Fletcher, J. F. (1981) Spatial Representation in Blind Children 3: Effects of Individual Differences. Journal of Visual Impairment and Blindness 75(2): 46–9.

[8] Kitchin, R. M. & Jacobson, R. D. (1997) Techniques to Collect and Analyze the Cognitive Map Knowledge of Persons with Visual Impairment or Blindness: Issues of Validity. Journal of Visual Impairments and Blindness 91(4): 360–76.

[9] Jacobson, D. R. (1998) Cognitive Mapping without Sight: Four Preliminary Studies of Spatial Learning. Journal of Environmental Psychology 18: 289–305,

[10] Millar, S. (1994) Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children. Oxford: Oxford University Press.

[11] Ungar, S. (2000) Cognitive Mapping without Visual Experience. in Kitchin, R. & Freundschuh, S. (eds) Cognitive Mapping: Past, Present, and Future , pp. 221–48.

London: Routledge.

[12] Warren, D. H. (1994) Blindness and Children. An Individual Differences Approach. Cambridge: Cambridge University Press.

[13] Bértolo, H. (2005). Visual imagery without visual perception?, Psicológica, 26, 173-188.

[14] Casey, S. M. (1978) Cognitive Mapping by the Blind. Journal of Visual Impairment and Blindness 72: 297–301.

[15] Rieser, J. J., Guth, D. A. & Hill, E. W. (1986) Sensitivity to Perspective Structure while Walking without Vision. Perception 15(2): 173–88.

[16] Spencer, C., Blades, M. & Morsley, K. (1989) The Child in the Physical Environment: The Development of Spatial Knowledge and Cognition. Chichester: Wiley.

[17] Juurmaa, J. (1973) Transportation in Mental Spatial Manipulation: A Theoretical Analysis. American Foundation for the Blind Research Bulletin 26: 87–143.

[18] Fazzi, E., Signorini, S. G., Bomba, M., Luparia, A., Lanners, J., & Balottin, U. (2011). Reach on sound: A key to object permanence in visually impaired children. Early Human Development, 87(4), 289–296. 2011.01.032

[19] Nordin, J. (2009). Indoor Navigation and Localization for Visually Impaired People using weighted topological map. J of Computer Science

[20] Ivanov, R. (2010). Indoor navigation system for visually impaired. … 11th International Conference on Computer Systems

[21] Speigle J, Loomis JM. (1993) Auditory distance perception by translating observers. Proceedings of the IEEE Symposium on Research Frontiers in Virtual Reality, San Jose, CA, October 25–26, 1993. Washington, DC: IEEE, 1993:92–9.

[22] Zwiers, M., Opstal, A. V., & Cruysberg, J. (2001). Two-dimensional sound-localization behavior of early-blind humans. Experimental Brain Research, 140(2), 206–222. doi:10.1007/s002210100800

[23] Hofman PM, Van Opstal AJ (1998) Spectro-temporal factors in two-dimensional human sound localization. J Acoust Soc Am 103:6234–6248

[24] Zwiers, M., & Van Opstal, A. (2001). A spatial hearing deficit in early-blind hu-

mans. Journal of Neuroscience.

[25] Hartmann, W. M., & Rakerd, B. (1989). Localization of sound in rooms IV: The Franssen effect. The Journal of the Acoustical Society of America, 86(4), 1366–1373. doi:10.1121/1.398696

[26] Dramas, F., Oriola, B., Katz, B. G., Thorpe, S. J., & Jouffrais, C. (2008). Designing an assistive device for the blind based on object localization and augmented auditory reality. In Assets ‘08: Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility. Presented at the Assets ‘08: Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility.

[27] Speigle J, Loomis JM. (1993) Auditory distance perception by translating observers. Proceedings of the IEEE Symposium on Research Frontiers in Virtual Reality, San Jose, CA, October 25–26, 1993. Washington, DC: IEEE, 1993:92–9.

[28] Ashmead DH, Davis DL, Northington A. (1995) Contribution of listeners’ approaching motion to auditory distance perception. J Exper Psychol: Human Percept Perf 1995;21:239–56.

[29] Speigle, J. M., & Loomis, J. M. (1993). Auditory distance perception by translating observers. In Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in (pp. 92–99). Presented at the Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in. doi:10.1109/VRAIS.1993.378257

[30] Petrie, H., King, N., & Burn, A. (2006). Providing interactive access to architectural floorplans for blind people. British Journal of Visual Impairment

[31] Origin sound by RutgerMuller, downloaded from http://www.freesound.org/people/RutgerMuller/sounds/51163/

[32] Origin sound by ToddBradley, downloaded from http://www.freesound.org/people/ToddBradley/sounds/32908/

[33] Origin sound by Anton, downloaded from http://www.freesound.org/people/Anton/sounds/137/

[34] Dryer. (1993). Cited in O’Shaughnessy & Stadler 2005

[35] M. Dempsey, (2003) “Indoor Positioning Systems in Healthcare”, Radianse Inc.

White Paper.

[36] Da Zhang, Xia, F., Yang, Z., Yao, L., & Zhao, W. (2010, March 9). Localization Technologies for Indoor Human Tracking.

[37] Isaac Amundson and Xenofon D. Koutsoukos, (2009). "A Survey on Localization for Mobile Wireless Sensor Networks", R. Fuller and X.D. Koutsoukos (Eds.): MELT 2009, LNCS 5801, 2009, Pages: 235-254

[38] http://cycling74.com/

[39] http://reactivision.sourceforge.net/

[40] http://www.ece.ucdavis.edu/binaural/binaural_tools.html

[41] Origin sound by ottophokus, downloaded from http://www.freesound.org/people/ottophokus/sounds/78402/

[42] Origin sound by sagetyrtle, downloaded from http://www.freesound.org/people/sagetyrtle/sounds/33986/

[43] Origin sound downloaded from http://www.sounddogs.com/sound-effects/2126/mp3/217440_SOUNDDOGS__fo.mp3

[44] Files downloaded from http://www.freesound.org/people/megamart/sounds/20736/ ; http://www.freesound.org/people/tweeterdj/sounds/30212/ ; http://www.freesound.org/people/sirplus/sounds/16593/ ; http://www.freesound.org/people/thereelfryboy/sounds/35024/

[45] Files downloaded from http://www.freesound.org/people/Anton/sounds/548/ ; http://www.freesound.org/people/ljudman/sounds/23388/ ; http://www.freesound.org/people/acclivity/sounds/24929/

[46] Files downloaded from: http://www.sounddogs.com/sound-effects/2226/mp3/504609_SOUNDDOGS__fo.mp3 ; http://www.sounddogs.com/sound-effects/2226/mp3/504610_SOUNDDOGS__fo.mp3 ; http://www.sounddogs.com/sound-effects/2226/mp3/504611_SOUNDDOGS__fo.mp3 ; http://www.sounddogs.com/sound-effects/2226/mp3/504612_SOUNDDOGS__fo.mp3 ; http://www.sounddogs.com/sound-effects/2226/mp3/504613_SOUNDDOGS__fo.mp3 ; http://www.sounddogs.com/sound-effects/2226/mp3/504614_SOUNDDOGS__fo.mp3

[47] Valente, Michael; Holly Hosford-Dunn, Ross J. Roeser (2008). Audiology. Thieme.

pp. 425–426. ISBN 9781588905208.

[48] Lloyd, Llewelyn Southworth (1970). Music and Sound. Ayer Publishing. pp. 169. ISBN 9780836951882.

[49] Plate reverb, in the style of Griesinger. Randy Jones rej@2uptech.com

[50] FAQ on acoustics, http://www.rockwool.com/acoustics/faq#f1069

Appendix I: Interview questions

1. Background information

- Did you have any visual experience before? If so, how long was it?
- How do you travel for daily activities? In group, couple or on your own? What about exploring unfamiliar areas or buildings?
- If you tend to travel in couple or group, are you motivated by functional or experiential purpose?
- What kind of tools will you travel with? (cane, dog, tracker, etc.)
- Do you have experience with tactile map, scale model or any navigational tools?

2. Strategies of orientation

- Do you have any plan before your travel? If you have, what kind of information or source do you want?
- Do you need a clear starting point to get started or prevent you from getting lost? If so, how to select the starting point?
- How do you memorize a route?
- How do you choose landmarks? And how do you recognize the landmarks (sound, touch, smell, etc.)?
- Do you have an overview of the route with landmarks in your mind? If you have, could you please describe it?
- Do you refer to any object or special spot to localize yourself during the navigation? If you do, what kind of reference do you use? e.g. turning points, starting points, etc.
- What will you do if you get lost? Sequentially trace back to the last reference point or directly to the starting point?

3. Tactile map

- Have you ever used a tactile map or tracker before? How do you feel about the experience?
- How can you link the points marked on the map with reality?
- What do you think about the scale of a tactile map? Is it an important element for you?

4. Experiences on spaciousness

- Are you aware of height and size of a space? If so, how do you measure or perceive it?

- Are you aware of the distance of a route? If so, how do you measure it?
- Do you have special acoustic memories of buildings?

5. Outdoor navigation

- How do you plan an outdoor route? What is your focus and concern?
- What kind of source(s) do you use during your way finding process? e.g. tools, starting point, reference point, landmarks, acoustics, etc.
- Will you use tactile tiles designed for guidance?

6. Indoor navigation

- Generally speaking, how do you evaluate the accessibility of public building?
- What kind of problems do you encounter? stairs, doors, automatic doors, elevators, etc.
- How do you plan a route inside a building? What is your focus and concern?
- Do you have personal preference when making the plan? Walking along the wall or going directly to destination?
- What kind of source(s) do you use during your way finding process? e.g. tools, starting point, reference point, landmarks, acoustics, etc.

7. Scale model

- Do you have experience with 3D model? How do you evaluate it?
- How do you compare the 2D display(tactile map) to 3D display?
- Are you aware of the material features of public building? How do you perceive them?
- Do you have preference of materials? Could you explain your preference in terms of material attribute? e.g. hardness, smoothness, flexibility, temperature, etc.
- Why are you familiar with certain materials? How do you recognize them?
- How do you associate the materials to other things?

Appendix II: Summary of user interviews

	Dick (congenitally blind)	Asha (Acquired blind)	
Background info	• 63 years old • no visual experience before • always travel with partner • use cane instead of dog • can feel the light • have interest in music • prefer to use quick and conveni-ent way to solve problem; not an explorer • doesn't have particular interest in materials	• 48 years old • 45 years blind • travel alone • use cane • totally blind • prefer wood and fabrics because of temperature • she dreams in smell, sound, and tactile	
Strategies of orientation	• memorize the landmarks • create a tactile picture in mind, not visually, but pieces of infor-mation; It needs practice • he memorizes the tactile maps by memorizing certain points in it • he doesn't measure the distance by feet	• safety is the most important • listen for directions; she would choose those permanent sound source • use cane for sound and detecting obstacles • need legend for maps to trans-late from two dimension to three dimension	

	Susanna (partially sighted)	Anja (Acquired blind)	Caroline (Visual impaired)
	• lost color perception since 5 yrs old with 3% sight remained • started learning braille when 7 yrs old • attended special primary school and regular high school • rich experience in using tactile map • running self employed business in tactile graphics training program for 3 yrs • love travelling • mostly travel independently with guide dog • she doesn't use tactile things much • she has an exploring character of taking risks • prefer smooth materials such as wood and glass; also like the sound and warmth of wood	• 65 years old with 40 years bad sight • Travel on her own • Want to be independent • Like walking in free nature • Use both guide dog and cane; the cane is not functional but signal • Using guide dog for the concern of safety • Use tracker for outdoor activities • Experienced with tactile map and scale model	• Became visual impaired since 7 years old • Travel alone, don't want to be dependent though sometimes travel in couple or group • Feeling safe when traveling in group • Use cane and sun glasses • Don't have experience with tactile map and scale model • Always stick to the thing she know, less adventurous
	• rely on the guide dog for obstacles • using a guide dog makes her put more energy on focusing on other things, looking further ahead and enjoying the environment (experiential aspect) • using a cane makes her focus on the route and obstacles (functional aspect)	• always make plan before traveling, basically about the travel information which could be acquired from internet • don't have a clear starting point during the navigation "I will never get lost" • She doesn't have a mind map, just follow her dog	• She doesn't have plan before visiting a building; just ask • She doesn't think that she has a starting point during the navigation. "maybe I do need it, but I didn't notice it." • She could memorize a route based on verbal based instruction • Became more sensitive to sound

	Dick (congenitally blind)	Asha (Acquired blind)	
Strategies of orientation	• he has a good memory to trace back and good sense of orientation • always remember the important objects • recognize the important objects by listening to the sound made by cane • acquire information from wall • never always follow the wall by hand • use the tactile tiles when there is no wall to follow • floor design is more important than wall: much more information • the starting point is very important for finding way back • trace back based on the last landmark, instead of the starting point • time is not important; landmarks are useful	• try to avoid using of stairs • use sound to recognize the material (mostly by tick of cane) • memorize the landmarks • doors (location and type) are most important when preparing for a trip • she can memorize the route as long as she can walk; she only memorize the way she walks • “feet are my eyes”: different materials tell me where I am • when she got lost, she would refer back to the previous landmark rather than the starting point • she creates mental picture • when facing difficulty, she would touch the wall by hand and hear the difference • she creates a 3D tactile model in her mind, which could be touched by hand	

	Susanna (partially sighted)	Anja (Acquired blind)	Caroline (Visual impaired)
	• she has different strategies for exploring unfamiliar areas or buildings depending on the specific purposes: for clear and functional purpose, she would find someone else to guide her to the exact destination; for those she have to travel alone for long time, she prefers to figure out the environment by herself: using cane or guide dog and walking after someone else by following the sound and voice • she has the intuition to create mental map by listening to the environment. The map illustrates an overview of the details of an environment, rather than a route map • she doesn't count steps • doesn't have plan before travelling except for using the website for planning a train route; always follow people, ask for directions and use tactile map • using elevators/stairs as landmarks • she said the learning abilities don't differ very much between born blind people and acquired blind people, but they store the information differently: the born blind rely on verbal description for route mapping and the acquired blind utilize their past memory of visual capabilities • In case of getting lost, she would get back to the mental map and go back until reaching the landmark she is familiar with • when using guide dog, she get information from the floor by listening to the acoustic feedback from her feet and the dog's feet	• she is sensitive to the moving tram, cars, smell and sound of shops, also she notices the way going up and down (crossing a bridge, etc) • She listens to the echo change when stepping on different floor covers • She uses the sound of a passing by tram as an indication for direction • She prefers to ask for help when getting lost, or just listen around to find reference point instead of going back.	• She has a mind picture, like a map, which is not visual like, but consists of a points • She notices the way going up and down and the material change • She would trace back to the last reference point when getting lost

	Dick (congenitally blind)	Asha (Acquired blind)	
Experiences on spaciousness	• The verbal based description could give him a holistic picture, which is hard to get by exploring by himself	• she has a notion of distance, but it is not an important issue for her • she could feel the material of floor by feet, but the pattern is hard to perceive • she creates general impression of a space based on landmarks, smells, sound	
Indoor orientation	• the functional aspects are more important	• echo plays an important role when measuring distance • she uses mostly the entrance as the starting point	
Outdoor orientation	• use a tracker for outdoor navigation • use sun to orientate • memorize the structure of the streets based on the orientation	• identify and remember landmarks by sound and smell	

	Susanna (partially sighted)	Anja (Acquired blind)	Caroline (Visual impaired)
	• use ear to feel distance and orientate • she estimates the distance by time and she would like to focus on the landmarks for most of the time • use echo to localize, but sometimes it is unreliable, such as in pool • Her mental map is 2D, like photo of printed map. Each photo represents one floor and she moves on the photo (Hannes's mental map is a movie) • she do notice the height of ceiling • the atmosphere (surroundings, people, whispers, etc), smell and acoustics create her memory of a building	• she is aware of height and size of a space by echo change • she is aware of time which help her estimate the distance of a route. She has intuition of time measuring • the nice listening experience of a building, such as the reflection from the wall, foot steps, will leave her special acoustic memories. • She uses echo for localization	• She could feel the height of a space, not with sound • She is able to aware the distance of a route by estimating the time or by route planner
	• she doesn't tend to walk along the wall and she doesn't touch the wall all the time • use sidewalk, door, corridor as reference, both acoustically and tactilely • use the height of ceiling as a reference point, based on acoustic change • she prefers to use the central point of a building as a starting point and a landmark • she doesn't use moving stairs because the guide dog is not allowed there; she would use it as a landmark • she notices different floor covers by echo change		
	• she has problem with bike road since it is hard to hear the traffic		

	Dick (congenitally blind)	Asha (Acquired blind)	
Tactile map	• the map makes him independent • the map should have sufficient information • a legend is necessary • need to indicate the starting point, which would be referred to afterwards • scale means nothing • distance means nothing • need direct and specific information to find my route • it's difficult to translate the 2D information to 3D information • the 2D information should be abstract: signs, descriptions • Less tactile information, more essence, more braille • verbal based information is preferred	• she needs to translate the 2D information to 3D image • 3D map is more helpful; she could freely explore it	
Scale model	• would be nice to have the missed information that sighted people have • 3D "real" building would be nice to feel and explore • more than 3 floors would be too much information	• • scale model for the blind is photograph for the sighted • • it is difficult to distinguish glass and metal (both cold and smooth) • • she doesn't have association to material; she focuses on physical feelings	

	Susanna (partially sighted)	Anja (Acquired blind)	Caroline (Visual impaired)
	• 2D map is more functional than 3D scale model • 2D map is easier to read • she needs a translation from 2D map to 3D information		• Too much information, which needs letter-by-letter reading. She couldn't go smooth and she doesn't like the material
	• 3D scale model has too much details for route planning • 3D scale model should include the architecture features, the route with doorway, room, corridor, elevators, etc involved. • the decorative things are not necessary to be included in 3D scale model • the scale of the model is not important, but the relation between the distance and size is necessary to be accurate • never mix two scales in one map • she would perceive the material features by listening to the sound and echo • like natural stuff • interesting to feel that the glass is cold but the temperature would change and tell you something about the environment	• she likes 3D models and she is looking for them all the time • she prefers 3D model since she could really feel it. • 3D model could give her a good overview. • 3D model is more realistic • she doesn't have preference of materials	

Appendix III: Results of Study I: Sound Perception

Recording 1

Partici-pants	Space	Record-ing Position	Sound Sources	Environment	Familiarity
1	Large spacious open space one floor high	Fixed	Cars, wind, door open (metal door), keys, bus stopping at bus stop, putting stuff on the ground	Spacious, not crowded, like factory building or warehouse	In IDE, work-shop
2	5m-6m high, as large as three studios	Moving	People walking, elevator, door opening (normal door)	Strong echo, marble floor	In IDE, work-shop
3	Very large space, around 5m high	Moving	Machine sound, lathing	Hard floor, slightly noisy, open spacious space	Not in IDE
4	Middle size, 5m-6m high	Fixed	Elevator, walking, stairs, door open and close, elevator door	Raining outside	In IDE, lift room
5	Very low	Fixed	Walking, door open (automatic door), moving chair	Empty	In IDE, toilet
6	Narrow, middle size, one floor high	Moving	Elevator, walking	Corridor, quiet, open windows	In IDE, lift room
7	Small size, 1 floor high	Moving	Elevator door, elevator moving, walking		In IDE, lift room
8	Large space, high corridor	Moving	Trolley, door	Hard floor, quiet	In IDE, big size lecture hall
9	Small room 20m2, 3m high	Fixed	Elevator, door, footstep		In IDE, lift room
10	Studio size, 1 floor high	Fixed	Loading trolley, footstep	Strong echo	Not in IDE

Recording 2

Participants	Space	Recording Position	Sound Sources	Environment	Familiarity
1	Smaller than Clip 1, narrow, one storey high	Fixed	People talking, clatter of cutler-ies, beep of chip card	Crowded, busy area, noisy, feel like waiting area/ queuing	In IDE, cof-fee corner
2	Higher than 1, large size	Fixed	Clatter of cutler-ies and dishes, people talking, cashier	Strong echo, queuing	In IDE, cof-fee corner
3	Large size, higher than 1	Cannot tell	Clatter of cutler-ies and dishes, people talking	Noisy, crowded	Not in IDE, Aula
4	Not very large, 3m high	Moving	High heel, order-ing food, walk-ing, dishes	Crowded, hard floor	In IDE, canteen
5	Middle size, very high	Fixed	Chairs (metal), clatter of cutler-ies, talking	Like restaurant	Not in IDE, restaurant
6	Small size	Fixed	Cleaning, trolley, clatter, pouring liquid	Noisy, crowded	In IDE, canteen
7	Large size, higher than 1	Fixed	Clatter, chip card beep, ordering, talking	Noisy, crowded	In IDE, canteen
8	Large size, higher than 1	Fixed	Clatter, tray,	Crowded	In IDE, canteen
9	Less than 3m high, small size	Fixed	Footstep, clatter, tray, talking	In the corner of a large space	In IDE, canteen
10	Large size, 5m-6m high	Moving	Cashier, chip card, walking, clatter	Noisy, crowded	Not in IDE, Aula

Recording 3

Partici-pants	Space	Record-ing Posi-tion	Sound Sources	Environment	Familiarity
1	Very large space, very high, at least three-storey high	Moving	People walking, talking, children playing, breathing	Weak echo, marble floor, walls are not sound absorbing	Not in IDE
2	Very large space (huge hall), multiple stories high	Fixed	People talking and walking by, revolving door	Main hall	In IDE, main hall
3	Narrow corridor, 3m high	Moving	Children shouting, talking from long distance	Probably open window on one side and wall on the other, along road, hard floor	In IDE
4	Very large space, very high	Moving	Moving stuff, people walking by	Crowded	In IDE, main hall
5	Very large open space	Moving	High heel, plastic clapping	Along road, like tram station, very open, floor is not cement	Not in IDE
6	Large size, not high	Moving	Walking, talking, dragging things, handbag rattling	Unfurnished, like garage, many people from far distance	Not in IDE
7	Very large, open space	Moving	High heel, children, talking from distance	Weak echo, hard floor	Not in IDE
8	Very large space, at least two floor high	Moving	People walking by,	Spacious, quiet	Not in IDE
9	Very large space, very high	Moving	Children, clapping, talking, revolving door, footstep	Open, semi-open space	In IDE, main hall
10	Narrow corridor, 4m wide, 1 floor high	Moving	Walking by, talking	Marble floor, like stairs	In IDE, stair near Balustrade

Recording 4

Partici-pants	Space	Recording Position	Sound Sources	Environment	Familiarity
1	As large as two studios, one storey high	Fixed	People walking by, talking, moving chairs, clatter of cutler-ies	Like large lecture hall, door is open or open space	In IDE, Studio near Balustrade
2	Large size, not high	Moving	Clatter of cutleries, chairs moving, talking	Open space, noisy, strong echo	In IDE, Balus-trade
3	Large size, around 5m high	Fixed	People talking	Like lecture hall	In IDE, big lec-ture hall
4	Very high, large size	Moving	People talking, moving chairs (metal)	Hard floor	In IDE, base-ment
5	Large space, 2-3 floor high	Fixed	People talking, sitting, friction	Many people, lecture hall before lecture, hard floor, whistle	In IDE, cafe
6	Large space, 5m high	Fixed	Talking, chairs, door open, pouring liquid	Many people	In IDE, studio
7	Large space, 5m high	Fixed	Dragging chairs, discus-sion, table, footstep		In IDE, coffee corner
8	Large space	Fixed	Dining, chairs, talking	Spacious	In IDE, dining area
9	Large space, 1 floor high	Fixed	Moving chairs and tables, talking		Not in IDE
10	Studio size	Fixed	Talking from distance, mod-elling		In IDE, computer studio

Recording 5

Partici-pants	Space	Record-ing Posi-tion	Sound Sources	Environment	Familiarity
1	Very high open space	Moving	High heel walk-ing, elevator ding, talking	Close to door, elevator hall con-nected to large hall, hard floor, open area, quiet, well sound-ab-sorbing	Not in IDE
2	Not very large, 4m-5m high	Moving	High heel walk-ing, elevator ding, women en-tering elevator, elevator door	Marble floor	Not in IDE
3	Large than an elevator	Fixed	Elevator ding, high heel	Probably inside an elevator	Not in IDE
4	Large space, very high	Moving	Elevator ding, talking, doorbell	Noisy, complicated	Not in IDE
5	Normal size for elevator hall	Fixed	Many elevators, frequent ding, elevator door, doorbell	Closed space, narrow	Not in IDE
6	Large space, very high	Fixed	High heel, talk-ing, elevator	Open space	Not in IDE
7	Large space, 1 floor high	Fixed	High heel, ding	Closed space	Not in IDE
8	Narrow corridor, 3m high,	Fixed	Elevator ding, talking	Open space, like hospital	Not in IDE
9	Small size room, not high	Fixed	Doorbell, foot-step, elevator	Narrow space	Not in IDE, apartment
10	Large space 20m high	Fixed	High heel, people from distance, bell (unfamiliar)	Like waiting area	Not in IDE, airport

Recording 6

Partici-pants	Space	Record-ing Position	Sound Sources	Environment	Familiarity
1	Hard to tell, too noisy	Fixed	Three kinds of talking (chatting; serious talking; service talking),	Open space near door, very crowded, very noisy, like restau-rant, reception	In IDE, dining area
2	Small size, 3m high	Fixed	Microwave oven or oven (ding), people talking	Very crowded, noisy	Not in IDE
3	Large space, 5m+ high	Moving	People talking and walking, ding (not elevator)	Very noisy	Not in IDE
4	Narrow, 5m high	Fixed	Talking, elevator (ding)	Open space	Not in IDE
5	Middle size, two floor high	Moving	Microwave or oven (ding), talk-ing	Crowded, like metro station	Not in IDE
6	Small open room in large space, 2 floor high	Fixed	Bell, open bags, putting things on table	Very noisy	In IDE, coffee corner
7	Middle size room, 1 floor high	Fixed	Bell, people talk-ing	Very noisy	Not in IDE
8	Large size, 5m high	Fixed	People talking, bell	Open space, like market, noisy	In IDE, coffee corner
9	Too noisy to tell	Fixed	People talking, bell	Sound source too close, too noisy	In IDE, coffee corner
10	Small size, not high	Fixed	Bell (familiar), people talking	Many people, very crowded, noisy	Not in IDE, in bus or tram

Appendix IV: Program update log of Concept 2

Version	Changes
1.0	• Origin program with color detection
2.0	• Added reverb effect of sound with for diameters: room size, echo decay time, high frequency damping, and sound diffusion • Added initializing function: it will automatic set the initial position of sound source • Added color background and description for the room setting
3.0	• Added footstep sound: it will change speed according to the movement of the cursor • Added initial value to reverb effect and volume: it will save time when initializing the program • Added auto detection of room (patch: changeroom): when cursor is detected in one room, the audio output would be switched off • Added initializing function to “ezdac~” • Added auto-stop function when cursor is removed: switch off the audio output after the cursor is removed from camera with a delay of 10 milliseconds • Fixed bugs
4.0	• Added the running mean to speed detection: to measure the average of 12 previous values in order to rule out the influence of unstable input • Adjusted the blue range of color detection (from 0.6-0.8 to 0.5-0.9): enlarge the range of blue to achieve a more stable input • Fixed bugs
5.0	• Adjusted reverb settings: more accurate reverb setting of the building • Removed one set of audio for “toilet” and “office” • Added random function and three new audio clips to “toilet” and “office”: change from manual set audio output to random audio • Adjusted radius of sound sources from 50 to 100: clear distribution of the sound sources • Fixed bugs • (Switched from fireware webcam to usb webcam)

6.0	• Added six new sound clips for footstep sound: three for left and three for right, it will play randomly • Added random function to footstep speed: the basic delay is based on the movement speed of cursor, the random delay will be added to the basic delay in order to create more real footstep sound • Combined patches of “stair”, “cafe”, and “reception” • Removed one set of audio for “stair”, “cafe” and “reception” • Adjusted the minimum volumes of “stair”, “cafe”, and “reception” from 0 to 0.1: in room 2, user can hear all three objects in low volume in order to get holistic view of the function areas of current room • Added mute function to the room 2: when approaching a certain sound source with a threshold of 50, the other two sound sources will be muted; when leaving the sound source to a distance of 50, the sources will be unmuted
7.0	• Redesigned the “addcursor” and “removecursor” function: adjust from average value to minimum and maximum value in order to rule out the influence of unstable input • Adjusted the initial value of speed detection: increase the footstep speed • Adjusted threshold of approaching distance of room 2 from 50 to 75 • Edited the sound files of room 2: a more harmony sound display • Fixed bugs
7.1	• Added a mouse detection function: in order to test without video input
8.0	• Added separate reverb settings to different sound sources • Changed the sound of “stair” to high heel sound in order to distinguish it from the footstep sound • (Added a wood figure to the model of reception in the scale model)

Appendix V: Research setup of §5.3

Research Guidance

The research consists of three parts. The first part is about exploring the 3D/2.5D scale model and then we will go downstairs to find the way in field. In the third part, I will ask you a couple of questions.

We will do the first time in this room. You are free to touch and explore the 3D/2.5D scale model in anyway you want. The 3D/2.5D scale model shows the overview of the ground floor plan of the city hall and I will explain the meanings of the objects on the model. You could find the starting point near the entrance, and your task is to find the wedding room, which is connected to the starting point by a route. Please try to memorize the route by using any strategy like keeping the turning points with relevant landmarks in mind. Let me know when you've memorized the information on the model.

In the second part, we will go to find the way to the wedding room in field. I'll take you to the starting point. During the process, please think aloud: say out what you are thinking about and how you connect the real situation to the model you just explored.

In the interview, I will ask you some questions about the experience of using the 3D/2.5D scale model and the way finding process. You are free to give any comment or tip.

I. Exploring the 3D/2.5D scale model

Time ________________

1. Could you find where the starting point is? How?
2. Could you find where the wedding room is? How?
3. Do you use any landmark to memorize the route? If you do, could you please indicate what they are?

II. Field exploration (Time __________)

III. Interview

3D scale model

1. Could you construct the environment in your mind by using the model?
2. Could you link the relevant landmarks to the route when using the model? If not,

why?

3. What do you think of the 3D/2.5D scale model? Please evaluate it from both the functional and experiential aspect.

Sound

1. Do you think sound is helpful or not? What objects did you recognize by sound?
2. Did you have any confusion about the sound?
3. What do you think of the sound? Please evaluate it from both the functional and experiential aspect.

Strategies in general

1. Do you have a mind picture of the route? If so, what does it look like?
2. How do you keep the right direction during the walk?
3. Did you encounter any problem during the way finding process? If so, how did you deal with it?
4. During the field way finding, how did you recognize the landmarks? Could you associate them to the ones on the 3D/2.5D scale model?

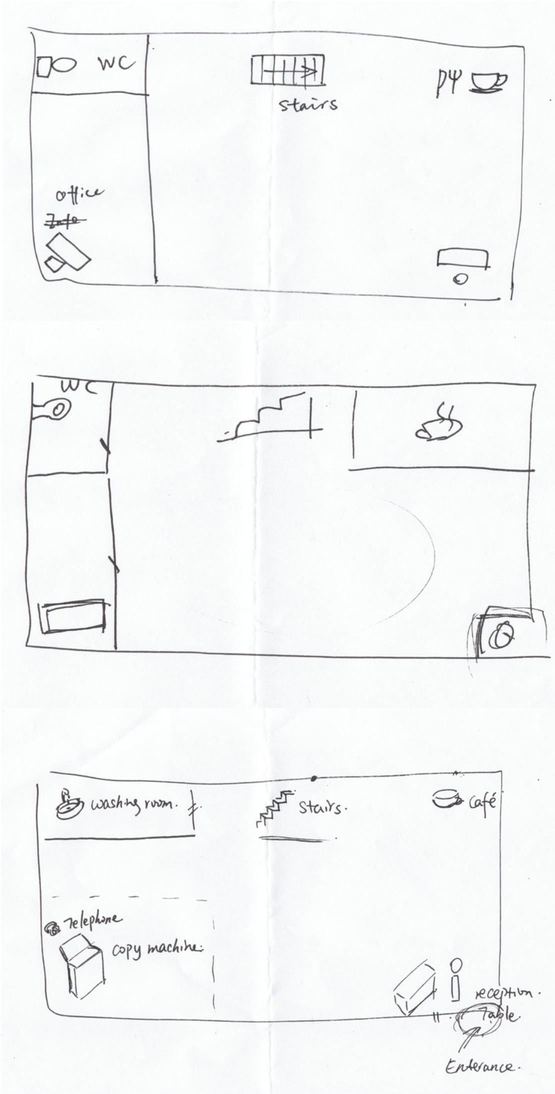
Appendix VI: Results of sound test

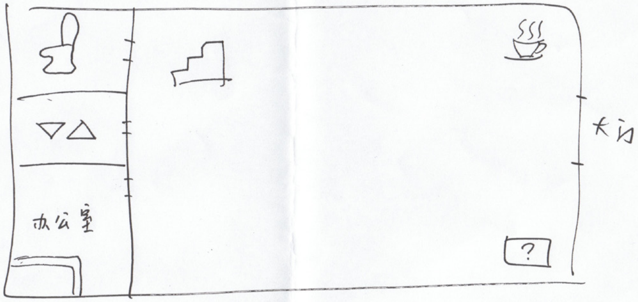
Landmark	Sound	P1 (age: 17)	P2 (*) (age:12)	P3 (age: 13)	P4 (*) (age: 13)	P5 (age: 13)
Toilet	Flush	8	3	2	5	2
	Running water	10	3	2	6	9
	Washing hands	5	7	6	7	4
Staircase	Low speed	dancing	running	3	X	hall
	High speed	dancing	4	school	17	running
Smoking room	Smoking	7	restaurant	12	6	8
	Lighter	outside	X	2	X	haircut
Reception	Female voice	11	4	X	X	shop
	Service bell	market	7	3	X	cashier
Office	Telephone	home	reception	anyroom	X	anyroom
	Printer	classroom	machine	2	X	5
	Typing	7	fixing bikes	3	X	3
	Environment	12	reception	lounge	X	living room
Elevator	Bell	X	12	X	X	alarm
	Door open	18	car	X	X	X
	Motor	train	20	washing machine	X	washing machine
Door	Automatic door	X	trashbin	converyor belt	X	digging
	Revolving door	ferry	platform	X	platform	station
	Lock	20	4	7	X	6
	Normal door	22	3	3	4	4
	Large door	6	5	6	toilet	bottle recycling
Canteen	Cutlery	kitchen	8	4	12	7
	Coffee machine	52	31	ball game	X	10
	Environment	50	3	6	7	party
	Cashier	supermar-ket	supermar-ket	supermar-ket	4	3

P6 (*) (age: 14)	P7 (age: 14)	P8 (age: 12)	P9 (age: 15)	P10 (*) (age: 20)	P11 (age: 12)	P12 (age: 59)
4	20	2	2	1	14	3
classroom	14	4	fountain	kitchen	3	7
4	12	3	3	3	27	2
room	room	room	2 people	room	room	6
hall	hall	X	2 people	2	corridor	2
4	sleeping	Blowing	Eating	3	snore	5
4	X	outside	X	3	X	5
7	shop	X	5	3	shop	shop
elevator	doorbell	home	home	clock	X	clock
anyroom	anyroom	anyroom	small room	1	1	1
7	X	X	home	anyroom	X	6
3	47	5	4	3	5	15
canteen	living room	home	big hall	3	12	classroom
X	any room	X	6	4	X	any room
X	X	7	6	5	X	X
4	X	3	X	2	phone	2
X	X	X	X	security check	X	heavy thing
building	cafe	station	big hall	train	restaurant	train
2		6	5	2	3	2
1	6	2	3	2	3	2
6	X	X	stairs	4	20	5
3	7	7	10	2	4	2
X	music	16	printer	10	X	10
4	13	3	4	3	4	3
6	shop	2	supermarket	3	2	shop

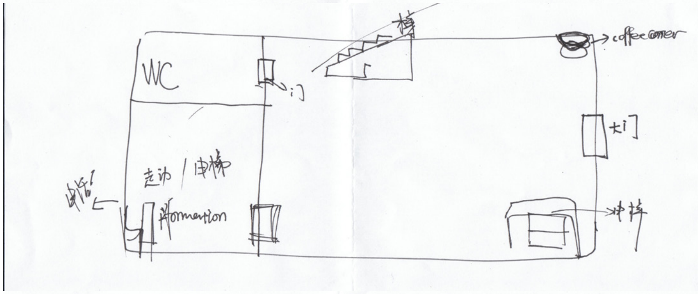
- Congenitally blind participant
- Acquired blind participant (* visually impaired)
- Correct answer (respond time in second)
- Correct object linked to wrong landmark (mismatched landmark)
- Failed to recognize the sound (answer, X=no clue)

Appendix VII: Maps drawn during user test of Concept 2

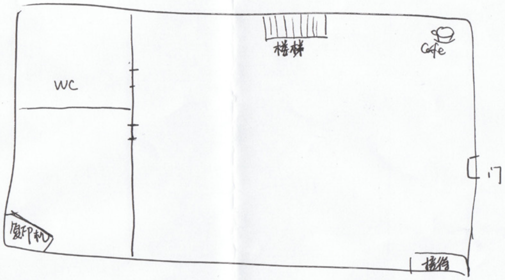




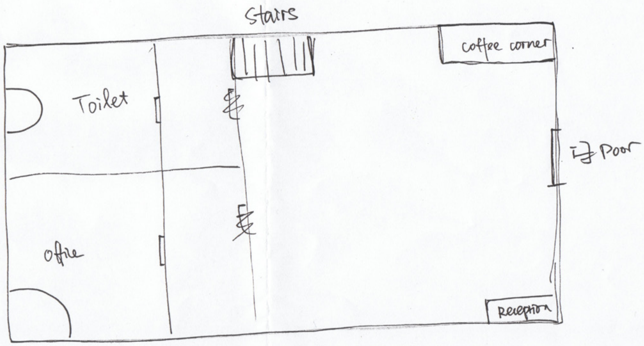
大门
办公室



coffee corner
WC
门
大门



楼梯
Cafe
WC
门
复印机
楼梯



Stairs
coffee corner
Toilet
门 Door
Reception

Appendix VIII: User test of Audigator

	Reception	Auditorium	Coffee	Toilet	Smoking
P1	6	33	6	60	
P2	54	X	80	10	
P3	5	X	X	17	
P4	23	X	12	20	
P5	2	X	X	X	X
P6	4	50	15	10	X
P7	5	X	X	49	
P8	5	X	X	89	
P9	X	X	X	10	
P10	10	X	X	52	X
average-control	12.7	41.5	28.3	35.2	
accuracy-control	0.9	0.2	0.4	0.9	
P11	7	30	X	23	
P12	4	X	25	7	
P13	4	20	15	21	
P14	45	X	11	18	
P15	X	32	X	39	
P16	5	52	X	5	
P17	X	46	X	15	
P18	13	32	9	18	
P19	2	28	13	16	
P20	3	60	17	8	
average-experimental	10.4	37.5	15.0	17.0	
accuracy-experimental	0.8	0.8	0.6	1	

P1 - P10: control group

P11-P20: experimental group

g	Office	Elevator	Canteen	Store	Entrance	Total
93	12	15	12	10	7	263
39	37	37	X	18	4	390
80	12	X	74	12	13	529
15	15	X	65	40	3	295
	X	X	X	X	4	389
	X	X	6	X	8	233
35	X	X	17	16	3	388
88	X	X	X	32	4	678
9	27	X	X	11	5	371
	62	X	14	X	4	573
51.3	27.5	26.0	31.3	19.9	5.5	410.9
0.7	0.6	0.2	0.6	0.7	1	
20	50	7	7	8	6	217
7	20	36	15	13	5	192
14	17	66	13	12	3	195
20	75	28	15	X	5	339
10	9	12	15	38	5	230
15	45	22	20	X	31	325
9	X	X	14	47	4	308
6	24	22	17	40	2	196
15	20	X	7	39	4	235
7	10	X	16	31	3	189
12.3	30.0	27.6	13.9	28.5	6.8	242.6
1	0.9	0.7	1	0.8	1	

X: incorrect answer or skipped tasks

13th May 2011 - 17th Jan 2012