Title	Navigation system based on depth camera and portable device for visual impairments
Sub Title	
Author	Zhang, Qin(Haruyama, Shin'ichirō) 春山, 真一郎
Publisher	慶應義塾大学大学院システムデザイン・マネジメント研究科
Publication year	2021
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2021年度システムデザイン・マネジメント学 第451号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40002001-00002021-0001

Master's Dissertation 2021

Navigation System Based on Depth Camera and Portable Device for Visual Impairments

Zhang Qin

(Student ID Number:81834535)

Supervisor Shinichiro HARUYAMA

September 2021

Graduate School of System Design and Management,

Keio University

Major in System Design and Management

SUMMARY OF MASTER'S DISSERTATION

Student Identification Number	81834535	Name	Zhang Qin

Title

Navigation System Based on Depth Camera and Portable Device for Visual Impairments

Abstract

The movement ability of visually impaired is limited by the disability, and the cost of the society cannot be neglected. Both traditional methods and the electronic assist devices have limits in many conditions. Their problems are mainly about inefficiency and importability of information for traveling or daily movement. This thesis is aiming for designing a wearable navigation system based on a depth camera for the visually impaired to have more sufficient information of obstacles' location. The navigation system is expected to enable the visually impaired to extend the range of their activities compared to that provided by conventional aid devices with more efficiency and be more user-friendly mainly for indoor purposes. To be more specific, a device telling the user with 5 different angles of the front by sound hints have been developed, which will be more informative and precise for user to avoid the obstacles. The system is based on a depth camera (the Intel Real Sense Camera D435), a Windows Stick PC, 5 speakers controlled by a microchip and a portable battery as the power supply. By hearing sound hints from the speaker band wearing on their heads to know the right direction and the layout of the obstacles in front of them with different sounds pitch and loudness. There are 3 levels of the distance circumstances: no more than 0.5m, from 0.5m to 1m, and more than 1m. All the calculation and data are processed by the stick PC in Windows 10. The experiment results show that the system can help the user to find a safe way in a clutter indoor environment and avoid the collision risk successfully with a better efficiency than traditional methods which is more intuitive and portable.

Key Word (5 words)

Depth Camera, Navigation, Portable Device, Sound Prompt, 5 Directions

Contents

List of Figure

List of Table

Chapter 1. The Current Situation of Visual Handicapped

In chapter 1, the paper will discuss about the problems of the visual impairment in daily life, and its influence on the society. It is also mentioned about present condition of assistive solutions for those people concisely. The importance of developing an assistive system for the visually impaired will be discussed and validated.

1.1. The Main Problem

According to the references from the World Health Organization, the number of people with visual disabilities is around 135 million, of which 45 million are blind [1]. This amount includes different scales of visual impairment, where the severest is blindness (visual acuity below 5%). This group represents a 13.6% of the visually impaired people (39 million people in the world) [2]. There are 2 kinds of main problem of those visual handicapped people. One is the autonomous movement, especially movements supporting their basic daily life. Another problem is the information gathering of the surrounding which may lead to some potential risk of collision and injury. What is more, this is not a simple medical problem, psychological issues also occur among the visual handicapped. They tend to have more negative emotions. The confidence and the self-esteem may also be severely challenged during the limited daily life.

Gathering information by vison is the prime method among the human senses to manage movements about the surrounding. It provides a feedback mechanism for balanced interaction with the environment and plays a vital role in the self-navigation. The blinds have to build another information loop to interact with the real world. How to find and build an alternative way with efficiency becomes the key point for researchers to develop systems for the handicapped people.

1.2 The Related Problems

Visual impairment causes a considerable impact on the lives of the blind individuals as well as their families and society.

In 2007, more than 1.64 million people are affected by visual issues in Japan and the total health cost was around 8.8 trillion JPY (72.8 billion USD). That is an equivalence of 1.7% of Japan's gross domestic product at that year. The loss of well-being (years of life lost from disability and premature mortality) cost 5.9 trillion JPY (48.6 billion USD). Direct health system costs were 1.3 trillion (11.1 billion USD). Other financial costs were 1.58 trillion JPY (13.1 billion USD), including productivity losses, care takers' costs, and efficiency losses from welfare payments and taxes. [3] Visual impairment imposes substantial costs on society, particularly to individuals with visual impairment and their families. Eliminating or reducing disabilities from visual impairment through public awareness of preventive care, early diagnosis, more intensive disease treatment, and new medical technologies could significantly improve the quality of life for people with visual impairment and their families, while also potentially reducing national health care expenditure and increasing productivity in Japan. If we can make a more friendly environment for the visual handicapped, the potential economy growth and the cost of the health welfare would become more positive in the future. [3]

1.3 Infrastructure and Assist Devices for the Visual Impairments

There are mainly 2 types of assisting facilities. The infrastructure and assisting devices (a typical integrated scenario shown in figure 1.1). Typical infrastructures are the blind sidewalk bricks, handrails and voice prompts about the traffic signals and other public facilities. The utility of those facilities is highly restricted by its location and the extensibility is almost zero since those are designed and installed at the very beginning of the implementation. For instance, the blind sidewalk bricks are already designed in a certain way when the road was finished. Only the main roads have enough design redundancy for the future construction, and it is impossible to extend it to all the destinations and locations where they want to go. What in more, the indoor assist facilities are severely inefficient for the potential users, many buildings have almost no assist devices or facilities for the visually impaired persons.

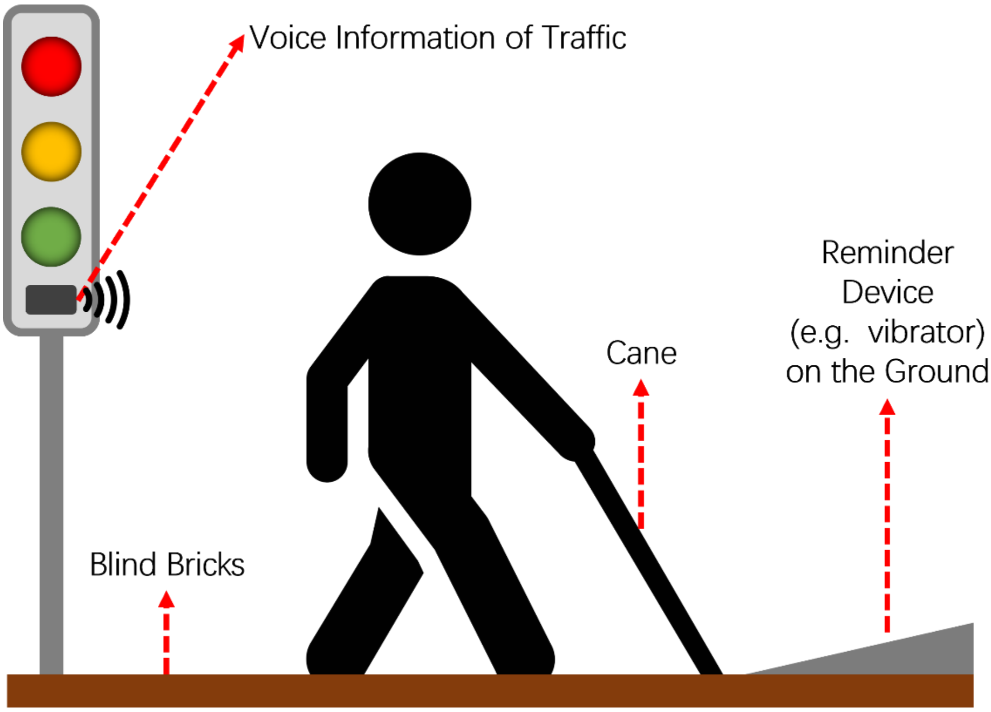


Figure 1.1 A typical integration using scenario of traditional assists

1.4 Negative Factors for the Visually Impaired

Visual perception contributes up to 80 percent of all information that human is receiving. Therefore, limitation or even full loss of vision influences the daily lives of people who are suffering from visual impairment severely. This leads not only direct consequences like loss of ability to see but also problems with movements and daily selfcare independently. As visually people are dependent on help of others, loss of self-confidence, negative self-concept and other problems on psychological level also occurs. Reducing necessity to be dependent on other people during movement has positive effect on multiple problems that visually impaired person is facing to.

Research have shown that vision loss severely lowers the mobility of the visually impaired [4]. As a result, approximately more than 30% of the blind population do not ambulate autonomously outdoors [5].

Visual impairment hinders a person's essential and routine activities [6]. Furthermore, low vision or complete vision loss severely lowers functional independence of the visually impaired. It also highly reduces their ability and willingness to travel independently. Most Existing navigation systems, especially indoor ones for the visual impairment are based on Radio Frequency Identification (RFID) or other similar methods which are highly rely on the indoor network infrastructure; therefore, the generality of those systems are highly limited.

Mainly, these kinds of devices are referred to a system that contains various cameras, GPS and laser sensors. Those Electronic Travel Assistants (ETAs) are replacing the traditional methods [7]. The cameras with computers have a higher detection and calculation ability that they can provide a wider angle of view and distinguish a far more sophisticated surrounding environment. There are many designs of visual assistant system based on this mechanism.

However, many ETAs work with a low performance, the accuracy and the calculation ability are too low for the user to have a comfort travel, what is more, many ETAs are using GPS as the locator which will significantly have a lower performance when it is indoor.

Both contrast sensitivity and visual acuity loss affect independently to deficits in performance on daily activities. Defining disability as deficits in performance relative to a social group, it is possible to determine visual acuity and contrast loss where most are influenced or challenged. Meanwhile, the cut-off points depend on the task, suggesting that defining disability using a single threshold for visual acuity or contrast sensitivity loss is independency [8].

Chapter 2. Conventional Methods

In chapter 2, both the traditional methods and new electronic methods of helping visually impaired people are introduced and compared. Each of the methods has its own advantages and disadvantages in different using scenarios. Also, several devices for visual impairments are picked and compared to figure out the structures and features so the view of current situation about assistant devices for blind people will be clear.

2.1 Traditional Assistant Methods

The most common methods of traditional assistant for the visual handicapped are the canes and guide dogs, and these 2 methods are still the main approach for the current potential users after considering the cost and usability of the new electronic devices. The interview [9] indicated that 12% and 55% of the visually impaired interviewees have used a cane or a guide dog as their aid for the navigation purpose.

2.1.1 Canes and Sticks

Almost all the blind people choose the canes as their first choice of the movement assistant tool. The obvious advantages of it were cheap and reachable. In some extreme conditions, a long wooden stick would work fine as well. The average length of the cane was 1.4 meter long, and the range for the user can detect is no more than its length. That means the movement speed will be highly restricted by the size of the cane. What is more, users still accompany a high risk of collision, because the user have a high risk of collision against obstacles to distinguish the routine.

2.1.2 Guide Dogs

According to the 2011 National Health Interview Survey, 21.2 million American adults aged 18 and older reported having vision loss. However, the guide dog has very limited accessibilities as only about 1,500 dogs are estimated to graduate from a dog-guide user program every year [10].

This situation severely limited the scale of widely application for the guide dog.

Also, there were three main behavioral reasons why guide dogs were withdrawn from service: environmental anxiety, training issues (a lack of willingness to work or confidence), and fear and aggression. Other reasons included chasing, attentiveness, social behavior, excitability, and distraction (table 2.1)[11].

Table 2.1 The Behavior Problems in Guide Dogs about Median Total Working Times and Age at Withdrawal [11]

Group	**Days**	**Years**	**Age**
Fear/Aggression	635	1.74	3.50
Chasing	739	2.02	3.98
Attentiveness	855	2.34	4.46
Social Behavior	945	2.59	4.39
Excitability	960	2.63	4.50
Distraction	986	2.70	4.71
Environmental Anxiety	1152	3.16	4.96
Willingness/Confidence	1617	4.43	6.44
Retired	3107	8.51	10.35

2.2 Electronic Travel Aids

2.2.1 Regular Computer Vision Applications for ETAs

Computer vision as a substitute for visual sense forms a powerful tool for developing assistive devices and applications for the visually impaired people. Applications based on computer vision enhance their ability of movements and orientation as well as object recognition and self-navigation. More recent systems even attempt to support social interaction [12].

There are various computer visions such as RGB vision, infrared vision, Depth camera vision and other types of vision. Mainly those are based on stereo camera views in order to measure the accurate distance to the object. The calculation is carried by laptops usually, which limited the mobility of the system and the level of user-friendly. Many of the conventional methods focused on 2-D image to 3-D mapping (see figure 2.1 and 2.2), then the system will provide navigation information for the users.

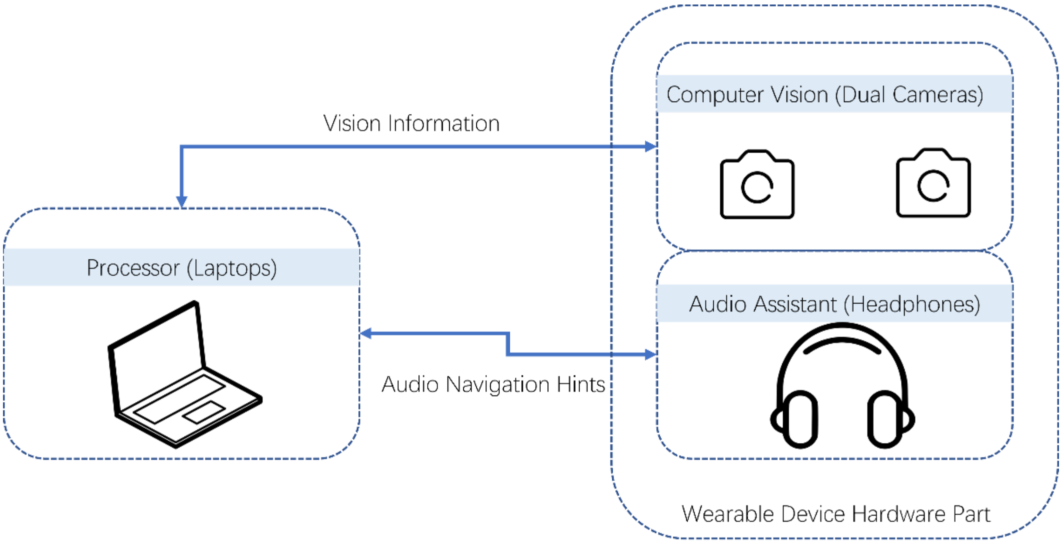


Figure 2.1 A Typical ETA System Components with Computer Vision

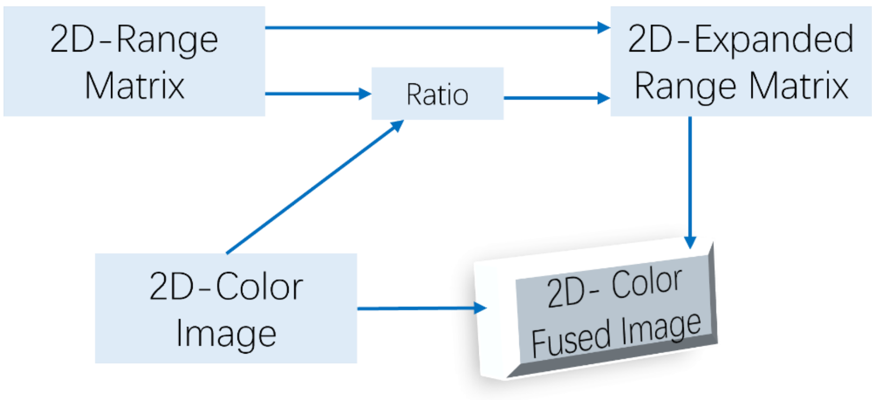


Figure 2.2 The Traditional Fusion Process from 2-D Image to 3-D Scale. [5]

2.2.2 Sonification Range Applications and Devices

Some researchers have developed a device by sonification of range detection to measure the 3-dimensional space as an ETA system [13] (figure 2.3 and table 2.2). Users can scan the space by wearing a device which could shoot a laser beam. The laser beam hits an object and generate a sound with certain pitch. The auditory signal represents the range.

Figure 2.3 The System Structure with Laser Range Finder [13]

Table 2.2 System Parameters Relevant to Both Mappings [13]

Range Readings per Second	8
Laser Range Finder Accuracy	+/- 10cm (the value may be slightly more)
Headphone Volume	Adjustable
Sound Note Duration	90ms - gap of silence between 2 sounds
Sound Note Amplitude (Volume)	Constant value of 80 (max amplitude is 127)

2.2.3 3D-map Guide Mobile Applications and Devices

Smart phones are also proper platforms for the visual handicapped people since the performance of them have dramatically increased in recent years. Many researchers choose phones as a highly integrated and mature platform to develop camera-based devices to detect environments [2]. Mainly they focus on the open spaces where the aerial obstacles located over the ground (figure 2.4).

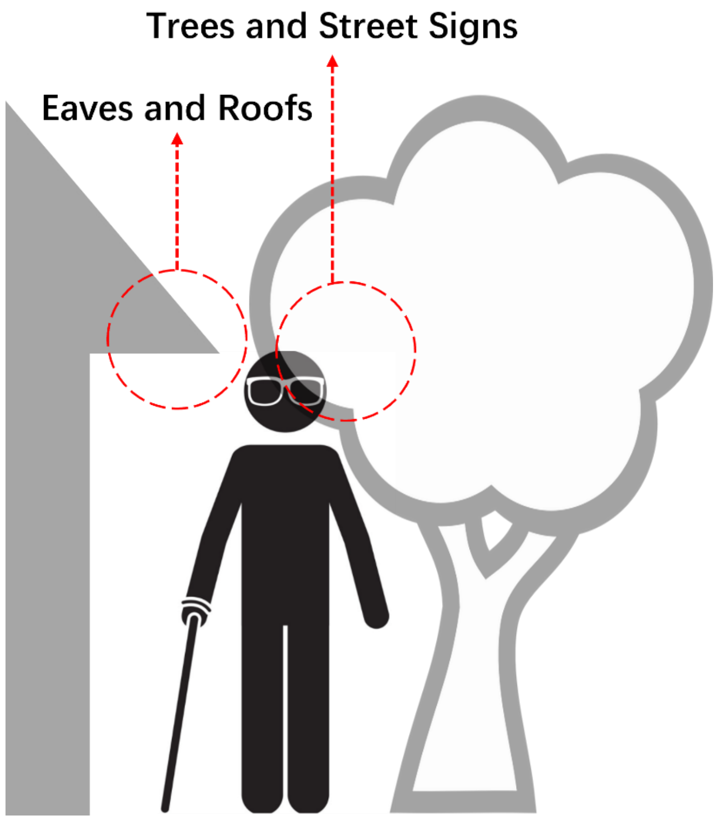


Figure 2.4 Aerial Obstacles Located Over the Ground

The obtained 3-D data are compressed and then linearized for detecting the incoming obstacles. Potential obstacles are tracked to gather and process enough evidence to warn the visually impaired people only when an actual obstacle is detected.

However, the current refresh rate is limited by the smart phone itself. The calculation performance of the processor is relatively low compared to the blooming camera resolution. Most of the system resource will consumed by the camera module. The refresh frame will significantly become low in order to sync with the camera frame for most models of smart phones.

2.2.4 Ultrasonic Echolocation System

The Sonic Eye provides a minimally processed input which, while initially challenging to utilize, has the capacity to be much more informative and integrate better with the innate human auditory sense. The echo sounds are captured by 2 ultrasonic microphones bilaterally installed, each installed inside an artificial pinna, also modeled after bat pinnae to produce direction-dependent spectral cues. The relatively raw echoes contain not only distance information but horizontal location information and also vertical location information (from the pinnae), as well as texture, geometric, and material cues. Behavioral testing suggests that novice users can quickly determine the laterality and distance of obstacles and objects, and with experience can also judge elevation (figure2.5). Thus, the Sonic Eye can become an assistive mobility device [14].

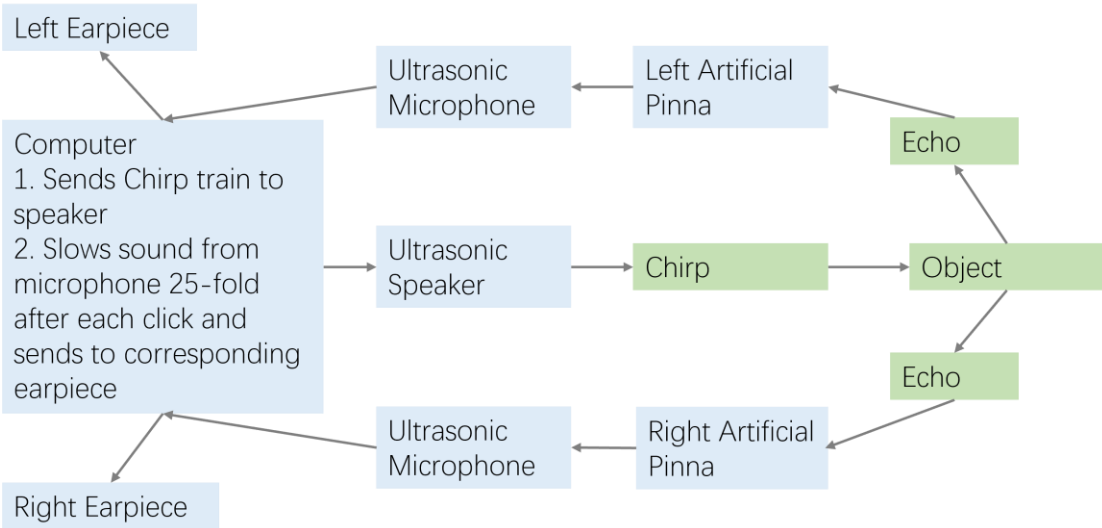


Figure 2.5 Diagram of Components and Information Flow of Ultrasonic Echolocation (the diagram is a modified version of the one in [14])

2.2.5 Indoor Mapping Matrix System

Some of the indoor navigation system establish multiple pivot points as the initial point to build the map. According to the topography map, the computer will generate a routine to the destination and it will send voice message to the user as the assistive information of the movement (figure 2.6).

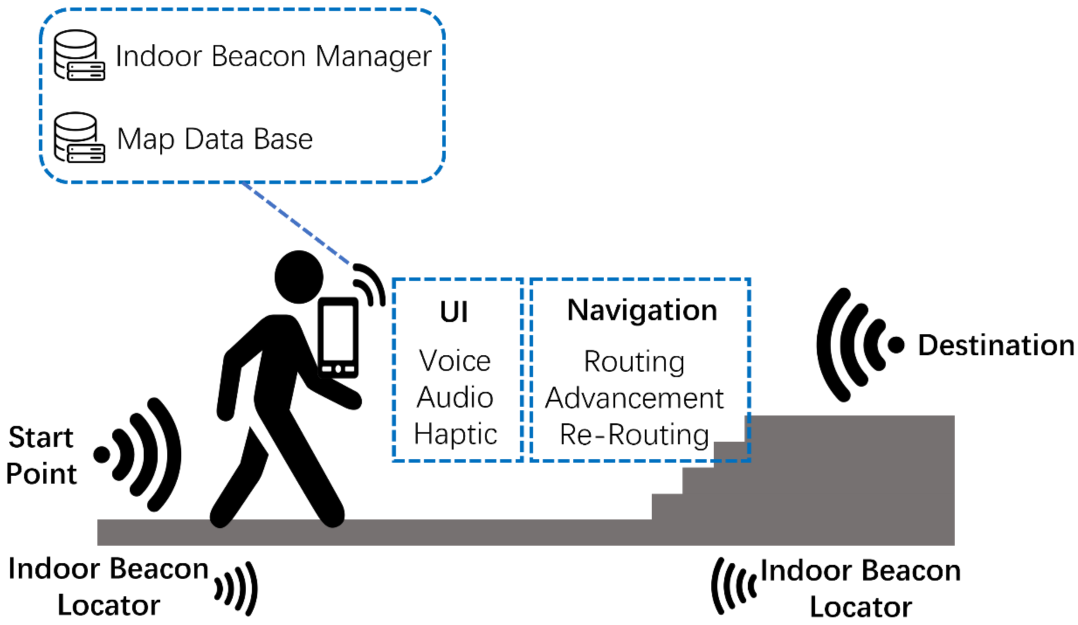


Figure 2.6 Typical Indoor ETA System Structure [14]

2.2.6 Navigation System Based on Haptic Interaction

Many designs chose haptic methods as the interaction method. A typical design uses vibrators as the indicator to tell the user whether he/she should turn left or right [15].

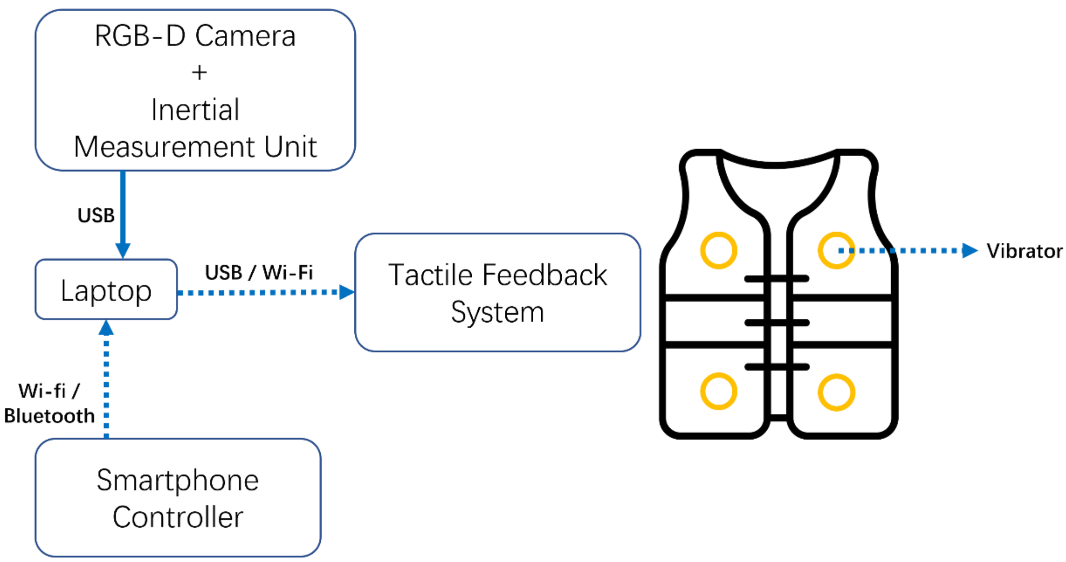


Figure 2.7 An Integrated Camera + IMU Sensor, A Smartphone User Interface, and Tactile Interface System [15]

This is a very typical structure of many popular system (figure 2.7). The system is not integrated well with many components still being placed apart. As we will discuss in chapter 3, the complexity of the system will cause an obvious inefficiency of daily use. One of the potential problems is the wires across the whole body. Our users are weak at the visual sensory, which means the user may accidentally twine the wires and cables at any time, especially in the situation of walking through a crowd or getting into a narrow space. It may also increase the difficulty of wearing it on and off by the visual handicapped person, which is an important aspect of improving the quality of daily life.

Another problem is the sustainability of the system. It is hard to image that visual impairment people have to carry all the devices including a laptop with many cables walking for a long time. The total weight of such kind of system are apparently not user-friendly,

What is more, the example design and many other similar designs are weak at information converting or interaction. There are only 2 dimensions of information for the user. Quantify should be take into consideration more rather than just qualitative with 2 kinds of signal. This should be the key issue to improve in our research.

2.3 The Advantages and Disadvantages of Conventional Methods

After researching various ETA systems and comparing them together of these conventional methods, there is a table shows the disadvantages and advantages of several typical ETA design (Table 2.3).

Table 2.3 The disadvantages and advantages of several typical ETA design [12]

System	Functionality	Interface and Components	Advantages and Disadvantages	Results
Virtual Acoustic Space (1999)	Vision substitute; orientation by constructing an acoustic perception of the environment	Stereo acoustic: two digital cameras embedded in eye glasses; headphones; portable computer	A: Portable, reduced form factor D: Blocks user's hearing; not tested in real environments	Six visually impaired and six sighted, showed >75% of object and distance detection
Tyflos (2008)	Obstacle detection by vibrations across chest through a vibrating vest	Vest with a 4 × 4 vibrator array; two digital cameras; laptop	A: Does not block hearing; detects obstacles at various height levels D: Needs more tests on real users	No experiments with visually impaired users
Kindetect (2012)	Person and obstacle detection by acoustic feedback	Acoustic depth sensor; computer	A: Easy to use; can detect head-level obstacles D: Blocks hearing; indoor operation only	Four blindfolded users detected obstacles along an indoor path
Vibratory Belt (2013)	Obstacle detection by vibrations across waistline through a vibratory belt	Belt with three motors; embedded computer; Kinect sensor	A: Easy to use; does not block hearing; detects head-level obstacles D: Generally, indoor operation only	18 blindfolded users performed comparisons between the white cane and the vibratory belt, resulting in similar travel time along an indoor path
Virtual White Cane (2013)	Obstacle detection by vibration of a Smartphone	Android Smartphone coupled with a laser pointe	A: Easy to use; does not block hearing; detects head-level obstacles D: Needs more tests on real users	18 blindfolded users detected obstacles along an indoor path

Chapter 3. The Analysis of the Design

After discussing about the current situation of visually impaired people and several typical conventional devices of electronic travel assistant devices, the demands and needs for potential users become clear. Meanwhile, the disadvantages of conventional devices are figured out. This chapter will introduce and analyze the user needs and design purpose of our system.

3.1 The Design Purpose

We are proposing a new device which can detect surrounding obstacles while provide enough initiative for the visually impaired people. Mostly, the current methods for the users are focused on the information gathering and topography map generation. These designs are weak at interaction channels between the devices and the users. The transition between the actual surrounding and the movement arrangement in mind is vague and indirect. Some of the design helps the user to build a routine to their destination with a loss of flexibility, other designs provide highly limited information about the direction with only 2 dimensions, which is left or right. Users need an interaction which accord with the intuition and perception of the human mind, while the device should provide a richer direction hint. Here we are going to provide information of obstacles in 5 directions.

3.1.1 User Persona and the Design Goal

According to the discussion of chapter 2, the main persona of our design is the people who are visually impaired or partially effected by the visual difficulties, containing illness such as cataract, glaucoma, or other eye diseases.

The main purpose of our design is to develop a portable device which could provide the information of distance and obstacle’s layouts of the surroundings in a more user-friendly way by sound hints. The vital interaction logic of the design is to convert the surrounding information into sound information, then all the sound information could become specific prompts to let the user to understand the real situation of their environment. The users can easily ‘decode those sound prompts to build a precise understanding of the surrounding, build an obstacle image in their mind as a normal person do. The ultimate objective of our design is to convert all the visual information to sound information in a very intuitive way, which may improve the efficiency of the data converting and help the user to have a much easier experience when they are moving.

By doing this, the user will finally gain an ability that they can replace their visual ability by hearing ability to some extent. Therefore, moving autonomously will no longer be a big problem for the visual handicapped people, especially there are no care workers nearby. The scale of their activity range would significantly expand; thus, the quality of their daily life will reach a new high level.

3.1.2 User Scenario

Considering all the demerits and merits of various designs we mentioned before, we mainly focus on the simplified indoor user scenario (figure 3.1). The reason is that most of the precise movement needs from the visually impaired people happen inside a building (moving from a room to another room, go up/downstairs, get into a building, moving across a big room or an area in a certain place, etc.). When the situation comes to outdoor, the movement needs are mainly relied on vehicles or other people’s transport service. Because the long-range travels for the blind people are severely dangerous and the time cost is dramatically high. Considering the movement speed is considerably low, it is almost impossible for them to travel alone by their foot or other traditional pedestrian methods.

The distance scale of outdoor movement is mainly in kilometers, the accuracy of the demand is relatively high. There are other services which can fit the demand well, such as GPS system. No matter who controls the transport system (other normal people or the device), this is just about a question of ‘95% of the travel distance’. In most of the cases this is not the main issue of our target users. The main issue is about ‘the last 10 meters.’ For example, a visual impaired person can achieve a travel from their home to the hospital front door by a taxi, however, it is impossible for them to precisely move to the right consulting room from the gate by still using the same navigation system of the taxi. In short, the measuring scale of the movement decreases from kilometers to meters, or even centimeters. This is the main design goal of our system.

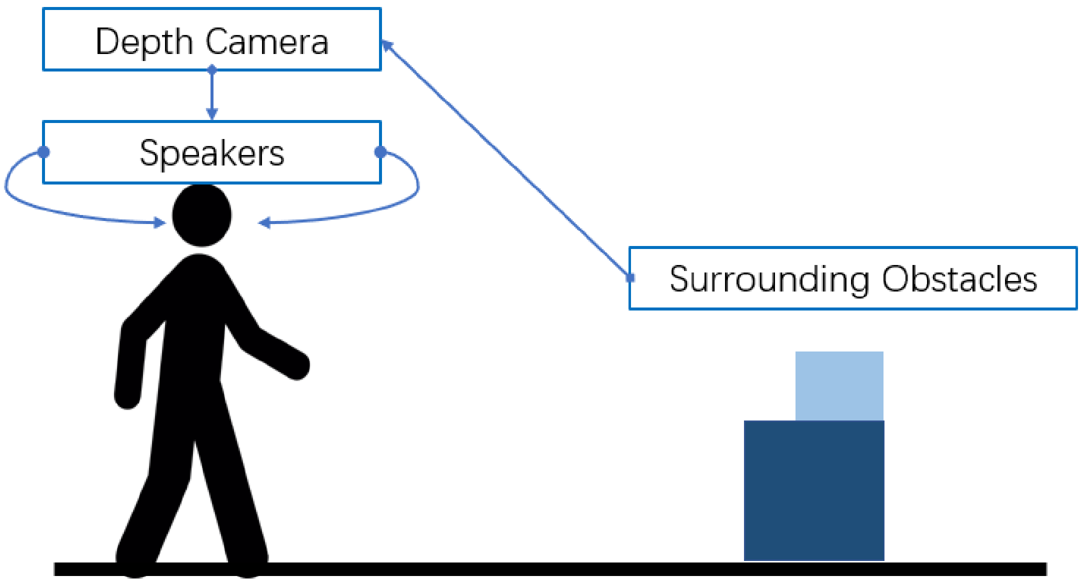


Figure 3.1 The Simplified User Scenario

3.2 The Innovation of the Design

3.2.1 The Function Innovation

According to our research and the discussion in chapter 2, it is not hard to let the user get to know the basic moving direction. What we want to do is make the level of accuracy and the information richness reach a new high level but still keep the information available to understand and utilize. The conventional methods are mainly just having 2 raw dimensions which is only right or left. It is hard for the user to quantify the angle when they need to make a turn.

We will divide the view angle of the front into 5 directions. The central front, slight right/left, left/right. This will provide a more precise information of the surrounding.

We may have more complex angles in future, but it should be done after a feasibility argumentation since too many angles may cause the misunderstanding and make the inefficiency of the system.

3.2.2 The Physical Structure Innovation

Most of the visual assistant devices contain multiple sub-systems and many other hardware. The integrity of those designs is relatively low. In many cases, the users must wear a vast, carry a computer, or even a whole computer with a carry bag. Both the size and the weight are not user-friendly enough for visual impaired people to equip by themselves. If the design is not a part of their daily life, the possibility of ignoring using the design will be higher than usual. This is not the original purpose of our design. What we are aiming for is to achieve a balance between the cost and the conveniency of the design. Just like the balance between the shortsightedness and the glasses, the ultimate goal of our operating cost will keep under a reasonable level, both economically and physically.

Therefore, the physical size of the device should not be too complex, and when the user wants to bring it on, the operation and the install should not be an issue of hesitation for them. The device should be integrated as a complete device and the weight should be limited (see figure 3.2).

3.3 The User Experience Design

We design the user experience flow as the chart. The system will gather visual information via dual depth cameras to generate depth images. Those images with depth data will be send to the processor, which is a PC, to execute following data processing steps. The PC converts the image information into values for controlling speakers, the speaker, as the final interaction component, will lead the user with its designated sound patterns. The whole experience is to transfer the visual information that visually impaired cannot utilize into sound information.

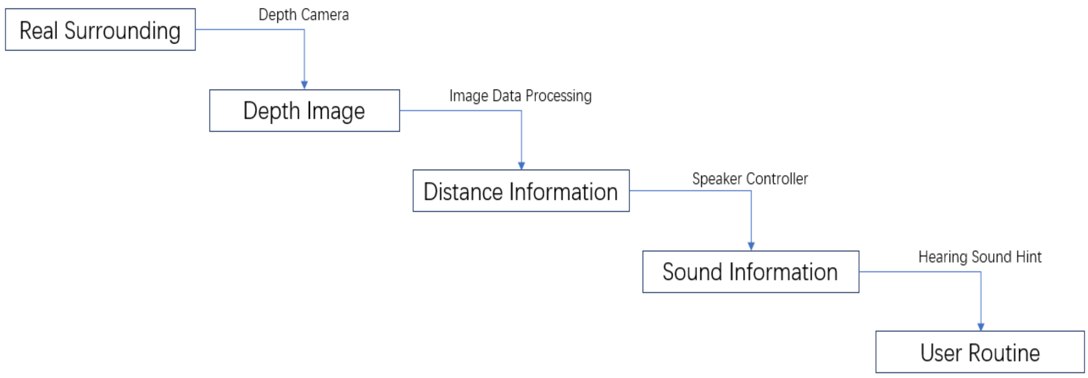


Figure 3.2 The User Experience Flow Chart

After the user switch on the device, the D435 stereo camera gather all the surrounding depth information into a depth image with distance value. Each of the image will be processed in order to find which is the nearest point currently. Then the location of the nearest point will be sent to the microchip to convert it to the signal of controlling the related speaker. The speaker related to the nearest direction will beep with certain sound of patterns which have different meanings of distance information. The user receives the sound by hearing, and they will understand the obstacles' layout of the surroundings. By executing the loop again and again, the user will refresh the information for them. Eventually they can find a safe way of the current routine.

Chapter 4. Enabling Technologies for our Proposal

4.1 The Depth Camera

4.1.1 Depth Sensor Mechanism

There are 2 main elements for a depth camera system to operate. An IR (Infra-Red) projector and an IR camera. The IR projector projects a certain pattern light dots of IR on the target object's surface with a preset density. Those dots cannot be captured by eyes since it is out of visible light range.

However, the IR camera can detect those IR dots on the surface. An IR camera is basically and essentially the same as a common RGB camera, with an ability of capturing images in the Infra-Red color range.

The camera sends its video feed stream of the distorted dot pattern into the processor of the depth sensors. Then the processor will work out each value of the dot based on the displacement of the dots. Objects nearby have a denser dot pattern while the far objects will have a sparser placement of the surface.

The stereo vision follows the same mechanism as the human eyes. Two cameras are installed in pair at a given distance. The distance is called as baseline. Mainly the two cameras will be triggered at a same time. By comparing two images from the two cameras at each time unit, point-to-point correspondences are calculated between the two images of the object. Given the roto-translation between the cameras and the lens distortion parameters, the point-to-point correspondences can be triangulated to obtain the respective location in the 3D space (figure 4.1 and 4.2) [16].

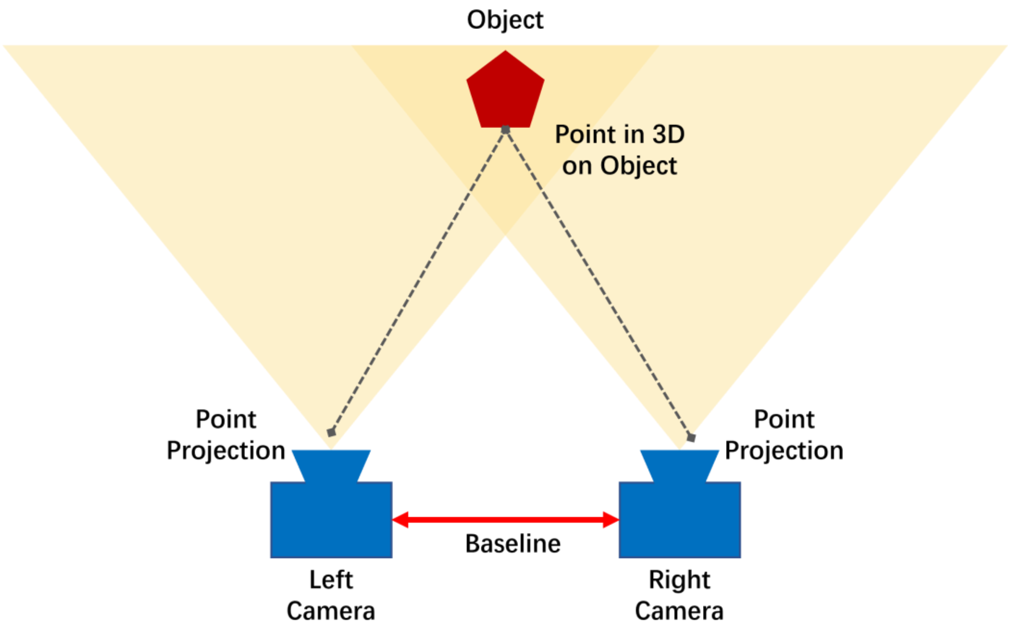


Figure 4.1 The Stereo Camera System [16]

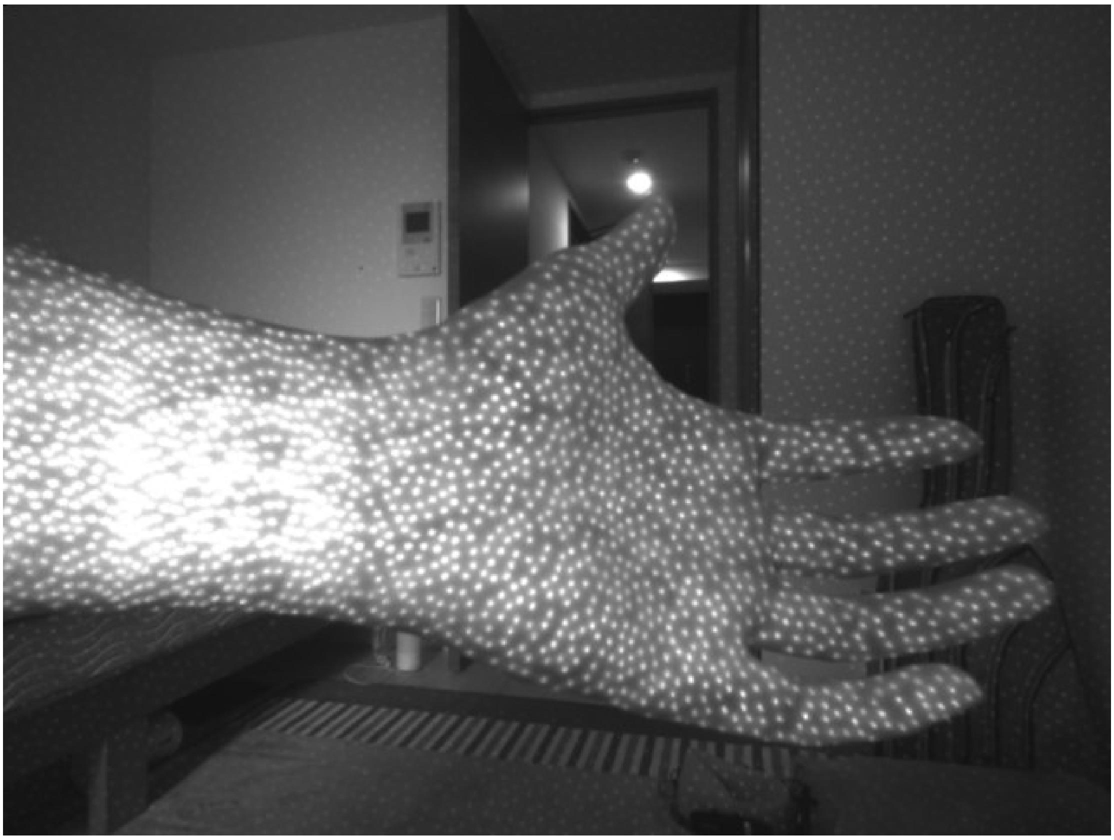


Figure 4.2 The IR Spot from the Depth Camera

Here we are using the Intel Realsense D435 depth camera as our image input. The Intel® RealSense™ D435 offers the widest field of view of all our cameras, along with a global shutter on the depth sensor that is ideal for fast moving applications.

4.1.2 The Depth Image Processing

Figure 4.3 shows the Depth Image rendered by OpenCV. The red represents near distance and the blue represents it is far away from the camera. The yellow means it is a modest range of distance. In default, the resolution of the image is up to 1280 x 720 with 30 fps. The distance data of each pixel location in the real world are equivalent to the pixel color value. That means, by processing the color data of the image, we could get the distance amount for each of the spot shown in the image.

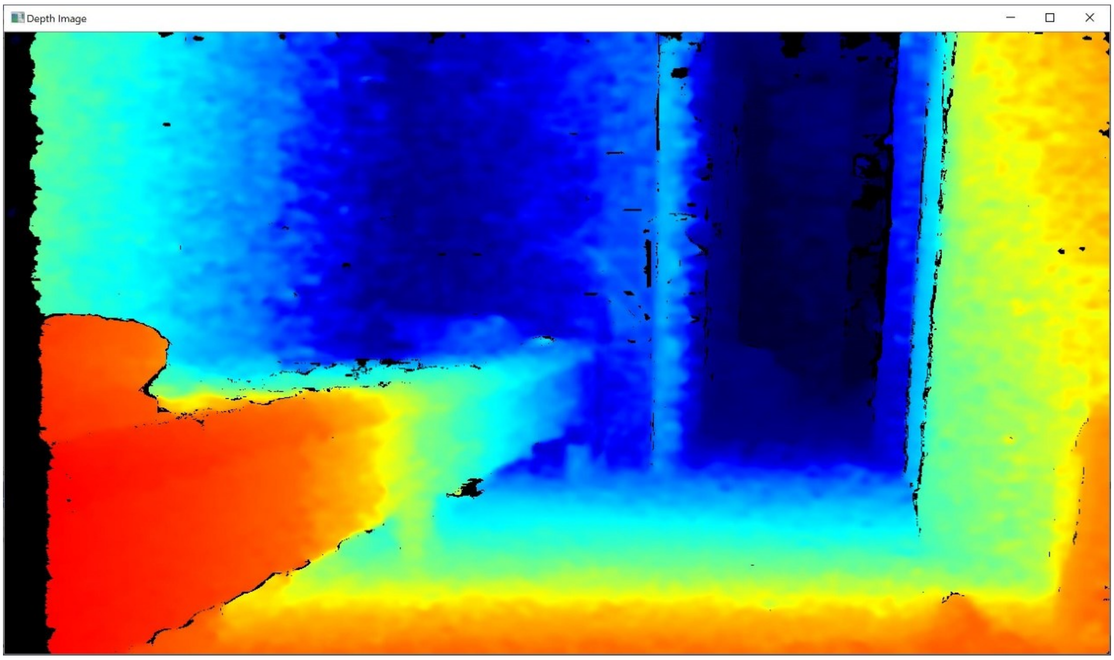


Figure 4.3 The Depth Image in RGB

4.2 The Sound Prompt for the Visual Impairment Person

Although the hearing ability of visually impaired are relatively as the same as common people, they have a better performance on distinguishing the sound directions due to their skilled daily life experience [12]. We can build a mapping relation between speaker at different angle to the locating direction of obstacle in real life. This would be more intuitive and ocular for the visually handicapped person. The principle of the design about the interaction is as simple and clear as possible to avoid potential misunderstanding and confusing.

According to the experience of daily life, the pattern of sound prompts should follow the rules which are:

The more of the number of sound cycles per second, the closer more dangerous it is; The higher of the pitch of sound, the closer and more dangerous it is.

The sound prompts should not interfere users about hearing surrounding background sounds. Visual impairments rely on these acoustics information to avoid hazard such as vehicle engine noise or voice signal sounds. The mapping layout is shown in figure 4.4

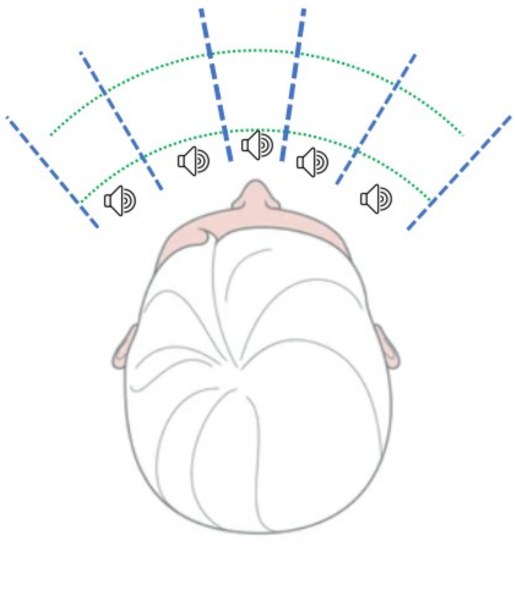


Figure 4.4 The Angle of the Direction Distribute for User Experience Design

Chapter 5. The Development of the System

5.1 The System Structure Brief

The structure of the whole system is shown below (figure 5.1). The whole system contains 2 parts, the software, and the hardware. Considering the system compatibility and stability, all the software parts are installed and processed inside the PC. The hardware components are implemented in one entirety to avoid long wires and complex plugins.

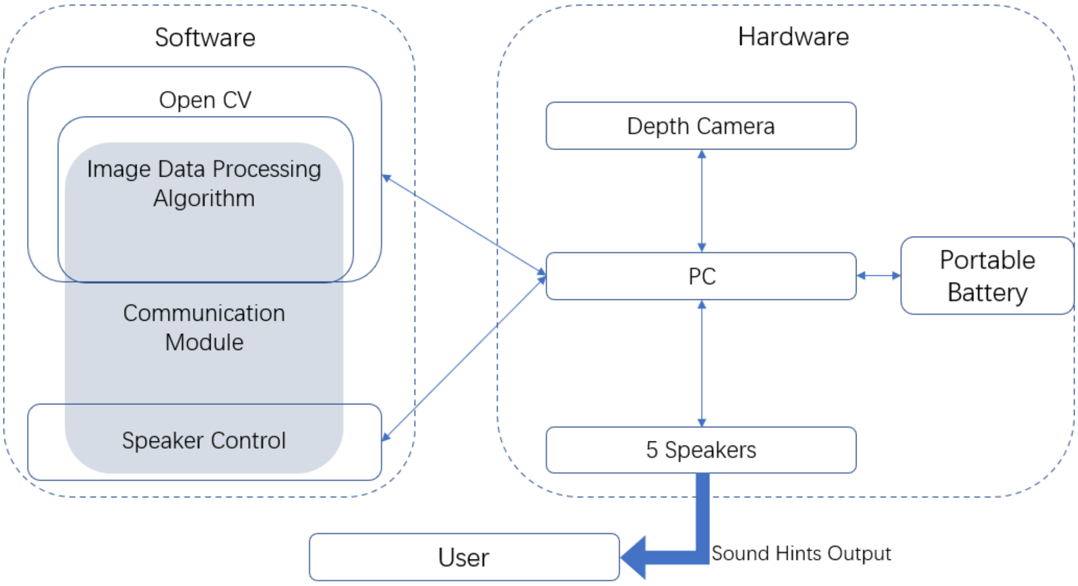


Figure 5.1 The System Structure

5.2 Hardware System

The prototype contains mainly 5 parts. The depth camera, the Microchip with 5 speakers, the PC stick, the portable battery, and the major physical structure (Hat with band). The total weight of the system is about 1.5 kg, which is user-friendly for long time travel (figure 5.2 and 5.3).

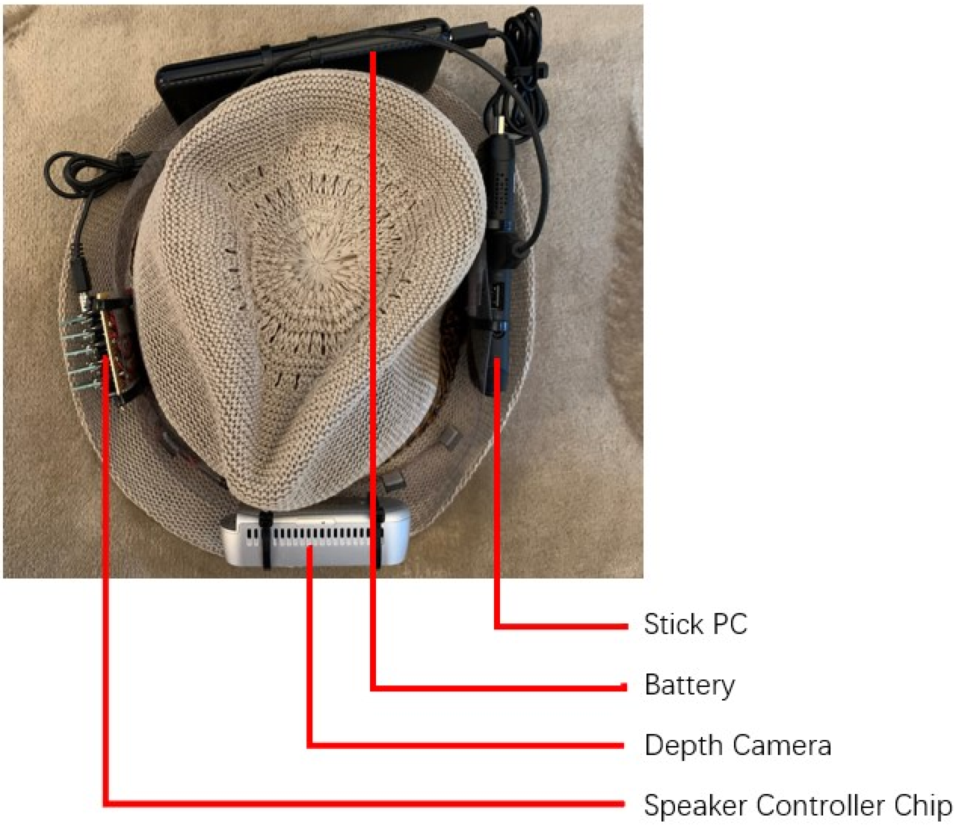


Figure 5.2 The Picture of the Prototype (Top view)

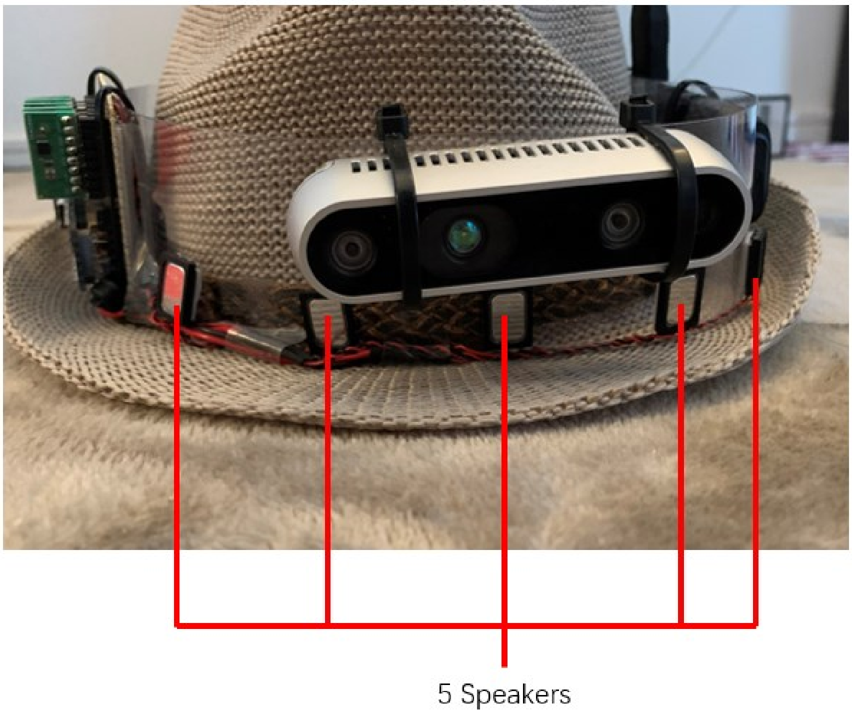


Figure 5.3 The Picture of the Prototype (Front view)

5.2.1 The Depth Camera

The camera (figure 5.4) is the only component gathering the information. The camera is placed in front of the center, and according to the spec, the front angle of view is 87° × 58°. The height of the camera is close to the position of eyes (about 8 cm above the eyes). The prototype is shown in figure 5.2, figure 5.3 and its spec is shown in table 5.1.

Table 5.1 The Main Spec of the Intel Realsense D435

Features	Use environment: Indoor/Outdoor	Ideal range: .3 m to 3 m	
Depth	Depth technology: Stereoscopic Minimum depth distance (Min-Z) at max resolution: ~28 cm	Depth Field of View (FOV): 87° × 58° Depth Accuracy: <2% at 2 m1	Depth frame rate: Up to 90 fps Depth output resolution: Up to 1280 × 720
Physical	Length × Depth × Height: 90 mm × 25 mm × 25 mm	Connectors: USB-C, 3.1 Gen 1	

Figure 5.4 The Image of the Depth Camera D435

5.2.2 The Microcomputer controller and 5 Speakers

5 speakers are evenly installed in the band which bind around the hat. The speakers can be controlled by the micro-chip with Pulse Width Modulation (PWM) signals. PWM compatible allows us to manipulate the pitch and the frequency. The chip PIC18f4550 supports direct program control when connecting to the PC via USB port, which means that we do not need to input the program code in advance. The code is running inside the PC system, and it is controlled through the connection wire. This was very helpful at the develop and test phase, we do not need to re-boost the code into the chip repeatedly when the code is modified.

5.2.3 The PC Stick

Differing from conventional systems which are using a laptop as the calculation processor, we choose a PC stick made by ASUS. This is a complete PC with the OS of Windows 10, 4GB of RAM. It supports all the develop environments and software with a significant small size apparently. The weight is only about 330 g. 2 USB ports can be expanded by USB hubs but still can fulfill the data transportation between different components.

5.2.4 The Portable Battery

We choose a mobile battery as our power supply. The demand is to reach a balance of durability and weight. Currently we are using a 10000 mAh which can last about 4 hours of energy durability for the entire system. It is about 400g, in order to keep the balance of the head band, it is fixed at the rear side of the system.

5.2.5 The Hat Band

We build this prototype based on a regular hat as our main structure. The weight is relatively low, but it can hold all the components firmly. In our user experience design part, the design should be user-friendly which means it is easy for user to wear. The hat fits all the demands we designed before.

5.3 The Software System

5.3.1 The Depth Data Gathering

The depth information is collected by the 2 stereo depth cameras. The frame rate of the depth image will refresh at 30 Hz. During the development, we found that the PC performance will strongly influence the refresh rate of the camera. Even though, due to the camera is integrated with many mature and fully developed SDK, the synchronization between the images and the data process could still work fine.

5.3.2 The Depth Data Processing

5.3.2.1 Pre-Processing of the Data

To reflect 5 directions with 5 speakers, we divide the whole image vertically into 5 blocks evenly. For each block represents an angle of 87 °/ 5 ≈ 17.4°. From left side to the right side, we mark each block from block 1 to block 5 (figure 5.7).

According to the early-stage test, the default image resolution scale (1280 x 720) will not have a proper performance due to the horizontal view angle is too wide, many obstacles from the selling such as bulbs will interference the data processing.

During the actual using of the system, the user definitely will not keep the camera angle as a horizontal line. They will tend to move their neck upward and downward, and this will make the minimum detect length fluctuate (figure 5.5).

As showed in the figure (figure 5.5 and 5.6), when the vision is looking down, the minimum distance of the bottom line becomes closer. It is necessary to set a limited 'meaningful section of angle' to guarantee most of the information gathered by the camera is valid. In another word, we need to set a filter at the camera view.

The common human eyes field of view is approximately 135 degrees. Comparing to the depth camera we are using, it is not necessary to cut off the angle of the view, and we need to remain the default maximum angle of the camera which is 87 degree.

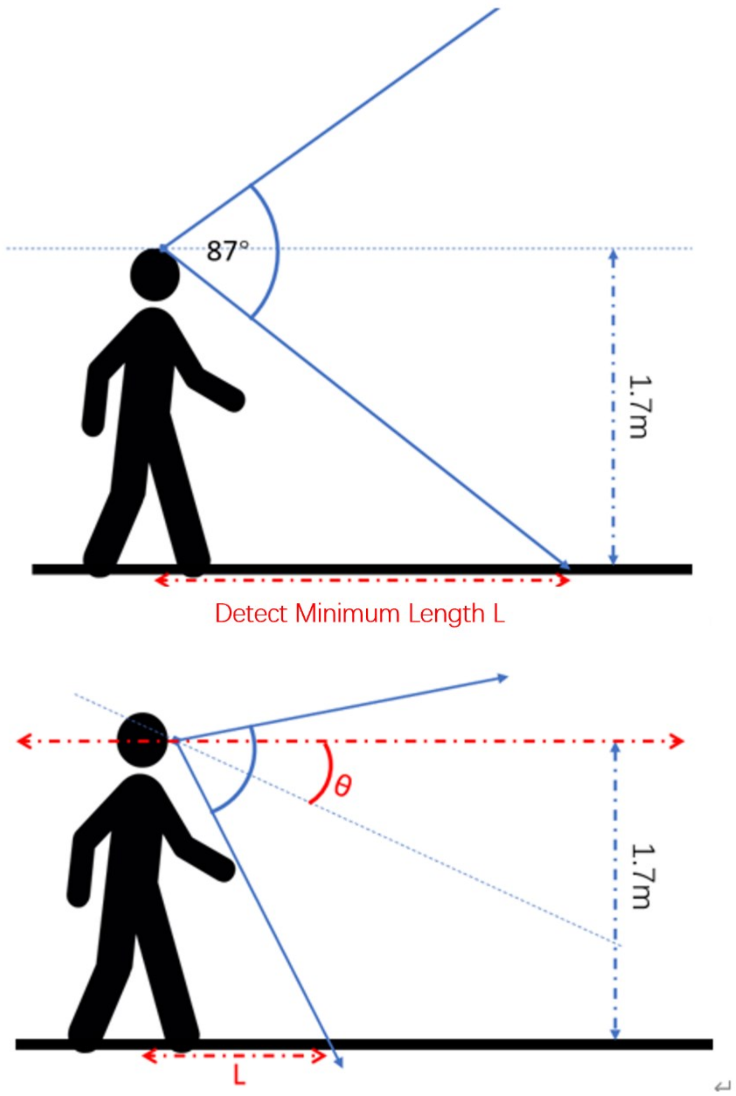


Figure 5.5 Simulation about the Influence of Vision Angle Change

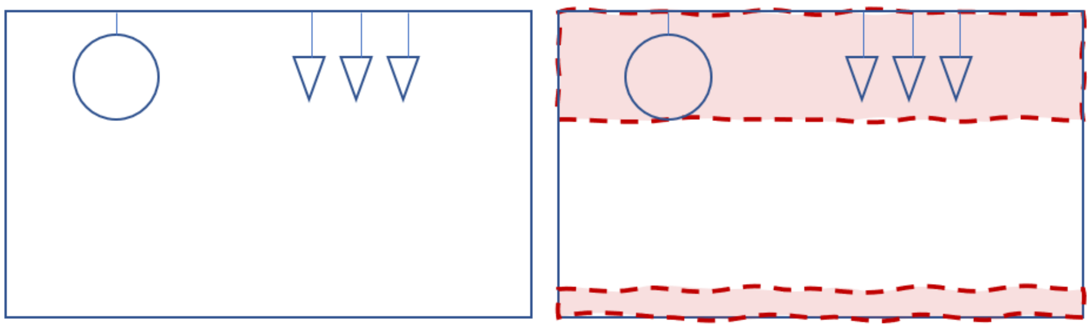


Figure 5.6 The Filter Shape of the Depth Image

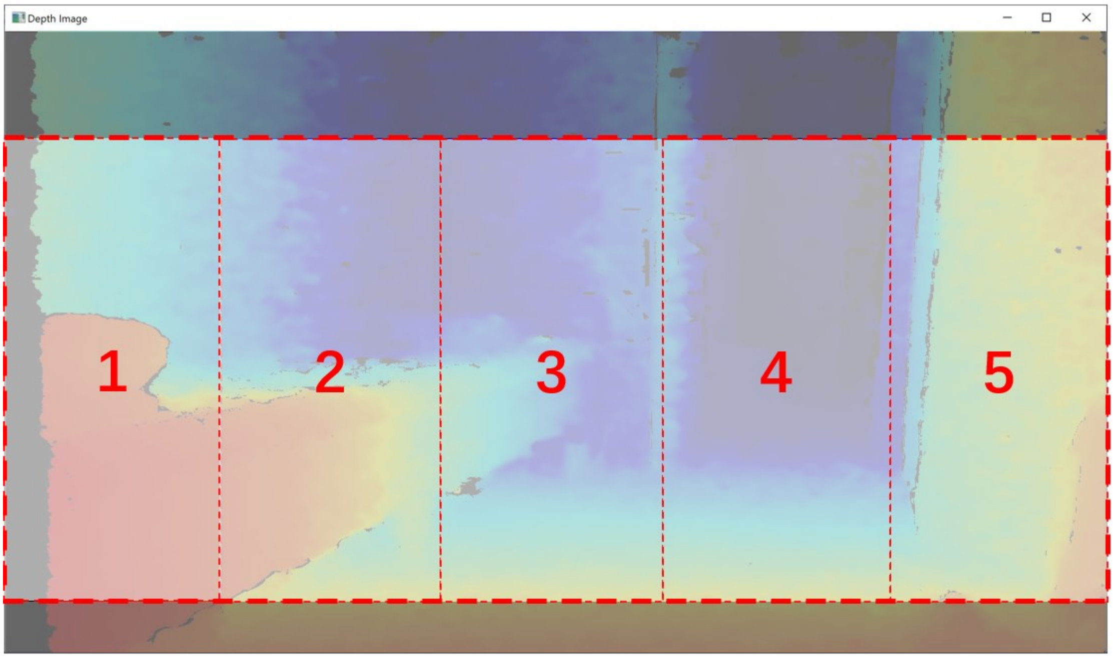


Figure 5.7 The Depth Image Divide into 5 Blocks Evenly

5.3.2.2 Processing of the Data (Algorithm)

Briefing

The main purpose of our system is to distinguish the distance in 5 directions and to avoid any potential collision of obstacles. The minimum distance value is the key number we need to focus on.

First, we get the minimum number of each block, then compare the least distance of all the 5 blocks and mark the index of the block.

The OpenCV provides a possibility that we can transform the image to meaningful values. OpenCV supports multiple programming language, and the compatibility ensures that we can process the data from the camera with an efficient and proper programming language. In our project, we choose C++ as our main coding language.

Algorithm

The principle of minimum value algorithm programing in One Block is:

1. Pick 2 elements (a, b), compare them. (say a > b)
2. Update the min variable by comparing (min, b)
3. Update the max variable by comparing (max, a)
4. Continue the comparison until visit all the elements in one frame of each block

The elements here refer to the distance value of each pixel, the sequence of the comparison stars from the left-up pixel to the right-down pixel one by one as the figure below (figure 5.8).

After getting 5 minimum distance value of each block, the algorithm will compare those 5 values and pick the most minimum value and its block index (figure 5.9).

Case Determination and Sound Pattern

Based on the user experiment design we discussed at preceding part of the thesis, there are 2 thresholds value for the system to distinguish the current distance situation into 3 kinds of cases. The safe distance, the cation distance and the dangerous distance. These 3 cases are mapping to different sound patterns (table 5.2).

Table 5.2 Cases Represents Different Situations

	Pitch	Frequency	Case
Dangerous Distance **(<0.5 m)**	HIGH	FAST	A
Caution Distance **(0.5m≤distance≤ 1m)**	LOW	LOW	B
Safe Distance **(>1m)**	MUTE	MUTE	C

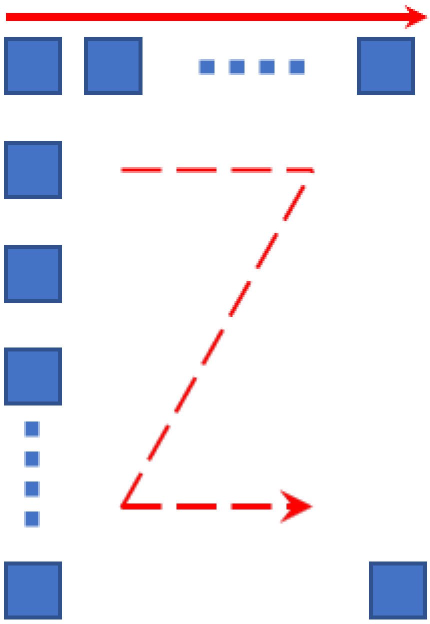


Figure 5.8 The Sequence of the Pixel Distance Value Comparison

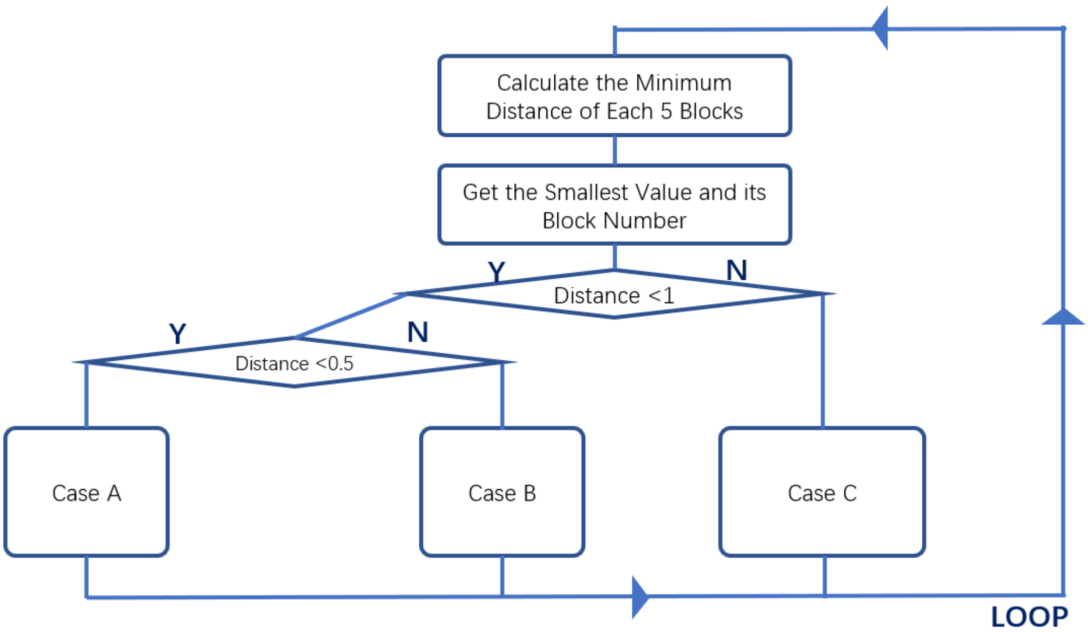


Figure 5.9 The Process Chart of Determine the Case

Algorithm Selection

There are 2 methods of getting the value of minimum distance, one is the average distance, and another is the current method of picking minimum value.

During the test, the average value of each block will not be taken into consideration. The reason for it is the 'spike shape' and the anomalous value.

If there is an obstacle in a shape of a spike, but the rest area of the block is still safe and far away, the average value of the distance of this block may still remain as a safe value (which is a big number relatively.) The result of the index of having minimum average distance will be one of the other blocks. This would become a dangerous situation; the user will be misled by the wrong information and an accidental collision might happen. Therefore, after the test of the engineering verification phase, we decide to use the minimum algorithm as the data processing method.

Output of the Data Processing

According to the table of Case Determination and Sound Pattern, there are 2 kinds of parameters should be transferred from image data processing algorithm module to the speaker control module of each case.

THE LEVEL OF THE DISTANCE and THE INDEX OF THE BLOCK. That means, the communication package between the 2 software modules contains 2 integers (figure 5.10).

The data package can be unpacked into meaningful information for the control module. The first digit refers to the index of the nearest block from 1 to 5, and it also refers to the speaker index which is going to beep. The second digit refers to the distance cases we set in the user experience design part from 0 to 2 by integers.

```
1block, min distance is 0.264
2block, min distance is 0.255
3block, min distance is 1.32
4block, min distance is 0.893
minimum distance of 5 blocks is 0.255, the index of minimum block is 3
0,0,
0block, min distance is 0.281
1block, min distance is 0.26
2block, min distance is 0.255
3block, min distance is 1.322
4block, min distance is 0.901
minimum distance of 5 blocks is 0.255, the index of minimum block is 3
minimum distance is less than 1m
3,2,
0block, min distance is 0.256
1block, min distance is 0.257
2block, min distance is 0.255
3block, min distance is 1.322
4block, min distance is 0.896
minimum distance of 5 blocks is 0.255, the index of minimum block is 3
0,0,
0block, min distance is 0.256
1block, min distance is 0.261
2block, min distance is 0.255
3block, min distance is 1.329
4block, min distance is 0.717
minimum distance of 5 blocks is 0.255, the index of minimum block is 3
minimum distance is less than 1m
3,2,
```

Figure 5.10 Data of the Data Process Program in Console

5.4 The Speaker Control Module

The speakers are controlled by the microchip of PIC18f4550. This chip supports PWM control signal, all the 5 speakers are paralleled link to the chip board. This was very helpful at the develop and test phase, we do not need to re-boost the code into the chip repeatedly when the code is modified. The microchip is linked to the PC with micro-B USB port and also supplies the power at the same time.

The speakers' tone can be controlled by the PWM signal, the pitch of the tone can easily be heard. Currently we set the high pitch at 294 Hz and the low pitch is at 147 Hz.

5.5 The Communication between Programs

The OpenCV program for the camera and the data processing are based on C++ while the microchip is running by the Visual Basic language. We need to build a channel for the data transfer. The communication between 2 different programing language within 1 PC environment was the most difficult problem during the whole development.

Socket programming is a way of connecting to nodes or process on a network to communicate with each other. One socket(node) listens on a particular port at an IP, which is the sender part of the code C++. Meanwhile socket reaches out to the other to form a connection which is the receiver of the VB program (figure 5.11).

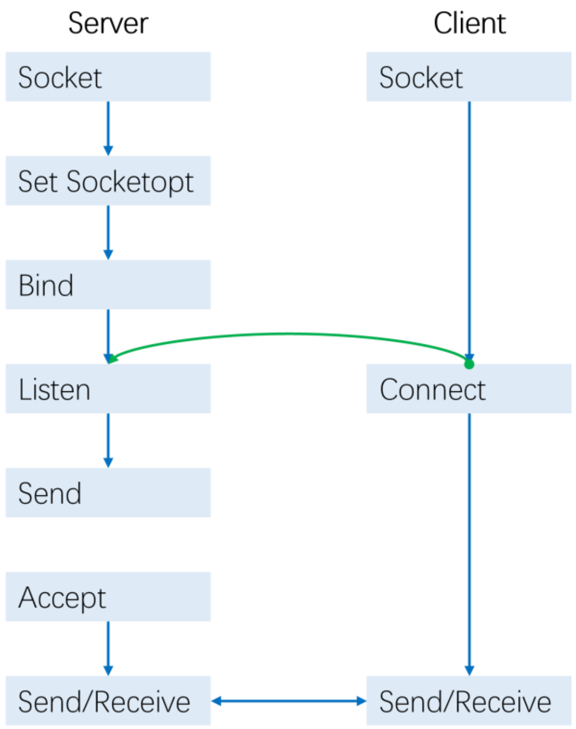


Figure 5.11 The Basic Process Chart of a Socket Structure

5.6 The User Interaction

As the article mentioned previously, the interaction between the user and the system is the sound information. We are aiming for converting visual information into acoustical information for the visual impairment people. The representation between the sound and the sight should follow the principle that is intuitive and no misunderstanding.

We build the mapping interaction as the figure below (figure 5.12).

The position of the speaker is accurately facing the angle of its related block area of the camera view. This can guarantee there is no confusing impacts on direction comprehension. Even though the user does not wear the device in a proper direction.

The angle of the sound source is still representing the obstacles information on that angle.

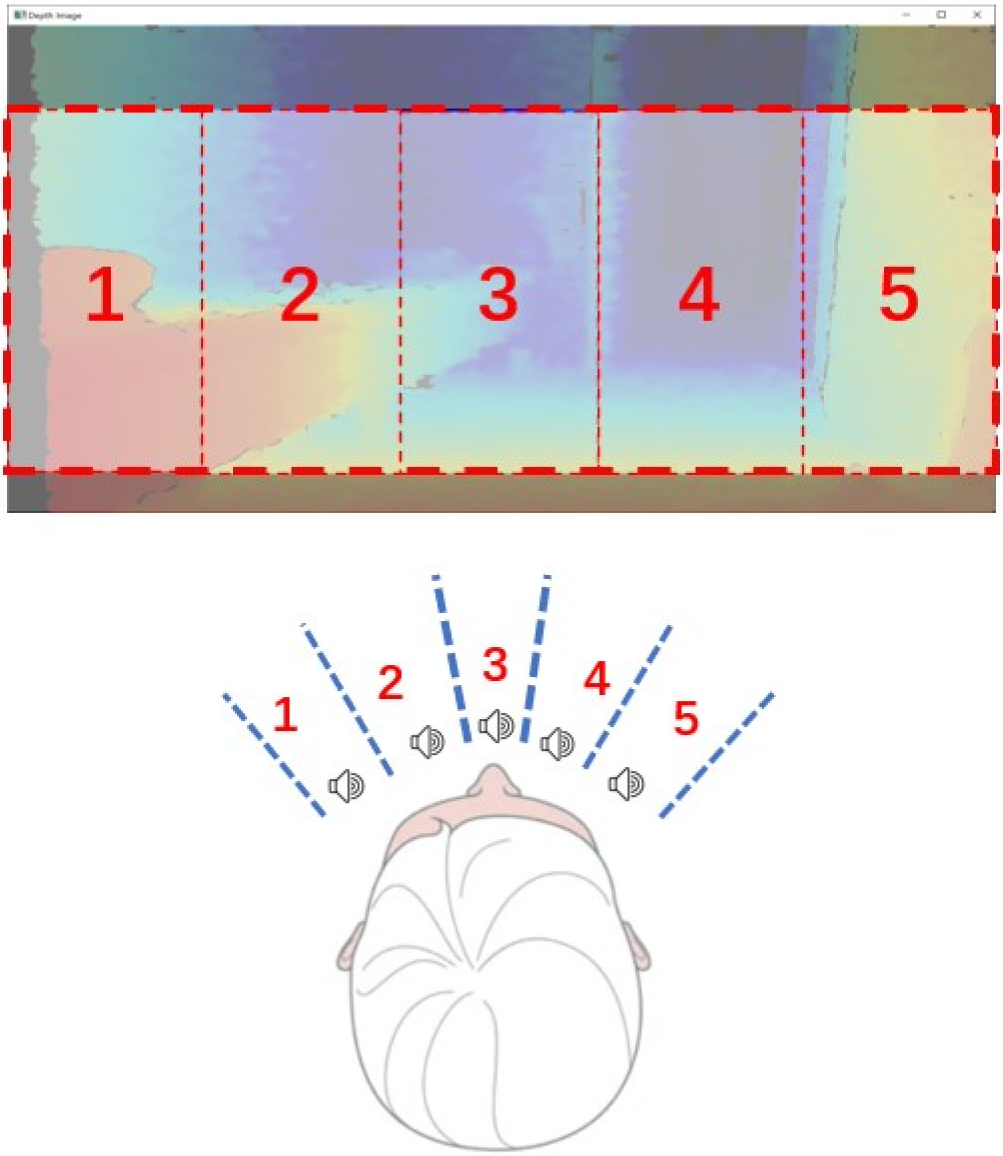


Figure 5.12 The Reflection Between the Block and the Angle

Chapter 6 The Verification of the System

6.1 Engineering Verification

During the development, we found that the performance heavily relies on the computational ability of the PC. The refresh rate of the system will influence the user experience and the system performance. We had to adjust the algorithm parameter and the coding method to establish the system as effective as we can. For example, the OpenCV RGB rendering should not be invoked when the device is on, but it is helpful when developing and debugging.

6.1.1 The Depth Camera Verification of Surface

According to the mechanism of depth camera, different surface will heavily influence the performance of the depth camera. The reflection rate of the IR will play a vital role in the whole depth camera system.

Although the camera spec shows that the resolution can up to 1280 X 720 and the frame rate can be up to 90 fps, there are still some ‘noise points’ which can interfere the distance calculation in some circumstances.

What is more, it is extremely important that we need to confirm performances of variable materials as many as possible. It will be a potential collision risk if there are some materials which the depth camera cannot detect very well. It will be somehow transparent in our system and that problem should be solved with an alternative way in advance.

To confirm the performance of the depth camera (here we are using Intel Realsense D435), we took pictures on different materials and check the image of each case.

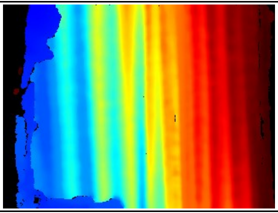
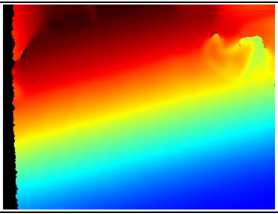
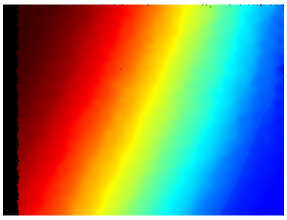
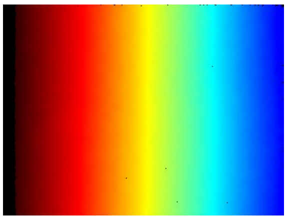
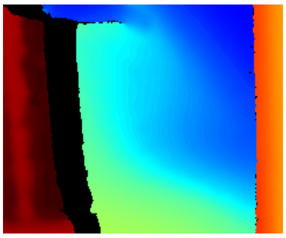
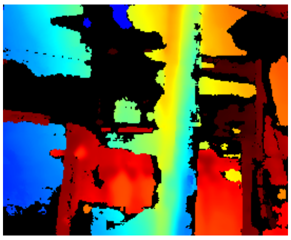
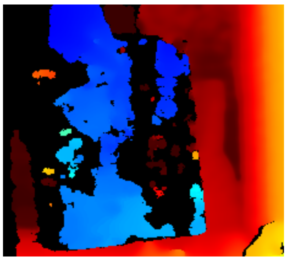
Result

The result shows that: for common indoor materials, such as woods, fabrics, metals, the depth camera can detect and measure the distance with a high efficiency. The edge can be clearly recognized and the distance value for each spot can also be measured. There are barely any noise point and the undetected edge zone which is black edge in the depth image are relatively small.

This proves that our design is verified for detecting tables, black boards, chairs and other indoor structures. Those materials contain most of the daily indoor cases.

However, due to the mechanism of the depth camera, the IR will penetrate through transparent materials such as glasses and plastics. The ratio of the IR reflection will just around 80% with an incident angle of 90 degree [17]. There will be some significant black spot which means the camera cannot detect the value of distance at that spot. What is more, the possibility of generating some noise point will get higher that other materials. Those noise point will influence the result of the minimum distance algorithm (table 6.1).

Table 6.1 Depth Camera Performance on Variable Material Surfaces

Material	Depth Image	Description
Fabric (Curtain)		Depth Camera works perfectly.
Fabric (Carpet)		Depth Camera works perfectly.
Concrete (Wall)		Depth Camera works perfectly.
Wood		Depth Camera works perfectly.
Metal		The black edge (vague area) is more obvious than other materials.
Glass		There might be some black area and the reflection ratio is relatively low.
Transparent Plastic Film		There might be some noise point, which may interfere the data process.

6.1.2 The Binaural Position Verification

The Experiment Design

We use a dummy head with 2 microphones placed at 2 ears separately. The 2 microphones will simulate 2 ears. Each of the ears will record the sound from the speakers individually to simulate the user's sense of hearing. The speaker will beep one by one as the device working, and the 2 microphones will record the sound level with 5 times. This experiment is the verification to confirm that the audio level gap is obvious enough for user to distinguish which speaker is beeping.

We need to find out that whether the two microphones can tell the difference. We need to measure the exact scale of the sound level difference and quantify the gap value to prove the verification for the engineering is passed.

We will start the test from the left side, which is the speaker of index 1(figure 6.1).

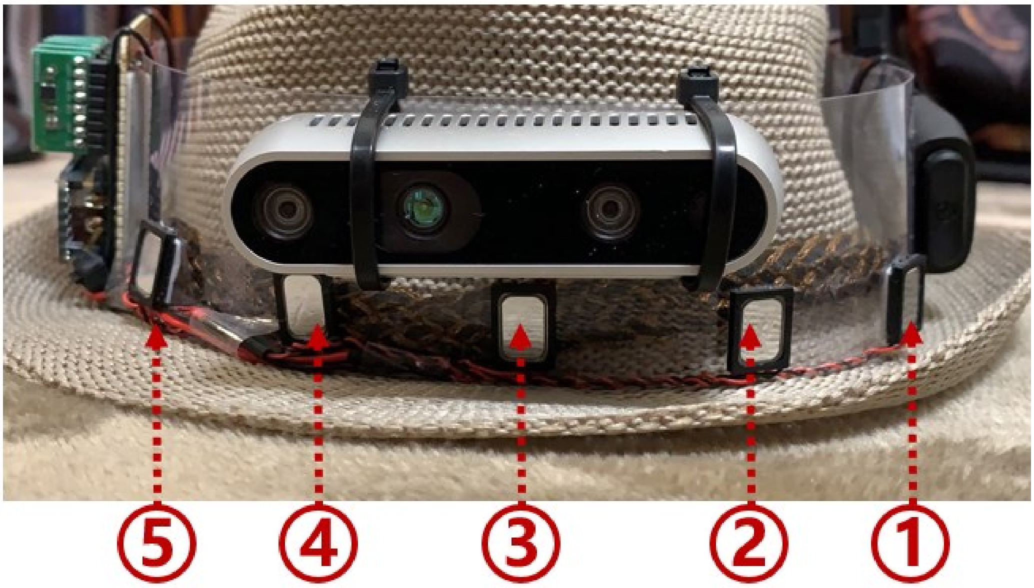


Figure 6.1 the Speaker Index Layout

The Result

To be brief, the sound level differences of each speaker recorded by two microphones can be obviously discovered. The maximum amplitude of the speaker is about 0.58, which is the situation of nearest speaker beeping (the most left speaker sound recorded by the left microphone). The minimum amplitude is approximately 0.1 (the most left speaker sound recorded by the right microphone).

This result proved that the engineering verification is qualified, and the system design is ready for the next step development and experiments (figure 6.2 table 6.2 and figure 6.3.1 to 6.3.5)..

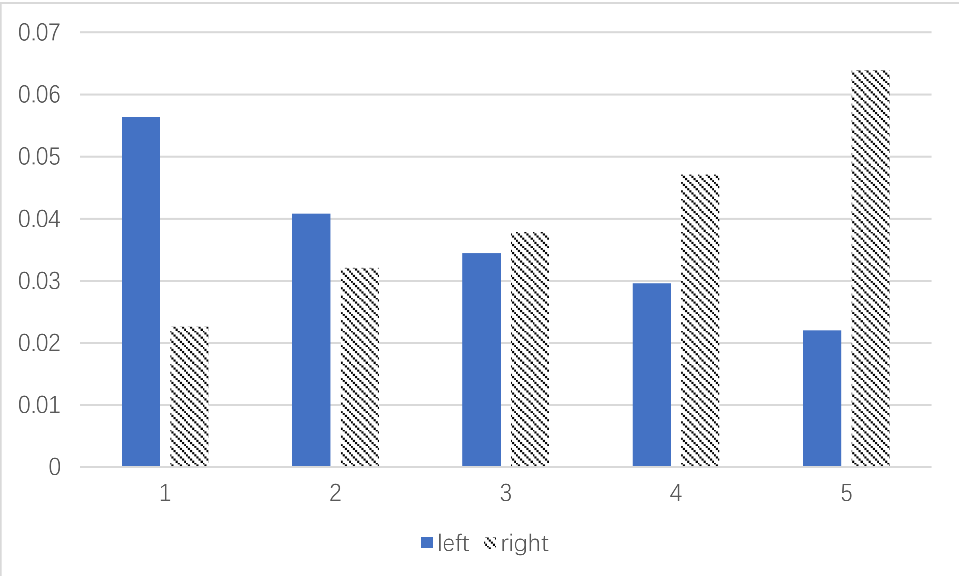


Figure 6.2 The Wave Amplitude of 5 Speakers in 2 Microphones

Table 6.2 The Sound Value of Each Speaker in Dual-Channel

	1	***2***	***3***	***4***	***5***
left	**0.0564**	**0.0408**	**0.0344**	**0.0296**	**0.022**
right	**0.0226**	**0.0321**	**0.0378**	**0.0471**	**0.0639**

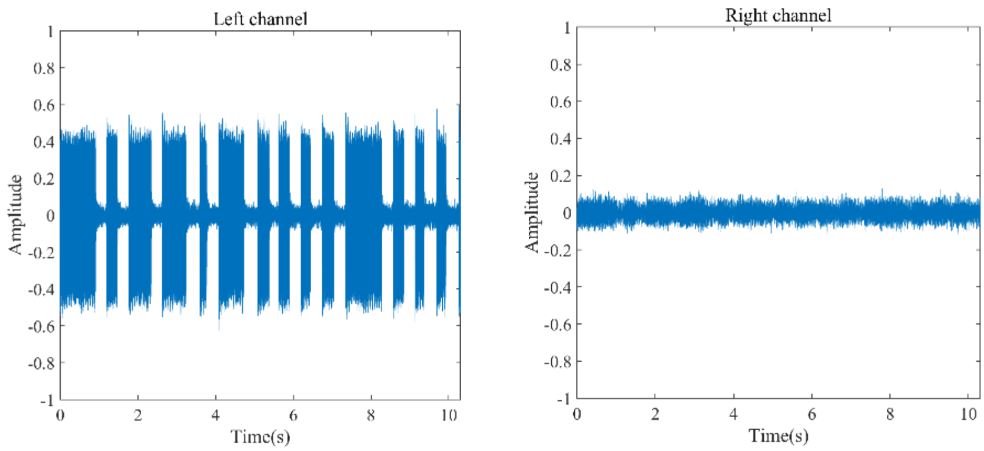


Figure 6.3.1 The Wave Amplitude in Direction of Speaker 1

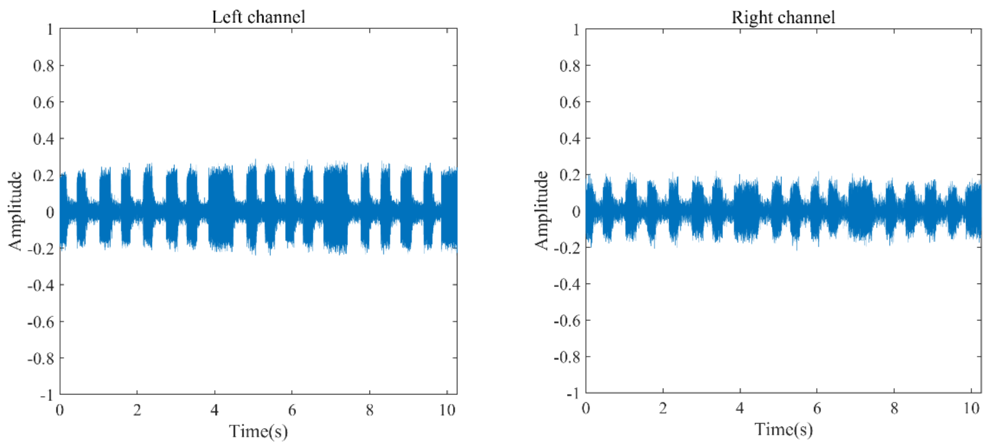


Figure 6.3.2 The Wave Amplitude in Direction of Speaker 2

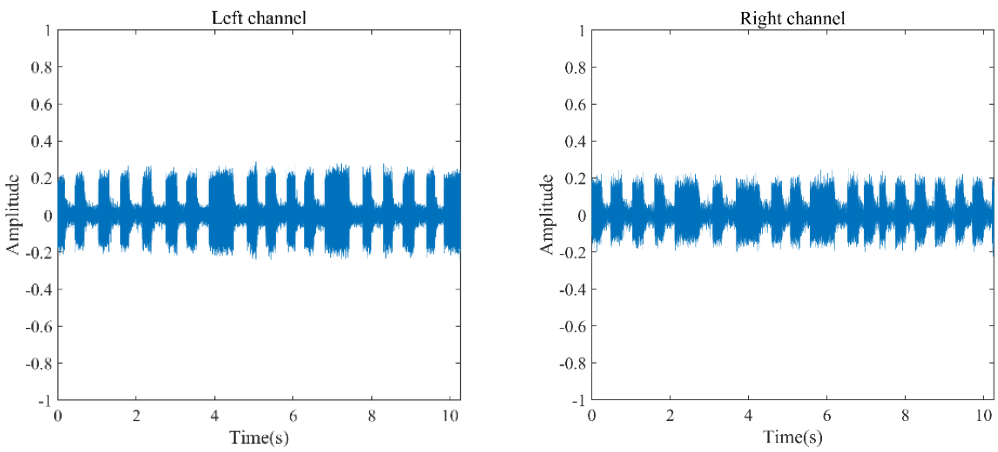


Figure 6.3.3 The Wave Amplitude in Direction of Speaker 3

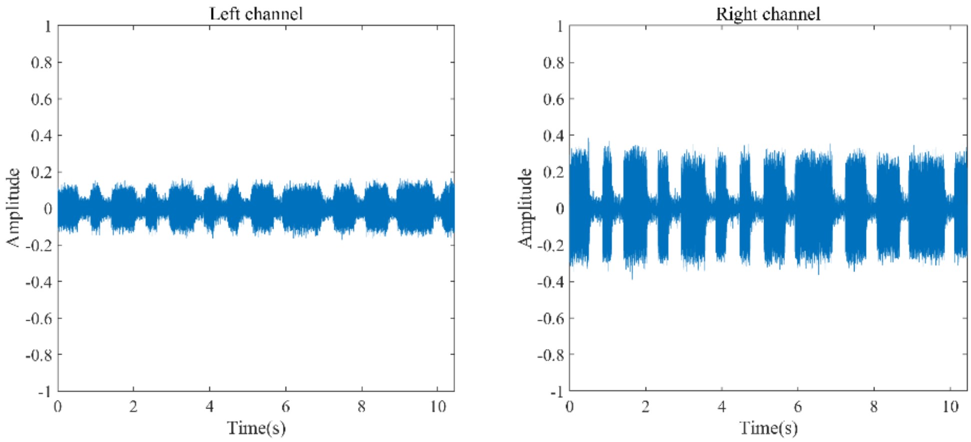


Figure 6.3.4 The Wave Amplitude in Direction of Speaker 4

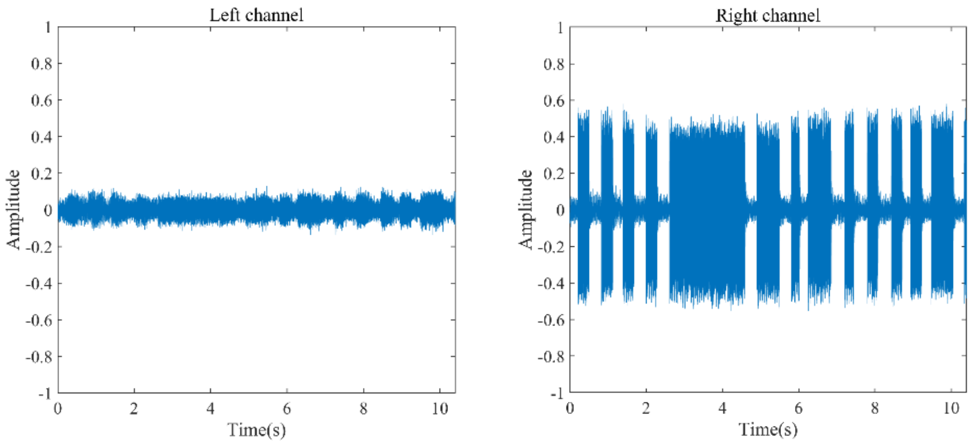


Figure 6.3.5 The Wave Amplitude in Direction of Speaker 5

Figure 6.4 Testing Each of the Directions

6.2 System Subjective Verification

6.2.1 The User Test of Distinguish the Direction

Conditions

The user should be able to tell the exact direction of the coming obstacles by hearing the related speaker beeping. This experiment is designed for verification of the system about the sound information accuracy. We need to have a clear result that user experience of the device.

According to the user experiment design in chapter 3, users should be able to tell which speaker is beeping and tell the exact direction of the nearest detected object. We will let some testers to do the experiment and estimate their subjective answer with the actual situation.

As the Figure 6.5 shows, the tester will sit on the chair to keep a steady posture and wear the depth camera sitting still. Their eyes are covered by an eye mask, so they cannot see where the objective is. The objective held by the researcher will change the location referring to the 5 directions in a random sequence. The only variable is the objective's direction, all the tests are in the case A (distance less than 0.5 meter). The sound prompts are in high pitch with a fast-looping speed. These limitations are settled for reducing contributory factors as many as possible.

Figure 6.5 Testing Distinguish the Speaker of Direction and Distance

The user sits still, and the environment background noise can be ignored. Each of the tester will answer total 25 times of the direction, each of the direction will be asked with 5 times in total randomly. Testers will answer with 'Left, Inclined-Left, Mid, Inclined-Right, and Right'. We will calculate the accuracy rate of each direction. The whole test will be held 2 times for each tester. We name the first test of tester A as **Tester A 1st Test** and so forth.

Result

Table 6.3 The Accuracy Result of 4 Tests about Distinguishing the Speaker Direction

	Left	*Inclined-Light*	*Mid*	*Inclined-Right*	*Right*
Tester A 1st Test	**4/5**	**2/5**	**5/5**	**4/5**	**5/5**
Tester A 2nd Test	**5/5**	**4/5**	**5/5**	**5/5**	**5/5**
Tester B 1st Test	**5/5**	**4/5**	**5/5**	**3/5**	**4/5**
Tester B 2nd Test	**5/5**	**5/5**	**5/5**	**4/5**	**5/5**

The result (table 6.3) shows that, the tester can distinguish the most Left / Right speaker beeping very well. However, it seems to be hard for testers to distinguish the inclined left / right at the first time. According to what tester said during the test, both testers claimed that it was very difficult for them to figure out what should the inclined Left/Right be. They felt very uncertain about the decision of the direction. It was vague to tell the difference about the most left/right with inclined left/right.

Because they cannot turn the head around to correct the right answer, the testers have to compare the sound fore and aft. They claimed that it becomes much easier when they can turn their head around.

Another conclusion is that the second test of each tester have a better result that the first time. The tester claims that they felt more confident than the first time. They can distinguish the difference better than before. We can see the result of accuracy supports this.

The result shows that our potential user can distinguish the direction of the beeping sound. The user interaction logic that converting sight information into audio information is practicable. The user can tell the 5 different directions of the obstacles.

6.2.2 The User Test of Distinguish the Distance

Conditions

This experiment's limitation is as same as the direction test above.

The user should also be able to distinguish the 3 distance cases: Safe (Case C), mid-range (Case B) and close range (Case A) by hearing different sound patterns from each speaker.

The testers will only need to hear and determine sounds of 2 different cases, and they should answer how near is it (Case A and B) since the case C will not make any sound.

Each of the tester will answer total 25 times of the distance, each of the distance case will be asked with 5 times in total randomly. The whole test will be held 2 times for each tester. We name the first test of tester A as **Tester A 1st Test** and so forth.

Result

Table 6.4 The Accuracy Result of Distinguishing the Speaker Distance Case

	Left	*Inclined Left*	*Mid*	*Inclined Right*	*Right*
Tester A 1st Test	**5/5**	**5/5**	**5/5**	**5/5**	**5/5**
Tester A 2nd Test	**5/5**	**5/5**	**5/5**	**4/5**	**5/5**
Tester B 1st Test	**5/5**	**5/5**	**5/5**	**5/5**	**4/5**
Tester B 2nd Test	**5/5**	**5/5**	**5/5**	**5/5**	**5/5**

The result (table 6.4) shows that the sound can be easily discriminated. What is more, according to the feedback of the tester, the sound patterns should not be over complexed such as 5 or 6 sound patterns since they are afraid to find out the difference from a large scale of sound patterns.

6.3 The System Validation Experiment

The final test of our design is to validate our device. We need to estimate the performance of our design for the visual impairments to help them spotting obstacles indoor in 5 directions and assisting them to establish their own way to their goal. We will simulate actual environment when visually handicapped people walking around in a building.

To simply the experiment, we set 3 scenarios of the obstacles. We need the tester move from the start zone to the finish zone and try not to hit the obstacles.

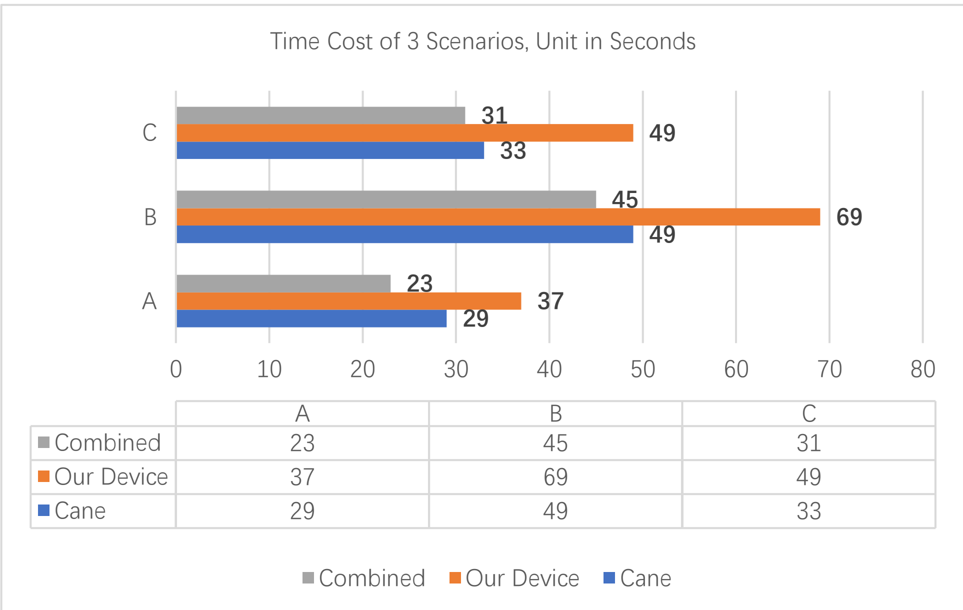
To compare different assistant methods for visually impaired, we set 3 different scenarios.

Test Condition 1. The traditional method (cane),

Test Condition 2. Our system of Research,

Test Condition 3. Two methods combined (our device with cane).

Result and Conclusion



	A	B	C
Combined	23	45	31
Our Device	37	69	49
Cane	29	49	33

Figure 6.6 Time Cost of 3 Scenarios of Cane and Our Designed Device

The functions of the system have already fully worked. It can easily tell the nearest directions of the front angle view (82 degree in 5 directions). The sound from the speaker represents the actual direction of the nearest angle is an intuitive way to help the user to distinguish the directions. The validation work has been proved (figure 6.6).

The result of the experience proves that the system is functionable. The early stage of validation and verification is passed. However, the current version of the system cannot lead the traditional method (Cane) with a huge advantage. The main reason of it is the current PC performance in our device is not enough for a high refresh rate. Users barely engage with the surroundings. It makes them feel unsafe and be reluctant to try to move faster. Meanwhile, users can go faster with the cane just because they will touch the ground and the obstacles hard without any hesitation because there is no injure risk.

When the condition comes to two methods combined, the user will have a slight better performance on the time cost comparing with using a cane. According to the feedback from the tester, they think that the cane provides a basic safety feeling underline which makes them become more confident about discovering the surroundings. The device then contributes on a more specific and precise perception of the obstacle's distribution. The experiment result supports this point that the cane and the device will have a better performance than any device working alone.

The flexibility of the system is better than other designs. At this part, the RGB sensor plays a role of the eyes, and user can rotate their head towards anywhere they want to check. This allows user to check half-height surroundings such as desks and chairs. This extends the using scenario and the future potential.

Chapter 7. Conclusion

The goal of this research is to develop a system which can assist visually impaired people to detect obstacles and navigate their routes indoors. Firstly, this research discussed the current situation of the visually impaired people. The challenges and needs of them are cleared and determined. After browsing and comparing the conventional methods, it shows that most of the ETAs can only provide 2 dimensions information for the users. Thus, to provide richer direction information to visually impaired people become the main design purpose. The depth camera is picked as our input device to gather distance information and generate depth image. Each frame of the depth image is divided into 5 blocks evenly to represent 5 angles of directions. All the pixels are processed by the minimum algorithm to find out the nearest direction and its related speaker. By converting the visual sense information into audition information, users can understand the obstacles and risks in surrounding area. This will be significantly helpful for them to get the destinations indoors. The system can detect all the static obstacles and objects with low speed such as walking human. The device of this research will be a useful assistant when the visually impaired people need to move across inside a building, and its performance would be improved and perfected when it is combined with traditional methods such as canes.

Through the whole experiment, there are some flaws of the system that the function of detecting the gaps and ditch is not considered. The algorithm is good at detecting protrude shapes rather than dent. The experiment shows that user performance would be improved with our device and a traditional method (which is a cane here) combined. We should improve it in the future work.

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my advisor Professor Shinichiro Haruyama for the continuous and great support of my study and research. It would be impossible to finish all the work without the support and the kindness from Professor Haruyama. Besides, I would like to appreciate Professor Toma's inspirational suggestions, which improved my research a lot. I would also thank committee of SDM, Keio University for your encouragement and understanding.

I would also like to give thanks to my friend Dr. Wu Nan, who helped me a lot during the research and development with outstanding knowledges and advice. Moreover, many of my friends supports me a lot for the final improvement and experiment. The research would not have been completed and improved without the help of them.

Once more, I feel grateful to my advisor Professor Haruyama and other Professors of SDM for their great teach and knowledge.

References

[1] Unite for Sight, World Health Organization Report, (http://www.uniteforsight.org/eye_stats.php), 2021 Jun.

[2] Juan Manuel Saez, Francisco Escolano, and Miguel Angel Lozano. Aerial Obstacle Detection With 3-D Mobile Devices. IEEE Journal of Biomedical and Health Informatics. 2015 Jan; 19(1):74-80.

[3] Chris B. Roberts, Yoshimune Hiratsuka, Masakazu Yamada. Economic Cost of Visual Impairment in Japan. Arch Ophthalmol. 2010 Jun; 128(6):766-771.

[4] Attitudes of Visually Impaired Persons Towards the Use of Public Transportation, Golledge, R.G, Marston, J.R, Costanzo, C.M. Journal of Visual Impairment Blindness. 1997 Sep; 91(5):446–459.

[5] Nikolaos Bourbakis. Sensing Surrounding 3-D Space for Navigation of The Blind., IEEE Engineering in Medicine and Biology Magazine. 2008 Feb; 27(1):49-55

[6] Mats E. NilssonBo N. Schenkman. Blind People Are More Sensitive Than Sighted People to Binaural Sound-Location Cues, Particularly Inter-Aural Level Differences. Hearing Research. 2016 Feb; Vol 332:223-232.

[7] Torque Stimuli, Jean-Philippe Choiniere and Clement Gosselin. Development and Experimental Validation of a Haptic Compass Based on Asymmetric. IEEE Transactions on Haptics. 2017 Jan-Mar; 10(1):29-39.

[8] Sheila K. West, Gary S. Rubin, Aimee T. Broman. How Does Visual Impairment Affect Performance on Tasks of Everyday Life, The SEE Project. Arch Ophthalmol. 2002 Jun; 120(6):774-780.

[9] Manduchi, R., Kurniawan, S. Mobility-related Accidents Experienced by People with Visual Impairment. Insight: Research & Practice in Visual Impairment & Blindness. 2011; 4(2):44-54.

[10] Schiller, J., Peregoy, J. Provisional Report: Summary Health Statistics for U.S. adults: National Health Interview Survey. National Center for Health Statistics. 2014 Feb;(260):1-161.

[11] Behavior Problems of Dogs (https://www.companionanimalpsychology.com/2016/07/behaviour-problems-in-guide-dogs.html), 2021 Jun.

[12] Juan R. Terven and Joaquín Salas, Bogdan Raducan. New Opportunities for Computer Vision– Based Assistive Technology Systems for the Visually Impaired. In Computer. 2014 Apr;47(4):52-58.

[13] Evangelos Milios, Bill Kapralos, Agnieszka Kopinska, and Sotirios Stergiopoulos. Sonification of Range Information for 3-D Space Perception. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2003 Dec;11(4):416-421.

[14] Jascha Sohl-Dickstein, Santani Teng, Benjamin M. Gaub, Chris C. Rodgers, Crystal Li,

Michael R. DeWeese, and Nicol S. Harper. A Device for Human Ultrasonic Echolocation., IEEE Transactions on Biomedical Engineering. 2015 Jun;62(6):1526 - 1534.
[15] Young Hoon Lee(B) and G´erard Medioni. Wearable RGBD Indoor Navigation System for the Blind. Computer Vision and Image Understanding. 2016 Aug; Volume 149: Pages 3-20.
[16] A Brief Introduction to 3D Cameras, (https://tech.preferred.jp/en/blog/a-brief-introduction-to-3d-cameras/), May 2021.
[17] T. B. Wang and Z. G. Liu. Relationship Between the Frequency of the Main LO Mode of Silica Glass and Angle of Incidence. Journal of Chemical Physics. 2003 July;119(1):505-508.