HandyEye: Reliable Shopping Experience for Visually Impaired Individuals

Alireza Namanloo
Computer science
Ontario tech university
Oshawa, Canada
naman.alireza@gmail.com

Farima Miri
Computer science
Ontario tech university
Oshawa, Canada
farimasadat.miri@uoit.ca

Nadia Maarfavi
Computer science
Ontario tech university
Oshawa, Canada
nadia.maarfavi@ontariotechu.net

***Abstract*—Shopping is a vital and fun activity. Visually impaired individuals usually face some difficulties and obstacles during shopping experience such as difficulty to find the item location or item details (like price and ingredients). In this paper, we proposed a new device to help VI to overcome such an obstacle and enjoy their experience. Our model consists of watch and earbud which work using Indoor Positioning System, Voice Recognition and Image Processing algorithms that tried to overcome the gaps in previously proposed models.**

***Index Terms*—Shopping, Visually impaired**

I. Introduction

Shopping is one of the crucial requirements for all of the people all over the world. We might be dealing with many problems while we are shopping that sometimes can be easy for us to resolve them and sometimes is not especially when we want to be dependent on someone else. This issue is concerning for the blind or visual impaired people as well. Obviously, most of them rely on their friends and family for shopping and most of the time they would not be able to shop independently. The question comes arise here is: what if the person who VI shoppers rely on them do not enough knowledge for helping them? Some staffers are somewhat unfamiliar about the stuff and store structure. Some of them are too busy and might ignore their questions or they can be easily angry or tired to help the others which most of the time are considered an extra work for them. In the big malls, many staffers do not even have enough language skills to read the ingredients of the products in a right way. Apart from the difficulties that were mentioned before, all of the people want to shop independently without relying on someone else no matter how much he/she has adequate English skills and is supportive toward Blind or visually impaired people. In order to get the difficulties that might arise in shopping process first we need to compare shopping levels between normal and VI or blind people. Overall, both groups do the same process for shopping. First, they prepare a shopping list. After preparation they are getting ready to go to supermarket, in order to get their desired product, they navigate themselves in the store and find the products in the related shelves. Then, they grab the products and read the price and ingredients if needed. After, they navigate themselves and getting to cash register, paying and then being prepared to go to exit door and finally going home. All these processes are the same for visual and visual impaired people. But there are many difficulties in navigation and reading the price and ingredient of the product s that sometimes make the shopping experience unpleasant for them. In the recent decade, a number of patents have been proposed in this regard to help VI and blind people in shopping groceries. The patents that have been proposed are divided into two groups: the first one put emphasize on indoor navigation and the second one addresses into outdoor navigation for VI and blind people. Our work is put into indoor navigation side and tries to resolve the issues that previous patents struggled with them and proposes new solutions for the previous problems. Moreover, in the payment section, we suppose that the products are being paid off through credit cards by VI people and blind people.

The contribution of this work is summarized as follows: In the first section we talk about related work, then we are going to talk about our methodology, after methodology in evaluation part, we compare our model with other models and justify the model using "Jakob Nielsen's 10 general principles". Finally, in the conclusion part we are going over to conclude what this model can bring for us to help VI and blind people.

II. Related work

In this section, we talk about several designs that have been proposed for VI and blind shoppers. We evaluate each model through six different design requirements which are vital in shopping process and are listed as follows : 1)Mobile Product Selection 2)Store Navigation 3)Product Search 4)Product Identification 5)Utilization of existing Devices 6)Minimal environmental adjustment [1]

A. RoboCart

It was one the earliest projects that was proposed in 2004 [2] in Utah university. They signed an agreement to do their experiment in supermarket named lee's supermarket mainly because of putting their experiment into action to evaluate their project through some tests in the supermarket. Their main goal was to design a robot to help blind and VI people while they are shopping. It was built in 2DX platform and it was included a laptop, a SICK laser rangefinder, a Radio Frequency Identification (RFID) reader, a 200mm x 200mm

RFID antenna, and a shopping basket as it is shown in (Fig. 1). They installed RFID tags in the different locations of the store. Robocart mainly relied on RFID tags for navigation. These RFID tags were placed in 5 different locations of the store: the beginning and end of each aisle and three different points within each aisle. The benefits of putting RFIDs in each aisle is they allowed the robot to know about its position and correct Localization errors through Global Markov model. Moreover, through HCD phases they test their projects with VI people. Their feedback causes them to correct some softwares and hardware sides especially in navigation parts which was led to smoother turns and better speed. They prepared some tasks for VI participants to do their experiments such as trying to get the Robocart in the designated field, using keypad to get some items, then directing themselves to different sections and get the desired product, placing the desired product into baskets and finaly directing themselves to a cash, puting the products there and paying not by cash and just with credit card and finally getting out of the store. As a result, we can say that Robocart can satisfy product selection which is not stationary due to having keypad which allows people to select their products by searching the product in that keypad. Moreover, it can satisfy store navigation with using robot that navigate them. Product search and product identification can be satisfied as well but not completely.

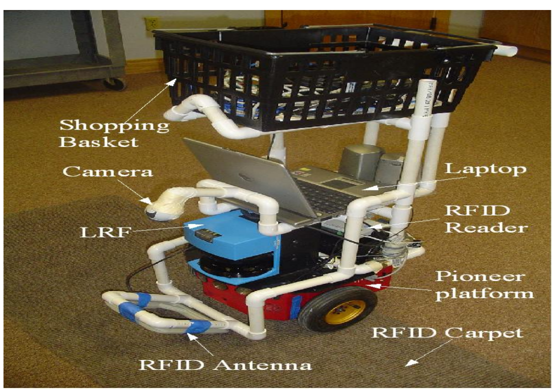

Fig. 1. RoboCart

B. ShopTalk

Shoptalk [3] was another project that was introduced for independent blind grocery shopping. They do not have any external server to support heavy computations. All the computations are being done through OQO 01 wearable computer. This computer was a solution with enough computational power. Shoptalk is composed of a Belkin keypad ,OQO computer , a barcode scanner which is a form of wireless and one base station (fig.2), and a USB in order to connect all of these things with together. They tried to solve the problems that were seen in Robocart through adding 2 plastic stabilizers that are put on the shelf to do alignment with the scanner and with barcodes that are exist in each shelf. Moreover, they have the other facilities for both left hand and right hand shoppers because all of the equipment goes into a backpack and there is numeric keypad that is connected to left or right shoulder strap. Apart from all of these facilities, the shoppers have a small headphone to manage all the instructions which are speech based. For having an independent blind shopping, Shoptalk have two main equipments namely shelf barcode scans and vocabular route commands. They concluded that VI shoppers can run vocabular route commands and item finding commands in supermarket with hundred percent accuracy and vocabular route navigations were enough for directing in the average malls by VI people and blind people. This system can satisfy mobile item choosing, item finding and navigation in the supermarket. However, identifying product can be meet partially as they need to scan the barcode of the product and this issue is getting difficult on the soft packages such as potato chips. Shoptalk cannot satisfy utilization of existing devices as it only uses the hardware component for computation and none of the existing devices are being used by VI shoppers through Shoptalk.

Fig. 2. ShopTalk

C. GroZi

This project [4] was run in California, San Diego and it has three main parts: 1) they have a website that all the VI users can go in there and prepare their shopping list. 2) in order to recognize the products, they have a computer vision approach. 3) having portable devices in order to run algorithms of computer vision and providing the user vocabular feedback. In the figures 3 and 4 we are seeing two devices named: MoZi Box and the GroZi hand glove with a camera which is small and has the portable capability and some motors that are vibarting. When they use GroZi, actually they are going to a website to prepare their shopping list and upload that list to the mentioned portable device. Moreover, this device can help the user to get directions for each item that has been written in their shopping list. For instance, if the VI people put on that gloves and show one aisle with that gloves, the portable component will direct them to the wanted product. This project was one of the projects that mainly focus on object recognition for object detection. In their design, they have 2 classification of images: in vitro and in situ. The first-class images are images of products that are being taken under normal lightning and normal conditions. The second-class images are achieved from real videos in the supermarket. In vitro images are used for training purposes and the other one is used for test purposes. Unlike two previous projects this system cannot meet mobile item choosing, it can support partially directing in the store, item search and identifying items (through different computer vision models). In addition, shoppers needs to put on a laptop on their backs and this device cannot satisfy utilization of the devices as well. However, it can satisfy minimal environmental adjustment.

Fig. 3. MoZi Box

Fig. 4. GroZi glove

D. iCare

It's another project [5] in one of the universities in the states named Arizona state University. This system mainly was designed for indoor navigation which is composed of PDA coupled with Bluetooth, a screen reader, a Wi-Fi and one RFID reader which has been placed into a hand glove (Fig. 5). There are some RFID readers in the hand gloves which help the shoppers to read ID's from the wanted items and then looking for them into a store database through using Wi-Fi connection and finally telling the user where to go. The main purpose of this project was searching into the store to get the products that VI shoppers are looking for them. When the shoppers shake their hands along the shelves, some messages such as "passing bread section," "passing juice section," etc. will be sent to the shoppers. While the shoppers need to know about the specific items, they will receive all the required information such as: price, ingredients and nutritional facts from the RFID tag which reads all of them from the package. Like Grozi this system cannot meet mobile item choosing. It can support store navigation and item search partially as RFID tags on each item can help them like navigation sign. However, the problem might arise here is how much is functional in real supermarket. Based on the hypothesis that RFID reader would be able to read the tags of every item, this system meet item identification. But it cannot satisfy utilization of existing devices in a complete way. Moreover, minimal environmental adjustment cannot be met by this system as they suppose that each item in the store has an RFID tag which is costly in big supermarkets.

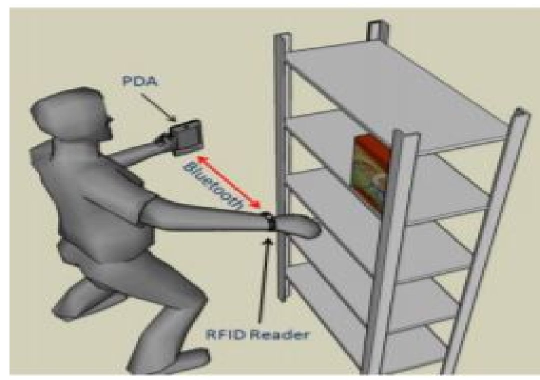

Fig. 5. iCare Framework

E. Trinetra

It [6] was run in Carnegie Mellon University and its main goal was to develop a system which is mobile and enables VI people to shop independently. This system includes barcodes and RFID tags. It is composed of Bluetooth headset and Nokia mobile phone, Barracoda pencil , BARAcoda IDBlue Pen, and finally a windows-based server (Fig 6). For scanning the barcodes VI shoppers rely on Baracode pencil and for scanning the RFID tags, they use Baracoda IDBlue. Firstly, shoppers scan a UPC barcode to get the information about the product and after getting related information, the data are passed to mobile phone wirelessly. We check a local cache to evaluate whether the barcodes have been scanned before or not. If not, the request is being sent to the remote server. And if it has been scanned earlier and it's in the server, the information is returned to the shopper phone. If the barcode is not available in the cache, a public UPC database is contacted by the server to retrieve the required information. This system can partially satisfy mobile item choosing. It cannot satisfy store directing and usually shoppers rely on other people to navigate themselves to get their products. It cannot meet item search and there is not any help for the shoppers at a run time. This system can partially meet product identification. Utilization of existing devices and minimal environmental adjustments can be satisfied by this project.

Fig. 6. Trinetra Hardware Components

F. IBM's Patent

This patent [7] has a unit which is portable and guides VI shoppers by telling them a speech synthesis and with data on their live locations from several barcode labels. Their database is responsible for storing the location of each barcode label. The portable unit is also responsible for creating a way between the current location of the shopper and the location of the item that the shoppers wants to find and also defining a way or route through using speech synthesis. The shopper would be able to choose their products from a shopping list on the portable unit. There are some ambiguous sections in this system such as whether UPC or MSI or both of them will be included in the directories or not. This system, does not meet mobile item choosing because it has mentioned that VI shoppers can select items from a shopping list that already exist on the device. Store navigation can be met by this system through vocabular route directions. Moreover, product identification can be satisfied through this system. It will meet utilization of existing devices requirements and cannot meet minimal environmental adjustment.

III. Methodology

HandyEye is a system consisting of a wrist watch, and a wireless earbud which has a microphone input and is connected to the watch.It works using an Indoor Positioning System (IPS), voice recognition and image processing algorithm.
IPS is a technology that can be used to allow users to allocate the location of people or assets inside a building using smart devices or other sensors. It gives the user an accuracy radius around 2 meters and without latency. It is accurate and reliable, so visually impaired people can use it to navigate complex buildings such as supermarkets and airports. We will use IPS through the deployment of Bluetooth beacons. These IPS use beacons and the user's Bluetooth-enabled smartwatch to find the connection. We will deploy beacon on each section of supermarkets such as dairy, bread and also it can be narrow down if needed such as baguette bread, toast bread, etc.

A. How the device works?

The user asks for a task with her voice through the microphone, the watch receives the task and recognizes which task it is, then returns the reply as a natural language voice to the earbud, so the user can hear it.

In greater details, we can define a shopping experience using HandyEye in the following steps. During the process the watch always waits and listens to the user's tasks, and any of the following steps can be interrupted by the user, when she asks a new task (Fig 7).
1. After the user enters the store, she asks for a trolley location, and she hears back the instructions on how she can get to the location to grab a trolley.
2. She asks for a product, like pasta, and she hears back the navigation instructions.
3. When she gets to the location of the requested product (1 meter accuracy using IPS), she points the watch to the product.
4. The watch uses image processing to identify the product. Then she hears: You found the product, now you can grab it. The watch waits for the person to provide the next task like: What is the price? What are the ingredients? or etc.

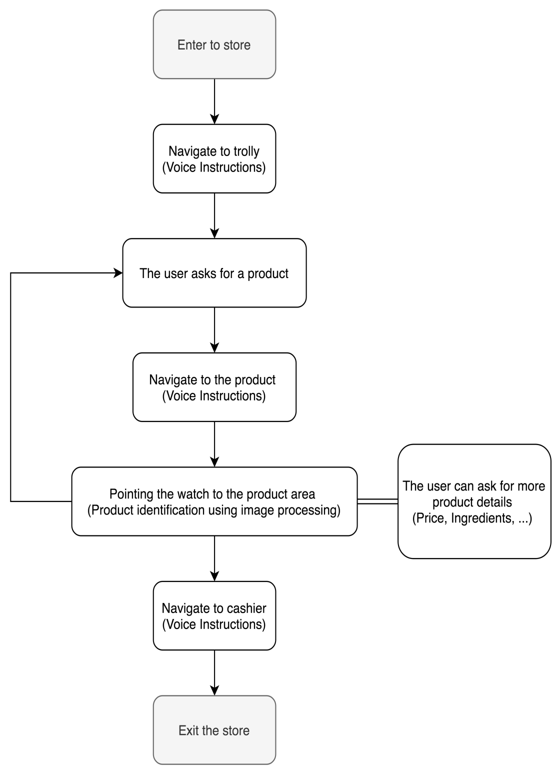

Fig. 7. Flowchart of HandyEye

B. What tasks device can perform?

There are many tasks that HandyEye is able to perform. However we can divide them into two main categories. The first category is navigation tasks, which guides the user the direction she needs to go from one location of the store to another location. The second category is product identification tasks, which recognizes a specific product and discovers exactly what it is. There are still some other tasks which can facilitate the shopping process and make it more enjoyable for the user and more flexible in a real-world shopping experience, like adding a product to a shopping list, removing a product during the process, and etc. HandyEye can store the users' shopping list in the memory and will guide the user from the nearest product to Farthest one. Also, it has the ability to create a favourite list, so the user can access them whenever she/he wants.

HandyEye is capable of helping users in the following tasks but not limited to it. What is the price? What are the ingredients? Forget about this product. I picked the product. Show me where I can find some cheese? Where is the cashier? Add this product to my favorite list. Recommend me an alternative product. Where is the next product (when there is a shopping route)

HandyEye is able to answer these various tasks through a number of devices, technologies, and methods. The watch and earbud, voice Recognition, IPS, Image processing, speech instructions, product database.

IV. Design Evaluation

For comparing our system with other systems, we used the design requirements metric proposed by [1]. Every shopping system must meet a specific requirement to be suitable and efficient. The design requirement for the shopping device is listed below and we will compare our model with others based on each criterion.

- Mobile product selection
- Store navigation
- Product search
- Product identification
- Utilization of existing devices
- Minimal environment adjustment

A. HandyEye compared to RoboCart

Our models allow the user to select the whole shopping list at the beginning of their shopping trip, or prepare the list even before getting to the store or chose any product at any time during the shopping using the voice command, while RoboCart on other hand gives the user ability to choose the item by list browsing using the attached keypad. In term of product search and product identification, in RoboCart the user has to scan the barcode shelf, retrieve the item and scan the barcode again to be sure of the selected item while in our model using high-speed image recognition technique the user can find the item by just moving his/her hand over the product and there is no need for all the trouble in finding the barcode. RoboCart will help the user to navigate to the shelf that contains the item but does not have a mechanism to help the user to find the item on the shelf. Regarding product identification, there wasn't any specific information on how RoboCart will retrieve the product information but our proposed model will do that by using voice command explained earlier. Doing some movement such as u-turn is one of Robocart's limitations while our model is light and does not limit any physical activity.

B. HandyEye compared to ShopTalk

In comparing our model with ShopTalk the most significant thing that points out is that our model is light and easy to use. Visually impaired people have to carry a lot of equipment such as white cans, guide dogs in their daily life and adding other items will be more a burden rather than help. In ShopTalk, the user has to carry a backpack with all the installed equipment which is an extra weight compared to our model which is just a light watch and earbud. Another difference between our model and the shoptalk is that ShopTalk requires the shopping list to be prepared and stored in a computational device which is not the case for our models. Our proposed model gives the ability to the user to prepare the list beforehand and also pick the item in real-time. In ShopTalk, an individual is required to scan a random item to find their location which is frustrating and time-consuming while our model guides the user by Bluetooth beacon installed in each different section and it can be done by just one command from the user.

C. HandyEye compared to iCare

Comparing our model with iCare we can mention that iCare does not fully address store navigation and product search and for doing so the user must move his/her hand along the shelf to receive instruction such as passing diary section, while our model will direct the user to the location of the item without bothering the user. In iCare, each product should have an RFID tag, this requires the store to go under so much change and not cover the minimum adjustment required for design while our model just needs a few Bluetooth beacons on each shelf. The most significant difference between our model and iCare is that iCare include a glove which a user must wear to be able to read the RFID tag in each product, this is not suitable for any weather or it might not be comfortable for all user while our model requires just wearing a light ware watch and earbud. iCare assumes that RFID readers can read all the RFID tags on the product, in case it fails there is no alternative to retrieve product information while our model has both image recognition and barcode scanner.

D. HandyEye compared to GroZi

In GroZi the shopping list is predefined and the user has not accessed to modify the list at the shop, unlike our model which gives the user the ability to predefine the list and choose the product at the shop as well. The instructions given to the user to find the product are limited to 5 orders and it might be not sufficient in most cases while our model uses voice command to make it interactive and easy to follow. To get the product identification in GroZi, the user must wear a powerful laptop to do heavy computation which is frustrating and not efficient while our model does the store navigation, product search and identification using the voice recognition and image processing algorithm embedded in a watch. The GroZi assumes that the user can navigate to each item independently which is the wrong assumption and needs some consideration. Like iCare, the GroZi requires the user to wear a glove that might not be comfortable in all weather and conditions compared to a watch which is stylish and can be worn in any condition.

E. HandyEye compared to Trinetra

The main difference between Trinetra and our models is that Trientra does not cover store navigation and product search while our model does. Trinetra does cover product identification by scanning the product barcode using Barcoda pencil or using Baracoda IDBlue Pen to scan the RFID tag. The fact that having two products to do the same task is confusing for users and doing either barcode scanner or RFID tag reader should be consistent and should be done using the same device, not multiple ones. Our model complies with this requirement by having an image recognition algorithm and a barcode scanner in the same watch. Trinetra is lacking so many tasks such as store navigation, product search which are fully covered in our model.

F. HandyEye compared to IBM patent

Like many other devices, IBM patent as well does not give the ability to the user to select the product at the store and the shopping list is predefined which is contrary to our model. Store navigation in IBM patent requires that the user scan some random products to alert the system for the current location which is frustrating to do a random scan just to find the current location and our models cover this part by using Bluetooth beacon installed in each section and does not require any action from the user. It is unknown if the IBM patent will do the product identification by scanning the barcode but our model does that. There is some aspect of the IBM patent that is unknown because it is just a patent and it has not been applied in any experiment situation.

V. Design Justification

For our design evaluation, we used Jakob Nielsen's heuristic, to identify any problems associated with our models in terms of the user interface. Jakob Nielsen's heuristics are the most used usability metric for user interface design. Jakob Nielsen's heuristics has 10 evaluation steps and we will justify our model based on these 10 steps.

1) Visibility of system status: which indicate that the user should be always informed about what is going on and the system should give information to the user within a reasonable time, our models meet this requirement by giving feedback on each stage to the user, such as "Your product has been found", "You reach your destination", "keep moving on this direction" etc.

2) Match between system and the real world: which indicates that the design should use words, phrases and concept familiar to the user and not a technical phrase which requires some prior knowledge to understand it. Our model meets this requirement because it speaks with the user using a system similar to Siri and Alexa and interacts easily with the user.

3) User control and freedom: which indicate that in case the user makes a wrong choice they can control and undo it. Our models meet this requirement in such a voice recognition algorithm in which a user can say a word such as "Cancel", "stop" to undo the request.

4) Consistency and standards: which indicate having consistency throughout the design and different words or actions should not imply the same meaning and direction. Our model meets this requirement by having friendly interaction between the user and the watch through the voice recognition algorithm which minimizes the misunderstanding.

5) Error prevention: which indicates that to check all the situations that error might occur and prevent them. In our model, we meet this requirement by having both image recognition and barcode scanner in case that image recognition is not working properly.

6) Recognition rather than recall: which indicates that the instruction for using the system should be easily retrievable and not require the user to memorize so much information, our design meets this requirement by storing the last shopping list of the user to the memory and also by providing step by step instruction to the user.

7) Flexibility and Efficiency of use: This indicates that the design should give flexibility to the naïve user and expert as well to choose their customization. In our model, we meet this requirement by having an easy, user-friendly watch that has a minimum button to choose from and it's not complex. The point is in our design there is no naïve and expert user, and all groups can use and access all features with minimum requirement.

8) Aesthetic and minimalist design: This indicates that the design should be minimalist and not include extra irrelevant information. Our model meets this requirement by having a few buttons on the watch and all require extra information that can be retrieved by voice command.

9) Help users recognize, diagnose, and recover from errors: which indicate that in case of error happening, the design should give appropriate message and instruction for the next step. Our model meets the requirement by instructing a reasonable time to the user in case the error happens, for example, if the user just went in the wrong direction, the route guide will tell the user that he is going wrong and give the substitute route. Or if the user picks the wrong product the voice will indicate to the user that the wrong product has been picked up and advise the user on further instruction.

10) Help and documentation: This indicates it's the best to have a working design that will be easy to use and not need much help and documentation but in case of need, this information should be available and easy to use. Our model meets this requirement to have instructions embedded in the system and can be retrieved by the voice command such as: how to check the price?

VI. Conclusion and Future Work

In this paper, we proposed a new device for assisting VI in their shopping experience. Our device consists of a watch and earbud that leverage using of IPS, voice recognition and image processing algorithms to aid VI. A comparison of our device with others indicates a promising result in terms of efficiency and usability. This work can be extended for future work in any direction. HandyEye can be used in other locations such as libraries or airports with minimal modification to the database. Also, we would like to enhance the image processing algorithms so it can detect multiple objects in the same frame, therefor gives the users more helpful information or even a comparison between products.

References

[1] V. Kulyukin and A. Kutiyanawala, "Accessible shopping systems for blind and visually impaired individuals: Design requirements and the state of the art," *The Open Rehabilitation Journal*, vol. 3, no. 1, 2010.

[2] V. Kulyukin, C. Gharpure, and J. Nicholson, "Robocart: Toward robot-assisted navigation of grocery stores by the visually impaired," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2005, pp. 2845–2850.

[3] J. Nicholson, V. Kulyukin, and D. Coster, "Shoptalk: independent blind shopping through verbal route directions and barcode scans," *The Open Rehabilitation Journal*, vol. 2, no. 1, 2009.

[4] M. Merler, C. Galleguillos, and S. Belongie, "Recognizing groceries in situ using in vitro training data," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2007, pp. 1–8.

[5] S. Krishna, V. Balasubramanian, N. C. Krishnan, C. Juillard, T. Hedgpeth, and S. Panchanathan, "A wearable wireless rfid system for accessible shopping environments," in *Proceedings of the ICST 3rd international conference on Body area networks*, 2008, pp. 1–8.

[6] P. E. Lanigan, A. M. Paulos, A. W. Williams, D. Rossi, and P. Narasimhan, "Trinetra: Assistive technologies for grocery shopping for the blind." in *ISWC*, 2006, pp. 147–148.

[7] O. K. S. T. J. Conzola Vincent Charles, Cox Aaron Roger, "Providing a location and item identification data to visually impaired shoppers in a site having barcode labels," U.S. Patent 6 497 367, December, 2002.