UNI FR
UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

University of Fribourg

Computer Science: Masters

Future User Interface: AR Glasses

Vision assistance for people with serious sight problems

Dr. Denis Lalanne

Fribourg, June 2016.

Moreno Colombo, moreno.colombo@unifr.ch
Marin A. Tomić, marin.tomic@unifr.ch

Table of contents

1. Introduction

a) Background and Problem Statement

People with physical disabilities are already using a lot of technologies to help themselves in everyday life with all kind of prosthesis, orthosis and/or some smart sensors which send alerts for various situation. Furthermore, there are many people who have vision problems and problems with their eyesight. Some of them cannot differentiate between colors, some have a hard time differentiating objects, some have different kind of dioptric problems and some have a mix of more of those problems or other kind of visual impairment.

Even though great help is provided to people with physical limitations by merging robotics and computer science, there is still a lack of devices to help visually impaired people. This paper will focus on specific wearable technology which uses some specific hardware with corresponding software to help people with serious sight problems.

b) Research Questions

Problems and challenges described in the section above can be put and described with answers to couple of precise questions. Questions which, if answered carefully and thoroughly, can explain a lot more than just assistive methods used for sight problems and complications influenced by them. The following few questions are made to do exactly that, to explain and elaborate more:

- Can wearable computing be used to help people with sight problems?
- How can such technology be implemented?
- Is it possible to augment reality through smart glasses in a way that helps specific users and provide multimodal intuitive interface?

The purpose of this report is to define and present possible solutions as well as to elaborate a methodology and describe the problems connected to it. By answering the questions and elaborating the whole aspect behind them, many of ambiguity will become much clearer. Exploring and understanding the specific person's problem from various points of view, computer science and user itself, will reveal more details and help in overall comprehension of this complex topic.

c) Project Description

Epson BT-200 smart glasses, which offer the possibility of augmenting reality, with a combination of Unity3D and plugins provide all necessary functions this project needs. The project aims to help persons who are visually impaired augmenting reality with simple red shapes. Hardware required for it consist of Epson BT-200 smart glasses and visual markers similar to QR codes. The software developed in Unity3D allows users to effectively perceive various objects (doors, remotes, sofas) with a visual marker put on them. Interaction with system is either through a special marker held by the user ("explorer mode") or head gestures. The system acknowledges commands via sound output (saying which object has been chosen), it then gives as output the position of the searched object, when this is in the field of view of the user, via a blinking red object shown at the same position of the searched object, combined with a sound telling the distance from it.

2. Architecture and Software

a) Epson BT-200

Epson BT-200 are binocular, transparent smart glasses with two 1cm wide screens implemented in each lens. BT-200 have VGA camera on the right side of the glasses, they have all of popular wireless and sensor technologies implemented and Android OS. Their processing power is 1.2GHz dual core with 1GB of RAM. With its see-through applications, BT-200 are made to be development platform for apps of the future and hands-free scenarios, delivering large, 2D or 3D, images with augmented reality.

All mentioned functionalities and features are part of BT-200 but also part of Epson "smartphone" which has to be connected with a cord to the glasses. "Smartphone" is actually a big touchpad with mentioned CPU and memory which does all the calculations and processing of the glasses.

b) Implementation

A couple of versions of the aforementioned idea were made. One implements input with "explorer" marker which, when shown to BT-200's camera, it allows the user to choose which object to search by removing the marker out of sight when hearing the wanted object. The other one consists in letting the user shake or nod his head to choose which object he wants to look for (this allows the user to choose between only two "favorite" objects). Both versions have audio feedback which depending on the distance from object outputs distance in meters or beeping sound. "Five" till "two meters" is output from 5 till 2 meters of distance. Distance less than 2 meters is represented as beeping sound whose frequency increases when user is getting closer to the object. Furthermore, same versions with different output method exist. To test and provide different option of evaluation, rather than hearing the distance from the object, these versions have just red object blinking depending on the distance from it. Choosing the objects with audio output still stays the same, with audio feedback.

An additional particular version has been adapted for a specific test case. Since the entire project start was conceived for a specific person with a very unusual and interesting state, with particular sight problems as well as cognitive ones, a simple version as a proof of concept has been made. This version has no inputs nor possibility of choice. Though, the outputs stay the same. Also, instead of red shape presenting the object, whole screen was made to be red when object is visible.

c) Special case

Due to privacy concerns, brief explanation without extensive details will be mentioned. A special version of the application, mentioned previously, was made to test and evaluate as well as to prove that the concept could be helpful. Tested person has a serious sight problem which prevents him to perceive almost anything but the red objects. Even though the person is not able to perceive it can see, but the brain cannot process what it sees. Therefore, the person cannot designate objects and shapes but it can differ the distance from that object. Moreover, the person has problems with memory and thus problem with learning. Due to these and connected issues, the version had to be as simple as possible.

3. Multimodal Interaction

a) CASE

All presented versions, except the “special” one are using exclusive multimodal communication type. The commands given by the user have no coreference and are given one at a time. User can choose between two modalities: either using “explorer” tag or moving the head.

Another input modality has been developed but, due to time limit and unexpected complications as well as main aim of the project (help a specific person), this modality has not been completely implemented for Epson BT200. The mentioned modality is speech recognition which could allow person to intuitively and more naturally choose between an arbitrary number of objects.

b) CARE

Usability and fusion in multimodal interaction can be characterized and assessed with the help of the CARE model. Equivalence property of the mentioned model is an accurate classification of all versions in the project. It is indeed necessary and sufficient to use any of the two provided modalities to reach the “goal”. The user has a choice between these two modalities and whichever he chooses, the goal remains the same.

c) Output

As mentioned before, there are basically three outputs provided by the different systems: audio, blinking red shape and combination of those two. Audio output for choosing the object, with head movement or the “explorer” marker, is present in every version. Additional audio output provides explicit information about the distance from the chosen object when this is in the camera’s field of view and is available in particular version. On the other hand, instead of audio feedback for distance, there is an option which provides blinking red object for it. Blinking speed depends on the distance of the user from the object. Finally, both blinking red shape and audio feedback have been merged in a unique system for the special case mentioned in previous paragraph. The goal of this was to increase the efficiency of the system for the user with serious sight problems. In fact, the user sometimes had some difficulties to perceive also the red color, so an audio feedback could have been a good addition to the system to help him react faster to the presence of the red color in his field of view.

4. Results and Conclusion

The application made for this project has been, from start, oriented towards special case of specific person's sight problem. Due to the specific problem, a qualitative evaluation has been made. With special case version of application, evaluation showed that user can detect bright red color. It can approximately determine where the object is and orient itself towards that object. Still, evaluation exposed important challenges which are necessary to be solved in future work. Challenges such as audio output optimization and synchronization, visible area of the implemented camera and duration of the appearance of the red color. Some of these problems are connected to hardware: the camera is located at the right side near right eye and it doesn't allow to naturally follow direction of the head, resolution of camera doesn't allow the system to recognize markers from further distance and in low light situation. Future work is to address these and related problems.

5. Future work

By proving the concept many areas and ideas have been elaborated. Speech recognition is one of the ideas, mentioned earlier in this paper. Even though speech recognition is not perfect, it could be used for simple few-words commands. Furthermore, by testing with different sizes and shapes of the presented red object, one could find interesting results. Changing the color or even the frequency of blinking can represent another interesting idea. Moreover, connecting smart glasses with smart bracelet or smart watch could be used to remove audio output which can be intrusive and uncomfortable. With these kind of wearable devices, one could use vibration as an unobtrusive output. Another idea is to implement simultaneous localization and mapping (SLAM) with built in inertial measurement unit (IMU) of the glasses which could map a specific area, localize the user in that specific area, and provide an output to lead the person in the wished point by using his position and the map of the room. Depth perception, even though not possible with Epson BT200, could also be used for various situations. It could be used to alarm user if something is in front of him, to scan different objects, to provide hand gestures recognition etc. Many other uses and ideas will eventually pop up during the future work and further implementations.

With the perspective they have, smart glasses can (should) be used for vision assistance for people with serious sight problems. One particular case has shown that they are helpful but further work is necessary. With this in mind, new projects and development have already started.