Developing vision substitution products using computer vision

Björn Gummesson

DEPARTMENT OF DESIGN SCIENCES
FACULTY OF ENGINEERING LTH | LUND UNIVERSITY
2019

MASTER THESIS

Developing vision substitution products using computer vision.

Björn Gummesson

LUND UNIVERSITY

Developing vision substitution products using computer vision

Copyright © 2019 Björn Gummesson

Published by
Department of Design Sciences
Faculty of Engineering LTH, Lund University
P.O. Box 118, SE-221 00 Lund, Sweden

Subject: Interaction Design (MAMM01)
Supervisor: Kirsten Rassmus-Gröhn
Co-supervisor:
Examiner: Charlotte Magnusson

Abstract

The purpose of this thesis is two-fold. The first is to investigate if cutting edge technology can be used to enhance the life for people with visual impairments. The second is to provide a simple example of how the design process can look from the beginning to a finished Hi-Fi prototype. This was achieved by gathering requirements for a product and creating a prototype,

This was carried out in three phases. First relevant literature was researched, and the discoveries were distilled into a set of design methods which was used. The second phase is the requirements gathering where the methods chosen in the first phase is used to elicit requirements for the prototype. The third phase is the implementation of said prototype with evaluation.

The overarching method used in the requirements gathering is the umbrella approach. In the beginning a very broad perspective is kept which iteratively shrunk until the requirements for the prototype is set. This follows the user centered approach to product design.

The functionality which was decided to be prototype was a hand gesture in front of a camera which could select certain areas of the video feed. This part of the image can then be analyzed in different ways, for example, with OCR. The functionality was eventually implemented by tracking different colored stickers placed on the tip of the index finger and thumb. The prototype was unstable which made the results from the user evaluations inconclusive.

The results of the thesis were semi-successful. Many different types of functionality were investigated but only superficially. The hand gesture functionality may be feasible from an implementation perspective, some implementation improvements possibilities is provided under discussion. It may also be feasible from a usability perspective; more evaluation needs to be done with a more stable prototype.

Keywords: product development, user centered design, interaction design, rehabilitation engineering

Acknowledgments

I would like to thank the people at Iris Hadar who welcomed me with open arms and made this thesis possible. I would also like to thank LVI, especially Henrik Blomdahl who was my contact person with the company and my supervisor Kirsten Rassmuss-Gröhn.

I would also like to acknowledge the motivational inspiration my family and my friend Pontus Håkansson has provided me when the task at hand have seemed impossible.

Lund, June 2019

Björn Gummesson

Table of contents

List of acronyms and abbreviations

LVI	Low Vision International
OCR	Object Character Recognition

1 Introduction

This introduction aims to make this paper easier to read by giving an overview of it and put it in a context. It contains some background information on which organizations I cooperated with and why this thesis is relevant. It also contains a brief overview of the goal of the thesis and how it was carried out.

1.1 Background

People with visual impairments are excluded from everyday things which people with full vision takes for granted. Meanwhile technology in general is improved rapidly. I wanted to investigate if I could find new ways of using cutting edge technology, to improve some aspect of the lives of people with visual impairments. In particular I wanted to investigate if image processing or computer vision could be used. To accomplish this, I acquired help from two organizations. Low Vision International and Iris Hadar.

Low Vision International is a world leading manufacturer of products for the visually impaired. All development and manufacturing are made from their Head Offices in Växjö.

Iris Hadar is an organization with the overall goal to create a more inclusive workplace. The division I visited in Växjö worked with people with visual impairments. They held different types of activities with the goal of helping people with visual impairments overcome the difficulties in the job market relating to their visual impairments.

1.2 Goals

The goal of this thesis is to investigate new ways of using image processing or computer vision to enable people with visual impairments to better access information in the environment. This will be investigated by creating a prototype for a product using a design process with heavy focus on usability. The degree of fidelity of the prototype was not decided beforehand.

I have identified one main challenge for this thesis, which is the additional difficulty which comes with designing a product for people with visual impairments. This put more restrictions on the product compared to designing for people without visual impairments.

1.3 Contributions

My hope is that this thesis will contribute to the industry in two ways. The first is to create a proof of concept of an innovative product which will benefit the lives of people with visual impairments. The intention is that if a prototype or product created during this master thesis is useful, LVI will pick it up and develop a market ready product from it.

The second is to provide an example of how the process of designing a new prototype can look. Many of the textbooks I have seen only provides partial examples or goes into much detail. Hopefully, this report will work as a simpler practical example for inexperienced people like me, of how they can design their ideas.

1.4 The method used

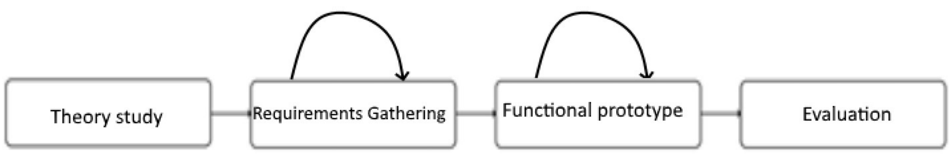


Figure 1

A visual overview of the method is shown in Figure 1. In the beginning I focused on studying the theory behind the design process. A lot of focus was put on different techniques to gather the requirements. I have described my findings during this phase in chapter 2. This chapter is a distillation of the knowledge I believed to be most useful for my situation and hope that readers which want to create their own product find this useful as well. In the User Study guidelines paper [1] they provide a practical example to each technique. This helped me understand the message better and in the same way, I have provided a small example of how many of the sections in chapter 2 was useful to me in my work.

Then when I felt I had gained enough theoretical knowledge, I started on gathering the requirements. This followed the umbrella approach; I started with a very broad problem space and successively narrowed this problem space down until I had found what functionality I had wanted to prototype.

1.5 The prototype

This functionality I wanted to implement, turned out to be a hand gesture for marking a certain area in a cameras field of view to be analyzed. Out of the different functionalities that was investigated during the user studies, this was the one I deemed could be most useful to most users which I had the time and resources to implement. More about this can be read in section 3.8. This gesture was opening the fingers see Figure 22.

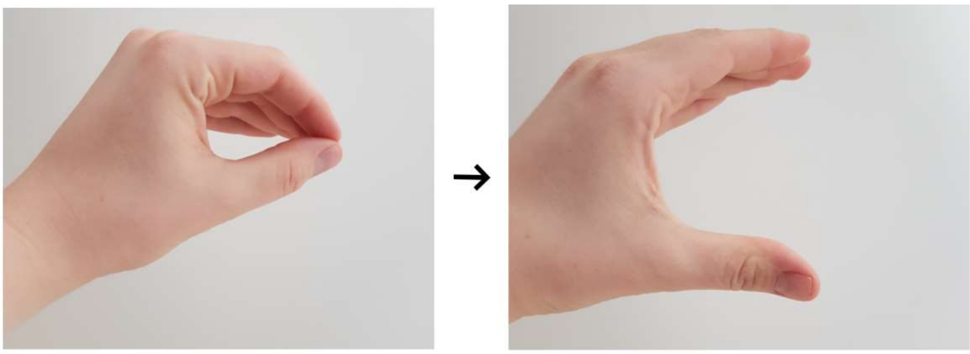


Figure 2

When I knew what to implement, the implementation phase was next. A lot of experimentation with different methods was done which did not generate any results and in the end, I tracked the movement by tracking colors of stickers I put on the index finger and the thumb (See Figure 33).

Figure 3

The were some issues with the stability of the implementation, this will be described in detail under chapter 4. Despite this, I managed to evaluate the prototype somewhat successfully, more is described in chapter 5.

2 Theory

In this section I provide the theoretical background for the work carried out in this paper. Sections 2.1, 2.2, 2.3 and 2.4 is related to the design process in general terms. These theories laid the foundation for my requirements gathering and might be needed for understanding Chapter 3. 2.5, addresses the different techniques I used for gathering data and 2.6 provides some technical detail which helps with understanding the implementation.

2.1 Design Process

When discussing the design process in general terms [2, p.15], Preece et al brings up four basic activities. These activities are intended to inform one another and to be repeated iteratively. The four activities mentioned are;

1. Establishing requirements
2. Designing alternatives
3. Prototyping
4. Evaluating

2.1.1 **Establishing requirements**

There are two main objectives when establishing requirements; to understand as much as possible about the users and produce a set of stable requirements that form a sound basis to start designing [2, p.350].

An oversimplification of the process to establish requirements can be described as three sequential activities;

1. Gather data
2. Analyze that data
3. Extract requirements

However, the reality gets a lot messier. For example, while analyzing data, often new insights arise which requires further data gathering. The iterative nature of the user-centered approach to design, addresses this concern.

2.1.2 **Designing alternatives**

Designing alternatives is the core activity of designing: actually, suggesting ideas for meeting the requirements [2, p.330]. Designing alternatives can be divided in to two sub-activities: conceptual design and concrete design.

During the conceptual design, the conceptual model for the product is created. It begins with designing some initial ideas for the conceptual model. When designing the initial conceptual model, Preece et al mentions three aspects that should be considered:

1. Which interface metaphors will be suitable?
2. Which interaction types would best support the user activities?
3. Do different interface types suggest alternative design insights or options?

[2, p.398]

The initial conceptual model needs to be further expanded and described in more detail before being prototyped. This means deciding which functions the product will perform, how these functions are related and what information is required to support them [2, p.403].

Concrete design is closely related to conceptual design. The difference between them is whether emphasis is put on conceptual issues or more concrete details. This is not a binary choice between one or the other. It is rather a gradual shift of focus and some concrete issues will come up during conceptual design and vice versa [2, p.405].

2.1.3 **Prototyping**

A prototype is one manifestation of a design that allows stakeholders to interact with it and explore it´s suitability. A prototype will usually emphasize one set of product characteristics and de-emphasize others [2, p.386].

2.1.4 **Evaluating**

Preece et al discuss evaluation from four different perspectives [2, p.452]. Why, What, Where and When of Evaluation?

Evaluating enables designers to focus on the real problems and needs of the user, rather than trying to guess or debate what these are. Evaluating a design before the product goes on sale, enables problems with the product to be detected early and therefore cheaper to correct.

What to evaluate depends on the product, the designers need to consider the goals and purpose of the product when deciding this. When what to evaluate is decided, where the evaluation shall take place needs to be adapted to this.

Evaluations may be done at several times during the design process. Anytime the designers want to test if their design really meets the user’s needs, evaluation can be done.

2.2 Data Gathering

Jenny Preece et al introduce 5 key issues that requires attention for user data gathering sessions to be successful, these issues are described briefly below [2, pp.222-230].

2.2.1 **Setting goals**

Before beginning it is important to identify specific goals for the study. The goals that are set will influence the data gathering sessions, the data gathering techniques and the analysis to be performed [2, p.222].

As the user study activities in this master thesis had different purposes, different goals were set up for the different activities.

2.2.2 **Identifying participants**

The goals for a data gathering session will affect the participants you choose. Most often you do not have access to the whole population you are trying to target so you need to choose certain participants. This is called sampling.

The sampling method used in this master thesis is convenience sampling. Some aspects of stratified sampling were considered as well. Convenience sampling is characterized by choosing participants based on who is available [2, pp.227,228]. I had some problems finding willing participants so choosing participants carefully to receive a more accurate picture of the population was not realistic. Stratified sampling relies on being able to divide the population into groups and then choose randomly within these groups [2, pp.228]. The aspects of this that was applied in my user studies, was normalizing the time I spent on participants with different types of visual impairments. For example, only one of the participants had a 100% visual loss. I spent more time with this participant to make sure this group was studied extensively as well.

2.2.3 **Relationship with participants**

The relationship between the person gathering data and the people providing data is a significant aspect of the data gathering process. Making sure the relationship is clear and professional will help to clarify the nature of the study. One way this can be achieved is to use a consent form. The details of consent forms often vary but usually asks the participants to confirm the purpose of the data gathering and how the data will be presented [2, p.228].

For logistical reasons an explicit consent form was not created for this user study. Since the participants had visual impairments, reading and signing on site would have taken an extra portion of time from the rest of the data gathering. I did however consider these aspects and verbally gave the participants the same information that could have been in the consent form.

Before every data gathering activity, I mentioned the following points:

1. The participant can withdraw at any time and if topics that the participant does not want to discuss comes up, they could simply say “I do not want to discuss this” or “ I don’t think I have anything to contribute to on this topic”
2. The data might be published publicly but will in that instance only identify the participant by age and type of visual impairment.
3. The goal of the current data gathering activity

A personal comment on this is that I believe establishing a professional relationship with the participants is especially important when the participants are differently abled. There is a risk for someone who is fully abled to offend participants unintentionally. People who are differently abled is a diverse group and in my experience in this thesis, my participants were comfortable with discussing their visual impairments. However, I suggest taking extra care in to giving participants the ability to opt out or decline to answer questions about certain topics at any time during the user studies. The issue about what information might be published is also very important.

2.2.4 **Triangulation**

Triangulation is a term used to refer to the investigation of a phenomenon from different perspectives. Preece et al mentions four types of triangulation. Two of which is relevant to this master thesis [2, p.228].

1. Triangulation of data means that data is drawn from different sources at different times. This was adopted by asking different participants the same questions during personal interviews.
2. Methodological triangulation means to employ different data gathering techniques. Which techniques were used in this master thesis will be described further on.

2.2.5 **Pilot Studies**

A pilot study is a small trial run of the main study. The aim is to make sure that the proposed method is viable before embarking on the real study. If it is difficult to find people to participate or if access to participants is limited, colleagues or peers can be asked to comment. The participants in the pilot study should not be involved in the main study [2, pp.229,230].

As mentioned earlier, I had a limited number of participants, so a substitute was made by asking my contact on LVI to examine the questions and come with improvements.

2.3 User-Centered Approach

Preece et al emphasizes throughout the book the need for a user-centered approach to development which focuses on the real users and their goals [2, p.327]. In relation to this, Preece et al brings up three principles which lead to a useful and easy to use computer system.

1. *Early focus on users and tasks*: Focus first on understanding who the users are.
2. *Empirical measurement*: Throughout the whole development, users' reactions to and opinions on different design alternatives should be recorded and analyzed.
3. *Iterative design*: Designs should be refined based on feedback.

This approach was used in this thesis. In the beginning of the requirements gathering, I only focused on understanding the users' needs from a broad perspective. The requirements gathering overall was also iterative. Each idea for a feature was tested by evaluating it instantly. One example of this, is the way I tried to disprove the utility of a new feature by asking things like "How often would have use for this feature?" and "How do you solve this need today and why would this feature be better?"

2.4 Conceptual model

A conceptual model is a high-level description of how a system is organized and operates. It describes what people can do with a product and what concepts are needed to understand how to interact with it [2, p.41]. It helps designers to straighten out their thinking before they start laying out their widgets.

Preece et al mention four core components of a conceptual model. By explicating these, it easier to discuss merits of different ones.

1. Metaphors explaining how to use the device.
2. The concepts that people are exposed to through the product.
3. Relationships between these concepts.
4. Mappings between the concepts and the user experience the product is designed to support or invoke.

Preece et al also emphasizes the importance of simplicity in conceptual models. The simpler they are, the easier they are for the user to use and hence how effective they are. These concepts where used to create a mental image in the beginning of the requirements gathering which I called product model.

2.5 Techniques

2.5.1 **Principles**

Principles in this context describes a set of pre-defined principles which can be helpful to product designers during the design process. There are several lists with principles from various organizations, in user study guidelines [1, p.17] “universal design” and “inclusive design” is mentioned. Universal design contains seven principles:

1. **Equitable Use:** The design is useful and marketable to people with diverse abilities.
2. **Flexibility in Use:** The design accommodates a wide range of individual preferences and abilities.
3. **Simple and Intuitive Use:** Use of the design is easy to understand, regardless of the user´s experience, knowledge, language skills, or current concentration level.
4. **Perceptible Information:** The design communicates necessary information effectively to the user, regardless of ambient conditions or the user´s sensory abilities.
5. **Tolerance for Error:** The design minimizes hazards and the adverse consequences of accidental or unintended actions.
6. **Low Physical Effort:** The design can be used efficiently and comfortably and with a minimum of fatigue.
7. **Size and Space for Approach and Use: Appropriate** size and space is provided for approach, reach, manipulation, and use regardless of user´s body size, posture, or mobility.

I simplified these principles in to two main principles in order to be more effective when discussing and thinking about design principles during user studies.

1. **Intuitive use:** The product shall be easy to understand for as many user´s as possible and provide a lot of feedback and clear information to the user. It shall also be physically easy to use.
2. **Tolerance of error:** The product should be able to handle user and other errors with as little manual error correction as possible.

2.5.2 **Scenarios**

Scenarios are fictional stories that is used to focus on user needs and thoughts [1, p.23]. In user study guidelines, the authors put forth an example of how the scenarios can be created. In this example, the authors implore the reader to focus on three things;

1. The situation that motivates the user´s need for a function.
2. How the user access that function.
3. The satisfaction the user receives from having performed said function.

An informal variation of this was used during meetings with LVI.

2.5.3 **Questionnaires**

Developing the questionnaires is often very time consuming but also very important [1, pp.33,34].

User study guidelines provides some guidelines for developing a questionnaire.

1. Provide a clear statement of purpose & guarantee participants anonymity.
2. Decide on whether phrases will all be positive, all negative or mixed.
3. Pilot test questions – are they clear, is there sufficient space for responses
4. Decide how data will be analyzed & consult a statistician if necessary
5. Offer a short version for those who do not have time to complete a long questionnaire.

User study guidelines also discusses the expected outcome. A questionnaire can be designed to gather different aspects of system design. These guidelines were used when designing the questionnaire used when gathering requirements. Further described in section 3.3.

2.5.4 **Diary Studies**

In a diary study, participants are asked to keep a daily record of their activities and thoughts from a certain perspective. This technique give room for longer contemplation for the user. It explicitly encourages people to look at and think about their lives which might introduce new ideas and perspectives which could have otherwise been missed. The principle of giving users time for contemplation was used in the user study sessions with Iris Hadar.

2.5.5 **Lo-fi prototyping**

A lo-fi prototype has been defined as an artefact that has been created independently from the likely form of a finished product [1, p.41]. Lo-fi stands for *Low Fidelity.* In other words, the fidelity of the prototype matching the product is low.

This technique is useful early in the design process as it is cheap to implement in terms of time and resources. It also tends to shift the designers focus from implementation to design. It is estimated that when making a lo-fi prototype, the designer spends approximately 95% of the time thinking about the design and about 5% of the implementation. When creating a more advanced hi-fi (high fidelity) prototype, the ratio is roughly reversed [1, p.41].

User Study guidelines provide an example of a work method when designing a lo-fi prototype called paper prototyping. In this method, a paper representation of the product is created. An example of this is the cardboard computer. The functionality in this case were simulated by switching index cards on the "screen" [1, p.41].

A similar representation of the product in this master thesis were created with a metal box, a lanyard and coins taped to the box. This should simulate a device with a camera lens that is hanged around the neck of the user. Further described in section 3.6.

2.5.6 **Brainstorming**

Brainstorming was first introduced as "organized ideation" [1, p.41]. The basic idea of brainstorming is to generate as many solutions to a problem as possible. Focus in this technique lies on quantity instead of quality. A group of people throws out as many ideas as possible and these ideas are then recorded in a way that suits the setting.

User Study guidelines provide some tips when conducting a brainstorming session:

1. Focus on quantity
2. Postpone Criticism
3. Encourage building on each other´s ideas
4. Stimulate idea generation

User study guidelines also provide an example where the problem was first explored in another setting. After this, a brainstorming session was held to generate solutions to the problem.

I used this structure in the first session with Iris Hadar. First a focus group was held where problems were discussed. Then a brainstorming session was held to generate solutions to these problems.

2.5.7 **Interview**

According to Preece et al, interviews can be thought of as a "conversation with a purpose" [2, p.232]. Interviews are the most traditional technique for gathering requirements. The general procedure is rather straight forward but there many techniques that can be useful to learn to carry out the interviews in an efficient way.

There are two general types of questions. Closed and open. Closed questions require an answer from a predetermined set of alternatives. Closed questions work well if the range of possible answers is known [2, p.234].

Open questions do not require the answer to be from a predetermined set of alternatives. They are more exploratory and therefor better suited where the goal of the session is exploratory [2, p.238].

Preece et al brings up different types of interviews. These are also reflected in User Study Guidelines. What Preece et al call Group interview, has its own section in User Study Guidelines where it is called Focus groups. This will be described in next section. These different types of interviews revolve around the use of open and/or closed questions.

2.5.7.1 Open or Unstructured interviews

In open interviews, as the name suggest, open questions are used. As they are more exploratory, they are good to use in the beginning of the design process. Even though they are unstructured, it is still advisable to have a plan of how to maintain focus on the main topics [2, p.233]. The interviewer needs to get the balance right between making sure that answers to relevant questions are obtained, while at the same time being prepared to follow new line of enquire that were not anticipated. Following a new line of enquire is often done by what is called probing questions (probes). For example, the interviewer will say: “Can you tell me a bit more about…”.

Open interviews gather rich data. As they are more exploratory in nature, they are often used early in the design process when the designers know little about the problems. It is good to have a plan before the interview about how to analyze the data.

2.5.7.2 Directive or structured interviews

In directive interviews, short and closed questions are typically used. Directive interviews is best suited for when the designer knows more about the problem.

2.5.7.3 Semi-directive or semi-structured interviews

Semi-structured interviews are a hybrid of the two types mentioned above. The semi-structured interview can be adapted to be more towards structured or unstructured depending on the goals of the interview.

2.5.8 **Focus Groups**

Focus groups is a group interview where participants are expected to be end users. At least one person from the group that conducts the user studies acts as moderator. The moderator keeps the discussion on track and introduces different topics. From the interviewer´s perspective, focus groups follow the same rules as open or semi-directive interviews [1, p.51].

2.5.9 **Workshops**

Workshops are hands-on sessions where small groups of professionals and/or non-professionals work creatively together. The key is the group work, allowing for an interaction between group members that can trigger ideas based on other people's ideas. Typical workshop activities include discussions and idea generation [1, p.51].

2.5.10 **Cognitive walkthrough**

A cognitive walk through is a usability inspection method that takes into account the first-time user and what kind of usability problems he or she might encounter. A cognitive walkthrough is typically performed by a person involved in the design project, not the end-user.

To perform a cognitive walkthrough, the designer creates typical use cases for the user. The designer then will "walk through" these use cases and investigate how hard it will be for a user to complete them.

A cognitive walkthrough can be used in the design phase with a paper prototype [1, p.29]. This was performed with the lo-fi prototype created when gathering requirements. See section 5.3 for more information.

2.6 Technical details

2.6.1 **OpenCv framework**

Opencv is an open source framework for computer vision first developed by intel in 1999. It is currently maintained under the BSD license. It has over 2500 optimized algorithms and leans towards real-time vision applications. I gained most knowledge about this framework from tutorials provided by the opencv community and from independent creators on Youtube.

2.6.2 **HSV color model**

The HSV color model is an alternative to the traditional RGB (Red Green Blue) color model. In the case of RGB, different colors are represented by different values of Red, Green and Blue. For example, black is 0,0,0.

HSV works the same way but instead of red, green and blue, the values represent hue, saturation and value. In this paper I use HSV, since it is more effective when tracking objects based on colors.

2.6.3 **Image processing and computer vision**

In the article *Computer vision-based object recognition for the visually impaired in an indoors environment: a survey* the authors make a distinction between computer vision and image processing [3, pp.6]. Image processing is any form of signal processing for which the input is and image. The output can be an image, or a set of characteristics or parameters related to the image. Computer Vision is often considered high-level image processing out of which some sort of computer intends to decipher the physical content of an image.

In this thesis the feature areas in the features document discussed in 3.7, reflects these distinctions. The area *Live Image manipulation* deals with image processing while the area *Identifying* has features which would be characterized as computer vison.

2.7 Assistive technology for people with visual impairments.

Electronic assistive technologies for people with visual impairments can be divided into three main categories; *Vision enhancement technologies, vision replacement technologies* and *vision substitution technologies.* [3, pp 3,4]

Vision enhancement technologies accept visual input from a camera, enhance it in some way, and output it on a visual display. *Vision replacement technologies* is different types of neuroprosthetic devices. *Vision substitution technologies* accept input from the user´s surroundings and decipher this to extract information which is supplied to the user by some sort of means e.g. audio.

Vision enhancement technologies is the corresponds to the feature area called *Live image manipulation* discussed in 3.7. *Vision substitution technologies* corresponds to the area called *Identifying.* This thesis will not investigate *Visual replacement technologies.*

3 Gathering requirements

During the phase of gathering requirements, I used the "umbrella approach" which entails looking at the problem from a very broad perspective in the beginning. Then you iteratively throughout the process narrows down the problem space to the issue you want to solve.

First, I conducted a pre study where I created an overview of different aspects of the product. I also created a plan for the user studies. The results from the user studies were to be recorded by continuously refine the work documents.

After the pre study I conducted the different user study activities that are described in chronological order. These are also divided in to three steps; planning, session and analyzing. The exceptions to this format are the sessions with LVI and the section describing the Lo-Fi prototype. The sessions with LVI was informal and had no structure. The section about the Lo-Fi prototype was extracted from the sessions with Iris Hadar to increase the readability.

The result section of this chapter gives an overview of the work documents which are also provided in appendix A, B and C. The last section motivates why I chose to implement the hand gesture functionality.

3.1 Pre study

3.1.1 **Product model**

A high-level description of the product was created (See Figure 4). It explicated different aspects of the product. The model was mainly created to easier communicate and discuss different aspects of the products with other people. But it also helped me when working analyzing data from user studies and creating the different sessions.

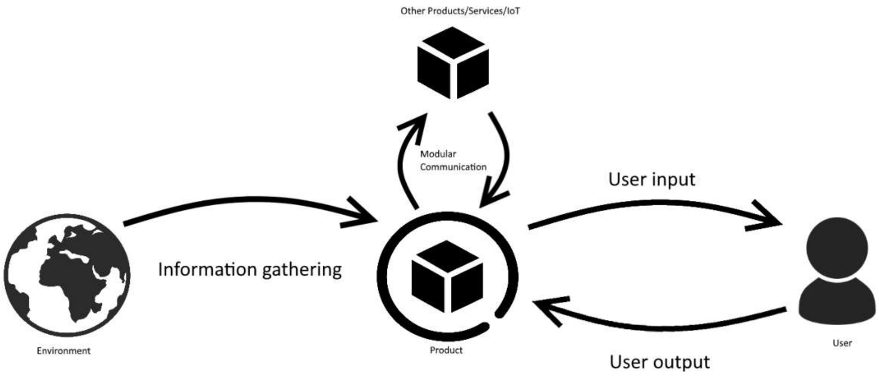


Figure 4

The product model described four entities that would be the foundation of the product.

These entities were;

1. User
2. Product
3. Environment
4. Other products and services

The product model also described the communication between the entities

The communication aspects with the product were the following:

1. Information gathering from the environment
2. Product output to user
3. User input to product
4. Communication with other services and products.

3.1.2 **User Studies plan**

With help from the product model an overview plan of the user studies was created. As suggested in the first step of data gathering in the theory part of this thesis, focus was on the goals of the user studies. Mainly on what information I wanted to gather.

Goals were created for two main categories; sessions with end-users and benchmarking. The goals were formulated as questions which queried for the desired data e.g. "What information can you not gather in your everyday life?".

The user sessions goals revolved around two things:

1. What information can you and can you not gather in your every day with and without the tools you use?

2. What tools do you have and always, sometimes, or never use and why?

The Benchmarking goals focused on:

1. What can and cannot existing products do?
2. Can they be divided into categories?

3.1.3 Work documents

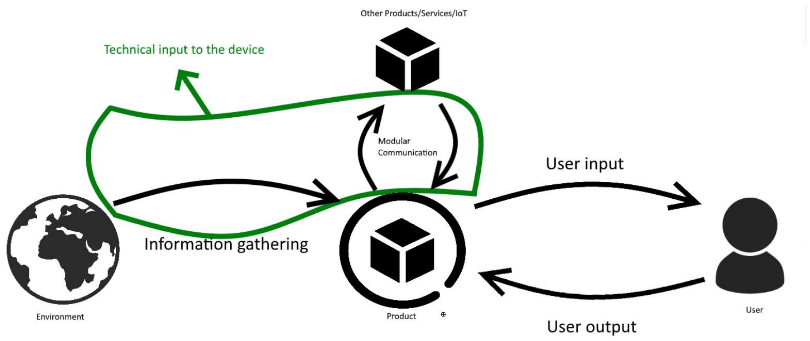
In the user study I used the principles of setting goals discussed in 2.2.1 and iteration in the design process discussed in section 2.3. I set a goal of producing a set of documents which would become the basis for deciding what features to be prototyped. I worked on these documents iteratively throughout the user study and added content and refined them as much as possible after each user study activity. This fitted the umbrella approach of starting broad and continuously narrow down the problem until I had enough information to decide on what feature to prototype. These documents are called Features, Output and input summary, and implementation options.

Features

The feature summary contained all features which I during the user study had discovered and could be of interest. I continuously revised the content throughout the process. Many features were renamed and grouped together to get a better result. This document is provided in Appendix A.

Output and Input

While the Features documents addressed the functionality important to the user, the Output and Input document focused on different types of methods to implement the features. Drawing inspiration from the product model I produced three categories in this document to be used for structuring the different methods. For *information gathering form the environment* and *communication with other services and products* from the product model, a category called *technical input to the device* was added. *Product output to the user* and *User input to the product* was added directly as two categories.



For each category I continuously added different technologies throughout the user studies and for each technology I took note of different advantages and disadvantages of using that technology.

For example, one of the technologies under the category *Output to user* was see through glasses. One advantage noted was that the user will have a better awareness of the environment compared to using screens that covers the eyes. A disadvantage was that it excluded people with no vision at all.

This document is provided in Appendix B.

Implementation options

The implementation options document was in a way, a merge of the two previous ones. The main effort on this was made during the later stages of the user study when a large portion of content had been added to the features document and the output and input document.

One headline was created for each of the categories in the features document. Live Image manipulation, Identifying, and information gathering. Then for each of these headlines, sub headlines were created for each category in the output and input document. For each of these sub headlines, technologies that could be used in this instance was added. For example, under the output to user part of the Live-image manipulation section See-through screens was an option.

This document is provided in Appendix C.

3.2 Sessions with Low Vision International

Throughout the process of gathering requirements, I had several sessions at LVI. These were more mentoring than information gathering in nature and carried out when needed. We met between each of the activities in the requirements gathering phase so that I could show the result and Henrik could provide guidance for the next activity. The meetings were informal and mostly consisted of Henrik helping me sort out my ideas and giving me insight into the industry. He also helped with planning the user studies and finding participants. We also communicated through email and phone calls between these meetings when I needed help with something minor.

There were some data gathering as well. We often used scenarios (2.5.2) to sort out ideas and generate new ones for coming data gathering activities. Simple personas were created when we wanted to explore different areas. Most often these were merely based on age and type of visual impairment. We then created scenarios for the area we wanted to explore and mentally put the different personas in these scenarios. Often the personas were given additional characteristics which would be relevant to the scenario. For example, how could an older person with tunnel vision use different technology for navigating through a workplace? An additional characteristic in this scenario were if the person would be new to the workplace.

3.3 Questionnaire

A questionnaire was created during the user studies. The sessions held with Iris Hadar generated would focus on qualitative data and I wanted to balance this with some quantitative as well. The focus of the questionnaire was existing tools and information that are hard to access in general for people with visual impairments.

Planning

The questionnaire was mainly created from the goals stated in the user study plan. As mentioned in 2.2.5, a pilot study should be carried out when doing questionnaires. However, a substitute for this can be to ask peers for input. This was done by asking Henrik at LVI to go through the questions. He came up with some improvements before the questionnaire was sent out.

Session

The questionnaire had a small explanatory introduction. This was created with inspiration from using a consent form (see 2.2.3). It briefly explained what the questionnaire was for. It also mentioned that the results can be published but will be anonymized to the degree that only age, and type of visual impairment will be linked to the answers.

The questions were heavily based on the goals in the user studies plan and can be summarized in the question below:

- What is your age and type of visual impairment?
- What information is hard to gather with/without tools?
- Which tools do you use, do you use a lot, don't you use, and which do you always carry with you when you are out? And why?

Analyzing

Four people answered the questionnaire. Not every question was answered by every participant. Here I will provide a summary of the results and the full results can be found in Appendix D.

The answers to what information are hard to ascertain without tools revolved around text information and obstacles while walking. There were similar answers to the question about what information was hard to ascertain with tools. One participant pointed out that he or she can gather most information with tools, but it takes time.

The most used tools were the white cane and different types of magnifying glasses. To the question about what makes the participant use some tools more than others, answers where about what type of situation the user was in. The answers were used to further add content to the feature summary.

3.4 Benchmarking

I wanted to benchmark already existing products to get a better understanding of what products already exists. I wanted to find out what functionality was missing and what functionality I might be able to improve.

Planning

When planning the Benchmarking session, the goals from the user studies plan was further developed in to three main goals;

1. Finding interesting features that could be implemented and/or improved upon.
2. Find interesting products which could be used as an external resource to the product.
3. Finding needs which the current products do not fulfill.

From these goals, a template to filled be out for each product was created (see Appendix E for an example). This template had the following items:

- Price and general description
- What are the products main features?
- How does it relate to the product model?
    - What information does it gather from the environment?
    - How does the user produce input to the device?
    - How does the user receive output from the device?
    - How does it communicate with other tools?
- What possible api´s or other ways are the to communicate with the product?
- What additional burdens comes with the product?

Session

Using the template, 15 different products was benchmarked. Information about the products came from various sources. Mostly from the manufacture's presentations about the product online. But also, from videos where the product was tested and presented, from youtube. For each product, one document was created from the template. An example of a document for one product can be found in Appendix E.

The products I have looked at are the following:

Cyber Eyez m300

Wearable glassess with a camera and screen that partially covers the users field of view. The screen can be positioned to reside in the desired position of the users field of view and can also be positioned on either right or left eye.

The main features are; offline and online OCR, Built in Alexa, scanning bar codes, color identifying, tell the mood of a person, object recognition, Skype and a flashlight.

Enhanced vision Jordy

Glasses with screens that covers the users field of view completely.

The main features are; magnifying, cctv, color select, freeze & focus, change brightness, HDMI-input for external sources.

Envision AI

An application for ios and android.

The main features are; OCR, magnifying, inverting colors, describing a scene, detect colors, scan barcodes, scan for and find objects.

Eyesight Eyewear 3

Glasses with screens that can be lowered or raised to cover the user's field of view completely, partially or not at all.

The main features are; Magnifying, changing contrast, freezing image, voice guidance during use, Airplay.

Iris Vision Live

A specialized version of the Samsung gear VR headset.

The main features are; Magnifying bubble and changing colors.

KNFB reader

Application that reads printed documents.

The main features are; OCR

NueEyes

Glasses with a camera and screens in front of the eyes.

The main features are; Changing contrast, changing color, OCR, Barcode scanner, Record video, Freeze image.

Orcam MyEye 2.0

Device with camera and speaker that can be fastened to a pair of regular glasses.

The main features are; OCR, tracking users finger pointing to a text, recognize faces, recognize products, recognize banknotes, recognize colors.

Samsung Relumino

A Samsung VR app.

The main features are; outline edges, magnify, change brightness, change contrast, different color filters, invert colors, screenshot, partial view bubble, partial view tunnel vision.

Seeing AI

An iphone app

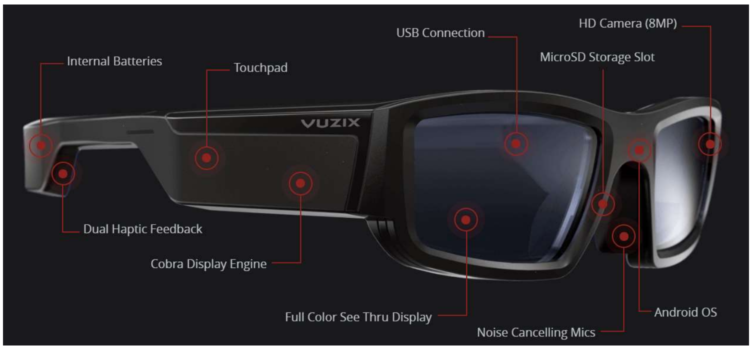
The main features are; scanning barcodes, OCR, recognize currency, describe a general scene, identifying a person, describe images.

VocalEyes

An object identifying app.

The main features are; object recognition and OCR.

Vuzix Blade



A pair of glasses with see-through screens.

The main features are; Apps can be installed for additional functionality.

Analyzing

The main features described above was added to the feature documents. These were the following:

- Inverting colors of a live image feed
- Changing brightness of a live image feed
- Changing color scheme of a live image feed
- Changing contrast of a live image feed
- Magnifying a live image feed

- Providing a magnifying bubble inside a live image feed
- Record a live image feed
- Freeze a live image feed
- Displaying certain parts of a live image feed on other locations of the users field of view.
- Scale down live image into the center of the field of view
- Text-to-voice documents
- Text-to voice hand writings
- Recognizing objects
- Recognize currency
- Scan barcodes
- Identify color
- Describing an image

The different types of input and output of the different devices was added to the output and input summary. The outputs were as follows:

- Screens in glasses that covers the users field of view completely
- Screens in glasses that covers the users field of view partially
- See-through screens
- Audio

The inputs were as follows:

- Physical buttons
- Hand gestures in front of camera
- Voice commands
- Touch screen
- Moving device around.

3.5 User studies with Iris Hadar

My user studies at Iris Hadar consisted of three sessions. The overarching method of these sessions followed the same umbrella principle as the rest of the thesis. In the first session, my scope was very broad but was narrowed down slowly towards the end of the third session. The insights gained from the sessions were used to further refine the work documents.

3.5.1 **Introductory meeting**

The goal of this first meeting was to establish a professional relationship with the participants and lay down a good foundation for future sessions. Some data gathering was done. Both to get the participants to start thinking about their needs, and to give me some initial data to work with.

Planning

During preparations for the first sessions with Iris Hadar, I focused on introducing me and my project, and getting to know the participants. I also prepared a boiled down version of the questions from the user studies plan. The idea was to use the time-dimension of Diary studies described in section 2.5.4 by introducing the participants to the topics of the user studies plan in an informal focus group. These topics could later be dived deeper into at the next session when the participants had had a couple of days to think about the topics.

Session

The session started with a meeting with the whole group over coffee and some sandwiches. During this I gave a small informal presentation about me and the master thesis. I also encouraged involvement from the participants by asking about the work Iris Hadar does.

After this, a focus group was held were discussion revolved around the topics brought up in the user studies plan. Then a brainstorming session was held to generate possible solutions to problems that had risen in the focus group. This structure was intentionally set up to be similar to the example under brainstorming 2.5.6 in the theory chapter.

Analyzing

No new input methods arose during the first session. Navigating and detecting obstacles was added as features. Haptic vibrations were added as output method. Some features were grouped together. Different ways of magnifying a live image was group in to *magnifying*. The different ways of changing colors, contrast and brightness etc. of a live image was grouped in to *changing image properties*. Scaling down a live image into the field of view and displaying a certain part of a live image in different areas of the users field of view became *partial view*.

3.5.2 **Continuing the data gathering**

During the second session with LVI, I focused on organizing the data gathered in a more manageable way for my purpose. Mainly by clarifying data from the previous session and probe deeper on ideas that came from previous session. In particular I wanted to further investigate different ways of analyzing text and objects.

Planning

During the second session I wanted to start the work on extracting the features I wanted to prototype. Some new group questions were created for the areas mentioned in the analyzing phase of the first session.

Most emphasis was however put on the individual interviews I was planning for the second session. These would be more directive than the focus group held in the first session but still contained some more open questions, hence they could be described as semi-directive interviews.

The open questions revolved around a regular day in the participants life and what hindrance their visual impairment brought to them in relation to this. Questions revolved around hobbies, evens/happenings and work life.

I also created some questions with the purpose of further refining the feature summary. By now, I had started to group similar features into different areas. Currently, the areas were Live image manipulation, identifying and navigation. A description of these areas can be found in 3.7.

Some of these questions was relevant for all areas while some were specific for certain areas. For example, a question relevant for all areas were “Have you used any product with similar functionality?”, while an example of a question specific to the identifying area is “In what types of situations do you have a hard time acquiring text information?”. Some questions with the purpose of gathering data for the output and input summary document was also created.

The questions which was relevant for all areas revolved around if the participant had used any tools with similar functionality and in what scenarios the participants could see themselves using the functionality.

Session

This session also started with a group meeting over coffee and some sandwiches. During this, another informal focus group, shorter this time, was held with focus on the new group questions.

After this I started preparing the participants for the individual interviews. This was done by giving a brief presentation of the different areas I had gathered.

Then three individual interviews were held. The goals were to interview more but the interviews took longer time than expected. I deduced that a large part of this was due to my inexperience as an interviewer. More time than planned was put on the open questions and I did not push forward to the next questions as quickly as I should have.

Analyzing

The focus of the analysis phase of this data gathering session was put on finding which features I had investigated would be most interesting to prototype. This was done by further refining the work documents *feature summary* and *input and output summary*. After this session, the features were grouped in to three groups. *Live image manipulation, identifying* and *Navigation.*

Live Image manipulation contained different ways to manipulate a live image e.g magnify, change contrast and change colors. *Identifying* contained three sub groups; OCR, Object recognition and Describe Scene. Navigating and detecting obstacles was grouped into *Navigating.*

After this session I decided to discard *Live Image Manipulation* as the area to implement. I could not come up with any significant improvement relating to this area compared to the products I had looked at during benchmarking.

At this point I also discarded the area *Navigation* as I had not come up with any concrete ideas for implementing a feature.

One issue that had risen with concerning products with features in the identifying group were the difficulties in aiming a camera at the object to be identified with adequate precision. This is something I wanted to investigate further during the next session.

3.5.3 More detailed data gathering

In the third session, most focus was on probing deeper on already discussed topics. This was the last session with LVI, and I wanted to gather enough data to be able to decide what to implement.

Planning

During the third session I also focused on extracting the features I wanted to prototype, this time in more detail than during the second session. Especially, I wanted to investigate if selecting a certain area to be analyzed could be done more efficiently. To do this, I created a Lo-Fi prototype, similar to a paper prototyped described in 2.5.5.

I experimented with this prototype myself as a preparation for the session and I came up with the idea of selecting an area by placing your index finger and thumb around the area to be analyzed. This I called “the crocodile hand”.

To further investigate the identification area, mainly OCR, I brought with me some everyday products to be “identified”. A bag of Whey protein powder and a can of Creatine.

Session

During the third session I held an individual interview with one of the participants which I did not have time for during the second. We also had a workshop with the prototype. During the individual interview, some of the questions from the previous session was asked but most time was spent letting the participant interact with the prototype. Most of this interaction was guided by different scenarios. Some were more direct where the participant was asked to use the products I had brought, and some were more abstract.

These scenarios were intentionally somewhat vague, an example of a scenario with the products given to the participant was; “ You want to know if you should buy this product, you also know that the prototype has a camera and functionality for helping you figuring it out, how would you proceed?”.

Then we held a dialogue where the participant explained her thought process and I asked probing questions and took notes. During the time the participant was interacting with the products and the prototype, I focused on the details of how the functionality could be implemented from the user´s perspective. For example, how could the user inform the prototype of what should be analyzed? I also wanted to test the usability of the crocodile hand.

The goal of the more abstract scenarios was to find instances where this product could be useful. For example, I asked the participant how this could be useful in different places like home, or in a store.

The workshop contained different activities like the one discussed in 2.5.9. One of these activities was to conduct a group interview with the questions I used in the individual interview. This time however, I took a step back and encouraged the group to discuss different ideas that came fourth.

To enhance the creativity, halfway into the group interview, I introduced a new question. “Can we build another prototype that looks completely different and better implements the features we are discussing?”. This sparked new discussions and we effectively went back and forth between this question and the previous ones and generated new ideas for both.

Analyzing

I received an initially positive result of using the hand gesture to select an area to be analyzed. When interacting with the Lo-Fi prototype and asked, “How will you show the product what to read on the bag?” one participant put the index finger and the thumb over the area to be read. During the group interview when I put forth the “crocodile gesture” the participants responded positively. I specifically asked if they thought they would be able to position the hand in a sufficient way and they agreed that they would. This was however not possible to fully test with the Lo-Fi prototype.

While experimenting with the lo-fi prototype in the more creative way, I realized that the product would require some sort of activation functionality for knowing when to take a shot of the selected area. The solution I came up with was to start with the index finger and thumb closed. Then separating them until the desired area had been chosen, at which point, the user would freeze the fingers in that position.

3.6 Lo-Fi prototype

Using the paper-prototype method described in 2.5.5, I created a Lo-Fi prototype. This prototype consisted of a metal box with 2 coins taped to it. A lanyard was also fastened to it. The coins represented a camera lens. When studying the theory for this master thesis, I found that communicating functionality to the user has a heavy focus during the design process. For example, one important purpose with the conceptual model is communicating to the user what to do (2.4). Therefor the main purpose of this lens was to more strongly communicate the functionality to the user. I also wanted to investigate one possible issue I had discovered. This issue was uncertainty of the capability of the user to capture the information to be analyzed within the cameras field of view. In other words, will a user with visual impairments be able to with enough precision direct the device at the correct angel?

As mentioned in 2.5.5, Lo-Fi prototyping affords focus on the design rather than implementation. I wanted this focus for both me when working with the design process, and for the participants during the user studies. Therefore, this Lo-Fi prototype was created. An image of the Lo-Fi prototype cane be seen in Figure 55.

Figure 5

3.7 Result

The way the results of the user studies were documented was the work documents described in 3.1.3. As mentioned, these documents grew to a finished result from iteratively refining them during the User studies. They are provided in appendix A, B and C. Below is a summary of these documents.

Features document

The main categories which the features document contained was called *Live image manipulation*, *Identifying* and *Navigation*.

For the *live image manipulation* area, I tried to gather what different ways the participant would like to manipulate a live image of their surroundings e.g. magnify, change contrast etc.

The area called *identifying* was divided in to two subcategories, OCR and object recognition. In the case of OCR, the goal was to ascertain in what type of scenarios it would be useful. In relation to Object recognition, the goal was also to find out in what type of scenarios it would be useful. Some additional focus was put on finding out if it would be possible to create a set of standard items that would encapsulate as many use cases as possible. And if there was some need for users to create a set of custom items which they could identify.

The area named navigation also had two subcategories. The first one concerned obstacles in the users way when walking. A scenario when the user was not carrying a cane and a scenario where the user was carrying a cane was considered. The second one was related to navigating in a sense of the user finding the correct path to their goal.

Output and Input document

As mentioned in 3.1.3 this document was divided into *Output to user*, *User Input* and *Technical input to the device.*

Under *Output to user* different types of screens in a glasses-like product, Audio and Haptic vibrations where analyzed. The main downside of the screens was that they cannot be used by people with full vision impairment. Audio would be easy to implement but has low bandwidth. It also removes auditory feedback from the environment which is extra important for people with visual impairments. Haptic vibrations do not have the issue with disturbing auditory feedback but might have a low bandwidth as well. As this was not the feature I chose to prototype, I did not investigate deeply into how advanced vibration patterns could be while still maintaining usability.

For the User input I considered physical buttons, hand gestures, voice commands, touch screens and moving the device around. Physical buttons have the advantage of providing physical feedback to the user, which becomes extra useful for people with visual impairments. I did however draw the conclusion that it has bandwidth limitations. Hand gestures have higher bandwidth capabilities than the physical buttons but is weaker in terms of fault tolerance and reliability. Voice commands are easy to use but I estimated that it would be even weaker than hand gestures in terms of fault tolerance and reliability, especially in noisy environments. The different ways people pronounce words for example, is a big problem for this technology. Touch screens are hard for people with visual impairments to use since they have trouble receiving the visual feedback it provides. I estimated that moving a device around could be more reliable than hand gestures in front of a camera. The use of an accelerometer and gyroscope does not require the device to be pointed in the correct angle like a camera for example. The bandwidth depends on the task to be performed. For example, the functionality of marking an area is easier with a hand gesture, especially when moving the fingers is included under hand gestures.

When investigating Technical input to the device I looked at three different ways. Camera, HDMI and output from other devices or applications. These were the different types of technological input I had found during my benchmarking. This section did not result in as clear advantages and disadvantages compared to each other. I estimated that this section more heavily was influenced by what functionality should be implemented. For example, the HDMI input could have been useful for a screen with image manipulation capabilities which is connected to a pc but would be rather useless if the other device was stationary and the feature to implement was warning for obstacles when walking to the store.

Implementation options document

As mentioned in 3.1.3, the implementation options document was a way to merge the two previous documents. It described how the different areas in the *Features* document could be implemented.

The output to user of the Live image manipulation was suggested to be done with some type of the screens discussed in the *Output and input* document. Even though these exclude users with grave vision impairments, these users have no use of the feature so that would not be a problem. I deemed the best user input method for this area was moving the device by aiming it at the object or area that should be manipulated. Other types of input would require some sort of motor on the device which would position the camera according to other input. This would be more difficult from an implementation perspective and not as intuitive for a user. If functionality like switching to a specific color filter was requested this should be done with physical buttons as this requires low bandwidth.

For the identifying area, the right method for output to user was not as clear. I looked at both Audio and haptic vibrations. Even though I drew the conclusions that both audio and haptic vibrations have low bandwidth in the *Output and Input document*, this also depends on what type of information it should communicate. For instance, considering the scenario of a user shopping in a store and wants to use a device to determine which products to put in the shopping basket.

If the user has a predetermined list of items to purchase and simply needs to confirm by scanning a barcode if the item scanned is on the purchase list, both a small haptic vibration and an auditory que like a bell could be use. In this instance, the haptic vibrations have the advantage of not disturbing other people in the vicinity and hence provide a more pleasing user experience.

However, if the user does not have a predetermined list and information communicated to the user must be more complex, e.g. the name of the product and its weight, a haptic vibration scheme for providing this complex information would not be usable. Hence, audio would be better in that instance.

A general conclusion that I drew from this is that if the information that can be produced to the user can be limited to a small set of instructions, haptic vibrations should be used. Otherwise, audio is better. This was relevant to identifying certain things in the surroundings under the identifying area in the *features* document.

For the navigation functionality, most focus was put on how output to the user should be implemented. Audio, haptic vibrations and screens was considered and the same advantages and disadvantages as before relating of these techniques was considered.

3.8 Deciding features to implement

My decision of what to implement grew continuously from the user studies. My decision was based on what I estimated to be within my time frame and what could be described as a simple formula; *The positive impact it will have on the users lives times the number of users it would affect.* I looked at the features document and the three categories; Live-image manipulation, identifying and navigation. During the user studies, the issue of costs of different tools had come to my attention. Many tools are excluded from a large part of visually impaired people due to the cost.

Every tool that was similar to what my ideas was for the Live-image manipulation area had been very expensive and my initial ideas for this area had only limited improvement capabilities compared to what already existed, so I did not choose anything in this area. The cost would make it accessible only to a small number of users. One note here is that I never investigated if there were a way to reduce the cost to make it more accessible.

I had no concrete ideas for the navigation area. This led me to settle on the identifying area where my discoveries on using hand gestures for this could be implemented. I estimated that this might be useful for users with varying degrees of visual impairments. Features within the identifying area in the features document could be useful for people with no vision at all. It could also possibly be implemented as an app on a phone which would eliminate the hardware costs since most user´s would already have a smartphone. Another reason for choosing the identifying area was that many of specific examples the participants came up with about what was hard for them could be solved by implementing products within this area. Some of these examples were; identifying colors when doing laundry, choosing correct product in store or when doing things at home like baking or preparing food, and figure out what settings the stove was turned to. The same was true for the navigation area, like for example navigating in crowded areas, but I did not find ways to solve as many needs in this area.

The feature I decided to implement was using the gesture of marking a certain area with your index finger and thumb. This area would then be captured and could be analyzed in different ways. This was new compared to the other products I had found due to the ability to only select a part of a cameras field of view. This would also be useful for many different types of identifying functionality.

Relating this to the product model, the idea was that in the future, this product could capture many different forms of information and produce to the user in an appropriate manner. The device would capture some form of information, e.g. a text, and produce to the user in an appropriate manner, e.g audio. I imagined if I could get this basic model to work, more advanced functionality could be added later. The goal from the beginning was therefore to implement both the gathering part, i.e. the gesture, and the output to the user, i.e. the reading out loud of the captured text.

However, the ocr module which I wanted to use turned out to perform very poorly on the images captured with my webcam. I did not have time to investigate this problem, so the prototype was only implemented with the gesture functionality.

4 Functional prototype

In this chapter I will describe the implementation of the functional prototype. The first section provides some information on how I decided what method and framework to use. The second describes in general terms how the prototype works both in words and with a visual explanation. I also provide some results in the third section.

.

4.1 Planning the implementation

I experimented with different functionality in opencv. Most experimentation was unsuccessful. At first, I tried to implement the feature by recognizing the hand using object identification, due to my limited time I gave this up as I did not progress fast enough. I also looked in to using background reduction combined with using different ways to blur the image, e.g. Gaussian Blur, but I did not make progress quick enough, so I discarded this method as well.

I settled on using object tracking by color in the HSV-format to be able to test the functionality. The goal was to implement the text to speech feature as well as the hand gesture controls but the module which I planned on using called pytesseract, did not do as well with the images from my program as the ones I had tested with before I started my prototype. Due to time limitations, I did not investigate further into this and only included the hand gesture functionality in the finished prototype.

4.2 Implementation

The functional prototype was implemented by recording video of the hand gesture. To be able to find the position of the fingers, different colored stickers on the top of the index finger and the thumb. The software tracked the movement of the thumb and the index finger based on these colors. These movements were then used to capture a certain part of the camera which should be analyzed.

4.2.1 Tracking the movement

Following the documentation for contours in opencv [4], to record the movement of the fingers, the following algorithm was used.

1. Convert each frame captured to HSV-color format
2. Create two masks, one for each color, which contains a lower and upper threshold for HSV-values close to the different colors wished to track.
3. Apply masks to the frame separately and color the pixels within the threshold white while all others are colored black.
4. Find the different contours in the image which encapsulate the different white areas.
5. Then to remove noise and only focus on the desired area want, i.e the tip of the finger, the contour with the largest area is selected and everything else is discarded.
6. A bounding box is then placed around the largest contour.

4.2.2 Using the movement

1. Capture the different positions of the fingers for the last 50 frames in a buffer
2. Calculate average distance between the fingers between the frames to determine if the fingers is still.
3. If the fingers are not moving, determine if they are in a closed or open position by using the distance between dem.
4. The whole gesture is then captured by looking for the pattern: “closed”, “moving”, “open”
5. When this pattern is detected, an area is captured as an image. This area is the distance between the fingers in length on the y-axis and 1.5 this length on the x-axis originating from the upper left corner of the top finger.

4.2.3 **Visual explanations**

The goal of this section is to show a visual explanation of how the software works. The following images is extracted in debugging the software, they are not part of what any user will see in the prototype.

In an instance where the finished product would be deployed on a pair of smart glasses, there might be some usability gained from showing the rectangle in the second image but for the purposes of testing with this prototype, this is only an explanation of how the software works.

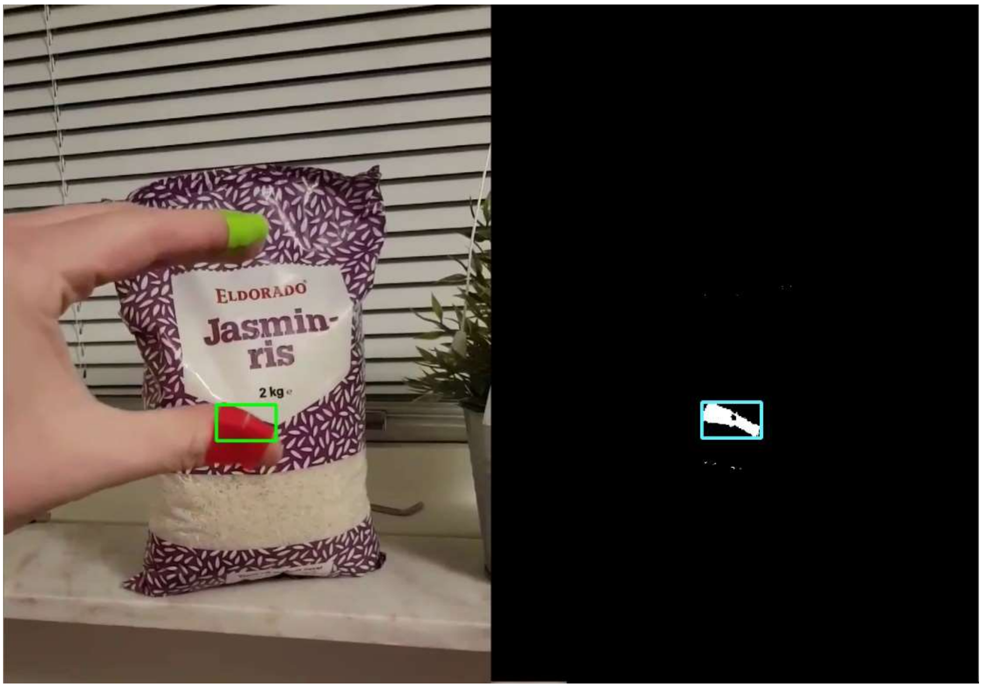


Figure 6

Figure 66 shows how tracking the fingers work. The left part is the recording of the webcam, the right part is the mask which sorts out the colors we do not want to track. The neon rectangle is the bounding box marking the area where the red sticker was detected. The turquoise rectangle has been added manually to better describe what is happening. The same is done for the neon sticker with its own mask.

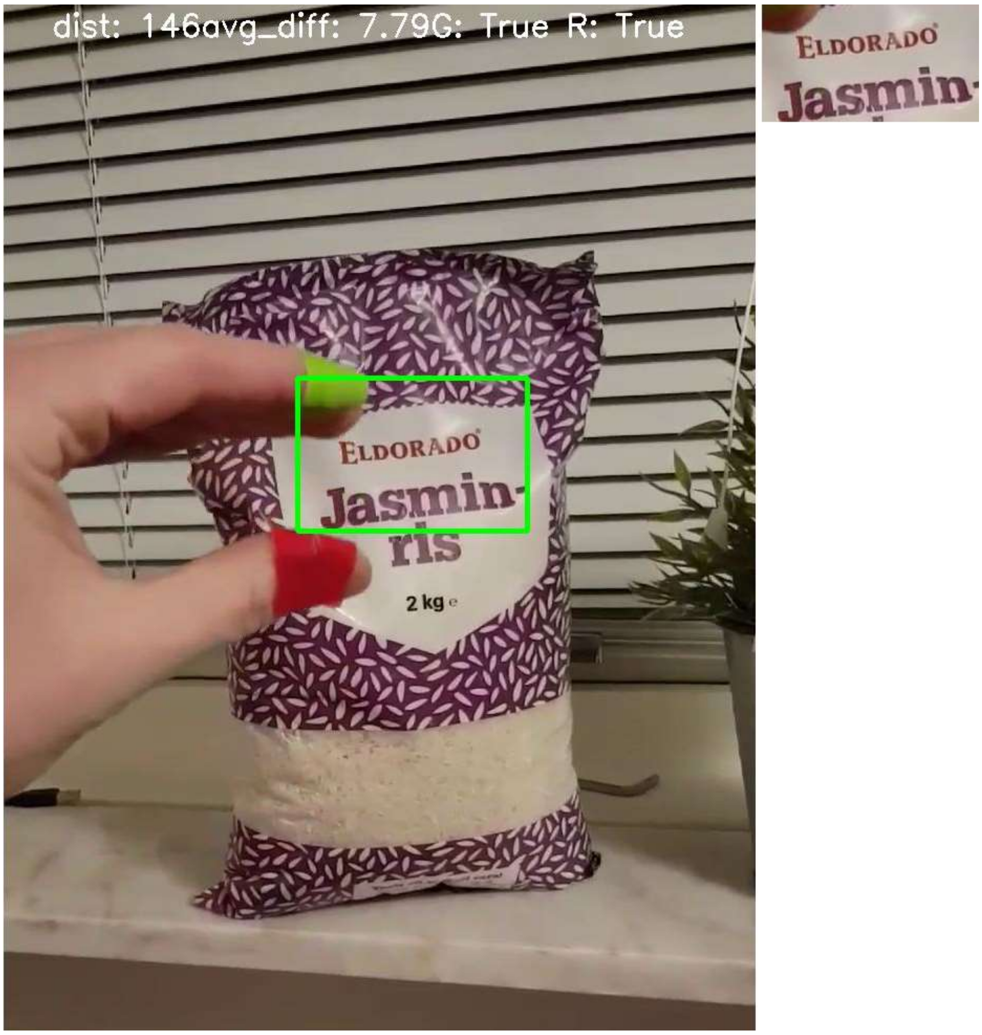


Figure 7

Figure 77 shows how the area is selected with the fingertips. The neon box in this image is not the same as the neon box in the previous image. This neon box shows the area which will be captured. To extract the requested area, the coordinates of the box are stored at first. Then the matrix with the image values which resides within these coordinates are simply copied to a new matrix which becomes the image in the top right corner of Figure 77.

4.3 Result

The way the functionality was implemented makes is unfeasible for a market ready product. I underestimated how much different light conditions drastically changed color values of the stickers on the fingers. When researching this method, I had found very little information about this. All I had found was recommendations to use the HSV-color space as this would be more immune to this compared to RGB. To be able to perform the usability tests, I needed to implement a calibrator with which I manually tuned the thresholds for before each test.

The first version of the product was very unstable, it would trigger constantly due to color noise which resided within the set threshold. Selecting the contour with the largest area remedied this heavily but there was still an issue of false triggers when the fingers was not present in the image. This was remedied by adding a threshold for the size of the area which could trigger the area selection. However, under certain light conditions which made colors in the surrounding environment close to the colors of the stickers, the product was still unstable.

Due to time limitations, this was done with a webcam connected to a computer. The goal was to implement this on a pair of smart glasses which LVI ordered but shipping delays caused them to arrive too late.

5 Evaluation

In this section I will describe the evaluation of the Hi-Fi prototype. The first section describes the use cases which were used in the following three sections. First a heuristic evaluation was made to find weak points in the prototype from a theoretical perspective. Then I performed a cognitive walkthrough to see if I could find any issues which could be interesting to keep in mind when evaluating with users. The user tests ran in to problems due to stability of the prototype which lead me to deem them inconclusive.

5.1 The use cases

For the different evaluation techniques, I used the same use cases. I first explicated the users. The users of this product are people with some degree of visual impairment. They vary in age and technological abilities. To be able encapsulate a large portion of the target user I created two personas that would come to reflect two different user segments.

A 20- year old person that had used visual aid tools all their life and a 60 year-old who´s visual impairment had affected them recently. The goal was to make the evaluation encompass users with different technology capabilities.

Two use cases were created, one where the user's goal was to analyze a moveable object, in this case, a small bottle about 8 cm high. The other use case was for analyzing the surroundings in the user's home, in this case, what temperature a stove knob was set to. This use case was inspired by one of the participants in the user studies who loves to bake but had large problems with the stove settings.

The first use case was applied in two different contexts, one where the moveable object is standing on a surface e.g a table, and the other, where the user holds the bottle in their hand.

5.2 Heuristic evaluation

For the Heuristic evaluation I used the heuristics described in User Study Guidelines under heuristic evaluation [1, p.39].

Here follows the heuristics and the conclusion from each one. They are prioritized in the order they are mentioned.

Visibility of system status

There are some limitations on this when the users are visually impaired. Progress bars and other common techniques is sub optimal. I therefore I decided on implementing sound clips that gives feedback on where in the activation sequence the user is.

User control and freedom

The user can cancel a started activation sequence by closing the fingers and achieve the “closed” state.

Error prevention

Since the prototype identifies the fingers based on the largest area which is within the hsv-treshold, when no fingers are present some noise might be big enough to trigger the activation sequence. A solution to this is discussed in the 5.1 implementation improvement possibilities.

Aesthetic and minimalist design

The prototype achieves a minimalistic design. No redundant information is presented to the user.

Help and documentation

Not created but should be available for finished product.

Recognition rather than recall

Would be useful to have some instruction on how the activation gesture works but the nature of the prototype does not allow for this in any way I have found. The lack of this might be mitigated with a good conceptual model.

Flexibility and efficiency of use

This prototype only has one function so not very applicable.

Match between system and real world

Not very applicable, there is not much information to be presented to the user in this prototype.

Consistency and standards

The prototype is not part of any other platform in a way that would affect the user. In relation to the finished product however, this might be relevant if for example implemented using an android app.

5.3 Cognitive walkthrough

The cognitive walkthrough (2.5.10) was performed by my me using the use cases described in 5.1. No new major defects in the product was discovered. The issue of the camera being directed in the correct angle was noted but this has been discussed before. When walking through the use case with the stove knob, I realized it might difficult to solely capture the desired area.

5.4 Controlled usability tests in the lab

To be able to simulate the experiences of a first-time user, I chose not to conduct the usability tests with Iris Hadar. I had discussed the functionality in depth during the user studies. This prior knowledge might have made it easier for them to use the product.

Since I did not have any access to eligible participants with visual impairments, I tested it on three friends instead. These friends had no visual impairments, but a blindfold was used to somewhat simulate the experience of a user with visual impairments. This was a compromise I had to make to be able to conduct the usability tests. As this was done in their homes, I discarded the use case with the stove since some difference between their stoves might have affected the result.

As mentioned, focus in the evaluation was on the accuracy and the usability of the functionality. To determine the accuracy of the method, I decided to record how many times out 5 tries would the capture be successful. To determine if it was intuitive, first-time users were at first given limited instructions.

The instructions were divided into two parts which varied in detail were I first tried a short description and if the user did not understand how to solve the task, a more detailed description was provided.

The first instruction was “The camera can detect where the stickers on your fingers are located. You want to mark the area you want to analyze with a gesture. This gesture is opening your fingers from a closed position to a still open position.”

The more detailed instruction included an explanation of how the product uses the activation sequence “closed”, “moving” and “open. It also included explaining that the closer the fingers were held to the camera; the larger area would be captured. After this more detailed instruction had been tried, I gave correcting instructions on what they did wrong.

Results

The results were heavily distorted due to the instability of the product. As mentioned, to assess accuracy of the method, I tried to record how many attempts out of five would be successful. Most failed attempts were caused by instability in the product and not wrongful use of the prototype. This led to the number of failed attempts not showing the usability of the product. I tried to assess how many of the failed attempts was caused by instability and how many was caused by the way the participant used the product. However, I could not determine this accurately. Hence, this data has been omitted. I could however see some difficulties the users had with using the product and these will be described below.

One common problem for participants when operating with few instructions was that they held their fingers to close to the object that should be analyzed, which lead to the image captured not covering the whole bottle. This was improved after receiving the expanded instructions. However, it still was a problem.

Another difficulty discovered was the understanding of how the bounding box which selects the area to be analyzed, was created. The way the prototype works, the bounding box appear to the right of the fingers. The fingertips create the top left and bottom left corner of the box. The participants seemed however to try to position the object to be analyzed between the fingertips.

6 Discussion

In this part I will discuss this master thesis in general terms. I will provide some conclusions I have drawn from the user studies method I used, and my thoughts on how the implementation of the prototype can be improved and viable in a real product. I will also present a section where I discuss the conclusions I have drawn about the usability of the gesture.

6.1 Reflections on the user studies

The method of working iteratively synergized well with my way of documenting the progress in the work documents. One of the large benefits I experienced was the way it helped me conduct the user studies. When planning the user studies, the goal of further refining the work documents simplified the process of planning the activities in the user studies. I could look at the documents and ask myself, "is this enough to decide what to prototype? What is missing?" and plan according to the answers to these questions.

As mentioned in 3.5, my inexperience as an interviewer lead to large portions of the interviews being spent discussing things which not progressed the user studies. I imagine this is somewhat always true in all user studies conducted, every word from the participants will not be relevant, but still, I believe this was more prevalent my case. Afterwards, when reflecting on the sessions, I realized that asking myself the question "does this contribute to the work documents", could have guided me away from these irrelevant topics.

The use of the product model helped me a lot during the user studies. Creating it in the beginning caused me as a designer to really think about the different aspects of the product. It also helped me when eliciting the different requirements during the user studies, especially during the bench marking. I could quickly assess which features could work together and which where new or small variations on something I had already seen.

One discovery I made when working with the work documents was related to bandwidth of different output methods to the user. I had assessed bandwidth capabilities on different output methods without considering what functionality they would be used with. Different methods of output to user synergize better with some functionality by providing higher bandwidth on said functionality. Take for example warning the user of an obstacle slightly to the left of their directional path. An intuitive way to communicate this to the user would be through a vibration applied to the body of the user in the direction of the obstacle. In this scenario this seems better than doing it through audio. For example, hearing "obstacle at 30 degrees to your left" would not be as simple to understand. However, in the case of communicating product name and weight in the scenario where the user is shopping, audio would probably be better. In other words, I specified bandwidth capabilities to early when I worked with the *output and input* document since I did not at that time know what functionality the techniques would implement.

6.2 Implementation improvement possibilities

The manual calibration of the color-threshold needs to be done automatically. This might be done with some sort of sampling of the current light conditions which then could be used to offset the light differences.

The issue of the implementation was the poor accuracy of tracking the fingers. As it was implemented, a combination of color threshold and size of the contours was used to determine the correct position. Using more ways of tracking the fingers and then draw a more educated conclusion from the different ways could result in a more precise tracking.

One suggestion to these ways is to use thresholds in more color spaces, e.g rgb. The use of background reduction might also be a useful technique. Background reduction works by removing the parts of a video that is still. If only the fingers are moving, I hypothesized it should be easy to sort them out from the rest of the image. However, I had not taken in to account the small movements caused by the fact that the camera is attached to the body when experimenting with this. These where large enough to cause too much noise in the image for the method to be ineffective. There might be some calibrating of the parameters, e.g how sensitive the algorithm is for movement, which would make it an effective method in combination with others.

My first idea was to use a type of object identifying method using convolutional networks. Time and resource limitation caused me to not be able to try this implementation as I would have needed to train a new network to recognize the pattern. After experimenting with the other methods, I believe that they are a more effective way, especially if the program should be able to run on portable hardware with a small computational power.

6.3 Usability

The usability was heavily affected by the instability in the product, this is described in more detail in 5.4. However, to assess the validity of the functionality of the prototype from a user perspective, this part of the discussion section tries to draw some conclusions from a scenario where the prototype was stable.

Starting with the first of the two simplified design principles defined in section 2.5.1 ;

Intuitive use: The product shall be easy to understand for as many user´s as possible and provide a lot of feedback and clear information to the user. It shall also be physically easy to use.

On a basic level, the functionality is easy to understand. You mark the section you want to analyze with your fingers. The stickers which was introduced to make the implementation easier, made this concept easier for the user to understand. During one of the sessions with Iris Hadar, I verbally tried to explain the concept of marking an area with your fingers. It was not until I explicitly guided them through the gesture that they understood what I meant. However, during the user studies, even with the limited instructions in the beginning, all participants understood the general idea. The hardship could be found in where to position the fingers when you opened them, e.g. that they should be to the left of the object. An effective way to teach this to the user would be to show the image in figure 2 from the debugging which shows the green rectangle that marks the coordinates. As the target user of this product is people with visual impairments, this is not optimal. It would work for users which have lost parts of their visions, but not those with grave or full impairments.

I looked in to creating a partial conceptual model like described in section .1.3.1 but could not come up with a suitable metaphor. I looked in to using something along the lines of snipping or pinching an area, but after trying the idea on people with no knowledge of the product, it seemed like it gave no cognitive value.

I deem limitations in accessibility low. Inexperience with technology would not bring with any large complications, marking an area with your fingers to a camera is not in its essence a complicated idea. The only thing I can come up with is the issue of positioning the fingers at a good length from the device which came up during the evaluation might be bigger for a user with low technical understanding. The user also needs to have one functioning hand with moderate dexterity.

After adding the sound ques for triggering the *closed* part of the activation sequence and the shot being taken, some user feedback is present in the system. I experimented with using a sound clip from an old camera as audio que for the moving part of the activation sequence but after trying it out myself, I found that it caused more distraction than proper feedback to the user in a cognitive sense.

I would also deem the product physically easy to use, the issue of the length between the fingers and the camera could fall under this category but otherwise I have found no physical difficulties in using the product.

Tolerance of error: The product should be able to handle user and other errors with as little manual error correction as possible.

As the functionality prototyped only was the marking of the area to be analyzed, no evaluation was made on the future product which would use it. Imaging for example text to speech features on text in the selected area, another gesture was discussed during the user studies which never made it into the prototype. I discovered that a regular issue with products offering text to speech functionality, is that the user needs to finish listening to a text which has been chosen to be read. A deactivating gesture of closing the fingers again or shaking the fingers was discussed but as mentioned, this was never tested.

7 Conclusions

In this section I will try to assess if my goal was successful and if the contributions I aimed to make was made.

Starting with the goal of investigating new ways of using technologies to enable people with visual impairments better access information in their environment. I would say the thesis have been semi-successful in achieving this. Various functionality and ways to implement said functionality has been investigated. However, only the surface of most of these areas has been scratched. There are many interesting topics which I would have liked to delve deeper into.

In the area I called live image manipulation I merely gathered all the different ways of manipulating a live image I found. I also found a need for navigational support. The white cane does miss certain obstacles, especially concerning obstacles on different heights. The area I chose to prototype was the one I called identifying. This revolved around analyzing different visual information in the environment, for example text on documents. The finished prototype only contained the functionality to select a certain area to be analyzed.

Regarding the contributions I wanted to make to the industry I would say I have been semi-successful as well. Starting of with providing an example of how to design a prototype for beginners. The design process I have followed in this thesis is simple and easy to follow. 2.1 Design process, 2.2 Data gathering, 2.3 User-Centered approach and 2.4 Conceptual model gives a brief overview to what I found to be the most useful theories of interaction design in my thesis. I chose techniques of varying nature to show different methods of data gathering, these can be exchanged for something else and still fit my example of a design process. The way I chose to record the results from the user studies by continuously refining a couple of documents synergized well the iterative nature of a user centered approach to design. It especially helped me with creating specific goals for each data gathering session, e.g. explore what other types of image manipulation could be useful in a product which manipulates a live image.

One thing I failed at was to find a concrete method for implementing a prototype. I just started to experiment with different ways of identifying the fingers and the whole process seemed somewhat arbitrary. It seems like creating a method for implementing all types of prototypes is hard. The method heavily depends on what functionality is to be implemented. There can however be guides or best practices for implementing certain types of functionality. In the case of object tracking based on colors, the creators of opencv provided examples of how their framework could be used for this. To find ways to implement functionality, investigating how similar functionality has been implemented can be useful.

Regarding the aim to create a prototype good enough to develop a market ready product from, I would say have not been fulfilled. The stability of the prototype makes it unusable and I have not managed to show that it is viable from a technical perspective. However, I have produced a way to implement the functionality if it can be made stable enough and I have also come up with suggestions which could improve the stability. The usability tests were inconclusive, there is not enough data to prove that it is viable from a usability perspective and more tests needs to be done with a more stable prototype.

8 Citerade verk

[1] C. Magnusson, K. Rassmus-Gröhn, K. Tollmar och E. Deaner, User Study Guidelines, Hapitmap, 2009.

[2] J. Preece, Y. Rogers och H. Sharp, Interaction Design: Beyond Human-Computer Interaction 4 edition, Wiley, 2015.

[3] R. Jafri, S. A. Ali, H. R. Arabnia och S. Fatima, ”Computer vision-based object recognition for the visually impaired in an indoors environment: a survey,” Springer-Verlag, Berlin, 2013.

[4] O. team, ”Contours in OpenCV,” OpenCv, 8 December 2015. [Online]. Available: https://docs.opencv.org/3.1.0/d3/d05/tutorial_py_table_of_contents_contours.html. [Använd 27 Juni 2019].

Appendix A Features Document

A.1 Live-image manipulation

- Invert Colors
- Brightness
- Color
- Black and While
- Outline edges
- Contrast
- Magnify
- Magnify part of screen
- Filters
- Freeze image
- Record video
- Partial view bubble
- Partial view Tunnel vision

A.2 Identifying

- OCR – Read text aloud
    - Instantly
        - Read instantly from a certain area
    - Scan document
        - Help with taking picture
        - “Session “with multiple pages scanned
        - Different parts of text e.g. paragraphs can be skipped etc.
    - Handwriting
- Object recognition
    - Recognize custom objects
    - Recognize standard/common objects
    - Recognize currency
    - Identify color

- Describe Scene
    - Take picture
    - From image in apps

A.3 Navigation

- Detect Obstacles
- Find correct path

-

Appendix B Output and Input document

B.1 Output to user

B.1.1 Screens in glasses that cover the user's field of view completely

Pros:

Allows for "clearer" image compared to partial view or see-through screens due to no "noise light "or user distractions from environment.

Easy to implement with existing VR-headset technology

Cheaper than see-through screens?

Cons

Limits the user's awareness of surroundings. Can be hard/impossible to use while walking around.

Does not work for users with extensive visual impairments (fully or close to fully blind).

B.1.2 See-through screens in glasses

Pros:

Give better awareness of surroundings compared to not see-though screens.

Cons:

Might be more challenging to implement from a hardware and cost perspective

Does not work for users with extensive visual impairments (fully or close to fully blind).

B.1.3 Screens that partially covers the user's field of view

Pros:

Better awareness of surroundings compared to screens that fully covers the user's field of view.

With the right physical implementation, the screen can be moved around in the user's field of view and hence used when walking around.

Easier to implement from a hardware perspective compared to see-through screens.

Cheaper then see-through screens?

Cons:

Not as good awareness of the surroundings as see-through.

Not the same "viewing capabilities" as full covering screens.

Does not work for users with extensive visual impairments (fully or close to fully blind).

B.1.4 Audio (information read to the user)

Pros:

Very easy to implement.

Works for users with extensive visual impairments.

Cons:

Low bandwidth

Removes, at least partially, the awareness of the surroundings a user receives from their hearing. Might be extra hindering for people with visual impairments, since this part of the population in higher degree relies on their hearing in everyday life.

If read aloud, users feel embarrassed or annoying to their surrounding

B.1.5 Haptic vibrations

Pros:

Removes little or none awareness of the surroundings (maybe some if the vibrations are distracting).

Should be relatively cheap to implement depending on how it should function. Vibration modules are cheap.

Easily implemented from a software perspective. All that is needed is pulse length, time and repetitions.

Cons:

Low bandwidth.

Learning curve for more advanced vibration patterns.

B.2 User Input

B.2.1 Physical buttons

Pros:

Very easy to use, especially for people with visual impairments because of the additional physical feedback

Cons:

Low Bandwidth

B.2.2 Gestures with hands in front of camera

Pros:

Easy to use (if it works)

Possible higher bandwidth than physical buttons

Cons:

Reliability and fault tolerance might be low. Further investigation possibly needed

Possibly computationally heavy. Further investigation might be needed.

Learning curve for more advance patterns. Might be mitigated with intuitive gestures.

B.2.3 Voice commands

Pros:

Easy to use, commands can be made very intuitive.

Cons:

Low bandwidth.

Draws attention

Reliability and fault tolerance might be low. Further investigation possibly needed

Possibly computationally heavy. Further investigation might be needed.

B.2.4 Touchscreen

Pros:

Probably easy to implement, a simple app for android and ios can do a lot.

High bandwidth

Low computational complexity.

Cons:

Hard to use for many people with certain visual impairments.

B.2.5 Moving device (e.g. head gestures for devices worn like glasses or maybe a "magic want")

Pros:

Probably more reliable than hand gestures in front of a camera. Further investigation might be needed into sensors.

Not computationally heavy.

Cons:

Low bandwidth for simple movements.

More advanced patterns might not be very reliable or fault tolerant.

Could be more complex to implement from a hardware perspective.

B.3 Technical input to the device

B.3.1 Video/image through camera

Pros:

Very high bandwidth

Cons:

In relation to this, some features might require complex implementations and heavy computations. Image recognition for example.

B.3.2 HDMI

Pros:

Very high bandwidth.

Easy to implement.

Cons:

Might not be very useful. Could be useful if other devices/services should relay video to the device and the device itself only focuses on other aspects, like presenting it to the user.

B.3.3 Output from other devices/applications/services

Pros:

Overall, good shortcut to interesting features. Further investigation needed.

Cons:

Might not be accessible in desired way.

Might be restrictions in form of copyright etc.

Appendix C Implementation options document

C.1 Live-image manipulation:

C.1.1 Output to user:

Can be done with:

1. See-though screens
2. Screens that cover the users' field of view completely
3. Screen that partially cover users' field of view
4. External screen

Notes:

Hardware on see-through might be inadequate for premium result. What benefits and disadvantages does an external screen have?

C.1.2 Input from user:

Can be done with:

1. Physical buttons
2. Gestures in front of camera
3. Voice commands
4. Touch screen
5. Moving device

Aiming device at area which should be recorded seems like the best option. Controlling different filters etc. should be used with physical buttons. I assess that the input in this feature requires only low bandwidth e.g. "change filter". Gestures in front of camera might be ideal as well.

C.1.3 Technical input to device

Can be done with:

1. Video camera

C.2 Identifying:

OCR, Object recognition, Describe scene

C.2.1 Output to user:

Can be done with:

1. Audio
2. Haptic vibrations
3. Screens

Notes:

Audio seems like the most versatile choice. Haptic vibrations have lower bandwidth but might still be a viable option in certain situations. Primarily when a small set of already known objects are wished to be recognized. E.g when shopping and 2 different vibrations patterns might be used for either confirming or denying that an item is on a predefined shopping list.

C.2.2 Input from user:

Can be done with:

1. Physical buttons
2. Voice commands
3. Touch Screen
4. Moving device
5. Geastures in front of camera

Notes: Physical buttons would be preferred in a scenario where user input does not need to be complicated. E.g activating detect object or color in front. OCR might benefit from geastures in front of camera since in this scenario, the user can point to a certain area of text.

C.2.3 Technical input to device

Can be done with:

1. Video Camera

C.3 Navigation:

C.3.1 Output to user:

Can be done with:

1. Audio
2. Screens
3. Haptic vibrations

Notes: Different types of navigation might afford different output options and is closely related to the user's capabilities. Warning the user of possible objects in their way can both be done with screens, haptic vibrations and audio. The audio has the disadvantage of drowning out audio feedback from the environment which a user with visual impairments relies on to a larger extent than a person without visual impairments. The haptic vibrations have the benefit that it servers users which grave impairments which the screens might not do.

C.3.2 Input from user:

Notes:

In navigation, input e.g. which objects to warn for and which destination to be put in, can largely be done prior to a journey. Input could hence be divided into "pre-journey" e.g. formerly mentioned examples and "during journey" e.g. deactivating warning system in crowded areas. The type of input to be used should reflect this division. "During journey", the user will not have the same interaction capabilities as during "pre-journey" and therefore the input for "during journey" has larger demand om usability.

C.3.3 Technical input to device

Can be done with:

C.4 Information Gathering:

C.4.1 Output to user:

Can be done with:

4. Audio
5. Haptic vibrations
6. Screens

Notes:

Audio seems like the most versatile choice. Haptic vibrations has lower bandwidth but might still be a viable option in certain situations. Primarily when a small set of already known objects are wished to be recognized. E.g when shopping and 2 different vibrations patterns might be used for either confirming or denying that an item is on a predefined shopping list.

C.4.2 Input from user:

Can be done with:

6. Physical buttons
7. Voice commands
8. Touch Screen
9. Moving device
10. Gestures in front of camera

Notes: Physical buttons would be preferred in a scenario where user input does not need to be complicated. E.g activating detect object or color in front. OCR might benefit from gestures in front of camera since in this scenario, the user can point to a certain area of text.

C.4.3 Technical input to device

Can be done with:

2. Video Camera

Appendix D Answers to questionnaire

Användarundersökning för hjälpmedel till synskadade

Detta är en användarundersökning för ett examensarbete av mig Björn Gummesson. Svaren kan komma att publiceras offentligt men deltagare är anonymiserade till den grad att endast ålder och styrka på synnedsättning kommer kunna publiceras.

Vilken typ av synnedsättning har du?

Glaucome

Om applicerbart, hur grov är din synnedsättning i %?

0% på vänster öga,20% på höger öga små synfält.

Hur gammal är du ?

- ○ 0-18 år
- ○ 18-25 år
- ○ 25-30 år
- ○ 30-40 år
- ○ 40-50 år
- ◉ 50-65 år
- ○ 65+ år

På grund av din synnedsättning och UTAN hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är svårt eller omöjligt att ta till dig?

Nyheter allmän info

På grund av din synnedsättning men MED hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är ändå svårt eller omöjligt att ta till dig?

Skyltinformation i utemiljö och allmänna lokaler.

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du i din vardag?

Iphone,förstoringsglas,optiskt förstoringsglas,solglasögon,talande våg,talande utomhustermometer,käpp.

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du mycket i din vardag?

Käpp,iphone,optiskt förstoringsglas.

Vilka typer av hjälpmedel relaterade till din synnedsättning har du men använder endast sällan eller aldrig?

Daisyspelare.

Om du använder dina hjälpmedel olika mycket, vad är det som gör att du använder vissa mer än andra?

Det beror på vilken aktivitet jag gör.

Om några, vilka hjälpmedel tar du alltid eller nästan alltid med dig när du ska utanför hemmet och varför just dessa?

Käpp för min rörlighet,solglasögon vid stark sol,iphone,förstoringsglas om jag måste läsa något.

Användarundersökning för hjälpmedel till synskadade

Detta är en användarundersökning för ett examensarbete av mig Björn Gummesson. Svaren kan komma att publiceras offentligt men deltagare är anonymiserade till den grad att endast ålder och styrka på synnedsättning kommer kunna publiceras.

Vilken typ av synnedsättning har du?

grav synnedsättning

Om applicerbart, hur grov är din synnedsättning i %?

vänster 0,07 (7%) höger 0,05 (5%)

Hur gammal är du ?

- ( ) 0-18 år
- (•) 18-25 år
- ( ) 25-30 år
- ( ) 30-40 år
- ( ) 40-50 år
- ( ) 50-65 år
- ( ) 65+ år

På grund av din synnedsättning och UTAN hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är svårt eller omöjligt att ta till dig?

om man inte har hfälpmedel så skulle det bli väldigt svårt för mig att ta mig runt utomhus, man skulle gå in i det mesta,och snava på kanter och upphöjningar, det skulle vara svårt att veta om man tar rätt produkter i affärer, färger.

På grund av din synnedsättning men MED hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är ändå svårt eller omöjligt att ta till dig?

stolpar, bilar, folk, spis, ugn, klockan, hålor, gropar, färger.

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du i din vardag?

käpp, fickminne, elektroniskt förstorningsglas, färgindikator, en plupp, kikare, glasögon med en slags plupp på, dc spelare, mobil, ipad, cc tv.

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du mycket i din vardag?

käpp, fickminne, elektroniskt förstorningsglas, färgindikator, en plupp, kikare, glasögon med en slags plupp på, dc spelare, mobil, ipad, cc tv.

Vilka typer av hjälpmedel relaterade till din synnedsättning har du men använder endast sällan eller aldrig?

kikaren

Om du använder dina hjälpmedel olika mycket, vad är det som gör att du använder vissa mer än andra?

vissa e mer användbara än andra, och andra är för klumpiga dör att bära med sig överallt.

Om några, vilka hjälpmedel tar du alltid eller nästan alltid med dig när du ska utanför hemmet och varför just dessa?

kepp, elektroniskt förstoringsglas, färgindikator, mobil, de hjälper mig mest för att kunna läsa om det behövs, känna efter faror, och veta om färger

Det här innehållet har varken skapats eller godkänts av Google.

Google Formulär

Användarundersökning för hjälpmedel till synskadade

Detta är en användarundersökning för ett examensarbete av mig Björn Gummesson. Svaren kan komma att publiceras offentligt men deltagare är anonymiserade till den grad att endast ålder och styrka på synnedsättning kommer kunna publiceras.

Vilken typ av synnedsättning har du?

Utsagen näthinna till 100 % på höger och till 55 % på vänster

Om applicerbart, hur grov är din synnedsättning i %?

Jag har 10 % synskärpa på den del av vänster som jag ser med. Skärpan ligger centralt och jag har tunnelseende.

Hur gammal är du ?

- ( ) 0-18 år
- ( ) 18-25 år
- ( ) 25-30 år
- ( ) 30-40 år
- ( ) 40-50 år
- (x) 50-65 år
- ( ) 65+ år

På grund av din synnedsättning och UTAN hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är svårt eller omöjligt att ta till dig?

All typ av skriftlig information, allt från samhällsinfo,tidtabeller prislappar mm mm

På grund av din synnedsättning men MED hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är ändå svårt eller omöjligt att ta till dig?

Jag kan ta till mig allt men det går långsamt, ibland extremt långsamt

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du i din vardag?

Förstoringsglas (trditionellt och elektroniskt), förstoringsprogram, en liten kikare för att kunna läsa skyltar och informationstavlor, läskamera. ljudböcker vi LegimusiPhone, specialtangentbord och naturligtvis vit käpp.

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du mycket i din vardag?

Samma som ovan. De hjälpmedel jag har använder jag flitigt.

Vilka typer av hjälpmedel relaterade till din synnedsättning har du men använder endast sällan eller aldrig?

Inga

Om du använder dina hjälpmedel olika mycket, vad är det som gör att du använder vissa mer än andra?

Det äår stundens behov som avgör

Om några, vilka hjälpmedel tar du alltid eller nästan alltid med dig när du ska utanför hemmet och varför just dessa?

Vit käpp, förstoringsglas (båda), kikare3n och iPhone.

Användarundersökning för hjälpmedel till synskadade

Detta är en användarundersökning för ett examensarbete av mig Björn Gummesson. Svaren kan komma att publiceras offentligt men deltagare är anonymiserade till den grad att endast ålder och styrka på synnedsättning kommer kunna publiceras.

Vilken typ av synnedsättning har du?

Gula fläcken

Om applicerbart, hur grov är din synnedsättning i %?

7%

Hur gammal är du ?

- ○ 0-18 år
- ○ 18-25 år
- ○ 25-30 år
- ○ 30-40 år
- ◉ 40-50 år
- ○ 50-65 år
- ○ 65+ år

På grund av din synnedsättning och UTAN hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är svårt eller omöjligt att ta till dig?

Att gå utan käpp

På grund av din synnedsättning men MED hjälpmedel, vilken eller vilka typer av viktig information i din vardagliga miljö är ändå svårt eller omöjligt att ta till dig?

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du i din vardag?

Vilka typer av hjälpmedel relaterade till din synnedsättning använder du mycket i din vardag?

Käpp och solglasögon

Vilka typer av hjälpmedel relaterade till din synnedsättning har du men använder endast sällan eller aldrig?

Elektroniskt förstoringsglas

Om du använder dina hjälpmedel olika mycket, vad är det som gör att du använder vissa mer än andra?

Käppen eftersom jag inte kan gå utan den

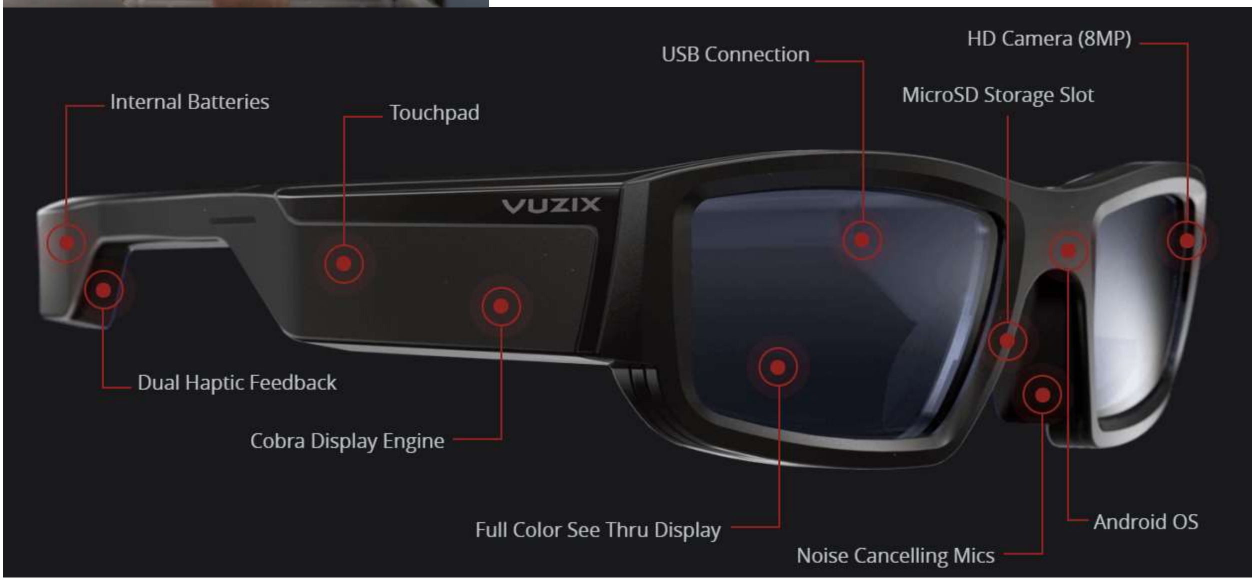
Om några, vilka hjälpmedel tar du alltid eller nästan alltid med dig när du ska utanför hemmet och varför just dessa?

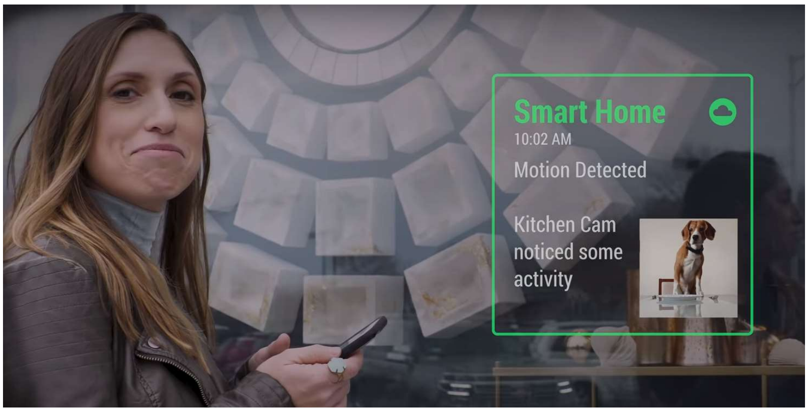
Både käppen och solglasögon

Det här innehållet har varken skapats eller godkänts av Google.

Appendix E Example of the benchmarking

E.1 Vuzix Blade





E.2 Price:

1000 $

8.1 Description:

A pair of glasses with built in see-through 854x480 pixel displays.

Not aimed for people with visual impairments.

“A handsfree extension of your phone that displays your digital life right in front of your eyes”

Proprietary OS based on android

Connects with your phone

Has HD camera

Touchpad on side of the glasses

E.3 What are its main features?

A variety of apps can be installed

USB and Bluetooth audio

Built in microphone

Open platform for 3rd party app integration compatible with IOS and android

Displays can be ordered with

- Prescription inserts
- Multiple lens color choice
- Photochromic
- UV protection

E.4 What information does it gather?

Capture images and video with a 8 mp camera with 780p video

E.5 How does the user produce input to the device?

Voice-contol

Touch pad on side of glasses

Head motion tracker

Remote control app for Android and IOS

E.6 How does the user receive output?

See-through full color 854x480 displays

Haptic vibration alert

E.7 How does it communicate with other tools?

Bluetooth

Wi-Fi

Micro USB

E.8 Possible api´s/ways to communicate with the product

The Android OS running on the Blade is a modified version of Android 5.1.1(Lollipop), tailored to the components and capabilities of the device.

The glasses uses existing android API's like :

- Camera: android.hardware.Camera or android.hardware.Camera2
- Sensors: SensorManager
- Bluetooth: BluetoothManager and BluetoothAdapter
- Database – Standard Android SQLite is supported

Some components of the blade require device-specific API´s. Vuzix has made SDK documentation available.

E.9 What additional user burden comes with the products?

The device has its own app store.