humans of AI

AI-powered smart glasses that empower visually impaired people to achieve greater independence and social integration.

master thesis by ferkan metin

march 2020

colophon

This is the final documentation of my graduation project titled *AI-powered smart glasses that empower visually impaired people to achieve greater independence and social integration.*

Executed in cooperation with the company *Envision.* This project completes my Master's program Design for Interaction at the faculty of Industrial Design Engineering in Delft.

contact

Ferkan Metin

Delft University of Technology

Faculty of Industrial Design

me@pherkan.com

+31 6 81 277 393

preface

I feel very privileged being in this field where human-centered design truly shows the positive impact it can have on people's lives.

The cooperation with Envision, Visio and the many visually impaired people opened a new world to me.

My aim within this project is to show how artificial intelligence is changing lives of visually impaired people and how the human aspect is necessary to improve future AI.

The future is bright and it is now.

envision

The company involved in this project is Envision AI. It is a company with six employees with their office in The Hague. Envision is a software platform that enables people with visual impairment to live more independently. Their app processes images taken by a user to intelligently extract useful information by using the smart phone camera.

Company Mentor

Karthik Mahadevan
Co-founder of Envision

www.letsenvision.com

tu delft

The master direction Design for Interaction, offered by the Industrial Design Engineering faculty at the TU Delft, specializes in the analysis of- and conceptualization for human-/product interactions. During the design process a clear understanding of the relation to the physical, cultural and societal user context is essential.

University supervisory Team

Chair: Dr. Dipl. - Des. Boess, S. U.
Applied ergonomics and design

Mentor: Dr. Lomas, J. D.
Cognitive Science and HCI

visio

Visio played a role as a stakeholder where multiple people of the organisation provided knowledge regarding visual impairment. Multiple meetings were held with Visio and its volunteers for conducting research with visually impaired people.

Stakeholders and volunteers

executive summary

This graduation thesis came into existence as a pilot of *Delft Design Labs* in collaboration with Envision AI. The bolded text in this executive summary represent the chapters in this report.

My graduation thesis started with the goal of looking at possibilities of artificial intelligence and its current use cases for visually impaired people, abbreviated throughout the report as VIP.

The report starts with an **introduction**, followed up by a **background analysis** on topics that are important for this subject such as visual impairment, artificial intelligence, inclusivity, diversity and accessibility.

Artificial intelligence is a very broad definition and we are still in the period of Weak AI. Most visually impaired people do or can not rely on sight. Diversity and inclusivity is about creating equal opportunities for everyone and is gaining more traction due to the population change, global talent and the shortage of it and non-traditional forms of work.

Subsequently, **contextual analysis** of VIPs was conducted to gain insights about their needs, wants and obstacles in daily life. This analysis consisted mainly of in-depth interviews and immersive research.

It showed that VIPs rely on technology as much as, or even more than, sighted people and do not want any special treatment. There is a significant difference between being born blind or becoming blind and whether the VIP is completely blind or still has some rest vision. Guidance from a sighted person, even if it is just a little, leads to a large gain of independence of the VIP.

Many daily tasks cost VIPs more energy, compared to sighted people, because the impairment is not supported or compensated.

The findings and insights from the contextual analysis are **synthesized** to personas: understanding who to design for and what their needs are.

In combination with artificial intelligence and visual impairment, focus areas were found in the field of computer vision, human assistance, accessibility outdoors and online, improvement of mobility and information gathering.

To further specify possible **design directions**, I conducted idea speed dating sessions and further research with Envision.

This showed that AI still lacks a lot of reliability and that the current solutions are cumbersome to use. It also showed that help of a person (human assistance) is required at times.

I specified my **design goal** as: designing a system that uses human assistance when AI is not reliable. An interaction vision was chosen for the final concept to feel like a kid that is being held by their parent: the concept should feel reliable, convenient and empowering.

In a pre-conceptualization phase, I conducted further **focused research** into the possible technological means to accomplish the design goal, while working closely with Envision. This replaced the wider ideation that normally follows goal forming. It resulted in a focus on supporting the AI with human help provided with smart glasses.

BeMyEyes and Aira are two existing tools that offer human assistance. Both are being used by a smartphone app and help visually impaired people in their daily tasks. Trying out these smartphone apps also showed its shortcomings in these areas: pointing the camera at the right angle, having another device in your hand as a VIP while having a white cane and/or a guide dog, the camera that might not show a lot of content.

Aira also has a service where it provides smart glasses for VIPs. This is preferred as it offers a handsfree experience, orientation is easier and the helper has an improved experience of seeing a point-of-view shot of what the VIP sees. I tested various smart glasses for their potential in supporting AI capabilities and human assistance.

Eventually the Google Glass is chosen for conceptualization: it is lightweight, comfortable, support from Google and a better CPU and camera compared to other smart glasses.

In the next chapter, **conceptualization**, I focused on creating the concept which consists of three elements: a smart glasses app for the VIP, an onboarding section in the current Envision app for the VIP and a *companion app* for sighted helpers.

The smart glasses are worn by a VIP and contains AI features such as reading text, identifying objects and finding people. Additionally, it contains possibilities of making a video call to a *companion* (human assistance). The VIP navigates through the menu with swipe movements, accesses features by double-tapping and gets feedback in audio (screenreader). By double-tapping on the side, a photo is taken with the smart glasses camera and the content is spoken out (*"Looks like a man sitting behind a laptop"*).

Companions are friends, family, unpaid volunteers or paid helpers. A companion signs up for the companion app, receives a video call from the VIP and gets visual input from the smart glasses camera.

The advantage of the companion app is the improved visual input (point-of-view and wide angle) and additional controls such as seeing the VIP's location, turning the flashlight on or off, taking a screenshot or messaging while video calling.

In my concept I also created a reciprocal learning model that focuses on improvement of the AI by VIPs and sighted people. The AI *improves itself* by guidance from the VIP and with the help of a sighted helper. The latter is done by adding data to the AI system within the Envision companion app.

I end this report with conclusions and recommendations. I conclude that AI-powered smart glasses with human assistance is an interim solution until AI becomes more reliable.

For now, it is the perfect solution for VIPs to achieve greater independence and social integration

More research is needed for the use of smart glasses by VIPs. Additionally, improvement of AI and smart glasses will lead to a future where this technology will create even more independence for VIPs.

To be a better version of themselves.

definitions and abbreviations

AI	-	Artificial Intelligence
CV	-	Computer Vision
DL	-	Deep Learning
D&I	-	Diversity and inclusivity
HI	-	Human Intelligence
ML	-	Machine Learning
MVP	-	Minimum Viable Product
TTS	-	Text to speech
UX	-	User Experience
VIP	-	Visually impaired person

Algorithm: process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

Artificial Intelligence: The theory and development of computer systems able to perform tasks normally requiring human intelligence.

Deep learning: a subset of artificial intelligence and a combination of ML-models. It uses statistical methods to enable machines to improve with experience.

Human Intelligence: Quality of the mind that is made up of capabilities to learn from past experience, adaptation to new situations, handling of abstract ideas and the ability to change his/her own environment using the gained knowledge.

Interaction Vision: A creative design technique to address already in an early phase of a design process the intended character of the interactions with a future design concept.

Machine learning: a subset of artificial intelligence that uses algorithms; a set of rules and statistical techniques used to learn patterns from data.

Minimum Viable Product: a version of a product with just enough features to satisfy early customers and provide feedback for future product development

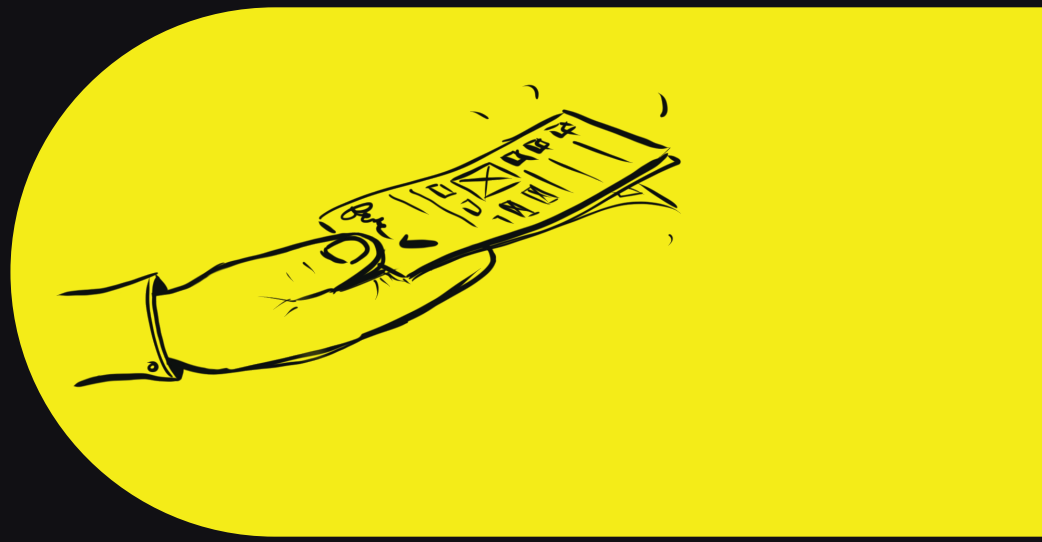
Speed dating: a design method for rapidly exploring application concepts and their interactions and contextual dimensions without requiring any technology implementation.

Text to speech: The technology of turning text into spoken words. TTS understands text and natural language to generate synthesized audio output.

table of content

a) introduction

This chapter introduces the project with the research topic and design goal. It introduces the involved stakeholders and provides an overview of the report.



in this chapter:

1. project

The MSc graduation project is the final project that completes the Master's degree *Design for Interaction* offered by the Faculty of Industrial Design Engineering (IDE) of Delft University of Technology (TU Delft). The project started in September 2019 and came to an end in March 2020 with this report and a public presentation given at the faculty.

This project is conducted in collaboration with Envision, Delft Design Labs Inclusive Design and Visio.

The aim of this project is to develop a design concept.

This development is realized by doing literature research, field research and working closely with Envision. Pages 5 and 6 show a detailed visual overview of the report.

This project started with the goal of **researching new use cases of artificial intelligence and developing a concept to aid visually impaired people**. This lead to conducting background analysis about the subjects involved. This chapter answered questions such as:

- What is visual impairment?
- What is AI and its differences?
- What role does inclusivity play?

After the background analysis, contextual analysis was conducted followed up by synthesis and exploring design directions. This formed towards a design goal to design a concept for visually impaired people that uses human assistance substitution when artificial intelligence is not reliable.

Design of a concept for visually impaired people that uses human assistance substitution when artificial intelligence is not reliable.

This project focuses on blind people; which means people with less than 5% vision.

Names of participants are changed to guarantee their privacy as well as their faces being blurred out.

The next chapter introduces the stakeholders and their roles, followed up by the visual overview of the report that represent all chapters included in this report.

2. stakeholders

The stakeholders for this research project are Visio and Envision.

3.1 Envision

Envision is a software platform that enables people with visual impairment to live more independently. With Envision, visually impaired users can shop in supermarkets, use public transport, read menu cards in restaurants, recognise their friends, find their belongings and much more. The app processes images taken by a user to intelligently extract useful information by using the camera of your smart phone.

Envision is founded in October 2017, which started as part of the incubation program of YES!Delft: a tech incubator that helps entrepreneurs to build and grow leading technology companies. Currently they are funded by 4impact: a Venture Capital fund investing in digital tech start-ups focusing on societal and sustainable impact and financial return.

The envision platform is as of today (March 2020) available as an app on the App Store (iOS) and Google Play Store (Android).

The two founders are Karthik Mahadevan, who finished his Master's degree in Integrated Product Design at Delft University of Technology and Karthik Kannan. Kannan finished his undergrad in India together with Mahadevan and currently works on the technical part of the app.

The company has 6 employees.

Envision as a company is interested in new use cases of their app for visually impaired people and bringing this same experience to more accessible hardware.

3.2 Visio

Visio, also known as Koninklijke Visio (literally translated as Royal Visio), is a Dutch expertise center for education, care and service for people with a visual limitation.

Visio offers information, advice but also different forms of research, counseling, rehabilitation, education and living. These services are for partially sighted or blind people, as well as those who suffer from intellectual, physical or sensory disabilities. People who are personally or professionally involved with partially sighted or blind people can also turn to Visio for information and professional development.

Visio is the current legal successor of the Royal Institute for Education of the Blind in Amsterdam, founded in 1808. They currently have around 3000 employees.

During this project Visio provided many of the knowledge regarding visually impaired people and the rehabilitation process. Additionally, design and testing sessions are conducted as a collaboration between Envision and Visio. Mainly the volunteers of Visio played a big role in this project, as without them a lot of research would have been not been possible.

Visio as an organisation is interested in

any research about technology that can help VIPs to live a more independent life.

Figure 1: Tools for visually impaired people at the Visio office in Rotterdam

Figure 2: Visio Headquarters in Huizen

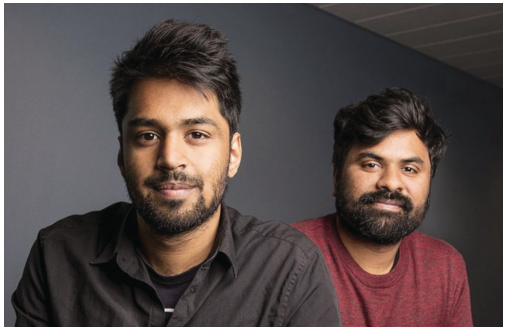


Figure 3: A picture of Karthik Mahadevan (left) and Karthik Kannan (right). (Source: Maurits Giessen for quotenet.nl)

Figure 4: Celebration with the investors at the YES!Delft office.

3. visual overview

Pages 5 and 6 show a visual overview how the report is set.

It starts with an introduction on the report, followed by background analysis on topics like artificial intelligence, visual impairment, inclusivity, diversity and accessibility. Consequently, contextual analysis is performed by doing immersion and in-depth interviews.

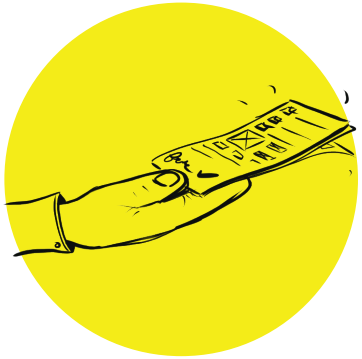
The data from this research was synthesized into personas and a design direction was explored. This led to the design goal, followed up by focused research and lead into the concept. Finally, recommendations are given what to focus on for future research.

The report is divided into 9 chapters, each chapter starting with a dark page and an introduction to the chapter. Every chapter, except the introductions ends with key take-aways of that chapter and what it contributes towards the design goal.

Insights are shown in blue blocks and quotes are shown in orange blocks.

A

introduction



Introducing the project with the research topic and design goal. It introduces the involved stakeholders and provides an overview of the report.

B

background analysis

Focuses on the different topics and provides background information regarding visual impairment, AI, inclusivity, diversity and accessibility.

C

contextual analysis

Insights by conducting interviews and immersion studies with the target group. These insights are of importance for a user-centered approach towards a final concept.

D

synthesis

Synthesizes the data from the background and field research. Additional research is conducted after the field research to find out about current technology (apps and tools) and needs of VIPs. Besides creating personas of the target group, this chapter also shows the first interesting

E

design direction

Focused research on Envision and speed dating sessions which provides the final direction and the revised design goal in the next chapter.

F

design goal

After conducting research in different fields, it shows that AI is not reliable and people are still needed. This chapter shows the revised design goal and the interaction vision for the final concept.

G

focused research

Focused research is conducted on apps that use computer vision for VIPs. Additionally, research on current similar tools (software and hardware) is conducted to gather insights and lead to a concept to test with VIPs.

H

concept

The concept focuses on AI-powered smart glasses, a companion app and the onboarding. A reciprocal learning model and human-ai interaction is explained. Different scenarios show how the concept is used with the use of a helper.

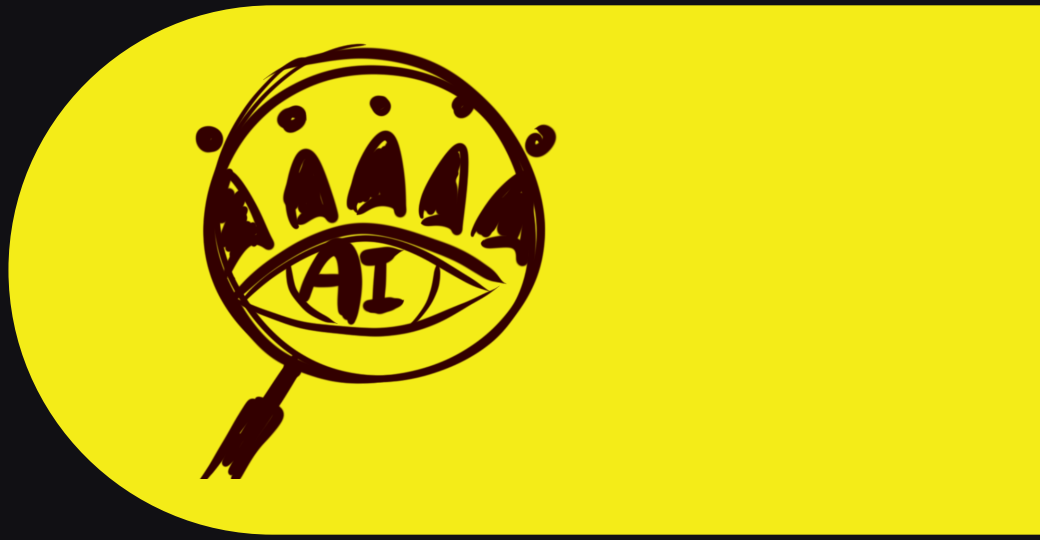
I

recommendations

Due to time and focus of the project, this chapter recommends different areas of conducting research for future AI for visually impaired people. Additionally, recommendations are given for a future of smart glasses.

b) background analysis

The second chapter of the report focuses on the different topics and provides background information regarding visual impairment, artificial intelligence, inclusivity, diversity and accessibility.



in this chapter:

1. artificial intelligence
2. visual impairment and blindness
3. inclusivity, diversity, accessibility
4. key takeaways and contribution

1. artificial intelligence

1.1 definition

Artificial Intelligence (AI) is the intelligence demonstrated by machines. It is an umbrella descriptor which refers to machines capable of performing things that humans typically would. It solves real-life tasks by emulating humans' perception, reasoning, learning, communicating, and acting against complex environments.

Steven Moore, who was dean at Carnegie Mellon School of Computer science, explained Artificial Intelligence as the science and engineering of making computers behave in ways that, until recently, we thought required human intelligence (Nilsson, 2010).

According to John McCarthy (Turing, 1950), Artificial Intelligence is the science and engineering that tries to make machines intelligent, trying to get them to understand human language, problems, goals and human beings.

To understand artificial intelligence, it is important to look at human intelligence, as these concepts should not be confused.

Human Intelligence is defined as the quality of the mind that is made up of capabilities to learn from past experience, adaptation to new situations, handling of abstract ideas and the ability to change his/her own environment using the gained knowledge.

1.2 difference with human intelligence

Below are a few examples why artificial intelligence is interesting in comparison to so-called human intelligence which is used for the final concept.

- **Speed of execution:** while one doctor can make a diagnosis in about 10 minutes, AI systems can make a million for the same time.
- **Less biased:** they do not involve biased opinions in decision making processes
- **Operational ability:** They do not expect a halt in their work due to saturation. Saturation would not be met by an operational malfunction.
- **Accuracy:** Preciseness of the output increases.

Artificial Intelligence has significant dominance in many tasks, especially when it comes to monotonous judgments.

Further on in this report, human intelligence will be discussed and clear distinctions will be given whenever a machine is preferable and whenever it is better to have actual humans conduct a task.

1.3 ai capabilities

After the introduction of Siri, many examples of AI appeared on the market. AI is bigger than ever, but why is it that AI is now finally getting the acknowledgement it deserves?

AI capabilities have been transformed due to:

- **Hardware:** Growth of AI-optimised hardware and thus having more computation power, especially making a difference for deep learning applications
- **Data:** Increased availability of data due to social media and the internet of

things

- **Algorithms:** Development of improved algorithms for AI
- **Open source:** Open source AI software frameworks that enable experimentation and to emphasise the business potential of sharing software source code which was introduced by Netscape for the first time in 1997.
- **Investment:** Increased investment in AI by venture capitalists and companies.
- **Awareness:** Greater awareness of AI among investors, executives, entrepreneurs and the public
- **Cloud-based AI services:** Catalyzing of developer adoption with Cloud-based AI services

This shows that AI will keep improving over time, meaning that it will also be more reliable.

1.4 computer vision

The design concept of this report focuses on computer vision. This is an interdisciplinary scientific field dealing with computers gaining high-level understanding from images. Computer vision is about automating tasks that the human visual system can do (Ballard, 1982).

Computer vision can be categorized under *Weak AI.* Figure 5 shows the differences of Weak AI and Strong AI. Figure 6 shows a timeline of when the different methods of AI were introduced.

For this project it is important to get a general grasp of AI and its possibilities: knowing what is possible now, what might be possible in a later stage or what will most likely never become reality.

Below an overview is given what AI is used for:

- Evolutionary Computation – e.g. genetic algorithms, genetic programming
- Vision – e.g. object recognition, image understanding
- Robotics – e.g. intelligent control, autonomous exploration
- Expert Systems – e.g. decision support systems, teaching systems
- Speech Processing– e.g. speech recognition and production
- Natural Language Processing – e.g. machine translation
- Planning – e.g. scheduling, game playing
- Machine Learning – e.g. decision tree learning, version space learning.

In the context of this project, this overview is used in a way to see what AI can be implemented in the design concept.

Besides the focus on computer vision to recognize objects or understand image, speech processing and natural language processing play a role in the final design concept.

Speech processing for a potential sound assistant for the concept.

Natural language processing for VIPs to understand visuals. This is shown in the concept whenever the user is scanning a document and is being informed whether it is a header or title that is being spoken out.

	Strong AI	Weak AI
Definition:	The form of Artificial Intelligence, which has the same intellectual abilities as human, or even surpasses him in it.	Weak AI is a machine intelligence that is limited to a specific or narrow area. It is not intended to match or exceed the capabilities of human beings.
Capabilities and domains:	Logical thinking, making decisions in case of uncertainty, plan and learn, communicate in natural language, use all these abilities to achieve a common goal.	Expert systems, navigation systems, voice recognition, character recognition, suggestions for corrections in searches
Found in:	Science Fiction	State of the art technology
Examples:	HAL 9000, The Terminator, Skynet, Marvin, Neuromancer, Glados, Commander Data	Alexa, Siri, IBM Watson, Google spam detection, Tesla autonomous car

Figure 5: Differences between strong AI and weak AI

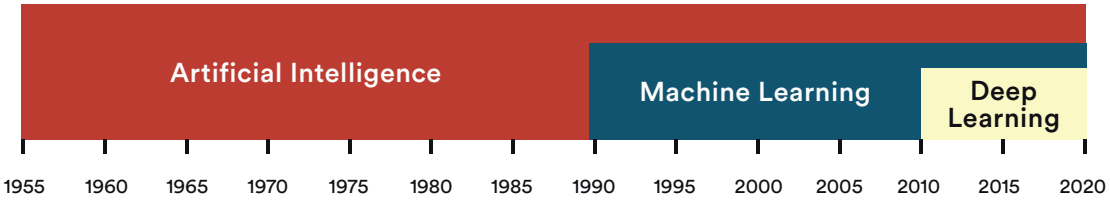


Figure 6: AI, ML and DL on a timeline.

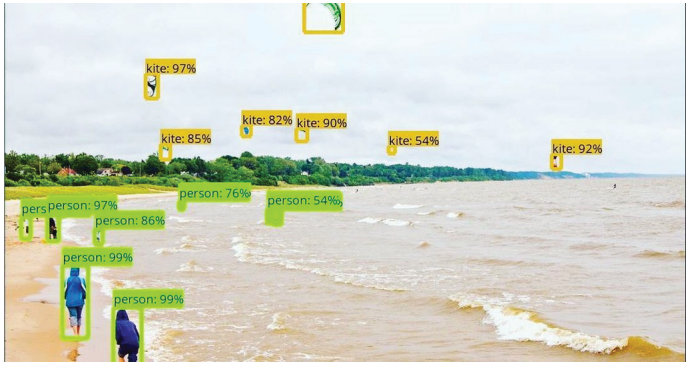


Figure 7: Two examples of computer vision that turns images into data with object detection

2. visual impairment and blindness

This graduation thesis report is focused on visually impaired and blind people. Whereas everyone knows what blindness is, it is often forgotten that there are many differences in visual impairment and how many people actually deal with a visual impairment.

2.1 Definition

Vision impairment, or low vision, means that even with visual correction, e.g. eyeglasses, contact lenses, medicine, or surgery, someone does not see well. Vision impairment can range from mild to severe.

Unlike low vision, blindness is strictly defined as the state of being totally sightless in both eyes. A completely blind individual is unable to see at all. However, the word blindness, is commonly used as a relative term to signify visual impairment (Dahl, n.d.).

This implies that the design concept will not be perceived visually. In addition, this also means that many of the design methods are not applicable when doing user research with VIPs.

Many of the design methods taught at the Design for Interaction faculty assume that the participants have sight: brainstorming, co-creation, affinity diagrams etc.

The focus of this project is on people that can not rely on their sight. Appendix 2 shows what causes visual impairment and different types of visual impairment.

2.2 The numbers

The "Nationale Rapportage Oogzorg" (2019) shows that in the Netherlands, about 250.000 people have a visual impairment. 150.000 of them are blind or partially sighted. In 2014, about 1.2 million people contacted the health insurance system for an eye condition and the waiting time for the eye doctor has gone up from 4-5 weeks (2010-2014) to 6.5 weeks (2016).

Looking at the numbers on an international level, between 300-400 million people are visually impaired due to various causes. Of this group, approximately 50 million people are totally blind and 80% of blindness occurs in people over 50 years of age.

Additionally, the ratio between male-female is 1:2 of being visually impaired and between 2010 and 2020 visual impairment increased by 20% (Oogfonds, 2019).

A report by "Vision 2020" (2019) shows that percentually the largest increase of visual impairment and blindess is in the age group between 55 and 74 years.

These numbers show what possible target groups to conduct research with. Additionally it also helps with making recommendations later in the report on the impact it can have worldwide.

2.3 Consequences

According to Runjić et al. (2015), visually impaired and blind young people have more problems with social skills, score lower on social competences than young people without a visual impairment. This leads to less dating, less friendship activities, less finished education and a job.

Furthermore, the lack of support services such as building access, transportation, information and communication can

make handicapped people overly dependent on their families, which prevents them from being economically active and socially included.

One study also talked about fatigue as an often mentioned symptom with irreversible visual impairment (Schakel, et al., 2017). Results indicate that fatigue related to having low vision is different compared to other patients and plays a role in the way patients react, adapt and compensate to the consequences of loss.

Visually impaired people often described it as a mental, daily and physical experience. High cognitive load, the amount of activities and the intensity of it were mentioned as causes.

Besides that, it is of huge importance to note that mobility is a big part of our lives; moving from one place to another.

The aforementioned research the lack of support services, fatigue, the high cognitive load and mobility shows possible opportunities and design directions for this project.

In 2007, Apple released its iPhone. The iPhone changed how we look at smartphones nowadays. This is not any different for VIPs, especially when Apple introduced VoiceOver two years later: a screenreader that speaks out the text what is shown on the screen. Figure 8 shows a screen on the iPhone where VoiceOver is activated and speaks out the text that is shown inside the block.

This element of not *feeling different* was not only gathered from research, but also showed with the first interactions with VIPs. This also formed the design directions in a later stage where VIPs prefer to use the same tools as anyone else.

Lastly, visual impairment is connected with a white cane. This is a cane that allows the user to scan their surroundings for obstacles or orientation marks, but is also helpful for onlookers in identifying the user as blind or visually impaired and taking appropriate care (Nichols, 1995).

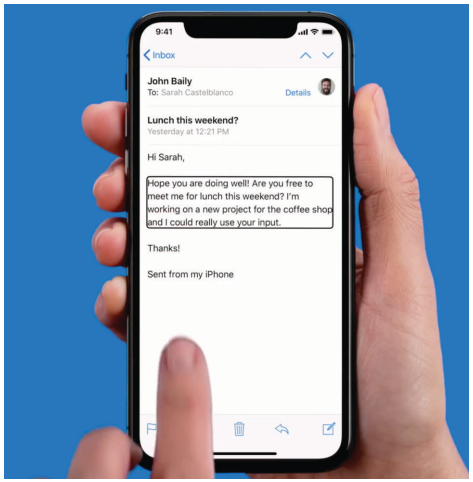


Figure 8: Person using iPhone with VoiceOver (source: Apple Support on Youtube)

Figure 9: Karthik and James navigating through the station while James is holding a white cane to scan his surroundings.

3. inclusivity, diversity, accessibility

This thesis focuses on aiding visually impaired people and therefore inclusivity, diversity and accessibility are topics that will play a major role.

3.1 definition

Inclusivity is the practice or policy of including people who might otherwise be excluded or marginalized. It is often used interchangeably with diversity. Although a diverse team is a team that has many different unique individuals. This does not only include the usual diverse selections such as race, age, sex and religion. Also additional unique personality characteristics such as introverts, extroverts, liberals and conservatives.

The difference between diversity and inclusivity could be explained by the fact that diversity is the *who or what* and inclusivity on the other hand is a *mindset*.

You can have a team of one, which is not diverse, but, can be inclusive. Someone who has an inclusive mindset behaves in a way that welcomes and embraces diversity.

This knowledge is crucial during the design process.

3.2 the importance

Diversity and inclusivity is about creating equal opportunities for everyone.

On a company level, more diverse workplaces are not just seen as a more desirable employer; they also outperform competitors and achieve greater profits.

Different reports state that companies with more diverse management teams have 19% higher revenue. (BCG) They also are more innovative, 1.7 times more likely to be innovation leaders in their market.

It leads to better decision making.

A blogpost by Glassdoor (2014) showed that 67% of job seekers said a diverse workforce is important when considering job offers.

Another study by McKinsey&Company (2014) show that companies with ethnically diverse companies outperform industry norms by 35%.

When looking at gender equality, .women account for 60% of college graduates but only 3% of leaders worldwide. Women and girls also represent two-thirds of the world's illiterate population.

Diversity can also be discussed on the workplace. It is a term which refers to a company which employs people of varying characteristics, such as gender, age, religion, race, ethnicity, cultural background, sexual orientation, religion, languages, education, abilities, etc. The necessity of diversity is demonstrated in direct and tangible benefits. It is not something that is just nice-to-have.

3.3 why now?

But why is diversity and inclusivity getting more attention?

Some reasons include:

- **Population profile changes:** the changing racial and ethnic mix of the U.S. population. The ethnic diversity among younger Americans through 2065. People between ages 15 and 24 make up close to 20% of the world's

population.

- **Millennials comprising three-quarters of global work force:** Research by EY (2015) shows that by 2025, millennials (those born between 1980 and 1996) are expected to comprise three-quarters of the global work force. Younger and increasingly diverse populations often bring with them evolving expectations and a willingness to bring diversity and inclusivity to the forefront of societal conversations.
- **Global talent:** workplaces are becoming more multicultural with global talent moving across countries and positions.
- **Non-traditional forms of work:** Non-traditional forms of work continue to gain popularity, such as freelancing, virtual work, and short-term project-based assignments.
- **Shortage of talent:** Especially acute in knowledge industries.

3.4 the role of inclusivity and accessibility

Inclusivity and accessibility plays a role in this project as it focuses on people with a visual impairment. Inclusivity eventually will lead to improved accessibility, whether that is for a deaf or blind person. Some examples can be seen in daily life, such as the Supreme Court handing victory to a blind man who sued Domino's website over site accessibility (Higgings, 2019).

In combination with AI, D&I technologies should be aware of AI amplifying stereotypes, adversely impacting underrepresented and marginalized populations.

AI algorithms have grown in complexity, opaqueness, ubiquity and exclusiveness.

The development of new technologies – specifically AI, machine learning, algorithms, text mining, sentiment analysis, and natural language processing – has provided novel capabilities which can be applied to diversity and inclusion challenges.

Additionally the role of inclusivity and diversity in this project is that the concept can be utilized by both blind and visually impaired people. Whereas blind people do not rely on sight, there are still a lot of VIPs that rely on some rest vision and should not be excluded.

Attention to accessibility is key in this project, because without an accessible interface the product could exclude a large part of the target group.

The events AI for Intelligent Vehicles (figure 10) and AiTech Symposium (figure 13) were visited to gain more information about AI: it helped with understanding in what stage AI is and partly about the human control of AI.

OpenVoice (figure 12) was visited with a VIP to conduct immersive research.

The Inclusive Design is Design event (figure 14) helped with seeing inclusivity in a different light (for deaf people). As a designer it made me contemplate about user experience of products and services: what might be a great UX for me, does not mean it might be great for others.

This is one of the key elements why human-centered design is so important.

September 19, 2019

AI for Intelligent Vehicles

Knowledge on artificial intelligence in vehicles and where AI stands currently.

Figure 10: Delft AI Meetup at Delft University of Techhnology

September 26, 2019

AiTech Symposium

Symposium about AI on how to keep intelligent autonomous systems under meaningful human control.

Figure 13: AiTech Symposium - all day event about AI.

September 20, 2019

Supertoegankelijk, The Checkout

Private event about accessibility in supermarkets and inclusive design.

Figure 11: Private event about accessibility in supermarkets organized by Visio.

October 31, 2019

Inclusive Design is Design

Event about inclusive design and how design affects the deaf.

Figure 14: One of the speakers talking about inclusive design and what it means for deaf people.

September 26, 2019

OpenVoice

Meetup about voice in the media. Visited this event with a VIP as part of immersive research.

Figure 12: OpenVoice Meetup in Amsterdam.

November 12, 2019

Visual revalidation on the move

Closed event about the revalidation of visually impaired people and how to get them moving.

Figure 15: Translation: *Today, Visio Leiden is on the move with and for professionals and experience experts in the region.*

4. key takeaways and contribution

- Artificial Intelligence as a definition is very broad.
- The focus of this report is on a subset of AI: computer vision
- Although the term was coined in 1956, AI is getting more popular nowadays due to hardware, data, awareness and more.
- There are all types of visual impairment, some more severe than others.
- About 250.000 people in the Netherlands have a visual impairment, worldwide about 50 million people are blind.
- Diversity and inclusivity is about creating equal opportunities for everyone.
- D&I are getting more traction due to the population change, global talent and the shortage of it and emerging conventionalization of non-traditional forms of work.

Chapter B contributes towards a final concept in a way that it is clear what the possibilities of AI are and in what stage we currently are.

Additionally, it shows that visually impaired people know how to use smartphones do not rely on visual input. This means that the AI-powered smart glasses should focus on other aspects such as form, haptics and sound and use current technology such as VoiceOver to make it accessible for VIPs.

The final concept should be accessible but also inclusive in a way that it can be used by everyone and no one is excluded. Not only age, gender and nationality play a role in this, but also the type of disability.

c) contextual analysis

This chapter shows interesting insights by conducting interviews and immersion studies with the target group. This was done to understand daily life from a VIP's perspective and identify challenges and opportunities for support through AI.

in this chapter:

1. interviews
2. immersion
3. key takeaways and contribution

1. interviews

In-depth interviews were conducted with ten visually impaired people. The participants were gathered through Envision, Visio, Oogvereniging and social media (figure 14).

1.1 the target group (figure 15)

- The age ranged from **16 to 58 years old** (average 34 yrs)
- 7/10 participants were **totally blind,** 3/10 participants still had **some rest vision**
- 6 male, 4 female
- 1 unemployed, 4 employed, 2 students and 3 volunteers
- 3/10 participants had a **guide dog**.

The goal of the interviews was to get a better understanding of VIPs and their daily obstacles with or without regards to technology. Likewise, it was to understand the needs and differences between visually impaired people, hence using a broad target group.

Figure 16: Tweet asking for visually impaired people to have a coffee with as part of the interviews.

The interviews were conducted in a trusted environment preferred by the VIP; in a café, at their work place or at their house.

Some of the insights also had overlap with the observation and immersion with the target group shown in the previous chapter.

The next pages show the participants and their quotes.

1.2 questions

Besides the general questions about age, occupation and what kind of visual impairment they had, below are some of the questions asked:

1. Do you use Envision or a similar app and how often?

2. What kind of technology changed your life?

3. What device are you the most work-productive on?

4. What kind of apps do you use?

5. Do you use special instances for traveling and are you part of a blind community?

1.3 findings and insights

The interviews showed the following interesting insights with corresponding quotes by the visually impaired people:

1. Lack of knowledge regarding technology

Whereas some of the users were very tech-savvy and up-to-date about the latest technology, some of the users were not aware of the possibilities. One example of a technological solution that is beneficial for visually impaired people is the bone-conduction headphones; a VIP is able to hear his surrounding while getting directions when commuting for instance.

"Is there a website with all the technological

solutions you have been talking about?" - Coen

2. Feeling *lucky* to be living right now

Four of the ten users also mentioned that they feel lucky because they are able to use a smartphone which makes them a lot more independent. A lot of times they mentioned that the *iPhone* changed their lives in combination with VoiceOver. Likewise, the addition of Siri and voice assistants in general.

"The iPhone is indispensable, it made me much less dependent of other people." - Christian

3. Living situation affects the chosen solution

Users living on their own or having less ability of relying on others were more dependent on technology than others. One user, which was also an expat, mentioned that he is forced to use apps like BeMyEyes: an app where volunteers can help you through your smartphone camera. This was due to the lack of family in the Netherlands that he could call if he would want to install something. During one session at Visio I met a woman who was revalidating, after offering different solutions to her problems she mentioned that she will not use any of them because she would totally rely on her husband for most things.

"I had to install a thermostat, but because I don't know anyone here in the Netherlands, I had to use BeMyEyes to have someone help me with it."

4. Different preferences of input

Another finding from the interviews was the difference in needs of possible solutions. Whereas some users thought haptics could play a bigger role in the future, some users mentioned that they have no feeling for haptics and much more prefer sound. Additionally, some of the participants would video call a friend or family member when in trouble, others never did.

"Using just sound to replace everything is tiring." - James

5. There is currently no all-in-one solution for accessible apps.

Findings showed that users do not rely on one specific app but use multiple instead. If Google Maps did not function the way it should, they would switch to Apple Maps. Besides that, some users relied on regular apps like Google Maps for navigation, where the younger VIPs used more advanced but primarily accessible options such as *BlindSquare*.

"I feel much more independent with apps like Envision, SeeingAI and Tap Tap See. If one of the three apps doesn't seem to work, I might call my mother." - Meryem

6. No special treatment

All VIPs made clear during the interviews that they do not require and want any special treatment. Often the VIPs mentioned that they do not want any special treatment. A lot of times they mentioned that they want to be normal

and look normal.

"I would not wear a helmet looking like a Terminator if it would help me to recognize text or anything."

7. Relying on technology

As mentioned before different solutions are preferred and used, which also shows the reliance on technology by visually impaired people. One of the participants, James, uses different apps like Google Maps, OV Maps and Envision on his phone to travel to work, school and back to home. All participants used an iPhone. When asked whether they rely more on technology or the government; most participants did not have any hope in the government improving their services but relied on technology much more.

8. Do as much as possible on their own

The majority of the users mentioned that they would want to do as much as possible on their own.

"It's tiring to ask everyone about everything." - James

9. Less energy than most people

Coen used to work 40 hours/week where I visited him for the interview. He works less hours now as it takes too much energy for conducting tasks while being blind. This is another insights gathered from most users and something that was not assumed beforehand. Being visually impaired means you have to think about every step you take, whereas people with sight can rely on their eyes if something is happening around them. This level of energy and focus of a visually impaired

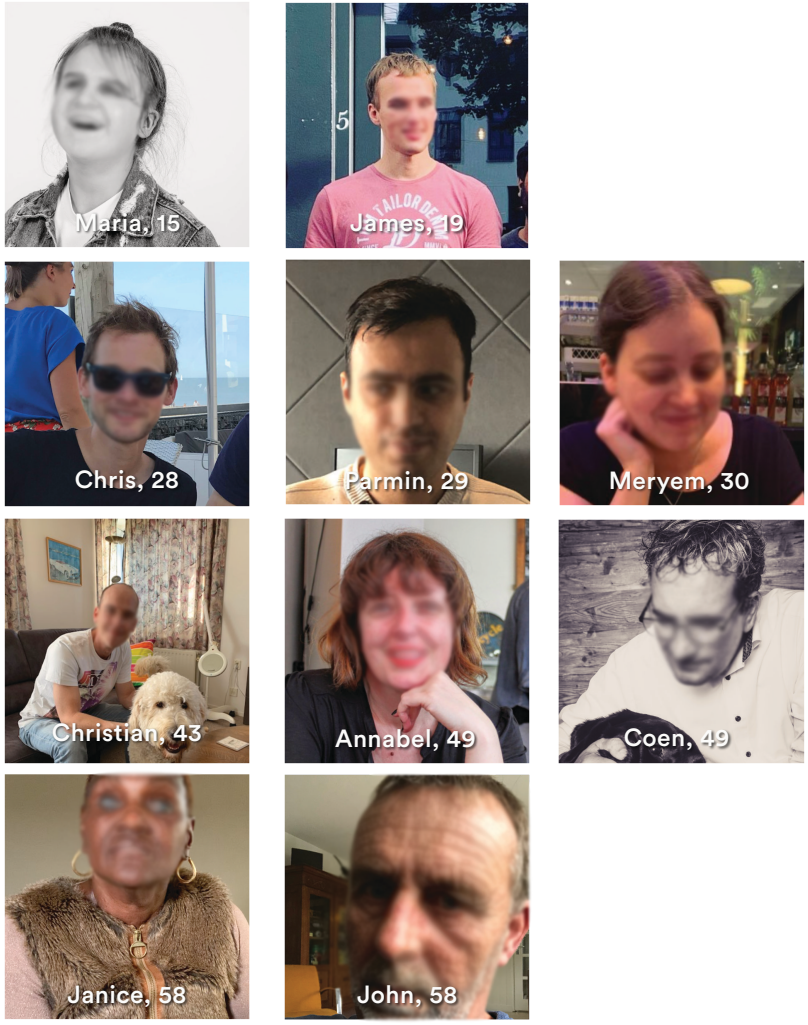


Figure 17: Interviews with visually impaired people.

Figure 18: Interview at Chris' work with his guide dog.

Figure 19: James showing how he uses the braille keyboard on his iPhone during an interview.

person is very important, especially when considering the design of products.

"I have to stay focused when I walk on the streets. Usually towards the evening I don't have a lot of energy left." - Janice

10. Change is not always good

For VIPs, change means that they are not able to complete a task in a way that is known to them. James told me that maintenance at a train station means that he has to video call one of his friends to get to his destination. Christian has a custom-made wardrobe so he knows what to wear. Janice keeps everything in separate boxes and puts everything back on the same spot. A software update for iOS on the iPhone made it impossible for Janice to send a voice message. It is therefore essential to look at what is known to the VIP and also understand that every change of a product/service can lead to uncertainties.

Besides the ten in-depth interviews with the aforementioned participants, smaller interviews were held which were mainly focused on using the Envision app. Likewise, at different accessibility events I spoke to visually impaired people and stayed in contact with them throughout the project. Many of those insights can be found in the next chapter that focuses more on observation.

2. immersion

The previous chapter focused on interviews with the visually impaired people. While interviewing is also very important, the information people provide during interviews is not always accurate or reliable. Observation did not only provide new insights but also verified some insights gathered from the interviews.

Living without a visual impairment means that it is difficult to imagine how the world is perceived by visually impaired people. Therefore, it is crucial to not make any assumptions and with each design step taking into consideration that the idea or concept is not going to be for people with sight. This mistake can be seen with many apps, of which one is the *OV-chipcard* app for public transport. The developer of the app did not add labels and thus it is inaccessible if you do not have sight.

Naturalistic observation and shadowing are used as part of the observation: studying the users in their natural environment (at their work, home or during an activity). This method is also known as an ethnographic interview. Depending on the situation, questions were asked during or after the observation.

The observation showed their pinpoints, product strengths and problems, struggles, physical and mental obstacles.

Staying in close contact with Visio, the eye institute in the Netherlands, and Envision allowed me to fully get immersed with the target group.

Gathering insights from observations was done by taking notes during the observation, but also analyzing the video recordings of the observation.

2.1 activities

The observations are carried out during the next activities:

1. One-on-one observation with the interviewees: Besides conducting one-on-one interviews with the users, observation was also a part of it.

2. From The Hague to Rotterdam with Janice: A 4-hour observation with Janice that started in the office of Visio in The Hague where I took a cab with her to Rotterdam. At her home she showed me the tools she would use, what kind of obstacles she runs into and what she does in her daily life. This is also called a *contextual inquiry*.

3. Visiting an event about accessibility with James: Another session where observation played a role was going with James to an event in Amsterdam. By using public transport to get to an unknown location, showed its own challenges and obstacles.

4. Salsa classes for blind people: To get a better grasp of what blind people are capable of, I joined salsa classes for visually impaired people. This was given on weekly basis in Rotterdam and was for both visually impaired people as sighted people.

5. Visiting a blind school: As part of the final observation research, I visited a 'Visio school': a school for visually impaired students, in the south of Rotterdam.

2.2 findings and insights

Immersion with the target group showed general insights about visually impair-

ment. Likewise, some insights were very straight-forward, where others were not.

The insights gathered from the observations:

1. Visually impaired people differ as much as everyone else from each other.

Although the findings are obvious, as *everyone* is different; it is still something to keep into consideration. This means that with the creation of a concept, there is not a solution that might fit all VIPs. Observation of the target group in particular showed the different levels of **tech-savviness**. But also **independence** played a huge role; whereas some of the VIPs barely needed help from others, some VIPs relied on daily basis on a friend or spouse. Additionally, some VIPs would not be able to live without a guide dog, whereas others did not see any added value.

2. The elderly visually impaired people rely more on different solutions of technology.

There was an obvious difference between elderly VIPs (40 years and older) and younger VIPs. Whereas the younger VIPs relied much more on just their smartphone, elderly relied on different technological solutions as can be seen in figure 18-4.

3. There is a difference between being born blind and becoming blind.

People that are born blind or have been blind from a very young age, usually are more at peace compared to people that used to have sight. If a VIP goes blind,

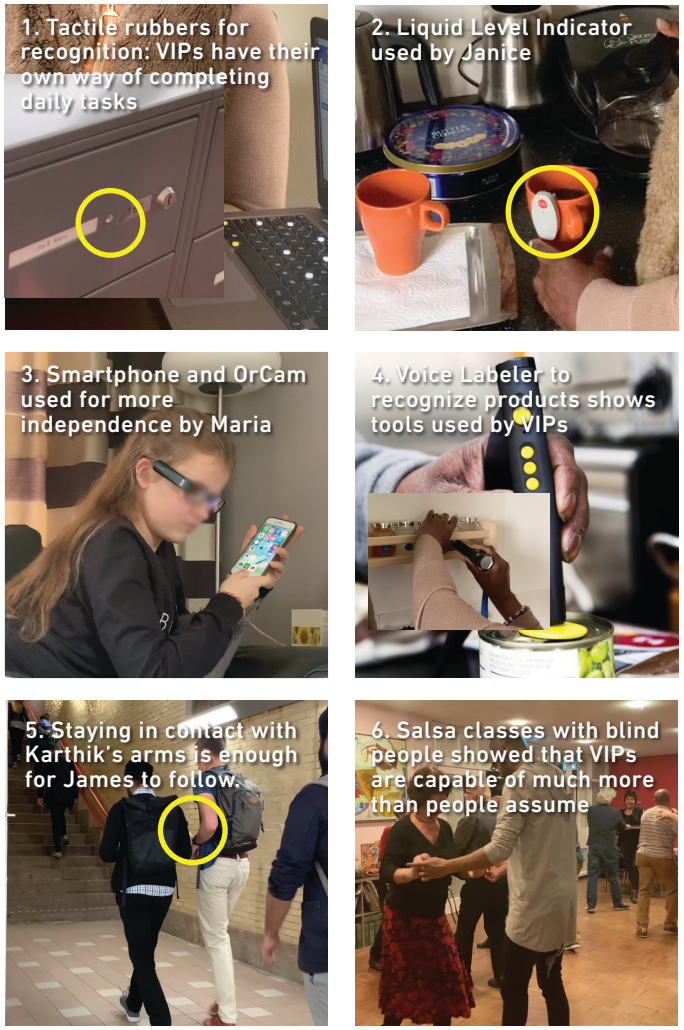


Figure 20: Pictures of immersion

they will still compare many daily tasks to the past where they would rely on their vision.

"The perfect world doesn't exist anymore. It's like gathering pieces of a puzzle, to complete a puzzle. Whereas the pieces are tools and software." - John

4. Difference between totally blind and having a bit of sight.

Another important element is the difference between being totally blind or still having a bit of rest vision. Immersion and interviews with VIPs showed that even with a little bit of rest vision, the VIPs would still as much as possible rely on their rest vision. A person that is totally blind does not even consider vision. This is an important insight because it means boundaries should be set for a final concept that it suitable for both visually impaired people with no vision and VIPs that still can rely a little bit on their rest vision.

5. Being blind requires a lot of energy.

Contrary to what people might think, VIPs consume much more energy compared to people with sight on daily basis. This is due to the fact that they have to think with each move whether they will not hurt themselves. As people with sight we rely so much on vision, that we don't have to think about running into someone. Being visually impaired means you are much more tired at the end of the day. Interviews that are conducted, shown in the next chapter, also showed that VIPs would have much less energy and also that they would bump more into obstacles after a long day. Therefore focus and clarity is of huge importance.

6. A little guidance leads to a lot of independence.

Another interesting insights is the fact that with a little bit of guidance, VIPs can be a lot more independence. This came to light when traveling with VIPs. Initially I thought much more guidance was needed, but figure 24-5 already showed that just a touch of the elbow is enough to navigate and for a VIP to follow.

7. No different than others

Although this insight is similar to the one gathered from the interviews, the observations once again showed that visually impaired people are no different than others. This also came to light when visiting the blind school for visually impaired students. Beyond the screen readers in the classrooms, it would be very difficult to say whether I found myself in a school for VIPs. This exceeds to a level where Maria uses Instagram, which is very visual, just because her friends use Instagram. And when asked if she would like it if her friends would put extra description for her, she mentioned that should be the job of Instagram.

"I don't want people to adapt to me, technology should." - Maria

8. The white cane as a symbol of visual impairment

Although research already showed that the white cane is also to let others identify the user as blind, this also showed

during the observation. When walking with James, people who saw James immediately stepped to the side when seeing him walk with his white cane.

9. Unforeseen obstacles

There are still unforeseen obstacles. For instance when arranging a meeting with Parmin in Amsterdam, it was immediately clear that it is still very difficult to meet at a location point without Parmin visually seeing me and knowing where I am. Likewise, with some other users the same was experienced and an exact location or meeting point was required to meet.

This insight once again shows the importance of not making any assumptions as stated before. What might be very common for a person with sight, might not be for visually impaired people. All of these insights should be taken into consideration when designing a product or service.

Besides that, short tests were conducted with visually impaired people to get a better understanding of their capabilities. One of this was to provide the VIP with 10 objects and ask whether they would know what it is. It showed that most items were easily recognisable by the VIPs by using their hands and feeling what it was. In both sessions packagings like a tea bag were not totally clear and the VIP used Envision and the OrCam to read what was on the package. This shows the significance of AI apps like Envision or a tool like the OrCam; portable, artificial vision glasses to understand text through audio feedback. Chapter G focuses more on this device.

Observation and general immersion with the target group allowed for direct insights. Additionally, it also helped to create an overview of the hardware and software used by visually impaired people and personas were created from these insights.

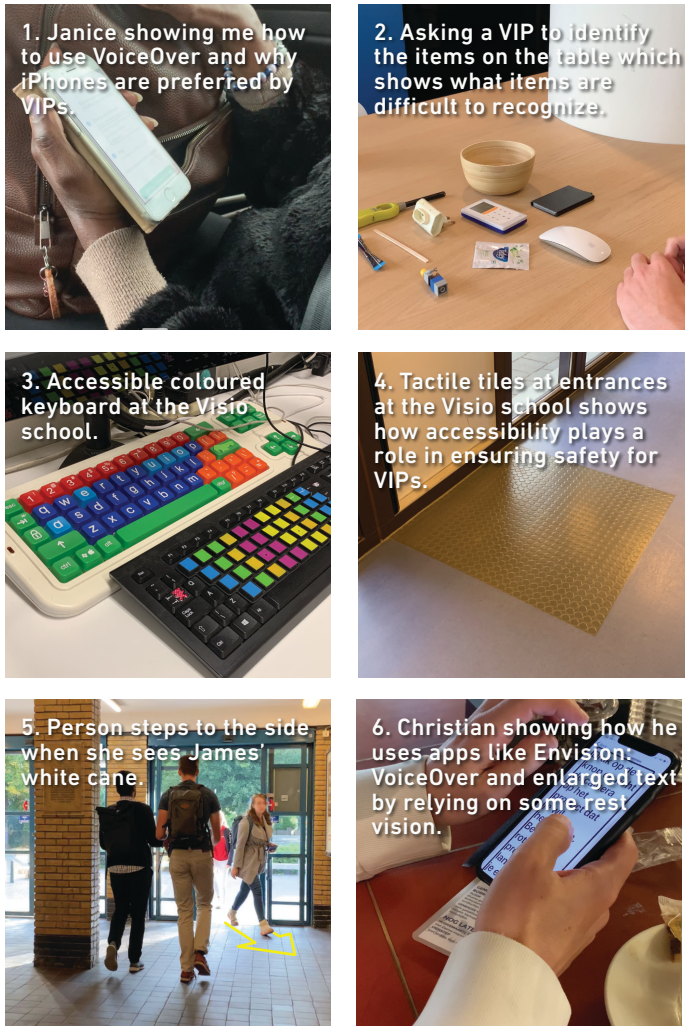


Figure 21: Pictures taken during immersion

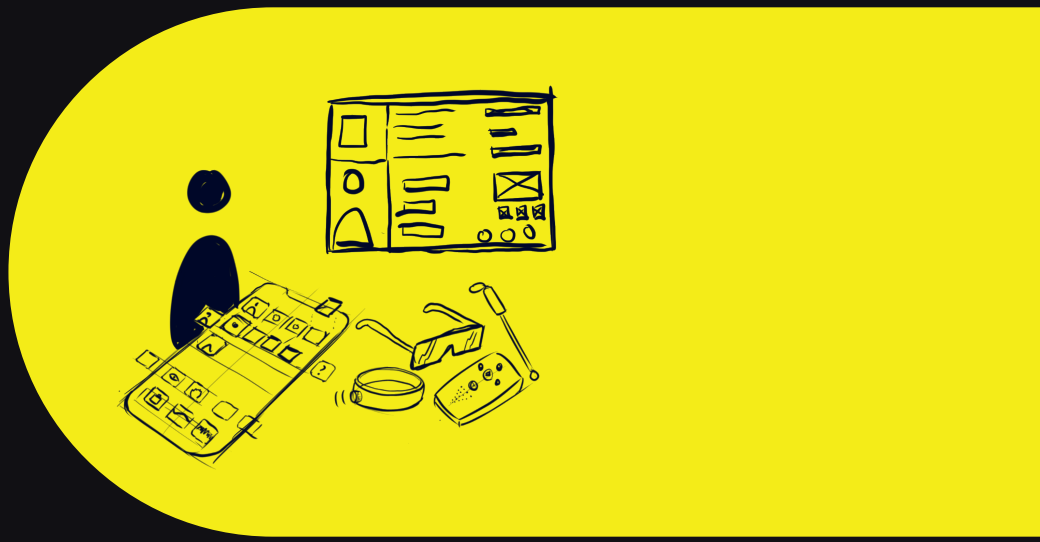
3. key takeaways and contribution

- Visually impaired people rely on technology as much as sighted people do.
- There is no all-in-one-solution for tools or apps.
- VIPs do not want any special treatment, nor do they want to be treated as different.
- Daily tasks take more energy as a VIP.
- Elderly people rely more on different tools, whereas younger people rely mostly on their smartphone.
- There is a significant difference between being totally blind and still having some rest vision.
- A little guidance leads to a lot of independence.

The insights gained from this contextual analysis contribute towards the research topic as it shows in what areas the VIPs can be helped and what a final concept should focus on. The concept should focus on guidance and see where it can replace existing (expensive) tools. Likewise, the interaction with the concept should not require a lot of energy (the amount of actions or the intensity of it).

d) synthesis

This chapter synthesizes the data from the background and field research. Additional research is conducted after the field research to find out about current technology (apps and tools) and needs of VIPs. Besides creating personas of the target group, this chapter also shows the first interesting directions.



in this chapter:

1. current technology for vips
2. needs
3. personas
4. key takeaways and contribution

1. current technology for vips

The previous chapter already showed the reliance on technology with the most important one being the smartphone. This reliance of technology is a huge factor as it makes sure that VIPs are mobile and independent; not having to ask someone else what train to take or ask to read a letter that they received.

But before smartphones, there already were correction devices and mobility aids: the history of lenses date back to 750 BC. The white cane was invented in 1921 and braille in 1824.

As stated in the previous chapter, elderly VIPs rely more on a lot of (expensive) tools such as a voice labeler or scanners. The immersion with the younger audience showed that all they used was their smart phone.

1.1 assistive technology

Field research and previous literature research Hersh et. al (2008) showed that assistive technology can be divided into the following categories:

- **Haptic aids**
- **Traveling aids**
- **AT for accessible information and communication**
- **AT for daily living**
- **Phone and tablet applications for blind and visually impaired people which will be discussed in the next chapter.**

The categories show possible design directions. Current technology also shows what VIPs need and what is used by them. Figure 22 shows an example of a voice recorder that is used for making lists or recording a certain route. The braille display is used to use a computer in a more productive way. A larger overview is shown in appendix 3.

1.2 apps

Field research with VIPs showed that the apps that are being used can be divided into the following categories:

Screenreaders for laptops, software that speaks out the content of a page. All VIPs had different preferences for what screenreader to use.

Reading apps, apps that were used to read books with Daisylezer, Voice Dream Reader and Amazon Kindle. All apps that are accessible with VoiceOver.

Milestone 112 Ace - Voice Recorder
(source: vision High-Tech YouTube)

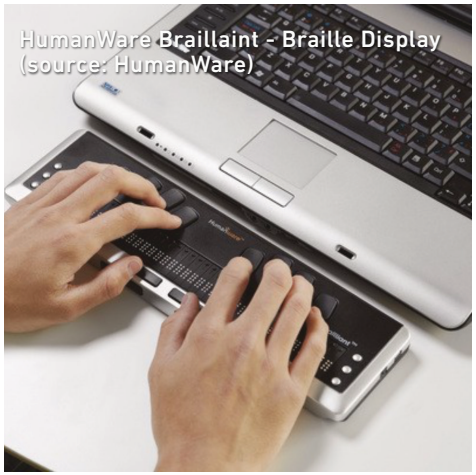

HumanWare Braillaint - Braille Display
(source: HumanWare)

Figure 22: Pictures taken during immersion

Transportation apps were used like many sighted people do. The only difference is the app *OV Info* used in the Netherlands. This is an app that is much more accessible compared to 9292.

Computer vision, the three apps that were mostly used to recognise text or objects were SeeingAI, TapTapSee and Envision.

Navigation apps, these were apps that were used to navigate outside and are mainly apps that are not needed by sighted people. With BlindSquare a visually impaired person is able to pinpoint a location on a map and knows what direction to go next. Soundscape by Microsoft turns location into 3D Sound so the user can navigate towards a location when wearing earphones.

Other apps include WhatsApp for messaging and video calling.

This overview shows once again what the concept should focus on and that accessibility is of huge importance. The final concept might have a great UX for sighted people, but that does not mean it also is great for VIPs.

An overview of these apps can be found in appendix 4.

2. needs

A human-centered approach to design a concept for visually impaired people requires to design for their needs and goals. Literature research, field research and the current technology showed what these overall needs are.

The needs are gathered and categorized into different categories:

- **Information:** the need for accessing (essential) information
- **Navigation/mobility:** the need to be mobile and go from A to B
- **Normative conformity:** the need to be accepted and liked on human level
- **Societal:** the need to be accepted and liked on societal level
- **Psychological/mental:** the need for being understood because of a visual impairment
- **Identification:** the need to identify objects and people

These needs are used throughout the design of the concept. The final concept focuses on needs like informing the user and helping with navigation and mobility. Additionally the concept focuses on identification, but makes sure that it does not look weird and looks accepted by others.

Figure 23 shows the categories of needs with examples gathered from user research.

Additionally, these needs are used in creating personas in the next chapter.

information

Having access to information, the need for...

...knowing the current location and time
...expiration date of food labels
...watching foreign movies
...reading letters
...knowing the correct bus stop and its arrival time

navigation/mobility

Going from A to B, the need for...

...avoiding obstacles and not bumping into things
...navigating outside of the house
...traffic lights with sound feedback
...sight for traveling in general
...walking straight

normative conformity

Being liked/accepted, the need for...

...not feeling *weird* with a white cane
...feeling the same as sighted people
...fitting in
...not being ignored
...doing/using the same things that sighted people do/use

societal

Regarding societal behaviour, the need for...

...accessible buildings, forms and websites
...feeling represented by society
...not being grabbed unwantedly
...non-overly helpful people

psychological/mental

The mind and emotional state of being/becoming blind, the need for...

...emotional support
...staying focused
...easier orientation
...not getting tired/fatigued

identification

Unability to identify things, the need for...

...knowing what an object is
...knowing the features of an object
...knowing the location of an object
...knowing who a person is
...knowing what a person's features are
...knowing what a person's emotions and facial expression are

Figure 23: An overview of general needs of visually impaired people gained from interviews, immersion and online research

3. personas

Personas are created based on the conducted online research, immersion with the target group and in-depth interviews.

The goal of the personas is to develop a concept based on the needs and goals of the VIPs. Additionally, the personas are created to express understanding and empathy with VIPs and also being able to communicate this with stakeholders. They helped to have a consistent and shared understanding of the users' values and needs (Boeijen, et al., 2017).

A few sessions at VIP also showed a specific target group that is not focused on during this project. This target group consists mainly of elderly people who went blind at a later age and still had some rest vision. Additionally, they mentioned the fact that they rely totally on their partner. This also meant that they did not or barely use a smartphone or knew how to use one.

Although truly interesting, designing a concept that uses AI for a target group that does not know the basics of using a smartphone or computer would affect the concept in a negative way regarding the timespan of the project.

The two personas also show the capabilities of the visually impaired and breaks stigma of people thinking that VIPs are not capable of doing much.

Staying in close contact with VIPs for over six months showed that they are capable of a lot more than most people assume. Technology should help to empower them and become even more independent.

Two personas are created; Christina and Amir.

3.1 persona amir

Amir is vey tech-savvy and totally relies on technology. He can't live without his smartphone and is always aware of the latest technology. He uses bone conduction headphones to be aware of the environment sounds and uses more accessible apps in comparison to regular apps. He loves doing extreme sports and does not think his visual impairment is a disability. Figure 24 shows the persona of Amir.

3.2 persona christina

In comparison to Amir, Christina is a little bit less tech-savvy although still relies a lot on technology. Whereas Amir mostly uses his iPhone, Christina uses more tools such as a voice labeler, a victor reader and a talking watch.

Additionally, Christina has a guide dog that increased her independence a lot.

Christina went blind at a later stage and still needs to adapt to being blind.

Although at times she struggles with being visually impaired, everyday she is trying to become more independent.

Figure 25 shows the persona of Christina.

Story

Amir studies Business IT at The Hague University. Every early morning he catches a **train** to university or his internship: accessibility testing at a big corporate. His dream is to be working at a big tech company and mainly work on accessibility for blind people.

After a full day of studying or working, Amir usually goes **running**, **sailing** or during the holidays he'll go **skiing**. And when he is not practising sports, he'll go for food with friends.

Amir was born blind but that does not keep him from doing whatever sighted people do.

Amir has no guide dog, but instead totally **relies on technology** with apps such as BlindSquare and Soundscape.

Amir

Uses daily

"can't live without"

iPhone X

Apple Watch

Bone conduction Headphones

White cane

WhatsApp

OV Info Travelplanner

Uses often

Braille Reader

BlindSquare GPS for blind

JAWS Screen Reader

Envision

Laptop

SoundScape 3D Sound

Independence

Tech-savviness

Housing situation
On his own

Civil status
Single

Blindness
Born blind

Position
Student

Education
Business IT

Age
21

Behaviour

- **Relies totally on technology**
- **Earger to learn new things**
- **Likes to do as much as possible on his own**
- **Does not miss sight because he never had it**
- **Tries to fit in and does not want to look or feel different**

Needs

- **Rely on technology**
- **Fitting in**
- **Stay focused**

Wants

- **Even more independence**
- **Technology that adapts to him**
- **Stop receiving unwanted help**
- **Doing everything that sighted people do.**

Figure 24: Persona of Amir, 21 year old business IT student.

Story

Christina is a sales manager at a big corporate in Delft. She used to be a project manager here, but after losing his sight she had to **relearn many things** which lead to a change in her job position.

She now works **part-time** instead of full-time and gets picked up by a **taxi**, instead of commuting by public transport. Her guide dog is always with her and whenever she really needs help she'll ask one of her colleagues.

Christina is from a small town nearby Rotterdam, where most people know her. She likes **reading books** and still loves biking. Although she can't ride on her own anymore after losing her sight, she still goes **cycling on a tandem bike**.

Christina

Uses daily

"can't live without"

iPhone SE

Guide dog

Bump dots for identifcation

White cane

WhatsApp

Talking watch

Uses often

Voice Labeler

Victor Reader Trek GPS + Audio

JAWS Screen Reader

Earcatch Audiodescription

Google Maps

Daisylezer Audiobooks

Independence

Tech-savviness

Housing situation
Husband / two kids

Civil status
Married

Blindness
Blind since 10 yrs

Position
Sales Manager

Education
Architect

Age
52

Behaviour

- **Loves technology but still struggles with it sometimes**
- **Finds it difficult to learn new things**
- **Does not mind help from others**
- **Compares herself often with sighted people**
- **Needs to adapt to being blind**

Needs

- **Needs some help with technology**
- **Help from others**
- **Relearn day-to-day tasks**
- **Stay focused**

Wants

- **More independence**
- **No change of environment**
- **Emotional support**
- **Society that adapts to him**
- **Help from others when needed**

Figure 25: Persona of Christina, a 52 year old sales manager.

4. key takeaways and contribution

- Current assistive technology can be divided into haptic aids, traveling aids, AT for accessible information and communication and AT for daily living
- Needs of visually impaired people can be divided into information, navigation, societal, psychological, normative conformity and identification
- Screenreaders are used by VIPs to turn text into speech and be able to use computers
- Most VIPs use apps like any other sighted person, whereas the more tech-savvy VIPs may use (more accessible) apps designed for VIPs
- Just like the iPhone replaced a lot of our tools, it did too for the VIPs

This chapter contributes towards the design goal in a way that it provides the first possible directions gained from the needs and current technology.

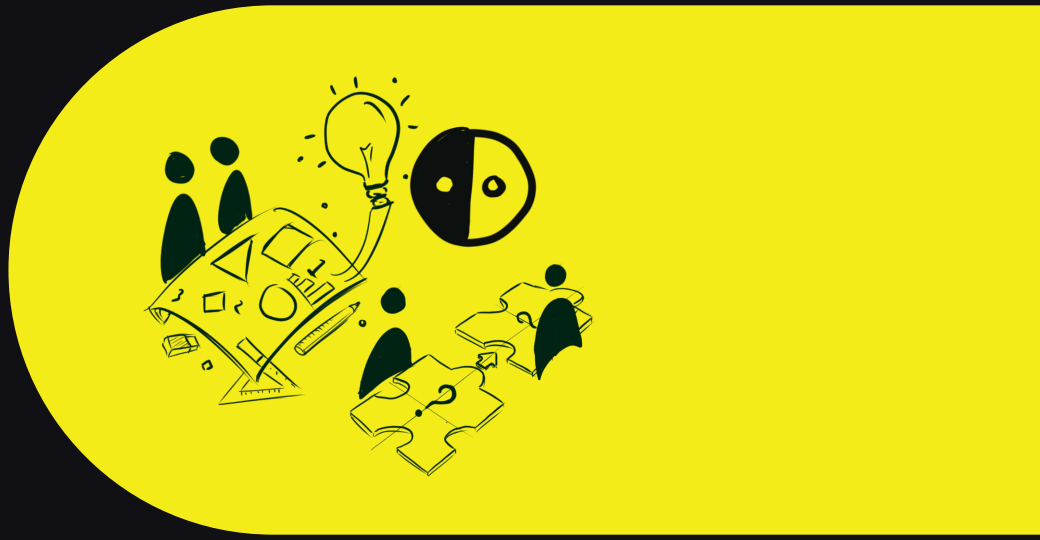
The needs show that the concept should focus on information, navigation and identification. Current technology shows that VIPs are already using apps that identify objects or help them with reading text.

The contribution of this chapter is understanding the users and coming up with possible directions: computer vision, human assistance, accessibility outside and online, mobility improvement and improved information gathering.

e) design direction

In the previous chapter contextual analysis was conducted and personas were created. This already showed a few interesting directions for the project.

This chapter focuses on research of Envision and speed dating sessions which provide the final direction and the design goal in the next chapter.



in this chapter:

1. speed dating
2. project requirements
3. envision
4. apps using computer vision for vips
5. key takeaways and contribution

1. speed dating

Speed dating is a design method for rapidly exploring application concepts and their interactions and contextual dimensions without requiring any technology implementation (Davidoff et al., 2007). Speed dating or so-called pre-ideation was conducted to gather insights about possible directions that would eventually lead to the revised design goal.

The speed dating sessions started with 18 ideas. Because of the large amount of ideas, only the two three VIPs rated all ideas and provided feedback on improving and discarding some of the ideas. This allowed to explore a large range of possible ideas and quickly narrow down with another three users to 9 ideas.

The sessions were conducted one-on-one by explaining the ideas verbally. The next chapters focus on what questions were asked and how it lead to interesting insights.

Appendix 5 shows all 18 ideas. These ideas differed from a focus on smart glasses to a smart phone application, smell, vibration, physical help and tools, services, buildings and internet.

After conducting the first part of the ideation discussed above (n=5), the ideas were narrowed down to 9 renewed ideas which were again discussed with new users (n=5).

The ideas were all asked in a way as *"Imagine there is..."*.

1.1 ideas

1. Smart glasses to have someone look with you: Imagine you have a pair of smart glasses with whom you can call random people or trained volunteers that are able to 'look' with you and tell you what you are looking at or help you with going from one point to another.

2. Application for improved scene description: Imagine there is an application that describes scenes by using your smartphone camera, but whenever the AI is not able to entirely recognise the scene, a trained volunteer can be called that explains the scene.

3. Smart material for object detection: Imagine there is an invisible piece of material that can detect obstacles. The closer an obstacle is to this piece of material, the faster it vibrates. You could add this piece of material to wherever you want on your body, so if you add it to the back and something is approaching you from the back it would vibrate quicker. Almost providing you with a 360 protective layer.

4. Magical button for physical help: Imagine there is a magical button that you can click on whenever you need help from a physical person. When you click on the button, one of the many trained volunteers in the Netherlands would come to help or guide you.

5. Insoles to help you walk straight: Imagine there are insoles that you can put in your current sneakers that will help you to walk straight. This will be done by hinting you with subtle vibrations that indicate whether you have to walk a bit more towards right or left.

6. Laser for distance and texture detection: Imagine you have a pointing device with whom you can detect distances and find out what the texture of materials is (e.g. you can detect the distance in front of you, but also see whether the ground in

front of you is muddy, consists of bricks or is wet).

7. Speech implementation in buildings: Imagine there are buildings which can detect your white cane and based on that will provide you information in speech about what building you are in and where you can go next.

8. Service for more accessible and standardised buildings: Imagine there is a service that standardises government buildings, making them more accessible. Accessible in the way of having the same tactile interior floor, same braille signs for rooms and toilets and the same appearance.

9. Online tool for redirection to accessible alternatives: Imagine there is an online tool that detects whether a website is inaccessible or misses information for VIPs, it gives you the option of redirecting you to similar more accessible websites.

Figure 26 shows quick sketches of the ideas. These were meant for my own understanding and discussing it with other sighted people.

This also showed to be a challenge to me as an industrial designer. Until now I would sketch out ideas and present this to the target group to make it clear what I wanted to create. With VIPs this was not possible and thus I had to be extra clear in my communication on what the idea would look like and do.

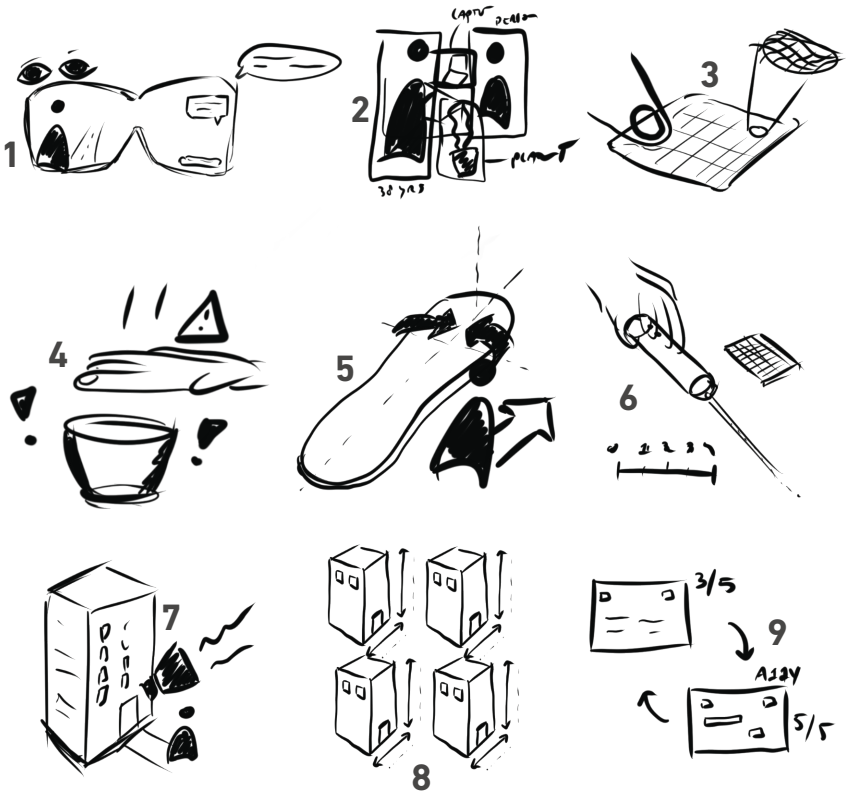


Figure 26: All 9 ideas discussed with VIPs.

1.2 Quantitative questions

Quantitative questions are asked to rate each idea. Although the amount of users is not very high, it still shows interesting directions, and are another reason to go for a certain direction.

The following quantitative questions were rated from 1 to 7, with 7 being the positive grade:

1. **On a scale from 1 to 7, how much do you like this idea?**
2. **How likely would you use something like this?**
3. **How exciting would this idea be if it was reality?**

The quantitative questions were asked to find out whether an idea was liked and if they would be excited about it. Figure 23 shows the results for these questions. The results show that the smartphone app and magical button is not liked as much as initially assumed. Likewise, it shows that standardized buildings and an accessible internet tool is liked a lot by the participants. The laser for detecting distances and textures also scored higher than average.

1. Standardization is key, but impossible.

The score for the accessible building corresponds with the earlier interviews, although it is liked very much by visually impaired people, they do not expect something like that to happen. Already during the interviews it became clear that the VIPs do not rely on the government to change and improve their lives, but instead technology.

1.2 Qualitative questions

Qualitative questions are asked to find out when an possible idea would be useful and how they would improve it.

1. **When would you use something like this?**
2. **How would you improve this?**

The last qualitative question is very useful as it turns the feedback sessions into co-creation sessions where the VIP is able to adjust an idea and tell how it could be improved. This provided great insights.

Qualitative questions helped to understand whether an idea is liked or disliked and why. Likewise, by asking where they would use the product, it was another confirmation whether their grade for likeability of using the product/service was justified.

2. Most ideas would be used outside.

At home, just like the previous chapter in research shows, everything is already clear and there are not many obstacles. It is after leaving the front door that problems or difficulties seem to occur.

"I'd use it when leaving the front door." - Christian

3. Needs to look normal.

The same insights that were gathered from the user research are again made clear by the VIPs here. In general the idea should not be looking weird and not make them stand out.

"It shouldn't look like a weird blind device." - Janice

4. Reliability is an important factor.

It was often mentioned that the ideas are interesting, but only if it works. I needed to mention "assume it really works in a way that I just proposed it". This is because of the many apps and tools that claim to be working in a way that usually does not meet expectations. This insight was gathered by using co-reflection as a method.

By all means it should be mentioned that the amount of users at this moment is very low that no direct conclusions can be drawn but it helps with the entire story.

2. project requirements

Besides insights from the questions, the ideas proposed during the speed dating were plotted against project requirements. These project requirements were chosen with the research project in mind and gathered from the earlier insights. Besides the criteria, also the importance was chosen, as the relevance of the solved issue is more important to the fact whether it fits with Envision as a company.

The following criteria is created with its corresponding importance level:

1. Relevance of the solved idea

How relevant is the solved idea? It is a small issue or a day-to-day struggle of a VIP?

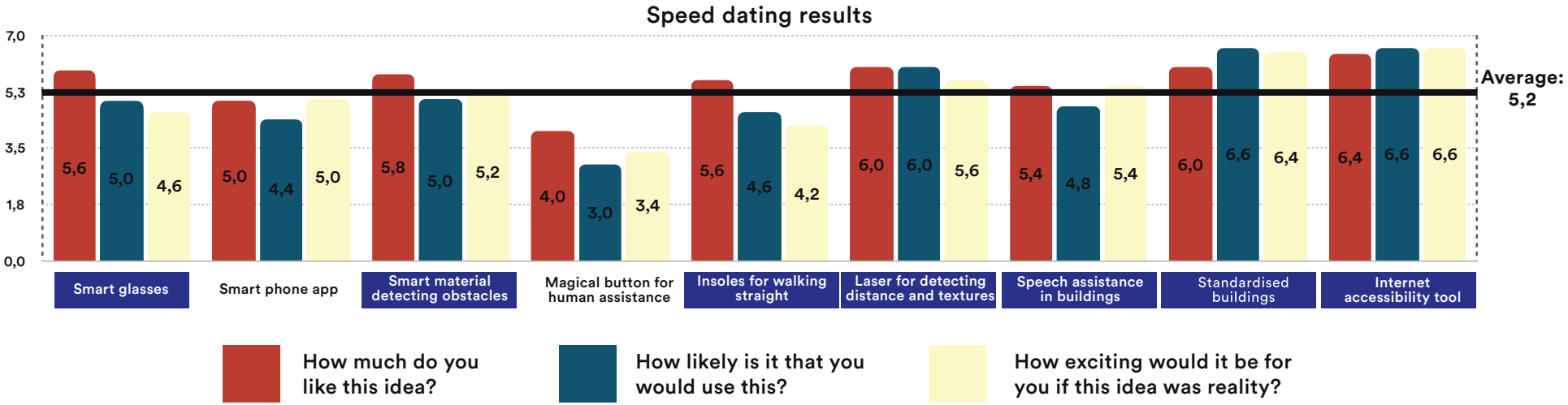


Figure 27: Speed dating results that show that a magical button for human assistance is not liked by the VIPs and that standardization is still wanted a lot.

[Importance: 2]

2. Implementation ease

How easy is it to implement this idea? This discerns the level of difficulty regarding technology and approval of external factors.

[Importance: 2]

3. Connection to AI

How much is the idea connected to or does it use artificial intelligence and therefore corresponds with the design goal to explore artificial intelligence for visually impaired people.

[Importance: 2]

4. Conventionality

How conventional is the solution? Does it make the VIP stand out from the crowd? This is one criteria created from the interview insights were most VIPs mentioned that they mostly would not wear something that looks silly.

[Importance: 2]

5. Size of impact

How large is the target group that can be impacted with this solution?

[Importance: 1]

6. Envision

How much does it fit with Envision, from a business and knowledge side?

[Importance: 1]

The average of the ratings are shown in figure 28. The individually chosen scores can be found in appendix 6.

The project requirements show that the three ideas fitted most are the smart glasses, the smartphone app and insoles for walking straight.

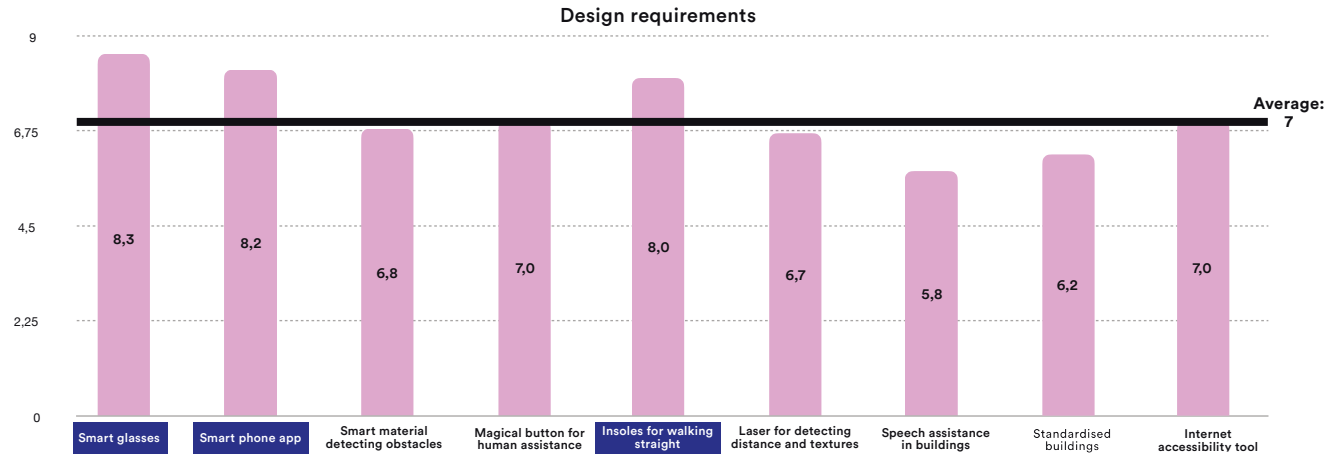


Figure 28: Design requirements based on each idea.

3. envision

The project remit was initially broader than to contribute directly to the Envision app. However, the contextual inquiry yielded many insights that were applicable to the Envision app. That is why at this point, I decided to select from the design directions those that contribute to Envision. To focus the project further, the Envision app is introduced in detail in the following.

3.1 What is it?

Envision is a smartphone app that empowers blind and low vision users to be independent by vocally describing the visual world around them. They have an app for both Android and iOS.

Envision uses mainly computer vision: an interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do.

3.2 Features

The features of the app are shown on the next page

1. **Instant Text:** Reads text instantly that is shown towards the camera and speaks it out
2. **Document scanning:** Helps the VIP to scan a document by edge detection and turns the document into speech
3. **Importing PDF/Image:** The possibility of importing PDFs or images allows the VIP to understand a picture that was sent to him through WhatsApp or any messaging app.
4. **Describe scene:** By taking a picture the scene is described such as what objects are in the picture.
5. **Detect colours:** By pointing the camera to an object, it speaks out the colours.
6. **Scan barcode:** The ability to scan a barcode so a VIP knows what a product in a super market
7. **Find people:** This feature is meant to find people, as soon as the camera detects a person it will provide a vibration.
8. **Find objects:** Just like the previous feature, this one is used to find objects ranging from a cup to a bench to a dog and more.
9. **Teach Envision:** By taking a picture of a person, you can teach the system a face so when using the find people feature it will speak out the name of the person.

Figure 29 shows an overview of the features of the Envision app.

3.3 Context of use

The Envision app is used in different situations and on different locations. The context of use is derived from user interviews, observation and online research. Being a graduate intern at Envision meant I stayed in close contact with the target group, knowing what they use Envision for and when they use it.

Some of the locations where Envision is used:

- **At work:** three of the ten visually impaired people mentioned that they

1. Instantly reading text from any surface
2. Reading complex documents and letters
3. Scan multiple pages and import PDF/images
4. Describe a scene
5. Detect colours of clothes and products
6. Scan barcodes
7. Find people
8. Find objects
9. Teach faces

Figure 29: All features of the Envision app as described in the text. Starting with the text tab, followed up by the general tab and ending with the find tab.

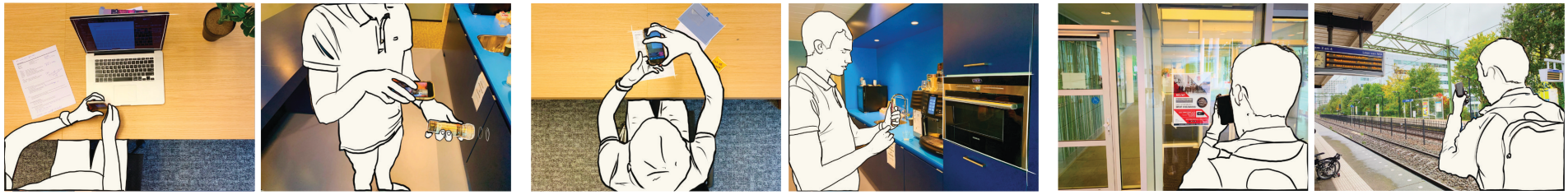


Figure 30: Context of use behind a computer, in the kitchen, at a desk, in the drawers, outside a building, outside on the train tracks.

Figure 31: Screenshot of a tweet by Morgan. She uses Envision to read a form she received in her mail.

Figure 32: Screenshot of a tweet by Steve. He uses Envision to read a greeting card.

would use the app to find out what was shown on their computer screen if the screen reader did not function. Peter, a visually impaired person, uses it to read labels on packages as seen on figure 30.

- **At home:** at home there are many use cases where the Envision app is used. For instance in the kitchen while cooking, to find out what spice is in a box. The immersion with Janice showed that the app can be used to detect colours and allow her to wash clothing without the need of someone else.
- **Outside:** Envision is used a lot outside. This can arrange from using the app at the train station to know what platform someone is at. But also pointing the camera towards an entrance, knowing what building a VIP is at or in what street even.

The context of use was also visible during online research and field research with the participants which can be seen in figure 30.

At this point of the research it was clear that computer vision could play a big role for visually impaired people as it replaces vision by using artificial intelligence.

Figure 34 shows that Jesse uses the Envision app to read the tracks and commute home. Though, he also mentioned that if something changed due to maintenance; he would video call his mother who would help him out.

Multiple VIPs mentioned that they prefer to do everything on their own as much as possible, but if it is not possible, than they would need some help from a human.

Whether that is a human physically helping them, or a human that helps them on distance with a video call.

This combination of using AI-powered tools in combination with the help of people led to the revised design goal.

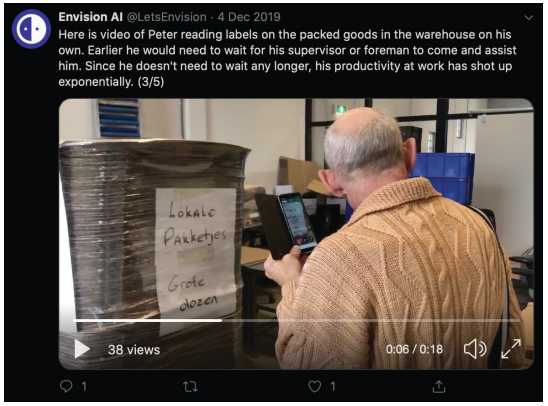


Figure 33: Screenshot of a tweet by Envision. It shows Peter who uses his phone to read labels on packed goods.

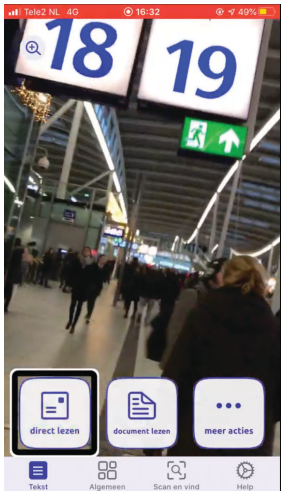


Figure 34: Screenshot of a video recording by Jesse who uses the Envision app to know at what platform he is at Utrecht Centraal.

4. apps using computer vision for vips

In the previous chapter I made an overview of the apps used by VIPs. In this chapter I introduce new apps that solely focus on computer vision apps like Envision.

This shows an overview of the features and what the apps focus on. By looking at the strengths and weaknesses of these apps, it shows what to focus on.

Apps created for the visually impaired is nothing new. Professor Jeffrey P Bigham introduced an app called *VizWiz*. It worked on the same principle as the currently available app BeMyEyes: a free app that connects blind and low vision people with sighted volunteers or company representatives for visual assistance through a live video.

Chapter D already showed that the App Store is filled with apps that make use of Artificial Intelligence and in specific computer vision. Besides that the previous chapter shows what the app Envision is capable of.

4.1 similar apps to Envision

An overview was created of similar apps that use computer vision to aid visually impaired people (appendix 7). The overview shows that most apps are meant for reading text or recognizing objects. The direct competitor of Envision when it comes to software is SeeingAI by Microsoft, both with 10 features.

Additionally, the list only includes apps that were actually functioning on the latest software (iOS13). Research into those apps showed that a majority of the apps are not reliable and doing what they claim. Besides knowing what kind of apps are on the market currently and how good they perform, it shows us in what stage of artificial intelligence we are that uses computer vision, whether that is to help visually impaired people or in any other field.

4.2 reliability

Besides knowing what kind of apps exist in the current market and how good they perform, it shows us in what stage of artificial intelligence we are that uses computer vision, whether that is to help visually impaired people or in any other field.

Reliability of artificial intelligence is a very important factor. During the speed dating sessions many of the users mentioned that most ideas would be interesting, but only if it was reliable and actually works.

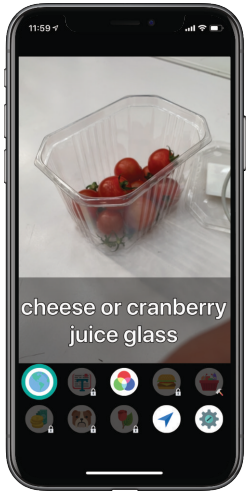
Figure 32 shows an example of a scene description made with different apps. It can be seen that every app is using its own AI engine and therefore provides different results.

This is of huge importance as scene description is getting better over time but some of the apps such as *Aipoly Vision* and *Triplens* are not doing a proper job at describing the scene at all.

Eventually this can lead to distrust and the VIP will not use it.

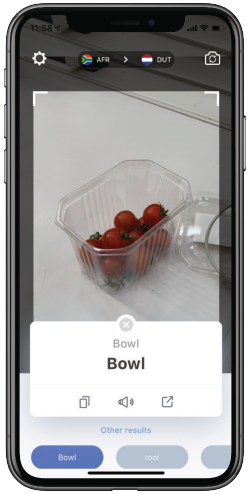
Besides reliability of AI and these apps, there are also tasks that are simply impossible with the help of some person. This can be physically someone helping a VIP or someone who is helping them with a video call.

Aipoly Vision



"Cheese or cranberry juice."

Triplens



"Bowl."

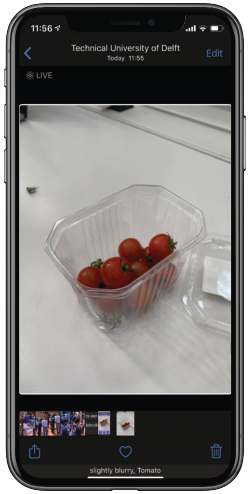
SeeingAI

"Probably a plastic container of food."

Envision

"It might be a bowl of fruit sitting on top of a plastic container."

Apple 'Siri'



"Tomato."

Figure 35: Picture taken with different apps which give different results.

5. key takeaways and contribution

- Standardization is key, but almost impossible.
- Inside the house VIPs do not struggle a lot, it is mainly outdoors.
- A possible solution should look normal
- Reliability is an important factor
- Computer vision is an interesting field for VIPs as it replaces vision that they do not have
- There are a lot of apps that claim to do something, but there is not an all-in-one-solution yet
- Envision and similar apps like SeeingAI is used in many different contexts by VIPs

This chapter shows that AI is not always reliable and as an intermediate solution human assistance should be used. Besides that it shows that a reliable system should be created and that the final product should not look odd and can be used outside.

f) design goal

The research started with looking for new use cases and developing a concept to aid blind people by using the possibilities of Artificial Intelligence. After conducting research in different fields, it shows that AI is not reliable and people are still needed. This chapter shows the revised design goal and the interaction vision for the final concept.

in this chapter:

1. revised design goal
2. interaction vision

1. revised design goal

As this is a Design for Interaction Master project, the overall aim is to find the design challenge that adds a benefit for the context of use. A concept being developed is evaluated according to this benefit. This can mostly translated into designing something that has a specific effect on someone in a certain way.

That *something* uses computer vision (AI) and human assistance.

The effect is the design goal, which is to step in when AI fails or when AI is not reliable enough to depend on.

Someone is visually impaired people and specifically blind people as stated earlier in the report; this is the group that can benefit most from the solution.

In a *certain way* can be translated into the interaction vision shown on the next page.

The goal is to use a combination of artificial intelligence and human intelligence. Human assistance can be seen as a VIP being helped by a sighted person that steps in whenever current AI technology is not able to fulfill a task for a VIP.

Design of a concept for visually impaired people that uses human assistance substitution when artificial intelligence is not reliable.

2. interaction vision

An interaction vision is a creative design technique to address already in an early phase of a design process the intended character of the interactions with a future design concept.

How should interacting with the concept be experienced by its user(s)? How should they feel when using it? Playful? Confronted? Energetic? Uncomfortable? And what means could designers apply to evoke or support this particular experience or feeling?

The interaction vision for this project was developed after the conducted research online and having talked to VIPs to get a clear understanding of their needs and wants.

The interaction with the product should be **reliable, convenient, empowering**. It should have the same feeling that you had as a kid walking or climbing on something while your father or mother would take your hand. The interaction vision (figure 33) shows that the kid can rely on his parent if something goes wrong. Besides that it shows the convenience of climbing up a tree. His facial expression shows the empowerment of climbing up a tree.

Convenient: Besides familiarity, the product should feel convenient. It should feel easy to use, without too many difficulties or distractions.

Reliable: The product should feel reliable to the user. Reliable in a way that the user is able to trust the device and could depend on it. But also in a way that if one solution does not function, it can rely on other solutions of the product.

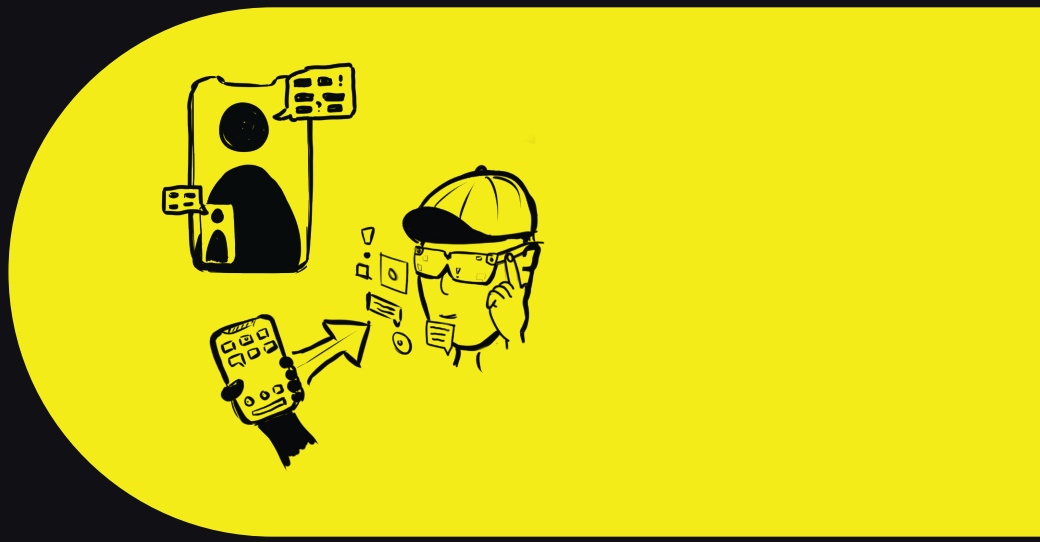
Empowering: Besides reliability, the product should empower the user: it should increase the degree of autonomy and self-determination, acting on their own authority. It should be a product that improves the user's quality of life.

The interaction with the concept should be like a kid being supported by his parent.

Figure 36: Interaction vision that shows reliability, convenience and empowerment by a kid holding his dad's hands.

g) focused research

To design a concept that uses computer vision and human assistance, focused research is conducted on apps that use computer vision for VIPs. Additionally, research on current similar tools (software and hardware) is conducted to gather insights and lead to a concept to test with VIPs.



in this chapter:

1. computer vision
2. human assistance
3. smart glasses
4. key takeaways and contribution

1. computer vision

In a pre-conceptualization phase, I conducted further focused research into the possible technological means to accomplish the design goal, while working closely with Envision. This replaced the wider ideation that normally follows goal forming.

As explained in the background analysis, computer vision is an interdisciplinary scientific field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do (Ballard, 1982).

To get a better understanding of computer vision, literature research was conducted. Among one of them a large-scale study by Bigham (2013), who presents findings of visual questions that blind people would like to have answered. Bigham's study lasted for over a year and more than 5.000 visually impaired people asked over 40.000 questions that they took with their phone using an app called *VizWiz Social*. And although this research is from 2013, it is still relevant for the final concept.

The study shows that the majority of the questions are identification questions (41%), followed by description questions (24%), reading questions (17%) and other questions (17%).

This shows the importance of features for the concept that answer identification questions (knowing what a certain product is) and description questions (what a person looks like or is wearing). An overview of this data can be seen in appendix 7.

Besides the previously conducted research on computer vision apps for VIPs, hardware solutions have been researched as well.

One of these solutions, known by many VIPs, is the OrCam (figure 38).

It is a device that can be attached to existing glasses and is priced at €4900,-.

OrCam is able to read texts, recognize faces and detect known objects. Thus, it does not offer as many features as Envision or Seeing AI. Though, research already showed that VIPs already have their hands full with a white cane and/or a guide dog. The reason to use the OrCam is because of its convenience.

The user can point to a piece of text and the OrCam will take a picture and turn the text into speech.

Figure 37: One of the interviewees with the OrCam.

Figure 38: OrCam device attached to glasses

2. human assistance

The previous chapter focused on solutions that use computer vision whether that is with an app or with a physical tool. To get a better understanding of human assistance and how it may help in a final concept, research was conducted to existing tool such as BeMyEyes and Aira.

2.1 BeMyEyes and Aira

Aira is a service that helps visually impaired people whenever they need help. Besides the app, they also have smart glasses called Aira Horizon. The VIP is able to wear their glasses and make a call in the same way that is being done with their app.

It is of importance to know that Aira does not use machine intelligence (AI) but mainly connects the VIP with a human. The only machine intelligence that is being used is that the agent is aware of the location of the VIP. This might help to tell the blind person where he or she is.

Aira is company based in the US and is also only available for people who live in the US, Canada, Australia or New Zealand. So officially not in the Netherlands.

BeMyEyes is almost similar to Aira, although it is a free service and does not have any hardware that improves the experience like Aira does.

BeMyEyes is a smartphone app that connects visually impaired people with volunteers who *lend their eyes.* Currently it has about two hundred thousand VIPs and 3.5 million volunteers that use their service.

2.2 Self-exploration with BeMyEyes

Self-exploration with BeMyEyes was conducted at different locations while being blindfolded.

The insights gained from self-exploration:

1. Although the task is simplified, it is not easy by all means

Living without a visual impairment means it is difficult to imagine how it is to live with one. Getting a coffee at the office immediately showed that even with someone else watching with you, it is not an easy task to fulfill. Although, it definitely made things easier or even possible to complete a task.

2. Having only one hand available

During self-exploration the first thing I stumbled upon is the fact that I was only able to use one hand, because with the other one I had to hold the phone. This is immediately an obstacle that is something to take into consideration.

3. Holding the phone correctly

Another insight from the self-exploration was to hold the smartphone in a way that it is easy for the other party to see what you are seeing. This was often really difficult and I had to be reminded many times to hold the camera straight. Besides that, one situation occurred where the helper was not able to see the touch screen due to sunlight as shown in figure 37 on the left.

4. Encouragement

Another interesting insight and which happened a few times during the calls was people encouraging you to do something. Somehow it felt good when another person complimented me on a task I completed. Although it has to be noted, that this might not be the case for VIPs.

5. It takes too much time.

Self-exploration also showed that it might take at least 10 seconds to get someone to help, up to a minute. Besides that, some volunteers picked up the video call and had to find a spot where they would be able to help.

2.3 Interviews with BeMyEyes volunteers

Another research focused on interviewing BeMyEyes volunteers (n=7). This was simply done by using the app BeMyEyes and asking whether the volunteer would be interested in participating in a short questionnaire. The questions were aimed to learn more about why volunteers would like to help, but also what their biggest challenge is when receiving a call from a visually impaired person.

1. **Why are you helping? What is the reason you are volunteering on BeMyEyes?**
2. **What is the maximum amount of time you would want to help?**
3. **Would you want to know more about the person who is calling you?**
4. **What is the biggest challenge for you as a companion to help a visually impaired person through the app?**

These questions showed that most volunteers simply love the action of helping another person. Four of the interviewees did have some connection with visual impairment, whether having a visually impaired friend or working in the field.

"I'm a retina specialist, we help blind patients. I like giving back. Also, since I work in that field."

Likewise it shows that as a volunteer for BeMyEyes you can always be contacted to

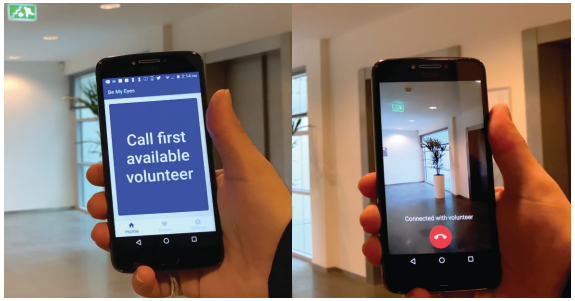


Figure 39: Self-exploration by calling a volunteer on BeMyEyes, asking where the elevator is.

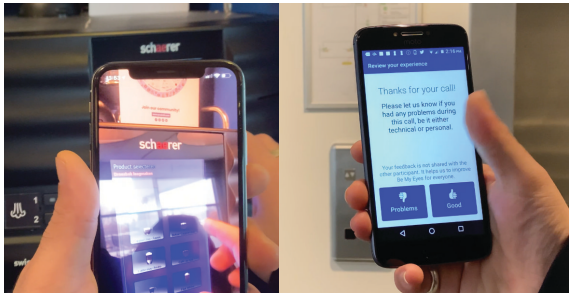


Figure 40: Left - Screen reflection while calling a volunteer, Right - End screen after a BeMyEyes call

Figure 41: Aira Glasses (Source: wearable-technologies.com)

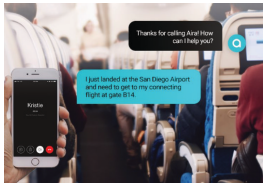


Figure 42: Picture of someone using Aira on the plane (source: www.aira.io)

help a VIP, which people did not mind. Although it should be added that there are more volunteers than VIPs using the app and therefore not many of the volunteers were called very often (one user mentioned that she did not receive a call yet after two months).

The maximum amount of time they would be able to help differed from 5 minutes up to 45 minutes with an average of 14 minutes.

Extra information of the VIP was not required and some of the challenges included orientation of the camera of the VIP but also helping with foreign documents.

"I had to figure out legal papers from the US, and I'm not from there, so that was challenging."

Due to the low amount of interviews, these results only show a few insights. To get a better understanding of unknown volunteers helping out visually impaired people, more research needs to be conducted.

2.4 Self-exploration with Aira

Additionally to the research with BeMyEyes, research was conducted with Aira. The insights of this analysis show the following:

More experienced

This showed in the way how they explained where something was. Instead of saying "It is a little bit on the right of you", they would use phrases like "Turn the camera two inches to the right." Besides that, they would initiate the call with their name so you know who you are talking to and asked what I would like to do today.

More technological solutions

Another insight showed the different tools that Aira agents have in comparison with BeMyEyes. During one session I claimed that a handwritten note with a phone number was left for me. The Aira agent asked whether he could take a screenshot so he had a better picture of the note. Secondly, the agent claimed that he would be able to speak up the number or that he would be able to text me the number. Additionally, it is possible to message the Agent beforehand what you are calling about which can be very handy if you are in a room where you can not talk out loud.

Implementation with other companies

Another difference with BeMyEyes are the implementations with other companies, such as Uber, Lyft and Microsoft. Aira agents are able to book an Uber for the VIP or as a VIP you can use Aira to talk to an expert from Microsoft when you have questions about their services and products.

Besides these differences, Aira is very similar to BeMyEyes and thus it was still very cumbersome to always have one hand occupied with your smartphone. Although, just like OrCam, they also have their own smart glasses which improve the experience of having a sighted person help out. The next chapter elaborates on this.

3. smart glasses

Research with Aira and OrCam showed advantages of smart glasses. At the same time Envision started exploring with bringing the Envision app to the glasses.

A minimum viable product with two features of the Envision app was created and tested with VIPs. This MVP was created by the engineering team of Envision and was a functioning Google Smart Glass. The first feature was the possibility of taking a picture and scene description being spoken out. The second feature being the possibility of taking a picture of text and having the text spoken out.

Figure 43 shows the VIP wearing the Google Glass. He taps on the side which is a touchpad and is able to take a picture of an item or a piece of text.

This MVP allowed early in the process to understand the interactions of VIPs using AI-powered smart glasses.

By working closely with Envision's team and having Visio as a stakeholder, I was able to test this first version with volunteers. In total three sessions were conducted where the volunteers tried out the smart glasses and provided feedback.

These sessions immediately showed the importance and advantages of wearing smart glasses contrary to using a phone, such as:

Handsfree experience

Already during the interviews and observations, it became clear that VIPs in general always have a white cane and sometimes even a guide dog with them. This showed to be a burden whenever they would want to use their phone. The smart glasses could offer a solution in a way that it is easy to use and interact with.

"I love the interaction with my AirPods, where I can tap twice on the side and use Siri."

Point of view

Self-exploration with BeMyEyes and video calling with VIPs also showed the difficulty in orientation and pointing the camera at the right direction. By having the camera on eye height, it is easier for another person to understand what the VIP sees, and aid him or her accordingly.

More information

Besides getting information from a point of view, most smart glasses also contain a larger angle camera which provides more information. This way the helper can be used more efficiently. Though, it should be mentioned that current smartphone cameras are also being equipped with a wide-angle camera.

Convenience

A lot of VIPs already wear glasses for protection. Smart glasses have the possibility of being a device that is always being carried with them. Additionally, it offers more convenience as the interaction can take place instantly by tapping on the side, whereas using a smartphone app requires getting the smartphone, unlocking it, opening the required app and using the wanted feature.

Figure 43: Reading text with the Envision smart glasses on products

Figure 44: Describing a scene with the Envision smart glasses

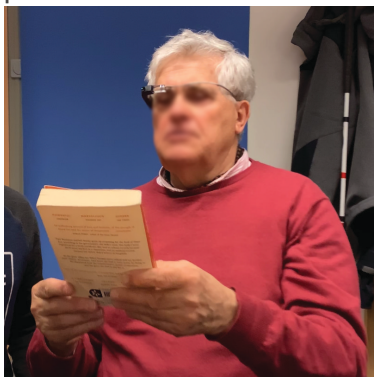


Figure 45: Reading text on a book

Figure 46: Scanning of a document to read the paper

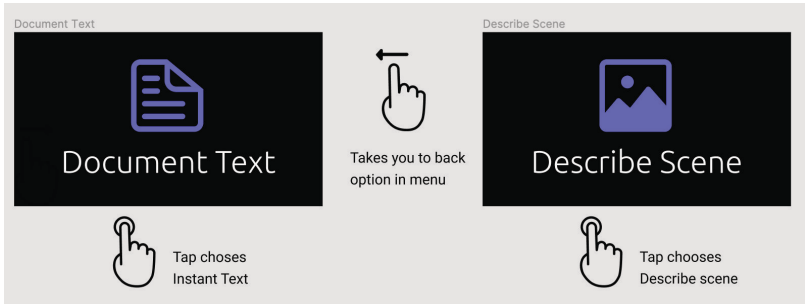


Figure 47: First version of the Envision app on the smart glasses that has the features of reading text (scanning a document) and describing a scene

Figure 43-46 on the next page shows the VIPs using the smart glasses. The sessions were conducted at Visio and with the *Haeghe Groep* (a pilot in collaboration with the municipality of The Hague). The current solutions and the sessions with Envision on the smart glasses showed my interest to create a concept based on the smart glasses.

The next chapter focuses on a concept made for the Google Glass (shown in figure 48). The Google Glass was chosen above the other smart glasses such as the *Eyecoming*, *Vuzix* and *Madgaze*.

This is due to the form and a requirement that it needs to look normal. Besides that, developing for the Google Glass is easier and the hardware is better than the other smart glasses.

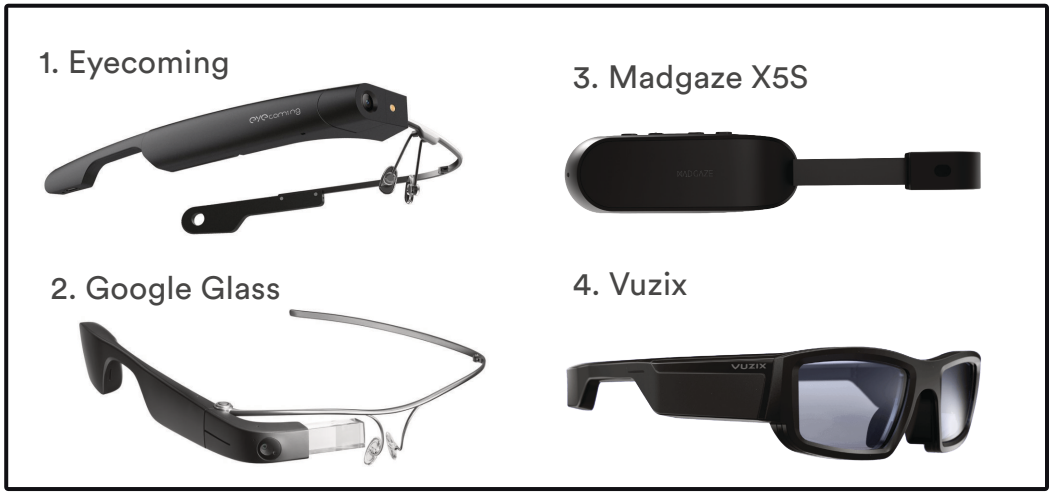


Figure 48: Overview of the smart glasses that are tested with Envision: Eyecoming, Google Glass, MadGaze and Vuzix

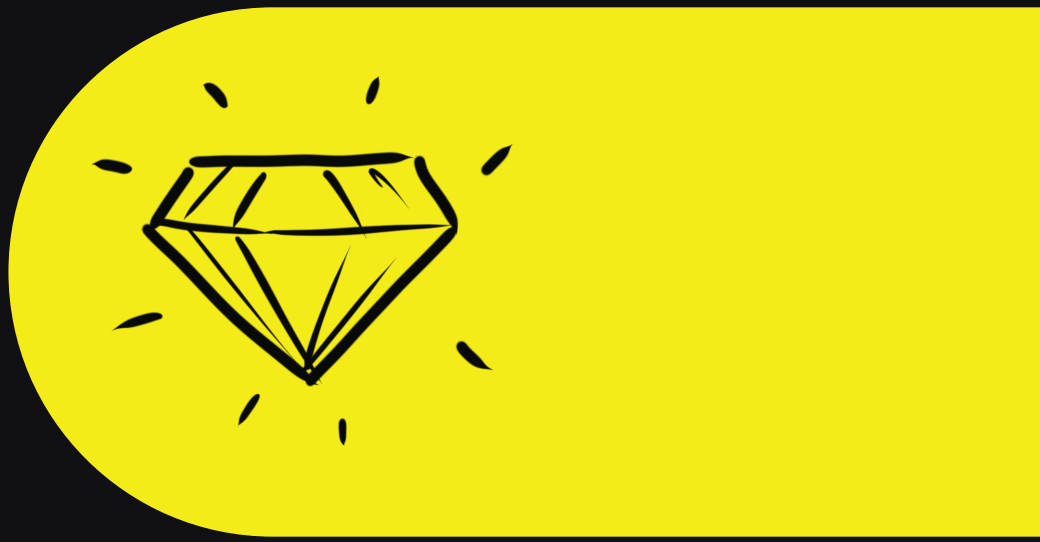
4. key takeaways and contribution

- Computer vision is used in combination with wearables such as OrCam
- Current human assistance for VIPs is being provided by Aira and BeMyEyes
- BeMyEyes is only used on a smartphone which might be cumbersome at times
- Aira is a paid service and always requires using a human
- There is currently no solution that combines artificial intelligence with possible intervention of humans
- Smart glasses are beneficial because of its handsfree experience, point of view camera, the amount of information it can provide and the convenience.

This chapter contributes towards the concept as it shows that it is of largest interest to design a concept on smart glasses and in specific the Google Glass. For the concept it is known who to design for as shown in the previous chapters, but also what the current market looks like and what to learn from them. Conducting user research with the MVP created with Envision provided many early insights that would be considered for the final concept.

h) conceptualization

This chapter shows the concept: AI-powered smart glasses that use human assistance whenever AI not reliable. The concept was tested both as a final concept and as an MVP created with Envision.



in this chapter:

1. prototyping

As mentioned in the previous chapter, the MVP of the smart glasses were tested with VIPs at Visio. During the first session the MVP consisted of reading text and describing scenes as seen in figure 47.

Together with Envision we worked on adding a video call feature to add human assistance to the smart glasses. Figure 49 shows this menu of the smart glasses created by Envision, which was tested during a second session with VIPs. This also meant that another app was needed for the sighted helper to receive a call from the VIP. In a very early stage, an MVP was created by Envision for a sighted helper to see what the VIP is *seeing*.

With the previous research and the two MVPs created by Envision, I started prototyping a concept of the smart glasses.

My goal was to improve interactions with the smart glasses and find out what is the best way to use the AI features and calling a person for help. I created a concept that used all features offered by Envision, while using the previously conducted research with the target group.

Figure 50 shows the Google Glass with its touch screen on the side and the screen. The screen is not of practical use as the VIPs only get feedback in the form of sound and they are not able to perceive anything displayed on the glass. The user has to interact with the touchpad on the side. The possible interactions are swiping up and down, front and back, tap and double-tap.

The prototype was created with a tool called ProtoPie. The main reason to use this prototyping tool is due to its accessibility to design for visually impaired people. Whereas all my previous prototypes were based on visual cues, this prototype needed to be build in a way as if the user was using a screenreader and could get feedback in the form of audio. None of the other prototyping tools where able to achieve this.

Prototyping allowed me to quickly test a design and gather feedback for a further design iteration.

Because the actual interaction touchpad is vertical, the users were told to use the phone vertically as shown in figure 52 and 53.

The prototype was tested with VIPs (n=5) at Visio.

Before testing the final prototype at Visio, earlier iterations of the prototype were tested at Envision (figure 52). This is also shown in the next chapters.

Due to time and complexity of the project, the other prototypes of the companion app and the onboarding are not tested with the VIPs. Thus, my recommendations shown in the next chapter also focus on conducting more research on these elements of the user experience.

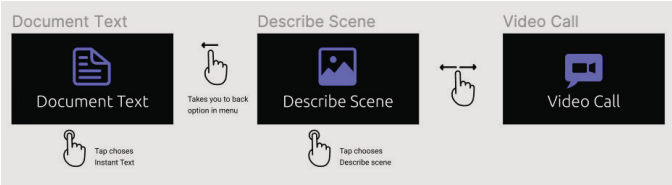


Figure 49: Envision glasses app during the second session

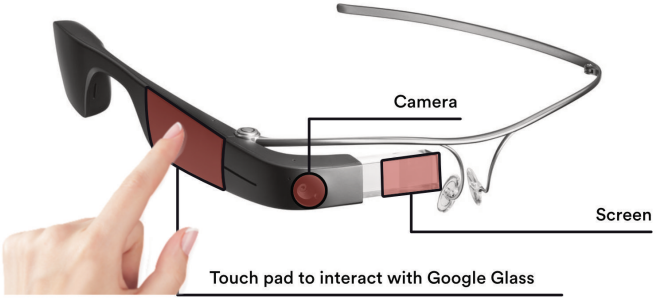


Figure 50: Google Glass screen, camera and touchpad.

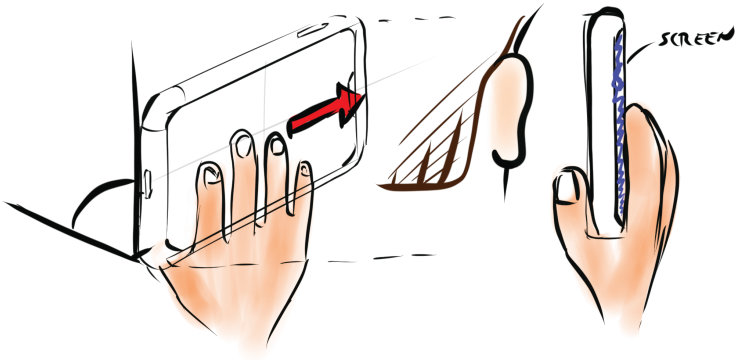


Figure 51: Prototype of the smart glasses tested in a way that the user holds the phone next to his ears vertically to mimic the glasses.

Figure 52: First design iteration tested by James

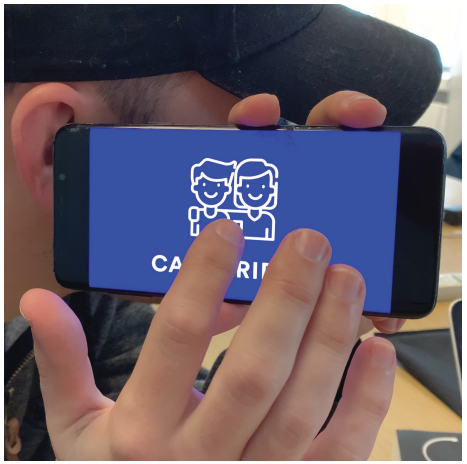


Figure 53: Testing the second design iteration with VIPs at Visio

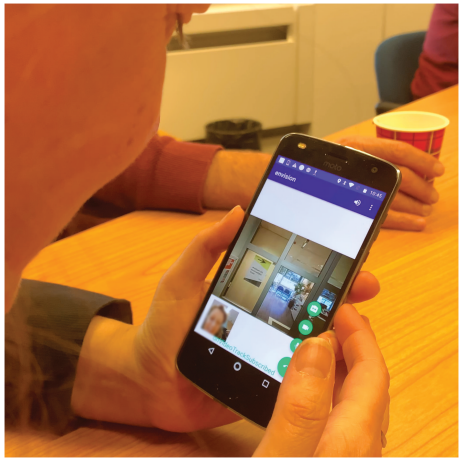


Figure 54: Gathering insights from sighted helpers on helping VIPs

Figure 55: Gathering insights from VIPs using the MVP

2. final concept overview

The previous chapter shows how the prototype is made and tested with the VIPs to iterate on a new design.

The first tests with VIPs already showed that interactions with the smart glasses can not be as complex as on a smart phone as the touchpad of the smart glasses is small compared to a smart-phone screen. This meant either looking for a way to tackle this complexity or creating another tool or app.

Research already showed that every VIP has their own preference of certain settings. Some prefer a really quick screenreader whereas others prefer a slower screenreader (the voice of elements being spoken out in a slower pace).

Creating a concept for video calling with smart glasses also meant that a concept was needed for the user who receives a call. This is also something that Envision was working on in a very early stage where it would only show the video input from the smart glasses on a separate app for the smartphone.

My concept combines knowledge gained from the Envision team, with previously conducted research with the target group and the tools BeMyEyes and Aira.

The in-depth interviews showed that there is not an all-in-one solution, with my concept I would like to tackle this issue.

It focuses on a concept that uses AI, but whenever it is not reliable the user should be able to get human assistance.

Therefore I decided to come up with a concept that is divided into three parts:

1. **Smart glasses app prototype:** the app where the VIPs use the features of the AI-powered smart glasses and human assistance on the Google Glass.
2. **Companion app prototype:** the app that is used by (sighted) helpers who can receive a call from a VIP and obtain additional tools to improve the AI.
3. **Onboarding app prototype:** the app that helps the user with onboarding and provides tutorials and settings regarding the glasses app. This can be advanced settings such as adding extra people as friends.

The next pages show an overview of the final concept. The consecutive chapters focus on the different elements of this concept.

The subchapter 3 to 6 focus on the smart glasses app prototype which was tested widely with the VIPs. Because this concept is not perceived visually by VIPs, the visuals are only meant as clarity for sighted people. Subchapter 7 deep dives into the interaction between humans and AI to see why the addition of human assistance is of such importance.

As a follow-up on this interaction, I created a reciprocal learning model (subchapter 8). This model explains how the AI can be improved and how to implement something like this in the companion app (subchapter 9). Unlike the smart glasses app, the companion app is used by sighted people and therefore the visuals are essential in understanding how to provide help.

In the companion app I also introduce a so-called Data Improvement System. Previous research showed that results

are not always accurate but can be improved if more data is available.

This Envision DIS system is a recommendation for the Envision team on how the AI can be improved and therefore the experience of the app and the smart glasses.

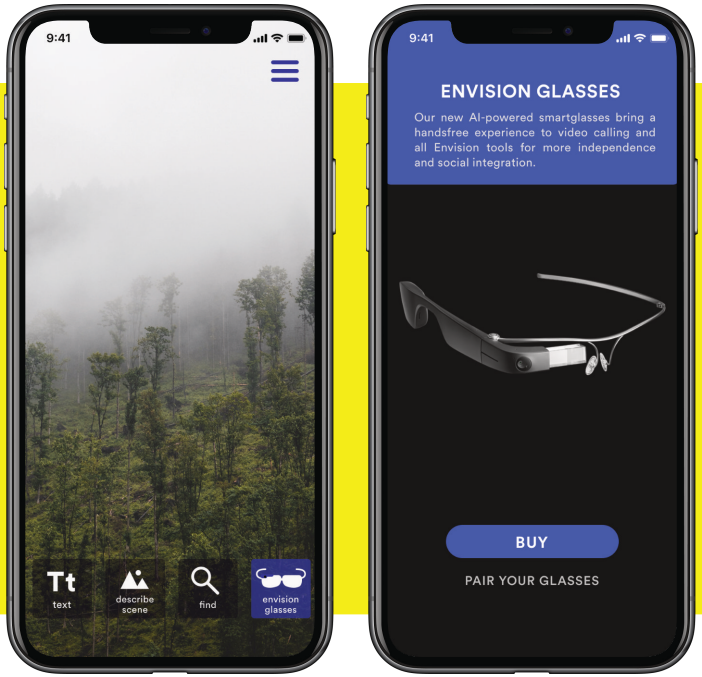
In subchapter 10, I created different scenarios that clearly show how the smart glasses are used in combination with the companion app. The scenarios connect the previous research with the created concepts.

It shows different causes of the AI failing, various reasons of calling a friend, volunteer or paid helper and situations gained from previous research with the VIPs.

Subchapter 11 shows the onboarding of the smart glasses.

Figure 56: Visualisation of the menu seen from the Envision glasses concept.

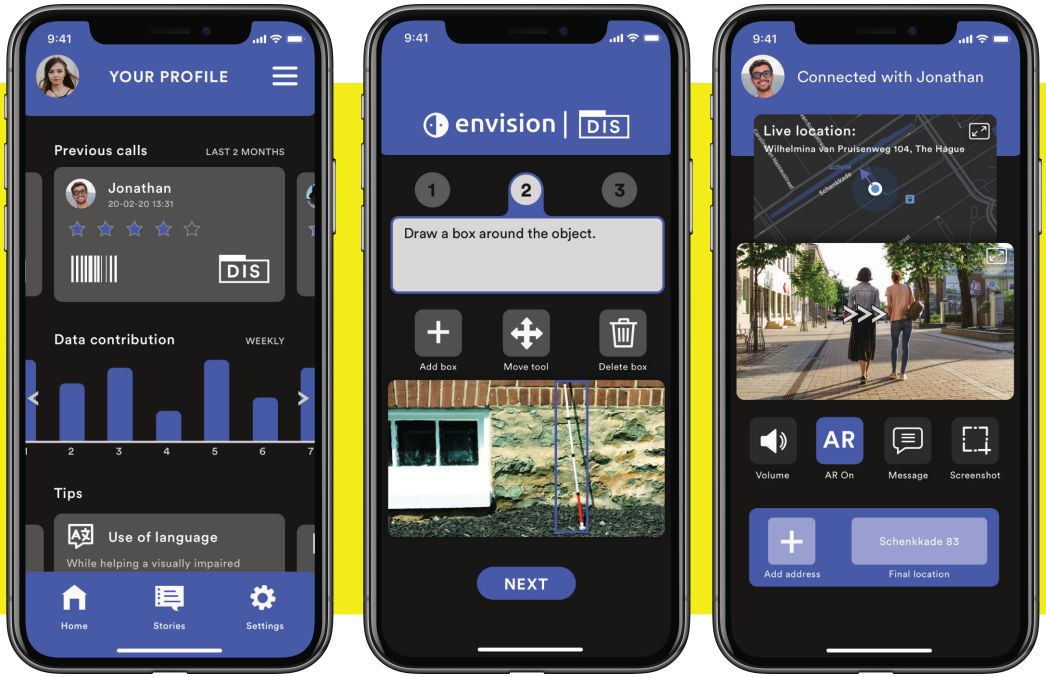
envision



current envision smartphone app for onboarding and advanced settings of the envision ai-powered smart glasses

Figure 57: The Envision app with the onboarding feature for the Envision glasses

envision COMPANION



envision companion app to help visually impaired people directly and indirectly

Figure 58: The companion app that allows to help a VIP and add data for AI improvement

3. navigation

3.1 categories

The interaction with smart glasses consists a lot of navigating through the different features. Two concepts were created to test with VIPs, to see which menu was preferred.

The first design (figure 60) focused on having all features next to each other and switching between these features by swiping front and back. The second design (figure 61) started with categories divided into the different features.

Both designs were tested by VIPs (n=3) and they all preferred the second design where the features are organized under the relevant categories.

3.2 swiping

As for switching between features, the original Google Glass uses swiping between features and this also seemed to offer the best clarity for VIPs while testing the glasses. This is also due to the form factor of the touch pad of the Google Glass.

3.3 infinity vs ending

Another importance is the scrolling through the menu. The first concept was created with an ending menu. VIPs mentioned that they preferred this type of menu above an infinite one. Though one of the VIPs mentioned that the ending sound was not totally clear whether it was an end or not. The next chapter focuses on these audio cues.

3.4 tap vs double-tap

Testing the smart glasses also showed known behavior by VIPs. As VoiceOver (screenreader) on the iPhone requires double-taps to enter a menu item, that is the first thing they indeed did. This lead many times to unwillingly entering a menu or using a feature.

Additionally, research also showed that the touch pad of the smart glasses are very sensitive. Sometimes a swipe to the front would end up being a tap instead.

Therefore, the second design changed the tap into a double-tap (figure 59).

Research was conducted with 5 VIPs. To confirm all tested details, a large study is recommended before the launch of an app designed for the Google Glass. This study should not only focus on people who used Envision before, but also those that are totally new to smart glasses or AI-powered apps like Envision and Seeing AI.

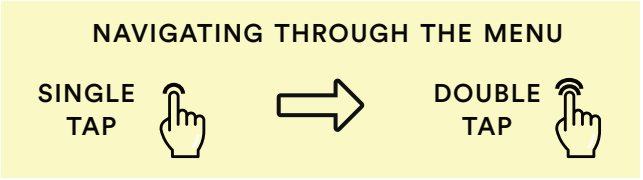


Figure 59: Change of a single tap to a double-tap for navigating through the menu.

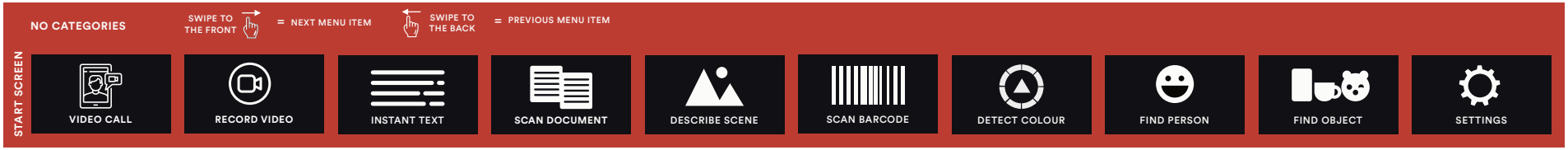


Figure 60: First design of the smart glasses app, menu without categories

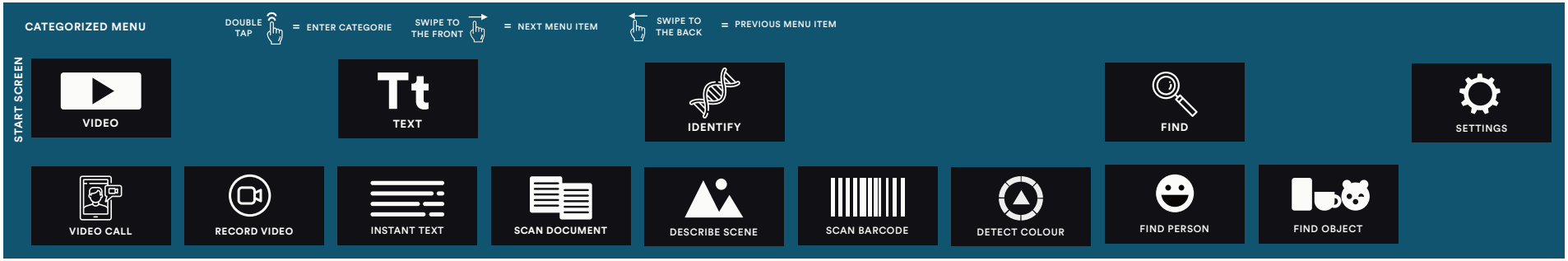


Figure 61: Second design of the smart glasses app, menu with categories

4. audio cues

Contrary to many products, the sound of the concept is of huge importance as the glasses are designed for visually impaired people. One of the VIPs mentioned that good UX for him means that the sounds are clear and always provide feedback whenever necessary.

Sound feedback was provided in the following elements.

- **Taking a picture:** some features rely on processing an image, therefore double-tapping means a picture is taken with the camera to be processed by the AI.
- **Video recording:** audio cues are provided whenever video recording starts and ends.
- **Navigating between features and selecting:** sounds that are played whenever the user switches from feature or selects one
- **Processing sound:** an auditory cue should be provided whenever the AI is processing data.
- **End of a menu:** an auditory cue to inform the user that he arrived at the end of a menu.

Taking pictures and video recording: the audio cues given for these two features are copied from existing sounds that smartphone apps already use. It showed that VIPs are familiar with these sounds and immediately it is clear what is happening when these familiar sounds were played.

Navigating between features and selecting:

As the VIPs rely on the screenreader voice, the first concept focused on leaving out any audio cues whenever the user would double-tap on a feature or swipe to another feature as the screenreader would already speak out where they were. Another design iteration focused on adding a very subtle tab-sound before the screenreader started (0.2s delay). The majority of the VIPs liked this feedback as it was a subtle but very clear reminder that their action (swipe or double-tap) succeeded.

Processing sound:

Another important sound is the processing sound. This is the sound that is being played whenever the AI is processing data before it provides a response. This sound is crucial as it tells the user that the AI is doing something and makes sure the user does not get confused.

This showed that in general the users liked sounds which were not too hard and more natural. Especially as the sounds are played constantly and should not be annoying.

Three different sounds were played and the users were asked which one they preferred:

1. Wood chimes (download link: http://bit.ly/hoai-wood)

2. Bell sounds (download link: http://bit-.ly/hoai-bells)

3. Natural sound sea waves (download link: http://bit.ly/hoai-waves)

The test with the VIPs showed that the first sound with the soft wood chimes was liked most by them.

"It is a soft and subtle sound. This is the least annoying."

End of a menu:

The previous chapter showed that VIPs preferred the design with an ending menu. Although for some the sound was not perfectly reflecting an end of a menu. Different sounds were played to ask what reflected more the end of a menu. Figure 62 shows where these sounds are played:

1. Subtle bass sound (download link: http://bit.ly/hoai-bass)

2. Reverse error pitch drop (download link: http://bit.ly-/hoai-reverse)

3. Futuristic hover sound (download link: http://bit.ly-/hoai-hover)

Once again, 4 of the 5 VIPs thought that the futuristic hover sound really reflected the end of a menu and was preferred over the other sounds. One of the early prototypes also

These subtle audio cues really show how it can affect the user experience of the concept. As VIPs rely a lot on sound in this concept, the sounds should be subtle and add something to the product. Additionally, control over the sound is of huge importance: controlling the volume of it and disabling any of the audio cues.

The user should be able to disable any audio cue and only rely on the screenreader, additionally the user should be able to only turn or off the processing sounds. This was given as feedback by the VIPs and also an option seen in other smartphone apps.

Figure 62: Showing an example of the menu where the audio cues are played for the ending of the menu.

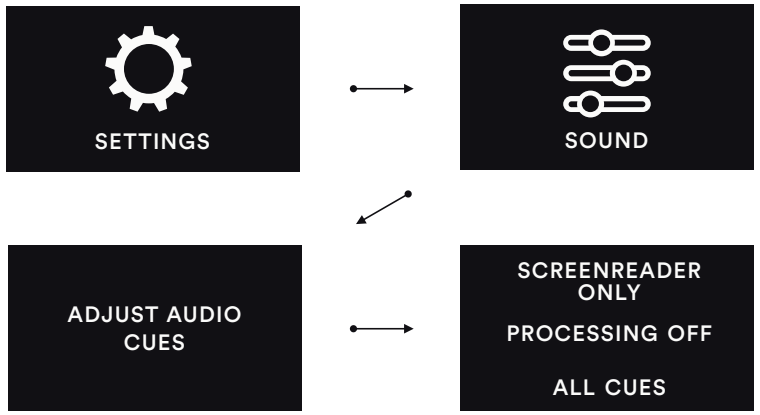


Figure 63: Showing an example of the menu where the audio cues are played for the ending of the menu.

5. information provision

The information that is provided should be in a way that it is clear for the target group to understand what each category and feature does.

As the smart glasses are designed for VIPs, the information should be provided with audio. Visual cues are redundant, especially since the screen on the smart glasses is too tiny for even most VIPs with rest vision to be helpful.

Information for the video is spoken in the next way:

"Video [1], Tab 1 of 5 [2], Videocall a person or record a video [3]. Double-tap to use a feature, swipe down to exit. [4]"

[1] Informs the user what feature or category the person is in

[2] Informs the user on how many more tabs there are

[3] Informs the user what features belong to this category

[4] Informs the user on the possible interactions

Providing this information can be done in different ways which was also tested with some of the VIPs:

1. **Information on the app:** No information is provided on the smart glasses, any information regarding the features and categories is all shown in the Envision app. With no information is meant that only the category or feature is spoken out [1].

2. **Always informed:** Extended information [1-4] is always provided on every feature and category.

3. **Two-finger tap:** Extended information is only provided whenever a user taps the feature with two fingers. Thus, tapping once would only speak out [1], whereas with two fingers it would provide all information [1-4].

4. **First use:** The first time the person enters a feature, extended information is given [1-4], whereas the other times it would only speak out the category or feature[1].

The VIPs all preferred the 4th option of have information on the first use only.

In addition, the clarity of naming the categories and features is equally important.

The first tests showed that most of the VIPs did not immediately know what was meant with an *"agent"*. When explained what that feature was, they all suggested and preferred something in the way of *"a paid helper"* as seen in figure 65.

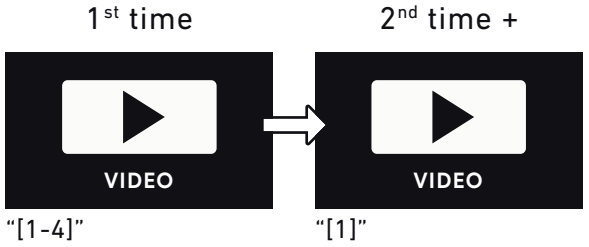


Figure 64: Information provision on the smart glasses

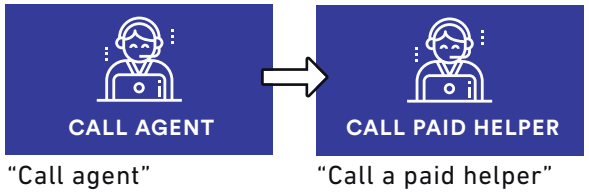


Figure 65: Naming features accordingly and clearly

Screenshot from the film: AI-powered smart glasses by Envision. This film was created with Envision and presented the presentation on the 6th of March 2020.

Taken by Sadjad Frogh (www.sadjadfro.com)

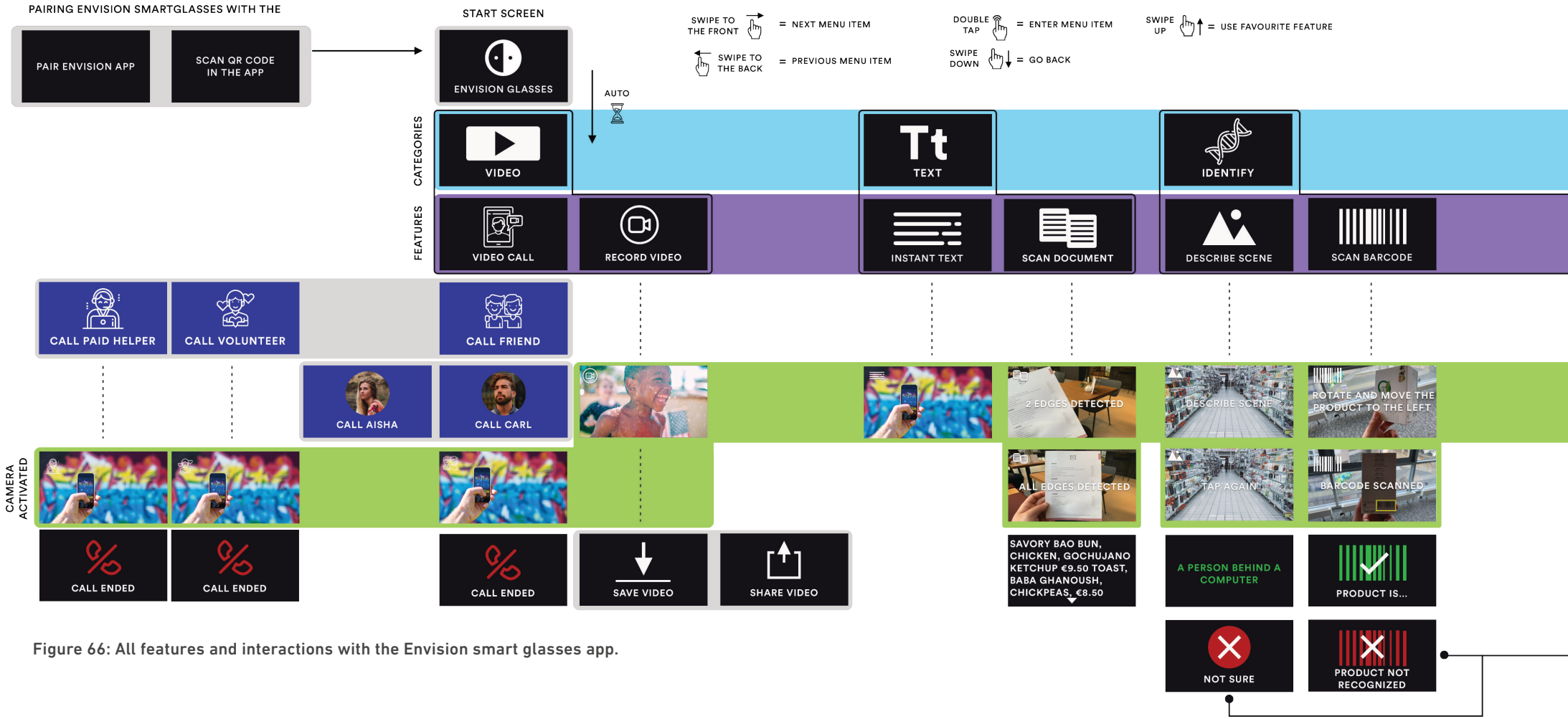
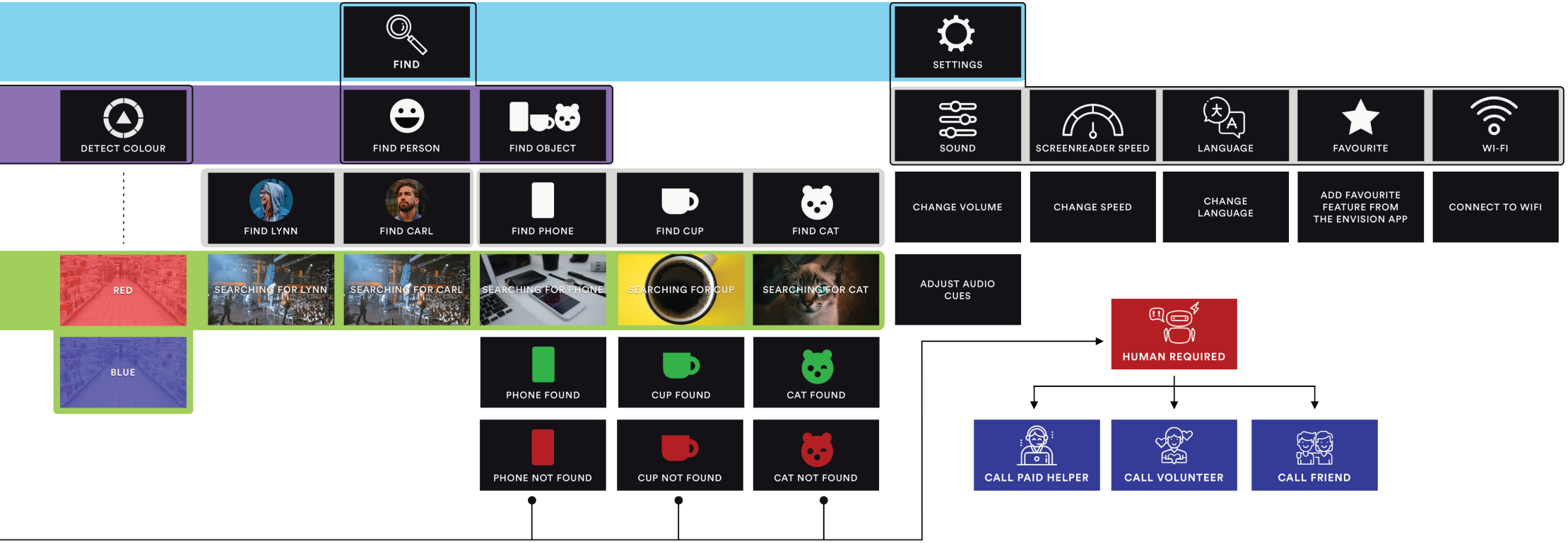


Figure 66: All features and interactions with the Envision smart glasses app.

6. smart glasses app flow

Figure 66 shows an overview of the smart glasses app features and interactions. Likewise, it shows what features connect the flow of the entire smart glasses app: the interactions with the smart glasses app, the features and how it connects to a possible intervention of human assistance. The visuals are created to provide clarity for sighted people. As shown in chapter 4, the users navigate through the menu with the audio cues and using the built-in screenreader.



7. human assistance implementation

The goal of these design interventions was to find out in what way the human assistance could be implemented and what possible scenarios play a role where AI is failing and therefore human intelligence/assistance is needed.

7.1 ai reliability

By staying in close contact with the engineering team of Envision, I was able to come up with a concept that involves the current Envision features, while implementing human assistance in the way of being able to video call another person.

AI features used in the Envision app are always coming with a reliability score. The AI system states in the back-end for instance whether *it is 80% sure* that it might be a mouse (or any other example). As a user you do not want to know what the percentages are, but you do want to be informed about the certainty.

Currently, the Envision app uses certain percentages in combination with the words that are being used. In the following example, assume that the Envision app is launched and the *scene description feature* is activated to take a picture of an apple.

81% - 100%: **It looks like** an apple.

61% - 80%: **Probably** an apple.

41% - 60%: **It might be** an apple.

0% - 40%: **I am not very sure**, but it might be an apple.

Whenever the AI starts using the words *It might be* or *I am not very sure* (thus an accuracy of lower than 60%) the result is not accurate. This was seen by taking pictures of random products and scenes. Figure 67 shows an example of a picture taken of a bin and each time the AI is not able to tell what the image is but instead thinks it is a cup of coffee or a suitcase.

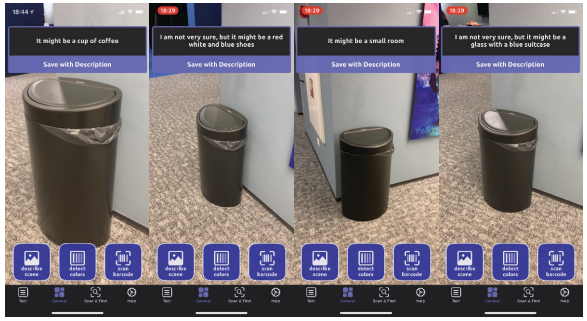


Figure 67: Screenshot of the Envision app using the feature Describe scene to take a picture of a bin

The current situation does not inform the VIP in a clear way that the AI might be wrong or provide the VIP another solution.

7.2 why we should implement human assistance?

There are different causes why to implement human assistance in a device that is capable of providing solutions with AI.

1. Not possible with AI: there are tasks such as avoiding obstacles on a way when going from one location to another that are (yet) impossible for the AI. Self-driving cars are becoming reality, but applying the same technology to pedestrians still needs a lot of research and development. Scenario 1 shows the interactions on how to tackle this situation with human assistance.

2. AI fails: the AI is not able to provide a solution due to lack or shortage of

trained data. Scenario 2 on page 81 shows a possible situation where this could be the case. Additionally, this also includes failures of the system. During self-exploration and user testing, now and then the app would stop recognizing text or describe scenes in a very *strange* way.

3. AI is not accurate enough: This shows for instance in reading texts. It might read the text of a menu, but it will not be accurate in a way that the AI is able to understand which dish correlates to what price and what exactly the headings are. This is shown in scenario 3.

In addition to the possible causes and human assistance as the only solution, there are also reasons for VIPs to prefer human assistance over AI:

- **Reliability:** being able to rely on a person and being assured of a correct result.
- **Speed:** it would be possible to find out what a product is with the reading text feature, but if time is an issue, video calling might be preferred.
- **Convenience:** simply the fact that you can call someone and that person will help you out.

Another reason to use video calling over an AI solution is to actually get in touch with a friend; showing another person where you are or what you are *seeing.*

The different scenarios are shown on page 80 to page 83.

7.3 human-ai interaction

The scenarios are shown in chapter 10.

Scenario 1 shows where AI and human intelligence *work together* for a better solution.

Scenario 2 shows where AI fails and human intelligence provides a solution and improves future AI.

Scenario 3 shows where AI is not accurate enough, and human intelligence provides a better solution.

Scenario 4 shows where AI firstly tries to get a better result by guiding the user, afterwards the AI still fails and switches to human intelligence.

The user of the smart glasses is guided towards using human assistance in three different ways:

- **By own choice:** the user switches to the video call feature and calls a person because of the aforementioned causes or reasons. This does not require any interference by the system.
- **Accuracy of AI fails:** the AI is not accurate enough and therefore hints the user to use video calling instead. The interference happens after the AI is not able to provide a solution that is of any use to the user.
- **Manual:** the AI is acting *weird* or it is not doing what the user wants, in this way the user manually swipes down to exit the menu and go to the video call feature. The interference here also happens on *human-level*.

Human assistance is divided in three categories:

- **Call a friend:** this is a friend or family

member of the VIP that has the companion app installed on his/her smartphone.

- **Call a volunteer:** this is a person who voluntarily applied to help VIPs on the Envision Companion app. The VIP can choose to provide all data and tools or only provide the volunteer with a video feed.
- **Call a paid helper:** this is a specialist provided by certain companies to help. They are being paid, but are also available 24/7. Mostly a paid helper is helpful for more difficult tasks or those that are related to a certain specialized field. Some examples: installing a Windows desktop with the help of a Microsoft specialist, understanding a new oven with the help of a Bosch employee.

These differences can also be seen in the scenarios.

Immersion and in-depth interviews with the target group already showed the different preferences regarding video calling people for help.

Some VIPs preferred calling a friend or family member because of trust and ease, while others preferred calling a volunteer and not *bothering* their family members.

"I don't want them to think, 'oh is he calling again?'"

In other cases some VIPs did not have a lot of family and thus already relied on volunteers and apps like BeMyEyes.

Additionally, some VIPs preferred paid helpers because of their experience and the fact that they are getting paid for it so they would not feel bad about it.

The different scenarios show the interactions on both the smart glasses user and the companion app user.

Figure 60 shows an overview of all features of the smart glasses and the possible interactions with it.

8. reciprocal learning

The scenarios on the previous pages already showed the differences in how a task is completed by a visually impaired person: with the use of artificial intelligence, by a human or a combination of both.

The scenarios show that people can help the AI or that the AI helps people *directly*.

Though, another possibility for reciprocal learning is the guidance of the user that helps the AI and therefore itself (figure 68).

Previous research already showed that a lot of times VIPs only need a little bit of guidance to get certain tasks done. Additionally it showed that sometimes the AI, or computer vision in this case, needs to be *steered* correctly for it to work.

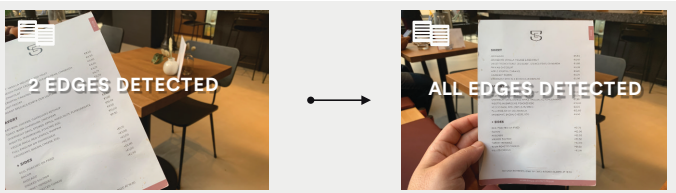
One example can be seen in the scanning document feature of the Envision app (figure 70). This tells the user whether all four edges are detected and therefore provides better results.

An improvement on this feature is to tell the user what edges are detected and provide feedback on how to move the paper (figure 62). The AI guides the user towards the best outcome, which leads to a reciprocal learning of both the human and AI: it detects that both left edges are not visible and therefore the paper needs to be put to the left.

Additionally, this system can also be used when a user would like to scan a barcode with the glasses (scenario 4, page 83). The AI is able to detect that there is a barcode available in the field of the camera on the smart glasses, but it needs a better angle to be accurate. The AI guides the user into getting the barcode placed correctly in front of the camera. This allows the AI to read the barcode correctly and provide the user with the product name and details (if the product is in the system).

A model was created for the improvement of this AI to show how AI and humans can work together for an improved solution (figure 71).

Figure 73 shows how a sighted helper that provides data to the AI which in turn helps the VIP.

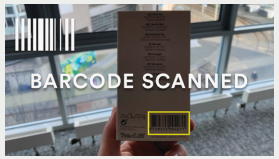


"Top-right edge and bottom-right edge detected. Move the paper to the right."

"All edges detected, scanning document."

Figure 68: AI guiding the user to take the best picture for an optimal outcome

"Barcode detected, rotate and move the product to the left."



"Barcode scanned correctly."

Figure 69: AI guiding the user to scan the barcode appropriately.

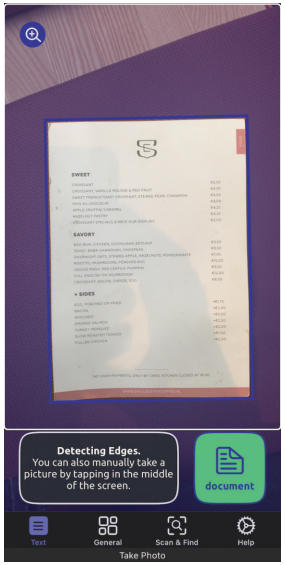


Figure 70: Detecting edges to scan a document on the Envision app

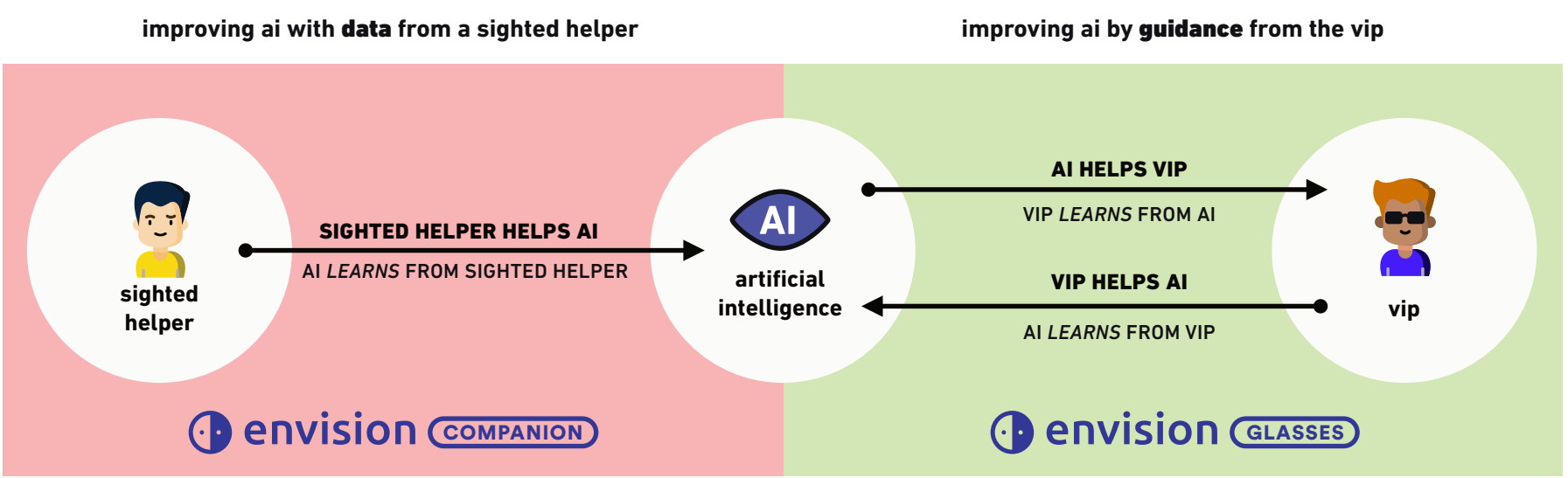


Figure 71: Reciprocal learning model between AI, sighted helper and visually impaired person as a user of the smart glasses

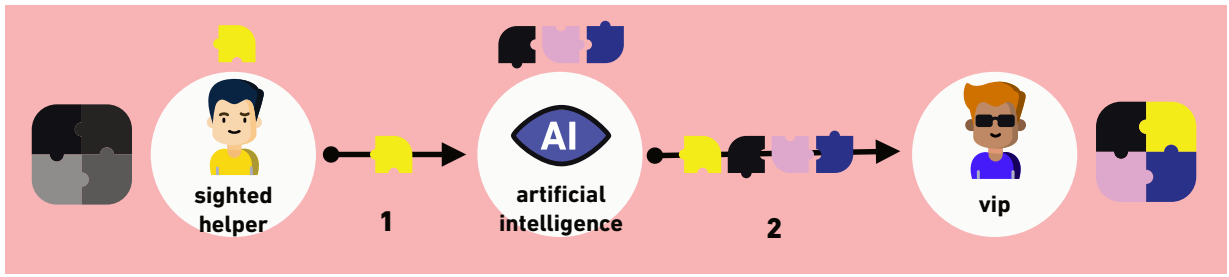


Figure 72: Sighted helper improving the AI by providing data added in the companion app which improves the solution for the VIP.

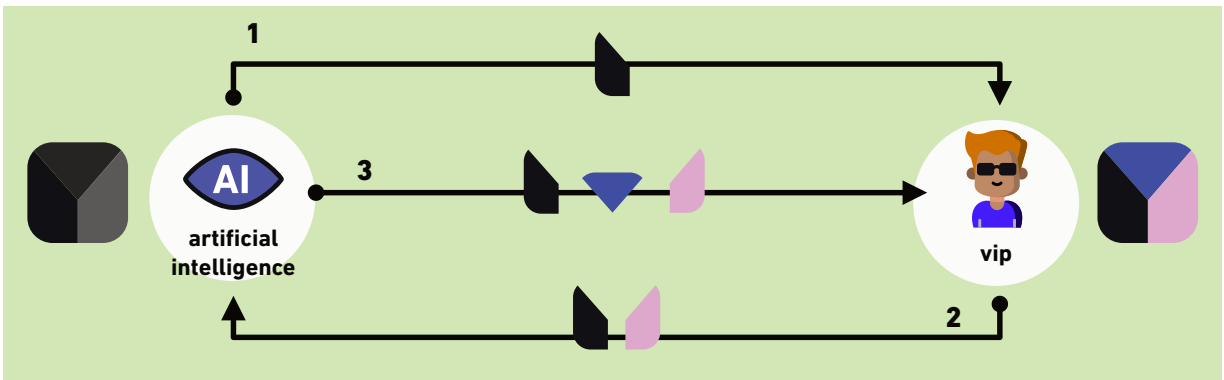


Figure 73: AI provides an unreliable solution to the vip, by guiding the AI, in specific computer vision, the AI improves which provides a better solution to the VIP.

9. companion app

The scenarios on page 80 to 83 show screens of the companion app. The goal of this companion app is to help a visually impaired person as much as possible. Additionally a companion can choose to be a helper only for (approved) friends and family or also be a volunteer for any other VIP that would like help from a random volunteer.

Previous research regarding these *helpers* showed that they like contributing towards a good cause, but that helping a VIP also requires some training. The companion app could offer this by showing their data contribution (figure 68, screen 4 and 5) or the amount of people they helped and tips on what language to use (Figure 68, screen 2). One of the last sessions at Visio showed that short and clear language is needed to guide a VIP from A to B.

Chapter 7 already showed that people can be a companion as family or friends, a volunteer or specialized help.

Besides the information for the visually impaired person, the information provided to the companion is equally important for a seamless user experience and the reason why video calling with the glasses is superior to video calling with a smart phone.

The scenarios show that as a companion you can have different tools to help a VIP.

Tools such as:

- **Video recording:** video recording the entire conversation. This can be handy to re-check the video later to tell a friend where he has to be careful next time when commuting.
- **Screenshot:** besides video recording the entire conversation, screenshots are a nice way to pause the screen and zoom in to get more detail. This is handy for checking a number.
- **Messages:** the messages section is a great way to provide the VIP information in written format. This way long pieces of information are accessible

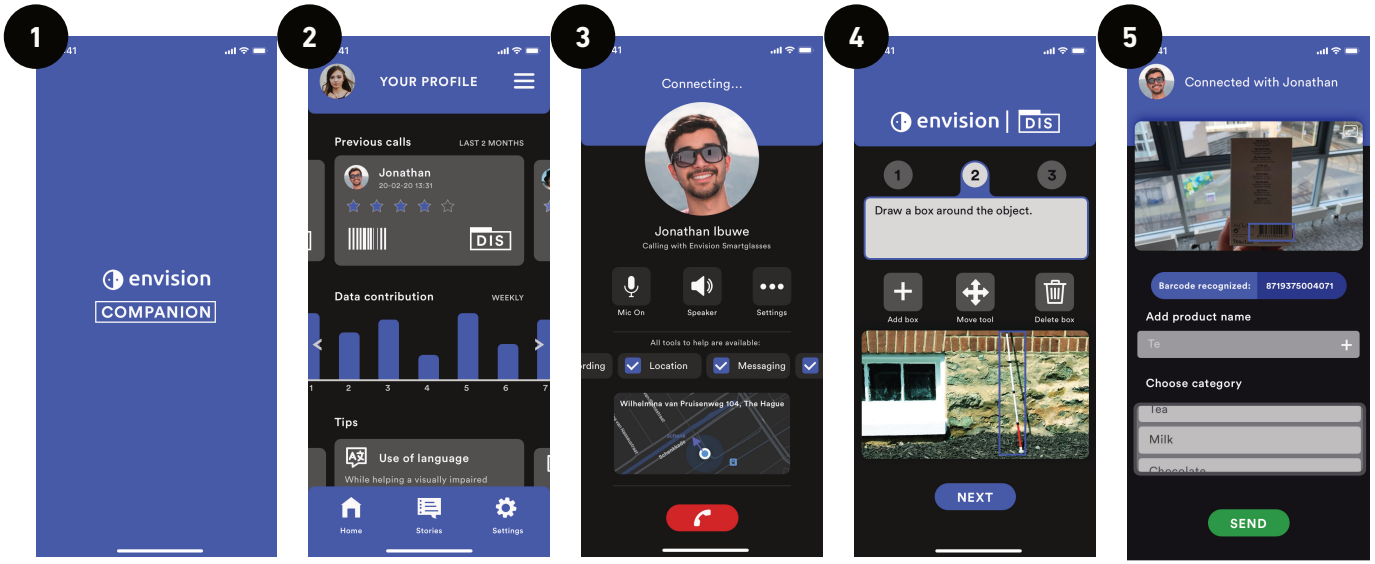


Figure 74: Screens from the companion app

for the VIP and there is no need to remember long pieces of text.

- **Location and AR:** another very helpful tool is to know the location of a person. This already showed during user tests with the MVP where I assumed that a VIP was at a certain building. If the location was visible, the VIP would have been lead to the right direction instead. Additionally augmented reality can play a role in giving extra information to the helper to understand how the VIP should walk towards his final location.
- **Flashlight:** simple but very effective is the possibility of turning on or off the flashlight. Although the current Google Glass does not have a flashlight, this could be a tool if the app is ever going to be released for other smart glasses.

Some of these tools are already used in existing similar apps such as the flashlight (BeMyEyes) and the ability to screenshot and message (Aira).

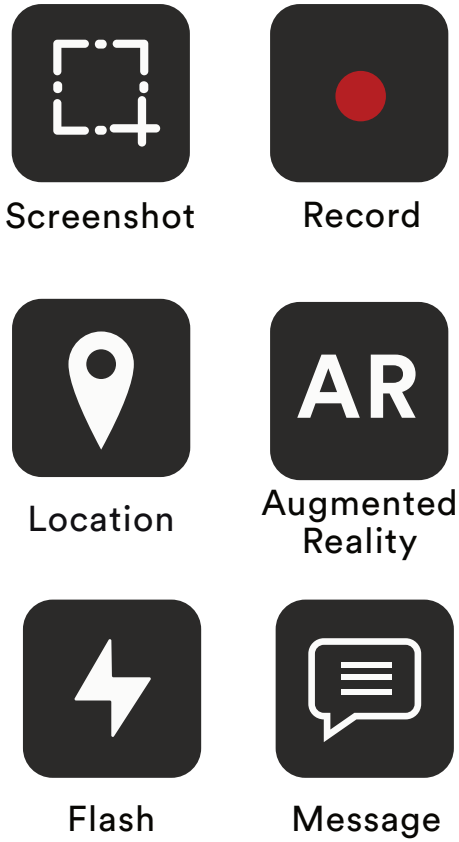


Figure 75: Tools for the companion app.

10. scenarios

The previous chapters show the smart glasses app and the companion app. This chapter focuses on scenarios that show the use of the smart glasses in combination with the companion app.

The scenarios are created by self-exploration with the BeMyEyes and Aira app, and conducting research with the target group.

The top part shows the smart glasses and the use of the features to call a person, whereas the bottom part shows the companion app and whenever the user is notified of a call.

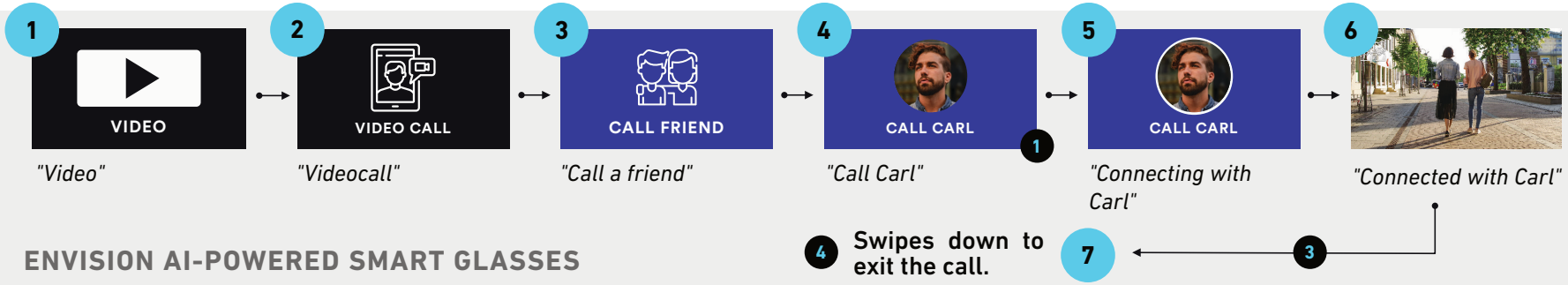
The scenarios focus on different reasons of using human assistance whenever AI fails.

Scenario 1: Getting from location to another while avoiding obstacles

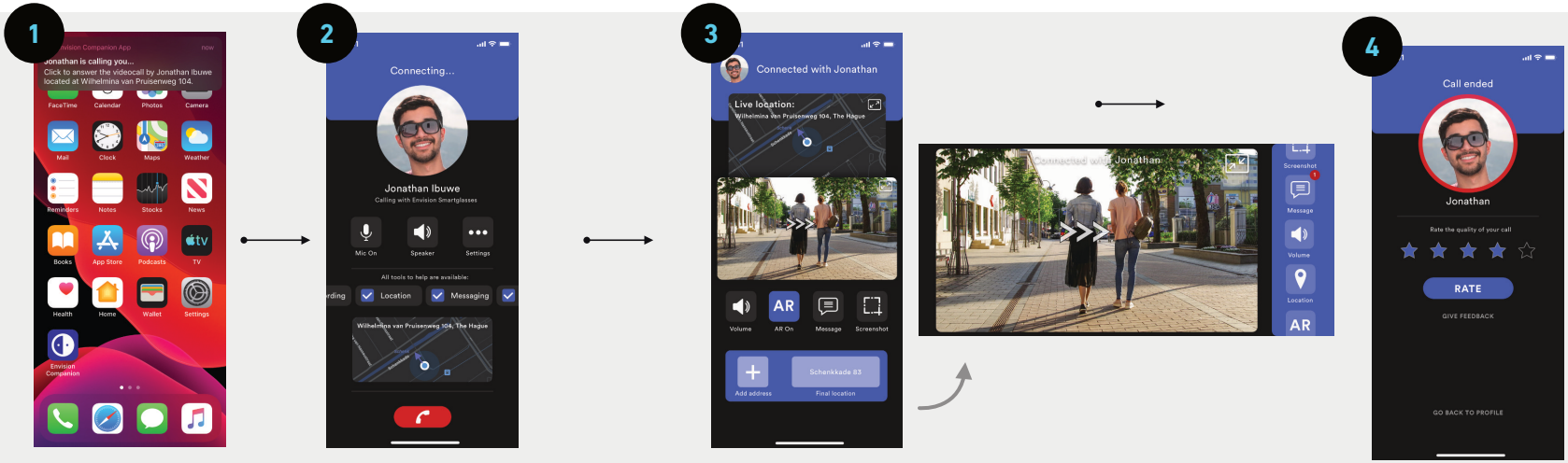
Jonathan has a job interview and is trying to get to the location. On his way he finds out there is a construction. He has his hands full with a bag and a white cane, therefore he calls Carl with the Envision smart glasses and asks if he could lead him around the construction.

Not possible with AI

Calling a friend



ENVISION AI-POWERED SMART GLASSES



Carl receives a notification from his friend Jonathan, he clicks on it and the Envision Companion app opens up.

The app is connecting with Jonathan and meanwhile Carl knows where Jonathan is and that he can use all the available tools such as location and messaging.

He is connected with Jonathan that tells him where he needs to be. Carl is not entirely familiar here so uses the AR to guide Jonathan to his job interview. He turns his phone to get a better full-screen overview of the situation.

Once arrived at the destination, Jonathan says bye and leaves the call. Carl leaves feedback regarding the quality of the call and closes the app.

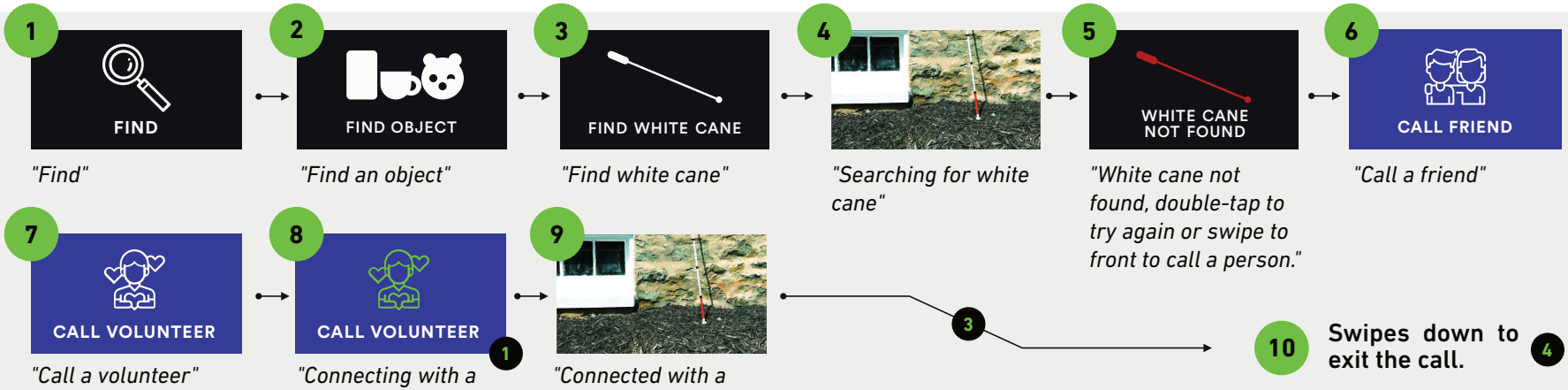
ENVISION COMPANION APP

Scenario 2: Finding an object after losing it.

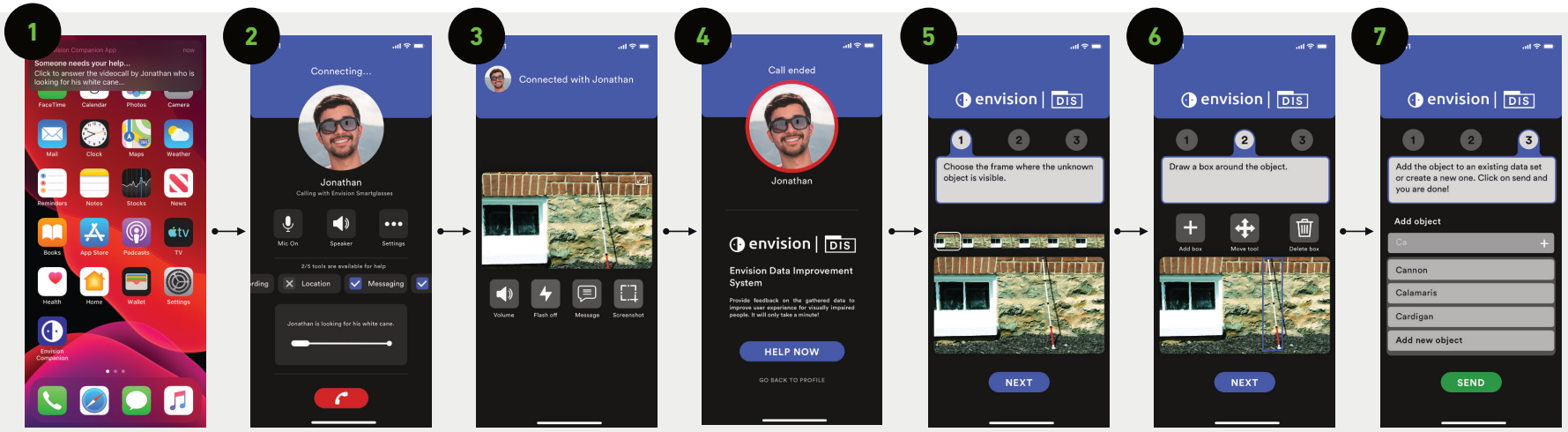
Jonathan sits outside in his garden, he wants to go back inside but he lost his white cane. He uses the Envision smart glasses to find where his white cane is. He is sure that he left it somewhere but the AI seems to fail. Therefore he calls a volunteer for help.

AI fails

Calling a volunteer



ENVISION AI-POWERED SMART GLASSES



Jennifer gets a call from a VIP looking for his white cane.
She clicks on the notification.

It states that Jonathan is not sharing his location, but Jennifer can use the flashlight if needed.

She connects with Jonathan, notices the white cane on a wall and leads Jonathan to it. They both hang up.

It shows that the call is ended, and Jennifer is keen on helping the AI with Envision DIS.

From the video that they took, she chooses a frame where the white cane is visible.

She draws a box around the white cane and clicks on next.

She adds 'white cane' as a new object and sends it to the system.

ENVISION COMPANION APP

Scenario 3: Reading a menu card at a restaurant

After a job interview, Jonathan wants to get some food at a restaurant nearby. He enters the restaurant and gets a menu card. He uses the Envision glasses to make sense of what they offer. Unfortunately he does not get any wiser nor does he know what each dish costs. The AI is not accurate enough to understand what price belongs to what dish.

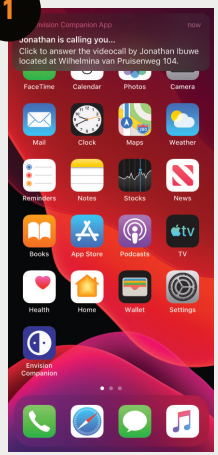
AI not accurate enough

Calling a friend

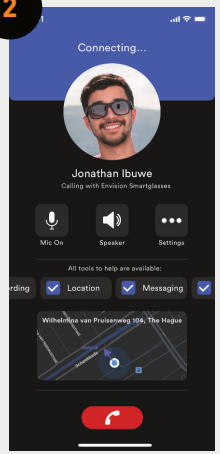
1. TEXT — *"Text"*
2. INSTANT TEXT — *"Instant Text"*
3.  — *"Sourd... 5.50.... Porridge... Scram..."*
4. TEXT — *"Text"*
5. VIDEO — *"Video"*
6. VIDEO CALL — *"Video call"*
7. CALL FRIEND — *"Call a friend"*
8. CALL CARL (1) — *"Call Carl"*
9. CALL CARL (2) — *"Connecting with Carl"*
10.  — *"Connected with Carl"*

→ 3 → Swipes down to exit the call. 11

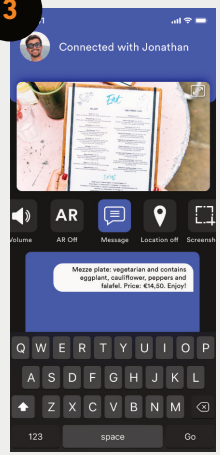
ENVISION AI-POWERED SMART GLASSES

1. 

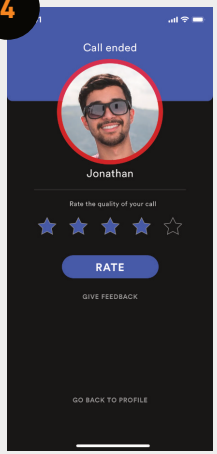
Carl receives a notification from his friend Jonathan, he clicks on it and the Envision Companion app opens up.

2. 

The app is connecting with Jonathan and meanwhile Carl knows where Jonathan is and that he can use all the available tools such as location and messaging.

3. 

He connects with Jonathan who is at a restaurant. He tried to use Envision to read the menu, but could not make anything of it. He asks whether Jonathan can help him and if he can write the vegetarian options in the message so he can read them back.

4. 

After writing the vegetarian options down, they hang up.

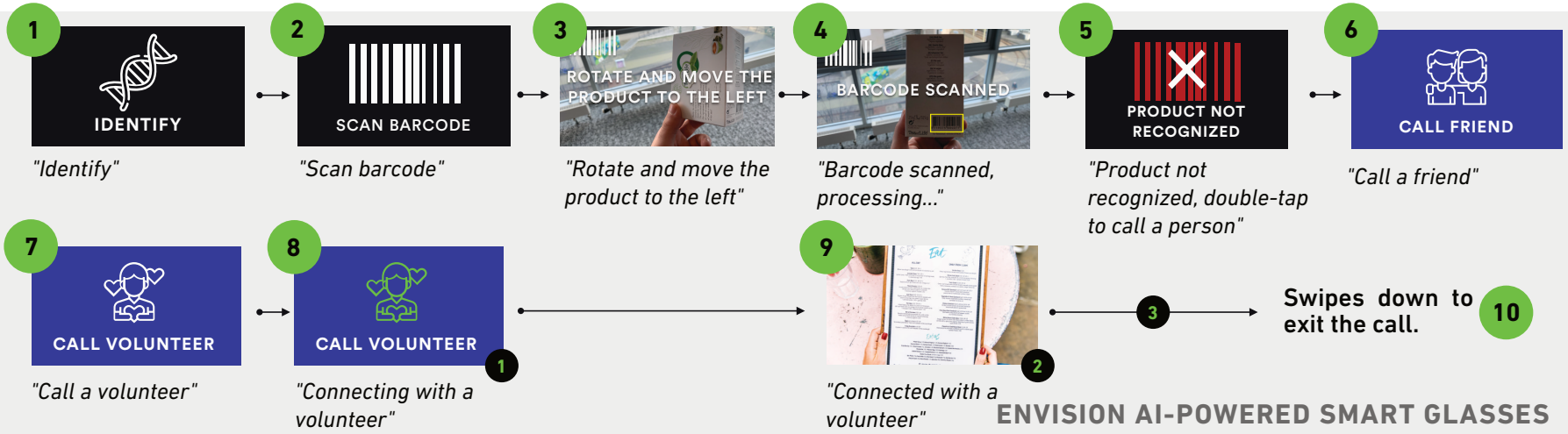
ENVISION COMPANION APP

Scenario 4: Scanning a product at a supermarket

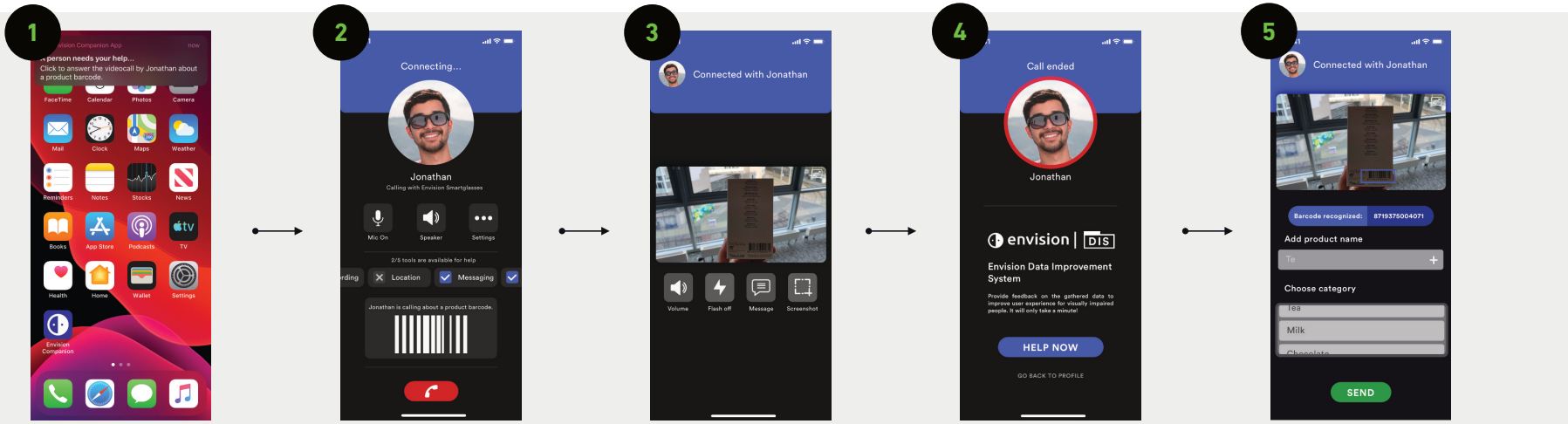
Jonathan grabs a pack of tea. Though he does not know which tea it is. The AI guides him towards getting the barcode right in front of the camera and the barcode gets scanned. Unfortunately this product is not recognized, thus he decides to call a volunteer to get some extra information about the tea.

AI fails

Calling a friend



ENVISION AI-POWERED SMART GLASSES



Laura receives a notification from the Envision Companion app. She clicks on it.

The app is connecting with Jonathan, Laura notices that he needs help with a product barcode.

Jonathan tells her that he found the barcode, but that is still not recognised. Laura explains what the product is.

After explaining the product, Jonathan thanks her and hangs up.

Besides helping Jonathan, Laura also helps the DIS by linking the product name and details with the barcode.

ENVISION COMPANION APP

11. onboarding

As mentioned in the first chapter, the glasses can not be a stand-alone device and it needs to be connected with a smartphone. In this case the Envision app is used to onboard the smart glasses concept.

The advantages of this is the ability to copy the same settings that the user might have with the Envision app and use it with the glasses. Settings such as the volume, faces of people they have trained in the app and screenreader speed.

The user opens the new Envision app and notices a new tab with the Envision glasses (figure 70,1). This is is done to show clearly that the user has to click on that tab but also to show other users that an additional device is available besides the app.

After opening that tab, the user is either able to buy the device or in this case pair his current Envision glasses. In the first step the user gets his pair of Envision glasses, he turns it on and scans the QR code on the app. This special code makes sure that the app is connected with the glasses. Additionally the user is moved towards customization of his Envision glasses that includes all advanced settings.

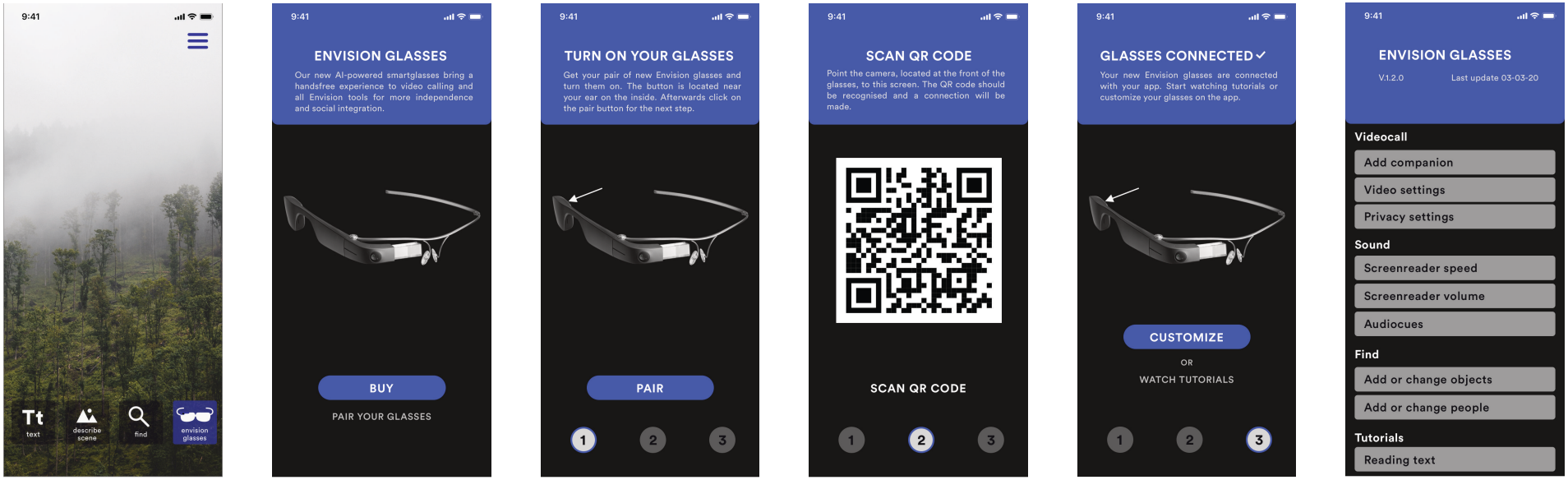
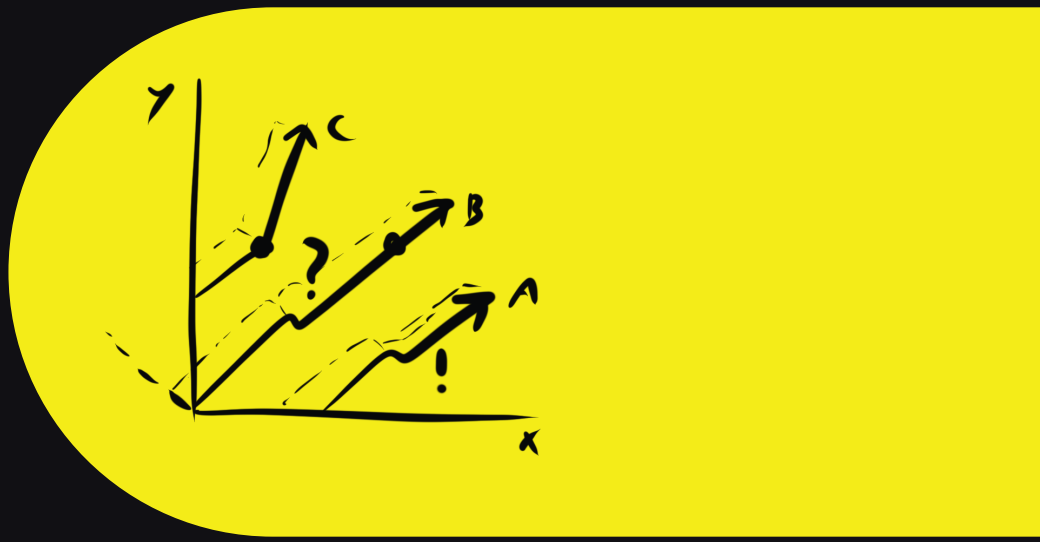


Figure 76: Screenshots of the Envision Companion app

i) conclusion and recommendations

The previous chapter ends with a concept of AI-powered smart glasses for visually impaired people. This chapter concludes the report and focuses on future research regarding AI, computer vision, VIPs and a combination within.



in this chapter:

1. target group
2. smart glasses
3. data and collaboration

1. target group

Artificial Intelligence can extend human frontiers and complement our capabilities. This project presents AI-powered smart glasses that empower visually impaired people to achieve greater independence and social integration. Since AI is a developing discipline, we enhance reliability by providing the opportunity for reciprocal learning between AI and its users. This graduation project is a collaboration between Envision (a leading computer vision app for people with a visual impairment) and the TU Delft department of Human-Centered Design.

What all started with exploring different use cases of AI for VIPs, ended in showing the importance of computer vision for improving the quality of life of VIPs.

The concept shows an interim solution until AI improves and is able to solve many more daily obstacles without the use of humans.

This report provides a clear overview of the target group: visually impaired people. The personas and data show a framework that can be used for future research.

Though, a small part of the target group was left out for this research: elderly people that are not tech-savvy and mainly rely on a friend or family member for most daily tasks.

I recommend to conduct another research with this target group to see how present-day technology, like the smart glasses, can also improve their quality of life.

In addition to a different target group, I also recommend to conduct user research on global level.

During a conference (February 2020) visited by Envision in Dubai this already showed differences compared to VIPs in the Netherlands:

- Wearing a burka meant that a lot of the VIPs were not able to interact with the touchpad on the Google Glass. A voice assistant would be preferred much more.
- Mandatory use of local language. Whereas a lot of the VIPs in the Netherlands were able to communicate in English, the people from Dubai relied totally on their local language.

Human-centered design shows in this case that nobody should be left out. Diversity and inclusivity should not be a side issue. It should be at the core of the company.

2. smart glasses

This report starts with exploring apps that use computer vision for visually impaired people. It can be concluded that this is very helpful, but to improve user experience smart glasses are introduced.

Smart glasses are superior to apps on smartphones. From the view of a VIP:

- a handsfree experience
- more convenience: not having to get your phone everytime to read a text
- a device that is always with you
- easier orientation.

From the view of a helper in combination with the companion app:

- point-of-view information (easier to help a VIP)
- improved help: provided with the companion app, such as location, the ability to screenshot or message and more.
- more information: provided with the wide-angle lens compared to many traditional smart phone cameras.

Though, this report shows that smart glasses are superior. Future research should also look at other options such as being able to offer the software to other cameras like a GoPro.

Additionally, the iPhone has a screen-reader (VoiceOver). The smart glasses tested with the concept (and MVP) do not have a screenreader yet and is being added manually. I recommend to conduct another research that solely focuses on using a screenreader with glasses for an optimal experience.

Besides testing the same software with other cameras, solutions could also be found in combining devices. Simple tests could conclude whether being able to use both the smart glasses camera in combination with the smart phone camera can help the VIP even better.

Research was conducted on current smart glasses in the market. It can be concluded that compared to smartphones, a lot of smart glasses are still very behind. Camera, CPU and other sensors are worse or non-existent compared to smartphones.

For Envision as a company it could be interesting to look at creating their own smart glasses optimized for visually impaired people. This means a light-weight solution that focuses more on providing a great user experience using audio. Additionally, a choice could be made to discard the screen because of its redundancy and high costs.

Future smart glasses could also contain an haptic engine to provide feedback with haptics instead of only with sound.

Additionally, sound is only coming from one side of the smart glasses whereas more research should be conducted towards having sound from two sides or hearing sound with bone-conduction like the Aftershokz, discussed earlier in this report.

The handsfree experience could be extended and improved with a voice assistant. Due to time limitations, there was no time to test whether voice assistants for smart glasses would be appreciated and preferred over using the touchpad on the side.

Although, previous research with the VIPs already showed they use Siri a lot to

control the smartphone or devices at home. When talking with VIPs during the last prototyping session, they also mentioned the possibility of having a voice assistant would be perfect.

3. data and collaboration

As mentioned in the previous chapter, data and privacy also plays a big role. Especially since the GDPR law in the EU was introduced in 2018. This law is a legal framework for collecting and processing data from individuals.

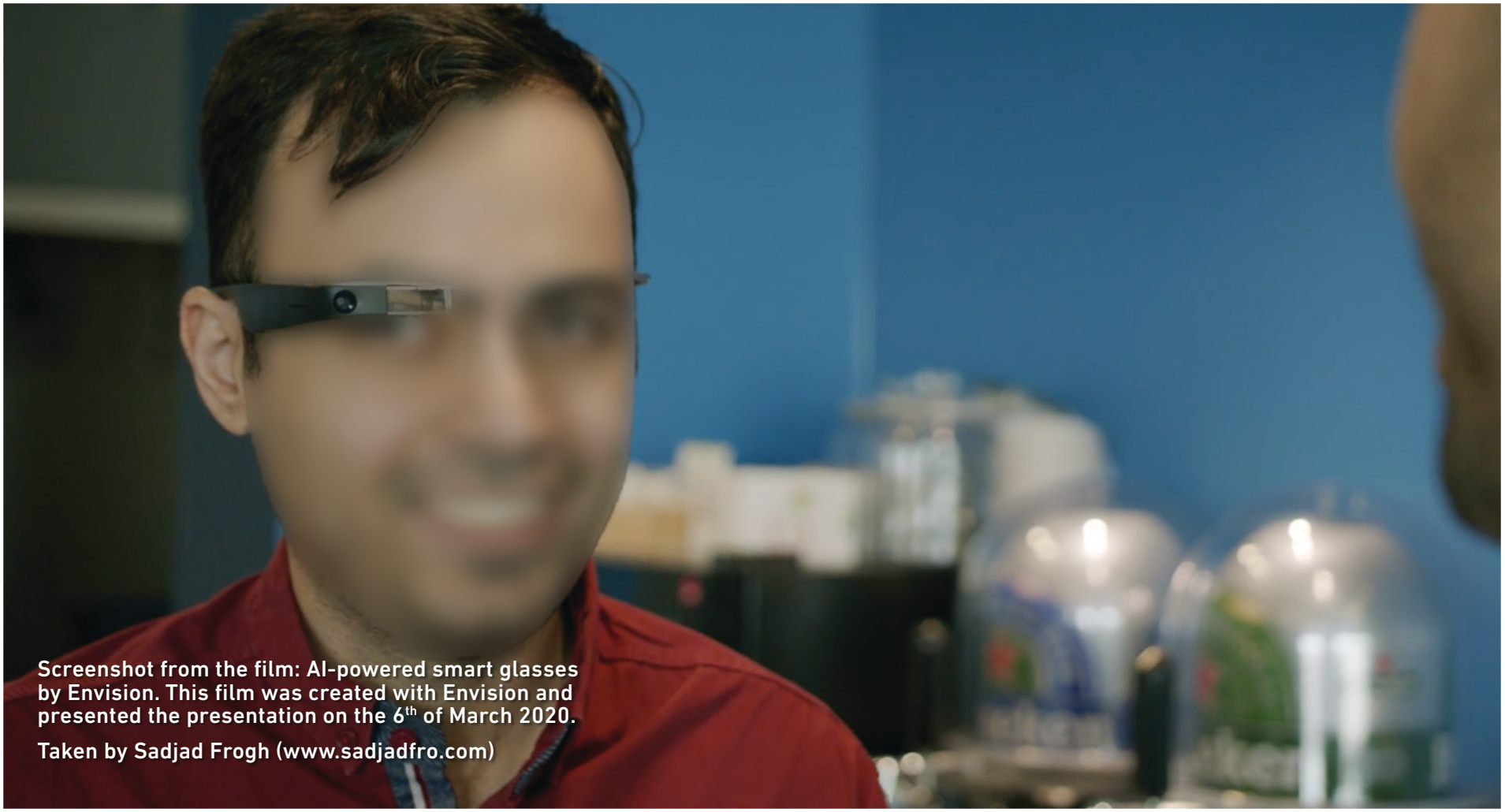
It is recommended to be loud and clear about the usage of data and what is being done with it. This means that the advantages should be explained such as improving the AI system. Especially if the concept introduces the so-called Envision Data Improvement System.

The previous chapter already showed the Envision companion app and the data that is allowed to be seen by the companion.

The VIP should always be in control on the usage of the data and who has access to it. Envision should be very transparent about what data is used and how it is used. Even if VIPs are not asking for it.

Lastly, the concepts focus on the different people that can join the companion app. I recommend for Envision to either build their own platform where volunteers can help or collaborate with like-minded companies and who offer similar solutions as discussed in the report: BeMyEyes and Aira.

Additionally, paid helpers could also be from other companies that can create more revenue for Envision to grow as a company.



Screenshot from the film: AI-powered smart glasses by Envision. This film was created with Envision and presented the presentation on the 6th of March 2020.

Taken by Sadjad Frogh (www.sadjadfro.com)

personal reflection

This report represents six months of hard work and I am happy to say that I am very proud of it.

This project made me rethink design in a life-changing way.

Design does not always have to be visual and user experience goes further than what the eye sees.

This project inspired me to be a better designer and question everything that I learned and know.

I have learned what motivates me and what my strengths are. I have also seen with my own eyes how this project can improve people's lives.

And that is what I will keep doing.

references

Boeijen, A., Daalhuizen, J., Schoor, R., & Zijlstra, Y. (2017). Delft Design Guide: Design Methods. Amsterdam: BIS Publishers.

Dahl, A. A. (n.d.). Blindness Symptoms, Treatment & Types. Retrieved from https://www.medicinenet.com/blindness/article.htm

Dana H. Ballard; Christopher M. Brown (1982). Computer Vision. Prentice Hall. ISBN 978-0-13-165316-0.

Davidoff, S. K., Lee, M. K., Dey, A. undefined, & Zimmerman, J. undefined. (n.d.). Rapidly Exploring Application Design Through Speed Dating. UbiComp 2007: Ubiquitous Computing Lecture Notes in Computer Science, 429–446. doi: 10.1007/978-3-540-74853-3_25

EY, "Study: Work-Life Challenges Across Generations," EY, 2015, https://www.ey.com/us/en/about-us/our-people-and-culture/ey-work- life-challenges-across-generations-global-study

Feiten en cijfers. (2019, March 22). Retrieved January 17, 2020, from https://oogfonds.nl/visuele-beperking/feiten-en-cijfers/

Hersh, M. A., & Johnson, M. A. (2008). Assistive technology for visually impaired and blind people. London: Springer.

Higgins. (2019, October 8). Supreme Court hands victory to blind man who sued Domino's over site accessibility. Retrieved February 26, 2020, from https://www.cnbc.com/2019/10/07/dominos-supreme-court.html

McKinsey&Company (2014, January). Why diversity matters. Retrieved February 24, 2020, from https://www.mckinsey-.com/business-functions/organization/our-insights/why-diversity-matters

Nationale Rapportage Oogzorg 2018. (2018). Retrieved January 14, 2020, from https://oogfonds.nl/fileadmin/oogfonds-/folderOoggezondheid.pdf

Nichols, Allan (1995), Why Use the Long White Cane?, archived from the original on 2010-03-30

Nilsson, N. (2010). The Quest for Artificial Intelligence, Cambridge University Press.

Recruiting a Diverse Workforce: Glassdoor for Employers. (2014, November 17). Retrieved February 24, 2020, from https://www-.glassdoor.com/employers/blog/diversity/

Runjić, Tina & Bilić Prcić, Ante & Alimovic, Sonja. (2015). The relationship between social skills and behavioral problems in children with visual impairment. Hrvatska Revija Za Rehabilitacijska Istrazivanja. 51. 64-76.

Schakel, W., Bode, C., Hilde P A Van Der Aa, Hulshof, C. T. J., Bosmans, J. E., Gerardus H M B Van Rens, & Nispen, R. M. A. V. (2017). Exploring the patient perspective of fatigue in adults with visual impairment: a qualitative study. BMJ Open, 7(8). doi: 10.1136/bmjopen-2016-015023

Stöppler, M. C. (2019, September 13). Vision Loss: Symptoms, Signs, Causes & Treatment. Retrieved from https://www.medicinenet.com/vision_loss/symptoms.htm

Turing, A. (1950). Computing machinery and intelligence. Oxford: Blackwell.

Vision 2020: The Right to Sight. (2019). Retrieved from http://www.vision2020.nl/contents/V2020NLrapport.pdf

j) appendices

Certain visualisations and raw data that are left out of the report.

appendix 1: extended version of AI explanation

The term artificial intelligence was first coined by John McCarthy in 1956 at the Dartmouth Conference. But the journey to understand if machines can truly think began even earlier. In 1950, Alan Turing, published a paper entitled Computing Machinery and Intelligence which opened the doors to the field that would be called AI. The paper itself began by posing the simple question, “Can machines think?”.

Turing then went on to propose a method for evaluating whether machines can think, which came to be known as the famous Turing test. The test, or Imitation Game as it was called in the paper, was put forth as a simple test that could be used to prove that machines could think. The Turing test takes a simple pragmatic approach, assuming that a computer that is indistinguishable from an intelligent human actually has shown that machines can think.

In **1955**, McCarthy organised the Dartmouth Conference: a summer workshop to clarify and develop ideas about thinking machines. He picked the name Artificial Intelligence for the new field. He chose the name partly for its neutrality; avoiding a focus on machines which imitated certain features of man (automata theory), and avoiding the science of communications and automatic control systems in both machines and living things (cybernetics) which was heavily focused on analog feedback.

Norbert Wiener is credited as being one of the first to theorise that all intelligent behaviour was the result of feedback mechanisms, that could possibly be simulated by machines and was an important early step towards the development of modern AI that we know now.v

In **1966**, the first natural language processing computer program was created at the MIT Artificial Intelligence Laboratory called ELIZA, by Joseph Weizenbaum, a German computer scientist.

In **1997**, IBM created Deep Blue: a chess-playing computer developed by IBM. It is known for being the first computer chess-playing system to win both a chess game and a chess match against a reigning world champion under regular time controls.

In **2005**, Stanford University’s Racing Team in cooperation with the Volkswagen Electronics Research Laboratory (ERL) created Stanley: an autonomous car. It won the 2005 DARPA Grand Challenge.

2010 was the year that Apple introduced Siri; a voice assistant that is one of the best examples of speech recognition.

Strong AI is capable of performing tasks that a human brain can or even surpass. Someone who is able to read Russian characters, also understands Russian speech and might also know something about the Russian culture and could even make recommendations which Russian cities to visit. In contrast, different AI systems would be needed for each of these tasks.

As stated before, AI is an umbrella descriptor and nowadays more research is based on so-called machine learning (ML) and deep learning (DL).

ML is a branch of AI, which is a method that feeds a lot of data into a machine to make it learn. It is used to handle data. By using ML, decision making can be improved, patterns and trends can be

uncovered in data and complex problems can be solved.

The field of Machine Learning seeks to answer the question:

"How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?".

Machine Learning uses algorithms, which is a set of rules and statistical techniques used to learn patterns from data. Some examples are linear regression, decision tree or a random forest. These algorithms are used to train a ML model.

Combining different ML-models is called Deep Learning. Likewise, it is a subset of AI which uses statistical methods to enable machines to improve with experience and makes the computation of multi-layer neural networks feasible.

Deep Learning is used because of the limitations of machine learning. Images for instance have a lot of pixels and therefore a lot of data. The computer is not sure what to look at.

Whereas DL models are capable of focusing on the right features by themselves, requiring little guidance from the programmer, it also solves the dimensionality problem that comes with Machine Learning. Despite being able to recognise flat 2D objects, it still struggles to recognise and analyse 3D objects.

It is inspired by our brain based on the concept of artificial neural networks.

Figure 6 shows a timeline of when the different methods of AI were introduced.

AI is present almost everywhere, such as healthcare, finance, banking, retail, manufacturing, energy, logistics, transport, insurance and tourism. Big tech companies have all adapted Artificial Intelligence into their business.

As mentioned before, Apple uses artificial intelligence in Siri for speech recognition. Some other companies that use AI are Google (spam detection), Tesla (smart vehicles powered by AI), Amazon (site predictions and suggestions), Netflix (predictive technology) and Facebook for adding attractive and relevant content to the News feed based on a user's preferences.

Another great example is the one of IBM Watson that studied 20 million oncology records to diagnose leukemia with patients.

This shows the importance of AI but it does not always benefit customers. For instance, MMC's report found that when companies do deploy artificial intelligence and machine learning, the use-cases are often quite banal. Some of the most popular ways the startups surveyed used AI included chatbots (26 percent of companies) and fraud detection (21 percent). In both cases, it is tricky to judge exactly how much this technology benefits customers.

Thus, when using AI we have to think about whether it is beneficial for the user. The goal is to improve independence of visually impaired people, whether that should be with AI or without, that is not the most crucial part of it.

appendix 2: different visual impairments and common causes

Diabetic Retinopathy

Cataracts

Glaucoma

Stargardt disease/macular degeneration

Retinal Detachment

Although it is not of huge importance for this field of study, it is still good to know about the common causes of blindness, which include diabetes, macular degeneration, traumatic injuries, infections of the cornea or retina, glaucoma (a disease in which the retinal neurons that send the signal from the eye to the brain die), and even the inability to obtain any glasses.

While there are many causes of macular degeneration, including genetic abnormalities such as Stargardt disease, age-related macular degeneration (AMD or ARMD) is by far the most common type. AMD is a disease associated with aging that gradually destroys the sharp central vision needed for seeing objects clearly and for common daily tasks such as reading and driving (Stöppler, 2019).

This is also why most visually impaired people are older than 50 years.

Cataract happens the most in the Netherlands, with an annual diagnosis of 70.000 people.

Besides the aforementioned types of blindness, there are other types of blindness such as color blindness (inability to perceive differences in various shades of colors,), night blindness (difficulty in seeing under situations of decreased illumination) and snow blindness (loss of vision after exposure of the eyes to large amounts of ultraviolet light).

If eyecare remains at the current level, due to ageing of the population, the amount of visually impaired and blind people will increase, in the Netherlands, to 300.000-440.000 people.

When interviewed, 82 percent of the Dutch population chose sight as their most important sense.

Looking at the numbers on an interna-

tional level, between 300-400 million people are visually impaired due to various causes. Of this group, approximately 50 million people are totally blind and 80% of blindness occurs in people over 50 years of age.

Additionally, the ratio between male-female is 1:2 of being visually impaired and between 2010 and 2020 visual impairment increased by 20% (Oogfonds, 2019).

A report by “Vision 2020” (2019) shows that percentually the largest increase of visual impairment and blindess is in the age group between 55 and 74 years.

Later in this chapter, personas are created which also reflect a part of this target group.

A recent study from the Organization for Economic Co-operation and Development (OCED) (2010) showed a huge labor market disadvantage. On average, the employment rate was 44% and 75% for people with and without disabilities, respectively. The inactivity rate was 49% and 20% for people with and without disabilities, respectively. Therefore, the inactivity rate for disabled people is about 2.5 times higher.

appendix 3: examples of technology used by vips

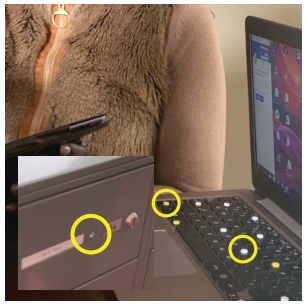
1. Penfriend 3 - Voice Labeling System (source: RNIB)

2. Victor Reader Trek - Talking Book Player and GPS System (source: RNIB)

3. Liquid Level Indicator

4. Milestone 112 Ace - Voice Recorder (source: vision High-Tech YouTube)

5. Talking Watch



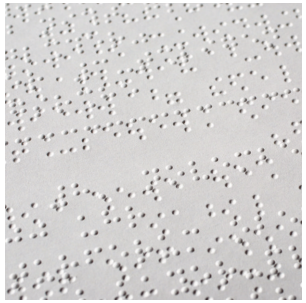
7. Tactile rubber marks for identification

8. Lechal - Insoles for navigation (source: Lechal)

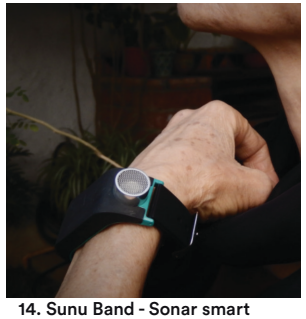
9. naviBelt - Tactile orientation aid (source: feelspace)

10. AfterShokz - Bone conduction headphones

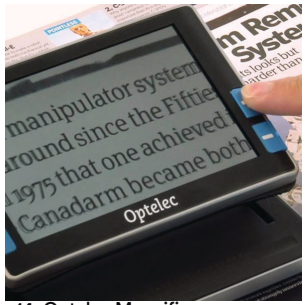
11. WeWalk - Smart white cane (source: wewalk)



12. Braille - Tactile reading and writing system (raised dots)



14. Sunu Band - Sonar smart band for detecting objects (source: Tech Xplore)



14. Optelec Magnifier - Electronic magnifier for reading (source: Optelec)

15. OrCam - Portable device to understand text and objects (source: OrCam)

15. HumanWare Braillaint - Braille Display (source: HumanWare)

appendix 4: apps used by VIPs

screenreaders

A screenreader is a software application that enables VIPs to use a computer. It "reads"/speaks out the content of a page.

NVDA (Windows)

Jaws (Windows)

Dolphin (Windows)

VoiceOver (macOS)

transportation

Apps used by VIPs to use transportation: public transport, car, taxi.

OV Info

9292

NS

Google Maps

Trevvel

reading

Apps used for reading books. All of them being accessible for visually impaired people.

Daisylezer

Voice Dream Reader

Amazon Kindle

computer vision

Apps used by VIPs that use computer vision which allows to recognise objects and people, speak out text and more.

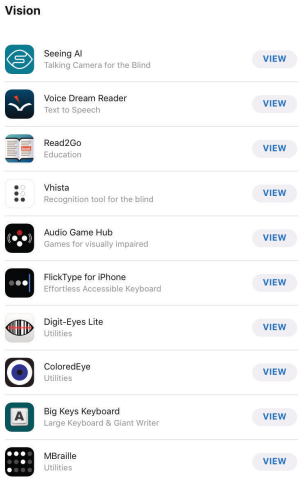
Seeing AI

TapTapSee

Envision

apple accessbility apps

A list of apps for accessibility shown in the App Store, below the apps for vision can be seen.



navigation

Apps used to navigate outside as a visually impaired person.

Blindsquare

Soundscape

Google Maps

others

Other frequently used apps for messaging and entertainment.

Earcatch audio-description

WhatsApp

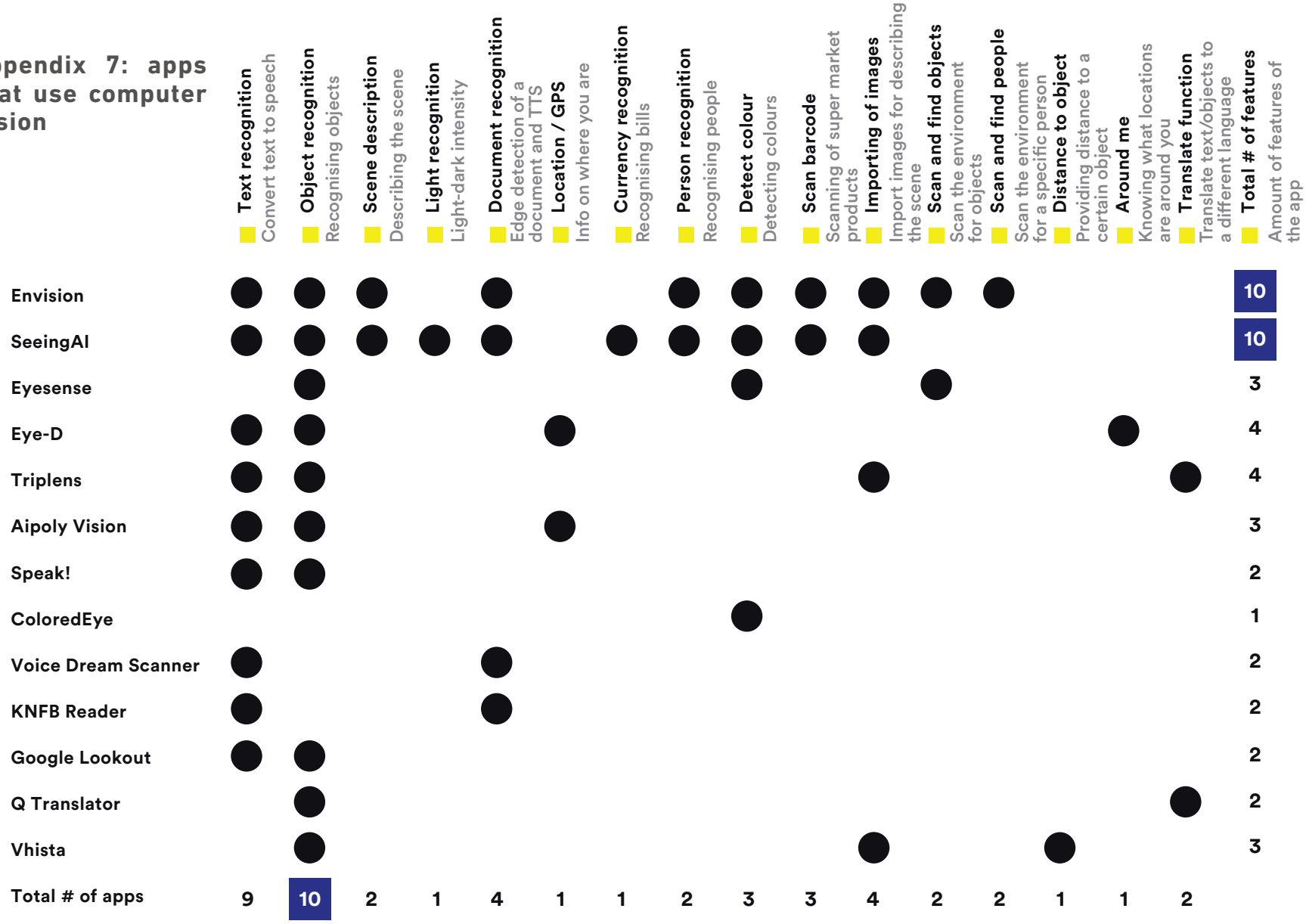
appendix 5: overview of all ideas

1. Imagine you have a pair of smart glasses with whom you can call random people that are able to guide you to a location.
2. Imagine you have a pair of smart glasses with whom you can have someone 'see' with you and tell you what you are looking at
3. Imagine there was an app on your phone that describes scenes, but whenever the scene is too difficult to understand it will call a volunteer who can help you
4. Imagine there was an app that describes a scene by taking a picture with your smart phone
5. Imagine there was an app on your phone that could be detected by government buildings so as soon as you enter a building it would use VoiceOver at the entrance on how to commute inside the building
6. Imagine there was a tool that spreads out a smell. So you could follow a smell to get from A to B to C. "Follow the smell of baked bread, after that follow the smell of gasoline, the destination is located at a smell of fries.
7. Imagine there is a device that makes smell clearer, so smell in general can be used to your advantage.
8. Imagine there was a vibration device that you attach to your buckle, which will vibrate whenever you encounter the possibility of bumping into an obstacle that is in front of you.
9. Imagine there is a suit that you can wear that is made from smart materials and as soon as something approaches you it starts vibrating at that position. Thus, if someone is approaching you from the right, the sensors attached to the right arm will start to vibrate. The closer someone is, the quicker it will vibrate.
10. Imagine there is a device that you can put on the outside of your arms and legs and it will vibrate whenever you are close to something.
11. Imagine there is a magical button you can click on whenever you are outside and a volunteer in the Netherlands could come to guide you.
12. Imagine a pair of shoes that help you to walk straight when you are walking on the side walk. (maybe with balanced/unbalanced vibrations and compass)
13. Imagine you enter a building, a camera detects your white cane and it tells you what building you are in and where you have to go next.
14. Imagine you have a laser with whom you can detect distances and find out what the texture of something is, e.g. bricks, mud, flat surface.
15. Imagine there is a service that certifies stores regarding its accessibility. E.g. '"this store is wheel-chair friendly" "this store is visually impaired friendly". The service focuses on getting a particular store or restaurant certified, like you have now with tripadvisor. This service could also offer a list of accessible restaurants and cafés.
16. Imagine there is a service that standardises government buildings, making them more accessible. Accessible in the way of having the same tactile interior floor, same braille signs for rooms and toilets and the same appearance.
17. Imagine there is an online tool that detects the website and if it is not accessible, it redirects you to a more accessible similar website.
18. Imagine there is a tool that measures your energy level and based on that slows down or speeds up the speed of the VoiceOver when browsing the internet

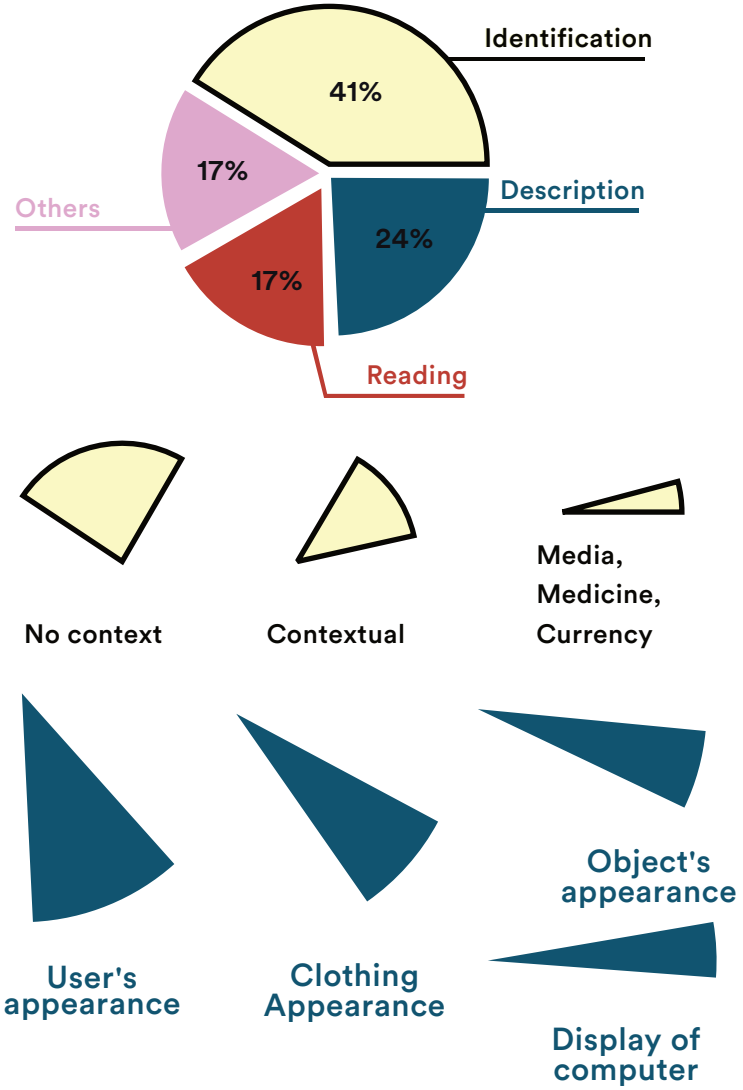
appendix 6: ratings for each category

	Relevance of the solved idea How relevant is the solved idea? Small or a day-to-day struggle of a VIP?	Implementation ease What is the level of difficulty of implementing this idea? (From technological point and the approval of external factors)	Connection to AI How much is the idea connected to/does it use AI?	Normality/ appearance How normal is the solution? Does it make the VIP stand out from the crowd?	Size of impact How large is the target group that can be impacted with this solution?	Envision How much does it fit with Envision, from a business and knowledge side?
Importance scale:	2	2	2	2	1	1
Smart glasses	5	4	6	4	5	7
Smart phone app	3	4	7	5	5	6
Smart material detecting obstacles	5	1	5	6	4	3
Magical button for physical help	2	3	6	5	6	4
Insole for walking straight	5	2	6	7	5	3
Laser for detecting distances and textures	4	4	6	2	3	5
Speech assistance in buildings	4	2	6	2	5	2
Standardised buildings	6	1	1	7	6	1
Internet accessibility tool	3	3	6	6	4	2

appendix 7: apps that use computer vision



appendix 8: a visual overview of the research conducted by bigham



looks like the back of a book