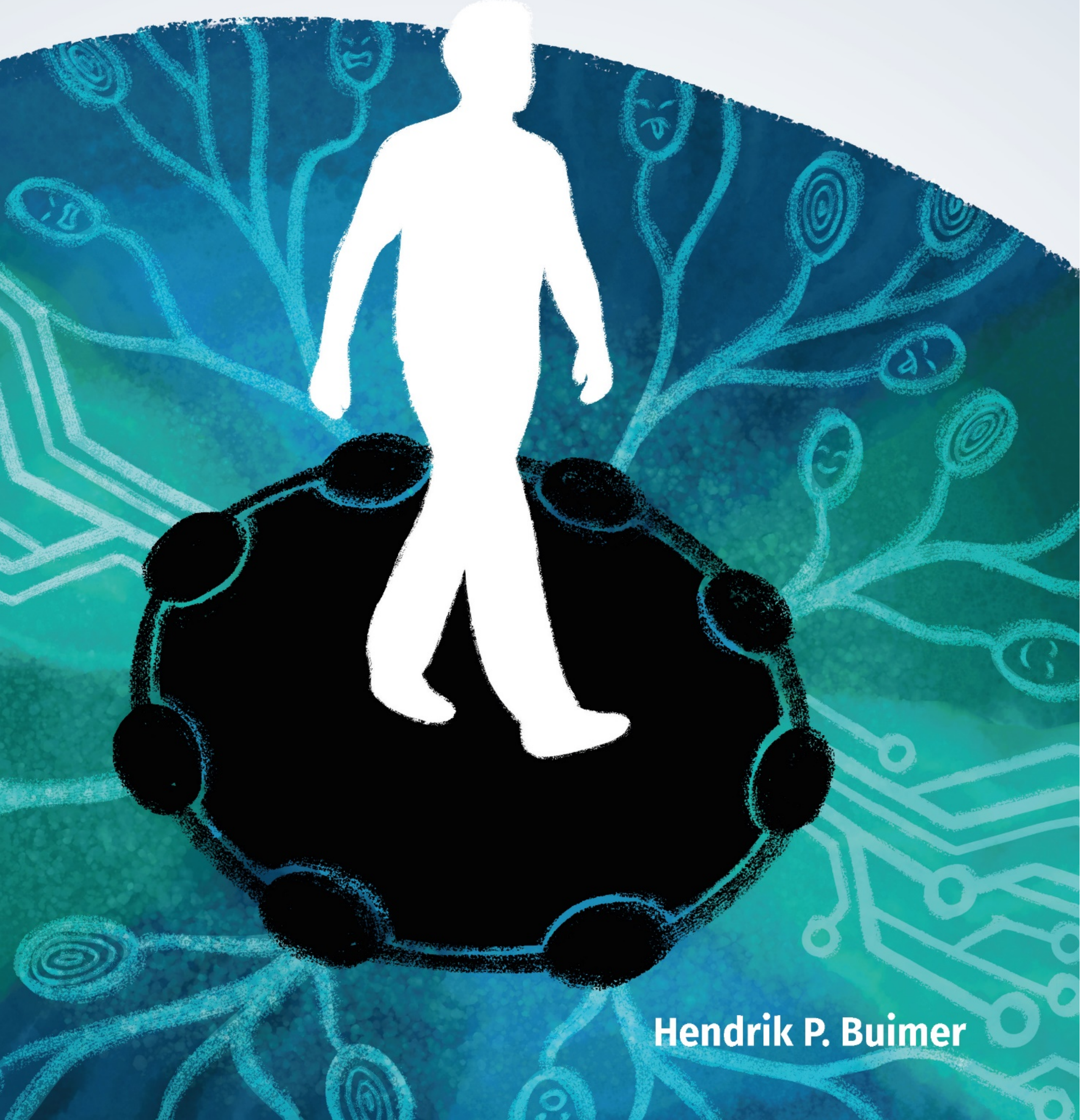
HUMAN-CENTERED DESIGN OF VIBROTACTILE WEARABLES FOR PERSONS WITH A VISUAL IMPAIRMENT



Hendrik P. Buimer

HUMAN-CENTERED DESIGN OF VIBROTACTILE WEARABLES FOR PERSONS WITH A VISUAL IMPAIRMENT

Hendrik Pieter Buimer

HUMAN-CENTERED DESIGN OF VIBROTACTILE WEARABLES FOR PERSONS WITH A VISUAL IMPAIRMENT

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. ir. A. Veldkamp,
on account of the decision of the Doctorate Board
to be publicly defended
on Friday 12 February 2021 at 12.45 hours

by

Hendrik Pieter Buimer

born on the 10th of May, 1989
in Hengelo (O), The Netherlands

This dissertation has been approved by

Supervisor

prof. dr. R.J.A. van Wezel

Co-supervisors

dr. ir. Y. Zhao

dr. T.M. van der Geest

Cover design: Artistic Leap Designs (artistic-leap.com)

Printed by: Ipskamp, Enschede

Lay-out: Hendrik Pieter Buimer

ISBN: 978-90-365-5121-2

DOI: 10.3990/1.9789036551212

Graduation Committee

Chair / secretary

prof. dr. J.N. Kok	University of Twente

Supervisor

prof. dr. R.J.A. van Wezel	University of Twente Radboud University

Co-supervisors

dr. ir. Y. Zhao	University of Twente
dr. T.M. van der Geest	University of Twente HAN University of Applied Sciences

Committee Members

dr. ir. M. Tabak	University of Twente
prof. dr. D.K.J. Heylen	University of Twente
prof. dr. P. Koenig	Osnabrück University
prof. dr. J. van der Steen	Erasmus MC
prof. dr. P.S. Sterkenburg	Vrije Universiteit Amsterdam

Samenvatting

De doelstelling van het onderzoek gepresenteerd in deze dissertatie is het verkennen van de mogelijkheden die draagbare consumentenelektronica (wearables en slimme brillen) bieden om slechtziende en blinde mensen te ondersteunen bij alledaagse activiteiten. In het bijzonder vanwege de multimodale interactiemogelijkheden van deze technologieën (prikkels via meerdere zintuigen tegelijkertijd – bijvoorbeeld de tril- en geluidssignalen bij het ontvangen van een bericht op je telefoon). In dit onderzoek hebben we gepoogd de mens continu centraal te zetten (*human-centered design*) bij onderzoek naar, en het ontwikkelen van, nieuwe draagbare technologieën om slechtziende en blinde mensen te ondersteunen in het dagelijks leven.

De studie omschreven in **hoofdstuk 2** had als doel om de prioriteit voor de ontwikkelings- en innovatieagenda van dergelijke technologieën in kaart te brengen, door in gesprek te gaan met de beoogde doelgroep van de te ontwikkelen oplossingen. Ondanks een breed scala aan beschikbare technologische oplossingen, identificeerden we verschillende urgente problemen en uitdagingen uit het dagelijks leven van slechtziende en blinde mensen. De moeilijkheden die men tegenkomt bij het bezoeken van onbekende omgevingen en locaties (inclusief het detecteren van obstakels), weerhoudt veel mensen met een visuele beperking van het zelfstandig bezoeken van nieuwe locaties. Daarnaast is het voor veel slechtziende en blinde mensen uitdagend of onmogelijk om non-verbale communicatie waar te nemen, wat impact kan hebben op sociale interacties en zelfs kan leiden tot (een gevoel van) sociale uitsluiting. Ook bleek uit de studie dat veel tekstuele informatie in de maatschappij slecht toegankelijk is voor slechtziende en blinde mensen. Deze eerste studie inspireerde ons om ons te richten op twee uitdagingen die veel slechtziende en blinde mensen dagelijks ervaren: het herkennen van non-verbale communicatie en het vinden van de weg in onbekende omgevingen.

In **hoofdstuk 3** worden de bevindingen omschreven van de ontwikkeling en evaluatie van een draagbaar systeem om de zintuigen te ondersteunen. Dit systeem bestaat uit een op het hoofd gedragen camera, een laptop met emotieherkenningssoftware en een riem die bestaat uit trilmotortjes om informatie via tactiele prikkels over te brengen naar de drager van het systeem. De reden om gebruik te maken van deze haptische riem in plaats van andere

manieren van zintuiglijke feedback geven, was om de ogen en oren van de gebruikers vrij te houden. Slechtziende en blinde personen zijn namelijk in hoge mate afhankelijk van hun gehoor – en soms overgebleven zicht – om hun directe omgeving waar te nemen. We hebben geëvalueerd of trilsignalen rondom het middel gebruikt zouden kunnen worden om zes basisemoties over te brengen en bruikbaar zijn en of een dergelijk hulpmiddel gewenst is voor dagelijks gebruik door slechtziende en blinde personen. Deelnemers aan het onderzoek waren na een korte trainingssessie goed in staat om de trilsignalen te interpreteren en de juiste emoties aan de trillingen te koppelen. De studie toonde dus aan dat deze trilsignalen, in een experimentele setting, gebruikt kunnen worden om de basisemoties direct over te brengen op de gebruikers.

In **hoofdstuk 4** beschrijven we het onderzoek waarin we het eerder ontwikkelde emotieherkenningssysteem onder meer realistische omstandigheden testten. Het onderzoek bevestigde dat deelnemers snel in staat waren om trilsignalen te leren, te onderscheiden en te koppelen aan de zes basisemoties. Deelnemers hadden het gevoel dat ze de trilsignalen konden gebruiken tijdens een gesprek met een acteur. Het kostte de deelnemers weinig moeite om de camera op de gesprekspartner gericht te houden. De software bleek de emotie blijdschap heel accuraat te kunnen detecteren, maar dit gold niet voor de andere vijf universele emoties. Op basis van dit onderzoek konden we concluderen dat het systeem nog essentiële verbeteringen nodig heeft op het gebied van emotieherkenningsprestaties en draagbaarheid, voordat het daadwerkelijk slechtziende en blinde personen kan ondersteunen tijdens alledaagse interacties. Desondanks zagen de deelnemers potentie in het systeem als hulpmiddel, ervan uitgaande dat het systeem in de toekomst aansluit bij de wensen van de eindgebruikers.

Het onderzoek dat wordt omschreven in **hoofdstuk 5** had als doel om te verkennen of een draagbaar GPS-navigatiesysteem met een trilriem slechtziende of blinde voetgangers kan ondersteunen om hun weg te vinden naar een bestemming in een onbekende omgeving. Er is onderzocht welke problemen men tegenkomt tijdens het navigeren en of de deelnemers de intentie zouden hebben om een dergelijk systeem te gebruiken in het dagelijks leven. Het prototype bestond uit een smartphone, een Raspberry Pi mini-computer en een riem met acht trilmotortjes. Het gidsen van de deelnemers door middel van trilsignalen werkte goed in realistische gebruiksscenario's.

Ondanks veel variatie tussen de deelnemers waren de signalen over het algemeen eenvoudig te interpreteren. Iedere deelnemer was na enkele instructies in staat om te navigeren met behulp van het systeem, al is de verwachting dat meer training tot nog betere resultaten zal leiden. De accuraatheid van GPS, welke in het beste geval ongeveer vijf meter was, was onvoldoende toereikend voor de doelgroep. Een mogelijke oplossing voor dit probleem zou zijn om het systeem wat flexibeler te maken, bijvoorbeeld door ervoor te zorgen dat routes zich snel aanpassen naar de locatie van de gebruiker, zodat deze continu aangepast worden op basis van de huidige positie van de gebruiker ten opzichte van de doellocatie.

Kortom, in dit onderzoek hebben we wensen onder slechtziende en blinde mensen geïdentificeerd. Deze wensen hebben met name betrekking op non-verbale communicatie en navigatie in onbekende omgevingen. Om aan deze wensen tegemoet te komen, is een draagbaar systeem ontwikkeld waarmee door middel van trillingen informatie over gezichtsuitdrukkingen van emoties en navigatierichtingen wordt overgedragen aan gebruikers. De verschillende studies hebben aangetoond dat deelnemers eenvoudig trillingen kunnen leren, interpreteren en gebruiken en dat ze enthousiast zijn over het concept. Ondanks dat er verschillende uitdagingen op het gebied van techniek en gebruiksvriendelijkheid aan het licht zijn gekomen die geadresseerd moeten worden bij toekomstige ontwikkelingen, heeft het onderzoekstraject waardevolle inzichten opgeleverd in de mogelijkheden die tactiele feedback kan bieden om slechtziende en blinde mensen in het dagelijks leven te ondersteunen.

Summary

The aim of the research presented in this dissertation was to explore the opportunities that novel technologies, such as wearable consumer electronics, offer to support persons with a visual impairment (PVIs) with activities of daily life. We used a human-centered design approach to investigate and develop new wearable technologies for persons with a visual impairment.

The objective of the first study in **chapter 2** was to assess priorities for the development and innovation agenda of these technologies, as expressed by prospective users of the technology applications, PVIs. Despite available technologies, a variety of urgent problems in the daily life of PVIs were identified. The difficulties PVIs experienced when finding their way in unfamiliar environments (including the detection of obstacles) withholds many PVIs to travel to new places independently. Also, the inability of PVIs to perceive nonverbal social cues has a direct impact on social interaction and leads to (perceived) social exclusion. Poor accessibility of texts encountered outside the home environment also proved to be a challenge. This study inspired us to focus on two challenges for PVIs: recognition of nonverbal communication and navigation in unknown environments.

In **chapter 3**, we reported the development and evaluation of a wearable system consisting of a head mounted camera, a laptop running emotion recognition software, and a haptic belt with small vibration motors to convey information to its users. The decision to use a haptic belt instead of other means of feedback was made to keep the eyes and ears of users free, as PVIs are highly dependent on their hearing – and in some cases remaining vision – to perceive their surroundings. We evaluated whether vibrotactile cues around the waist could be used to convey six basic emotions to users and whether such a device is desired by PVIs for use in daily living situations. Participants were able to interpret emotion cues accurately with the device after a short training session. The study showed that, under experimental conditions, vibrotactile cues can be used to convey facial expressions of basic emotions to PVIs in real-time.

In **chapter 4**, we describe the research in which we tested the emotion recognition system that was previously developed under more realistic conditions. We confirmed that participants were quickly able to learn, distinguish, and remember vibrotactile signals associated with the six emotions.

Participants felt they were able to use the vibrotactile signals during a conversation with an actor. Participants had no difficulties in keeping the camera focused on the conversation partner. The emotion recognition was very accurate in detecting happiness, yet performed unsatisfactorily in recognizing the other five universal emotions. From this study, we conclude that the system required essential improvements in emotion recognition performance and wearability before it is ready to support persons with visual impairments in their daily life interactions. Nevertheless, the participants saw potential in the system as an assistive technology, assuming that the system would meet the requirements of the end-users in the future.

In **chapter 5**, we report a study in which we explored whether a wearable GPS based navigation system with a vibrotactile belt could support pedestrians with a visual impairment to find their way towards a target destination in unfamiliar outdoor environments. It was investigated which problems they encounter while doing so, and if the users showed intention to use such a system in real life. The prototype consisted of a smartphone, a Raspberry Pi mini-computer, and a waist belt with eight vibrotactors. Vibrotactile signaling to guide PVIs through an unfamiliar environment worked well in real world usage scenarios. Despite a lot of variation between the participants, the signals conveyed were generally easy to interpret. All participants were able to navigate with the prototype after some instructions, although extensive training is likely to lead to better results. The GPS accuracy, which was approximately five meters at best, was too limited for PVIs. A solution would be to make the prototype less unforgiving, for example by adding flexibility by supporting adaptable routes that adjust to the position of the user compared to the target destination.

We have identified needs amongst PVIs, especially related to nonverbal communication and navigation in unknown environments. To address these needs, we have developed a vibrotactile wearable system to convey information about the emotion of faces and navigation directions. Our studies showed that participants can learn, interpret, and use the vibrotactile signals conveyed by the device easily and that they are enthusiastic about the concept. Although various technical and usability challenges were identified that need to be addressed in future development, the conducted research has resulted in valuable insights in the possibilities that tactile feedback offers to support PVIs in their daily lives.

Table of Contents

The research was funded by: Stichting voor de Technische Wetenschappen (STW): Take Off – Grant number: 15666; Oost NV: INTERREG-project MIND (Medische Innovaties Nederland Duitsland) – Grant number: 122035; and The Netherlands Organisation for Health Research and Development (ZonMW): Inzicht – Grant number – 94211004. We would like to thank the funding organizations for their support, especially considering the uncertain outcomes that were described in the original research proposal. In addition, we would like to thank all partners (Stichting Bartiméus, Koninklijke Visio, VicarVision, Cinoptics, Noldus Fachhochschule Münster) with whom we successfully collaborated during the project.

1 General Introduction

1.1 Smart wearables for persons with a visual impairment

With the introduction of Google Glass in 2013 and the Apple Watch in 2014, the world became widely familiar with a new generation of wearable consumer electronics. These devices combined the wearability of regular accessories, such as bracelets, watches, and glasses with the computing power and connectivity that smartphones offer. Such devices are often equipped with a wide variety of sensors (i.e. camera, microphone, motion sensors, compass), connectivity options (i.e. GPS, Bluetooth, 4G, Wi-Fi) and offer interaction through various modalities (i.e. auditory, visual, tactile, speech). It is multimodality that makes this new generation of wearable computers potentially interesting for a wide audience, including persons with sensory impairments such as persons who are blind or visually impaired (PVIs). However, research was needed to gain insights into the needs, wishes and usage of wearable technologies for persons with visual impairments [1].

Worldwide, there are an estimated 36 million persons who are considered fully blind (who have a visual acuity of less than 0.05), while there are another 217 million persons who have visual impairments that result in a visual acuity of less than 0.3 [2]. In the Netherlands alone, there are about 300,000 persons with a visual impairment [3]. Since many visual impairments are age-related and the population is aging, the number of PVIs is expected to increase in the future [4]. The type and severity of vision loss that visual impairments inflict are highly diverse. However, PVIs have in common that they are continuously faced with challenges in their everyday life due to their limitation to perceive visual information. Accessibility for PVIs in our society is often not optimal, because humans are highly visually oriented and have the tendency to use a lot of visual information for communication. As a result, it is difficult for some PVIs to participate in society to their satisfaction, with the possible consequence of perceived (social) exclusion, and possibly reduced independence and well-being.

In recent years, new legislation such as the Convention on the Rights of Persons with Disabilities has been introduced which forces society to be more inclusive, also for PVIs.[1] This includes adjusting the living environment to match the needs of PVIs (i.e. ticker in a traffic light). Alternatively, PVIs can equip themselves with tools to acquire visual information that they otherwise cannot perceive. Over the years, plenty of assistive aids have been developed for PVIs – such as (smart) canes, GPS trackers, and screen readers –, some of which have made it into the standard inventory of PVIs. However, many assistive aids are costly and often only available through health insurance. Consequently, there is a need amongst PVIs for more widely available and cheaper technologies, such as smartphones, to replace costly assistive aids [5].

Therefore, the aim of our research was to explore the opportunities that novel and widely available wearable consumer electronics offer to support PVIs to acquire visual information they encounter during activities of daily life. Hereby a strong emphasis was put on a human-centered design approach to avoid working towards technologies that are not wished for or needed by PVIs.

1.2 Human-centered design

We aimed to contribute to a more accessible society for PVIs, by exploring the opportunities that novel and consumer electronics offer together with the target group. To ensure that what we developed meets the needs (and real problems) of PVIs, rather than following a technology push, we put strong emphasis on following a human-centered design (HCD) approach. Ideally this means early involvement and understanding of end users, empirical testing with prototypes, and iterative design of a prototype, in this case to support PVIs with activities of daily life [6]. One of the aims of the HCD approach is to ensure an easy to use and useful system by putting an emphasis on user needs and requirements, to ensure that a system is beneficial for the tasks its intended users wish to use the system for [7]. Additionally, earlier research showed that technologies which are developed without consulting PVIs, are more likely to end up on a shelf in the attic rather than in the hands of the intended user [8]. It often happens that technologies are developed without the end-users in mind. Krishna and colleagues [9] for example, described a major pitfall in the development of navigation aids for PVIs. A review by Hakobyan and colleagues showed auditory

---

[1] UN CRPD. Ratified in The Netherlands since 14-07-2016.

feedback is common practice for mobile assistive technologies [10], even though using headphones means users cannot hear their surroundings, which is crucial for safe navigation of PVIs. Therefore, one of reasons to use the HCD approach was to avoid developing technological solutions that were not desired or needed by PVIs, with the risk of developing a technology that would not benefit the intended users in their daily lives. Already before developing a system prototype, user characteristics of PVIs, context of use, and end user needs should be explored [11]. Consequently, the first thing that was done in this project was to familiarize with the PVIs by conducting interviews with PVIs to gain insights in the biggest problems they face in their daily lives. Based on these findings, it was decided to focus on the possibilities that the (vibro)tactile modality offer to support PVIs during activities of daily life.

The user involvement should not end after this initial phase, as the essence of human-centered design is to involve end-users to enable them to influence a design [6, 12]. Also in the later stages of development and prototype testing, end users were actively involved. Mixed research methods were used to gain insights in both the functioning of the prototype and the experiences from PVIs using them. During the studies with emotion recognition and navigation prototypes, end users were actively involved and proved to be valuable sources of feedback. The final principle is the one of iterative design, which means that system development should quickly adapt to the findings and experiences of users [6]. Because of the rigidness of scientific research, it is not always possible to have short iteration. However, between studies, the prototypes were further developed based on both the technical performance of the system and the experiences from the participants. Generally, PVIs' experiences and needs formed the basis of important decisions, including the general research direction presented in this dissertation.

1.3 Sensory substitution and multimodality

In case of a visual impairment, the remaining functional senses are essential to perceive the world. This also means that much of the information one wishes to convey to persons with a visual impairment, be it temporal or permanent, should be translated to, and conveyed through auditory or tactile modalities. This is where sensory substitution comes into play.

The most famous and successful form of sensory substitution is Braille, which translates visual information (characters) into tactile information (dots), which allows PVIs to acquire information from written texts. Digital sensory substitution devices often consist of artificial receptors, such as cameras or microphones, to register real-world information, which is then conveyed to the user of the device through a different modality (i.e. visual information is conveyed to a user using a tactile display) [13]. A cornerstone of sensory substitution is brain plasticity, which is the notion that the central nervous system has an ability to adapt to signals received from different receptors [13]. With some training, it can learn how to make sense of signals that are acquired through artificial receptors. According to this idea, a PVI should for example be able to "see" a cup of coffee through tactile information, as was first proved by Bach-y-Rita in the sixties [14]. Bach-y-Rita and colleagues tried to help persons who were visually impaired to perceive and interact with various small objects by using tactile information. The participants were able to do so, by sitting in a chair which was equipped with a grid of vibrotactors, located in the back of the chair. On this grid, tactile representations of objects in the physical phase were presented to the users who, after extensive training, were able to grasp cups and moving balls. By doing so, Bach-y-Rita was able to substitute vision with haptics and is widely acknowledged as the first researcher to successfully build a personal sensory substitution device [14]. In more recent years, research towards sensory substitution devices has continued and has, amongst other applications, resulted in a tactile vest [15], tactile glove [16], and various belt applications [17–20]. These and other studies have reaffirmed that persons are able to give meaning to signals such as vibrotactile signals, and that the brain can adapt to artificial receptors, such as auditory or tactile interfaces [13]. In addition, research by Van Erp and colleagues has shown that tactile cues can be easily interpreted, even in cognitively demanding conditions [21], making multimodal user interfaces very interesting opportunities for sensory substitution systems.

In our research, we have put a strong emphasis on vibrotactile feedback to convey information. Vibrotactile feedback refers to the use of vibration to encode information [22]. In consumer electronics, vibration signals are often used to inform about events on the device, such as receiving text messages or getting news updates. A great advantage of tactile feedback is that it is simple, fast, and direct. Furthermore, tactile feedback does not reduce the hearing capacity of the user of a technology. This is particularly a benefit for PVIs, who are highly dependent on their hearing to perceive their surroundings in order to stay safe. Additionally, in the case of ambiguous auditory cues, tactile cues even help to process the incoming information nonetheless [23]. We investigated to what extent PVIs could benefit from multimodal information by providing vibrotactile cues to improve their ability to acquire information about their surroundings during two different activities of daily life: social interactions and during navigating.

1.4 Social interactions

During interactions between two or more persons, most information (approximately 65%) is transferred through nonverbal communication cues, such as nodding, gestures, and facial expressions [24]. This means that persons without vision are generally unable to perceive most of the information exchanged during social interactions. Persons without any sensory impairments rely heavily on both vision and hearing during social interactions. Thus, it is not hard to imagine that it is much more challenging, or even impossible, for PVIs to perceive nonverbal cues in interactions. Ultimately, this could lead to feelings of social exclusion. What if it is possible to use a different sense to make up, at least partly, for this loss? What if we could use vibrotactile signaling, and instead of the visual sense use the tactile sense to perceive nonverbal information during social interactions?

In recent years, various scientific studies have been conducted towards sensory substitution systems to support PVIs during social interactions, such as the social interaction assistant [17], which uses vibrotactile signals to guide PVIs around social venues and determined and conveyed the location that potential conversation partners would be at. Other efforts include a glove with vibration motors to convey facial expressions [16], and vision to audio substitution systems [25]. However, for a device to be used in real-life social interactions, we believe it should be unobtrusive, and even more so, keeping the ears and hands of its users free, as PVIs rely on their remaining senses to perceive their environment and use a variety of other assistive devices too.

In this dissertation, there is a focus on conveying emotional states of a conversation partner through a wearable device. Generally, emotional states are to a large extent communicated and perceived through visual cues. Of course, it is often possible to derive emotions from verbal cues such as speech or sounds associated to happiness (laughter) or sadness (crying and sniffing). However, it does make it a lot easier to determine whether someone is feeling happy if one can see a facial expression. Particularly, when conveyed emotions are ambiguous, or when the only way to determine how someone feels is visual – during quiet moments, or when PVIs are speaking themselves – it can be impossible for PVIs to determine how the other is feeling. For example, it can be very difficult to determine whether someone is silent because they are carefully listening or bored.

To assist PVIs with determining facial expressions of their conversation partners, we have developed a prototype of a wearable sensory substitution system. This system consists of a spectacles-mounted camera, a tablet running emotion recognition software [26,27], and a waist-worn vibrotactile belt. The principle is that the camera functions as artificial receptors which records (eyes that see) the surroundings of the PVI wearing the system. Then, the emotion recognition software recognizes a face and determines whether facial expressions are shown by the conversation partner of the PVI. Once a facial expression is detected and classified as an emotion, the detected emotion is conveyed to the wearer of the system through vibrotactile signals. By doing so, the tactile sense replaces a function of the visual sense, leading to sensory substitution.

1.5 Navigation

In addition to social interactions, a problem that receives attention in this dissertation is the mobility of PVIs. Being able to walk around wherever and whenever one wants is easy to do in a safe and efficient manner for (most) sighted persons. Even for PVIs, navigating in familiar surroundings is do-able. In their own house and neighborhood most people can navigate freely with or without mobility aids.

Independent mobility gets particularly difficult for PVIs when they are outside their familiar surroundings. To understand why this is the case, it is important to know what encompasses wayfinding. Wayfinding is made up of two important tasks. The first one is to sense the environment for obstacles and hazards, while the second one is to navigate towards destinations beyond the directly perceptible environment, which includes a continuous awareness about one's own position in relation to a travel destination [28]. If we look at the most used mobility aids amongst PVIs, these are white canes and guide dogs [29]. Of course, white canes are of great help for PVIs to detect obstacles and other hazards (e.g. cyclists, cars, stairs). Furthermore, guide dogs are impressive and can help to navigate familiar routes, yet their main purpose when it comes to mobility is to avoid obstacles. In recent years, there have been many developments applying different sensors, such as sonar, infrared, optical, and GPS. Additionally, the auditory [28,30–32] and tactile [21,33-44] modalities have been used to support PVIs to improve the ability of PVIs to navigate independently [10,29,45,46]. Despite these developments, only few of these systems have been commercialized and are widely used in the everyday life of PVIs [46]. As a result, there remain unmet needs when it comes to PVI mobility aids, particularly when it comes to awareness about one's location and walking unfamiliar routes, be it due to changes in familiar environments or exploring unfamiliar environments [5,47].

Important design requirements for navigation aids are that they should keep ears and eyes free, should be comfortable to wear, reliable, cheap, user friendly, robust, wireless, and able to convey information in real time [29]. Furthermore, in earlier research, scholars have identified a wish from PVIs for the use of widespread technologies as assistive aids [5]. To assist PVIs with navigation, we have evaluated whether it was useful to apply the same vibrotactile belt that was used for emotion recognition as a wearable navigation aid. Apart from the belt, the core component of the device was an Android smartphone, which was equipped with the necessary sensors to support wayfinding. A combination of a mainstream smartphone with a vibrotactile belt meets all the aforementioned requirements for a navigation aid and therefore was a worthwhile prototype to investigate.

1.6 Chapter contents

At the start of the research project, we elicited priorities for the development and innovation agenda of smart wearable technologies from prospective users of the to be developed technologies, PVIs. Based on the findings in this study, which are presented in **chapter 2**, we chose to channel our efforts towards solutions for problems encountered during wayfinding and social interactions. More specific, to support PVIs with waypoint navigation and with accessing facial expressions of emotions. The core of the developed prototype consisted of a tablet or smartphone to run task-specific software (e.g. navigation software or emotion recognition software) and a belt with nine vibrotactors to communicate from the system to its user. For the recognition of facial expressions of emotion, a USB-camera attached to spectacles was added to the system.

In **chapter 3**, we present the research in which we evaluated the prototype through user evaluations with both sighted and visually impaired participants. We sought to determine whether the device could improve one's ability to determine the facial expressions of others and whether such a device is desired for use by the target group. With the prototype, the PVIs were significantly better in recognizing the facial expressions of emotions than without it. Participants were quickly able to learn and apply the signals communicated by the prototype. While the findings were promising, the study was conducted in lab conditions that did not resemble reality (i.e. under ideal lighting conditions, validated sets of pictures and videos), meaning the findings were not generalizable to real-life conditions. Therefore, a follow-up study was deemed necessary to evaluate the prototype under more realistic conditions.

In **chapter 4**, we describe whether the emotion recognition system is also applicable and useful in conversations. To do so, we answered the following questions. First, are users able to continuously point the camera at the face of the conversation partner? Second, does the software reach a satisfactory recognition rate when it is confronted with an unstable video stream and suboptimal lighting conditions? Third, we wished to know whether people can interpret vibrotactile cues while engaging in a meaningful conversation. Finally, it is essential to determine whether the system lives up to the prospective users' requirements and whether they believe they would adopt such technologies in their daily lives.

Besides difficulties to determine the facial expressions of conversation partners, we found a need amongst PVIs for support during wayfinding in unfamiliar surroundings. In **chapter 5**, we explored whether a wearable GPS based navigation system based on a vibrotactile belt with nine vibrotactors could guide pedestrians with a visual impairment towards a target destination in an outdoor environment. PVIs were invited to perform various navigation tasks at the university campus with and without the wearable navigation system, which provided directions to the wearer in compass-like fashion, while being continuously observed by researchers. The aim was to investigate whether such a device would be useful, as well as to determine what sort of issues PVIs experience while visiting unfamiliar surroundings.

1.7 References

1. Williams MA, Buehler E, Hurst A, Kane SK (2015) What not to wearable. In: Proc. 12th Web All Conf. - W4A ’15. ACM Press, New York, New York, USA, pp 1–4
2. Bourne RRA, Flaxman SR, Braithwaite T, et al (2017) Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Heal 5:e888–e897. DOI 10.1016/S2214-109X(17)30293-0
3. Keunen JEE, Verezen CA, Imhof SM, et al (2011) Increase in the demand for eye-care services in the Netherlands 2010-2020. Ned Tijdschr Geneeskd 155:1–6
4. VISION2020 (2005) Vermijdbare blindheid en slechtziendheid in Nederland
5. Schölvinck AFM, Pittens CACM, Broerse JEW (2017) The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 111:201–217
6. Gould JD, Lewis C (1985) Designing for usability: key principles and what designers think. Communications of the ACM 28(3):300-311. DOI 10.1145/3166.3170
7. ISO (2019) Systems, Ergonomics of human-system interaction — Part 210: Human-centred design for interactive. https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:en. Accessed 18 Dec 2019
8. Phillips B, Zhao H (1993) Predictors of assistive technology abandonment. Assist Technol 5:36–45. DOI 10.1080/10400435.1993.10132205

9. Krishna S, Colbry D, Black J, Balasubramanian V, Panchanathan S (2008) A systematic requirements analysis and development of an assistive device to enhance the social interaction of people who are blind or visually impaired. Workshop on Computer Vision Applications for the Visually Impaired, James Coughlan and Roberto Manduchi, Marseille
10. Hakobyan L, Lumsden J, O'Sullivan D, Bartlett H (2013) Mobile assistive technologies for the visually impaired. Surv Ophthalmol 58:513–528. DOI 10.1016/j.survophthal.2012.10.004
11. Van Velsen L , Van der Geest TM, Klaassen R, et al (2008). User-centered evaluation of adaptive and adaptable systems: a literature review. The Knowledge Engineering Review 23(3):261-281. DOI 10.1017/S0269888908001379
12. Abras, C, Maloney-Krichmar, D, Preece, J (2004) User-centered design. In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications
13. Bach-y-Rita P, Kercel SW (2003) Sensory substitution and the human-machine interface. Trends Cogn Sci. DOI 10.1016/j.tics.2003.10.013
14. Bach-y-Rita P (1967) Sensory plasticity: applications to a vision substitution system. Acta Neurol Scand 43:417–426. DOI 10.1111/j.1600-0404.1967.tb05747.x
15. Novich SD, Eagleman DM (2015) Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. Exp Brain Res 233:2777–2788. DOI 10.1007/s00221-015-4346-1
16. Krishna S, Bala S, McDaniel T, et al (2010) VibroGlove: an assistive technology aid for conveying facial expressions. Proc SIGCHI Conf Hum Factors Comput Syst Ext Abstr 3637–3642. DOI 10.1145/1753846.1754031

17. Panchanathan S, Chakraborty S, McDaniel T (2016) Social interaction assistant: a person-centered approach to enrich social interactions for individuals with visual impairments. IEEE J Sel Top Signal Process 10:942–951. DOI 10.1109/JSTSP.2016.2543681
18. McDaniel TL, Villanueva D, Krishna S, et al (2010) Heartbeats: a methodology to convey interpersonal distance through touch. In: Proc. 28th Int. Conf. Ext. Abstr. Hum. factors Comput. Syst. - CHI EA ’10. pp 3985-3990
19. McDaniel T, Krishna S, Balasubramanian V, et al (2008) Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. In: HAVE 2008 - IEEE Int. Work. Haptic Audio Vis. Environ. Games Proc. pp 13–18
20. McDaniel TL, Krishna S, Colbry D, Panchanathan S (2009) Using tactile rhythm to convey interpersonal distances to individuals who are blind. Proc CHI 2009 Ext Abstr 4669–4674. DOI 10.1145/1520340.1520718
21. Van Erp JBF, Van Veen HAHC, Jansen C, Dobbins T (2005) Waypoint navigation with a vibrotactile waist belt. ACM Trans Appl Percept 2:106–117. DOI 10.1145/1060581.1060585
22. MacLean KE, Schneider OS, Seifi H (2017) Multisensory haptic interactions: understanding the sense and designing for it. In: Handb. Multimodal-Multisensor Interfaces Found. User Model. Common Modality Comb. - Vol. 1. ACM, pp 97–142
23. Oviatt S (2017) Theoretical foundations of multimodal interfaces and systems. In: Handb. Multimodal-Multisensor Interfaces Found. User Model. Common Modality Comb. - Vol. 1. ACM, pp 19–50
24. Knapp M, Hall J, Horgan T (2013) Nonverbal communication in human interaction. Cengage Learning, Wadsworth
25. Meijer PBL (1992) An experimental system for auditory image representations. IEEE Trans Biomed Eng 39:112–121. DOI 10.1109/10.121642
26. Van Kuilenburg H, Wiering M, Model MJDUA, et al (2008) Facereader 6.1 – How do they do it?
27. Noldus FaceReader. http://www.noldus.com/human-behavior-research/products/facereader?gclid=CNPm3enUrM8CFYYcGwodyp0BJQ. Accessed 26 Sep 2016

28. Loomis JM, Golledge RG, Klatzky RL (1998) Navigation system for the blind: auditory display modes and guidance. Presence Teleoperators Virtual Environ 7:193–203. DOI 10.1162/105474698565677
29. Tapu R, Mocanu B, Tapu E (2015) A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments. In: 2014 11th Int. Symp. Electron. Telecommun. ISETC 2014 - Conf. Proc. IEEE, pp 1–4
30. Roentgen UR, Gelderblom GJ, de Witte LP (2011) Users' evaluations of four electronic travel aids aimed at navigation for persons who are visually impaired. J Vis Impair Blind 105:612–623. DOI 10.1177/0145482X1110501008
31. Ran L, Helal S, Moore S (2004) Drishti: an integrated indoor/outdoor blind navigation system and service. In: Second IEEE Annu. Conf. Pervasive Comput. Commun. 2004. Proc. IEEE, pp 23–30
32. Wilson J, Walker BN, Lindsay J, et al (2007) SWAN: System for wearable audio navigation. In: 2007 11th IEEE Int. Symp. Wearable Comput. IEEE, pp 1–8
33. Velázquez R, Pissaloux E, Rodrigo P, et al (2018) An outdoor navigation system for blind pedestrians using GPS and tactile-foot feedback. Appl Sci 8:578. DOI 10.3390/app8040578
34. Meier A, Matthies DJC, Urban B, Wettach R (2015) Exploring vibrotactile feedback on the body and foot for the purpose of pedestrian navigation. In: Proceedings of the 2nd international Workshop on Sensor-based Activity Recognition and Interaction (iWOAR '15), pp 1–11. DOI 10.1145/2790044.2790051
35. Dobbelstein D, Henzler P, Rukzio E (2016) Unconstrained pedestrian navigation based on vibro-tactile feedback around the wristband of a smartwatch. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16), pp 2439–2445. DOI 10.1145/2851581.2892292
36. Spiers AJ, Dollar AM (2016) Outdoor pedestrian navigation assistance with a shape-changing haptic interface and comparison with a vibrotactile device. In: 2016 IEEE Haptics Symp. IEEE, pp 34–40
37. Kärcher SM, Fenzlaff S, Hartmann D, et al (2012) Sensory augmentation for the blind. Front Hum Neurosci 6:37. DOI 10.3389/fnhum.2012.00037

38. Tsukada K, Yasumura M (2004) ActiveBelt: Belt-type wearable tactile display for directional navigation. In: Davies N, Mynatt ED, Siio I (eds) UbiComp 2004: Ubiquitous Computing. UbiComp 2004. Lecture Notes in Computer Science, vol 3205. Springer, Berlin, Heidelberg. DOI 10.1007/978-3-540-30119-6_23
39. Pielot M, Henze N, Heuten W, Boll S (2008) Evaluation of continuous direction encoding with tactile belts. In: Haptic Audio Interact. Des. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–10
40. Heuten W, Henze N, Boll S, Pielot M (2008) Tactile wayfinder: A non-visual support system for wayfinding. Proc 5th Nord Conf Human-computer Interact Build Bridg, pp 172–181. DOI 10.1145/1463160.1463179
41. Adame MR, Jing Yu, Moller K, Seemann E (2013) A wearable navigation aid for blind people using a vibrotactile information transfer system. In: 2013 ICME Int. Conf. Complex Med. Eng. IEEE, pp 13–18
42. Cosgun A, Sisbot EA, Christensen HI (2014) Evaluation of rotational and directional vibration patterns on a tactile belt for guiding visually impaired people. In: 2014 IEEE Haptics Symp. IEEE, pp 367–370
43. Flores G, Kurniawan S, Manduchi R, et al (2015) Vibrotactile guidance for wayfinding of blind walkers. IEEE Trans Haptics 8:306–317. DOI 10.1109/toh.2015.2409980
44. Dura-Gil JV, Bazuelo-Ruiz B, Moro-Perez D, Molla-Domenech F (2017) Analysis of different vibration patterns to guide blind people. PeerJ 5:e3082. DOI 10.7717/peerj.3082
45. Giudice NA, Legge GE (2008) Blind navigation and the role of technology. Eng Handb Smart Technol Aging, Disabil Indep 479–500. DOI 10.1002/9780470379424.ch25
46. Gori M, Cappagli G, Tonelli A, et al (2016) Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. Neurosci Biobehav Rev 69:79–88. DOI 10.1016/j.neubiorev.2016.06.043
47. Quiñones P, Greene T, Yang R, Newman M (2011) Supporting visually impaired navigation: A needs-finding study. Chi2011 1645–1650. DOI 10.1145/1979742.1979822

2 Setting The Development And Innovation Agenda For Technology Applications For Persons Who Are Blind Or Visually Impaired: A User-Centered Approach [2]

Abstract

The objective of the study was to elicit priorities for the development and innovation agenda of (smart) accessible technologies, as expressed by PVIs. Open interviews were conducted with 26 PVIs (mean age: 48.8, SD: 17.6) to explore the biggest problems that PVIs encounter in their daily life, and what technological aids they used to cope with these problems. Structured phone interviews were conducted with 27 PVIs (mean age: 41.7, SD: 15.5). During this interview, 59 activities of daily life, derived from the problems mentioned during the first interview and visual functioning questionnaires, were rated on difficulty and importance. The biggest problems mentioned during the first interviews were related to mobility (e.g. obstacle detection, orientation, and wayfinding in unfamiliar environments), recognition of faces, and acquiring written information. During the telephone interviews, recognition of persons at a distance, navigating unfamiliar surroundings, and reading labels on product packaging, medication, and phone books received the highest ratings in terms of difficulty and importance. Despite available technologies, a variety of urgent problems in the daily life of PVIs were identified. The difficulties PVIs experience when finding their way in unfamiliar environments (including the detection of obstacles) withholds many PVIs to travel to new places independently. Also, the inability of PVIs to perceive non-verbal social cues has a direct impact on social interaction and leads to (perceived) social exclusion. Poor accessibility to texts encountered outside the home environment also proved to be a challenge.

---

[2] Buimer HP, Zhao Y, Wentzel J, Van der Geest TM, Van Wezel RJA (2017) Setting the development and innovation agenda for technology applications for persons with visual impairments or blindness: A user-centered approach. *Available on ResearchGate*. DOI 10.13140/RG.2.2.16247.04002

Future development and innovation should focus on these problems. We hope to inspire the community that develops assistive technologies for PVIs to start developing novel assistive technologies with an emphasis on the end-user in every stage of the development, from concept to product to ensure useful solutions to relevant issues.

2.1 Introduction

Around 285 million persons worldwide are severely visually impaired in both eyes, 39 million of whom are fully blind [1]. The daily lives of PVIs are often strongly affected by their visual impairment, resulting in challenges during activities of daily life (ADL). Although these challenges are often different for persons with low vision or blindness [2], they cause a risk for exclusion from school, work, and society [3–5].

Assistive aids to cope with such challenges are often costly and thus PVIs wish for common technologies (such as smartphones) to replace them [5,6]. Recent advances in smart technology show interesting features that might be useful as assistive aids [7–9]. Already, many PVIs choose smartphones over task-specific assistive technologies available, despite limited accessibility features of smartphones [2,10]. Advances in smart mobile technologies, computer vision, and wearables might further boost the development of novel assistive technologies [9].

The last years research and development in the multidisciplinary field of assistive technology for PVIs expanded rapidly to a community that can be roughly divided in four sections: multisensory research, accessible content processing, accessible user interface design, and mobility and accessible environments research [9]. For comprehensive overviews of available assistive technologies for PVIs we refer to [7–9,11,12].

As researchers we wish to contribute to the development of novel assistive technologies that support PVIs. Many studies pose solutions for everyday problems of PVIs, often initiated with novel technologies in mind. As a result, technical solutions for a specific activity or usage context are presented. These studies are often focused on themes such as accessibility, independent travelling, in- and outdoor navigation, object detection, and reading written information [9]. A glance at the titles and abstracts of recent editions of major conferences in the field of assistive technologies for PVIs (as presented in [9]), shows a similar trend where many contributions were focused on technological solutions for (indoor) navigation [13–16], object recognition [17,18], accessibility of websites and code [19], 3D-printed or otherwise tangible maps [20,21], and even guide drones [22]. Only few studies seem to start with an assessment of the needs of PVIs [23,24].

We believe that PVIs will only gain advantage from technological advances if these address relevant issues. Involving prospective users throughout research and development, increases the chances of an acceptable, useful, and easy to use technology application [8,25–27]. Therefore, it is crucial to determine what activities are difficult and to understand the urgency of solving these problems, before developing technical solutions. Potentially difficult ADL can be derived from visual functioning questionnaires (VFQ), which serve to assess the visual functioning of PVIs. VFQ determine the visual acuity in specific eye diseases (e.g. glaucoma [28,29], cataract [30–33], macular degeneration [34,35], and general low vision [36,37]). After an assessment of the difficulty and importance of ADL with PVIs, an empirical user-centered foundation for the design and development of novel technologies can be formed.

The objective of the current study was to elicit priorities for the development and innovation agenda for (smart) accessible technologies, as expressed by PVIs. Not to establish just a wish list, but to create an overview of problems that PVIs experience in their daily life to inspire technology developers and designers to create relevant technologies and applications.

2.2 Study 1: Exploring daily life problems of PVIs

2.2.1 Objective

The goal of the first study was to explore the biggest problems that PVIs encounter in their daily lives, and to determine which technological aids (e.g. smartphone apps) they used to cope with these problems. To identify these problems, interviews were conducted with visitors of an assistive technology exhibition in the Netherlands.

2.2.2 Participants

Eligibility to participate was checked based on three criteria: persons had to be partially sighted or blind, aged 18 or older, and had no other cognitive or sensory impairments. 26 persons participated (9 females, 17 males), who varied in age (mean=48.8, SD=17.6, min=19, max=87). Seven persons were fully blind, whereas the others varied in type and severity of vision loss (e.g. nearly blind, blurred vision, loss of peripheral vision, loss of central vision). Fifteen participants had congenital visual impairments, whereas ten others acquired their visual impairment later in life. One person did not disclose this information.

2.2.3 Materials

The semi-structured interview questions prompted participants to describe the biggest problems that PVIs experience in their daily lives. Furthermore, participants were asked to describe which assistive technologies they used to cope with these problems and to what extent these technologies enabled them to do what they desire (Table 2.1 for a full list of the interview topics).

Table 2.1. Topics of the face-to-face interviews.

1. Respondent demographics
    a. Gender
    b. Age
    c. Type of visual impairment?
    d. Congenital or later developed?
    e. Are you using a smartphone?
2. Questions
    a. Can you briefly describe the biggest problem that you encounter in daily life due to your visual impairment?
        Of which problems, do you believe they should be fixed by means of existing technology?
    b. What assistive technology do you use to counter these problems?
        What are the pros and cons of these assistive technologies?
    c. (If the answer to 1d is yes) Do you use your smartphone as an assistive technology to assist in performing this activity?
        What apps do you use?
    d. Are there daily activities that you cannot/will not do because of your visual impairment, which you would like to do independently?

2.2.4 Procedure

Visually impaired visitors of the exhibition were approached to participate in the interview. After obtaining informed consent of the participant the interviews were conducted. The informed consent and interview were audio-recorded, which resulted in a total 2 hours and 46 minutes of recorded interviews (6 minutes per interview). While participants were asked to describe only the biggest problem in their daily life, some participants used the occasion to explain a variety of problems that impacted their daily lives.

The study protocol was evaluated and approved in accordance to the Declaration of Helsinki by the Ethics Committee Behavioral Science of the University of Twente.

2.2.5 Data analysis

The interviews were transcribed, anonymized, divided into episodes, and analyzed by a single coder. The often-specific problems described by the participants were analyzed, grouped and labeled in more generic activities of daily life. Also, the use of technological aids was analyzed.

2.2.6 Results

Problems with orientation and mobility.

The most frequently mentioned problematic activities were getting around in unfamiliar outdoor environments, orientation, and detecting obstacles in your way (all mentioned five times). The issues were diverse and often closely related to other problems. Orientation and mobility have been recognized as problems for PVIs earlier and thus this finding is not surprising, and in line with earlier research, innovations, and developments regarding this topic [2,12,38–40]. Mobility and accessible environment research was even considered as one of the four major research communities in assistive technology research [9].

Getting around in unfamiliar outdoor environments was mentioned as the most critical problem by five participants. A 41-year-old (congenital low vision, late blind) stated a desire to be able to visit unfamiliar places without a lot of effort (e.g. go shopping wherever and whenever she wants). Currently, this requires too much effort and withholds her from going to new places and exploring new venues. Wayfinding in familiar surroundings was far less often reported as a problem. One participant explained this was due to his ability to define landmarks, such as lampposts, in his familiar surroundings.

However, PVIs often reported an inability to detect and recognize obstacles, which even made wayfinding in familiar surroundings difficult sometimes. A 29-year-old postnatally fully blind participant told that pavements, often full of shop signs, bikes, and other obstacles, pose a tripping danger if obstacles are non-detectable with her cane. Besides canes and guide dogs, plenty of assistive technologies aimed at mobility and orientation were used by the participants (Trekker Breeze, Ariadne GPS, Blindsquare, Navigon, ViaOpta Navigator for wayfinding; the Dutch public transport planning website 9292OV.nl; and the iCane for both wayfinding and obstacle detection). However, due to the inaccuracy of GPS based technologies and the fact that not all maps are optimized for pedestrians, participants complained about situations where they could not find the location they were looking for. One participant even described how she was guided onto a motorway by her navigation aid.

Problems with seeing facial features/expressions up close.

An urgent problem that only received little attention in literature is the inability to recognize persons and their facial features. However, four participants described this as their biggest problem. A 29-year-old participant (congenitally partially sighted) said she could not recognize her professional contacts on conferences. Consequently, she has to wait for others to come to her, while she wants to take the initiative sometimes. Furthermore, she wished to know who is around. For two participants, who saw silhouettes and were able to locate persons, their most pressing problem was the inability to distinguish details about facial features and expressions. A 44-year-old congenitally partially sighted participant especially missed such information in social interactions at work or in pubs.

Applications for person or emotion recognition are slowly hitting the consumer market (e.g. Orcam). However, only few studies have been done to test their effectiveness for people with visual impairments [41–45]. None of the mentioned applications were used by the participants in this study.

Problems with written information.

Reading written information was often mentioned as a major problem. Information found in small texts (e.g. information on product packaging), menu cards, public transport information screens, and street signs are highly inaccessible for PVIs. This is particularly the case outside the home environment, where it is often impossible or impractical to make use of available aids. A 34-year-old congenitally partially sighted participant described the difficulty faced when trying to find specific products on a supermarket shelf. Another participant (30-year-old partially sighted from birth) reported he was not able to find and read street signs announcing construction zones, which led to uncertainty and might lead to potentially dangerous situations.

The participants reported various assistive technologies used to access written information, of which digital and non-digital magnifiers and screen readers on PC were mentioned most often. In addition, smartphone applications such as VoicEye and KNFB Reader were popular. Besides text-to-speech functionalities, these applications include other functions like magnification and contrast enhancement. Participants also used (multi-purpose) assistive aids, such as a daisy player (to listen to audiobooks), Orion Webbox (to hear TV subtitles), and Bones Milestone which can be used for audiobooks, as a memo-recorder and as an alarm clock. Finally, participants used the speech interface of their smartphones, which enables them to use phone functions such as its memo recorder or calendar.

Problems with technology use.

The experiences with technology were mixed. Only 16 out of 26 participants were using a smartphone, and thus PVIs are not equally technology-ready, which is important to know for developers of technological aids. Also how they experienced technology use differed, best illustrated by touch screen usage, for which the accessibility was perceived differently between usage scenarios. One person (42 years old, postnatally partially sighted) could not understand how touch screens could possibly help PVIs, while others stated that touch screens on smartphones work better than desktop computers with assistive software, due to the limited information shown on the small screen. Problems mentioned when using a computer involved using screen magnifiers and inaccessible websites. Magnifiers are difficult to work with since zooming in on content comes at the expense of losing a general overview of the screen, as was emphasized by a 76-year-old partially sighted participant. Three participants reported using appliances with voice feedback, such as a talking thermostat or weighing scale.

2.2.7 Summary

Various problems were identified that PVIs encounter in their daily lives. Most of the problems were associated to mobility. Mobility related ADLs described by PVIs were orientation, the detection of obstacles, stairs, and curbs, as well as finding your way in unfamiliar surroundings. The technological aids participants used to tackle these problems, were mostly aimed at wayfinding. Only the iCane offered obstacle detection through the cane and through a sensor to detect obstacles above the waist. A second issue that was often mentioned by the participants was the inability to recognize facial expressions and features, which led to awkward social interactions. No aids were used by PVIs to acquire such information. Finally, written information is difficult to access for PVIs, especially when information must be acquired from non-digital sources such as product packaging, newspapers or correspondence. Despite available technologies such as OCR and text-to-speech, accessing written information remained a problem.

2.3 Study 2: Zooming in on problematic ADLs

2.3.1 Objective

The goal of the second study was to create a list of ADLs that require attention of developers and designers of (smart) assistive technologies and applications. The list, which was based on VFQs and extended with findings from the earlier conducted interviews, included activities that were rated by the participants on perceived difficulty and importance during a telephone interview.

2.3.2 Participants

Participants were recruited from a pool of PVIs who participated in the first study, or were in a participant pool of the Royal Dutch Guide Dog Foundation or the Accessibility Foundation. 27 PVIs participated (14 female, 13 male), from all age groups (mean=41.7, SD=15.5, min=17, max=67), 8 of whom were fully blind, while 17 varied in types and severity of vision loss (e.g. nearly blind, blurred vision, loss of peripheral vision, loss of central vision). The visual impairment of the remaining two persons was not disclosed.

2.3.3 Materials

To ensure the structured interview covered a wide variety of relevant activities, the ADLs from the first study were expanded with activities from VFQ. These questionnaires address many specific ADLs (e.g. cooking, grooming, reading small texts), making them very useful to determine which activities cause problems for PVIs. Although the VFQs do not include a (comprehensive) construct aimed at technology usage, the participants in the interviews reported urgent problems associated to this topic. Therefore, problematic activities discussed in the first interviews and the VFQ were merged to create a list of ADL, after which duplicates were removed. After adding technology-related activities, a list of 59 ADLs was formed (the activities and their source can be found in Table 3).

2.3.4 Procedure

After oral informed consent and permission to record the interview were obtained, the interview was conducted via telephone. All ADLs were read out loud to the participant one-by-one. Participants were asked to rate each of the activities on difficulty in performing the ADL on a five-point scale (1 = not difficult, 5 = impossible to do), and in on importance of the ADL in their daily lives on a four-point scale (1 = not important, 4 = extremely important) in line with the scales used in the earlier mentioned questionnaires [29,30,35,36]. Participants were asked to rate the activities while keeping in mind the assistive technologies they would normally use during the activity. Additionally, ADLs could be marked as irrelevant when participants would not bother to do the ADL even if they could (scored as null data). Participants could elaborate on the score they gave each item, leading to an average interview time of 47 minutes, with the shortest just above 20 minutes and the longest being almost 80 minutes.

2.3.5 Data analysis

The ratings provided by participants resulted in difficulty and importance scores for each of the ADL. For each activity, an overall score was calculated by multiplying the mean scores for difficulty and importance and the number of persons who deemed the activity relevant (mean difficulty × mean importance) × N, in line with earlier research [46]. This score was used to order the list of activities.

2.3.6 Results

The results of the telephone interview can be found in Table 3, where the 15 ADLs with the highest score are presented. Like the face-to-face interviews, the biggest problems were related to mobility, social interactions, and accessibility of printed and written texts. The ADLs did not differ much in importance ratings, and were all rated somewhat between important and extremely important. However, the ranking of ADLs once difficulty and relevance were taken into account was different. When looking at the overall scores, recognizing persons at a distance, getting around in unfamiliar surroundings and identifying products on a full shelf were amongst the most highly ranked activities.

Table 2.2. Overview of the ADL ranking.

			Difficulty		**Importance**		
Activity	n	Source	M *	SD	M *	SD	Score
Recognizing persons at a distance	27	*A,B,C,D, E,F,I,J, F2F*	4.4	0.89	3.2	1.04	382.2
Getting around in unfamiliar surroundings	27	*D,G,I,J, F2F*	3.6	1.22	3.8	0.40	366.2
Identifying products on a full shelf	25	*E,F,I,J*	4.0	1.00	3.4	0.77	344.0
Reading labels on product packaging/ medication/ phone book	26	*C,E,G,I, J,K,F2F*	3.8	1.39	3.5	0.71	342.7
Getting around in unfamiliar buildings	27	*D,I,J*	3.4	0.97	3.5	0.75	320.3
Reading street names	25	*E*	4.1	1.22	3.1	1.09	318.2
Recognizing persons nearby	27	*A,B,C,F, I,J, F2F*	3.4	1.34	3.4	1.01	310.1
Reading menus	26	*I,J*	3.8	1.47	3.0	1.08	300.8
Getting around in a supermarket	27	*F2F*	3.1	1.40	3.5	0.75	295.6
Finding small dropped objects	27	*A,B*	3.4	1.11	3.2	0.62	289.9
Operating machines	26	*new*	3.1	1.23	3.5	0.65	283.1
Moving around in a crowd	27	*I, J*	2.9	1.33	3.5	0.70	275.0
Reading television subtitles	24	*C,I,J,F2 F*	3.5	1.47	3.2	1.05	266.0
Identifying products in your hand	27	*F2F*	2.7	1.26	3.5	0.64	260.4
Identifying landmarks	27	*new*	2.6	1.19	3.7	0.48	256.7

Reading print in letters/ mail/ invoices/ flyers	27	*D,G,I,J, F2F*	2.8	1.39	3.4	0.88	256.1
Identifying building entrances	27	*New*	2.4	1.15	3.8	0.40	251.8
Noticing objects off to either side	26	*A,B,E,K*	3.0	1.43	3.1	0.89	243.1
Travelling independently	26	*D*	2.3	1.26	3.9	0.27	239.3
Noticing objects in your way	27	*A,B,I,J,F 2F*	2.5	1.28	3.5	0.70	239.3
Cooking	26	*K,F2F*	2.5	1.14	3.6	0.57	235.0
Using public transportation	25	*G,I,J*	2.4	1.23	3.8	0.47	234.2
Playing table and card games	24	*I,J*	3.0	1.04	3.2	0.76	231.2
Awareness of current position	26	*F2F*	2.4	1.17	3.7	0.56	230.2
Crossing a road	27	*B,D,I,J,F 2F*	2.3	1.02	3.7	0.45	228.2
Reading road signs	21	*E,I,J,F2 F*	4.2	1.18	2.5	1.25	224.6
Visiting a website on a smartphone	26	*F2F*	2.5	1.45	3.4	0.85	223.4
Using the computer to find information on a website	26	*K,F2F*	2.4	1.24	3.6	0.70	221.8
Operating household appliances	27	*G,J*	2.1	1.03	3.7	0.47	214.8
Using the computer to order/buy something from an online shop	23	*K*	3.0	1.11	3.1	0.87	212.9
Matching clothing	26	*E,I,J*	2.4	1.06	3.4	0.81	212.2
Identifying stairs curbs and using them	27	*A,B,D,K*	2.2	1.19	3.5	0.75	211.1
Reading ordinary sized prints in books/ newspapers/ magazines	26	*A,B,C,E, F,I,J,K*	2.5	1.53	3.2	0.92	205.0
Payment using deposit money	26	*I*	2.2	1.13	3.6	0.90	203.9
Reading texts in dim light	16	*F2F*	4.1	1.36	3.1	1.06	202.1
Internet banking	25	*new*	2.2	1.27	3.5	0.92	197.1
Using contactless payment cards	25	*new*	2.2	0.96	3.6	0.71	195.8
Using a remote control	26	*new*	2.4	1.21	3.0	1.08	191.4
Keeping your clothes clean	27	*I*	2.1	1.00	3.4	0.93	188.7
Housekeeping	26	*K*	2.1	0.99	3.3	0.63	184.0

Watching television	26	*G,I,J*	2.4	1.30	2.8	1.16	179.3
Going from light to dark surroundings or vice versa	21	*A,B,I,J,K*	3.0	1.16	2.8	1.03	174.2
Doing needlework and handicraft	20	*C,K*	3.1	1.23	2.8	0.95	170.8
Seeing in a dim light environment	18	*A,B,D*	3.2	1.17	2.9	1.00	170.8
Getting around in familiar surroundings	27	*D,G,I,J*	1.6	0.69	3.7	0.53	164.6
Taking pictures with a smartphone	23	*new*	2.9	1.18	2.5	0.95	163.6
Grooming	27	*E,G*	1.7	0.68	3.6	0.79	163.3
Watching the clock	26	*I,J*	1.8	1.23	3.3	0.94	157.3
Operating computers with touch screens (e.g. tablets)	24	*F2F*	1.8	0.98	3.6	0.77	155.9
Getting around in familiar buildings	27	*D,I,J*	1.7	1.03	3.4	0.79	155.0
Holding place while reading	21	*I,J*	2.5	1.60	2.8	1.12	148.9
Maintaining an agenda	26	*new*	1.5	0.90	3.5	0.95	141.5
Gardening	18	*I,J*	3.0	1.19	2.6	1.25	138.0
Using Whatsapp/ SMS	27	*new*	1.4	0.64	3.5	0.80	137.2
Playing computer games	18	*new*	3.2	1.40	2.2	1.15	125.7
Calling with a smartphone	27	*new*	1.3	0.71	3.5	0.85	119.6
Sending and receiving email	27	*new*	1.2	0.51	3.6	0.84	118.6
Watching teletext	17	*new*	3.1	1.65	1.8	1.07	96.6
Descending stairs in dim light	22	*new*	1.8	1.05	2.4	1.18	94.5

Difficulty: (1) not difficult (2) somewhat difficult (3) difficult (4) extremely difficult (5) impossible
Importance: (1) not important (2) somewhat important (3) important (4) extremely important
n = Number of participants who found activity relevant (total n=27)
Source: F2F: Derived from the face-to-face interviews; new: Not derived from literature or face-to-face interviews; A: [28]; B: [29]; C: [30]; D: [31]; E: [32]; F: [33]; G: [34]; H: [35]; I: [36]; J: [37]; K: [47]
M = Mean
** = Means are rounded to one decimal*
SD = Standard Deviation
*Score: (M Difficulty * M Importance) * N. Note that the mean scores presented in the table are rounded off to one decimal. Scores were calculated with non-rounded off mean score.*

2.4 Conclusions and recommendations

The primary objective of the study was to elicit priorities for the development and innovation agenda of accessible technologies, as expressed by PVIs. The findings from this study can be beneficial for anyone willing to contribute to smart technology applications aimed at PVIs. Several ADLs were identified that require attention of designers and developers of smart technology. These activities can be roughly divided into three themes, which are: mobility, social interactions, and written information. While mobility and written information issues get a lot of attention in the literature and on conferences, this is not the case for social interactions.

To establish a thorough development and innovation agenda, it is insufficient to only determine which ADL is most difficult or important. Instead, both difficulty and importance, as well as relevance of a specific activity should be considered. Combining difficulty and importance scores led to new insights and to a ranking of ADLs that differed from a difficulty-only ranking (with extreme differences such as *playing computer games*, -41; and *reading texts in dim light*, -31).

Furthermore, participants were asked to consider their use of assistive technologies when determining the difficulty of activities. By doing so, some activities were possibly rated lower on difficulty than they would be if assistive technologies were not taken into account. For example, potentially difficult activities like maintaining an agenda, watching the clock, and internet banking received low difficulty scores, as suitable alternatives are available or successful assistive aids exist to support these activities.

2.4.1 Mobility

Wayfinding in unfamiliar surroundings, both in- and outdoors, was identified as a difficult and important problem which even withholds PVIs from going to new places. PVIs cannot familiarize themselves with new surroundings in the same way sighted people can [48,49], which makes it difficult to orientate themselves. PVIs usually plan their travels beforehand and have the apps to do so. However, wayfinding often relies on descriptions of landmarks that are difficult or impossible to explore beforehand for PVIs, as regular maps do not include the type of landmarks needed for PVIs [6,40]. PVIs are unable to create a detailed mental map before visiting an area, which limits them when visiting unfamiliar places [40,41], and thus some participants avoid visiting new environments.

Furthermore, some participants emphasized their struggles with navigation technologies. GPS based navigation technologies are often too inaccurate for pedestrian wayfinding, nor do they provide the information needed to get around, which has been proposed in earlier research too [6,40]. Therefore, it is important to think of ways that improve the accuracy of wayfinding information for pedestrians with a visual impairment.

Activities related to wayfinding in familiar environments were rated relatively low in study 2 (45th and 50th respectively). However, as mentioned in the first interview, also familiar public spaces are often subject to change (e.g. road constructions, obstacles), sometimes rendering familiar routes and localization cues useless [40]. Also other obstacles faced in public space cause difficulties (e.g. parked bikes, poles). Even with a cane, some PVIs have difficulty to safely pass such obstacles. Earlier research identified a wish from the PVI community to enhance navigation systems with information about such obstacles in public space, both indoors and outdoors [6]. Besides obstacles, changing lighting conditions and positioning of furniture proofed to be challenging [4]. Therefore, solutions for mobility issues, especially to enhance visits to unfamiliar environments (both indoors and outdoors) and detection of obstacles, are activities worthy of further technological research and developments. This was also identified as a future trend of research in the field of assistive technologies for PVIs in [9].

2.4.2 Social interactions

There seems to be a wish for technological aids that support the recognition of persons and their facial features and expressions. It was known that PVIs feel excluded due to their inability to access nonverbal information, such as facial expressions, hand gestures, knowing who is around, and orientation to these persons [4,41,50]. Considering the importance of nonverbal visual communication in social interactions and the impact it has on social inclusion [4], it is not surprising that this study is not the first to recognize this issue. Both in professional (e.g. conferences) and in informal social interactions (e.g. pubs), PVIs find their inability to perceive nonverbal cues a problem. Some studies have tried sensory substitution devices to convey information such as person identity and the recognition of emotions of conversation partners [41,43,45,51,52]. However, it does seem that this line of research has not yet been embraced by a larger community as a possible assistive technology [9].

More developments are needed before such technologies are ready to hit the consumer market.

2.4.3 Written information

In both studies, various activities related to the inability to access written or printed information from the direct environment were discussed and marked as difficult (including reading street names, menus, as well as small texts as found on product packaging/ medication/ phone books), in line with earlier studies [6,53,54]. While participants used a variety of assistive technologies to access written information, these often did not suffice. Many of the technologies used are either not portable (screen readers), or translate information into another modality before presenting it to the users with visual impairments of the technology (audiobooks, spoken TV subtitles). These technologies are often not usable outside home environments, which is where the most problematic reading activities seem to occur. Although more wearable devices (e.g. Orcam) and smartphone apps (e.g. KNFB reader, Be My Eyes and VizWiz [54]) do exist, it remained a struggle to access written information and texts.

2.4.4 Limitations

The study was aimed, designed and executed to provide an insight in the problems PVIs experience in their daily lives. Therefore, the participant sample is not big or diverse enough to apply statistical analyzes to form definitive conclusions on the priority list. Hence, no comparisons between the rankings of various demographic characteristics could be done.

2.4.5 Recommendations

The priority of the future research and development should be aimed at the previously discussed topics (social interactions and wayfinding). It is important to keep in mind that the involvement of prospective users in the research must not end after an agenda-setting study. Following the principles of user-centered design, prospective users must be involved in each iteration in the development of technology aids from the concept to the release on the market (or even after that) to be able to monitor and alter directions of the development so that it would fit the needs of PVIs.

2.5 References

1. World Health Organization (2012) Visual impairment and blindness 2010. World Heal Organ 96:1. DOI 10.1007/SpringerReference_61253
2. Szpiro S, Zhao Y, Azenkot S (2016) Finding a store, searching for a product. Proc 2016 ACM Int Jt Conf Pervasive Ubiquitous Comput - UbiComp '16, pp 61–72. DOI 10.1145/2971648.2971723
3. VISION2020 (2005) Vermijdbare blindheid en slechtziendheid in Nederland
4. Naraine MD, Lindsay PH (2011) Social inclusion of employees who are blind or low vision. Disabil Soc 26:389–403. DOI 10.1080/09687599.2011.567790
5. Schölvinck AFM, Pittens CACM, Broerse JEW (2017) The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 111:201–217
6. Broerse J, Scholvinck AF, Essink D, Pittens C (2015) Een onderzoeksagenda vanuit cliëntenperspectief. Amsterdam
7. Ye H, Malu M, Oh U, et al (2014) Current and future mobile and wearable device use by people with visual impairments. In: Proc. 32nd Annu. ACM Conf. Hum. factors Comput. Syst. - CHI '14. ACM Press, New York, New York, USA, pp 3123–3132
8. Hakobyan L, Lumsden J, O'Sullivan D, Bartlett H (2013) Mobile assistive technologies for the visually impaired. Surv Ophthalmol 58:513–528. DOI 10.1016/j.survophthal.2012.10.004
9. Bhowmick A, Hazarika SM (2017) An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends. J Multimodal User Interfaces 11:149–172. DOI 10.1007/s12193-016-0235-6
10. Grussenmeyer W, Folmer E (2017) Accessible touchscreen technology for people with visual impairments. ACM Trans Access Comput 9:1–31. DOI 10.1145/3022701
11. Csapó Á, Wersényi G, Nagy H, Stockman T (2015) A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research. J Multimodal User Interfaces 9:275–286. DOI 10.1007/s12193-015-0182-7

12. Tapu R, Mocanu B, Tapu E (2015) A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments. In: 2014 11th Int. Symp. Electron. Telecommun. ISETC 2014 - Conf. Proc. IEEE, pp 1–4
13. Sato D, Kitani K (2017) NavCog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In: Proc. 19th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS ’17. ACM Press, New York, New York, USA, Baltimore, pp 270-279
14. Cassidy CT, Tyler C (2017) Identifying visual cues to improve independent indoor navigation for blind individuals. In: Proc. 19th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS ’17. ACM Press, New York, New York, USA, pp 413–414
15. Guerreiro J, Ahmetovic D, Kitani KM, Asakawa C (2017) Virtual navigation for blind people: Building sequential representations of the real-world. In: Proc. 19th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS ’17. ACM Press, New York, New York, USA, pp 280–289. DOI 10.1145/3132525.3132545
16. Freeman E, Wilson G, Brewster S, et al (2017) Audible beacons and wearables in schools. In: Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI ’17. ACM Press, New York, New York, USA, pp 4146–4157
17. Kacorri H, Kitani KM, Bigham JP, Asakawa C (2017) People with visual impairment training personal object recognizers. In: Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI ’17. ACM Press, New York, New York, USA, pp 5839–5849
18. Sosa-García J, Odone F (2017) “Hands On” visual recognition for visually impaired users. ACM Trans Access Comput 10:1–30. DOI 10.1145/3060056
19. Królak A, Chen W, Sanderson NC, Kessel S (2017) The accessibility of MOOCs for blind learners. In: Proc. 19th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS ’17. ACM Press, New York, New York, USA, pp 401–402
20. Götzelmann T (2016) LucentMaps. In: Proc. 18th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS ’16. ACM Press, New York, New York, USA, pp 81–90

21. Ducasse J, Macé MJ-M, Serrano M, Jouffrais C (2016) Tangible reels. In: Proc. 2016 CHI Conf. Hum. Factors Comput. Syst. - CHI ’16. ACM Press, New York, New York, USA, pp 2186–2197
22. Avila Soto M, Funk M, Hoppe M, et al (2017) DroneNavigator: Using leashed and free-floating quadcopters to navigate visually impaired travelers. Proc 19th Int ACM SIGACCESS Conf Comput Access - ASSETS ’17, pp 300–304. DOI 10.1145/3132525.3132556
23. Szpiro S, Hashash S, Zhao Y, Azenkot S (2016) How people with low vision access computing devices. In: Proc. 18th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS ’16. ACM Press, New York, New York, USA, pp 171–180
24. Hamilton-Fletcher G, Obrist M, Watten P, et al (2016) “I Always Wanted to See the Night Sky”: Blind User Preferences for Sensory Substitution. In: Proc. 2016 CHI Conf. Hum. Factors Comput. Syst. - CHI ’16. ACM Press, New York, New York, USA, pp 2162–2174
25. Velázquez R (2010) Wearable Assistive Devices for the Blind. In: Lay-Ekuakille A, Mukhopadhyay SC (eds) Wearable Auton. Biomed. Devices Syst. Smart Environ. Springer, Berlin, Heidelberg, pp 331–349
26. Williams MA, Buehler E, Hurst A, Kane SK (2015) What not to wearable. In: Proc. 12th Web All Conf. - W4A ’15. ACM Press, New York, New York, USA, pp 1–4
27. Maus M, Belza B, Chi N, et al (2016) Wayfinding technologies for older adults with visual impairments: Ideas for future directions. Journal of Visual Impairment & Blindness 110(6):477-480
28. Gothwal VK, Wright TA, Lamoureux EL, et al (2009) Catquest questionnaire: Re-validation in an Australian cataract population. Clin Exp Ophthalmol 37:785–794. DOI 10.1111/j.1442-9071.2009.02133.x
29. Khadka J, Pesudovs K, McAlinden C, et al (2011) Reengineering the glaucoma Quality of Life-15 questionnaire with Rasch analysis. Investig Ophthalmol Vis Sci 52:6971–6977. DOI 10.1167/iovs.11-7423
30. Lundström M, Pesudovs K (2009) Catquest-9SF patient outcomes questionnaire. Nine-item short-form Rasch-scaled revision of the Catquest questionnaire. J Cataract Refract Surg 35:504–513. DOI 10.1016/j.jcrs.2008.11.038

31. Pesudovs K, Wright TA, Gothwal VK (2010) Visual disability assessment: valid measurement of activity limitation and mobility in cataract patients. Br J Ophthalmol 94:777–781. DOI 10.1136/bjo.2009.169490
32. Pesudovs K, Gothwal VK, Wright TA, Lamoureux EL (2010) Remediating serious flaws in the National Eye Institute Visual Function Questionnaire. J Cataract Refract Surg 36:718–732. DOI 10.1016/j.jcrs.2009.11.019
33. Gothwal VK, Wright TA, Lamoureux EL, et al (2009) Catquest questionnaire: Re-validation in an Australian cataract population. Clin Exp Ophthalmol 37:785–794. DOI 10.1111/j.1442-9071.2009.02133.x
34. Lamoureux EL, Pallant JF, Pesudovs K, et al (2008) Assessing participation in daily living and the effectiveness of rehabiliation in age related macular degeneration patients using the impact of vision impairment scale. Ophthalmic Epidemiol 15:105–113. DOI 10.1080/09286580701840354
35. Schmier JK, Halpern MT, Covert D (2006) Validation of the Daily Living Tasks Dependent on Vision (DLTV) questionnaire in a U.S. population with age-related macular degeneration. Ophthalmic Epidemiol 13:137–143. DOI 10.1080/09286580600573049
36. Stelmack JA, Massof RW (2007) Using the VA LV VFQ-48 and LV VFQ-20 in low vision rehabilitation. Optom Vis Sci 84:705–709. DOI 10.1097/OPX.0b013e3181339f1a
37. Stelmack JA, Szlyk JP, Stelmack TR, et al (2004) Psychometric Properties of the Veterans Affairs Low-Vision Visual Functioning Questionnaire. Investig Opthalmology Vis Sci 45:3919. DOI 10.1167/iovs.04-0208
38. Williams MA, Galbraith C, Kane SK, Hurst A (2014) "Just Let the Cane Hit It." In: Proc 16th Int ACM SIGACCESS Conf Comput Access - ASSETS '14 217–224. DOI 10.1145/2661334.2661380
39. Williams MA, Hurst A, Kane SK (2013) "Pray Before You Step out": Describing personal and situational blind navigation behaviors. In: Proc 15th Int ACM SIGACCESS Conf Comput Access. DOI 10.1145/2513383.2513449

40. Quiñones P, Greene T, Yang R, Newman M (2011) Supporting visually impaired navigation: A needs-finding study. Chi2011, pp 1645–1650. DOI 10.1145/1979742.1979822
41. Bala S, McDaniel T, Panchanathan S (2014) Visual-to-tactile mapping of facial movements for enriched social interactions. In: 2014 IEEE Int. Symp. Haptic, Audio Vis. Environ. Games Proc. IEEE, pp 82–87
42. Krishna S, Bala S, McDaniel T, et al (2010) VibroGlove: an assistive technology aid for conveying facial expressions. Proc SIGCHI Conf Hum Factors Comput Syst Ext Abstr, pp 3637–3642. DOI 10.1145/1753846.1754031
43. Panchanathan S, Chakraborty S, McDaniel T (2016) Social interaction assistant: a person-centered approach to enrich social interactions for individuals with visual impairments. IEEE J Sel Top Signal Process 10:942–951. DOI 10.1109/JSTSP.2016.2543681
44. Tanveer MI, Anam ASMI, Yeasin M, Khan M (2013) Do You See What I See?: Designing a sensory substitution device to access non-verbal modes of communication. In: Proc 15th Int ACM SIGACCESS Conf Comput Access. DOI 10.1145/2513383.2513438
45. Buimer HP, Bittner M, Kostelijk T, et al (2018) Conveying facial expressions to blind and visually impaired persons through a wearable vibrotactile device. PLoS One 13:e0194737. DOI 10.1371/journal.pone.0194737
46. Bruijning JE, Van Nispen RMA, Van Rens GHMB (2010) Feasibility of the Dutch ICF Activity Inventory: a pilot study. BMC Health Serv Res 10:318. DOI 10.1186/1472-6963-10-318
47. McAlinden C, Khadka J, De Freitas Santos Paranhos J, et al (2012) Psychometric properties of the NEI-RQL-42 questionnaire in keratoconus. Investig Ophthalmol Vis Sci 53:7370–7374. DOI 10.1167/iovs.12-9969
48. Loomis JM, Klatzky RL, Golledge RG, et al (1993) Nonvisual navigation by blind and sighted - Assessment of path integration ability. J Exp Psychol 122:73–91. DOI 10.1037/0096-3445.122.1.73
49. Banovic N, Franz RL, Truong KN, et al (2013) Uncovering information needs for independent spatial learning for users who are visually impaired. In: Proc 15th Int ACM SIGACCESS Conf Comput Access - ASSETS ’13 Article No. 24. DOI 10.1145/2513383.2513445

50. Krishna S, Colbry D, Black J, Balasubramanian V, Panchanathan S (2008) A systematic requirements analysis and development of an assistive device to enhance the social interaction of people who are blind or visually impaired. Workshop on Computer Vision Applications for the Visually Impaired, James Coughlan and Roberto Manduchi, Marseille
51. McDaniel TL, Krishna S, Colbry D, Panchanathan S (2009) Using tactile rhythm to convey interpersonal distances to individuals who are blind. Proc CHI 2009 Ext Abstr 4669–4674. DOI 10.1145/1520340.1520718
52. McDaniel T, Krishna S, Balasubramanian V, et al (2008) Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. In: HAVE 2008 - IEEE Int. Work. Haptic Audio Vis. Environ. Games Proc. pp 13–18
53. Strumillo P (2010) Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation. In: 3rd Int Conf Hum Syst Interact HSI'2010 - Conf Proc, pp 17–24. DOI 10.1109/HSI.2010.5514595
54. Brady E, Morris MR, Zhong Y, et al (2013) Visual challenges in the everyday lives of blind people. In: Proc SIGCHI Conf Hum Factors Comput Syst, pp 2117–2126. DOI 10.1145/2470654.2481291

3 Conveying Facial Expressions To Persons Who Are Blind Or Visually Impaired Through A Wearable Vibrotactile Device [3]

Abstract

In face-to-face social interactions, PVIs lack access to nonverbal cues like facial expressions, body posture, and gestures, which may lead to impaired interpersonal communication. In this study, a wearable sensory substitution device (SSD) consisting of a head mounted camera and a haptic belt was evaluated to determine whether vibrotactile cues around the waist could be used to convey facial expressions to users and whether such a device is desired by PVIs for use in daily living situations. Ten PVIs (mean age: 38.8, SD: 14.4) and 10 sighted persons (SPs) (mean age: 44.5, SD: 19.6) participated in the study, in which validated sets of pictures, silent videos, and videos with audio of facial expressions were presented to the participant. A control measurement was performed to determine how accurately participants could identify facial expressions while relying on their functional senses. After a short training, participants were asked to determine facial expressions while wearing the emotion feedback system. PVIs using the device improved their ability to determine which facial expressions were shown. A significant increase in accuracy of 44.4% was found across all types of stimuli when comparing the scores of the control (mean±SEM: 35.0±2.5%) and supported (mean±SEM: 79.4±2.1%) phases. The greatest improvements achieved with the support of the SSD were found for silent stimuli (68.3% for pictures and 50.8% for silent videos). SPs also showed consistent, though not statistically significant, improvements while supported. Overall, our study shows that vibrotactile cues are well suited to convey facial expressions to PVIs in real-time. Participants

---

[3] Buimer HP, Bittner M, Kostelijk, T, Van der Geest TM, Nemri A, Van Wezel RJA, Zhao, Y (2018) Conveying facial expressions to blind and visually impaired persons through a wearable vibrotactile device. PLoS ONE, 13(3), [e0194737]. DOI 10.1371/journal.pone.0194737

became skilled with the device after a short training session. Further testing and development of the SSD is required to improve its accuracy and aesthetics for potential daily use.

3.1 Introduction

A wide range of daily life activities cause major problems for PVIs, including wayfinding in unfamiliar surroundings, detecting objects and persons, and recognition of faces and facial expressions [1–4]. One such activity is face-to-face interaction: when they take place between two sighted persons (SPs), much information is exchanged nonverbally via body posture, gestures, interpersonal proximity and facial expressions. For example, facial expressions are believed to be closely related to one's emotions and provide information about the message one is trying to convey [5]. Because of their inability to fully access nonverbal information, PVIs who lost vision early in life can experience adverse effects on their social development, ultimately impacting their social inclusion as adults [3,6–8]. Despite demand from the PVI community [1,2], to our knowledge, there are only few assistive technologies available that attempt to support PVIs in accessing nonverbal communication in real time during social interactions.

In the absence of the ability to see, the human brain can learn to process information normally acquired through vision by using other senses, such as the auditory or haptic systems [9,10]. For PVIs, this means that information such as color, written information, nonverbal cues, or landmarks, can be obtained through auditory or haptic cues. The most well-known example is Braille, which is widely used amongst PVIs to interpret written information [9]. A study in the late 1960's showed that it was possible to convey visual information to PVIs using a haptic display built into the back of a chair, a so-called sensory substitution device (SSD) [11]. This system, which translated visual information directly to haptic patterns, enabled PVIs (after extensive training) to pick up objects. More recently, several other SSDs were presented that use audio or tactile tongue displays to convey visual information through another sense [12–16]. However, in the case of conveying information during social interactions, the use of audio or tactile tongue displays seem unsuited, for these concepts interfere with hearing and speech, which needs to be avoided during social interactions.

When it comes to conveying social information, it has been demonstrated that it is possible to convey spatial information (such as the location of and distance to other persons), walking directions, person identity, and social cues to PVIs through a vibrotactile belt, using variations in vibration location, frequency, and intensity [17–22]. Various studies have presented a tactile grid in the back of a chair to convey facial expressions (amongst others the Haptic Face Display (HFD)) [8,23,24]. While the HFD conveyed information with a high level of detail (the device used 48 tactors to display 15 vibration patterns), it was not mobile, as it required users to sit in the chair for it to be effective [8,23]. Furthermore, a vibrotactile glove was developed to convey Ekman's facial expressions of emotions plus neutral expressions through seven different vibrotactile patterns displayed on 14 tactors mounted on the back of the fingers [25]. In each of these studies, participants were quickly able to learn and interpret complex vibrotactile patterns conveyed. Both studies did focus on methods to convey information about facial expressions. However, these studies did not present a fully functional system that is capable of recognizing facial expressions and conveying these to its users in real time.

In this paper, a wearable SSD designed to support PVIs in determining the facial expressions of other persons is presented. The SSD classifies facial expressions into emotions, which are then conveyed using vibrotactile stimuli provided by a belt worn on the waist underneath clothing. Through user evaluations by PVIs and SPs, we sought to determine whether such a device could improve one's ability to determine the facial expressions of others and whether such a device is desired for use by PVIs.

3.2 Materials and Method

3.2.1 Participants

Twenty participants were included in the study including ten PVIs and ten SPs (see Table 1 for an overview of the participant characteristics). To maximize the number of potential users, we choose to include a group of PVIs (age: 38.8, SD: 14.4, range = 18–58) with a wide range of visual impairments who reported difficulties recognizing facial expressions and consisted of both early and late blind persons. The control group which was composed of SPs (age: 44.5, SD: 19.6, range = 20–68) who were each gender and age matched to one of the PVIs, creating two groups with reasonably similar compositions. Exclusion criteria included other cognitive or sensory impairments besides visual loss. The study was approved by the ethical committee of the Faculty of Electrical Engineering, Mathematics and Computer Science of the University of Twente and conducted in accordance with the guidelines of the Declaration of Helsinki. Data acquired from the study were only used after obtaining oral informed consent from the participant. Participants were told they could quit participation at any moment, without having to provide a reason for doing so. There were no drop-outs after informed consent was obtained.

Table 3.1. Overview of the PVI participant characteristics.

Visually impaired group			
Gender	Age	Description of Vision Loss	Time of Occurrence
Male	27	Fully blind	Late blind
Female	18	Fully blind	Early blind
Male	58	Light perception	Congenitally blind
Female	23	Central vision loss: Stargardt Disease	Late blind
Female	44	Left: Light perception; Right: Tunnel vision	Late blind
Male	58	Blurred vision, Severe near-sightedness	Late blind
Female	50	Familial Exudative Vitreoretinopathy	Congenitally blind
Male	27	Peripheral tunnel vision	Early blind
Female	43	Left: Blurred vision; Right: Light perception, Glaucoma	Congenitally blind
Female	40	Light perception, Retinitis Pigmentosa	Early blind

3.2.2 Apparatus

The SSD used in the study is shown in Figure 3.1. The various components of the device were controlled and linked via custom software on a Microsoft Surface Pro 4 tablet (6th Gen 2.2-GHz Intel Core i7-6650U processor with Intel Iris graphics 540, Windows 10 operating system). Users wore a Logitech HD Pro Webcam C920 mounted on a baseball-cap to record images in the gaze direction. The detection of faces and facial expression recognition from this live video stream was achieved using FaceReader 6™ (Vicar Vision, Amsterdam, The Netherlands). This software uses a robust real-time face detection algorithm to detect a face from the video stream [26] and an artificial deep neural network that can classify facial expressions into one of six basic emotions (anger, disgust, joy, fear, surprise, sadness) as well as a neutral facial expression [5,27].

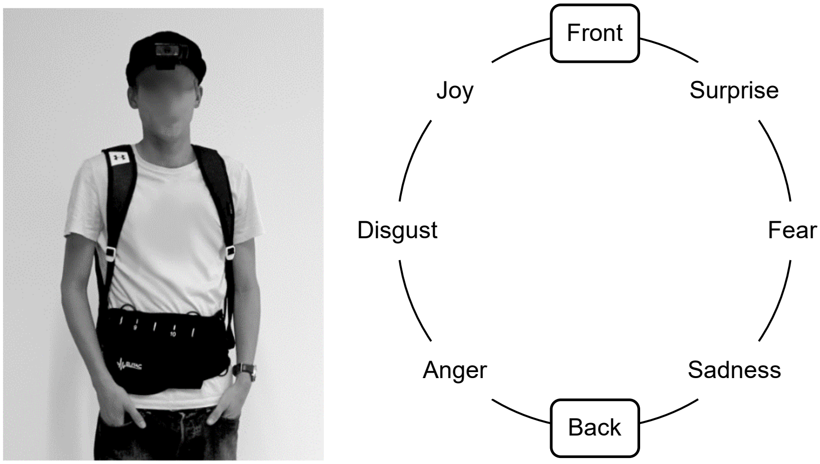


Figure 3.1. An overview of the sensory substitution worn in the study. Left: person wearing the device which consisted of a webcam mounted on a cap (1), a tablet in a mesh backpack (2), and an vibrotactile belt (3). Right: the six basic emotions and their placement on the vibrotactile belt worn around the waist of the user. More positive emotions were positioned in the front, whereas negative emotions were conveyed on the back.

The detected facial expressions were conveyed to the user by a series of vibrating motors (tactors) which were connected to the tablet via a Bluetooth connection. These tactors (3V pancake direct current unbalanced motors with a maximum rotational speed of 150 cycles/s and maximum vibration strength of 158.3 ± 2.4 Hz) were attached to a fabric belt with Velcro worn around the waist (Science Suit, Elitac, Utrecht, The Netherlands). The waist was chosen as it is not often used for social interactions, unlike for example the hands. Furthermore, the waist provides sufficient space to place multiple tactors (sized 34 × 16 × 11mm) at the spatial distance required to ensure that people could easily distinguish vibrations from different tactors [28,29]. Vibrotactile signals on the torso can be distinguished with an acuity of 2 to 3 cm [30]; an even lower acuity can be achieved on the back near the spine, where it is likely that distances lower than 1.3 cm are distinguishable [31]. The six tactors used in the current study were placed at least 4 cm apart, meaning the minimal distinguishable distances were amply complied with.

Each tactor was coupled to one of six basic emotions [5]. More positive emotions were positioned toward the front whereas negative emotions were positioned toward the back (see Figure 3.1 for tactor placement) in line with expressions of emotions ("butterflies in the stomach" or "stabbed in the back/talking behind one's back" [32]). The one to one association of each tactor to an emotion was purposely chosen to make the task of learning and interpreting the vibrations very easy for PVIs. As the ultimate goal was for PVIs to use such a system in daily life situations during which they may face other sensory information and/or use other assistive devices that require their attention, the system was designed to avoid unnecessary sensory and cognitive overload in real-life situations.

Upon detection of a face, the user was alerted with two 150ms vibrations on all tactors with a 50ms break in between. After another 200ms, the tactor associated with the recognized facial expression vibrated as long as the expression held. This feedback was only provided if the facial expression detected deviated from the neutral expression. A long 300ms vibration on all tactors was used to indicate when the software no longer detected a face.

3.2.3 Materials

Three types of materials were used as test stimuli to determine how accurately persons could identify facial expressions: pictures, silent videos, and videos with audio. The pictures and videos were derived from validated sets of pictures from the Warsaw Set of Emotional Facial Expression Pictures (WSEFEP) [33] and videos from the Amsterdam Dynamic Facial Expression Set (ADFES) [34], and included facial expressions of joy, surprise, fear, sadness, anger, and disgust [5]. Audio-visual stimuli were created by combining silent videos from the ADFES with (very obvious) annotated non-linguistic affect bursts (i.e. short bursts of sounds persons made while expressing an emotion) from other validated sets [35,36]. Although the set of Hawk and colleagues [35] did contain both audio and video files (not combined), video files from the ADFES were used due to their better image quality. Audio and video files were matched based on the annotated emotion and its intensity and combined using video editing software. All syncing was done manually to ensure the beginning of the facial expression and the affect burst matched.

To see how FaceReader performed when it was confronted with stimuli that were not directly loaded into the software but which were subject to head movement and lighting conditions, only stimuli were used for which it was known that the software could detect the correct emotion. Therefore, prior to the experiments, all visual stimuli to be used in the experiment were loaded directly into FaceReader to verify that the software could determine the correct facial expressions in each of the stimuli under optimal conditions. Stimuli that were not correctly detected were excluded from the study. For comparison, earlier studies showed that the accuracy of SPs in determining facial expressions from these sets was 87% for the ADFES and 82% for the WSEFEP. Similarly, FaceReader achieved an accuracy of 88% and 89% for these sets, respectively [37].

3.2.4 Experimental design

The experiment was divided into three phases (an unsupported control phase, a training phase, and a supported phase) (see Figure 3.2). The visual stimuli were projected on a wall, two meters in front of the participant (see Figure 3.3), while audio was played at a volume that all participants could clearly hear from the speakers of a laptop placed right behind the participant. The size of the projected face was slightly bigger than a normal face would be at a two-meter distance to create face sizes like those encountered in normal social interactions.

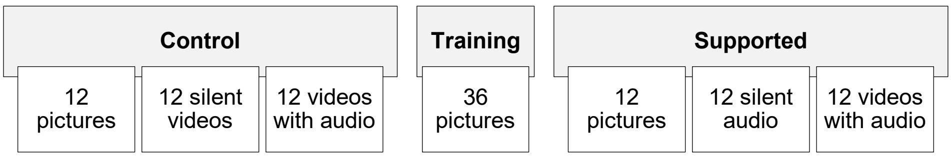


Figure 3.2. Study design. The experiment was divided into a control, training, and supported phase, each with 36 stimuli consisting of pictures, silent videos, and videos with audio.

Figure 3.3. Experimental setup. The participant was positioned in front of a projector, which projected stimuli on a wall two meters away.

During the control and supported phases, 12 pictures, 12 silent videos, and 12 videos with audio were presented to the participant. The order in which stimuli were presented within each set of 12 was randomized beforehand to ensure that participants could not guess which emotion was presented next (the order of stimuli did not vary between participants). To avoid order bias, the stimuli sets were presented in reverse order for half of the participants (5 PVIs and 5 SPs). The training session included only pictures, the first 12 of which were presented in an order of emotion, whereas the remaining 24 were presented in a random order.

The control phase was used to ascertain how accurately subjects could identify the emotions displayed in the stimulus sets whilst relying only on their functional senses. A short beep was used to indicate when a stimulus was about to be presented, after which each stimulus was displayed for six seconds. Following each stimulus, the participants were instructed to indicate which emotion was expressed. The participants were made aware that no new stimulus would be displayed until they finished giving their answer. If a PVI was unable to detect the first three stimuli of a set, the session continued to the next set of stimuli.

During the training phase, which lasted for about 20 minutes to half-an-hour, participants were introduced to the SSD and received instructions on how to interpret the vibrotactile cues. After measuring waist circumference to ensure correct tactor spacing and placement, the minimum perceivable and maximum comfortable vibration strengths of each user were determined and the upper and lower boundaries of the tactor vibrations were programmed accordingly. Participants were then instructed which emotion was assigned to each tactor location. To familiarize participants with the device, three sets of 12 pictures were shown while they received the corresponding tactile cues on the belt. SPs were asked to close their eyes during training to ensure attention was directed to the vibrotactile cues. During the first 12 pictures, the participants were told which emotions were conveyed by the belt. For the second set of pictures, answers given by the participants were either confirmed or corrected by the examiner. For the final 12 pictures in the training set, participants practiced completely without receiving feedback.

In the last phase of the experiment, trained participants were supported by the device and asked to identify the emotions from a stimulus set consisting of 12 pictures, silent videos, and videos with audio. In addition to the questions asked during the control measurements, participants were also asked to report the location of the vibrating tactor for each stimulus.

3.2.5 Data analysis

Trials with correctly identified emotions were scored with a 1, whereas incorrect answers were scored with 0. For each measurement, mean performance scores were calculated. The performances of both the participant and FaceReader were rated in this way. For the user performance, it did not matter whether mistakes were made due to wrong interpretation of the vibrotactile cue by the respondent or to a misclassification by the FaceReader software causing the device to convey the wrong vibrotactile cue. The between-subject effects of group (SP or PVI), and the within-subject effects of phase (control—no SSD, supported—with SSD) and stimuli (pictures, silent videos, video with audio), were analyzed using repeated measures ANOVA with an alpha of 0.05 in IBM SPSS Statistics 22. A similar analysis was also performed to analyze how participants performed when the FaceReader software misidentified the emotion shown. In this case, the between-subject effects of group (SP or PVI), and the within-subject effects of stimuli (pictures, silent videos, videos with audio) and FaceReader accuracy (wrong, correct) was analyzed. Because data for some conditions were slightly skewed towards a performance 100%, the assumption of normality was violated, which should be considered when interpreting the results. Furthermore, Mauchly's test was used to test the assumption of sphericity. If the assumption of sphericity was violated, the degrees of freedom were corrected using Greenhouse-Geisser or Huynh-Feldt corrections. Post hoc comparisons were performed using Bonferroni adjustments. Finally, the performance of the participants was compared to that of FaceReader to determine the extent the SSD contributed to the performance of the participants. Unless otherwise stated, descriptive statistics are represented by the mean±standard error of the mean.

3.3 Results

An overview of the performances for both participant groups under the different experimental phases can be found in Figure 3.4 and Table 3.2. Overall, higher performance levels were achieved with the support of the SSD compared to control for both SPs and PVIs across all types of stimuli. Statistical analysis showed significant main effects of phase, stimuli, and group on performance. Significant two-way interactions were found between phase and group, stimuli and group, and phase and stimuli. A significant three-way interaction also existed between phase, stimuli, and group.

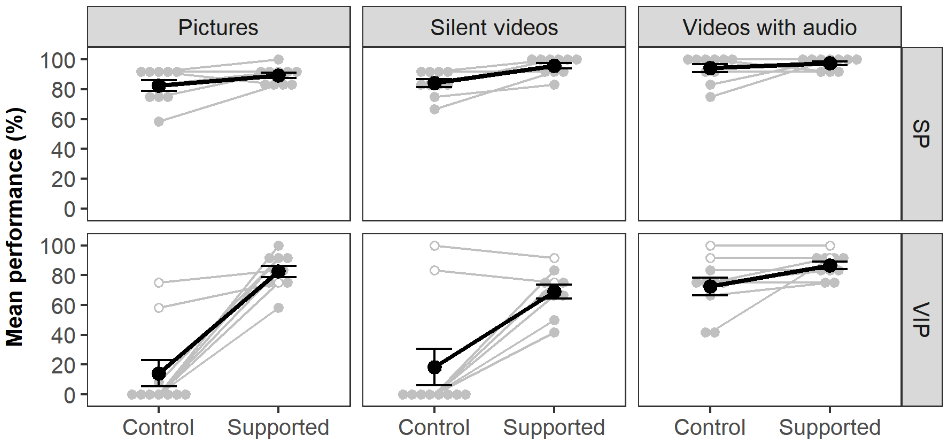


Figure 3.4. Scores for each phase, stimuli, and group. Black dots and lines are associated with the mean score of the subgroups whereas grey dots and lines are associated with individual participants. Grey dots with white filling were visually impaired participants who had sufficient remaining vision to detect stimuli unsupported. Error bars show standard error. SP: sighted persons, PVI: visually impaired persons.

The SSD had a significant effect on performance ($F(1,18) = 39.59$, $p < .001$) for all participants combined: average performance scores differed significantly (p

< 0.001) between the supported phase (86.8±1.3%) and the control phase (61.0±1.8%). Note that Mauchly's sphericity test was not applied for this within-subject factor, as there were only two levels (unsupported and supported).

Table 3.2. Overview of the mean performance across different phases, stimuli, and groups. The table presents the mean performance and standard error of the mean across different phases and stimuli for SPs and PVIs separately (top 2 sections) as displayed in Figure 3.4 and that for all participants combined (bottom section). In addition, the table also indicates the differences between the mean performances of the control and supported phases.

Group	Stimuli	Phase				
		Control		Supported		Difference
		Mean	SEM[a]	Mean	SEM[a]	
PVI (N = 10)	All stimuli	35.0	2.5	79.4	2.1	44.4 *[b]
	Pictures	14.2	3.2	82.5	3.5	68.3 *** [b]
	Silent videos	18.3	3.5	69.2	4.2	50.8 *** [b]
	Videos audio	72.5	4.1	86.7	3.1	14.2 * [b]
SP (N = 10)	All stimuli	86.9	1.8	94.2	1.2	7.2
	Pictures	82.5	3.5	89.2	2.8	6.7
	Silent videos	84.2	3.3	95.8	1.8	11.7
	Videos audio	94.2	2.1	97.5	1.4	3.3
All participants (N = 20)	All stimuli	61.0	1.8	86.8	1.3	25.8 * [b]
	Pictures	48.3	3.2	85.8	2.3	37.5 * [b]
	Silent videos	51.3	3.2	82.5	2.5	31.3 * [b]
	Videos audio	83.3	2.4	92.1	1.8	8.7 *** [b]

[a] *SEM = Standard error of the mean*
[b] **** p < .001, * p < .05.*

In addition, there was an significant interaction effect for condition and group ($F(1.18) = 20.55, p < .001$): whereas the mean score of PVIs increased significantly ($p < 0.05$) from 35.0±2.5% during control to 79.4±2.1% when supported; SPs also achieved an improvement (from 86.9±1.8% to 94.2±1.2%), which was not a statistically significant difference. These results suggested that participants were more capable of identifying facial emotions whilst supported by the SSD.

3.3.1 Effect of sightedness on performance

The between-subject effects of group (SP or PVI) had a significant effect on performance ($F (1,18) = 42.311, p < 0.001$). The SPs were overall significantly better ($p < 0.001$) in detecting facial expressions than their PVI counterparts: the average performance across both phases was 90.6±1.1% for SPs and 57.2±1.8% for PVIs. Without the support of the SSD in the control phase, eight of the 10 PVIs could not identify emotions at all, while two PVIs were able to use their remaining vision to achieve performance scores above chance level. However, this sample size is too small and diverse (one person had tunnel vision in one eye and light perception in the other, whereas the other had only peripheral vision due to Stargardt disease) to conduct separate statistical analysis. The differences in accuracy between PVIs and SPs were much bigger in the control phase (35.0±2.5% for PVIs vs 86.9±1.8% for SPs) than in the supported phase (79.4±2.1% for PVIs vs 94.2±1.2% for SPs). In fact, the performance of PVIs when supported by the SSD reached a level that was not significantly different from those of SPs in the control phase ($t(18) = 2.061$, $p = 0.054$). Altogether these findings emphasize the beneficial effects of using the SSD for PVIs in recognizing emotions.

3.3.2 Effect of stimulus type on performance

For the within-subject effects of stimulus, Mauchly's test showed a violation of the assumption of sphericity ($\chi^2(2) = 6.044$, $p < .05$). Therefore, Greenhouse-Geisser corrections were applied ($\varepsilon = .77$). The type of stimulus presented to the participant (picture, silent videos, or videos with audio) had a significant effect on performance ($F(1.54,27.71) = 50.259$, $p < 0.001$). Post hoc tests showed that performance for videos with added audio (87.7±1.5% was significantly higher than for pictures (67.1±2.1%) and silent videos (66.9±2.2%). No significant performance difference was found between pictures and silent videos. Thus, participants found it easiest to identify emotions when additional auditory cues were provided.

For the two-way interaction between the type of stimuli and phase, Mauchly's test showed that the assumption of sphericity was violated ($\chi^2(2) = 7.817$, $p < .05$) and Greenhouse-Geisser corrections were applied ($\varepsilon = .731$). A significant two-way interaction was found for phase and stimuli ($F(1.46,26.30) = 23.05$, $p < 0.001$). Furthermore, a significant three-way interaction effect between the phase, type of stimulus, and participant group was found ($F(1.46, 26.30) = 16.35$, $p < 0.001$). For SPs, there were no significant performance differences between control and supported phase for each type of stimulus. In contrast, PVIs showed significant performance improvements for pictures (68.3% increase, $p < 0.001$), silent videos (50.8% increase, $p < 0.001$), and videos with added audio (14.2% increase, $p < 0.05$) compared to control. In the control phase, PVIs were significantly better at determining the expressed emotions from videos with audio than those from other types of stimuli ($p < 0.001$). With the support of the SSD, the mean performance difference between pictures and videos with audio was no longer statistically significant. The difference between silent videos and videos with audio remained significant, possibly due to the inherent performance of FaceReader (see below). These results suggest that vibrotactile cues could enhance the recognition of facial expressions, especially in the absence of auditory cues.

3.3.3 Interpretation of the vibrotactile signals

The accuracy of FaceReader was a limiting factor in the performance improvements of the PVIs. Overall, the software achieved an average accuracy of 73.6±1.6% in classifying the facial expressions from the experimental stimuli. SPs outperformed FaceReader overall, whereas PVIs performed only as well as FaceReader in absence of auditory cues whilst outperforming FaceReader (65% vs. 86.7%) when auditory cues were also provided (see Figure 3.5).

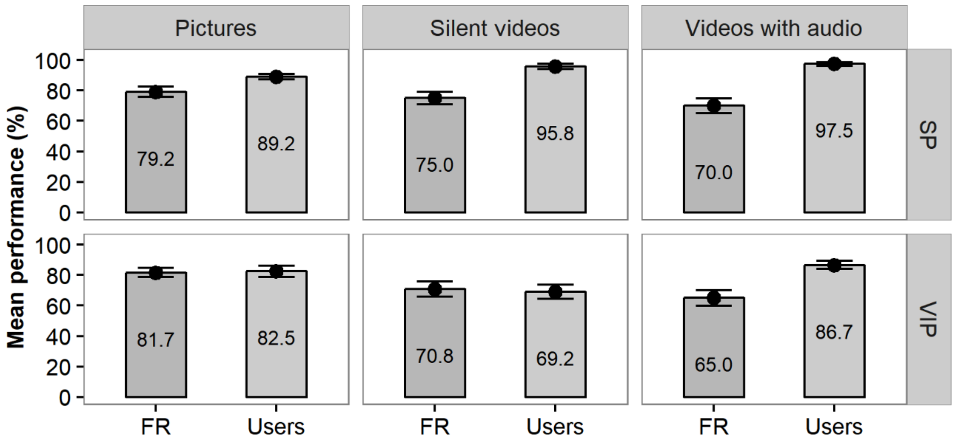


Figure 3.5. FaceReader (FR) accuracy versus user performance in the supported phase. This graph shows the mean performance in the supported phase of the participants (right, also shown in Fig 3.4 and Table 3.2) compared to FaceReader (left). Error bars represent standard error.

To examine how participants dealt with inaccuracies in FaceReader, an additional analysis was conducted to compare how participants performed when FaceReader was correct with how they performed when FaceReader was incorrect (see Figure 3.6 and Table 3.3). Mauchly's test of sphericity showed that the assumption of sphericity was met for stimuli ($\chi^2(2) = 1.275$, $p < .528$) and the interaction between stimuli and FaceReader ($\chi^2(2) = 4.389$, $p < .111$.

Similar to before, the main effects of group ($F (1,16) = 26.474, p < .001$), stimuli ($F (2,32) = 7.921, p < .01$) and the two-way interaction between stimuli and group ($F (2,32) = 3.752, p < .05$ were significant. Notably, the main effect of FR ($F (2,32) = 25.456, p < .001$) and the interaction effect between FaceReader and group ($F (1,16) = 14.946, p < .002$ were also significant. No significant interaction effects were found for the two-way interaction between stimuli and FaceReader ($F (2,32) = 3.194$, ns) and the three-way interaction of stimuli, FR, and group ($F (2,32) = 2.016$, ns).

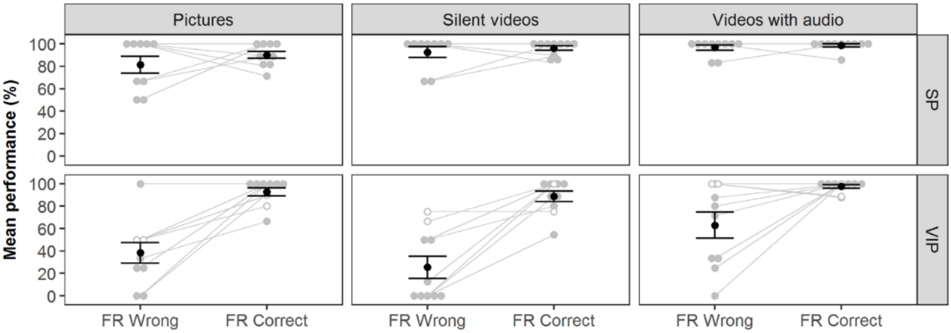


Figure 3.6. Performance of participants in recognizing facial expressions in relation to FaceReader (FR) accuracy. Performance of participants in recognizing facial expressions in relation to FaceReader accuracy. The average accuracy of SPs and PVIs when FaceReader correctly identified the facial expression compared to the performance if FaceReader misidentified the facial expression. Error bars represent the standard error. Grey dots with no filling correspond to PVIs who had sufficient remaining vision to detect stimuli unsupported.

The performance of SPs did not depend on the success of FaceReader, which suggested that SPs were able to correct for the mistakes of FaceReader using visual and auditory cues. PVIs, however, performed significantly worse across all stimulus types when FaceReader conveyed the incorrect emotion. While unsurprising for stimuli lacking audio (pictures and silent videos), PVIs also performed significantly worse for videos with audio (66.7±7.4 vs 72.5±4.1). Although their performance for videos with audio exceeded that for pictures and silent videos, auditory cues were insufficient for PVIs to fully correct for FaceReader mistakes. PVIs seemed to be unable to correct for mistakes by the system using the auditory stimuli. Nevertheless, the performance of PVIs remained higher when the system was used in combination with auditory cues (86.7±3.1), because of the near perfect performance in cases where FaceReader was correct (97.4±1.8).

Table 3.3. Performance of participants in relation to FaceReader accuracy.

Group	Stimuli	FR	N	Performance Mean	SEM[a]	Mean difference
Sighted	Pictures	Wrong	25	84.00	7.48	14.2
		Correct	95	90.53	3.02	
	Silent videos	Wrong	30	93.33	4.63	3.6
		Correct	90	96.67	1.90	
	Videos with audio	Wrong	36	94.44	3.87	2.4
		Correct	84	98.81	1.19	
PVI	Pictures	Wrong	22	36.36	10.50	54.4 ***[b]
		Correct	98	92.86	2.61	
	Silent videos	Wrong	35	25.71	7.50	63.3 ***[b]
		Correct	85	87.06	3.66	
	Videos with audio	Wrong	42	66.67	7.36	34.6 *[b]
		Correct	78	97.44	1.80	

The table shows the mean and difference in mean participant performance, when FaceReader correctly and incorrectly detect the emotion shown. Furthermore, the table shows the number of times FaceReader was wrong/correct for each condition (N).
[a] *SEM = Standard error of the mean*
[b] **** $p < .001$, * $p < .05$.*

3.4 Discussion

The objective of the study was to determine the feasibility of using a wearable device to convey facial expressions of emotions through vibrotactile feedback. By combining various existing technologies, we developed a wearable SSD that conveys facial expressions to its users in real time through a vibrotactile belt. This study showed that participants could easily distinguish and interpret vibrotactile stimulation associated with the six basic emotions in real time. In fact, PVIs significantly improved their ability to determine facial expressions while wearing the SSD for all types of stimuli, reaching an overall accuracy of 79.4%. As participants were still able to use their senses of hearing and sight, if any, in determining the facial expressions, the SSD also did not interfere with other sensory modalities. Thus, our study confirms the conclusions of previous studies that haptic cues can be a beneficial tool for conveying visual information [8,18].

In line with earlier studies using the same annotated sets [33,34,37], SPs reached a performance mean of 86.9% without the help of the SSD. In the control phase, PVIs were able to determine the facial expression in 72.5% of the videos with audio, which is in line with the expected performance for the affect bursts [36]. Previous studies showed that the FaceReader software can recognize facial expressions from annotated sets of stimuli with an accuracy close to 90% [27,37]. In our study, the software reached lower accuracy averaging 73.6% (range: 65%-81.7%). This discrepancy may be explained by the fact that the stimuli presented in our study were not directly loaded into FaceReader. Instead, these were fed from a live video stream, and therefore subject to head movements, changing focal length, and changes in lighting and luminance.

The results of the studies are consistent with the general principles of multisensory integration. According to the Bayesian view on multimodal cue integration, perception is probabilistic and in order to form a coherent percept of the world, cues from different sensory modalities are combined in such a way as to favor the most reliable (or least uncertain) cues [38]. With limited sight, PVIs therefore generally rely on auditory and haptic cues (e.g. text-to-speech and braille). As shown in the study, PVIs achieved a high degree of accuracy in trials with auditory stimuli, even without additional haptic cues, since the auditory cues used in the study were very unambiguous and easy to interpret. The performance of PVIs to detect the correct emotion from auditory stimuli significantly improved as soon as (the even more unambiguous) haptic cues were added.

This improvement of performance was highly dependent on the accuracy of the software, as inaccuracies of the software were hardly corrected for by the participants with visual impairments. Even when auditory stimuli were presented, mistakes by the software led to a significant decrease in performance, resulting in a performance that was lower than performance in auditory stimuli only. Overall, however, the accuracy of the software was sufficient to improve the performance of PVIs for all types of stimuli, including auditory.

Furthermore, it is important to take into account that during social interactions in real-life conditions, emotions are often conveyed without auditory cues (e.g. smiling, frowning). In such cases, PVIs are forced to rely on the haptic cues conveyed by the vibrotactile device. It is promising that the performance of PVIs for videos with audio without haptic feedback was comparable to that for silent videos with the support of the vibrotactile belt, meaning that in the absence of auditory cues their ability to detect the correct emotion did not decrease. Moreover, the performance with both auditory and haptic feedback was higher than that with either sensory modality alone.

In contrast with previous SSD studies [8,11–16,21,22], the system presented here is fully wearable, does not interfere with other sensorimotor functions used in social interactions, namely touch of hand, speech, and hearing, and provides simple cues to represent a basic set of emotions to avoid cognitive overload in more realistic usage situations [39]. Furthermore, the full working prototype was tested in real time with PVIs, showing that the system could process live visual input and convey facial expressions of emotions in an easily interpretable fashion.

The responses to the device were generally positive amongst the PVIs. PVIs described various scenarios in which the device could be beneficial including face-to-face and group meetings (in line with [21]) and were willing to try the device in such settings. Nevertheless, the participants stated that the device required some alterations before they would use it over an extended period. Participants namely had reservations about the weight and fit of the cap-mounted camera and were concerned that the SSD would bring unwanted attention to their impairment. Finally, to make it more worthwhile for PVIs to wear the device for the entire day, the participants desired additional features, such as those within the domain of social interactions or beyond, such as outdoor navigation or the access to public transport information [1,2,4,40].

3.4.1 Limitations of the study

One could argue that it is perhaps unsurprising that PVIs were able to learn how to use the SSD within a short training period and achieve significant performance improvements, considering results from earlier studies [30,31]. Indeed, only six tactors placed at least four centimeters apart were used on the waist, while the spatial acuity of the torso is between 2 to 3 cm [30]. Moreover, participants were only required to learn one to one associations between six tactors and emotions, while earlier research has shown that persons are able to learn far more complex (vibrotactile) cues [8,9,11,12,15]. Nevertheless, the ultimate goal was to create a system that PVIs could easily use in real-life situations. According to [41,42], the ability to process tactile information is likely to be significantly worse in real-world conditions where other sensory inputs are competing for attention. The simplicity of the system and the sensory mapping was therefore intentional lest the vibrotactile cueing becomes unnecessarily difficult in daily life.

Another drawback of the study is its potential lack of generalizability to real-life situations, a concern that was also raised in [41,42]. In the experimental setup, the stimuli were presented on a fixed position and participants were instructed as to where the stimuli were presented. In real-life, users would have to localize and aim the camera towards the targeted person on their own. The lighting and gaze direction of the conversation partner were also stable in the experimental setup. In a physical conversation lighting and gaze direction would change continuously, thus impacting the quality of the analysis of facial expressions. Furthermore, it is important to thoroughly determine how well device users can interpret the vibrotactile cues in real-life situations, where other sensory stimuli might compete for attention and cognitive overload might become an issue. Second, there was a purposeful 550 ms delay between the displayed stimulus and the vibrotactile cues associated to the displayed facial expression to warn users that a face was recognized. This caused for the fact that audio and visual cues were often interpreted before vibrotactile information was conveyed. In such cases, the vibrotactile cues were merely used to confirm or adjust already made decisions and caused more confusion than clarity.

3.4.2 Conclusions

This study showed that a SSD like the one presented, using vibrotactile cues at the waist, was a feasible method to convey information about facial expressions to PVIs, which may lead to improvements in their social interactions [2,18]. Participants were quickly able to learn how to interpret the cues conveyed by the device and combined this with information acquired from other functional senses. Furthermore, PVIs saw potential use of the device in real situations. Nevertheless, for the device to be readily adopted and accepted by PVIs as a daily life assistive technology, a more aesthetically pleasing design is required (e.g. smaller camera, less weight, unobtrusiveness) and more usage goals should be addressed (e.g. navigation). Finally, studies that are more closely resembling realism are needed to determine the accuracy of the device in real-life situations and user acceptance of the technology over time.

3.5 References

1. Van der Geest TM, Buimer HP (2015) User-centered priority setting for accessible devices and applications. In: Weisbecker A, Burmester M, Schmidt A (eds) Mensch und Comput. 2015 – Work. Oldenbourg Wissenschaftsverlag, Stuttgart, pp 383–389
2. Krishna S, Colbry D, Black J, Balasubramanian V, Panchanathan S (2008) A systematic requirements analysis and development of an assistive device to enhance the social interaction of people who are blind or visually impaired. Workshop on Computer Vision Applications for the Visually Impaired, James Coughlan and Roberto Manduchi, Marseille
3. Naraine MD, Lindsay PH (2011) Social inclusion of employees who are blind or low vision. Disabil Soc 26:389–403. DOI 10.1080/09687599.2011.567790
4. Broerse J, Scholvinck AF, Essink D, Pittens C (2015) Een onderzoeksagenda vanuit cliëntenperspectief. Amsterdam
5. Ekman P (1970) Universal facial expressions of emotions. Calif Ment Heal Res Dig Dig 8:151–158.
6. Papadopoulos K, Metsiou K, Agaliotis I (2011) Adaptive behavior of children and adolescents with visual impairments. Res Dev Disabil 32:1086–1096. DOI 10.1016/j.ridd.2011.01.021
7. Fischer ME, Cruickshanks KJ, Klein BEK, et al (2009) Multiple sensory impairment and quality of life. Ophthalmic Epidemiol 16:346–353. DOI 10.3109/09286580903312236
8. Bala S, McDaniel T, Panchanathan S (2014) Visual-to-tactile mapping of facial movements for enriched social interactions. In: 2014 IEEE Int. Symp. Haptic, Audio Vis. Environ. Games Proc. IEEE, pp 82–87
9. Bach-y-Rita P, Kercel SW (2003) Sensory substitution and the human-machine interface. Trends Cogn Sci. DOI 10.1016/j.tics.2003.10.013
10. Jóhannesson Ó, Balan O, Unnthorsson R, et al (2016) The sound of vision project: On the feasibility of an audio-haptic representation of the environment, for the visually impaired. Brain Sci 6:20. DOI 10.3390/brainsci6030020
11. Bach-y-Rita P (1967) Sensory plasticity: applications to a vision substitution system. Acta Neurol Scand 43:417–426. DOI 10.1111/j.1600-0404.1967.tb05747.x

12. Meijer PBL (1992) An experimental system for auditory image representations. IEEE Trans Biomed Eng 39:112–121. DOI 10.1109/10.121642
13. Hanassy S, Maidenbaum S, Tauber D, et al (2013) EyeMusic: A colorful experience for the blind. Multisens Res 26:116–116. DOI 10.1163/22134808-000S0084
14. Abboud S, Hanassy S, Levy-Tzedek S, et al (2014) EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution. Restor Neurol Neurosci 32:247–257. DOI 10.3233/RNN-130338
15. Kaczmarek KA (2011) The tongue display unit (TDU) for electrotactile spatiotemporal pattern presentation. Sci Iran 18:1476–1485. DOI 10.1016/j.scient.2011.08.020
16. Danilov Y, Tyler M (2005) Brainport: An alternative input to the brain. J Integr Neurosci 04:537–550. DOI 10.1142/S0219635205000914
17. McDaniel TL, Villanueva D, Krishna S, et al (2010) Heartbeats: a methodology to convey interpersonal distance through touch. In: Proc. 28th Int. Conf. Ext. Abstr. Hum. factors Comput. Syst. - CHI EA '10, pp 3985-3990
18. McDaniel T, Krishna S, Balasubramanian V, et al (2008) Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. In: HAVE 2008 - IEEE Int. Work. Haptic Audio Vis. Environ. Games Proc., pp 13–18
19. Kärcher SM, Fenzlaff S, Hartmann D, et al (2012) Sensory augmentation for the blind. Front Hum Neurosci 6:37. DOI 10.3389/fnhum.2012.00037
20. Krishna S, Little G, Black J, Panchanathan S (2005) A wearable face recognition system for individuals with visual impairments. In: Proc 7th Int ACM SIGACCESS Conf Comput Access - Assets '05, pp 106. DOI 10.1145/1090785.1090806
21. Panchanathan S, Chakraborty S, McDaniel T (2016) Social interaction assistant: a person-centered approach to enrich social interactions for individuals with visual impairments. IEEE J Sel Top Signal Process 10:942–951. DOI 10.1109/JSTSP.2016.2543681

22. McDaniel TL, Krishna S, Colbry D, Panchanathan S (2009) Using tactile rhythm to convey interpersonal distances to individuals who are blind. Proc CHI 2009 Ext Abstr, pp 4669–4674. DOI 10.1145/1520340.1520718
23. McDaniel T, Bala S, Rosenthal J, et al (2014) Affective haptics for enhancing access to social interactions for individuals who are blind. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), pp 419–429
24. Ur Réhman S, Liu L (2008) Vibrotactile rendering of human emotions on the manifold of facial expressions. J Multimed 3:18–25. DOI 10.4304/jmm.3.3.18-25
25. Krishna S, Bala S, McDaniel T, et al (2010) VibroGlove: an assistive technology aid for conveying facial expressions. Proc SIGCHI Conf Hum Factors Comput Syst Ext Abstr, pp 3637–3642. DOI 10.1145/1753846.1754031
26. Viola P, Jones MMJ (2004) Robust real-time face detection. Int J Comput Vis 57:137–154. DOI 10.1023/B:VISI.0000013087.49260.fb
27. Den Uyl MJ, Van Kuilenburg H, Lebert E (2005) The FaceReader: Online facial expression recognition. In: Noldus LPJJ, Grieco F, Loijens LWS, Zimmerman PH (eds) Proc. Meas. Behav. Wageningen, pp 589–590
28. Van Erp JBF (2002) Guidelines for the use of vibro-tactile displays in human computer interaction. Proc Eurohaptics, pp 18–22. DOI 10.1016/j.ajodo.2008.04.017
29. Wang Y, Millet B, Smith JL (2016) Designing wearable vibrotactile notifications for information communication. Int J Hum Comput Stud 89:24–34. DOI 10.1016/j.ijhcs.2016.01.004
30. Van Erp JBF (2005) Vibrotactile spatial acuity on the torso: Effects of location and timing parameters. In: First Jt. Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst. IEEE, pp 80–85
31. Jóhannesson ÓI, Hoffmann R, Valgeirsdóttir VV, et al (2017) Relative vibrotactile spatial acuity of the torso. Exp Brain Res 235:1–11. DOI 10.1007/s00221-017-5073-6
32. Nummenmaa L, Glerean E, Hari R, Hietanen JK (2014) Bodily maps of emotions. Proc Natl Acad Sci U S A 111:646–51. DOI 10.1073/pnas.1321664111

33. Olszanowski M, Pochwatko G, Kuklinski K, et al (2015) Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. Front Psychol 5:1–8. DOI 10.3389/fpsyg.2014.01516
34. Van der Schalk J, Hawk ST, Fischer AH, Doosje B (2011) Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). Emotion 11:907–920. DOI 10.1037/a0023853
35. Hawk ST, Van Kleef GA, Fischer AH, Van der Schalk J (2009) "Worth a thousand words": absolute and relative decoding of nonlinguistic affect vocalizations. Emotion 9:293–305. DOI 10.1037/a0015178
36. Lima CF, Castro SL, Scott SK (2013) When voices get emotional: a corpus of nonverbal vocalizations for research on emotion processing. Behav Res Methods 45:1234–45. DOI 10.3758/s13428-013-0324-3
37. Lewinski P, Den Uyl TM, Butler C (2014) Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. J Neurosci Psychol Econ 7:227–236. DOI 10.1037/npe0000028
38. Ernst M (2006) A Bayesian view on multimodal cue integration. In: Hum. Body Percept. From Insid. Out. Oxford University Press, New York, New York, USA, pp 105–131
39. Elli GV, Benetti S, Collignon O (2014) Is there a future for sensory substitution outside academic laboratories? Multisens Res 27:271–291. DOI 10.1163/22134808-00002460
40. Hakobyan L, Lumsden J, O'Sullivan D, Bartlett H (2013) Mobile assistive technologies for the visually impaired. Surv Ophthalmol 58:513–528. DOI 10.1016/j.survophthal.2012.10.004
41. Spence C (2014) The skin as a medium for sensory substitution. Multisens Res 27:293–312. DOI 10.1163/22134808-00002452
42. Jones LA, Sarter NB (2008) Tactile displays: Guidance for their design and application. Hum Factors J Hum Factors Ergon Soc 50:90–111. DOI 10.1518/001872008X250638

4 Opportunities And Pitfalls In Applying Emotion Recognition Software For Persons Who Are Blind Or Visually Impaired: Simulated Real Life Conversations [4]

Abstract

Many communication cues exchanged between persons are nonverbal. PVIs are often unable to perceive these cues such as facial expressions of emotions. Earlier, we have determined that PVIs are able to recognize facial expressions of emotions from validated pictures and videos by using emotion recognition software and vibrotactile signals associated to six basic emotions. In this study, we determined whether the previously tested emotion recognition system worked as well in realistic situations as it did under controlled lab conditions. The system consists of a camera mounted on spectacles, a tablet running facial emotion recognition software, and a waist belt with vibrotactors to provide haptic feedback representing Ekman´s six universal emotions. Eight PVIs (mean age: 46.75, age range: 28-66) participated in two training sessions and one experimental session. During the experiment, participants engaged in two 15-minute conversations, in one of which they wore the emotion recognition system. Exit-interviews were conducted after the experiment to assess the experiences of the participants. Due to technical issues with the registration of the emotion recognition software, only six participants were included in the video analysis. We found that participants were quickly able to learn, distinguish, and remember vibrotactile signals associated with the six emotions. Four participants felt they were able to use the vibrotactile signals in the conversation. Five out of six participants had no difficulties in keeping the

[4] Buimer HP, Schellens R, Kostelijk T, Nemri A, Zhao Y, Van der Geest TM, Van Wezel RJA (2019) Opportunities and Pitfalls in Applying Emotion Recognition Software for Persons With a Visual Impairment: Simulated Real Life Conversations. JMIR Mhealth Uhealth;7(11):e13722. DOI 10.2196/13722

camera focused on the conversation partner. The emotion recognition was very accurate in detecting happiness but performed unsatisfactorily in recognizing the other five universal emotions. The system requires essential improvements in performance and wearability before it is ready to support PVIs in their daily life interactions. Nevertheless, the participants saw potential in the system as an assistive technology, assuming their user requirements can be met.

4.1 Introduction

4.1.1 Background

A large number of communication cues exchanged between persons are nonverbal (e.g., gestures, facial expressions, and gaze direction). As a result, the inability to perceive these cues leads to a loss of communication effectiveness. Not perceiving these nonverbal cues is particularly present in PVIs and can lead to feelings of exclusion [1,2]. Previous inventory studies of assistive technology needs [3,4] found that a need among the community of PVIs still exists for a solution that makes nonverbal signals accessible.

To convey such information and make it accessible for PVIs, visual information can be translated into auditory or tactile cues, which is the foundation of SSDs. The SSD developed by Bach-y-Rita [5], often considered as the first SSD, translated a video feed into a grid of vibration motors attached to the back of a chair. PVIs trained to use this system were able to pick up an object. More recently, the vOICe system (which translated a video with very high contrast directly into auditory cues) and a Tongue Display Unit were developed (which translated video into electric signals on the tongue) [6–8]. However, these systems rely largely on senses and body parts that are crucial for interpersonal communication, that is, mouth and ears. In the last years, various researchers have acknowledged the issue of nonverbal information and worked toward a range of sensory-substitution designs: information such as interpersonal distance, location of others, facial expressions of emotions, and gestures was conveyed using various vibrotactile devices, including the back of a chair [9,10], a glove [11], and a vibrotactile belt [12,13].

In addition, recent advances in machine learning and computer vision technologies, which improved the ability of computers to recognize patterns from images, allow for very specific recognition tasks. One of the opportunities that arose from these advancements is the possibility to train a computer to recognize faces and facial expressions from a video. This idea was pioneered by Bartlett et al. in a collaboration with Paul Ekman [14], leading to effective emotion recognition software. Among its numerous applications, such software is particularly interesting for people who are unable to detect and interpret facial expressions of emotions, such as PVIs. Given the need from the community and the potential of emotion recognition software in meeting these needs, we investigated to what extent such emotion recognition software can support PVIs in their daily lives.

Previously, we presented a wearable system for facial emotion recognition consisting of a head-mounted camera and a vibrating belt [15]. It was designed to help PVIs perceive facial expressions of the six universal emotions [16]. Via a tiny camera attached to standard spectacles, pointing in the direction where the wearer is looking, a continuous video stream was available for processing. To determine which emotions were expressed, FaceReader facial expression recognition software was applied. This software can be used to detect Ekman's universal emotions from pictures and videos [16–18]. In this study, FaceReader was applied to analyze the video stream, originating from the video camera, in real time. Once a face is detected, the software attempts to create a map of the face including almost 500 facial landmarks. If a face muscle movement is detected, the deviation from a baseline neutral face is recognized as a facial expression and then classified as an emotion. Depending on the occlusion of the faces, either a standard model (called 3D Active Appearance Model) or proprietary deep neural networks are used to assign a score to the intensity of the facial expression (emotion). Once an expression was detected as being one of the six universal emotions, vibrotactile signals on a waist belt conveyed this information to the wearer of the system. Previous research has shown that vibrotactile signals were interpretable and usable in cognitively and physically demanding tasks such as flying a helicopter or steering a boat [19]. By signaling emotions through haptic feedback at the waist during social interactions, the ears of the users remained free to engage in a conversation.

This system was previously tested with PVIs in a controlled laboratory setting [15], with pictures from the Warsaw Set of Emotional Facial Expression Pictures (WSEFEP) [20] and videos from the Amsterdam Dynamic Facial Expression Set (ADFES) [21]. Such validated sets of images and videos of actors showing strong emotional expressions in controlled laboratory conditions play an important role in the training and development of emotion recognition software. However, these sets do not entirely resemble realistic facial expressions [22]. The results of our first study were promising, as the FaceReader software reached high recognition scores and users were able to easily learn how to interpret and use the vibrotactile signals from the system [15]. That the system worked well under controlled laboratory lighting in classifying very expressive facial expressions was expected, as emotion recognition software often achieved high recognition scores when tested with validated sets [18,23]. On validated sets such as the ADFES and WSEFEP, FaceReader was known to achieve recognition rates of 89% and 88%, respectively [17]. However, these sets are often acquired during photoshoots with actors who express strong emotional states under optimal lighting conditions. Facial expressions in real life are often much subtler than those found in the validated sets, and lighting conditions also differ between real life and controlled conditions [22,24,25]. As a result, we cannot be certain that the recognition scores achieved from these validated sets can be replicated once such software is faced with genuine facial expressions of emotions, expressed in natural conversations. Thus, to determine whether the technology applied is useful for this purpose, its performance must be further tested under real-world conditions [26].

4.1.2 Objectives

To investigate whether the system is usable and useful during realistic conversations, we aimed to examine the following questions: (1) were PVIs able to keep the camera focused on the face of the conversation partner?; (2) did the software reach a satisfactory recognition rate when confronted with an unstable video stream and suboptimal lighting conditions?; (3) were PVIs able to interpret the emotion cues while engaging in a meaningful conversation?; and (4) did the system live up to prospective users' expectations, and did they believe they could use such technologies in their daily lives?

4.2 Methods

4.2.1 Ethical approval

The study was designed in accordance with the Declaration of Helsinki and approved by the ethical committee of the Faculty of Electrical Engineering, Mathematics and Computer Science of the University of Twente, Enschede.

4.2.2 Participants

A total of eight PVIs (four females and four males; mean age 46.75 years, age range 28-66 years) were included in the study. Moreover, four of the participants were visually impaired from birth, whereas the others became visually impaired later in life but had experience with vision for at least ten years. All participants reported difficulties with the recognition of facial expressions and did not suffer from any other cognitive or sensory impairments (Table 4.1).

Table 4.1. Participants.

ID	Age (years)	Gender	Visual impairment	Sight	10 or more years of vision	Emotion logs available[a]
1	28	Male	Neurological damage to eye nerves	Tunnel	No	Yes
2[b]	28	Male	Persistent fetal vasculature	None	No	No
3[b]	66	Male	Retinitis pigmentosa	None	No	No
4	59	Female	Retinitis pigmentosa, glaucoma, and cataract	None	Yes	Yes
5	47	Male	Cone dystrophy	Peripheral	Yes	Yes
6	64	Female	Congenital rubella syndrome	None	No	Yes
7	50	Female	Retinitis pigmentosa	None	Yes	Yes
8	32	Female	Aniridia	<5%	Yes	Yes

[a] *See the section Data Collection and Analysis.*
[b] *No emotion logs were collected during the experiment.*

4.2.3 Apparatus

The experiments were conducted in an observation room at the HAN University of Applied Sciences (Arnhem, The Netherlands), where it was possible to discreetly observe the experiments from an adjacent viewing room and record with three different cameras and two microphones for post analysis.

The emotion recognition system used in the study consisted of a small USB camera clipped on spectacles; a Microsoft Surface Pro 4 tablet running the FaceReader 6 SDK emotion recognition software (running the general face model); and a chain of nine vibrotactile stimulators attached to a Velcro belt, 6 of which were used (see Figure 4.1). To ensure steady connections between the components, cable connections were used.

The FaceReader system classified facial expressions into one of the six of Ekman's universal emotions: happiness, sadness, anger, fear, surprise, and disgust [16]. When an emotion was recognized, feedback was given to the participant via the vibrotactile stimulators worn around the waist. Each emotion was assigned to a single vibrotactile stimulator (Figure 4.2). The decision to use a belt with vibrotactile stimulators and let it be worn on the waist had several reasons. First, the tactile modality was chosen because the use of the auditory modality is undesirable during social interactions, in which auditory information is very important for PVIs and should, therefore, not be interfered with. Second, the tactile resolution of the waist is enough to place six vibrotactile stimulators without the risk of users being unable to distinguish them. A belt can be worn unobtrusively underneath clothing and wearing it at the waist means that the system does not interfere with the conversations in the sense that it keeps your eyes, ears, and hands free.

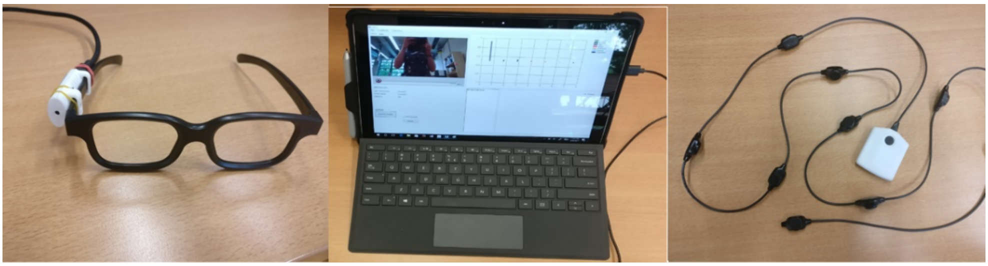


Figure 4.1. Schematic overview of the used system.

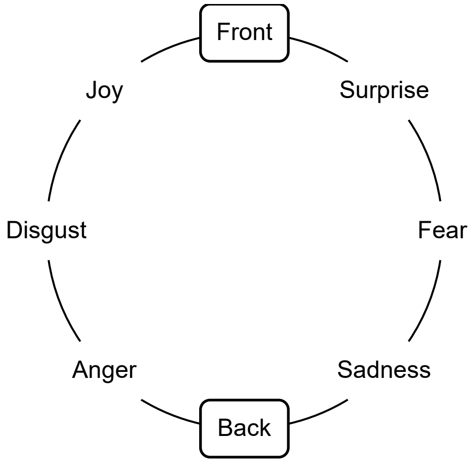


Figure 4.2. Emotion mapping. The mapping of Ekman's universal emotions on the waist band.

Vibrotactile stimulators were activated when the associated emotion was detected above a certain threshold of certainty, after which the vibration strength indicated the certainty that the software had detected an emotion. The standard threshold to initiate a vibrating pulse was 0.3 (on the FaceReader scale from 0 to 1), with two exceptions. In pretests with an actor, the FaceReader software was more sensitive to some emotions than to others. For example, when the actor was concentrating on listening and looking interested, his facial expression was at times classified as angry. Therefore, the threshold for the emotion anger was raised to 0.6. A threshold of 0.7 was chosen for the emotion happiness as it was rapidly detected even if the expression was barely visible.

4.2.4 Procedure

The study consisted of four different phases: initial training, refresher training, the experiment, and an exit interview.

First, the participants were trained to interpret the vibrotactile signals during two separate training sessions (Table 4.2). The first session took place about two weeks before the experiment. After a brief introduction to the tactile mapping around the waist, the training included application of four sets of 36 vibrotactile stimuli. After each stimulus, the participant was asked to indicate which emotion was signaled. These stimuli were initially accompanied with matching auditory cues that matched the vibrotactile feedback. The number of audio-accompanied stimuli steadily decreased (100%, 50%, 33%, and 0%), resulting in a full dependency on vibrotactile signals in the fourth set. In these sets, each emotion was equally represented. The fifth and last set included only vibrotactile feedback. The minimal number of stimuli presented in this set was 20, whereas the maximum number of stimuli was set at 96. It was ended when participants achieved a 95% success rate over the last twenty stimuli that were presented to them. The initial training phase lasted approximately 65 min.

The refresher training session took place just before the experiment and consisted of three sets. The first two sets included 24 stimuli in which the emotions were equally represented. The first set was fully accompanied by audio, whereas the second set included only vibrotactile feedback. The third set was the same as the final set of the initial training. It included 96 vibrotactile-only stimuli and was ended when participant achieved a 95% correct rate in the last twenty stimuli presented to them. The refresher training had an average duration of approximately 30 minutes.

After training was completed, the participants engaged in two conversations with an actor for approximately fifteen minutes. The conversations were set up as a mock job interview in which the actor played the role of the director of a fictional company. The actor and the participant were seated at a table approximately 1.5 meters opposite of each other.

The first conversation was colloquial, whereas the second was more formal. This was done to put the participants at ease and because we expected that the quantity and type of facial expressions of emotions might differ between the two.

The actor followed a similar structure for all conversations with the eight participants. He was instructed not to exaggerate any facial expressions to keep the conversations as realistic as possible. A total of four participants wore the system in the first conversation, whereas the remaining four wore it during the second conversation.

After the experimental session, a semi structured exit interview was conducted to allow participants to share their thoughts and ideas about the system. It addressed various themes associated with system functionality, user experience, and technology acceptance (see Table 4.3).

Table 4.2. Stimuli distribution. Overview of the number of stimuli that were presented to the participants in each training session and the percentage of audio-accompanied stimuli.

Session	Total number of stimuli (percentage audio accompanied stimuli), n (%)				
1	36 (100)	36 (50)	36 (33)	36 (0)	20-96 (0)[a]
2	24 (100)	24 (0)	20-96 (0)[a]	— [b]	—

[a]*Ended after a 95% correct score in the last twenty stimuli.*
[b]*The second session included only three sets of stimuli.*

Table 4.3. Exit interview topics. An overview of the topics discussed in the exit interview, the associated questions, and the subthemes of interest.

Topic, exit interview question	Subthemes
Experience	
How was your experience using the system?	Accuracy, usefulness, feedback, fun of use, and ease of use
What was the value of the system during the conversation?	Additional value, distraction from the conversation, and interpretation of the system
Potential utility	
Would you like to use the system in your daily life?	System usage, situations of use, and recommendation to others
Imagine you have a job interview next week and we will lend you the system. Would you use it during the job interview?	Usage in job interview
Adjustments to the system	
If you are the manager of the team that develops this system, what do you think they should tackle first?	Improvements of the system and additions to the system
Social acceptance	
How do you expect people in your surroundings (eg, family, friends, and colleagues) will react to you using a system like this?	Reaction of surroundings and introduction of the system

4.2.5 Data Collection and Analysis

After the experimental sessions, five sets of data were acquired. First, the training data were collected and analyzed by tallying the number of correct and wrong answers for each set of stimuli. Second, to determine the performance of the emotion recognition software, video recordings and emotion logs were saved and reviewed. The actor and the participant were filmed from 3 different camera positions (one ceiling camera and one stationary camera aimed at the participant, and one ceiling camera focused on the actor). The emotion logs kept a record of the emotion recognition software by saving timestamps together with the strength of a detected emotion (a value between 0 and 1; an example of such a log can be found in Table 4.4). Each time the software detected an emotion above threshold for at least five timestamps within a time span of one second, a fragment was labeled as an emotion signal. Due to technical issues, the emotion logs of only six out of the eight experiments were analyzable.

Table 4.4. Example of an emotion log. The emotion log provides a value between 0 and 1 for each emotion at each timestamp (hours [hh]:minutes [mm]:seconds[ss]:milliseconds[ms]).

hh	mm	ss	ms	Neutral	Happiness	Sadness	Anger	Surprise	Fear	Disgust
14	1	24	226	0.848577	0.261171	0.045551	0.037921	0.066328	0.101152	0.034363
14	1	24	279	0.853517	0.246223	0.052067	0.039493	0.064576	0.089338	0.034737
14	1	24	354	0.860102	0.227184	0.061851	0.045071	0.060259	0.075662	0.033203
14	1	24	439	0.868132	0.206724	0.068306	0.054500	0.054876	0.063745	0.032779
14	1	24	550	0.875653	0.181231	0.069503	0.060954	0.048684	0.050882	0.034567
14	1	24	617	0.866770	0.157769	0.065179	0.063756	0.041683	0.042549	0.050491
14	1	24	673	0.850679	0.149551	0.064955	0.064503	0.038547	0.039288	0.058762
14	1	24	742	0.834428	0.144484	0.065088	0.065431	0.036603	0.038934	0.061535
14	1	24	798	0.807851	0.137719	0.065118	0.067514	0.033447	0.040757	0.061940

The third dataset consisted of the relationship between the emotion log fragments and the video recordings. For every instance FaceReader recognized an emotion, as was derived from the emotion logs, a video fragment was created, resulting in 166 video fragments that required annotation across all videos. A total of two independent coders then analyzed each fragment to determine if they recognized a facial expression of any of the six universal emotions (happiness, sadness, anger, fear, surprise, and disgust) and annotated the video fragments accordingly. In cases where the coders did not identify one of these emotions (be it seeing another emotion or seeing no emotion at all), the fragment was coded as "none/other". Eventual disagreements were discussed in person to reach consensus.

To analyze how well the software detected emotions, a fourth set was used. This set consisted of a sample of every third minute that was extracted from the videos. This resulted in a total of 29 minutes of video that was analyzed for facial expressions of emotions, independent of the emotion logs. A first coder analyzed the videos and identified fragments in which facial expressions of emotions were thought to be visible. A second coder analyzed the same sample of randomly selected videos while seeing the selected fragments by the first coder. The second coder checked whether there was an agreement with the expressions identified by the first coder or if new expression were to be added. Again, any disagreements were discussed to reach consensus. Finally, the fragments in which facial expressions were detected by the two coders were compared with the emotion logs to analyze how well the system detected facial expressions of emotions during the conversations.

The fifth and final set were the exit interview transcriptions. The exit interviews were recorded and transcribed verbatim, which resulted in seventy pages of text. A content analysis was performed to extract the most important information fragments from the transcriptions. Codes were assigned to all fragments, based on the question to which the fragments provided an answer. To structure the answers, the subthemes mentioned in Table 3 were assigned to the answers.

4.3 Results

4.3.1 Learn to Interpret Vibrotactile Signals

Both training sessions successfully taught participants how to interpret the vibrotactile signals. Participants were quickly able to correctly identify most stimuli. The average performance gradually improved over each set of stimuli. Moreover, six out of the eight participants achieved perfect scores in the final set (see Table 4.5).

The refresher session, right before the experiment, showed even better results, ending with a perfect recognition rate for all participants in the third and final set (Table 4.6).

It can be concluded that during the training session, most participants were able to learn how to identify the correct emotions from the vibrotactile signals.

Table 4.5. Number of errors during the initial training. The table shows the numbers of errors of all participants during the initial training session.

ID	Set 1	Set 2		Set 3		Set 4	Set 5	Total
	A[a] n=36	A n=18	T[b] n=18	A n=12	T n=24	T n=36	T n=20-96	
1	3	0	4	0	0	1	0	8
2[c]	0	0	0	0	0	0	0	0
3[c]	10	3	0	1	0	1	7	22
4	0	0	0	0	1	0	0	1
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	1	0	2	0	0	0	0	3
8	0	0	1	1	0	1	2	5
Total	14	3	7	2	1	3	9	39

[a]*A: auditory accompanied vibrotactile stimuli.*
[b]*T: tactile only stimuli.*
[c]*No emotion logs were available for this participant in the experiment.*

4.3.2 Camera Pointing

A prerequisite for the emotion recognition system is that participants can focus the camera on the face of the conversation partner. Generally, the participants were able to focus the camera on the actor who sat on the other side of the table. In five out of the six videos analyzed, the face of the actor was out of sight for only a couple of seconds out of the approximately fifteen minutes that the conversation lasted. Moreover, one participant was an exception to this rule. This participant tended to slightly lift her head upward. As a result, the face of the actor was out of the field of view of the camera many times, rendering the emotion recognition software useless in our context.

Table 4.6. Number of errors during the refresher training session. The table shows the number of errors of all participants during the refresher training session.

ID	Set 1, A[a] (n=24)	Set 2, T[b] (n=24)	Set 3, T (n=20-96)	Total
1	1	0	0	1
2[c]	0	1	0	1
3[c]	6	0	0	6
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0
8	2	0	0	2
Total	9	1	0	10

[a]A: auditory accompanied vibrotactile stimuli.
[b]T: tactile only stimuli.
[c]No emotion logs were available for this participant in the experiment.

4.3.3 Emotion Recognition Performance

Agreement With Software-Detected Emotions

This analysis was based on the video fragments in which the software had detected an emotion. A total of two coders annotated 166 fragments for which the emotion log values exceeded threshold levels for both emotion value and duration. Initially, the coders disagreed thirty times, but mutual consensus was achieved after reviewing the fragments together.

The agreement between the codes assigned by the coders and emotion recognition software was poor. In only 54 out of the 166 fragments, the coders and FaceReader classified the same universal emotions (Table 4.7).

In thirteen cases the coders agreed upon seeing one of the six universal emotions but differed from the emotion detected by the software. For most fragments (99), however, the coders disagreed with the software in the sense that the software detected a universal emotion, whereas the coders did not. In this analysis, the coders and the software showed most agreement on the emotion happiness, whereas the performance for other emotions is far worse. It can be concluded that the coders disagreed with the signals that FaceReader conveyed for five out of the six emotions. Only when it came to the signaling of happiness, the coders largely agreed with the software.

Table 4.7. Crosstabs of agreement between coders and software. The table shows a tally of the number of time the coders and FaceReader classified a fragment as a particular emotion. The diagonal shows the number of times that the coders and FaceReader classified.

FR	Coders							
	Anger	Disgust	Happiness	Sadness	Fear	Surprise	None/other	Total
Anger	2	0	0	0	0	0	8	10
Disgust	1	0	0	0	0	0	12	13
Happiness	0	0	42	0	0	1	3	46
Sadness	1	0	2	4	0	1	41	49
Fear	0	0	2	0	0	1	4	7
Surprise	0	0	4	0	0	6	31	41
Total	4	0	50	4	0	9	99	166

4.3.4 System Performance to Detect Facial Expressions of Emotions

To determine how well the system derived facial expressions from the actor, we checked the random sample of every third minute. In this time, the two coders detected 72 facial expressions of basic emotions and another twelve expressions not necessarily being associated with one of Ekman´s emotions. The software detected only 44 in the same time frame. In twenty cases there was an agreement between the human coders and the system detection. A total of 24 times the human coders and FaceReader disagreed about emotions being expressed. The remaining 51 facial expressions of universal emotions that were detected by the coders (31 instances of happiness and 20 instances of surprise) were not recognized by FaceReader (Table 4.8).

These results show that in the random sample of videos, mostly happiness and surprise were expressed according to the coders. In the 44 cases where the software did detect an emotion, the coders disagreed in almost half of the cases. The only exception seemed to be happiness, which, if it was detected, was agreed upon in 17 of 19 cases. However, the coders detected many more instances of happiness that the software was unable to detect. This was similar for facial expressions of surprise.

Table 4.8. Overview of the agreement between human coders and the software for the facial expressions detected. The table shows the classification of detected emotions.

Emotions	Agreement	Disagreement	Not detected by FR[a]
Anger	0	5	0
Happiness	17	2	31
Disgust	0	2	0
Sadness	0	10	0
Fear	0	2	0
Surprise	3	3	20
Total	20	24	51

[a] *The coders identified 11 other facial expressions that did not classify as 1 of the universal emotions.*

User Experience

Besides the objective measures to analyze the technical performance of the system in a realistic condition, we asked the participants how their experience was using the system and whether they saw potential utility of the system in their daily lives.

How did persons perceive the vibrotactile cues during the conversation?

> *"There was no doubt about the transferred information. Once you know the location of emotions, the usage of the system is very convenient." [P2, male, aged 28 years, fully blind]*

A total of four participants even believed they were able to subconsciously interpret the vibrotactile signals:

> *"I am consciously participating in the conversation while I am taking the feedback subconsciously into account, so the vibrations did not distract me, absolutely not." [P4, female, aged 59 years, fully blind]*

Moreover, three others found it difficult to focus their attention on the signals and struggled to interpret them. Participant 2 (male, aged 28 years, fully blind) felt that the prolonged stimuli really disrupted his ability to engage in the conversation with the actor. Participant 5 (male, aged 47 years, partially sighted) found it particularly difficult to interpret the signal while he was speaking himself, although he felt that the interpretation of “happiness” and “surprise” was easier than the other emotions as these occurred more frequently than other emotions. Indeed, during his conversation, happiness and surprise were conveyed quite often (seven and five times, respectively), although sadness was conveyed as often as surprise.

In addition, three participants did not realize that vibration strength was linked to the intensity of the expressed emotion and explicitly requested for such a functionality. Participant 1 (male, aged 28 years, partially sighted) suggested to use pulsating vibrations instead to convey intensity. Thus, to convey such information other methods are wished for.

A total of seven participants were skeptical about the accuracy of the conveyed emotions, as these did not always correspond to the vocal cues and the atmosphere during the conversation. Participant 4 (female, aged 59 years, fully blind), for example, reported to have received signals of a disgusted facial expression, whereas she felt there was absolutely no reason to believe the actor looked disgusted. In the video analysis of her interview with the actor, no facial expressions of disgust were identified by the two coders, whereas there were seven instances where the software conveyed disgust. Participant 1 (male, aged 28 years, partially sighted) felt that the system did consistently provide information at the moments it was most necessary. Moreover, two persons reported receiving only limited signals.

Despite the inaccuracy, six participants were convinced about the usefulness of the concept of an emotion recognition system. Participant 7 (female, aged 50 years, fully blind) believed such a system would especially be useful in the absence of vocal cues, where it would be the only source of information. In addition, confirmation of the emotions derived from vocal cues led to more confidence. Furthermore, one participant (Participant 5, male, aged 47 years, partially sighted) mentioned that the system could be valuable while the user himself is talking, as feedback on emotional expressions of the conversation partner in such situations can make the user more confident about the story he or she is telling. On the downside, two persons reported that because of inaccuracies, the system caused more confusion than clarity. Participant 4 (female, aged 59 years, fully blind) mentioned that she believed she could detect faulty signals and therefore was not so much deterred by the errors. It could be that some participants are more willing to accept errors as these did not weigh up against the experienced usefulness of the system.

A total of four participants wanted to use the system in their daily life, whereas three others wished for more time with the system before they could decide. Moreover, two persons were convinced that the system, as it is, would be an added value to their life. In addition, three others believed that the system would have an added value after some adjustments.

Besides the obvious accuracy improvements, participants wished for improvements in the wearability of the system and requested additional functionalities. A total of seven participants said they would like a more portable system, which could be achieved by going fully wireless. All but one participant requested for a camera that is integrated in the spectacles. The one participant who did not request for a camera integrated in spectacles (P5, male, aged 47 years, partially sighted) suggested that the camera should be clearly visible to make conversation partners aware they are being filmed. Moreover, two participants suggested to get rid of the spectacles completely and allow users to clip the camera anywhere on their clothing.

A total of seven out of the eight participants believed that a person recognition function would increase the usefulness of the system. Participant 1 (male, aged 28 years, partially sighted) even stated that emotion recognition and face recognition are inseparable, as in a busy environment you need to know who you look at before emotion recognition becomes relevant. Such technologies could also be applied to locate a specific person that you are looking for. Related to this, participant 2 (male, aged 28 years, fully blind) wished for a function to request for feedback from the system at any time to check whether the system is still on and focused after long periods of inactivity. Such a trigger was also suggested by participant 1 (male, aged 28 years, partially sighted), who wished for a trigger to switch the feedback on or off to avoid an information overload from the system.

4.4 Discussion

4.4.1 Principal Findings

The main objective of this study was to determine whether the system presented in an earlier study [15], which was successfully tested in laboratory conditions, would work in a realistic conversation situation. The study affirmed that a positive laboratory test does not necessarily mean that a system is ready for real-life usage [26]. Although the concept of our system seems to have potential, it requires essential improvements before it is ready to support PVIs in real life.

We have confirmed findings from a previous study [5] and showed that it is very easy to learn how to interpret vibrotactile cues signaled with a haptic waist belt. That in itself is not a surprise – both older [5,27] and more recent [28,29] studies have shown that much more complex tactile patterns can be successfully learned. Furthermore, it was already known that healthy persons can interpret vibrotactile cues in demanding environments [30]. However, what is interesting is that we now have indications that users of the system with a visual impairment were able to use information conveyed through vibrotactile signals while they engaged in a conversation.

Most of the participants were able to keep the camera focused in the direction of the actor for most of, if not all, the time. That means that the principle of putting a camera on spectacles will provide a sufficient camera feed for such computer vision analysis. Moreover, one person, however, aimed the camera above the actor for most of the time. Before using such a system, this person would require training on how to aim the camera to benefit from the system. Another option would be to redesign the camera and spectacles in such a way that the camera angle can be altered to match the wearer's preference. This is a feature that some mobile eye trackers already have.

The accuracy of the software was not sufficient to convey all of Ekman's [16] six universal emotions under our experimental conditions. Earlier studies on the performance of the software were conducted using validated sets composed of pictures including high-quality full-face pictures in ideal lighting conditions. In such conditions, the software used in the study achieved accuracy scores as high as 88% [17]. In our experiment, however, the software was confronted with suboptimal lighting conditions and unstable shots because of a moving camera and actor. In these conditions, the only emotion that was accurately recognized was happiness, which was also reported by the participants. This is if the emotion was recognized, as the coders detected 31 more instances of a facial expression of happiness that the software did not detect. The other five emotions were poorly recognized. As a result, the performance achieved in the experiment was too poor to be a support for PVIs in real life.

It is unlikely that the system presented in the study would be used by the participants in their daily lives in its current form. Too many adjustments were suggested to believe that implementation would be easy (e.g., emotion recognition accuracy, portability, person recognition, and an on/off switch). Nevertheless, there is reason to believe that such a system might be used, once the required adjustments are adapted.

4.4.2 Limitations

The emotions were not uniformly represented in the conversations. It is clear that the emotion happiness was far more often detected and conveyed via vibrotactile feedback than the other emotions. Therefore, it is difficult to draw definitive conclusions on the accuracy of the other emotions.

During the experiment, it was impossible to objectively measure if and how well participants were able to perceive the vibration signals. Therefore, we had to rely on the reports by the participants when it comes to the interpretation of signals.

We chose not to encode any other emotions than the six universal emotions. As a result, we cannot provide an overview of all the facial expressions of emotions that were expressed in the conversations. Thus, we cannot state which emotions are more often present in realistic conversations and could possibly be added to the emotion recognition software.

4.4.3 Conclusions

For systems such as the one presented in this study, the main improvements to be made should be sought in the direction of both the wearability of the system and improvement of the emotion recognition software. The solution for wearability would be to make components of the system smaller (from tablet to smartphone) and by making all connections between system components wireless. To improve emotion recognition software, we suggest training with more ambiguous and subtle facial expressions under different lighting conditions. We do believe that if these concerns are taken care of, systems such as the one presented here might be of great benefit for PVIs in their future daily social interactions.

4.5 References

1. VISION2020 (2005) Vermijdbare blindheid en slechtziendheid in Nederland
2. Schölvinck AFM, Pittens CACM, Broerse JEW (2017) The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 111:201–217
3. Van der Geest TM, Buimer HP (2015) User-centered priority setting for accessible devices and applications. In: Weisbecker A, Burmester M, Schmidt A (eds) Mensch und Comput. 2015 – Work. Oldenbourg Wissenschaftsverlag, Stuttgart, pp 383–389
4. Krishna S, Colbry D, Black J, Balasubramanian V, Panchanathan S (2008) A systematic requirements analysis and development of an assistive device to enhance the social interaction of people who are blind or visually impaired. Workshop on Computer Vision Applications for the Visually Impaired, James Coughlan and Roberto Manduchi, Marseille
5. Bach-y-Rita P (1967) Sensory plasticity: applications to a vision substitution system. Acta Neurol Scand 43:417–426. DOI 10.1111/j.1600-0404.1967.tb05747.x
6. Ward J, Meijer P (2010) Visual experiences in the blind induced by an auditory sensory substitution device. Conscious Cogn 19:492–500. DOI 10.1016/j.concog.2009.10.006
7. Meijer PBL (1992) An experimental system for auditory image representations. IEEE Trans Biomed Eng 39:112–121. DOI 10.1109/10.121642
8. Kaczmarek KA (2011) The tongue display unit (TDU) for electrotactile spatiotemporal pattern presentation. Sci Iran 18:1476–1485. DOI 10.1016/j.scient.2011.08.020
9. McDaniel T, Bala S, Rosenthal J, et al (2014) Affective haptics for enhancing access to social interactions for individuals who are blind. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), pp 419–429
10. McDaniel T, Tran D, Chowdhury A, et al (2019) Recognition of tactile facial action units by individuals who are blind and sighted: A comparative study. Multimodal Technol Interact 3:32. DOI 10.3390/mti3020032

11. Krishna S, Bala S, McDaniel T, et al (2010) VibroGlove: an assistive technology aid for conveying facial expressions. Proc SIGCHI Conf Hum Factors Comput Syst Ext Abstr, pp 3637–3642. DOI 10.1145/1753846.1754031
12. McDaniel TL, Krishna S, Colbry D, Panchanathan S (2009) Using tactile rhythm to convey interpersonal distances to individuals who are blind. Proc CHI 2009 Ext Abstr, pp 4669–4674. DOI 10.1145/1520340.1520718
13. McDaniel T, Krishna S, Balasubramanian V, et al (2008) Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. In: HAVE 2008 - IEEE Int. Work. Haptic Audio Vis. Environ. Games Proc., pp 13–18
14. Bartlett MS, Hager JC, Ekman P, Sejnowski TJ (1999) Measuring facial expressions by computer. Psychophysiology,Cambridge Univ Press 36:253–263
15. Buimer HP, Bittner M, Kostelijk T, et al (2018) Conveying facial expressions to blind and visually impaired persons through a wearable vibrotactile device. PLoS One 13:e0194737. DOI 10.1371/journal.pone.0194737
16. Ekman P (1970) Universal facial expressions of emotions. Calif Ment Heal Res Dig Dig 8:151–158.
17. Lewinski P, Den Uyl TM, Butler C (2014) Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. J Neurosci Psychol Econ 7:227–236. DOI 10.1037/npe0000028
18. Den Uyl MJ, Van Kuilenburg H, Lebert E (2005) The FaceReader: Online facial expression recognition. In: Noldus LPJJ, Grieco F, Loijens LWS, Zimmerman PH (eds) Proc. Meas. Behav. Wageningen, pp 589–590
19. Van Erp JBF, Van Veen HAHC, Jansen C, Dobbins T (2005) Waypoint navigation with a vibrotactile waist belt. ACM Trans Appl Percept 2:106–117. DOI 10.1145/1060581.1060585
20. Olszanowski M, Pochwatko G, Kuklinski K, et al (2015) Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. Front Psychol 5:1–8. DOI 10.3389/fpsyg.2014.01516

21. Van der Schalk J, Hawk ST, Fischer AH, Doosje B (2011) Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). Emotion 11:907–920. DOI 10.1037/a0023853
22. Dhall A, Goecke R, Lucey S, Gedeon T (2011) Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark. In: 2011 IEEE Int. Conf. Comput. Vis. Work. ICCV Work. IEEE, pp 2106–2112
23. Kharat GU, Dudul SV (2009) Emotion recognition from facial expression using neural networks. In: Human-Computer Syst. Interact., pp 207–219
24. Mehta D, Siddiqui MFH, Javaid AY (2018) Facial emotion recognition: A survey and real-world user experiences in mixed reality. Sensors (Switzerland) 18:1–24. DOI 10.3390/s18020416
25. Dhall A, Goecke R, Gedeon T, Sebe N (2016) Emotion recognition in the wild. J Multimodal User Interfaces 10:95–97. DOI 10.1007/s12193-016-0213-z
26. Pawluk DT V, Adams RJ, Kitada R (2015) Designing haptic assistive technology for individuals who are blind or visually impaired. IEEE Trans Haptics 8:258–278. DOI 10.1109/TOH.2015.2471300
27. White B (1970) Perceptual findings with the vision-substitution system. IEEE Trans Man Mach Syst 11:54–58. DOI 10.1109/TMMS.1970.299962
28. McDaniel T, Tran D, Devkota S, et al (2018) Tactile facial expressions and associated emotions toward accessible social interactions for individuals who are blind. In: Proc. 2018 Work. Multimed. Access. Hum. Comput. Interface - MAHCI'18. ACM Press, New York, New York, USA, pp 25–32
29. Bala S, McDaniel T, Panchanathan S (2014) Visual-to-tactile mapping of facial movements for enriched social interactions. In: 2014 IEEE Int. Symp. Haptic, Audio Vis. Environ. Games Proc. IEEE, pp 82–87
30. Van Erp JBF (2005) Presenting directions with a vibrotactile torso display. Ergonomics 48:302–313. DOI 10.1080/0014013042000327670

5 Guiding Pedestrians Who Are Blind Or Visually Impaired In Unfamiliar Environments With Vibrotactile Feedback: Five Case Studies

Abstract

The objective of the study was to explore whether a wearable GPS-based navigation system with a vibrotactile belt could support pedestrians with visual impairments to find their way towards a target destination in outdoor environments, what problems they encounter while doing so, and if the users showed intention to use such a system in real life. The prototype consisted of a smartphone, a Raspberry Pi mini-computer, and a waist belt with 8 vibrotactors. Five PVIs (three blind, two low vision; three late onset, two congenital) participated in the study. The participants were observed while navigating three routes on a university campus. The participants were informed about the course of the first route beforehand, and used their regular navigation aids. The second route was also sent to the participants before, and besides their regular navigation aids the prototype was worn to complete this route. Participants did not receive any information about the third route, and were dependent only on the prototype. Vibrotactile signaling to guide PVIs through an unfamiliar environment worked well in real world usage scenarios. Despite a lot of variation between the participants, the signals conveyed were intuitive and all participants were able to navigate with the prototype after some instructions, although extensive training is likely to lead to better results. The GPS accuracy which was approximately 5 meters at best, was not precise enough for visually impaired pedestrians. A solution would be to make the prototype less unforgiving by adding flexibility in the sense of adaptable routes.

5.1 Introduction

"It is surprising that in this day of advanced technology, the blind are still moving about in the world using a cane, a guide dog, a sighted companion, or an outstretched hand" [1]. Many years after the above quoted statement was made, white canes and guide dogs are still amongst the most popular navigation aids for PVIs [2]. This does not mean that there have not been developments to enhance the navigation abilities of PVIs. On the contrary, many electronic travel aids (ETA) have been developed over the years, applying different sensors (e.g. sonar, infrared, optical, GPS) and feedback modalities (audio, tactile) to support PVIs to improve their ability to navigate independent of helpers (see for reviews: [2–5]. Despite these developments, few of these systems have been commercialized and are widely used in the everyday life of PVIs [4]. As a result, one of the biggest challenges that PVIs face in daily life remains to be navigation, both indoor and outdoor, in familiar and particularly in unfamiliar places [6].

What makes it so difficult to develop navigations aids for PVIs? Firstly, PVIs are far from a homogeneous group. Not only do PVIs differ in type and extent of visual impairment. They also differ in age of onset, level of education, intelligence, as well as in their mobility and orientation skills [7]. Furthermore, 81% of PVIs are over fifty years old [8], which typically means that new technology might be more slowly accepted as older persons are known to be generally slower at adopting new technologies. In that context, we believe it is important to work following universal design principles to assess user needs and that devices developed for PVIs are adaptive to end user preferences (e.g. in- and output modality, relevance of feedback), which might vary across usage contexts [9]. Therefore, we chose to test our new prototype at a relatively early stage of development and get feedback from a diverse group of PVIs.

In this study we present a prototype that supports PVIs to independently navigate a route towards a destination, stay on track while en route, and to determine whether they reached their destination. What should such a device encompass? Schölvinck and colleagues found a wish amongst PVIs for the use of existing technology platforms, such as smartphones, to replace standalone assistive devices which are often costly [10]. Using a smartphone as the core component of an ETA has the advantage that it is already equipped with a set of sensors that can power navigation applications (GPS, internet, gyroscope). It also allows the user to carry only one device, a crucial feature for PVIs who might be using canes and/or a guide dog in addition to their ETA. In recent work, these obvious advantages have led to smartphones being successfully used as the core component for an indoor-navigation assistant [11,12]. There are several design requirements for ETAs that support outdoor navigation: they should be wearable (hands and ears free), portable (comfortable to carry), reliable, low-cost, user friendly, robust, wireless, and convey information in real time to not affect the user's walking speed [2]. Previous research has studied supporting pedestrians through various feedback methods, such as tactile feedback from shoe soles [13,14], wrist bands [15] and shapeshifting devices [16], as well as auditory feedback [17–20]. A proven hands and ears free method to convey directional information is the use of vibrotactile signals on a waist belt [14,21–29]. The combination of an everyday smartphone with such a belt should make for a relatively low cost and accessible device. However, most studies are conducted with (in some cases blindfolded) sighted participants in order to test the system functionality. This does not generalize well to pedestrians with visual impairments who have completely different perception of the world surrounding them and who face unique navigation challenges, as societal infrastructure is still primarily designed around vision.

With PVIs, few studies compare existing navigation strategies to the tested device. In this study, we allowed participants to plan routes in advance using their regular strategies, and to use their regular travel aides for some routes, with and without our system. This allowed us to get a clearer sense of the added value of our system in the daily life context of our participants.

Therefore, the objective of our study was to explore whether a wearable GPS based navigation system with a vibrotactile belt could support visually impaired pedestrians in finding their way outdoors, with varying levels of prior planning. We also tested how our navigation system interacted with regular travel aids, whether electronic or cane/guide dog. Finally, during and after the experiments, we inquired what problems were encountered, and if users showed intention to use such a system in real life.

5.2 Materials and Methods

5.2.1 Ethical Approval

The study protocol was evaluated and approved by the ethics committee of the Faculty of Electrical Engineering, Mathematics and Computer Science at the University of Twente (Enschede, The Netherlands).

5.2.2 Participants

PVIs from an established pool of past participants and interested individuals were approached to take part in the study. Participants were invited to take part if they were of adult age, were visually impaired to an extent that caused difficulties in navigating unknown environments and did not have other sensory and/or cognitive impairments. Furthermore, the participants were screened to guarantee at least three participants that were congenitally blind and three that acquired blindness later in life. Oral informed consent was sought and recorded on video prior to formal inclusion in the study. Eventually, seven persons took part in the study, two of whom were excluded because data collection was incomplete due to equipment failure in recording the vibrotactile inputs. The excluded participants performed similarly to other participants with a comparable visual impairment.

5.2.3 Wearable Pedestrian Navigation System

For this study, we developed a custom wearable navigation system for visually impaired pedestrians. User-centered design principles guided our development process, with priority given to ease-of-use and minimizing barriers to adoption. This resulted in the choice of an Android-powered smartphone as core hardware and computing platform. This allowed for a highly mobile system with sufficient processing power to support real-time applications in a mid-range device (Huawei P8 Lite). In addition, recent smartphones contain several communication channels and sensors such as Wi-Fi, Bluetooth, GPS, and a gyroscope.

As many PVIs are keen smartphone users, our visually impaired participants had at minimum some familiarity with the user interface, with some displaying expert skills. To generate navigation instructions, we designed a custom app that comprises a route generator based on the principles of turn-by-turn navigation. The route generator allows users to search for a destination and receive the quickest route. With an active data connection, the app can also recalculate a new route if the user diverges too far from the initial route. To calculate and generate routes, OpenStreetMap and an open source routing engine for road networks (Open Source Routing Machine; OSRM) with a bicycle profile were used. Preliminary tests showed that these tools generate sufficiently accurate routes for our experiment.

To be able to compare performance between participants in the present study, we predetermined and cut up the entire route into parts, guiding the PVIs from waypoint to waypoint. Information concerning the direction where to go was conveyed to the users through a vibrotactile interface. The interface we chose was used in a belt configuration, worn around the waist. The belt is composed of nine spin-up vibration motors (dimensions: 34 × 16 × 11mm; maximum vibration frequency: 158.3±2.4Hz), of which eight were used in the study. They were linked to a control module (size: 56 × 48 × 14mm) used to communicate with our GPS software. The control module runs on a li-ion battery (3.7V, 800 mAh) and has an operation time of four to eight hours (Science Suit, Elitac). This enabled intuitive communication of directional instructions to the user, while leaving their hands free to use other assistive methods such as canes and guide dogs for near obstacle detection and avoidance. With the Elitac belt, direct wireless connection to the smartphone was not possible for our application, since a PC-only USB dongle was required to establish a Bluetooth connection. Therefore, to connect the vibrotactile interface to the smartphone, while ensuring mobility of the system, a Raspberry Pi was used as a workaround. The Pi was USB-connected to the haptic belt and opened a Wi-Fi network that the smartphone could connect to wirelessly. A small powerbank was attached to the Pi to ensure several hours of autonomy. The complete prototype weighed about 600 grams, including 350 grams for the Pi and powerbank (Figure 5.1). No discomfort was reported by the users. A next generation system running on the end-user's smartphone would mean that the vibrotactile belt can be worn without the powerbank and Pi, and thus weigh no more than a t-shirt, making it seamlessly wearable.

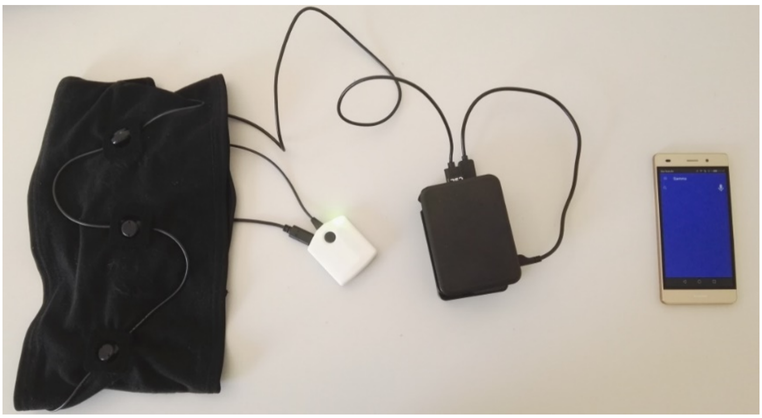


Figure 5.1. Overview of the complete system. Left to right: Haptic belt with 9 tactors and control module (approximately 90 g), Raspberry Pi (126g), PowerBank 10000 mAh (230g), and Huawei P8 Lite (132g).

5.2.4 Haptic Signaling

Eight tactors were used to convey directional information. The tactors were positioned evenly on a belt (see Figure 5.2). Because participants varied significantly in waist circumference, we used an elastic belt, which resulted in variation of the inter-tactor distance across participants. While different parts of the waist have some variation in their sensitivity to vibration, those differences were not relevant in the context of this experiment, since we operated well above detection threshold.

To drive the tactors during navigation, we have created a model that maps directions onto the tactors. Assuming the belt is a circle, we defined the upper intersection of a perpendicular straight through the center as the reference direction, with a value of zero degrees. This point is also equivalent to the completely straightforward direction. All other directions can then be represented as angles in this reference system (e.g. 90 degrees for right, 180 for back and 270 for left). Therefore, tactors were placed on the four cardinal directions plus the primary intercardinal directions to guide users [29–31]. Because it is a general model, a variable number of tactors can be used. The more tactors on the belt, the better the resolution, up to the perceptual limit of about fifteen tactors, where it becomes difficult to distinguish which tactor vibrates [32]. This reference system also nullifies the waist circumference differences between users, as they all share the same reference. A schematic of the model is shown in Figure 5.2.

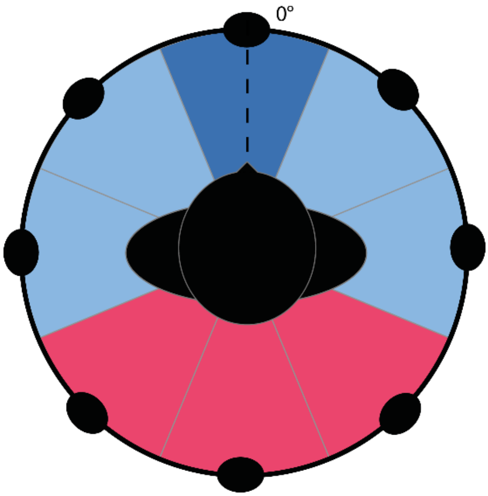


Figure 5.2. Tactor placement and direction coding of the haptic navigation interface. Eight tactors were worn on the waist. The tactor in the dark blue area signals if the user is correctly en route towards the next waypoint. If the tactors in light blue areas vibrate, the user should adjust his or her walking direction. When tactors in the red area vibrate, the user is walking away from the next waypoint and has to turn around. The dotted line represents the 0-degree line.

In addition, the belt's resolution for giving navigation directions was improved by interpolating between directional tactors. That is, if the navigation direction is exactly between two tactors, they would both vibrate at half intensity. When the navigation direction gets closer to one tactor, vibration intensity increases for that tactor, while decreasing for the next closest tactor. Vibration is then at maximum intensity for perfect alignment [33]. No additional vibrations patterns were used. The system conveyed short vibrotactile signals in the direction the user was supposed to be going, relative from where they were facing, every two seconds.

5.2.5 Experimental Procedure

Participants were asked to perform three different navigation tasks in an environment unfamiliar to them (University of Twente campus). During these navigation tasks at least two researchers were present. One was present to monitor and capture the experiment, while the other shadowed the participant and made sure the participant stayed safe throughout the experiment. The shadowing researcher also held the smartphone running the GPS software. Since the prototype used the smartphone sensors and compass, it was required to keep it roughly horizontal throughout the experiment.

Several days before the experiment, participants received information about two of the three routes in the form of an email with Google Maps links for these routes. This gave them the opportunity to prepare at home, on their own or with assistance. This allowed us to evaluate their existing resources and capabilities to plan and then move around in an unfamiliar environment. In contrast, they were not informed about the route for the third task.

During the first route (offered for preparation), participants were asked to use any assistive navigation technologies they would normally use (e.g. a smartphone app, or dedicated navigation device). For the second route (offered for preparation), the usual aids were used in combination with the prototype. The third and final route (not offered for preparation) was done with the prototype only. Participants were allowed to use their canes or guide dogs to detect and avoid obstacles at all time. Safety was ensured by an experimenter walking slightly behind the participant. The whole experiment was video-recorded by a second experimenter from a 5 to 10-meter distance. Following task completion, which took about one hour, the participants were briefly interviewed about their experience.

Routes

Several factors were considered when determining the routes (e.g. safety, comfort, complexity). Rather than following the shortest possible route from starting point to objective, the included routes were designed to present participants with a mixture of straight paths, corners, and obstacles, hence why automated routing could not be used. Furthermore, the likelihood of encountering cyclists on the paths – the campus is home to many shared spaces between pedestrians and cyclists – was minimized by avoiding paths with heavy traffic. Finally, to limit fatigue and stress amongst participants, the chosen routes offered plenty of opportunities to rest. As a result, the routes were walked in the same order for each participant. Although this might lead to a potential order effect, we believed the convenience of the participants outweigh the limiting effect that a possible order effect might have on the results. Each route had some more tricky areas to understand specific challenges that visually impaired pedestrians encounter during wayfinding. The first route included plenty of shared spaces as well as straight and wide paths. The second route was a little bit more challenging in the sense that the paths were less wide and guided the users in between multistory buildings, which could lead to poor GPS signals. The third route started on a parking lot, after which participants were asked to find their way through a small park, a wooden bridge, and a small pedestrian path. During both the second and the third route, participants encountered two crossroads within a couple of meters of which we knew they would be challenging for both the system and the participants.

Because routes were designed for the above-mentioned features and were as a result often not the shortest path between start and arrival, routes were annotated manually and navigation was done in the app's offline mode. Switching to live tracking of the user on the routes was still possible in principle. However, this would generally imply a deviation from the desired route and was therefore not used in this study. In addition, manual waypoints came at the cost of the possibility of rerouting in case of wrong turns. The routes used for this study are shown in Figure 5.3. After each route, participants were guided towards the starting point of the next route. Participants could be asked to stop with their current route if the system was continuously providing them with ambiguous information, when they were wandering and going nowhere near the aimed for destination or when they were heading towards danger such as a body of water or a ditch.

Figure 5.3. Routes included in experiment. The three pre-programmed routes on the University of Twente campus. The marker S is the start of the route, while F marks the finish: (1) Own resources, offered for preparation, 521m, 31 waypoints; (2) Own resources + prototype, offered for preparation, 420m, 19 waypoints; and, (3) Prototype, not offered for preparation, 503m, 32 waypoints.

5.2.6 Interviews

To conclude the experiment, semi-structured interviews were conducted with the participants. The interviews were used to gain insights in the ways PVIs usually navigate, and to find out their initial impression about the haptic navigation system. The guiding questions can be found in Table 5.1. If deemed necessary, follow-up questions were asked. Answers to these questions were used to put the case study results into perspective.

Table 5.1. Questions in the exit-interview.

Exit-interview guiding questions
1. To what extent are you visually impaired? What caused it and since when?
2. How do you usually get around?
3. How do you plan your trips?
4. What kind of difficulties do you encounter moving around?
5. What are you first impressions of the system?
6. Was the belt helpful to you? (Likert scale)
7. Did the belt increase your feeling of safety while moving? (Likert Scale)
8. What would we need to change to make the belt more valuable and useful to you?

5.2.7 Data Collection and Analysis

User position data were recorded by the smartphone sensors, while the experiment was video-recorded from a distance of a few meters by an experimenter. For each path, failure/success to complete the task was evaluated. We then derived the time to completion, as well as the walked distance. For each participant, average speed and path accuracy (deviation from intended path) were also calculated. Furthermore, the directions that were conveyed to the users at any point during the experiment were recorded. The exit interviews were also video-recorded and analyzed.

5.3 Results

In this section, we first examine the effect of real-time localization accuracy on the performance of PVIs using a navigation system. We then describe five cases in which users displayed a variety of strategies with and without input from our wearable pedestrian navigation system.

5.3.1 GPS Accuracy and User Performance with Navigation Systems

Navigation systems are built around three core components: a map, a routing method, and a real-time positioning signal. The accuracy of real-time positioning is often the main limiting factor to performance during navigation. While sighted individuals can compensate for positioning errors, it is not the case for visually impaired individuals. Indeed, sighted individuals can compare the sensor-based localization data on the map, and their actual localization based on integrating visual input such as landmarks with the map. In turn-by-turn navigation for sighted pedestrians, this error compensation makes systems usable even with 10- to 20-meter errors. On the contrary, visually impaired individuals rely for the most part on navigation instructions linked to sensor-based localization. Therefore, in a south-north street, a 10-meter error will cause them to attempt a turn 10 meters before or after the crossing. Consequently, providing information only at turns might be insufficient for PVIs.

The GPS accuracy varied significantly across different areas of the campus (see Figure 5.4), with approximately 5 meters being the most accurate. The architecture of the buildings on campus is very modern. Several buildings have glossy facades that reflected and deflected satellite signals, which caused poor GPS accuracy in some locations. This had a significant impact on the accuracy of the route instructions. Especially in the spaces between buildings the accuracy deteriorated. This phenomenon, called 'urban canyon', is a common problem for many GPS devices in urban environments like the university campus.

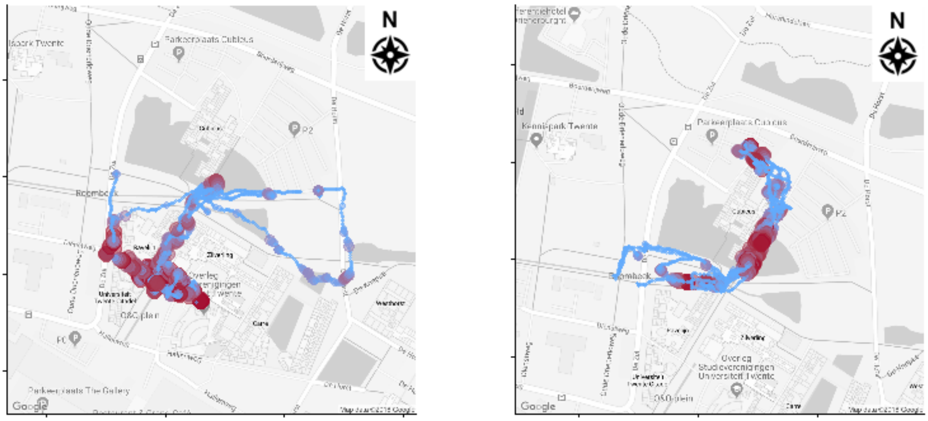


Figure 5.4. Overview of GPS accuracy on the campus. The accuracy is indicated by both the size and the color of the dots on the map. Blue, tiny dots represent relatively high GPS accuracy (close to 5 meters), while the red big dots resemble areas in which the GPS accuracy was relatively poor.

5.3.2 Case Studies

Case 1

The first participant was a female and was between thirty and forty years old. She became fully blind around about seven years prior to the study. For navigation, she used a long cane with assistance from a sighted friend. Other than that, she did not use any navigation aids, since she never goes out alone because she feels she cannot find her way independently. Therefore, she decided not to walk the first route, as she would be helpless while doing so (offered for preparation, regular aids).

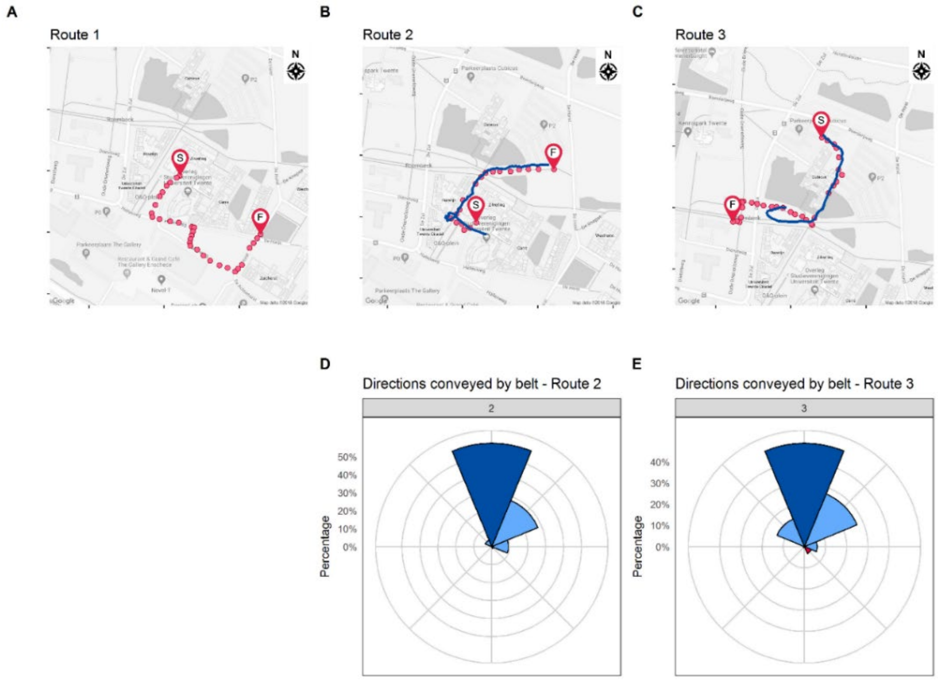


Figure 5.5. Observation figure from participant 1. A, B, and C show the route that the participant was supposed to take (red dots) and the route that the participant actually took (blue line). The marker S is the start of the route, while F marks the finish. D and E show how often each vibrotactor signaled from a particular direction.

Table 5.2. Observation data of participant 1. [a] During Route 1 (offered for preparation, regular aids) participants used their own assistive technologies, while Route 2 (offered for preparation, regular aids, and prototype) and Route 3 (not offered for preparation, prototype only) were supported by the vibrotactile belt. [b] Speed was calculated after removing speed = 0. [c] Was successful, until poor GPS signals caused a detour. * The first route was walked to gain insight in how participants used assistive technology during navigation.

Route [a]	Distance (m)	Time (mm:ss)	Speed (km/h) [b]	Successfully completed
1 *	n/a	n/a	n/a	No
2	353.36	14:36	2.47	Yes
3	324.31	17:31	2.13	No [c]

At the start of the second route (see Figure 5.5B, offered for preparation, regular aids and prototype), it took her some time to get used to the belt. After that, she walked slowly and carefully, constantly concentrating on the signals sent by the belt. As Figure 5.5B shows, the participant took a small detour at the first turn and then took all following turns according to the pre-programmed route. By keeping a low pace, she was able to follow the belt instructions, and thus the path, very accurately. Figure 5E shows the direction conveyed by the vibrotactile belt. This participant received almost only signals from tactors placed on the front of the waist, meaning she was walking straight towards the next waypoint most of the time.

At the starting location of the third route (Figure 5.5C, not offered for preparation, prototype only), the system suffered from GPS inaccuracy, which caused the belt to convey inaccurate information. After getting away from the starting location, the participant encountered an area with foliage right next to a building, affecting the GPS accuracy of the system. She was able to cope with this inaccuracy without any help from the observing researchers. At the point of the route where two crossroads follow in quick succession, she deviated from the pre-programmed route as she walked past the first turn to the right and took the second turn. After she realized she must have missed a turn, she tried to get back on the pre-programmed route. Due to GPS inaccuracy, it was not possible to get back on the pre-programmed route, and she chose to take an alternative route instead. As is visible in the bottom-left corner of Figure 5.5C, she took a small detour to ultimately get back on the pre-programmed route.

Post Hoc Interview

According to this participant the system "helps me to choose the right direction. It really helps to go straight forward. It gives me more security, but you cannot only rely on the system." She was pleased with the location on the body where the signals were conveyed. With vibrotactile signals she felt less overloaded with information than she experienced with auditory information. She believed that the system would be more valuable if turns were more accurately conveyed.

Case 2

The second participant was a male between twenty and thirty years old. He suffered from heavy tunnel vision since birth. He used an identification cane to warn others of his visual impairment. For navigation, he usually used the smartphone application *BlindSquare*.[5] Additionally, he used apps which provide

---

[5] http://blindsquare.com

access to public transport information. Generally, he encounters most navigation problems in heavy traffic areas. This participant was familiar with the haptic belt as he had participated in an earlier study, in which the same belt was used to convey non-navigational information. The participant did not prepare the first route (see Figure 5.6A, offered for preparation, regular aids). Nevertheless, after setting up the GPS coordinates of the destination in *BlindSquare*, he was able to reach the destination of the first route. The participant was comfortable using the software and walked at high speed towards the target waypoints. There were no difficulties finding turns, but his way of turning was very angular, almost robotic. Due to the fact that *BlindSquare* used the shortest path, which was different from the pre-programmed route, the participant deviated quite a lot from the planned route.

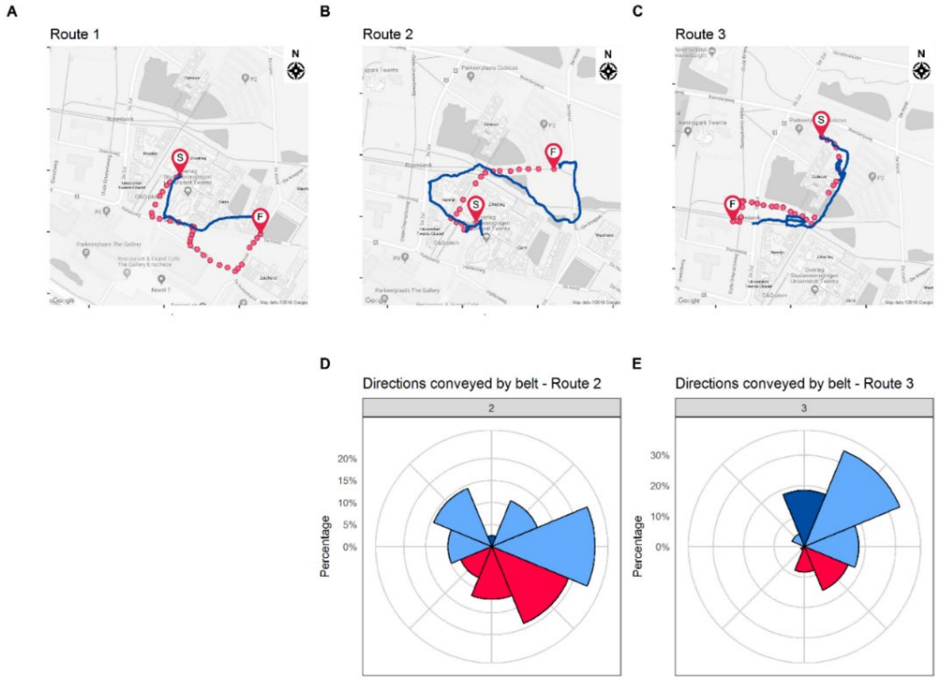


Figure 5.6. Observation figure from participant 2. A, B, and C show the route that the participant was supposed to take (red dots) and the route that the participant actually took (blue line). The marker S is the start of the route, while F marks the finish. D and E show how often each vibrotactor signaled from a particular direction.

Table 5.3. Observation data of participant 2. [a] During Route 1 (offered for preparation, regular aids) participants used their own assistive technologies, while Route 2 (offered for preparation, regular aids, and prototype) and Route 3 (not offered for preparation, prototype only) were supported by the vibrotactile belt. [b] Speed was calculated after removing speed = 0. * The first route was walked to gain insight in how participants used assistive technology during navigation.

Route [a]	Distance (m)	Time (mm:ss)	Speed (km/h) [b]	Successfully completed
1 *	330.05	05:16	3.76	Yes
2	680.93	10:39	4.16	Yes
3	523.84	08:50	3.74	No

During the second route (provided for preparation and navigated with own aids and the prototype, see Figure 5.6B), it was obvious the participant did not trust the information conveyed by the belt. Only at the start of the route, he used the directions conveyed by the belt to get going, after which he took matters in his own hands and just headed wherever he felt he should go to. Figure 5.6D shows that the system mostly conveyed signals in directions other than at 0 degrees, meaning that the participant was almost never faced directly at the next waypoint. The participant seemed to take guidance from the belt every now and then, and with a giant detour he reached the destination of the second route. Interestingly enough, the route taken was neither the route we chose, nor the route calculated by *BlindSquare*.

During the third route (see Figure 5.6C, not offered for preparation, prototype only), the participant started walking with a high pace again, and again only used the belt as loose guidance. Due to this, he missed the first waypoint, and had to take a small detour to get back on the route and continued his way. Then, like the participant in case 1, he encountered difficulties at the double crossroads. Dependent on the vibrotactile belt, he was not able to find the right turn, even after multiple attempts. Therefore, it was decided to stop the experiment. Overall, Figure 5.6E shows that the participant was better able to follow the route than in the previous path.

Post Hoc Interview

The participant found the feedback mechanisms intuitive but had a difficult time to use the signals due to a lack of precision. He suggested a vibration method that would adapt to the users' walking speed, being intermittent while walking and continuous while standing still. Another improvement to the system was to make it more adaptable to route changes. That way, the system would be able to deal with people who take a detour from their initially planned route.

Case 3

The third participant was female and between twenty and thirty years old who became fully blind at fourteen years old after an accident. In daily life she usually uses a long cane and occasionally *BlindSquare*. She feels confident when visiting familiar places, which she can visit independently. In her daily life, she usually never goes to unfamiliar places alone because she feels it can be dangerous sometimes. Thus, she was a bit scared of what would happen during the experiment. After a thorough explanation of the experiment, she was confident enough to try the navigation tasks by herself.

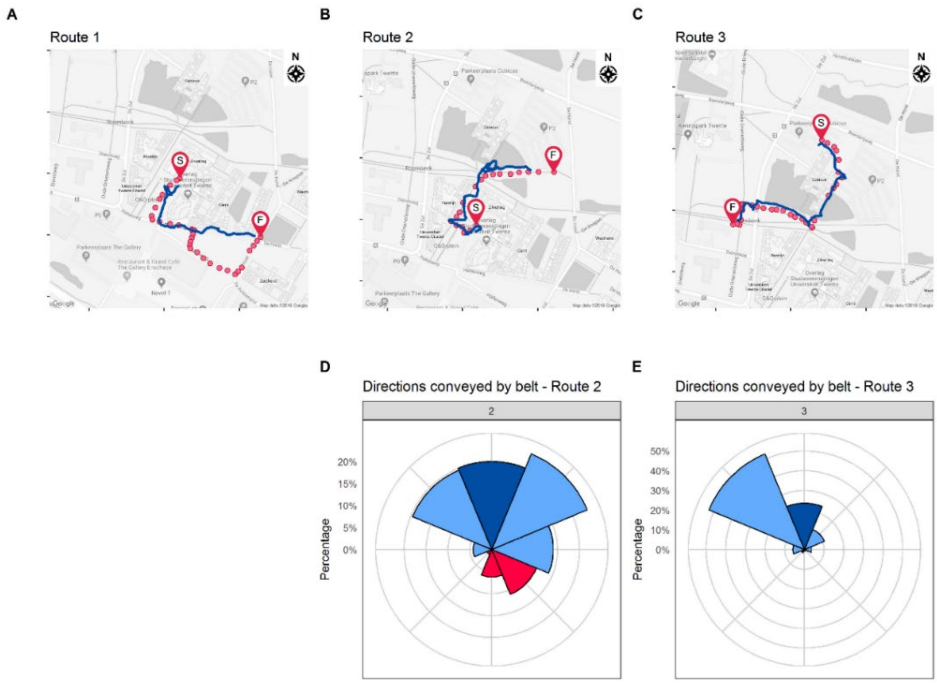


Figure 5.7. Observation figure from participant 3. A, B, and C show the route that the participant was supposed to take (red dots) and the route that the participant actually took (blue line). The marker S is the start of the route, while F marks the finish. D and E show how often each vibrotactor signaled from a particular direction.

Table 5.4. Observation data of participant 3. [a] During Route 1 (offered for preparation, regular aids) participants used their own assistive technologies, while Route 2 (offered for preparation, regular aids, and prototype) and Route 3 (not offered for preparation, prototype only) were supported by the vibrotactile belt. [b] Speed was calculated after removing speed = 0. * The first route was walked to gain insight in how participants used assistive technology during navigation.

Route [a]	Distance (m)	Time (mm:ss)	Speed (km/h) [b]	Successfully completed
1 *	336.76	08:33	2.75	Yes
2	452.62	12:45	2.63	Yes
3	437.93	12:14	2.89	Yes

She did not prepare for the first path. Therefore, we helped her set up *BlindSquare* to calculate a route from the starting point to the first destination. As with participant 2, this meant she took a route that differed from the scheduled path. Problems occurred at the start, as she had trouble finding which direction to depart from the starting point. After explaining the clock-like description of directions, she was able to follow the edge of paths. Detecting corners and making turns took some time, as she was not very experienced with her cane.

For route 2 (see Figure 5.7B), she used *BlindSquare* together with the haptic belt. The belt ensured that she was much better capable of finding the right starting direction. It seemed she trusted *BlindSquare* more than the belt, which led to confusion when the instructions provided by the belt and *BlindSquare* were not in line. Ultimately, she was able to find the route that *BlindSquare* conveyed without help.

In the third route (see Figure 5.7C, not offered for preparation, prototype only), the participant was more carefully paying attention to the belt. Due to this, she was able to follow the belt signal very well and only needed few stops to concentrate on the directions signaled by the belt.

Post Hoc Interview

After initial skepticism, as the participant was an inexperienced independent traveler and was scared to use the belt, all three tasks were completed successfully and with minimal errors. In everyday life, this participant does not travel to unfamiliar environments independently. If she must, she would rather travel with her boyfriend or someone else. The participant used the belt thoroughly and sometimes stopped to take time to interpret the signals and act accordingly. When she used the belt together with her own navigation software (*BlindSquare*), it caused confusion because the system and *BlindSquare* provided incongruent information. After a while she got used to the system and could use it, although she would have liked the system to vibrate more intensely at corners to indicate that she had to turn.

Case 4

The fourth participant was a male between forty and fifty years old. He was fully blind from birth. He uses a guide dog and is an experienced train traveler who commutes between Germany and the Netherlands for his job. As a result, he has extensive knowledge of the Dutch railway map. In addition, he uses *Reisplanner* (an app of the Dutch Railways) to check departure times and location of trains at the station he is at. Furthermore, he uses a *Kapten Mobility* device (Kapsys), which is a GPS based assistive aid that provides spoken navigation information.

He did not prepare for the first route (offered for preparation, regular aids), as the information provided to explore this route was not compatible with his *Kapten*. Therefore, the coordinates of the destination waypoint were entered into his GPS device, which then calculated the shortest route towards the destination which differed from pre-programmed route.

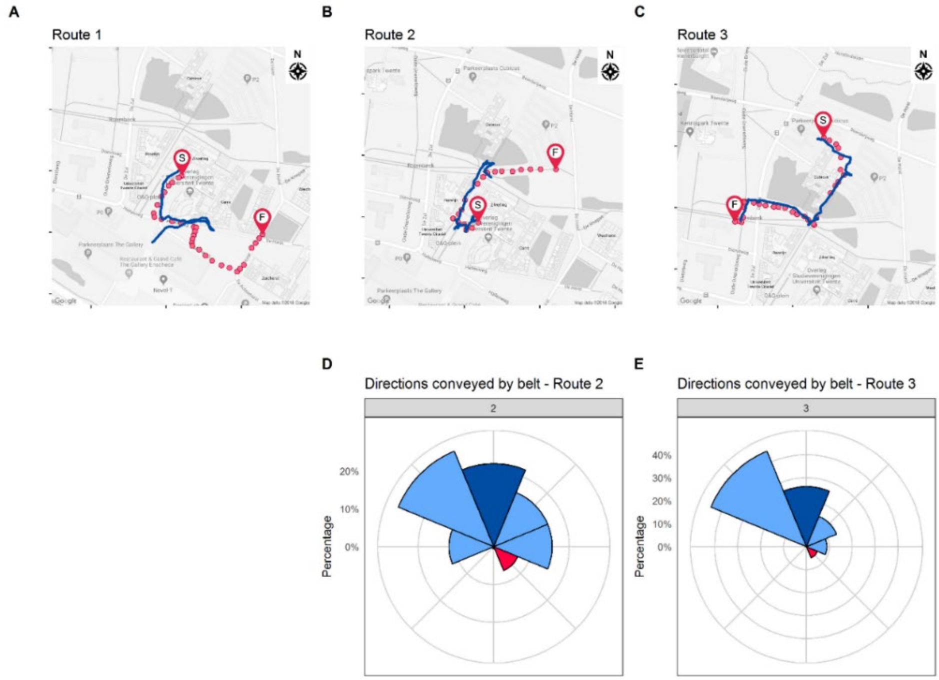


Figure 5.8. Observation figure from participant 4. A, B, and C show the route that the participant was supposed to take (red dots) and the route that the participant actually took (blue line). The marker S is the start of the route, while F marks the finish. D and E show how often each vibrotactor signaled from a particular direction.

Table 5.5. Observation data of participant 4. [a] During Route 1 (offered for preparation, regular aids) participants used their own assistive technologies, while Route 2 (offered for preparation, regular aids, and prototype) and Route 3 (not offered for preparation, prototype only) were supported by the vibrotactile belt. [b] Speed was calculated after removing speed = 0. * The first route was walked to gain insight in how participants used assistive technology during navigation.

Route [a]	Distance (m)	Time (mm:ss)	Speed (km/h) [b]	Successfully completed
1 *	364.2	05:39	3.00	No
2	349.95	08:20	3.72	No
3	439.59	07:19	4.19	Yes

From the start of the first route (see Figure 5.8A, offered for preparation, regular aids), it was clear that the participant was an experienced traveler. He was confident, had good control over his guide dog and was able to maintain a high pace while walking. However, near the end of the route he got confused because his GPS device provided ambiguous information, making him believe that he had missed a turn.

On the second route (see Figure 5.8B, offered for preparation, regular aids and prototype), getting the starting direction straight was fairly difficult due to poor GPS reception. This meant that the directions provided by the belt were poor. Also, there was a small bridge that had to be crossed, that the dog tried to avoid because of the metal grid on the bridge. At the double crossroads he took a small detour. However, he was able to correct for it and got back on the right path. The participant was stopped before reaching the final waypoint, as due to poor GPS signal this waypoint was located in the middle of a body of water. Consequently, this route was not completed successfully.

When navigating the third path, there were two small detours at the park and the following parking lot. At this point, the GPS accuracy was not specific enough to be able to guide him through the tight turns found in this area. As he relied on his guide dog, it was difficult to find the right route out of the many paths available. Another minor hiccup occurred when trying to find the correct turn at the double crossroads. Despite the difficulty he was quickly able to find the correct path. Before reaching the destination waypoint, the participant had to cross a street, which caused issues as he was unable to find the pedestrian route across the street. At that moment, his guide dog corrected him and guided him safely off the street.

In general, the participant was able to follow the belt-signals alright (see Figure 5.8D & 5.8E). The signals conveyed to the user were primarily on the side of the waist and mostly on the left-hand side. This was possibly due to the guide dog guiding the participant. While observing this participant we were able to recognize him as an experienced traveler. He walked reasonably fast and confident even in an unfamiliar environment. He not only has good control over his dog, they are a team and work very well with each other.

Post Hoc Interview

Initially the participant was overwhelmed with the amount of vibrations conveyed by the system. He was positive about the fact that the belt provides information in an intuitive way, without “disturbing your thoughts”. With some training, he believed he would be able to instinctively follow the vibrations. Hence, he was convinced about the usefulness of the system. A problem that requires improvement according to this participant was the inaccuracy of the system. Due to poor GPS coverage, and subsequently poor GPS accuracy, the belt was often vibrating at the wrong place. He experienced a couple of times that he felt that he was on track, but that the belt kept signaling he was wrong, which was confusing.

Case 5

The final participant was a male between forty and fifty years old, who had heavy tunnel vision with a rapid decline over the last few years. He expected to be fully blind within five years. As for now, he had very little vision left and was in orientation and mobility training to learn how to successfully use a long cane. He had no experience with navigating with GPS based devices or applications.

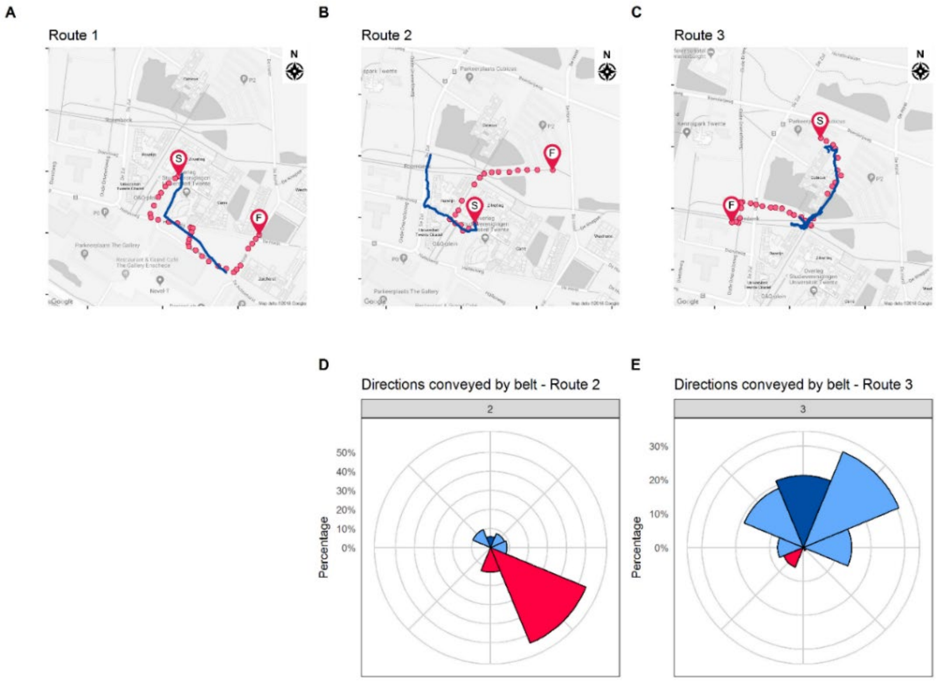


Figure 5.9. Observation figure from participant 5. A, B, and C show the route that the participant was supposed to take (red dots) and the route that the participant actually took (blue line). The marker S is the start of the route, while F marks the finish. D and E show how often each vibrotactor signaled from a particular direction.

Table 5.6. Observation data of participant 5. [a] During Route 1 (offered for preparation, regular aids) participants used their own assistive technologies, while Route 2 (offered for preparation, regular aids, and prototype) and Route 3 (not offered for preparation, prototype only) were supported by the vibrotactile belt. [b] Speed was calculated after removing speed = 0. * The first route was walked to gain insight in how participants used assistive technology during navigation.

Route [a]	Distance (m)	Time (mm:ss)	Speed (km/h) [b]	Successfully completed
1 *	287.04	13:34	2.65	No
2	231.9	04:12	3.35	No
3	416.5	15:49	3.06	No

As many other participants, he did not prepare for the first path. However, he was willing to try Google Maps for the first and second task. For him, it was his first time using Google Maps and he was confused about how to navigate using a map that shows your own position and the surroundings. After explaining how it worked, he was able to navigate towards the destination waypoint, only to stop one waypoint before reaching the actual destination.

On the second path (Figure 5.9B), the participant started well, walking towards the first waypoint immediately. However, upon arrival there he started to ignore the belt signals and walked wherever he felt like. At first, he did not realize that a signal in the back meant to turn around. After an additional explanation he did not seem to care about the belt at all. This is also clearly represented in Figure 5.9E, which shows that the participant almost only received signals from the tactors on his back.

During the third and final path (Figure 5.9C), he decided to walk slower, more careful, and made more stops to try and use the belt signals. However, when walking, he was unable to react to the belt signals quickly enough, which caused him to miss turns at the double crossroads. With his remaining vision he was trying to correct for the system inaccuracy and his delayed reactions, which proved to be very difficult. After four failed attempts to find the correct route on the crossroads it was decided to help him find the correct turn. After continuing his path, he ignored the belt again and was walking towards a water body. Before reaching the water, the task was stopped.

None of the three navigation tasks were successful for this participant. It was clear this participant was not as experienced and comfortable using his cane, compared to the other participants. As he had not finished his orientation and mobility training, he tried to rely as much as possible on his remaining vision. He seemed overwhelmed by the sensory input of the belt, cane, and vision together and was not able (or not willing) to interpret the belt signals continuously.

Post Hoc Interview

Generally, this participant does not go to unfamiliar places out by himself. In his hometown he has no difficulty to find his way, because he knows the exact location of places. However, unfamiliar places are a whole different story. Concerning the belt, he was positive. He was positive about the vibrations conveyed by the belt, with the side note that the information conveyed should be correct (with poor GPS accuracy, it is not). Furthermore, he stated that he was content with the constant feedback from the belt, as it kept him updated even when walking in the correct direction.

5.3.3 Likert-scale questions

In the post hoc interview, the participants were asked, through the use of a Likert-scale, whether they felt that the prototype was helpful to them, and whether it increased their feeling of safety when navigating. Three out of five participants agreed that the prototype was helpful for navigation purposes, while one strongly disagreed with this statement. The other respondent remained neutral. However, the prototype did not seem to have any influence on the feeling of safety of the participants. There was one participant who strongly disagreed with the statements that the prototype enhanced feelings of safety.

Participants were generally positive about the helpfulness of the prototype. All but one were positive about the continuous feedback to show that they are still on track and to know that the system is still functional. Furthermore, the participants were positive about the fact that the haptic feedback did not interfere with hearing and their thoughts, making it easier to concentrate on near field obstacles. The second question about perceived safety was a lot harder to answer. The participants required more time with the belt to become able to address this question properly.

5.4 Discussion

The objective of our study was to explore whether a wearable GPS based navigation system with a vibrotactile belt could support outdoors wayfinding for visually impaired. In addition to the tasks, exit-interviews were conducted to determine whether the participants showed intention to use such a system in real life. Additional goals were (1) to test the system at an early development stage to ensure actionable user feedback before further development and (2) to use prior planning of routes and regular navigation aids as a control condition to compare our system's performance to.

The principle of guiding PVIs in unfamiliar outdoor environments with a vibrotactile belt is feasible and shows promise in real world usage scenarios. The developed prototype could be operated from a smartphone and conveyed wayfinding information through tactile feedback, hereby not hindering other senses required for navigation and social interactions. In line with lab-tests [28,32], and what was demonstrated in practice in later studies [21,27–29,33,34], we found that a waist belt with eight vibrotactors provides sufficient information to guide pedestrian PVIs along a route. The continuous signaling of the belt, which guided users from waypoint to waypoint, was experienced positively by the participants. It made it easier to keep on track and indicated the system was still working.

By conducting case studies in realistic conditions rather than fully controlled conditions, we were able to observe whether pedestrians with a visual impairment could use the system and what problems they encountered in various scenarios. This study setup provided some interesting insights that could not be gained with sighted participants or in a confined space.

Some of these insights were visible at the individual level. Indeed, successful use of the belt seemed very much dependent on individual characteristics. The participants in the study were diverse and their strategies and actions were quite varied. For example, some participants had increased walking speeds while wearing the belt, while others were slower than with their regularly used aids. We also noticed that participants who lost visual function later in life were more at ease in building a spatial map of their surroundings, often using strategies like following edges with their cane in a systematic way. Among the two participants with a guide dog, their level of control over the dog directly affected their ability to process input from the vibrotactile belt, with more control associated with more bandwidth for other inputs. Finally, participants with tunnel vision appeared to rely mainly on their remaining vision, at the expense of other information, however useful it may be to them. Because our number of participants was limited, these insights are not conclusive, rather the basis for future research hypotheses.

Generally, the system appeared to be very intuitive and suitable for users with varying levels of navigation skills. Three participants were able to interpret and use the vibrotactile signals after only a short explanation and were able to follow the first pre-programmed route well. The performance of the two others in the second route (offered for preparation, regular aids, and prototype) seemed to be very much dependent on the willingness of participants to trust the belt over their regular aids. On the third route (not offered for preparation, prototype only) these participants performed similar to the others. These results showed that the device can be used by a broad group of PVIs. Additionally, the results also emphasize the importance of proper training for users of electronic travel aids to use and trust such a system. Users should be well instructed, trained, and made aware of possible issues they might encounter while using the system. This includes learning how to cope with signaling errors and inaccuracies.

The largest problem encountered during the observations was insufficient accuracy of GPS. Smartphones have a known accuracy of 4.9 meters in optimal conditions (i.e. under open sky) [35,36]. The GPS accuracy maps (see Figure 5.4) showed several points on the campus where GPS inaccuracies caused difficulties. The GPS accuracy on the campus was approximately 5 meters at best, which proved to be sometimes not sufficiently inaccurate for visually impaired pedestrians, who need more precise positioning to navigate in complex environments.

System errors are not necessarily perceived as bad, as long as users are able to cope with these errors [37]. Coping strategies can be acquired by training and familiarity with the system. Therefore, it is important that users of the system undergo training for how to recognize and cope with these issues. In addition, the system could inform users in a context-dependent manner depending on GPS accuracy. For instance, in situations where users have to deal with multiple turns/corners/decisions below the current accuracy limit (i.e. 5 meters) users should be warned to pay extra attention or be provided with an audio description of the upcoming situation. Alternatively, users could get updates on GPS signal accuracy.

In the study, pre-programmed routes were followed so that the participants had limited exposure to dangerous situations, such as paths or roads that were shared with cyclists or cars. To do so, for each route a fixed number of waypoints was setup that had to be followed consecutively. This came with the disadvantage that the prototype was unable to recalculate a new route towards the destination once participants missed a waypoint. This caused difficulties for three participants who missed a waypoint while walking the route in which they relied on the prototype only. Rather than calculating a new way towards the destination, the system kept on signaling in the direction of the missed waypoint. This, in combination with the earlier mentioned poor GPS accuracy, led to unnecessary confusion and struggles for participants, who were unable to get back to the missed waypoint. Sometimes they got closer to the destination in the meantime without being aware of it. Besides missing waypoints, there are plenty of other reasons to imagine, which might force a detour from the original route. This was also posed by Williams and colleagues [38], who listed various other reasons why the primary route is unavailable (e.g. construction, weather conditions). Therefore, it is essential for a navigation system for visually impaired pedestrians to be flexible and able to adapt to route changes.

In the current prototype, we use the GPS position in relation to the next waypoint to find the direction, then use the orientation of the smartphone to map this direction into the coordinate system of the person (e.g. translate from 5° north to turn 20° left). The big advantage is that it does work when standing still. It can update all the time when a person turns on the spot. The other possibility of calculating the facing based on current and last GPS reading only works if you move a significant distance, and with the bad GPS that doesn't work well. For the system to provide good direction information using the smartphone sensors, it was supposed to be handheld tilted 30-40˚ facing the user. As some participants used other hand-held assistive technologies or aids such as canes and guide dogs, it was not always possible for them to hold the smartphone. Therefore, one of the researchers walked alongside the participant holding the phone mimicking their every move. In an ideal setup, adding a gyroscope to the front of the belt would solve the issue and mean that a user is not required to hold a phone anymore.

5.4.1 Conclusion

The type of system presented in the current study, which guide users using a vibrotactile belt, seems to be appropriate in addressing wayfinding issues that PVIs encounter in their daily life. We have found that a participant who would never go out to unfamiliar environments was able to find her way around on a campus she had never been to. In general, our participants showed heterogeneity in how they navigate outdoors, and design requirements should take that heterogeneity into account. The compass-like way of presenting directional information on a vibrotactile belt worked intuitively and such belts can be worn unobtrusively, indicating they might be ready for adoption by PVIs already. The most challenging aspect of further development seems to be how to deal with the limits of GPS inaccuracy. It seems to be only a matter of time before more accurate localization methods, such as the Galileo global navigation satellite system developed by ESA with a supposed accuracy of approximately 1 meter [39], are available to the public. Until then, other methods should be explored to allow for more accurate localization of pedestrians, such as Bluetooth beacons which can achieve average accuracy errors of 1.65 meters [40] and might prove to be sufficient to guide pedestrians with visual impairments. This technology has already been investigated to support PVIs with important information in locations with insufficient GPS coverage [41]. In the meantime, training and specific feedback on accuracy to the user can help in places where guiding is less reliable.

5.5 References

1. White B (1970) Perceptual findings with the vision-substitution system. IEEE Trans Man Mach Syst 11:54–58. DOI 10.1109/TMMS.1970.299962
2. Tapu R, Mocanu B, Tapu E (2015) A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments. In: 2014 11th Int. Symp. Electron. Telecommun. ISETC 2014 - Conf. Proc. IEEE, pp 1–4
3. Giudice NA, Legge GE (2008) Blind navigation and the role of technology. Eng Handb Smart Technol Aging, Disabil Indep, pp 479–500. DOI 10.1002/9780470379424.ch25

4. Gori M, Cappagli G, Tonelli A, et al (2016) Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. Neurosci Biobehav Rev 69:79–88. DOI 10.1016/j.neubiorev.2016.06.043
5. Hakobyan L, Lumsden J, O'Sullivan D, Bartlett H (2013) Mobile assistive technologies for the visually impaired. Surv Ophthalmol 58:513–528. DOI 10.1016/j.survophthal.2012.10.004
6. Quiñones P, Greene T, Yang R, Newman M (2011) Supporting visually impaired navigation: A needs-finding study. Chi2011, pp 1645–1650. DOI 10.1145/1979742.1979822
7. Schinazi VR, Thrash T, Chebat D-RR (2016) Spatial navigation by congenitally blind individuals. Wiley Interdiscip Rev Sci 7:37–58. DOI 10.1002/wcs.1375
8. Bourne RRA, Flaxman SR, Braithwaite T, et al (2017) Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Heal 5:e888–e897. DOI 10.1016/S2214-109X(17)30293-0
9. Wentzel J, Velleman E, Van der Geest T (2016) Wearables for all. In: Proc. 13th Web All Conf. - W4A '16. ACM Press, New York, New York, USA, pp 1–4
10. Schölvinck AFM, Pittens CACM, Broerse JEW (2017) The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 111:201–217
11. Ahmetovic D, Gleason C, Ruan C, et al (2016) NavCog: A navigational cognitive assistant for the blind. In: Proc. 18th Int. Conf. Human-Computer Interact. with Mob. Devices Serv. - MobileHCI '16. ACM Press, New York, New York, USA, pp 90–99
12. Sato D, Kitani K (2017) NavCog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In: Proc. 19th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS '17. ACM Press, New York, New York, USA, Baltimore, pp 270-279
13. Velázquez R, Pissaloux E, Rodrigo P, et al (2018) An outdoor navigation system for blind pedestrians using GPS and tactile-foot feedback. Appl Sci 8:578. DOI 10.3390/app8040578

14. Meier A, Matthies DJC, Urban B, Wettach R (2015) Exploring vibrotactile feedback on the body and foot for the purpose of pedestrian navigation. In: Proceedings of the 2nd international Workshop on Sensor-based Activity Recognition and Interaction (iWOAR '15), pp 1–11. DOI 10.1145/2790044.2790051
15. Dobbelstein D, Henzler P, Rukzio E (2016) Unconstrained pedestrian navigation based on vibro-tactile feedback around the wristband of a smartwatch. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16), pp 2439–2445. DOI 10.1145/2851581.2892292
16. Spiers AJ, Dollar AM (2016) Outdoor pedestrian navigation assistance with a shape-changing haptic interface and comparison with a vibrotactile device. In: 2016 IEEE Haptics Symp. IEEE, pp 34–40
17. Loomis JM, Golledge RG, Klatzky RL (1998) Navigation system for the blind: auditory display modes and guidance. Presence Teleoperators Virtual Environ 7:193–203. DOI 10.1162/105474698565677
18. Roentgen UR, Gelderblom GJ, De Witte LP (2011) Users' evaluations of four electronic travel aids aimed at navigation for persons who are visually impaired. J Vis Impair Blind 105:612–623. DOI 10.1177/0145482X1110501008
19. Ran L, Helal S, Moore S (2004) Drishti: an integrated indoor/outdoor blind navigation system and service. In: Second IEEE Annu. Conf. Pervasive Comput. Commun. 2004. Proc. IEEE, pp 23–30
20. Wilson J, Walker BN, Lindsay J, et al (2007) SWAN: System for wearable audio navigation. In: 2007 11th IEEE Int. Symp. Wearable Comput. IEEE, pp 1–8
21. Kärcher SM, Fenzlaff S, Hartmann D, et al (2012) Sensory augmentation for the blind. Front Hum Neurosci 6:37. DOI 10.3389/fnhum.2012.00037
22. Tsukada K, Yasumura M (2004) ActiveBelt: Belt-type wearable tactile display for directional navigation. In: Davies N, Mynatt ED, Siio I (eds) UbiComp 2004: Ubiquitous Computing. UbiComp 2004. Lecture Notes in Computer Science, vol 3205. Springer, Berlin, Heidelberg. DOI 10.1007/978-3-540-30119-6_23

23. Pielot M, Henze N, Heuten W, Boll S (2008) Evaluation of continuous direction encoding with tactile belts. In: Haptic Audio Interact. Des. Springer, Berlin, Heidelberg, pp 1–10
24. Heuten W, Henze N, Boll S, Pielot M (2008) Tactile wayfinder: A non-visual support system for wayfinding. Proc 5th Nord Conf Human-computer Interact Build Bridg, pp 172–181. DOI 10.1145/1463160.1463179
25. Adame MR, Jing Yu, Moller K, Seemann E (2013) A wearable navigation aid for blind people using a vibrotactile information transfer system. In: 2013 ICME Int. Conf. Complex Med. Eng. IEEE, pp 13–18
26. Cosgun A, Sisbot EA, Christensen HI (2014) Evaluation of rotational and directional vibration patterns on a tactile belt for guiding visually impaired people. In: 2014 IEEE Haptics Symp. IEEE, pp 367–370
27. Flores G, Kurniawan S, Manduchi R, et al (2015) Vibrotactile guidance for wayfinding of blind walkers. IEEE Trans Haptics 8:306–317. DOI 10.1109/toh.2015.2409980
28. Van Erp JBF, Van Veen HAHC, Jansen C, Dobbins T (2005) Waypoint navigation with a vibrotactile waist belt. ACM Trans Appl Percept 2:106–117. DOI 10.1145/1060581.1060585
29. Dura-Gil JV, Bazuelo-Ruiz B, Moro-Perez D, Molla-Domenech F (2017) Analysis of different vibration patterns to guide blind people. PeerJ 5:e3082. DOI 10.7717/peerj.3082
30. Pielot M, Poppinga B, Heuten W, Boll S (2011) A tactile compass for eyes-free pedestrian navigation. In: Campos P, Graham N, Jorge J, Nunes N, Palanque P, Winckler M (eds) Human-Computer Interaction – INTERACT 2011. INTERACT 2011. Lecture Notes in Computer Science, vol 6947. Springer, Berlin, Heidelberg. DOI 10.1007/978-3-642-23771-3_47
31. Pielot M, Poppinga B, Heuten W, Boll S (2012) PocketNavigator. In: Proc. 2012 ACM Annu. Conf. Hum. Factors Comput. Syst. - CHI '12. ACM Press, New York, New York, USA, pp 3131
32. Cholewiak RW, Brill JC, Schwab A (2004) Vibrotactile localization on the abdomen: Effects of place and space. Percept Psychophys 66:970–987. DOI 10.3758/BF03194989

33. Heuten W, Henze N, Boll S, Pielot M (2008) Tactile wayfinder: A non-visual support system for wayfinding. Proc 5th Nord Conf Human-computer Interact Build Bridg, pp 172–181. DOI 10.1145/1463160.1463179
34. Xiao JZ, Joseph SL, Zhang XC, et al (2015) An assistive navigation framework for the visually impaired. IEEE Trans Human-Machine Syst 45:635–640. DOI 10.1109/THMS.2014.2382570
35. The National Coordination Office for Space-Based Positioning Navigation and Timing GPS.gov: GPS Accuracy. https://gps.gov/systems/gps/performance/accuracy/. Accessed 6 Aug 2018
36. Van Diggelen F, Enge P (2015) The world's first gps mooc and worldwide laboratory using smartphones. In: Proc. 28th Int. Tech. Meet. Satell. Div. Inst. Navig. (ION GNSS+ 2015). Tampa, Florida, pp 361–369
37. Abdolrahmani A, Easley W, Williams M, et al (2017) Embracing errors: Examining how context of use impacts blind individuals' acceptance of navigation aid errors. In: Proc 2017 CHI Conf Hum Factors Comput Syst - CHI '17, pp 4158–4169. DOI 10.1145/3025453.3025528
38. Williams MA, Hurst A, Kane SK (2013) "Pray Before You Step out": Describing personal and situational blind navigation behaviors. In: Proc 15th Int ACM SIGACCESS Conf Comput Access. DOI 10.1145/2513383.2513449
39. ESA (2011) Galileo Open Service (OS). https://gssc.esa.int/navipedia/index.php/Galileo_General_Introduction. Accessed 14 Aug 2018
40. Sato D, Kitani K (2017) NavCog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In: Proc. 19th Int. ACM SIGACCESS Conf. Comput. Access. - ASSETS '17. ACM Press, New York, New York, USA, Baltimore, pp 270-279
41. Van der Bie J, Visser B, Matsari J, et al (2016) Guiding the visually impaired through the environment with beacons. Proc 2016 ACM Int Jt Conf Pervasive Ubiquitous Comput Adjun - UbiComp '16, pp 385–388. DOI 10.1145/2968219.2971387

6 Discussion

6.1 Discussion

The aim of the research presented in this dissertation was to explore the opportunities that novel technologies, such as wearable consumer electronics, offer to support PVIs with activities of daily life. These devices were particularly interesting as they combined the essential wearability, processing power, connectivity, sensors, and multimodality to benefit PVIs. By following a human-centered design approach, we have worked towards two wearable sensory substitution devices that should support PVIs when encountering two specific issues in their daily lives. One supports PVIs to recognize facial expressions of emotions to enhance their social interactions, and one assists them to find their way in unfamiliar surroundings. Throughout the project, we had interesting experiences, encountered challenging issues, and collected some worthwhile insights that are worth sharing and to be reflected on.

6.1.1 Human-centered design

Throughout our studies, we have continuously followed a human-centered design approach. According to the International Organization for Standardization [1], the aim of such an approach is to ensure a usable and useful system by putting an emphasis on user needs and requirements. By doing so, the idea is that the developed system is beneficial for the tasks its intended users wish to use the system for, and also ensures that the system is accessible and safe to use. In this project, we intended to develop a system that was easily accepted and adopted by PVIs and thus we involved representatives from this target group at all studies conducted within the project.

In all our studies, we have combined quantitative and qualitative methods to ensure the voice of PVIs was heard in the project. On the one side, we valued the objective measurements of technical performance. On the other side, the opinions of participants proved to be very valuable too. For example, during the study described in chapter 3, the technical performance of the system gave not much reason for optimism, as the emotion recognition did not perform as well as we expected. However, the participants were enthusiastic about the system and saw potential for it if the performance improved. As a result, we concluded that the concept of the system was feasible, and that PVIs are likely to be receptive to such a system. It seems to be only a matter of time before similar systems reach satisfactory performance and can become of value in the daily lives of PVIs.

During the project, we have gained several insights that rely heavily on the input of the target group representatives. In fact, the decision what to develop in later stages of the project was derived from the study presented in chapter 2. During the project, we continuously checked whether the developed systems were still wished for by PVIs and where our focus should be for future developments. For example, in the study described in chapter 3, the participants were happy with the size of the vibrotactile belt. However, they emphasized that the vibrations were uncomfortably strong and explained an urgency to develop a smaller camera, as the one used was too heavy and uncomfortable.

Due to the chosen human-centered approach, a strong connection was established between the researcher and the target group, and a continuous tension between human-centered and scientific research emerged. It was often difficult to distinguish the scientific (discovering opportunities for sensory substitution in consumer electronics) and societal (working towards an assistive technology) goals that existed alongside each other in the project. Due to the project setup with strong involvement of end-users, the data and information were collected directly from the target group of the technology. Conducting the research creates expectations amongst participants, who sometimes even asked for an estimated time-to-market of the prototype. In a way, the impact of the participants on the scientific research was much bigger than the impact that the results of the research could have on their daily lives.

Unfortunately, the research never got to the point in which it was able to implement the technology in the daily life of PVIs, which would have been the final test for the technology before finalizing the prototype. More development, particularly on emotion recognition software, is required before it would make sense to test the prototype in daily life.

6.1.2 Vibrotactile sensory substitution and multimodality

There are various modalities to convey information to users of technology. In the case of SSDs, a very important design decision must be made concerning how information is conveyed between the device and its users. First, when dealing with PVIs, a decision must be made whether to convey information through audio or haptics. In this project, we chose to use haptics, as we observed that PVIs are very much dependent on their ears to perceive the world surrounding them, particularly in the use cases we were intending to assist them with. Interfering with hearing could lead to dangerous situations while navigating, and to inconvenience during social interactions.

Ideally, the system would provide the tactile information so quickly after the auditory and remaining visual information that it seems to be perceived at the same time. In the studies described in the third and fourth chapter, this was not yet achieved for two reasons: accuracy and time. During the study in presented in chapter 3, emotion recognition by the system was highly accurate, yet there was a purposeful delay of 550 milliseconds between the audiovisual and the tactile stimulus. This meant that sometimes the emotion was already recognized by the user before any tactile signals were given, which created an odd experience. In the more realistically setup study presented in chapter 4, the SSD delay was shortened. However, the emotion recognition accuracy was poorer. Consequently, situations occurred in which the information conveyed to the user did not match the sounds or haptic signals they perceived at the same time. For example, a participant in the study presented in the fourth chapter, stated she felt the SSD signaling a disgusted facial expression, which was not in line with how she perceived the conversation or any auditory cues that were present. Nevertheless, we believe that if we were to increase the accuracy of the system, the speed with which information is processed and conveyed is enough, even for high demanding tasks such as social interactions.

A concern that might be raised when it comes to tactile feedback is its relatively small sensory bandwidth. The visual sense has a much larger bandwidth than the auditory and somatosensory sense, which seriously constrains the nature and amount of information that can be *translated* from one sense to another. To enable access to relevant information, SSDs should be task-focused, meaning that only limited information (that which is required for a specific task) should be conveyed to its users. In this dissertation, it was decided to reduce the to be conveyed information to a minimum. However, because it was so easily interpreted by the users of the device, even the limited amount of information proved to be valuable. This showed that despite the limited information that can be conveyed through vibrotactile signals, the information that can be conveyed can be very meaningful for PVIs. Since vibrotactile signals can be very meaningful and useful for PVIs and leave eyes and ears free, the (vibro)tactile modality show large potential to assist PVIs in their daily lives.

6.1.3 Pros and cons of using of existing technology

One of the cornerstones of the research project was to make use of already, or almost, available technologies. By doing so, we believed we would be able to develop solutions that might be beneficial for PVIs and which could be relatively easy brought to the market to support PVIs in their daily lives. However, because we were developing for a group with special needs when it comes to technology, we encountered considerable challenges during the research. There were various moments where we have encountered the boundaries of the capabilities that every day technology has. Initially, the research was inspired by the opportunities that smart glasses could have for PVIs. Quickly after the start, we found out that the available smart glasses ran insufficient battery power to use it for processor heavy applications – especially since we were to develop solutions that should be able to support PVIs during their daily lives. As a result, the prototype used during the first experiments was composed of a regular webcam, a tablet, and the vibrotactile belt. While still portable, this meant the prototype was significantly bigger and far less unobtrusive than what we, and most likely PVIs, wished for. Over the course of the project, we were able to develop a smaller and more wearable prototype by decreasing the camera size and substituting the tablet for a smartphone. Nevertheless, some issues remained even at the end of the research project. In both the emotion recognition and the navigation system, inaccuracy of the software resulted in issues that are likely to be problematic for user adoption in the long run. The studies described in chapter 3 and particularly 4 showed that the emotion recognition software was not yet able to recognize emotions from facial expressions in a level that would benefit PVIs in their daily lives, while the study described in chapter 5 showed that GPS accuracy is insufficient to support PVIs with only vibrotactile feedback.

6.1.4 Conveying facial expression information

Over the course of the research projects of this dissertation, two studies have been conducted aimed at making facial expressions of emotions accessible for PVIs. During the first study described in **chapter 2**, participants were wearing the first prototype of the SSD and were confronted with sets of previously validated pictures and videos of actors expressing a variety of emotions, part of which were accompanied with validated auditory cues of emotions. The early prototype consisted of a vibrotactile belt, a tablet running FaceReader [2,3], and a webcam mounted on a baseball cap. The study showed that persons can learn how to interpret the vibrotactile signals conveyed by the system and improved significantly in their ability to determine the correct emotions from the facial expressions shown on the pictures and videos. However, the study setup was highly artificial and therefore it was impossible to generalize the findings to real-world conditions [4]. Thus, a closer-to-realism study was deemed necessary to further validate the usefulness of the SSD.

A follow-up study, presented in chapter 3, was conducted with an updated version of the SSD and in a setting which more closely resembled a realistic conversation between two persons. Before the study commenced, the SSD was updated to improve its wearability by exchanging the webcam on a cap with a small USB camera, to be mounted on spectacles. The PVIs were using the new prototype during a conversation with an actor. The facial expressions of the actor were as real as possible. The study showed great promises for SSDs in the real world, as PVIs were able to keep the head-mounted camera oriented in the direction of the conversation partner, making it a stable video feed for video processing software to work with. On the downside, the only emotion that was satisfactorily recognized was happiness. As for the other basic emotions, the quality of the current recognition software was insufficient for this real-life condition.

The tactile information conveyed to users in the research presented in this dissertation was deliberately made as easy as possible to explore whether it was possible for PVIs to use them in social interactions. Therefore, it was decided to convey the six universal emotions of Ekman through six vibrotactors that were placed on the waist at a distance that exceeded the spatial acuity of the skin there [5,6]. In the literature that is more focused on the tactile interfaces, we have seen various methods to convey information about facial expressions of emotions that seem to be more complex yet can provide more detailed information to PVIs [7–9]. One of the elements that might be interesting for future research, is whether it is possible to provide more detailed information to PVIs following the principle of facial action coding system, as was presented in a controlled environment by McDaniel and colleagues [9], and whether this is feasible in real life applications. By doing so, one gives PVIs more room to interpret facial expressions, such as a raised eyebrow or tightened lips. Our current system does not leave such room for interpretation of users, as the computer algorithms determine which emotion is associated to a facial expression, after which this emotion is conveyed to the user. Naturally, PVIs should be consulted about their wishes for the complexity of the information they would wish to receive.

6.1.5 Privacy issues

A different theme that was not yet discussed in the papers, yet was raised often during the project, and should therefore not be neglected, is the privacy and handling of the information collected by the system. If one is walking around with a camera mounted on spectacles, this might lead to privacy concerns for the people surrounding them, for they do not know whether and what the camera is filming and for what purpose. Unsurprisingly, when the first smart glasses were about to hit the consumer market in great numbers, a lot of security and privacy issues were raised. However, if future applications of such a device are proven to be beneficial for PVIs in their daily lives, should we deny access to it because of these privacy concerns? Especially in the context of an inclusive and accessible society, denying PVIs such assistive aids would not make a lot of sense. Therefore, a societal discussion is needed before technologies such as the one presented in this dissertation are ready to hit the consumer target. One of the biggest questions should be whether the accessibility of information of persons with a visual impairment outweighs the privacy of others in public space. Even if we know that the camera is not continuously filming, and no recordings are saved, how can we ensure others know about this purpose and are fine with it? Legislation is required to ensure what is possible and what is not, before technologies like the one presented in this dissertation can be successfully implemented in the daily lives of PVIs.

6.1.6 Conveying navigation information

For navigation purposes, the vibrotactile belt was also used in a compass-like way, which was used to guide PVIs through an unfamiliar environment. Despite a lot of variation between the participants, the signals conveyed were intuitive and all participants were able to navigate in unfamiliar environment. The results were achieved only after brief instructions. More extensive training is likely to lead to better results. The GPS accuracy, which was approximately five meters at best, was often insufficient for pedestrians with visual impairments to determine when to turn, particularly in places with multiple route options. Additionally, the system used in our study was very rigid. Users were expected to follow the desired route and were guided back to the initial route after deviating from it. As such, it was not designed in a way that it could deal with system and user errors. Earlier research showed that errors are accepted by users of navigation aids as long as dealing with the errors does not draw much attention from bystanders [10]. However, the system design resulted in participants somewhat wandering around trying to get back to the previous waypoint. Usage of adaptable routes, an easy fix which would recalculate a new route after a wrong turn has been taken, could make our navigation system for PVIs much more effective – especially since there are many reasons for regular routes to be unavailable, such as construction works or weather conditions [11]. Additionally, there is reason to believe that the major problem faced by many GPS navigation aids for PVIs, including ours, being the GPS inaccuracy, will be history in a few years thanks to developments in localization technologies [12]. Finally, a challenge remains to provide PVIs with information about their surroundings [13], something that was not covered in our studies, yet is wished for by PVIs. Currently, plenty of navigation aids rely on audio to convey such information. As mentioned earlier, doing so might block hearing, which can cause dangerous situations. However, the lack of accuracy without any further instructions might also not be an ideal situation, as PVIs do not get any spatial awareness from such a device. Bone conducted audio cues for detailed instructions, with a tactile compass like the one presented in our work, might proof to be the solution for safe guidance of PVIs.

6.1.7 Diverse target group

During the project, we have encountered various challenges due to the size and diversity of the population of PVIs. Because of the diversity in the population of PVIs, all the conducted studies deliberately included a mixed sample, including persons with a mixture of age of the onset of the visual impairment as well as severity. Due to the limited size and the large diversity within the samples included in our experiments, unfortunately we could not compare performances based on individual characteristics such as congenital visual impairment. Consequently, various reviewers demanded increased sample sizes to be able to compare performance between various clusters of PVIs, or to improve the generalizability of the studies. Nevertheless, we have experienced that the qualitative feedback received from participants proved to be valuable for the aim of our research, which was to develop vibrotactile wearables that suit the needs and wishes of PVIs. Despite the small sample sizes, the collected information saturated quickly, and because of the diversity within the sample, we were confident that our findings and design decisions would benefit the development of a system that would eventually benefit the PVI community in their daily lives.

6.2 Conclusion

Inspired by new many new developments in consumer electronics, particularly smart glasses and other wearables, we investigated how PVIs could be supported by such technologies during activities of daily life. We have identified needs amongst PVIs, especially related to nonverbal communication and navigation in unknown environments. We developed a vibrotactile wearable system to convey information about the emotion of faces and navigation directions. Our studies showed that participants can learn, interpret, and use the vibrotactile signals conveyed by the device easily and that they are enthusiastic about the concept. However, in the various studies, technical and usability issues were identified that need to be addressed in future developments. Looking back at the research project, we can proudly say that we have come a lot closer towards a working SSD to support PVIs with activities of daily life. We have gained valuable insights in the possibilities that tactile feedback offers to support PVIs. None of the work presented in this dissertation could have been done without the commitment of our participants, for which we are very grateful.

6.3 References

1. ISO (2019) Systems, Ergonomics of human-system interaction — Part 210: Human-centred design for interactive. https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:en. Accessed 18 Dec 2019
2. Van Kuilenburg H, Wiering M, Model MJDUA, et al (2008) Facereader 6.1 – How do they do it?

3. Noldus FaceReader. http://www.noldus.com/human-behavior-research/products/facereader?gclid=CNPm3enUrM8CFYYcGwodyp0BJQ. Accessed 26 Sep 2016
4. Spence C (2014) The skin as a medium for sensory substitution. Multisens Res 27:293–312. DOI 10.1163/22134808-00002452
5. Jóhannesson ÓI, Hoffmann R, Valgeirsdóttir VV, et al (2017) Relative vibrotactile spatial acuity of the torso. Exp Brain Res 235:1–11. DOI 10.1007/s00221-017-5073-6
6. Van Erp JBF (2005) Vibrotactile spatial acuity on the torso: Effects of location and timing parameters. In: First Jt. Eurohaptics Conf. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst. IEEE, pp 80–85
7. Krishna S, Bala S, McDaniel T, et al (2010) VibroGlove: an assistive technology aid for conveying facial expressions. Proc SIGCHI Conf Hum Factors Comput Syst Ext Abstr, pp 3637–3642. DOI 10.1145/1753846.1754031
8. Krishna S, Little G, Black J, Panchanathan S (2005) A wearable face recognition system for individuals with visual impairments. In: Proc 7th Int ACM SIGACCESS Conf Comput Access - Assets '05 106. DOI 10.1145/1090785.1090806
9. McDaniel T, Tran D, Chowdhury A, et al (2019) Recognition of tactile facial action units by individuals who are blind and sighted: A comparative study. Multimodal Technol Interact 3:32. DOI 10.3390/mti3020032
10. Abdolrahmani A, Easley W, Williams M, et al (2017) Embracing errors: Examining how context of use impacts blind individuals' acceptance of navigation aid errors. In: Proc 2017 CHI Conf Hum Factors Comput Syst - CHI '17, pp 4158–4169. DOI 10.1145/3025453.3025528
11. Williams MA, Hurst A, Kane SK (2013) "Pray Before You Step out": Describing personal and situational blind navigation behaviors. In: Proc 15th Int ACM SIGACCESS Conf Comput Access. DOI 10.1145/2513383.2513449
12. ESA (2011) Galileo Open Service (OS). https://gssc.esa.int/navipedia/index.php/Galileo_General_Introduction. Accessed 14 Aug 2018

13. Schölvinck AFM, Pittens CACM, Broerse JEW (2017) The research priorities of people with visual impairments in the Netherlands. J Vis Impair Blind 111:201–217

Publications

Papers

Buimer HP, Zhao Y, Wentzel J, Van der Geest TM, Van Wezel RJA (2019) Setting the development and innovation agenda for technology applications for persons with visual impairments or blindness: A user-centered approach. Available in open repository ResearchGate. DOI 10.13140/RG.2.2.16247.04002

Buimer HP, Bittner M, Kostelijk T, Van der Geest TM, Nemri A, Van Wezel RJA, Zhao Y (2018) Conveying facial expressions to blind and visually impaired persons through a wearable vibrotactile device. PLoS ONE 13(3):e0194737. DOI 10.1371/journal.pone.0194737

Buimer HP, Schellens R, Kostelijk T, Nemri A, Zhao Y, Van der Geest TM, Van Wezel RJA (2019) Opportunities and Pitfalls in Applying Emotion Recognition Software for Persons With a Visual Impairment: Simulated Real Life Conversations. JMIR Mhealth Uhealth 7(11):e13722. DOI 10.2196/13722

Buimer HP, Nossek F, Nemri A, Van der Geest TM, Bauer G, Van Wezel RJA (n.d.) Guiding Pedestrians with Visual Impairments in Unfamiliar Environments with Vibrotactile Feedback: Five Case Studies. *To be submitted*

Outside this dissertation

Buimer HP, Tabak M, Van Velsen L, Van der Geest TM, Hermens H (2017) Exploring Determinants of Patient Adherence to a Portal-Supported Oncology Rehabilitation Program: Interview and Data Log Analyses. JMIR Rehabil Assist Technol 4(2):e12. DOI 10.2196/rehab.6294

Conference contributions

Buimer HP, Van der Geest TM, Nemri A, Schellens R, Van Wezel RJA, Zhao Y (2017) Making Facial Expressions of Emotions Accessible for Visually Impaired Persons. Poster session presented at 19th International ACM SIGACCESS Conference on Computers and Accessibility, Baltimore, United States, pp 331-332. DOI 10.1145/3132525.3134823

Buimer HP (2017) Making facial expressions accessible to blind and visually impaired persons using a wearable vibrotactile device. Oral presentation at 12th International Congress of the International Society for Low Vision Research and Rehabilitation, Vision 2017

Buimer HP, Bittner M, Kostelijk T, Van der Geest TM, Van Wezel RJA, Zhao Y (2016) Enhancing emotion recognition in VIPs with haptic feedback. In C. Stephanidis (Ed.), International Conference on Human-Computer Interaction, HCI 2016 (pp. 157-163). (Communications in Computer and Information Science; Vol. 618). Switzerland: Springer. DOI 10.1007/978-3-319-40542-1_25

Van der Geest TM, Buimer HP (2015) User-centered priority setting for accessible devices and applications. In: Mensch und Computer 2015 - Workshopband (pp. 383-390). (Mensch & Computer – Tagungsbände / Proceedings ; Vol. 2015). Walter De Gruyter.

Dankwoord

Arnhem, 20 december 2020

Het was december 2014 – op dat moment werkte ik druk aan mijn thesis bij Roessingh R&D om mijn master Communication Studies aan de Universiteit Twente (UT) af te ronden – toen ik solliciteerde op een baan als junior onderzoeker bij de UT. Ik werd enthousiast door de omschrijving van het onderzoek naar toepassingsmogelijkheden van *smart glasses*, waarbij eindgebruikers centraal zouden staan (het schijnt zelfs dat ik deze slimme bril in mijn slaap weleens genoemd heb). In mijn sollicitatiebrief noemde ik het voordeel van de tijdelijke aard van het contract. Ik overwoog toentertijd om in de toekomst een promotieonderzoek te doen, maar had nog wel mijn twijfels of een voltijdbaan als onderzoeker bij mij zou passen. Uiteindelijk startte ik in maart 2015 als student-assistent bij de vakgroep Biomedical Signals & Systems (BSS) en begon ik in mei 2015 officieel als junior onderzoeker op het onderzoek dat uiteindelijk zou uitmonden in de dissertatie die nu voor u ligt. Op het moment van schrijven van dit dankwoord is 2020 alweer bijna voorbij en is het einde van mijn traject als promovendus in zicht. Dankbaar ben ik voor alle mensen die ik in de afgelopen jaren heb mogen ontmoeten, die mij door alle ups en downs van het promoveren hebben geholpen en die mij hebben laten zien dat er helemaal niet zoveel mis is met een baan als onderzoeker.

Om te beginnen, zou ik graag alle mensen willen bedanken die hebben deelgenomen en bijgedragen aan de studies. Een belangrijk uitgangspunt tijdens onze onderzoeken was om de mens centraal te stellen in de ontwikkeling van applicaties voor slimme brillen voor blinde en slechtziende personen. Om deze reden is het onderzoeksproject gestart met het ophalen van wensen en uitdagingen die mensen met een visuele beperking in het dagelijks leven tegenkomen, tijdens de ZieZo-beurs in 2015. Ook in alle opvolgende studies hebben de beoogde eindgebruikers een cruciale rol gespeeld. Ik denk dat ik namens het hele onderzoeksteam spreek als ik zeg dat we zonder jullie hulp, enthousiasme en kritische feedback nooit de stappen hadden kunnen maken die we in de afgelopen jaren hebben gemaakt!

Dan de mensen uit dat team van onderzoekers waarvan ik onderdeel heb uit mogen maken. Richard, heel erg veel dank voor de tijd en energie die je in mij

hebt gestoken. Je deur stond altijd open voor vragen en hulp wanneer dat nodig was. Ik vond het erg fijn dat je mij heel vrij hebt gelaten om zelf te ontdekken wat onderzoek doen inhield. Op momenten dat ik het even niet meer kon overzien of gestrest was over of alles wel liep zoals het zou moeten, was je een baken van rust en gaf je me altijd het gevoel en vertrouwen dat alles wel goed zou komen. In de zomer van 2016 gaf je mij de kans om van het onderzoek dat we deden in Twente een promotietraject te maken, wat ervoor zorgde dat ik uiteindelijk bij de Radboud Universiteit terecht kwam. Door deze stap heb ik weer veel nieuwe mensen leren kennen, ben ik naar Arnhem verhuisd en heb ik kennis kunnen maken met biofysisch onderzoek. Dit zijn dingen die ik nooit had verwacht toen ik in april 2015 mijn masterdiploma Communication Studies ontving. Bedankt voor deze verrassende wending in mijn leven en al je hulp gedurende de afgelopen jaren.

Thea, naast de 'usual suspects' ben jij denk ik wel de meest stabiele factor geweest van mijn tijd als (pre)masterstudent, onderzoeker en promovendus. Mijn premastercertificaat ontving ik uit jouw handen, in mijn master volgde ik meerdere colleges bij jou, je was mijn afstudeerbegeleider bij mijn masterthesis en dan ook nog eens als copromotor betrokken bij mijn tijd als promovendus. Met je kritische blik zorgde je ervoor dat ik altijd scherp bleef, zowel inhoudelijk als tijdstechnisch. Ik vond het altijd erg fijn om samen te sparren over het opzetten van onderzoeken en het schrijven van artikelen. Ons tripje met Jobke naar Toronto – inclusief de rondleiding door de stad, het bezoek aan een wedstrijd van de Toronto Blue Jays en de roadtrip naar de Niagara Falls – zal ik nooit vergeten. Ook de lunchsessies bij je lectoraat van de HAN waren een erg waardevolle toevoeging aan mijn dagelijkse werkzaamheden. Dankjewel!

Yan, when writing the acknowledgements I was unsure whether to write your part in Dutch or English. Our communication has always been mostly in English, although I believe Dutch would have been fine too. Thank you for your support as a daily supervisor in Twente, and during the later stages of the research in Nijmegen. You made me feel welcome right after I started at the University of Twente. Your daily support was great and you helped me a lot with including a quantitative component to the research conducted. I remember I was always welcome to ask questions and you took the time to explain and teach me about data analysis. It has been a lot of fun to work together with you and I have a lot of great memories from that time. Thank you for your time and effort.

After moving to the Radboud University, a change in daily supervision was required. Abdel, thank you for all the interesting conversations we have had during the research. Your directness and eagerness to question everything were more than welcome and have resulted in new and interesting perspectives on research and life.

Jobke, dankjewel voor de leuke samenwerking die we in Enschede hebben gehad. Met veel plezier heb ik met jou gewerkt aan ons onderzoek op de UT waarin we bekeken hoe mensen met een visuele beperking navigeren in bekende en onbekende omgevingen. Ondanks dat het in de zomer erg leeg was op de campus, hebben we volgens mij beide veel van dit onderzoek opgestoken. Daarnaast zal ik ons bezoek aan Toronto waar ik eerder al aan refereerde niet meer vergeten. Het was mijn eerste conferentiebezoek met collega's en samen met Thea hebben we er een erg leuke trip van gemaakt.

Naast mijn team van begeleiders wil ik ook graag iedereen bedanken met wie ik samen heb mogen werken tijdens mijn verschillende studies. Marian, Renske, Frederic, Otto, Thom en Iwan, heel veel dank voor de samenwerking gedurende jullie bachelor- en masterstages. Hopelijk hebben jullie deze samenwerking als net zo leerzaam en gezellig ervaren als ik dat heb gedaan. Veel succes gewenst met jullie toekomstige loopbaan.

Promoveren houdt (gelukkig) veel meer in dan alleen het doen van onderzoek. Zowel bij de Universiteit Twente als bij de Radboud Universiteit heb ik nieuwe mensen ontmoet en vriendschappen gevormd. Allereerst wil ik graag mijn kamergenoten en collega's van de vakgroep BSS aan de UT bedanken. Frauke, Kees, Ben, Rick, Sabine, Akshita, bedankt voor de vele discussies, geouwehoer, koppen koffie en gezelligheid. Het was een feest om een werkkamer met jullie te delen. Daarnaast wil ik ook Ainara, Henk, Fokke, Frederiek, Dirk en alle andere collega's bedanken voor de fijne tijd die ik bij BSS heb gehad met vele lunchwandelingen, koffiepauzes, promotiefeestjes, donderdagmiddagborrels en de gezellige avonden met Mexicaans eten. Verder wil ik in het bijzonder nog Wies en Sandra bedanken. Jullie deur stond altijd open en jullie waren erg behulpzaam. Zelfs nadat ik weg was uit Enschede kon ik nog altijd met vragen bij jullie terecht en als ik jullie weer zag, waren jullie altijd in voor een praatje. Bedankt daarvoor!

Na mijn fijne tijd in Enschede heb ik ook in Nijmegen erg leuke collega's leren kennen. Annemiek, Elisabeth, Arjen en Loes, bedankt voor de geweldige tijd die we hebben gehad in Nijmegen. Ik heb veel van jullie kunnen leren en veel met jullie gelachen. Ik vind het dan ook erg fijn dat ik mijn paranifmen uit jullie midden heb kunnen kiezen. Voor mijn gevoel konden we het met elkaar over alles hebben en hebben we elkaar, zowel op werkgebied als persoonlijk vlak, veel kunnen geven. Peter, Zahra, Lei, Snandan, Ahmed, Dominik, Katharina, Seba, Hans and all the others at Biophysics, thank you all for an amazing time! I will always remember the Friday bowling sessions, wakeboarding, the day-out to the Ooijpolder and the Honigfabriek, the legendary Christmas celebrations with the department, and an amazing weekendtrip to Belgium. Judith en Marie-Louise, heel erg bedankt dat jullie altijd voor iedereen klaarstaan. Nina en Richard, op het moment dat het allemaal even niet meer lukte, waren jullie er om mij erop te wijzen dat het soms verstandig kan zijn om een pas op de plaats te maken. Een moment waarvoor ik jullie niet genoeg kan bedanken.

Op het moment van schrijven van dit dankwoord werk ik alweer bijna twee jaar als onderzoeker bij Vilans. Wederom een bevestiging dat de eerdere twijfels die ik in mijn sollicitatiebrief naar Richard uitsprak nergens voor nodig waren. Ik wil graag in het bijzonder mijn collega's van het voormalige team technologie en medepromovendi bedanken voor het fijne welkom. Sanne, Henk Herman, Dirk, Sandra, Lotte, Nick, Liesbeth, Kiyomid, Doortje, Johan en Johan, Riske, Esther, Lisa, Xandra, Minke en Nienke. Dankzij jullie voel ik mij op mijn plek bij Vilans en met velen van jullie heb ik ook de laatste loodjes van het afronden van mijn proefschrift kunnen bespreken. Dank daarvoor.

Dan de mensen zonder wie dit boekje nooit had bestaan. André en Carolien, papa en mama. Jullie hebben mij onvoorwaardelijk gesteund, waardoor ik nooit iets tekort heb hoeven komen. Jullie hebben mij altijd vrij gelaten en alles zelf laten ontdekken. Ik ben ook dankbaar dat jullie steeds weer interesse in mijn onderzoek toonden, maar dat jullie deze interesse op momenten dat ik hier niet op zat te wachten ook meteen parkeerden. Dries en Marloes, ondanks dat jullie niet altijd helemaal doorhadden wat ik nou precies uitspookte, waren jullie erg betrokken en zorgden jullie natuurlijk voor de broodnodige gezellige afleiding. Oma Hil, ook al was het soms lastig te onthouden wat ik nou precies deed, met alle liefde legde ik jou opnieuw uit wat voor onderzoek ik deed. Hopelijk ben ik nu ook volgens jouw richtlijnen echt afgestudeerd en ik beloof dat ik, ook met mijn nieuw behaalde titel, niet zal lopen pronken.

Tot slot wil ik graag de persoon bedanken van wie ik buiten het onderzoeksteam zeker weet dat ze mijn complete proefschrift heeft gelezen. Lieve Emily, ik herinner me nog goed dat ik bij jou thuis was toen ik hoorde dat ik aan de slag kon gaan als junior onderzoeker. Nu, zes jaar later, wonen we alweer enkele jaren met veel plezier samen in Arnhem. De afgelopen jaren heb ik altijd op jou kunnen bouwen en we hebben fantastische dingen beleefd zoals onze vakanties, sport- en kookavonturen en onze reis naar Centraal-Amerika. Na tegenslagen, zoals vervelende reviews, wist je mij altijd op te vrolijken en te motiveren om door te blijven gaan. Als het aan mij ligt, gaan we in de toekomst nog veel meer avonturen beleven samen. Bedankt voor alles!

About The Author

On May 10 1989, Hendrik Pieter Buimer was born in Hengelo (Overijssel), the Netherlands. In 2007 he finished secondary education at the OSG Erasmus in Almelo. After this, he moved to Groningen to follow a Bachelor's program in Communication Systems at the Hanze University of Applied Sciences, which he finished in 2012, after writing a thesis about the requirements of a portal to support renal transplant patients in their lives after a transplantation at the University Medical Center Groningen. It is here where he got inspired to do research, and decided to continue studying at the University of Twente in Enschede. After finishing a pre-master in Communication Science, he received a Master of Science in Communication Studies at the University of Twente in 2015 under supervision of dr. Thea van der Geest. He graduated on a study at Roessingh R&D in which he explored determinants of portal usage and patient adherence to a blended-care rehabilitation program. The moment he graduated from the University of Twente, Hendrik started to work as a junior researcher at the Biomedical Signals & Systems group at the same university under supervision of prof. dr. Richard van Wezel, dr. Thea van der Geest, and dr. Yan Zhao on a project in which the opportunities of smart wearables for persons with a visual impairment were explored. Medio 2016, Hendrik moved to the Biophysics group at the Radboud University in Nijmegen to continue the work he started earlier in Enschede, which eventually resulted in the research described in this dissertation. Since January 2019, Hendrik works as a researcher with a focus on eHealth at Vilans in Utrecht, the national center of expertise for long-term care in the Netherlands, where he still works to date.