UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI PSICOLOGIA GENERALE

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE PSICOLOGICHE

CICLO XXVIII

TOWARD MULTIMODALITY: GESTURE AND VIBROTACTILE FEEDBACK IN NATURAL HUMAN COMPUTER INTERACTION

Direttore della Scuola: Ch.ma Prof.ssa Francesca Peressotti

Supervisore: Ch.mo Prof. Luciano Gamberini

Dottoranda: Valeria Orso

Abstract

In the present work, users' interaction with advanced systems has been investigated in different application domains and with respect to different interfaces. The methods employed were carefully devised to respond to the peculiarities of the interfaces under examination. We could extract a set of recommendations for developers. The first application domain examined regards the home. In particular, we addressed the design of a gestural interface for controlling a lighting system embedded into a piece of furniture in the kitchen. A sample of end users was observed while interacting with the virtual simulation of the interface. Based on the videoanalysis of users' spontaneous behaviors, we could derive a set of significant interaction trends. The second application domain involved the exploration of an urban environment in mobility. In a comparative study, a haptic-audio interface and an audio-visual interface were employed for guiding users towards landmarks and for providing them with information. We showed that the two systems were equally efficient in supporting the users and they were both well-received by them. In a navigational task we compared two tactile displays each embedded in a different wearable device, i.e., a glove and a vest. Despite the differences in the shape and size, both systems successfully directed users to the target. The strengths and the flaws of the two devices were pointed out and commented by users. In a similar context, two devices supported Augmented Reality technology, i.e., a pair of smartglasses and a smartphone, were compared. The experiment allowed us to identify the circumstances favoring the use of smartglasses or the smartphone. Considered altogether, our findings suggest a set of recommendations for developers of advanced systems. First, we outline the importance of properly involving end users for unveiling intuitive interaction modalities with gestural interfaces. We also highlight the importance of providing the user the chance to choose the interaction mode better fitting the contextual characteristics and to adjust the features of every interaction mode. Finally, we outline the potential of wearable devices to support interactions on the move and the importance of finding a proper balance between the amount of information conveyed to the user and the size of the device.

Sommario

I sistemi computazionali hanno ormai da tempo abbandonato lo scenario immobile della scrivania e tendono oggi a coinvolgere sempre di più ambiti della vita quotidiana, in altre parole pervadono le nostre vite. Nel contesto del *pervasive* o *ubiquitous computing*, l'interazione tra l'utente e la macchina dipende in misura sempre minore da specifici sistemi di input (per esempio mouse e tastiera) e sfrutta sempre di più modalità di controllo naturali per operare con i dispositivi (per esempio tramite i gesti o il riconoscimento vocale). Numerosi sono stati i tentativi di trasformare in modo sostanziale il design dei computer e delle modalità di interazione tra cui l'impiego di sistemi per il riconoscimento dei comandi gestuali, dispositivi indossabili e la realtà aumentata. In tali contesti, i metodi tradizionalmente impiegati per lo studio della relazione uomo-macchina si rivelano poco efficaci e si delinea la necessità di una adeguata revisione di tali metodi per poter indagare adeguatamente le caratteristiche dei nuovi sistemi. Nel presente lavoro, sono state analizzate le modalità di interazione dell'utente con diversi sistemi innovativi, ciascuno caratterizzato da un diverso tipo di interfaccia. Sono stati inoltre considerati contesti d'uso diversi. I metodi impiegati sono stati concepiti per rispondere alle diverse caratteristiche delle interfacce in esame e una serie di raccomandazioni per gli sviluppatori sono state derivate dai risultati degli esperimenti. Il primo dominio di applicazione investigato è quello domestico. In particolare, è stato esaminato il design di una interfaccia gesturale per il controllo di un sistema di illuminazione integrato in un mobile della cucina. Un gruppo rappresentativo di utenti è stato osservato mentre interagiva con una simulazione virtuale del prototipo. In base all'analisi dei comportamenti spontanei degli utenti, abbiamo potuto osservare una serie di regolarità nelle azioni dei partecipanti. Il secondo dominio di applicazione riguarda l'esplorazione di un ambiente urbano in mobilità. In un esperimento comparativo, sono state confrontate un'interfaccia audio-aptica e una interfaccia audio-visiva per guidare gli utenti verso dei punti di interesse e per fornire loro delle informazioni a riguardo. I risultati indicano che entrambi i sistemi sono ugualmente efficienti ed entrambi hanno ricevuto valutazioni positive da parte degli utenti. In un compito di navigazione sono stati confrontati due display tattili, ciascuno integrato in un

diverso dispositivo indossabile, ovvero un guanto e un giubbotto. Nonostante le differenze nella forma e nella dimensione, entrambi i sistemi hanno condotto efficacemente l'utente verso il target. I punti di forza e le debolezze dei due sistemi sono state evidenziate dagli utenti. In un contesto simile, sono stati confrontati due dispositivi che supportano la Realtà Aumentata, ovvero un paio di smartglass e uno smartphone. L'esperimento ci ha permesso di identificare le circostanze che favoriscono l'impiego dell'uno o dell'altro dispositivo. Considerando i risultati degli esperimenti complessivamente, possiamo quindi delineare una serie di raccomandazione per gli sviluppatori di sistemi innovativi. Innanzitutto, si evidenzia l'importanza di coinvolgere in modo adeguato gli utenti per indentificare modalità di interazione intuitive con le interfacce gesturali. Inoltre emerge l'importanza di fornire all'utente la possibilità di scegliere la modalità di interazione che meglio risponde alle caratteristiche del contesto insieme alla possibilità di personalizzare le proprietà di ciascuna modalità di interazione alle proprie esigenze. Infine, viene messa in luce le potenzialità dei dispositivi indossabili nelle interazioni in mobilità insieme con l'importanza di trovare il giusto equilibrio tra la quantità di informazioni che il dispositivo è in grado di inviare e la dimensione dello stesso.

Table of contents

Introduction 8

Chapter 1 — **Drawing the field of work**

1. Computers pervading our lives………………………………. 13
2. Gesture-based interactions………………………………… 13
3. Tactile interfaces………………………………………….. 14
4. Wearable computers………………………………………. 15
5. Augmented Reality………………………………………... 17
6. Multimodal interfaces…………………………………….. 19
7. Challenges in evaluating advanced systems………………. 20
8. How to deal with the methodological issues………………. 23
8.1 Informing the design of advanced systems……………… 23
8.2. Evaluating advanced systems…………………………… 24

Chapter 2 — **How would you do that? Exploring users' spontaneous gestural commands**

1. Motivation and aim………………………………………... 27
2. Related work………………………………………………. 27
3. Equipment…………………………………………………. 27
4. Setting……………………………………………………... 28
5. Experimental procedure…………………………………… 29
6. Sample…………………………………………………….. 31
7. Analysis……………………………………………………. 32
8. Results……………………………………………………... 32
8.1 A taxonomy of the gestures……………………………… 32
8.2 Frequency of the hand gestures…………………………... 41
8.3 Movements……………………………………………….. 46
8.4 Additional observations………………………………….. 55
9. Discussion…………………………………………………. 57

Chapter 3	**Exploring urban landmarks: comparing a haptic-audio and a audio-visual interface**	
	1. Motivation and aim………………………………………...	58
	2. Related work……………………………………………….	59
	3. Equipment………………………………………………….	60
	3.1 Tactile glove………………………………………………	60
	3.2 Baseline application………………………………………	61
	4. Method……………………………………………………..	62
	4.1 Experimental design………………………………………	62
	4.2 Setting…………………………………………………….	62
	4.3 Sample…………………………………………………….	63
	4.4 Experimental procedure…………………………………..	63
	4.5 Materials…………………………………………………..	64
	5. Results……………………………………………………...	65
	5.1 Performance………………………………………………	65
	5.2 Questionnaires…………………………………………….	67
	5.3 Interviews…………………………………………………	68
	6. Discussion………………………………………………….	70
Chapter 4	**Following the vibes: a comparison between two tactile displays in a navigational task in the field**	
	1. Motivation and aim………………………………………...	71
	2. Related work……………………………………………….	72
	3. Equipment………………………………………………….	73
	3.1 Tactile glove………………………………………………	74
	3.2 Tactile vest………………………………………………..	74
	4. Method……………………………………………………..	75
	4.1 Experimental design………………………………………	75
	4.2 Setting…………………………………………………….	75
	4.3 Sample…………………………………………………….	76
	4.4 Experimental procedure…………………………………..	77

	4.5 Materials……	77
	5. Results……	78
	5.1 Performance……	78
	5.2 Questionnaires……	80
	5.3 Interviews……	83
	6. Discussion……	84
Chapter 5	**Augmenting the Reality: a comparison between a smartphone and a pair of smartglasses**	
	1. Motivation and aim……	86
	2. Related work……	87
	3. Equipment……	88
	4. Method……	89
	4.1 Experimental design……	89
	4.2 Setting……	89
	4.3 Sample……	90
	4.4 Experimental procedure……	90
	4.5 Materials……	91
	4.6 Data collected……	94
	5. Results……	94
	5.1 Performance……	94
	5.2 Interaction behaviors……	97
	5.3 Questionnaires……	101
	5.4 Interview……	106
	6. Discussion……	107
Chapter 6	**General discussion**	108
	Conclusions	116
	References	118

Introduction

Since the advent of mobile devices and services coupled with computers, computational systems have left static workstations, trending towards an unlimited spread into everyday activities, a phenomenon that is termed pervasive computing. Pervasive computing is characterized by the effort of distributing computational systems into our environment (Dix 2009). This view assumes that, when interacting with computers, individuals act less through specific input device (e.g., keyboard and mouse) and more employing intuitive ways of controlling devices (Fleer and Leichsenring 2012). In this framework, numerous attempts have been made to transform substantially the design of computers and the interaction modalities between humans and machines. The most advanced ones include computers disappearing into garments, i.e., wearable computers, and into the environment, i.e., ambient interfaces (Barfield and Caudell 2001; Baber 2001). In such contexts, the interaction modalities rely on the behavioral repertoire the individual would naturally realize in her everyday interactions, mainly referring to the use of speech and gestures (Ballendat, Marquardt and Greenberg 2010; Möller, Krebber and Smeele 2006). On the other hand the output sent by the computer is no longer constrained within visual or auditory icons, rather the system is capable of delivering unobtrusive feedback to the user, i.e., tactile interfaces (Krishna, Bala and Panchanathan 2010). Ultimately, the methods and techniques traditionally used in the field of human computer interaction (HCI) are challenged and require to be properly adjusted to respond to the new trends. The present work has as a general research question how to address new standards for the interaction and the evaluation of advanced systems. This issue has been tackled through four experiments, each involving a different technology (gestural interface, vibrotactile interface, augmented reality and haptic-audio interface) and representative samples of potential end users. The aim was to answer to the following specific research questions:

- How can a gestural interface be designed to be effectively integrated in a domestic environment like the kitchen? – Experiment 1

- Is a haptic-audio interface a valuable tool for supporting a tourist in a free exploration task as compared to a conventional audio-visual system? – Experiment 2
- Are vibrotactile stimuli an effective navigational cue to guide users in an urban environment? - Experiment 3
- Given a mobile AR application for the exploration of the surroundings in a touristic context, how does a wearable device perform as compared to a hand-held device? – Experiment 4

In the present work, user interactions modalities with advanced systems have been investigated in different application domains, as mentioned above, exploiting and integrating different methods of analysis.

Chapter 1 describes first the field of work and the different type of interfaces that have been analyzed in the present work. The challenges posed by the special characteristics of such interfaces for the traditional HCI methods are reported. Possible ways for overcoming such shortcomings are outlined.

The first application domain is analyzed in Chapter 2 and pertains the domestic environment. In particular, the design of a gestural interface for controlling a lighting system embedded into a piece of furniture in the kitchen is investigated. The majority of the gestural interfaces developed involve interactions with large displays, and the commands are consistent with that size. However, when interacting with home appliances, small displays are more likely to be embedded and appropriate gestural commands need to be employed (Garzotto and Valoriani 2012). A total of 30 participants volunteered in the study. Their task was to perform the gestures they would spontaneously use to control specific functions of the lighting system in front of a virtual simulation of the system and their interactions were recorded by means of four video-cameras. Based on the videoanalysis of users' spontaneous interactions, we were able to extract the general trends in participants' gestures, despite the known lack of affordances in gestural interfaces (Norman 2005; Greenberg, Marquardt, Ballendat, Diaz-Marino and Wang 2011). The experiment reported in this chapter is the result of the collaboration with the Global Technology Center of Eloctrolux S.p.a., located in Porcia (Italy).

The third chapter explores users' interaction on the move, in particular the touristic context is addressed here. Tactile interfaces have been experimented for supporting information discovery in the mobile context (Robinson, Eslambolchilar and Jones 2009), since they present considerable advantages over audio-visual interfaces. They are in fact useful when the visual and auditory attention is limited and they can deliver precise spatial information in an unobtrusive manner over 360° around the user (Srikulwong and O'Neill 2011). No previous studies have addressed a comparative evaluation of a tactile-audio interface and a touristic guide based on a traditional audio-visual interface. In this experiment, a tactile glove was integrated with audio contents and was paired with an application allowing the serendipitous discovery of points of interest in the surroundings of the user. Such system was compared to an analogous application running on a smartphone developed *ad hoc* for resembling popular smartphone-based applications for exploring points of interest in the urban heritage. A total of 20 participants volunteered in the study, which followed a between-subjects design with the device as the independent variable. The data collected included video-recordings of users' behaviors and post-experience questionnaires and interviews. The results show that participants were able to correctly indicate the point of interest advised by the both devices within a similar amount of time. The experiment was conducted in collaboration with the Ubiquitous Interaction Research Group head by professor Giulio Jacucci and based at the Helsinki Institute of Information Technology (Finland). In particular, the tactile glove was designed and developed by Yi-Ta Hsieh and Antti Jylhä; the baseline application was developed by Salvatore Andolina. For this study Renato Mazza, a master's student, was responsible for the data collection, under the writer's supervision.

Chapter 4 investigates the employment of tactile interfaces as navigational tools. In the scientific literature, several studies have investigated the potential of tactile interfaces for delivering navigational directions (Tsukada and Yasumura 2004), however no previous work has directly compared the different devices employed (mainly wearable devices to be put on the torso or to be held in the hand; Prasad, Taele, Goldberg and Hammond, 2014; Jacob, Mooney and Winstanley 2011). Two different wearable computers with an embedded tactile interface, i.e., a tactile glove (the same employed in Experiment 2) and a tactile vest, were compared as navigation tools in an urban environment. The experiment had a within-subjects design, with 24 participants using both the tactile glove and the tactile vest to walk through two different, yet analogous, routes in the center of Padua. Their performance and experience were recorded by means of datalogs, questionnaires and interviews. The results indicate that the two devices are equally efficient in providing navigational indications to users. The two devices were both praised for different reasons: users appreciated the fact that the tactile glove was light-weighted and unbulky, however the tactile vest was praised for the vibrotactile cues being clearly perceivable. This experiment was conducted with the collaboration of Ann Morrison and Walther Jensen, from the Department of Architecture and Media Technology at Aalborg University (Denmark), who designed and developed the tactile vest. For this study Renato Mazza was responsible for the data collection, under the writer's supervision.

The fifth chapter explores Augmented Reality (AR) technology. AR has gained a lot of interest from researchers, especially in the touristic domain (Noh, Sunar and Pan 2009) and several applications running on hand-held devices have been developed. However the need of continuously switching the attention between the screen and the environment has been addressed to be one of the main limitations of visual AR in a mobile context (Dünser, Grasset, Seichter and Billinghurst 2007). Head-worn viewers supporting AR have the potential to overcome this drawback allowing the user to benefit of enhanced contextualized information, yet without the need to focus on the screen of a hand-held device. Users' interaction with a head-worn viewer allowing stereoscopic view of AR applications was compared with users' interactions with a smartphone. The study compared the use of a commercial AR application for locating places in the surroundings of the user running on the head-worn viewer or on a smartphone. The participants involved (N= 36) were asked to complete three common tasks in the touristic application domain, i.e., reach a pre-fixed destination, access and read additional information and find the nearest points of interest within a certain category. The trails were video-recorded. Participants were able to successfully accomplish all the tasks with both devices, however a difficulty in reading the text emerged when they were using the head-worn viewer. The video-analysis suggests that the head-worn viewer caused a smaller number of interruptions in the tasks execution, supporting a smoother interaction, compared to the smartphone. Overall, the aspects related to the experience (e.g., comfort of use, acceptability, pleasantness of use) were positive for both the head-worn viewer and the smartphone. For this study Andrea Beretta, a master's student, was responsible for the data collection, under the writer's supervision.

In Chapter 6 the overall findings from the four experiments reported are discussed and a set of general recommendation are extracted and described.

Finally, concluding remarks are reported in Chapter 7.

Chapter 1

Drawing the field of work

1. Computers pervading our lives

About twenty years ago Mark Weiser envisioned a coming era where computation would be seamlessly embedded in the world. He termed this "ubiquitous computing". In the evolution of computing systems, this era would be the arrival point, following first two phases: the mainframe and the personal computer. The mainframe era is characterized by a considerable distance between the average user and the computer, as the computational machines are confined in special places and can only be run by experts. A second trend is the PC era, where every the average user owns and uses a personal computer. Here the relation between the human and the machine has become personal, since we entrust our data to the computer. Finally, in the ubiquitous computing era, each of us is sharing many computers, which are integrated in everyday objects, connected to one another and operated through a central controlling station (Weiser 1997; Buxton 1997). In Weiser's view, computers are completely, yet invisibly, integrated with the environment. They empower a smoother shift from what is only marginally attended (e.g., the periphery) and what is in the central focus of our attention, allowing the user to have control of a larger amount of stimuli without being overwhelmed by information (Weiser 1997).

Today many of the predictions made by Weiser have come true, with computers that have effectively leapt off the desktop and have contaminated nearly all the aspects of our lives. The term ubiquitous coined by Weiser meant to embed computing in everything, not only to bring computers everywhere. In this respect, the research on ubiquitous computing has evolved as an extremely multidisciplinary field, with contributions from mobile computing, wearable computing, augmented reality, nearfield communication (Greenfield 2010; Abowd 2012).

Regardless of the shape the technology takes, what is really relevant is the relation we, as users, have with it (Weiser 1997). A significant effort has been expended in trying to transform the traditional "Windows Icons Menu Pointing" (WIMP) interaction styles into more natural interaction modalities. In this framework, the increased naturalness is realized by replacing artificially defined commands with input taken from everyday behavioral repertoire, e.g., speech, hand gestures, body movements and proxemics. These emerging interaction styles are grounded into the user's pre-existing knowledge of the real world and melt themselves into her daily practice (Jacob et al. 2008; Abowd 2012; O'Hara 2008). Each of these new interaction modes alone can hardly be the only communication channel with the computing system. Let us consider the case of a wrist-worn activity tracker. Once a user has achieved her daily fitness goal, she receives a vibrotactile notification and by swiping her finger on the display she can glimpse summary data. But if she wishes to have a detailed monitoring of her activity, she has to access the application coupled with the activity tracker running on her smartphone, with which she interacts through touch-based gestures.

Here we argue that it is an effective integration of diverse and intuitive interaction modalities within the same system that can build an interaction with "beautiful seams" (Weiser 1993). However this integration raises a number of issues from a human factors perspective. If computational systems lie in all the objects of our lives, it can be the case that we engage with one (or more) of them unintentionally. We need therefore to uncover appropriate ways to connect with only the desired piece, to (in)activate a specific function and to gently disengage from it (Greenfield 2010). To say it with Norman's (2010) words "we need new standards".

2. Gesture-based interactions

As human beings, we extensively make use of gestures for communicating with each other, for signifying objects and for interacting with the environment (Wachs, Kölsch, Stern and Edan 2011). Being a constituent component of our everyday behavioral repertoire, hand gestures qualify as a valuable interaction means for natural interaction with computers embedded in our environment (O'hara, Harper, Mentis, Sellen and Taylor 2013; Marquardt and Greenberg 2012). Gestural input has been experimented in many different forms either mediated by handheld devices (e.g., Shoemaker, Tsukitani, Kitamura and Booth 2010) or as direct touch-based gestural commands (e.g., tabletop surfaces; Morris, Huang, Paepcke and Winograd 2006). Touch-less hand gesture input has several advantages over mediated interactions, firstly allowing the user to access information in total sterility, because there is no contact with surfaces. Secondly, information is accessible from a distance, enabling people with physical impairments to control devices afar. Finally it facilitates the exploration of big data, since it involves no movement constraints and allows direct manipulation of information (Wachs, Kölsch, Stern and Edan 2011; O'hara, Harper, Mentis, Sellen and Taylor 2013). However, such interaction modalities have clear limitations. First, making the possible actions visible challenges touch-less gestural interfaces (Norman 2010; Marquardt and Greenberg 2012). In addition, providing meaningful feedback regarding the appropriateness of the commands prompted is difficult, yet crucial in this context (Norman 2010). This raises questions of what kind of feedback to provide. Some researchers have proposed to exploit proxemics, as introduced by Hall (1966), to further characterize touchless gesture input. According to Hall[1], physical distance between individuals correlates with social distance and is an implicit communication message. In parallel, a system can automatically modulate its state by reacting according to the user's distance from it and to her behaviors, revealing interaction possibilities only when the user is nearby (e.g., Prante et al., 2003) and providing contextual feedback (Ballendat, Marquardt and Greenberg 2010).

Finally, despite their flexibility and potential intuitiveness (Stern, Wachs and Edan 2008), gestural interactions are strongly related to the application in use (Wachs, Kölsch, Stern and Edan 2011) and the cultural practices (Norman, 2010). This implies a conscious involvement of end users in the design of interaction vocabularies (Ruiz, Li and Lank 2010) and highlights the need to develop new conventions (Norman 2010).

3. Tactile interfaces

When an interface conveys information through the sense of touch it is generally termed a haptic interface. Haptic is a broad term including both cutaneous and kinesthetic touch. The former applies to the perception of the surface features and tactile

---

[1] Hall (1966) categorized four discrete distance zones affecting relations and communication among

stimulation. The latter refers to our interpretation of the location of our limbs in the space and is conveyed by muscles and tendons. Tactile interfaces are a subcategory of haptic interfaces and involve the stimulation of mechanoreceptors in the skin. Touch is one of the most informative senses and mechanical interaction is vital when a sense of presence is desired. The first experimentations with tactile stimuli involved in fact virtual environments (O'Malley and Gupta 2008; Chouvardas, Miliou and Hatalis 2008).

Tactile interfaces can deliver stimulation via different modalities, i.e., static pressure or vibration, electric field or thermal flow (Chouvardas, Miliou and Hatalis 2008). Despite a variety of stimulation possibilities exist (e.g., focused ultrasound, surface acoustic waves), for the scopes of the present work mechanical vibration-based are of interest.

The primary interest in developing tactile displays was in fact in allowing the access to graphical user interfaces to blind and deaf people (Chouvardas, Miliou and Hatalis 2008). Later, tactile interfaces have been experimented for replacing visual and acoustic information when the device was used in mobility by able-bodied users. Interacting with a device in mobility (e.g., a smartphone) poses several challenges that can be assimilated to temporary impairments involving the sight, hearing or dexterity (Hoggan, Anwar and Brewster 2007; Hoggan and Brewster 2006). Tactile displays have proved to be an effective and powerful means for communicating information to users. On the other hand, users can reliably comprehend messages enclosed in the tactile mode (Hoggan, Anwar and Brewster 2007). Structured vibrotactile messages conveying information non-visually, namely tactons, have been proposed. By manipulating the parameters of the vibration, i.e., rhythm, roughness and spatial location, different information can be encoded in tactons (Brown, Brewster and Purchase 2006). Tactile displays have been experimented also as a means for delivering feedback in touchless gestural interfaces, that is to say for enhancing touchless interfaces with a touch feedback (Freeman, Brewster and Lantz 2014).

The majority of the applications employing a tactile interface are devised for the palm or the fingertips (e.g., Ahmaniemi and Lantz 2009), which are the most sensitive part of the body, with the exception of the lips. The application of tactile displays on the torso has also received attention by researchers (e.g., Van Erp, Van Veen, Jansen and Dobbins 2005), as the surface of the skin of the back can covey twice the information as the fingertips (O'Malley and Gupta 2008; Chouvardas, Miliou and Hatalis 2008). Tactile interfaces have also been experimented on handheld devices (e.g., Wei, Ren and O'Neill 2014) and wearable devices (e.g., Weber, Schätzle, Hulin, Preusche and Deml 2011). Among all the possible stimulations that tactile interfaces can deliver, vibration is a suitable candidate for being integrated into garments, because of the average size of tactile displays. Tactile interfaces are in fact commonly embedded into wearables: gloves (e.g., Carton and Dunne 2013), vest (e.g., Prasad, Taele, Goldberg and Hammond 2014) and shoes (e.g., Watanabe and Ando 2010).

4. Wearable Computers

Current trends in computing have brought to the design and development of systems allowing the user to interact outside of the static context of the office desktop and personal computer. Computing devices have evolved in form factors that are increasingly portable (Gemperle, Kasabach, Stivoric, Bauer and Martin 1998). Furthermore, it emerges a need for integrating computational intelligence in the surrounding environment. To this end, clothing appear to be an ideal place to set computers, as they could expand the user's capabilities at any given moment without the need to refer to external devices (Cho, Lee and Cho 2009). In Barfield et al.'s words, wearable computers are "fully functional, self-powered, self-contained computers that allows the user to access information anywhere and at any time" (Barfield et al. 2001). A further distinction can made in (a) computers that can be worn, that are small PCs that are configured to fit one's body; (b) information appliance that can be worn, which refers to technological devices the user can wear with a specific set of functions; (c) computers as clothing, that are computationally simple devices aiming at monitoring the environment; and (d) smart clothing, in which the electronic components are sewn with the textile (Baber 2001; Gemperle, Kasabach, Stivoric, Bauer and Martin 1998). In sum, the defining characteristics of wearables are that they mediate the communication between the wearer and the world and that they allow the wearer to manipulate information. Following Weiser's thought seeing ubiquitous computing "weav[ing] themselves into the fabric of everyday life until they are indistinguishable from it" (Weiser 1991), Baber (2001) sees wearables as an ideal form factor to host computational systems, because they are almost unnoticeable for the casual observer and, on the other hand, they are not only physically, but also cognitively and socially blended with the user.

Ubiquitous computational systems purely meant as computers embedded into the environment raises several issues. Privacy is the first concern. In fact, the data sensed and recorded by the system are never only owned by the user, but they are always also stored in the environment itself, and many are the examples in which a user wouldn't trust the environment (e.g., the customer-provider relation is always delicate on this respect). In addition, the information stored in the user's profile is limited to the data registered in a specific environment and is always, at least in part, partial. For a fully customized profile, one would need to constantly update her information almost every time she accesses a new place. When wearable sensors are put on the person, rather than in the room, they always move with the wearer. Therefore information recorded is essentially always firstly in the user's hand (consistently with the application settings). More importantly, the data stored in the user's profiles always travels with the wearer, evolving thus with her and removing the need for transferring it to new environments (Rhodes, Minar and Weaver 1999).

In this perspective the pervasiveness of computers is achieved not by computers being seamlessly embedded in all the environments, rather by the computational devices being fully integrated and always accessible on the user's body.

5. Augmented Reality

According to the seminal definition given by Azuma (1997), Augmented Reality (AR) is characterized as a combination of digitally created images super imposed onto the real-world surroundings, where the alignment between the real and virtual contents occurs in 3D and that is interactive in real time. Thus, by its very definition, AR supplements the user with contextualized information regarding her environment. In the past, almost all AR applications were developed merely as mobile variants of the existing desktop versions and later were deeply dependent on the infrastructure supporting the tracked model (Wagner 2005; Schmalstieg and Reitmayr 2010). Advances on the software, with AR applications running without the need of fiducial markers, and on the hardware sides, with mobile devices (either hand-held or mounted on the head) with ever increasing processing power, have finally released AR to be fully mobile (Huang Alem and Livingston 2013). AR applications that are currently available on the market rely on location-based information resulting in the merging of virtual with real-world objects in an impressive and realistic entity (Schmalstieg and Reitmayr 2010). Digital annotations, which are labels providing relevant information regarding the real-word surroundings, are now a consolidated component on current mobile AR applications (Li and Du 2013). It is evident that AR could not exist independently from portable devices (either wearable or not), however it enables the user to shift the focus of her attention between information, therefore augmenting her view without intruding with the task. Finally, it can be argued that AR literally expands our perception of and our interaction with the real world, encouraging us to look outward rather than turning inward into an artificial world (Buxton 1997).

6. Multimodal interfaces

Multimodal interfaces entail more transparent, flexible and powerfully expressive interaction modalities as compared the conventional graphical user interfaces (GUI). An interface is defined multimodal when it can recognize two or more user input modalities in a coordinated manner with multimedia output from the system (Oviatt 2003). Some of the modalities experimented so far include gestures, speech, eye blinks, haptics and pen input, each eventually coupled with a proper input device (Oviatt 2003; Jaimes and Sebe 2007). Likewise human-human communication, a more effective dialogue between the user and the system is likely to occur when different input modalities are used in combination, especially exploiting speech and hand gestures (Jaimes and Sebe 2007; Wachs, Kölsch, Stern and Edan 2011). The presentation of the first multimodal system, "Put That There" (Bolt 1980), brings us back to early 80ies. That application combined speech and pointing to create and move objects on a 2D display. Since then substantial progresses have been made in the field. Such interest is largely motivated by the potential that multimodal interfaces have to be accessed by users with a wide range of different skills. Furthermore, the flexibility these systems provide allows accommodating the continuously changing conditions of the mobile context of use; mobile users have been termed in fact "situationally impaired" (Hoggan and Brewster 2006). While walking, for instance, the user experiences a reduced dexterity and capacity of clearly discriminate visual objects; or when driving, one may only marginally look at a display, likewise a visually impaired person (Oviatt 2003; Gandy, Ross and Starner 2003). The different features of the modalities considered together with the characteristics of the situation need to be carefully accounted, in order to determine which modalities are functionally better suited to the situation in exam (Oviatt 2003; Jaimes and Sebe 2007; Li and Duh 2013; Perry Dowdall, Lines and Hone 2004).

Besides their flexibility, multimodal interfaces have proved to be less error prone and to smoothly support the user to recover from errors. Users in fact can select the modality they consider more effective and also they tend to switch modality when they have realized that an error has occurred, thus facilitating error recovery (Oviatt 2003; Jaimes and Sebe 2007; Oviatt and Cohen 2000). Evidence also suggests that multimodal interfaces can reduce the user's cognitive load. As the task complexity increases, users can autonomously manage their resources by distributing information across the different modes, consequently enhancing the task performance. The alternation between modalities was also shown to prevent physical fatigue or overuse, especially over long interactions (Oviatt 2003).

In addition to the strengths illustrated above, the integration of different interaction modalities makes the communication between the user and the machine more transparent and natural, and ultimately it tends to resemble more a conversation than a command and control dialogue (Oviatt and Cohen 2000). There is thus a growing interest in designing such systems, exploiting both explicit user commands (e.g., hand gestures, speech) and implicit behaviors (e.g., eye gaze, facial expression), especially in the applications of ubiquitous computing (Oviatt 2003).

7. Challenges in evaluating advanced systems

The concept of usability was born and developed with reference to systems relying mainly on graphical user interfaces (GUIs). Likewise, the methods for inspecting the usability were devised for interactions based mainly on visuals. Researchers usually refer to the definition provided by the standard ISO 9241, according to which the usability of a system is "the effectiveness, efficiency and satisfaction with which specified users achieved specified goals in particular environments[2]". Here the focus is mainly on the user's performance, intended as the accuracy and completeness of the task achievement (effectiveness) and as the resources expended to accomplish the task (efficiency). Attention is given also to the comfort and acceptability experienced by users while busy at solving the task (satisfaction). Based on these essential metrics, a wide and well-consolidated body of research exists collecting methodologies and recommendations for conducting effective usability testing investigating both aspects relative to the performance and to the experience of use. The User Experience (UX) refers to the ensemble of emotional and affective aspects deriving from the interaction with a system and that are usually identified with the pleasantness and playfulness of the experience (Hassenzahl 2004).

Evaluation methods can be broadly divided into formative and summative. The former refers to tests carried out iteratively in the early stages of the design life-cycle, while the latter regards mainly the tests of a definitive version of the protoype. Even if the UX can be investigated throughout the projects it is evident that a more structured understanding can be gathered when the system approaches the final stages of development.

Typically, usability tests take place in a controlled laboratory setting often arranged to resemble a familiar environment (e.g., a living room; Orso, Spagnolli, Gamberini, Ibañez and Fabregat 2015), it involves participants that are representative of the end users (Kjeldskov and Stage 2004; Rubin and Chisnell 2008; Dix 2008). It is common practice to dedicate a part of the test to the investigation of the experience users had with the system.

A number of differences however exist between conventional systems, for which the interaction relies mainly on graphics and advanced systems, these differences make traditional evaluation methods often inadequate (Dix 2008). In fact, advanced systems generally imply a shift away from a static context of use, e.g., wearables or smart appliances. Conventional laboratory evaluation can become unsuitable, as it fails to consider contextual aspects that can significantly impact on the interaction. Ideally the tests should take place in the setting in which the system is expected to be used by end users, that is to say in the field (Gorlenko 2003; Kjeldskov and Stage 2004; López-

---

[2] http://www.w3.org/2002/Talks/0104-usabilityprocess/slide3-0.html
Last access January 7th 2015.

Cózar and Callejas 2010; de Sá and Churchill 2013). To this end, *smart rooms* are sometimes employed. However, these settings cannot represent a fully familiar environment able to trigger genuine everyday behaviors from users. *Living labs*, that is houses embedded with sensing technology, have proved to be an advancement in this respect, as they can unobtrusively record users' interactions for an extended period of time (Dix 2008; López-Cózar and Callejas 2010). Additionally, advanced systems are often developed to be used in mobility, requiring mobile testing in field. Such evaluations however pose several difficulties. First it can be complicated to organize realistic studies that can trigger and capture key interaction situations. Furthermore, some well-established data collection techniques, e.g., think-aloud protocols, are unsuitable to be employed in public spaces. In real-world setting a number of unpredictable and uncontrollable variables are likely to intervene and affect the interaction (Kjeldskov and Stage 2004). Another relevant characteristic of those devices is that they are often implemented to be used while the user is busy in another activity. This pushes us to leave the rigidly organized nature of task analysis (Dix 2008), favoring scenario-based tasks. In scenario-based tasks detailed narratives depict specific situations requiring the user to take precise actions, yet without given her directive instructions (Ko, Chang and Ji 2013; de Sá and Churchill 2013). This adaptation can also account for the many different uses that one can do of the same application, depending on the characteristics of the environment (Gorlenko and Merrik 2003). In fact, it builds a situation in which the user is required to act and of which the researcher is fully aware. The need for evaluating in a realistic environment may also require the researcher to conduct different test trails under different conditions (Gorlenko and Merrik 2003).

The use of these innovative interaction modalities out of the workplace or the house, has highlighted new aspects that can affect the usability, that is social acceptance (de Sá and Churchill 2013). According to Suero Montero, Alexander, Marshall and Subramanian (2010), social acceptance refers to the users' feeling of (dis)comfort in interacting with a certain device in private and public spaces. A growing number of studies mainly regarding wearables (e.g, Costanza, Inverso, Pavlov, Allen and Maes 2006) and gestural interfaces (e.g, Profita et al. 2013) have started to address social acceptability. Even if this aspect has been investigated also indirectly by presenting users with sketches

(Koelle, Kranz and Möller 2015), it is evident that in this case only a realistic context can provide meaningful results.
Regardless of the specific implementation characteristics, advanced systems are usually the result of a complex integration of different components. For instance, a mobile AR application for browsing the events in a city entails a menu, a feature for filtering relevant categories and a feature for managing the size of the area of interest (Cabral et al. 2014). López-Cózar and Callejas (2010) suggested evaluating the single components separately in the early phases of the design, in order to gain a clear understanding of which parts are puzzling.
Notably, the investigation of the user experience with the system is paramount, as the qualities of the interaction between the user and the device are considered to affect the usability more significantly than the qualities of the computing device alone (Gorlenko and Merrik 2003; López-Cózar and Callejas 2010).
As final remark it is worth noticing that the preliminary phases of the design lifecycle, where the requirements of the system are drawn, should be grounded on a deep understanding of the users' everyday practice, since these systems aim at integrating unobtrusively within their environment (Dix 2008).

8. How to deal with the methodological issues

From a methodological standpoint formalized procedures for the evaluation of advanced systems do not exist yet. Researchers need to piece together elements from the already established approaches for addressing the issues that are identified from time to time (Lim 2008). A systematic analysis of all the methods and the adjustments that can be adopted is above the scope of the present work. Here we describe the main actions that the researcher can take to tackle the challenges posed by advanced systems outlined above.

8.1 Informing the design of advanced systems

In this kind of less structured domains, where the interaction is richer and less constrained, it is crucial to understand people's everyday practices to inform the design of the system. In order to unveil a user's basic daily operations, ethnographic investigation is becoming increasingly deployed in the field. In this context, the goal of ethnographic research is to provide a detailed and contextualized description of the actors, the practices and the places in which such operations occur. In addition, it should also highlight the eventual relation between the people that may participate in the operation or share the setting (Greenfield 2006; Dix 2010; see for example Obrist, Bernhaupt and Tscheligi 2008).

Participatory design methods (e.g., focus group, affinity diagram) can be also useful in this phase to gather users' firsthand perspective. However they must be used cautiously, as they tend to address directly topics that are likely to be unfamiliar for the participants. As a consequence, such methods may not be truly effective. It has been suggested that representing the matters through illustrated materials (e.g., sketches, video fragments) can help the user to conjure up the topic of discussion (Orso et al., 2015; Litosseliti 2003).

8.2 Evaluating advanced systems

Regarding the evaluation phase, there are a number of tools that can be employed to deal with the constraints posed by the genuine characteristics of such systems described above.

The essence of a usability evaluation consists of observing participants while they are engaged in accomplishing a realistic activity and build an understanding of what in the system may hamper the interaction. To do that users are given a set of assignments, or tasks, that they are asked to achieve. In order to build the activity to be realistic, the researcher presents the user *scenario based tasks*, that are narratives capturing the essence of the interaction and providing a vivid context, thereby engaging the participant to perform a task as if she was at home or in the office (Rubin and Chisnell 2008; Go and Carroll 2004). A well-established procedure is to video record users' interactions with the system and then to analyze their behavior off-line (Heath, Hindmarsh and Luff 2010). It is common practice to list a set of interaction behaviors realized by the user (i.e., defining a coding scheme) and then note on a dedicated grid the frequency of a certain behavior and their duration. This allows to have an objective measure of the user's performance, for instance the number of errors occurred or the time required to complete a certain operation (Rubin and Chisnell 2008). For the video-analysis of systems based on natural interaction modalities, the analysis of breakdowns has been proposed as more informative. Action breakdowns are interruptions or slowdowns in the projected course of the action, as defined by Gamberini et al. (2013). This method consists of a first identification of the breakdown episodes and then in the further specification of its characterizing elements (e.g., eventual tools and operations involved). This method is particularly appealing because on the one hand it allows the researcher to observe what was the user's actual understanding of the system. On the other hand it allows to reveal aspects related to the naturalness of the experience from the user's perspective, since the issues are not *a priori* defined by the experimenter. In both cases the video-analysis is usually conducted with the support of video-coding software (e.g., Noldus The Observer XT[3], Elan[4]). Video-recording users' interactions allow the researcher to move the experimental setting out the laboratory, allowing also to capture the richness of context.

Besides analyzing the overall interaction behaviors, it can be useful to get an insight on how demanding is for the user to use a system from a cognitive standpoint. Well-established self-reported instruments exist for estimating the cognitive workload (e.g., NASA-TLX; Hart and Staveland 1988). Despite their spreading, such tools rely on the respondents' memories of the interaction and fail to record online workload (Ozok 2009). An objective measure of the cognitive load is the pupil size. It is in fact well-known that a restriction of the pupil size correlates to an increase of the cognitive workload (Duchowski 2007). Since the pupil diameter is subject also to affective responses and changes in the ambient light, it is suggested to use such metric in conjunction with other data-collection techniques (e.g., the NASA-TLX survey cited above). Modern eye-tracking systems can reliably detect changes in the pupil diameter also in full mobility, as the recording components are mounted on glass-like support and the computational component run on small and light devices (e.g., Pupil Labs[5], Tobii Pro Glasses 2[6], SMI Eye Tracking Glasses 2[7]). In addition, eye-tracking measures can be employed to identify what parts of the interface the user identifies as perceived affordances. Users tend in fact to fixate the part of the interface they are about to

[3] http://www.noldus.com/human-behavior-research/products/the-observer-xt
[4] https://tla.mpi.nl/tools/tla-tools/elan/
[5] https://pupil-labs.com
[6] http://www.tobiipro.com/product-listing/tobii-pro-glasses-2/
[7] http://www.eyetracking-glasses.com

manipulate. If the affordances placed by the designers fail to be noticed, i.e., fixated, it is likely that the user will operate on the wrong component to accomplish her goals (Manhartsbergen and Zellhofer 2005).

Despite the inherent drawbacks of self-reported methods, questionnaires remain probably the most frequently employed tool to inspect the usability and also the user experience. Questionnaires allow to collect a large amount of data in a relatively short time and in an inexpensive manner. In addition, questionnaires administered before and after the experience can be used to compare the expectations users had of the system after a quick approach and their actual opinion following a session of use (Turunen et al., 2009). Often questionnaires fail to capture the genuinely subjective aspects of the experience and tend to align the opinions of the different users. To prevent this loss of information, brief interviews are increasingly integrated in the experimental procedure. This combination provides the researcher with a richer body of information, comprising both quantitative and qualitative data. Such a dataset can, in turn, unveil diverse aspects of the interaction with the system.

In brief, an effective evaluation of advanced systems should be arranged as a realistic situation in order to gain meaningful results. This includes:

a. the early evaluation of the single components;
b. a realistic setting;
c. scenario-based tasks;
d. triangulation of objective and subjective metrics.

How would do that?
Exploring users' spontaneous gestural commands

1. Motivation and aim

Interaction techniques designed for traditional computers are not always appropriate for the use with other form factors and in some contexts where hand-free interaction is more appropriate (e.g., when the user has dirty hands; Shoemaker, Tsukitani, Kitamura and Booth 2010). Hand gestures are widely used in our everyday interactions and may be a valuable tool for interacting with technology (Ballendat, Marquardt and Greenberg 2010) that can be embedded in the environment to leverage a user's expected routines supporting what a person would naturally do in that location (Greenberg, Boring, Vermeulen and Dostal 2014). This kind of interaction raises a set of technical and social issues related to unintended activations (Greenberg 2010), social acceptance (Montero, Alexander, Marshall and Subramanian 2010) and the pervasiveness of the technology (Greenberg, Boring, Vermeulen and Dostal 2014). Furthermore, commands and gestural sets are strongly dependent on the task and on the context of use (Stern, Wachs and Edan 2008). It is, therefore, of paramount importance to define a set of gestural commands that is appropriate for the user in first place but also for the context of use. The aim of the present experiment was to develop a gestural vocabulary involving target users. The case study examines the integration of infra-red (IR) technology into a hood for the control of the spotlight.

2. Related work

In defining gestural vocabulary for interacting with machines two approaches have been suggested: one focuses on the constraints imposed by the technology (i.e., Technology

Based Gesture Vocabulary), the other one involves the user to inform the design on her (Human Based Gesture Vocabulary). The former ensures that the gestures are easily recognized by the system, however the single commands are likely to be difficult for the user to be learnt and executed. On the other hand, the Human Based Gesture Vocabulary approach makes users participate in the act of defining the gestures and linking them to the different functions, resulting in a set of gestures that is usually easier to master for the final user (Nielsen, Störring, Moeslund and Granum 2008; Ruiz, Li and Lank 2011). An alternative method was proposed by Brereton, Bidwell, Donovan, Campbell and Buur (2002). They played short video fragments illustrating people engaged in different situations gesticulating. In a following discussion, participants identified five themes representing the different gestures. Stern, Wachs and Edan (2008) introduced an automatic tool for the classification of hand gestures for controlling a computer game. Despite Human Based Vocabulary Approach should be favored (Nielsen, Störring, Moeslund and Granum 2008), often technological constraints prevail and the set of commands is defined by researchers. Previous studies that have addressed gestural interfaces have mainly involved the interaction with tabletop surfaces (e.g., Rekimoto 2002; Epps, Lichman and Wu 2006) and mobile devices (e.g., Ruiz, Li and Lank 2011; Park et al. 2011). Regarding the interaction hands-free in the domestic context, Do, Jang, Hoon Jung, Jung and Bien (2005) introduced “Intelligent Sweet Home”, a system capable to recognize predefined users’ hand gesture for selecting and operating different home appliances (e.g., opening and closing the curtains). Bobeth et al. (2012) compared different freehand interaction techniques for operating a TV addressed to older adults. Garzotto and Valoriani (2013) proposed a predefined set of gestural commands for operating an oven based on the system’s technical characteristics.

3. Equipment

A virtual simulation of the prototype was developed using Flash CS5.5. The virtual simulation consisted of the representation a hood and a hob according to the standard proportion. The simulation included a dark bar resembling an IR detector and the reflection of a spotlight on the stove (Figure 1). The virtual simulation allowed the experimenter to control the position, the size and the brightness of the spotlight on the stove.

The virtual simulation was rear projected (for a detailed description please refer to the Setting section) on a dedicated white screen (2.5 m height by 2.1 m width), by means of an Epson EH-TW3600 projector. Three video cameras were employed for recording users' interaction and the screen projection was also video-recorded. The four video streaming were combined by means of 4-channel Edirol mixer.

Figure 1 The virtual simulation of the prototype of the lighting system integrated into the hood.

4. Setting

The experiment took place in one of the laboratories of HTLab, located in Padua. The dedicated room was arranged with a white canvas for the rear projection, a projector and three cameras for the recording (Figure 2). One camera (Number 1 in Figure 2) provided a pan framing, which yielded a clear view of the wider movements participants made. Another camera was placed on the right side of the screen at an approximate height of 1.40 m and provided a clear view of the participant's hands[8] (Number 2 in Figure 2). A third camera was placed on the left-hand top corner of the screen at an approximate height of 2 m and provided a framing from above (Number 3 in Figure 2). The virtual simulation projected on the screen was also recorded and the four video framings were combined together (Video recording station, Number 4 in Figure 2). The projector was placed 2 m behind the screen (Number 5 in Figure 2).

[8] The height of the camera was adjusted for each participant, in order to ensure a clear view of participants' hands.

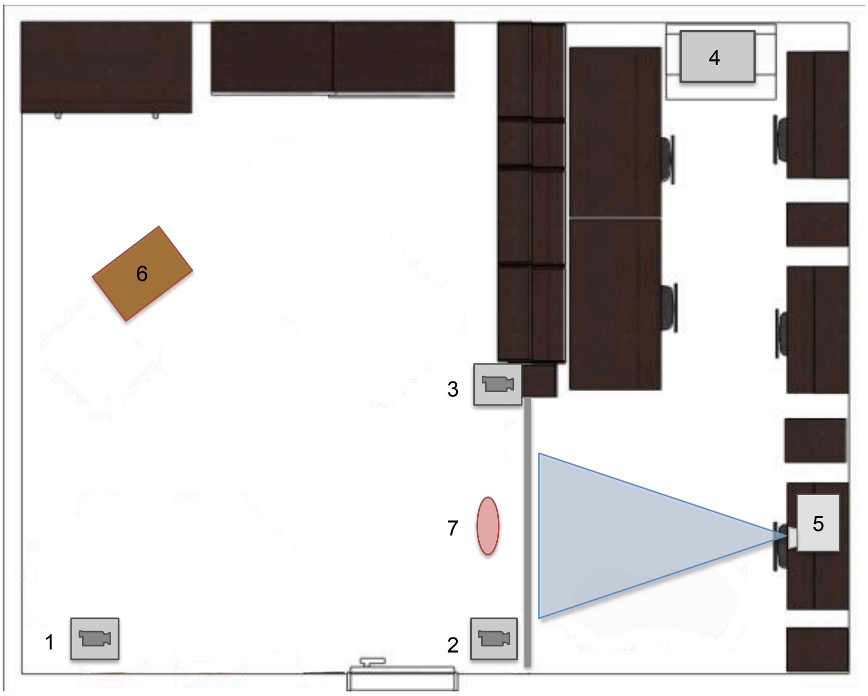

Figure 2 –A sketch of the testing room. Numbers 1, 2 and 3 represent the cameras, Number 4 represents the video recording station. Number 5 is the projetor. Number 6 is the experimenter's position and Number 7 is participant's position.

During the entire experimental session, the experimenter sat out of participants' view, occupying the position indicated by the spot Number 6 in Figure 2. Participants stood in front of the screen (Number 7 in Figure 2), the only instruction they received was that the gesture detection bar would not be able to recognize gestures when the hands were at more than 10 cm.

Such setting allowed users to stand in front of a natural size (60 cm width and 150 cm height) simulation of a hood and to interact with it (Figure 3).

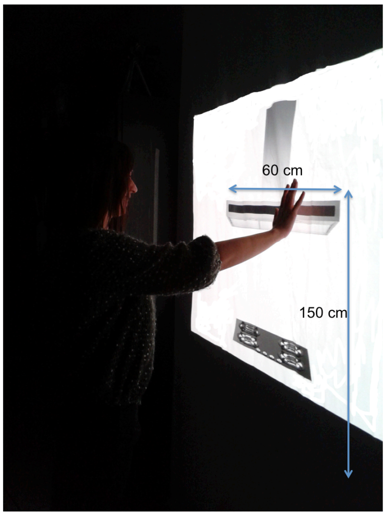

Figure 3 A user interacting with the virtual prototype.

5. Experimental Procedure

On the day of the test, the participant was first welcomed in the laboratory and received a brief introduction regarding the activity. After, she read and signed the informed consent and a non-disclosure agreement. The experimenter then invited the participant to move in front of the canvas. Instructions were read aloud and when the user confirmed that everything was clear, the experiment began. The experimenter sat out of the participant's sight for the entire experimental session.

Before each animation, the participant was asked to watch carefully the animation on the canvas. When the animation ended, the experimenter shortly described what has just happened and asked the experimenter to perform the gesture she would spontaneously do to activate that specific function. In total, eight functions were addressed: (a) power on; (b) power off; (c) moving the spotlight rightward; (d) moving the spotlight leftward; (e) increasing the size of the spotlight; (f) decreasing the size of the spotlight; (g) increasing the brightness of the spotlight; (h) decreasing the brightness of the spotlight. To avoid order effect, the animations were proposed in a counter-balanced order.

Finally, participants had to fill in a brief questionnaire assessing their previous experience with touchscreens and gestural interfaces. Moreover, they were asked to list the devices they thought had influenced the gestures they did.

6. Sample

A sample of 34 people (15 females) was recruited for in the present experiment. The mean age of the user group was 32.2 years old (*SD*= 14.20). All the participants but two were right handed, one was left-handed and one was ambidextrous. None of them had previous experience with touch-less interfaces. Participants were recruited by word of mouth.

7. Analysis

The video-recordings obtained during the experimental trails were analyzed offline by means of the software the Observer XT 11 by Noldus, which allows a systematic coding of all the events occurring in a video.

The events selected for coding were (a) the shapes of the hand(s) and the movements employed to perform each gesture; (b) the direction of the movements or the area on the bar to which users directed the gestures; (c) the hand that was used to perform the gesture; (d) if participants interacted correctly with the interface, i.e. they maintained a distance lower than 10 cm, if they touched the interface or made the gestures at a distance greater than 10 cm.

8. Results

8.1 Taxonomy of the gestures

The gestural commands prompted by participants during the testing trails were grouped in three categories: (a) gestures made with the fingers; (b) gestures made with the hands and (c) gestures made with the arms. The first category comprises the gestures involving specific strokes with fingers (e.g., pointing). In the second category are included the gestures performed with the whole hand (e.g., tapping on the detection bar). Finally, in the third category, the gestures made with wide movements of the arm(s) were scored.

Gestures with Fingers

Within the present category are included the gestural commands made with a single finger (namely, *Finger Pointing*), with a combination of two fingers (namely, *Thumb-Index* and *Pinch*) and by snapping the thumb and the index finger (namely, *Fingers Snapping*).

With *Finger Pointing* we denote the gestures participants prompted with the one finger (index or middle finger) pointing toward a specific area of hood (Figure 1). Within this category a main distinction between static commands and movements can be made.

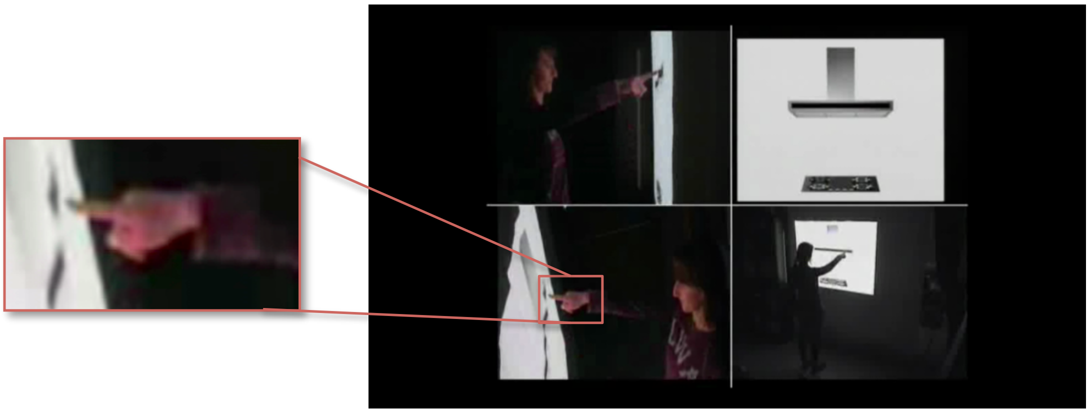

Figure 1 An example of a hand gesture in the category *Finger Poiting*

In the former sub-category we scored:

- *Simple Pointing*, that is the user moved finger pointing to indicate a specific area of the interface (e.g., the user pointed the right-hand corner for moving the spotlight to the right);
- *Pressure*, that is the user mimics to press a button on the interface (e.g., the user simulate to press a button placed on the central area of the hood);

In the latter sub-category we scored:

- *Pointing with Horizontal Movement*, that is the user pointed one finger towards the interface and then moved the hand horizontally (e.g., the user pointed the finger to the central area of the interface and then moved the hand to the right-hand corner of the interface for shifting the spotlight to the right);
- *Pointing with Vertical Movement*, that is the user pointed one finger towards the interface and then made a vertical movement (e.g. the user pointed one finger to

the central area of the interface and then moved the hand upwards for increasing the light intensity).

With the label *Thumb-Index* we refer to the gestures participants made using both the thumb and index fingers. Within this category two types of gestures can be distinguished on the basis of the movements participants made: *Thumb-Index* (Figure 2), when participants simulated to manipulate a small leverage and *Pinch* when participants opened and closed the two fingers.

In the former sub-category we scored the following gestures:

- *Thumb-Index with Horizontal Movements*, that is the user simulated to manipulate a small leverage making either a wide or small horizontal movement (e.g., the user simulate to move the leverage slightly rightward for switching the hood's spotlight on);
- *Thumb-Index with Vertical Movements*, that is the user simulated to manipulate a small leverage making small vertical movements (e.g., the user simulate to move a small leverage slightly upwards to increase the intensity of the light).

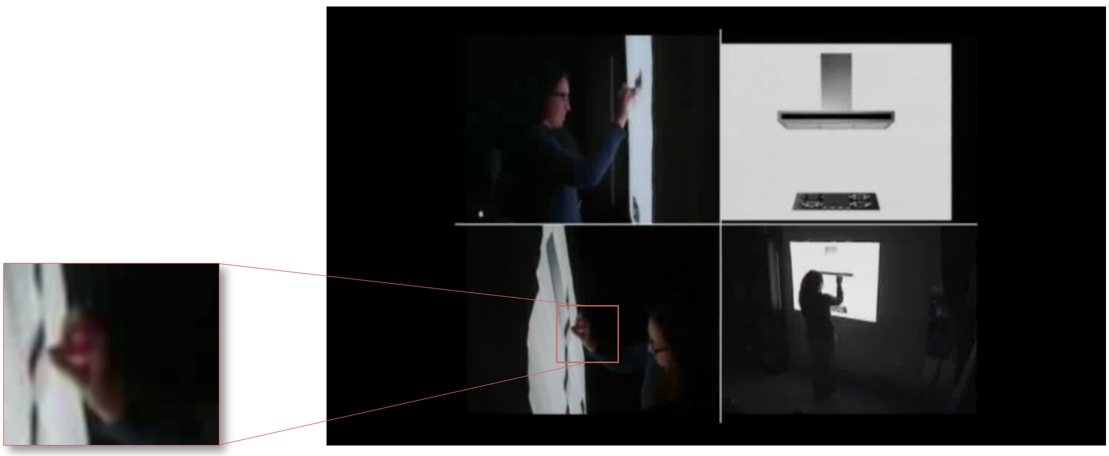

Figure 2 An example of hand gesture in the category *Thumb Index*

In the sub-category labeled *Pinch* (Figure 3) we scored:

- *Pinch Opening*, that is the thumb and index fingers are placed in front of the interface touching each other then are moved apart (e.g., the user put the thumb and index on the central area of the interface and then move them apart for increasing the size of the spotlight);
- *Pinch Closing*, that is the thumb and index fingers are in front of the interface and then on finger is moved towards the other (e.g., the user moved the thumb and index to the central area of the hood and then moved the two fingers to connect them for reducing the size of the hood's spotlight).

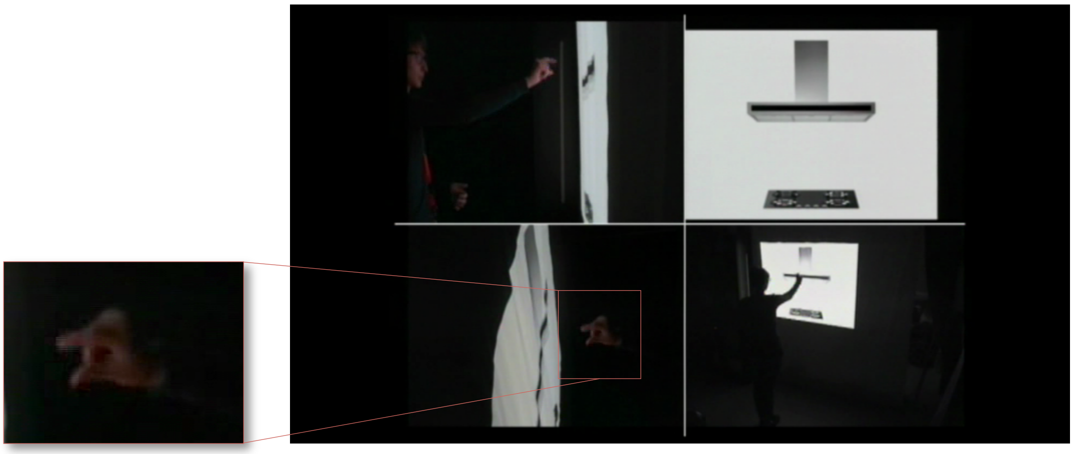

Figure 3 An example of hand gesture in the category *Pinch*

Figure 4 An example of a hand gesture in the category *Fingers Snapping*

Finally, the category *Fingers snapping* (Figure 4) includes those gestures participants made snapping the thumb and index finger in front of a specific area of the hood (i.e., the user snapped the fingers in front of the left-hand corner of the interface to shift the spotlight to the left).

Gestures with hands

The general category of gestures made with the hand include gestures made with the hand open and facing the interface (namely *Vertical Hand*); gestures made with the hand placed perpendicularly with respect to the interface (namely *Hand Upright*); gestures made with the hand open and the palm upwards (namely *Palm Upwards*); and finally commands input with the hand shaping a semicircle (namely *Semicircular Hand*).

Commands within the category labeled *Vertical hand* (Figure 5) include:

- *Vertical Hand with Horizontal Movement*, that is the participant put the open hand in front of the hood and then moved it horizontally (e.g. the user moved the hand horizontally towards the rightwards for moving the spotlight to the right);
- *Vertical Hand with Vertical Movement*, that is participant put the open hand in front of the hood and then moved it upwards or downwards (e.g., the user put the hand in front of the central area of the interface and then moved it upwards to increase the brightness);
- *Vertical Hand with Circular Movement*, that is the user put the open hand in front of the interface and then moved it clockwise or counterclockwise (e.g., the user put the hand in front of the central area of the hood and the moved it counterclockwise to switch the light off);
- *Vertical Hand with Pressure*, that is the user put the hand in front of a specific area of the interface and simulated to press a big button (e.g., the user put the open hand on the left-hand corner of the interface and simulated a prolonged pressure for moving the spotlight to the right);
- *Vertical Hand Approaching*, that is the user put the hand in front if the interface and then moved it closer to the interface (e.g., the user put the hand in front of

the central area of the interface and then moved it closer for increasing the brightness);

- *Hand Tapping*, that is the user placed the hand in front of the interface and tapped slightly towards the interface (e.g., the user put the hand in front of the central area of the interface and then tapped to decrease the brightness);
- *Opening/Closing the Hand*, that is the participants put the hand in front of the interface and then opened or closed it (e.g., the user put the open hand in front of the central area of the interface and then closed it for reducing the size of the spotlight);

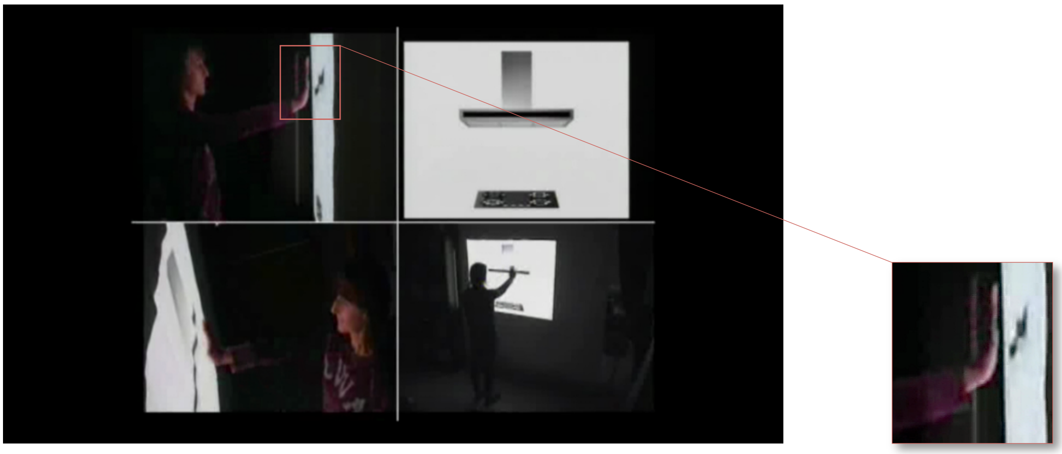

Figure 5 An example of hand gesture in the category *Vertical Hand*

- *Grabbing*, that is participants put their hand with the fingers bent and then moved it vertically or horizontally, as if they were grabbing something (Figure 6; e.g., the user put the hand with the fingers bent in front of the central area of the interface and the moved it to the right-hand corner of the interface to shift the spotlight to the right);

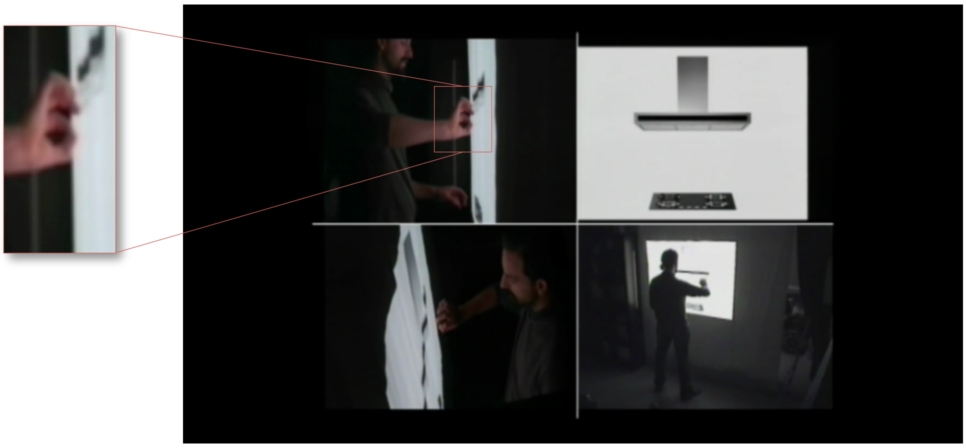

Figure 6 An example of hand gesture in the category *Grabbing*

- *Vertical Hands Moving Horizontally*, that is the user put the hands in front of the central area of the interface and then moved them to the interface's corners or vice versa (Figure 7; e.g. the user put the hands on the central area of the interface and then moved them to the interface's ends to increase the size of the spotlight).

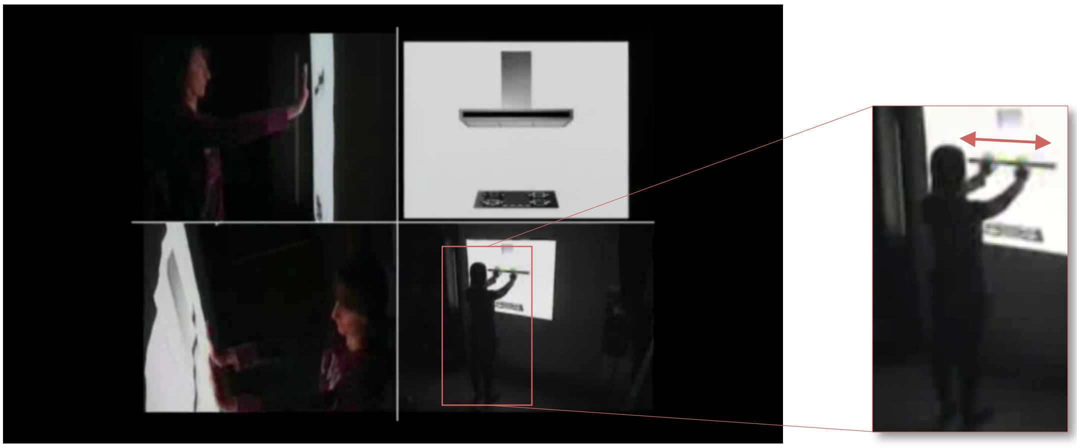

Figure 7 An example of a hand gesture in the category *Vertical Hands Moving Horizontally*

The sub-category *Semicircular Hand* (Figure 8) includes the following:

- *Semicircular Hand - Clockwise Rotation*, that is the user put the hand in front of the interface and then simulate to turn a knob clockwise (e.g., the user directed the hand shaping a semicircle to the right-hand area of the interface and then turned it clockwise to increase the light intensity);

- *Semicircular Hand - Counterclockwise Rotation*, that is the user put the hand in front of the interface and then simulated to turn a knob counterclockwise (Figure 8; e.g. the user directed the hand shaping a semicircle on the central area of the interface and then turned it counterclockwise to decrease the light intensity).

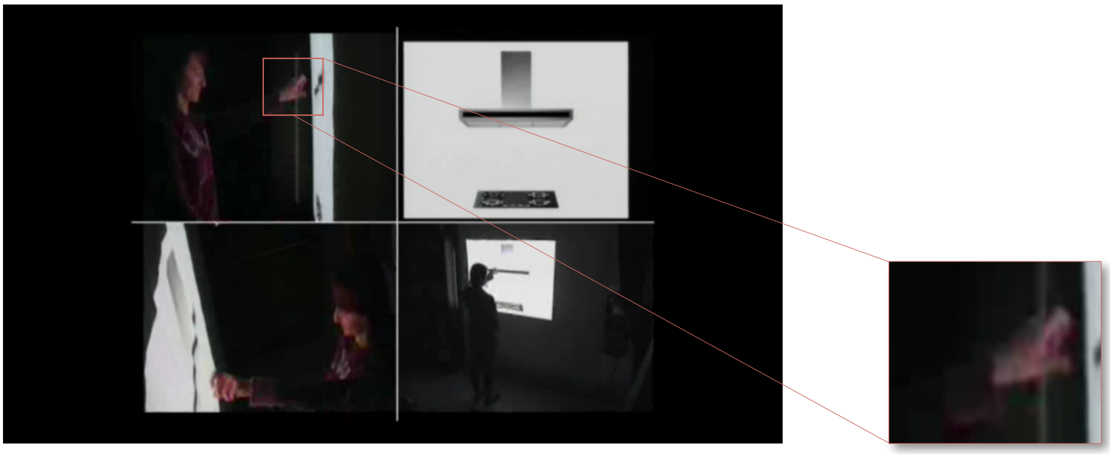

Figure 8 An example of hand gesture in the category *Semicircular Hand*

The sub-category named *Hand upright* (Figure 9) include the following:

- *Hand Upright with Horizontal Movement*, that is the user put the hand upright in front of the interface and then moved it horizontally (e.g. the user put the hand upright and then moved it to the left-hand corner of the interface to move the spotlight to the right);

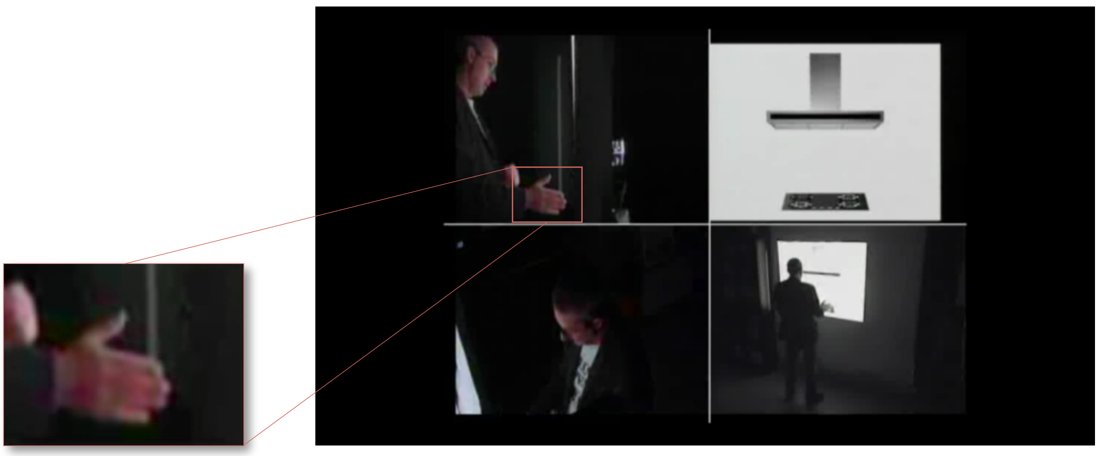

Figure 9 An example of a hand gesture in the category *Hand Upright*

The sub-category *Palm Upwards* (Figure 10) includes:

- *Palm Upwards with Vertical Movement*, that is the user placed the hand with the palm upwards and then made a vertical movement (e.g., the user put the hand in front of the central area of the interface and the moved it upwards for increasing the light intensity).

Figure 10 An example of a hand gesture in the category Palm Upwards

Gestures with the arms

In the category of gestures prompted moving the entire arm falls only one command that has been labeled as *Wave*. This gesture was scored when participants made a wide movement with the arm in front of the hood as if they were drawing a portion of circumference (Figure 11).

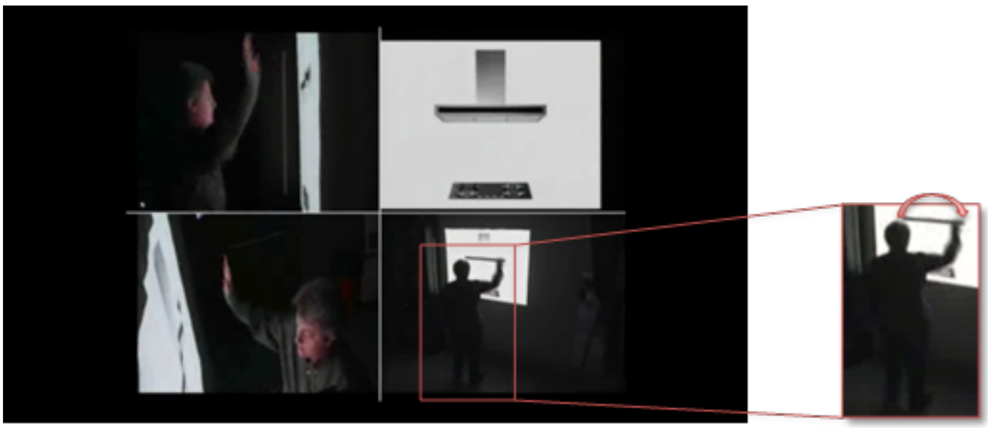

Figure 11 An example of a hand gesture in the category *Wave*

8.2 Frequency of the hand gestures

The frequency with which the different hand gestures were prompted is now considered with respect to the different functions to be activated.

Power on and power off

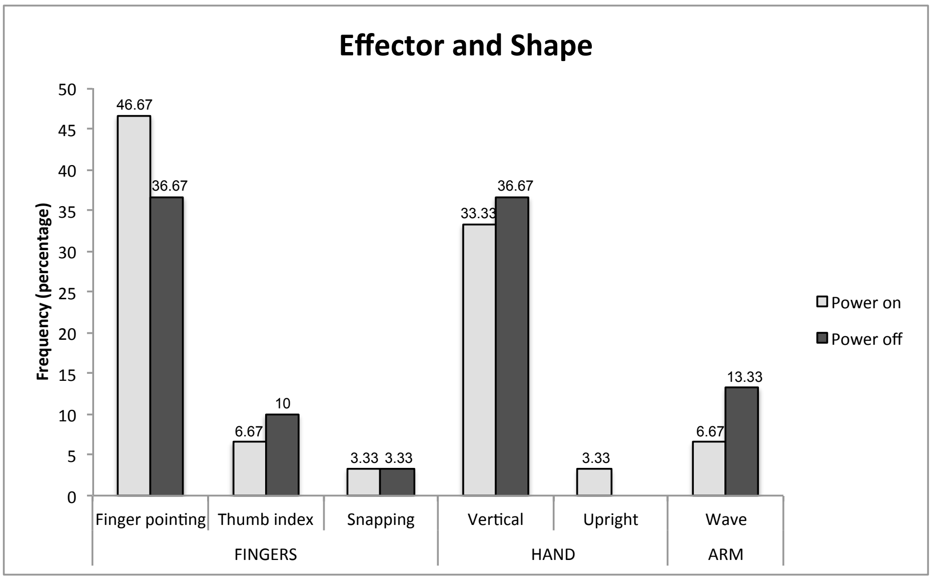

Figure 12 The frequency of the effector used and the hand shape emerged for the task of switching the light on and off

The most frequent gesture for switching the light on is *Finger Pointing*, with the 46.67%; the second most frequent gesture is *Vertical Hand* (33.33%). The 6.67% of participants prompted a gesture either in the *Thumb-Index* category or in the *Wave* category. Finally, the 3.33% of the sample either snapped the thumb and index (*Fingers Snapping*) in front of the interface or put the hand perpendicular with respect to the interface (*Hand Upright*).

For what concerns the switching off of the light, the most frequent gesture is again *Finger Pointing* (33.67%) together with *Vertical Hand* (33.67%). The 13.33% of the users prompted a gesture in the *Waving* category. The 10% of the sample made a gesture in the *Thumb-Index* category. Finally, the 3.33% of the participants snapped the thumb and index fingers (*Fingers Snapping*).

It is worth noticing that the 70% of the users prompted gestures belonging to the same category for switching the light on and off.

Shifting the spotlight

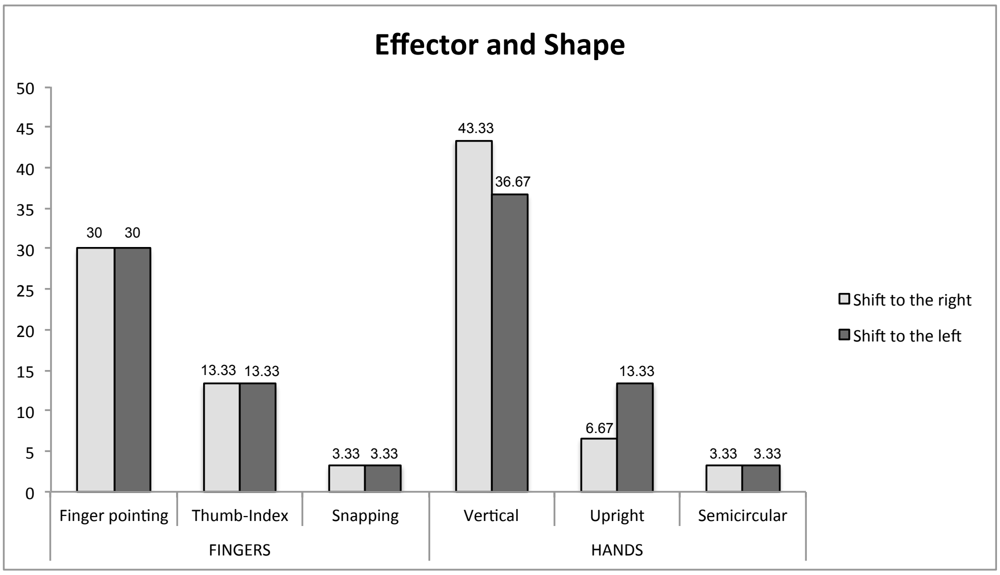

Figure 13 The frequency of the effector used and the hand shape emerged for the task of shifting the spotlight rightwards and leftwards

When asked to move the spotlight to the rightwards, the 46.67% of the users made a gesture with the *Vertical Hand*. The 30% of the users prompted a gesture within the *Finger Pointing* category; the 13.33% of the participants made a gesture belonging to the *Thumb-Index* sub-category. The 6.67% of the sample made a gesture with *Perpendicular Hand*. Finally, the 3.33% of the users input a gestural command within the *Fingers Snapping* sub-category.

Consistently, the most frequent gesture for shifting the spotlight leftwards belonged to the *Vertical Hand* category (36.67%). The second most used gesture (30%) involved *Fingers Pointing*. The 13.33% of users either used a *Thumb-Index* or a *Perpendicular Hand* gesture. Finally, the 3.33% either snapped the fingers (*Fingers Snapping*) or made a gesture within the *Semicircular Hand* category.

Again, participants were highly consistent in the gestures they prompted for moving the spotlight in the two directions: the 93.3% of the users made the same type of gesture in both cases. Please note that here the direction of the movement or the area on the interface on the which the gesture was performed is not considered here; it will be treated in section 8.3.

Modulating the intensity of the light

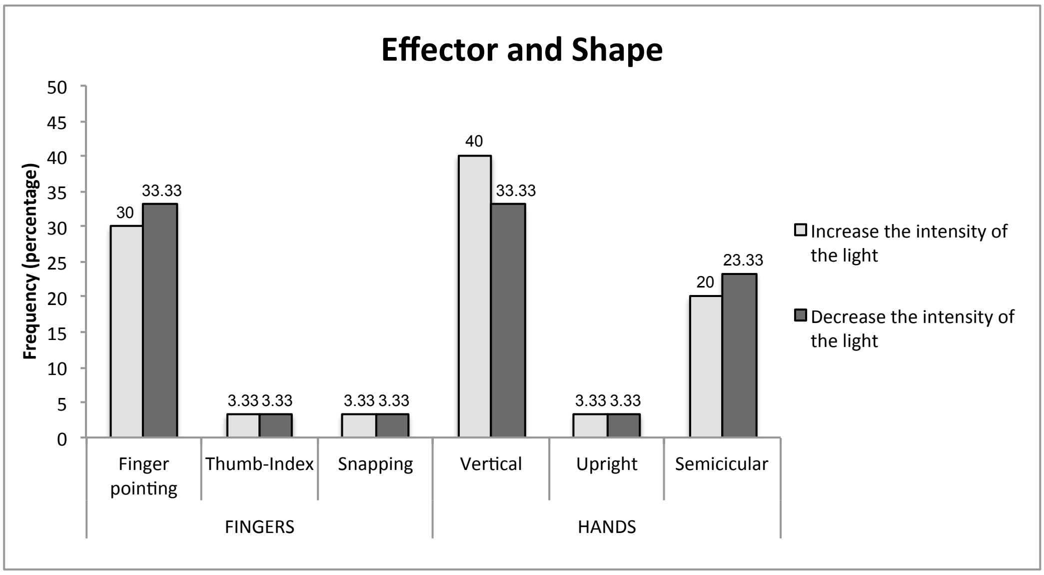

Figure 14 The frequency of the effector used and the hand shape emerged for the task of modulating the intensity of the light

For increasing the intensity of the light, the most frequent gestural command prompted by users belonged to the *Vertical Hand* category (40%). The second most frequent gesture used was *Finger Pointing* (30%). The 20% of the gestures made fell within the *Semicircular Hand* category. Finally, the 3.33% of the participants either made a *Thumb-Index* gesture snapped the fingers (*Fingers Snapping*).

For decreasing the intensity of the light, the most frequent gestures were in the *Vertical Hand* and in the *Finger Pointing* categories (33.33% each). The 23.33% of users prompted a gesture in the *Semicircular Hand* category. Again, the 3.33% of users either made a *Thumb-Index* gesture snapped the fingers (*Fingers Snapping*).

Again, participants tended to be consistent in the gestures they prompted for both increasing and decreasing the intensity of the light: the 73,3% of participants made consistent gestures.

Modulating the intensity of the light

For both increasing and decreasing the size of the spotlight, the most frequent gesture (53.33%) belonged to the *Vertical Hand* category. The second most used gestures prompted involved the fingers, either with *Finger Pointing* and *Thumb-Index* gesture (16.67% each) The 6.67% of the users made a gesture within the *Semicircular Hand* category. Finally, the 3.33% of the users either snapped the fingers (*Fingers Snapping*).

In this case, the 100% of the participants made the same type of gesture for increasing and decreasing the size of the spotlight.

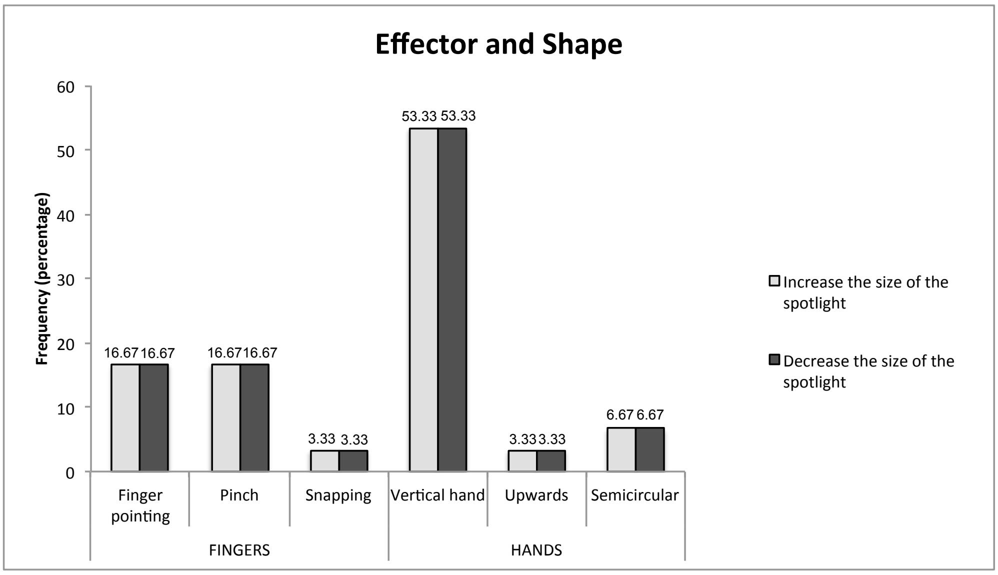

Figure 15 The frequency of the effector used and the hand shape emerged for the task of modulating the size of the spotlight

Summary

Overall the two most frequent hand gestures emerged were *Vertical Hand (*42%*)* and *Finger Pointing* (26.25%). For each of the functions considered, more than the 50% of the gestures emerged belonged to either the *Vertical Hand* or the *Finger Pointing* categories. It is worth noticing that some gestural commands emerged only associated with specific functions, for instance the *Wave* gesture is associated only with the task of switching the light on and off.

Interestingly, when asked to modulate the intensity of the light, the 20% prompted a gestural command that resemble the movement usually done for turning a knob (*Circular Hand*).

Finally, when asked to modulate the size of the spotlight a considerable percentage of users (16.67%) made a *Thumb-Index* gesture resembling the command used on touchscreen interfaces for controlling the zoom functions.

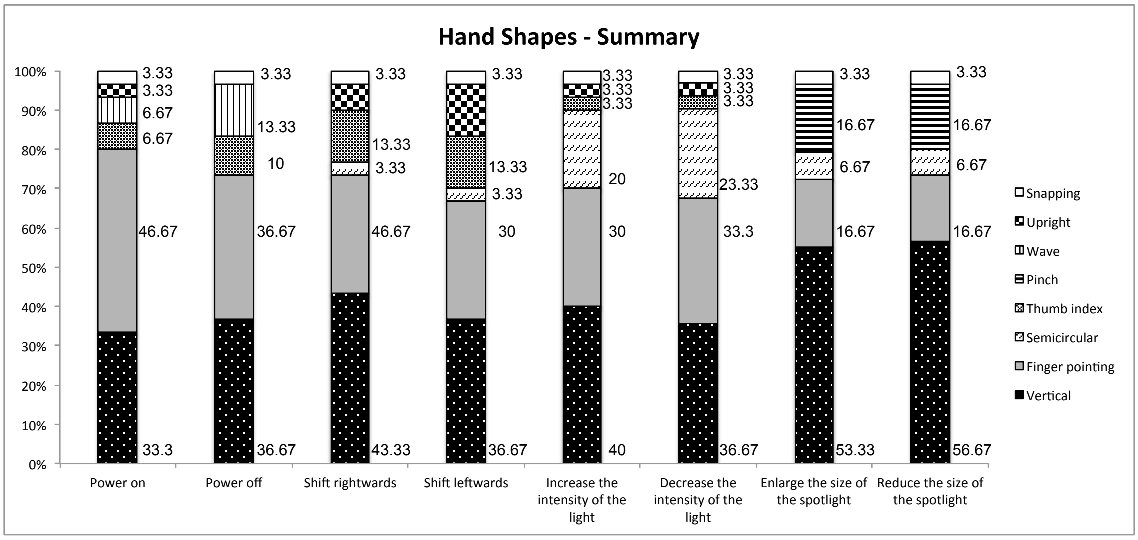

Figure 16 A summary of the different hand shapes emerged for all the functions examined

Following Ruiz et al. (2011), an agreement score for each task was calculated considering only the diverse shapes of hands emerged. The agreement score was calculated as follows:

$$A = \sum_{p_i} \left(\frac{|p_i|}{|p_t|} \right)^2$$

where, t is a task in the set of all tasks T, P_t is the set of proposed gestures for t, and P_i is a subset of identical gestures from P_t. The range for A is [0, 1].

Overall, the agreement scores for the shapes of the hands were quite low, ranging from 0.15 for the Power Off to the 0.29 of the Increasing the Brightness. This finding suggests a high variability in the way in which users approached the detecting bar, and it is consistent with Figure 17, depicting the heterogeneity of the hand postures emerged.

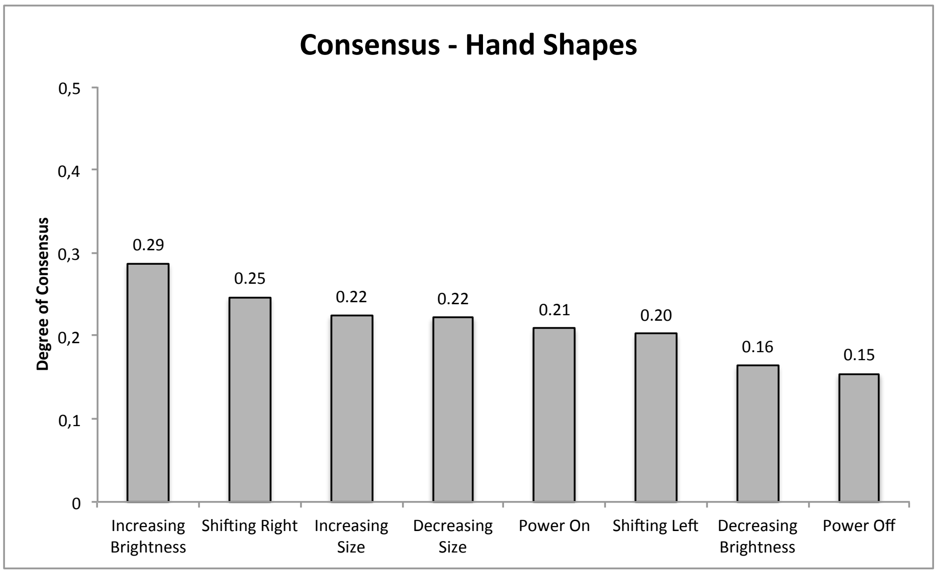

Figure 17 The consensus scores computed for all the functions examined

8.3 Movements

The gestural commands prompted by users can be divided in *Strokes*, i.e., static postures, and *Movements*. Within the latter category three directions were identified: *Horizontal Movements*, *Vertical Movements* and *Rotational Movements*.

All the gestures participants made by simply aiming at the detection bar without any movement, were categorized as *Strokes*. Three areas on which the gestures were directed can be distinguished: the left side of the interface (namely *Left*), the central area of the interface (namely *Center*) and the right side of the interface (namely *Right*).

The *Horizontal Movements* category comprises all the movements made horizontally along the interface. Two directions were identified: from the central area to the right side of the interface (namely *From the center to the right*) and from the central area to the left side of the interface (namely *From the center to the left*).

The category *Vertical Movements* include all the commands participants input by moving the hand(s) on the vertical plane, that is to say upwards or downwards. Two sub-categories were recognized and scored: one includes the movements made from the lower area of the interface to the top (namely *From the bottom to the top*) and one includes the movements users made from the top area of the interface to the lower part (namely *From the top to the bottom*).

The category *Rotational Movements* includes the gestures made to mimic a knob. It includes two subcategories: *Clockwise Rotation* and *Counterclockwise Rotation*.

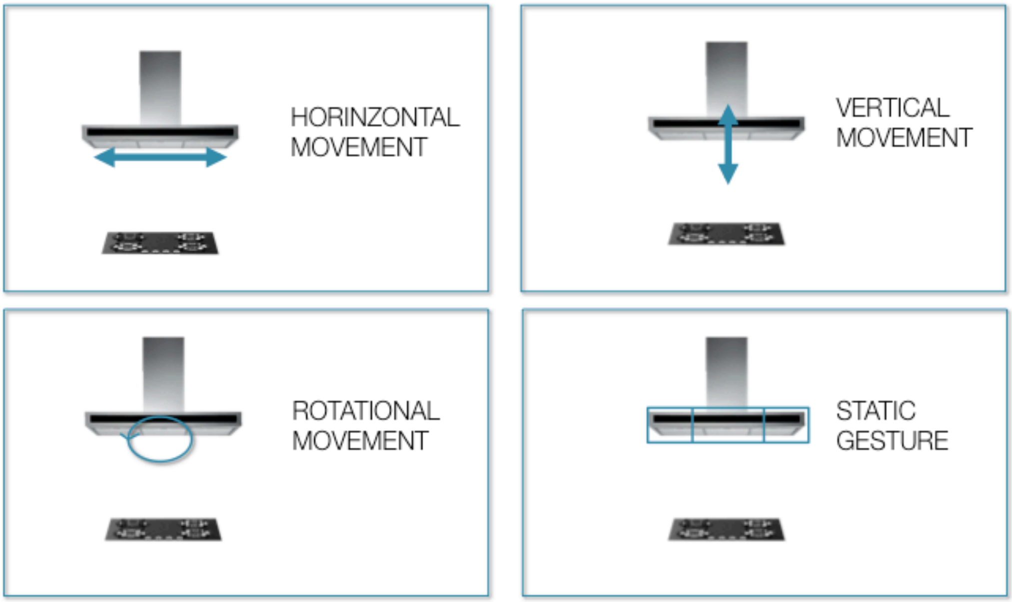

Figure 18 An overview of the categories of movements identified

Power on and power off

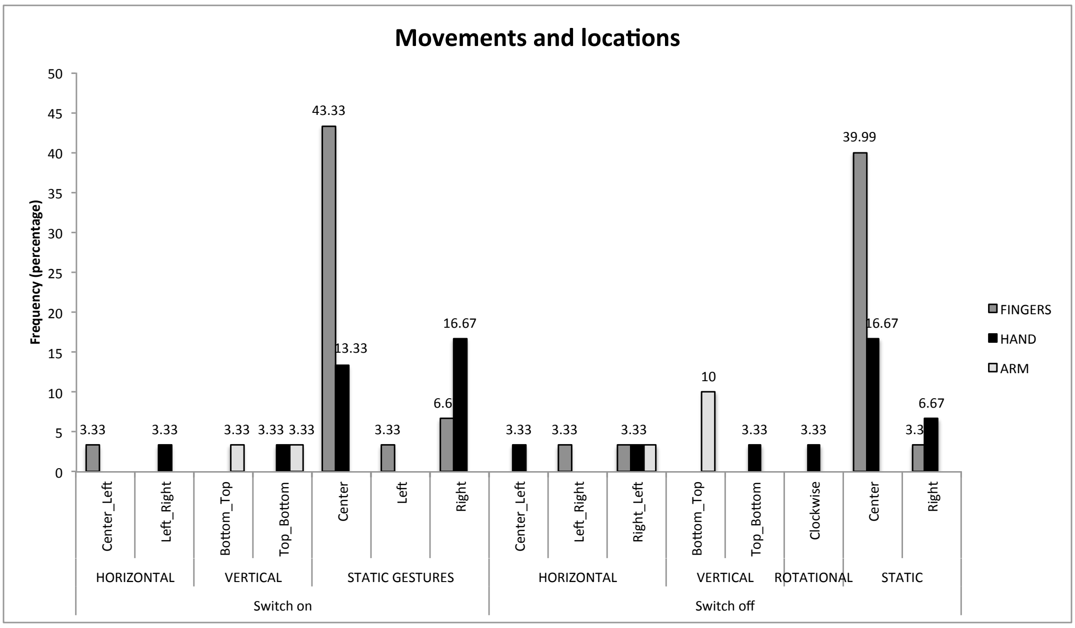

Figure 19 The movements and the location of static gestures for the task of switching the light on and off

For both switching the light on and off the most frequent gestural command chosen was a *Stroke*, mainly addressing the central of the detection bar. In particular, for power on the 56.66% of the gestures prompted were Strokes addressing the central area of the detection bar (43.33% made with the finger and 13.33% made with the hand). The

23.34% of users addressed a static gestural command toward the right-hand side of the detection bar (16.67% with the hand and 6.67% with the fingers). A similar picture emerges for power off: the 56.66% of the static gestures were addressed toward the central area of the interface (39.99% with the fingers and 16.67 with the hand). To a lesser extent compared to power on, participants also addressed the gestures to the right-hand side of the detection bar (6.67% with the hand and 3.33% with the finger). Interestingly, the 10% of the users prompted a gesture with a vertical movement from the lower to the upper part of the gesture detection bar.

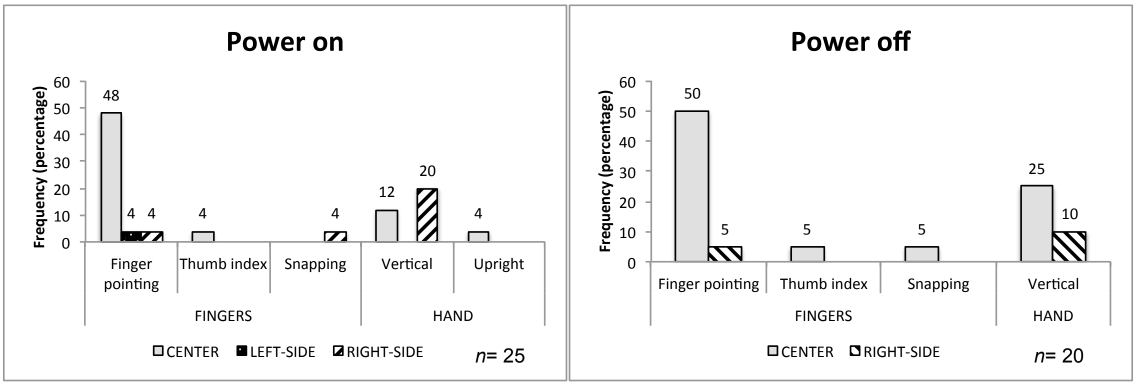

Figure 20 The most frequent hand shapes emerged for switching the light on (Left) and off (Right)

The most frequent hand gesture prompted was *Finger Pointing* toward the central area of the detection bar for either switching on and off the light (respectively 48% and 50%). The second most used gesture was *Vertical Hand,* interestingly it was addressed mainly to the right side of the detection bar when it was employed for power on (20%) and to the central area when it was used for power off (25%).

Shifting the spotlight

For this function, a further distinction is made between *Wide Movements* and *Small Movements*, as it was noticed during the video-analysis that some actions involved the entire width of the detection bar (i.e., *Wide Movements*) while others were oriented on a smaller area (i.e., *Small Movements*).

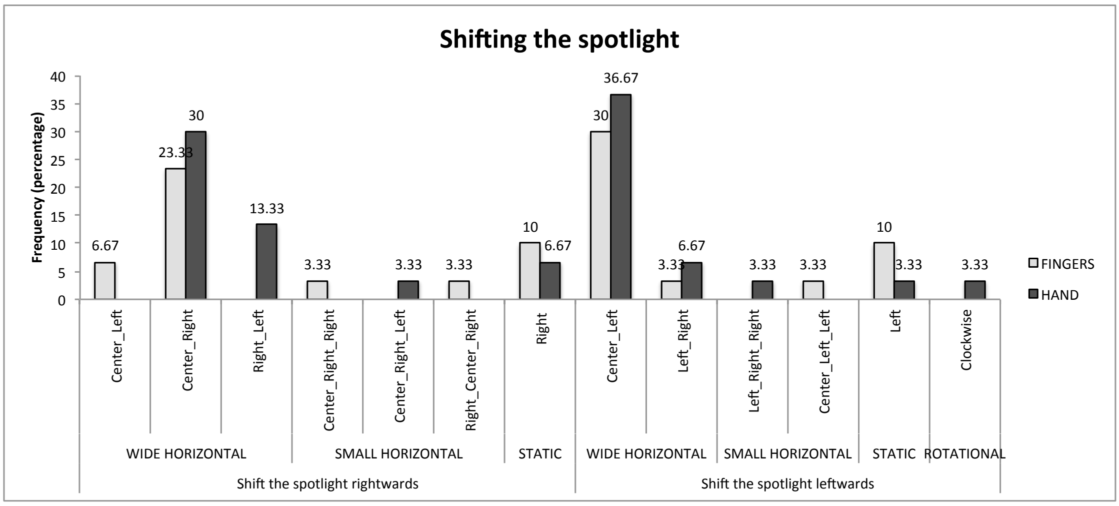

Figure 21 The movements and the location of static gestures for the task of shifting the spotlight leftwards and rightwards

Both for shifting the spotlight leftwards and rightwards wide horizontal movements prevailed, mainly involving the hand. It is also worth noticing that the direction of the movement tended to be consistent with the target position of the spotlight. For shifting the spotlight rightwards, users tended to make a *Wide Horizontal Movement* with the hand (30%) or with the fingers (23.33%). Consistently, for shifting the spotlight rightwards the most frequent movement was a *Wide Horizontal Movement* with the hand (36.67%) or with the fingers (30%).

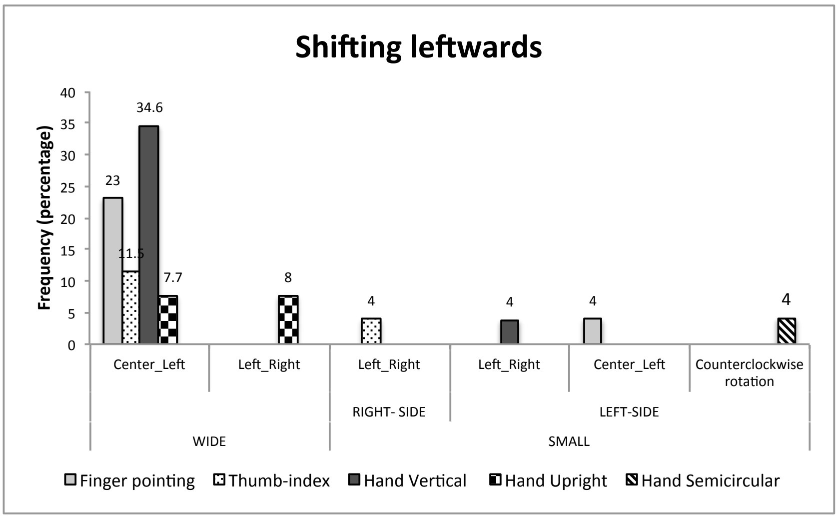

Figure 22 The movements, the locations and the hand shapes prompted for shifting the spotlight to the left

Interestingly, the most frequent hand gesture involved a wide movement from the central area of the detection bar to the left belonged to the *Vertical Hand* category (34.6%), followed by *Finger Pointing* (23%). To a minor extent the 11.5% of participants prompted a *Thumb-Index* gesture and the 7.7% placed their hand upright.

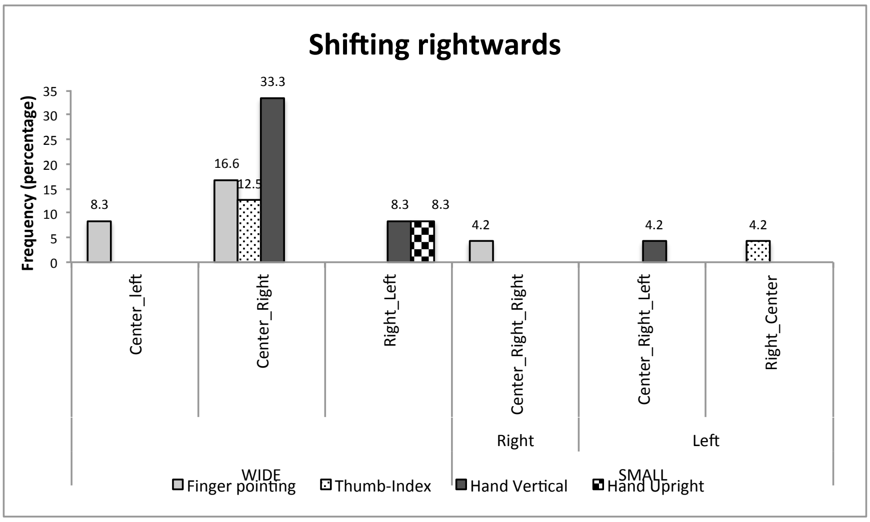

Figure 23 The movements, the locations and the hand shapes prompted for shifting the spotlight to the right

Consistently with the shifting in the opposite direction, the most frequent gesture involved a *Wide Horizontal Movement* from the central area of the detection bar to the right side prompted with the *Vertical Hand* (33.3%). A smaller amount of participant prompted a gesture in the same direction but involving *Finger Pointing* (16.6%) or *Thumb-Index* (12.5%).

Modulating the intensity of the light

Either for increasing and for decreasing the brightness, more than the 50% of the participants prompted a *Stroke* gesture. In particular, for increasing the light intensity the 23.3% of the sample either used the fingers or the hand; while for decreasing the intensity, the 26.7% used the fingers and the 23.3% used the hands. For this function, also *Vertical* and *Rotational Movements* were prompted. For increasing the brightness, the 16.7% of the sample performed a gesture with the hands from the lower to the upper part of the interface and the 10% did the opposite for decreasing the brightness. The

23.3% of the sample prompted a *Rotational Movement* either clockwise (13.3%) or counterclockwise (10%) for increasing the brightness.

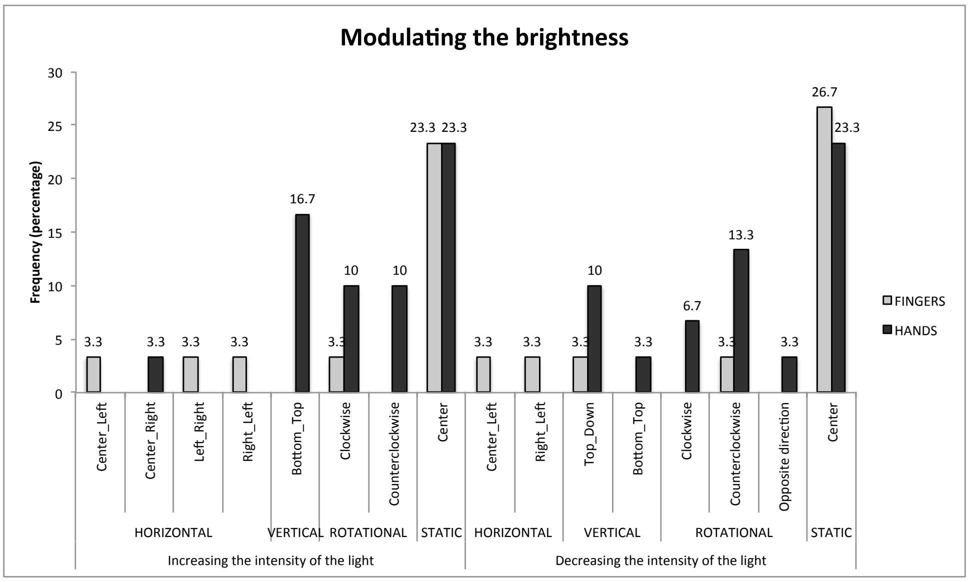

Figure 24 The movements and the location of static gestures for the task of modulating the brightness of the light

Finally, the 16.7% of the participants performed a counterclockwise gesture for decreasing the intensity and the 6.7% did the same in the opposite direction.

Looking more in detail the hand postures associated to the different movements, it is evident that the 56.66% of the sample prompted a gesture with their hand for increasing the brightness. In particular, the 23.33% of them prompted a *Stroke* and the 20% prompted a gesture in the *Semicircular Hand* category. The 13.3% of the sample made a vertical movement with the hand placed vertically. One third of the sample prompted a gesture within the *Finger Pointing* category, the 20% of the participants simply pointed the detection bar and the 10% pointed one finger and the moved the hand horizontally. A similar picture emerges for decreasing the brightness: the 53.33% of the participants prompted a gesture with their hand, the 23.33% of them simply placed the hand facing vertically the detection bar and the 10% moved it upwards or downwards. Finally, the 20% of participants performed a gesture in the *Semicircular Hand* category and made a *Rotational Movement.* Again, one third of the sample prompted a gesture in the *Finger Pointing* category, the 23.3% of them simply pointed at the detection bar and the 6.7% made a vertical movement while pointing.

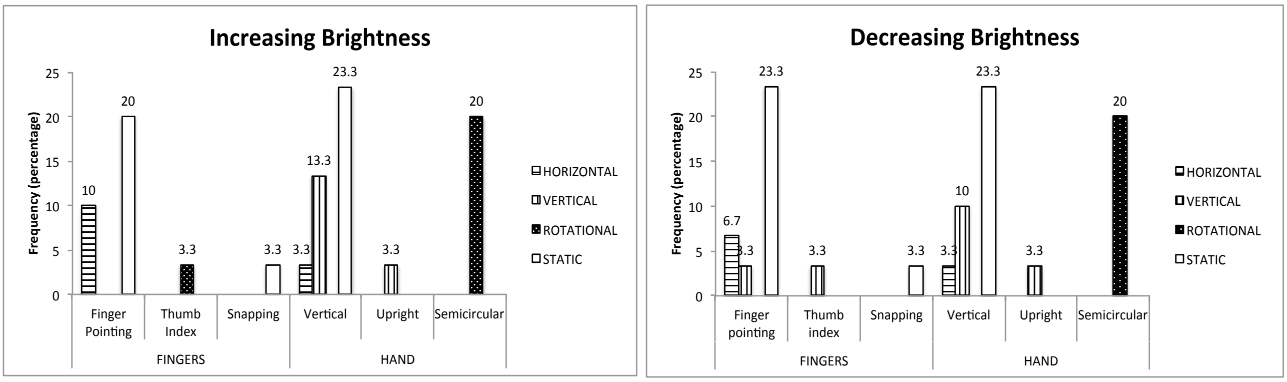

Figure 25 The movements, the locations and the hand shapes prompted for increasing (Left) and decreasing (Right) the brightness of the light

Modulating the size of the spotlight

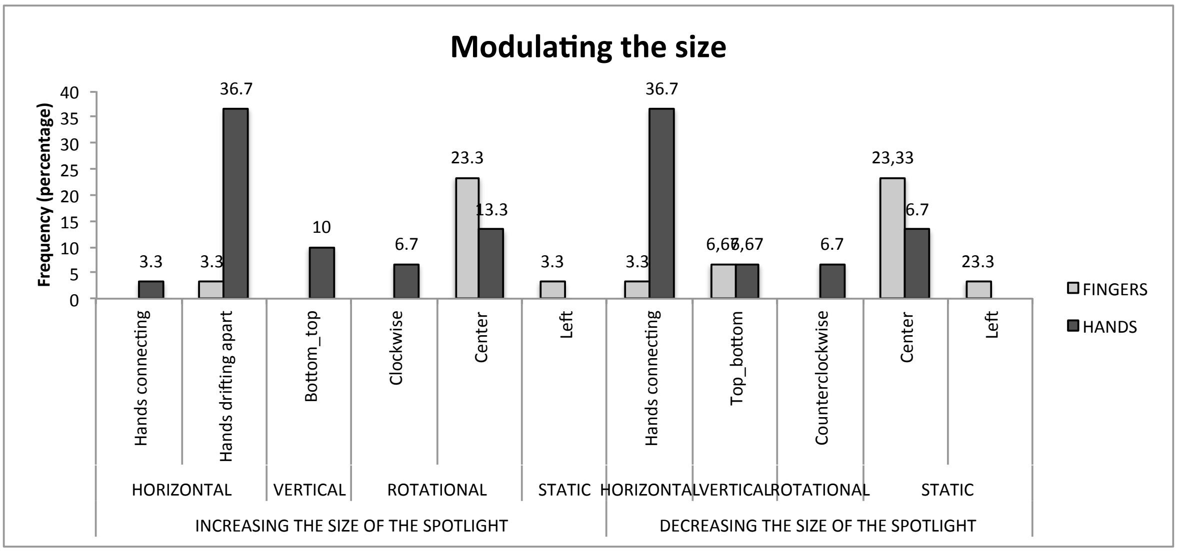

Figure 26 The frequency of the movements and and the directions of the gestures emerged for modulating the size of the spotlight

For modulating the size of the spotlight, the most frequent gesture performed involved a horizontal movement along the detection bar with both hands. In particular, for increasing the size of the spotlight the 36.7% of the participants drifted apart their hands starting from the central area of the detection bar. Exactly the same percentage of participants did the same gesture, though in the opposite direction, for decreasing the size of the spotlight.

Approximately one-third of the users prompted a *Rotational Movement* with the fingers (23.3%) or with the hand (13.3%) for increasing the size of the spotlight. For decreasing the size of the spotlight, one-third of the users prompted a *Stroke* towards the central area of the detection bar with the fingers (23.3%) and with the hand (6.7%).

Interestingly, in this case there was no clear prevalence of movements or strokes for increasing the size of the spotlight. Almost all participants who prompted a stroke for increasing the size of the spotlight addressed the central area of the detection bar either making a Pinch-like gesture (33.3%), placing the vertical hand (33.3%) ore pointing the fingers (16.7%). Those who made a movement, mainly drifted their hand apart starting from the central area of the detection bar (55.6%).

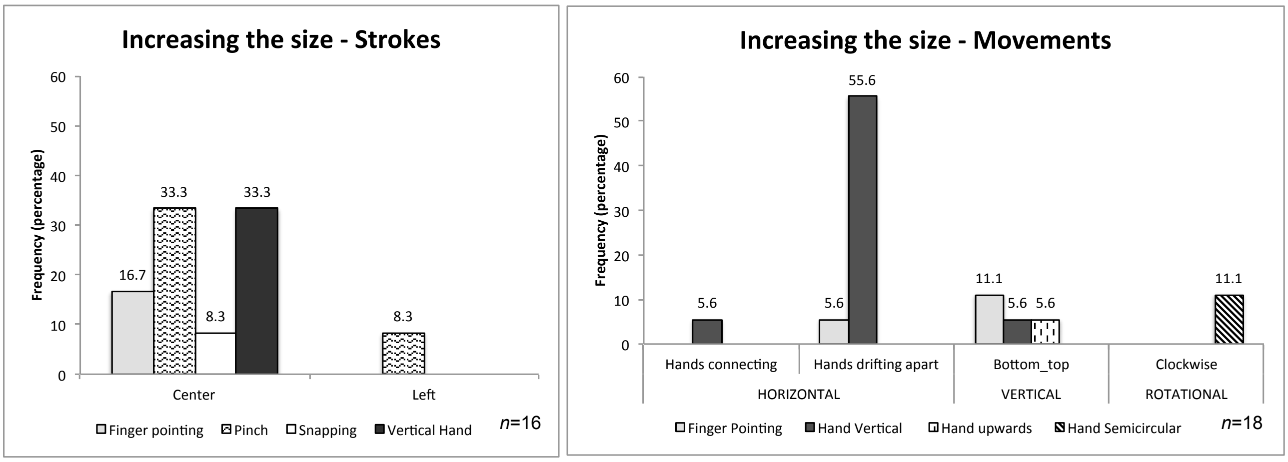

Figure 27 Increasing the size of the spotlight: location of static hand gestures (Left) and type of movements (Right)

A similar picture emerges also for decreasing the size of the spotlight: the participants who prompted a stroke for decreasing the size of the spotlight addressed the central area of the detection bar and mainly made a *Pinch*-like gesture (33.3%), placed the vertical hand (33.3%) or pointed the fingers (16.7%). Those who made a movement, mainly connected their hand on the central area of the detection bar (55.6%).

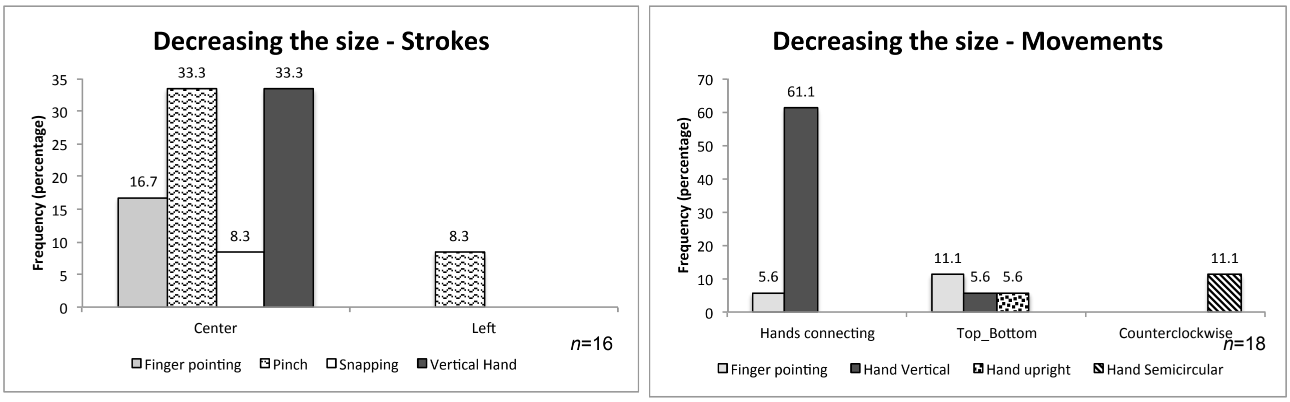

Figure 28 Decreasing the size of the spotlight: location of static hand gestures (Left) and type of movements (Right)

Summary

The gestures prompted for switching on and off the light are strongly characterized by strokes.

Similarly movements along the horizontal dimension of the interface are strongly associated with the task of shifting the spotlight to the sides of the interface. For what concerns the modulation of the light intensity, static gestures are the most frequent, however rotational and vertical movements, even if to a minor extent, emerged. Similarly for the task of modulating the size of the spotlight, about a half of the sample made a static gesture (often directed towards the target position of the spotlight) and the half of the participants made a horizontal movement (differently from the shifting task, they used both hands).

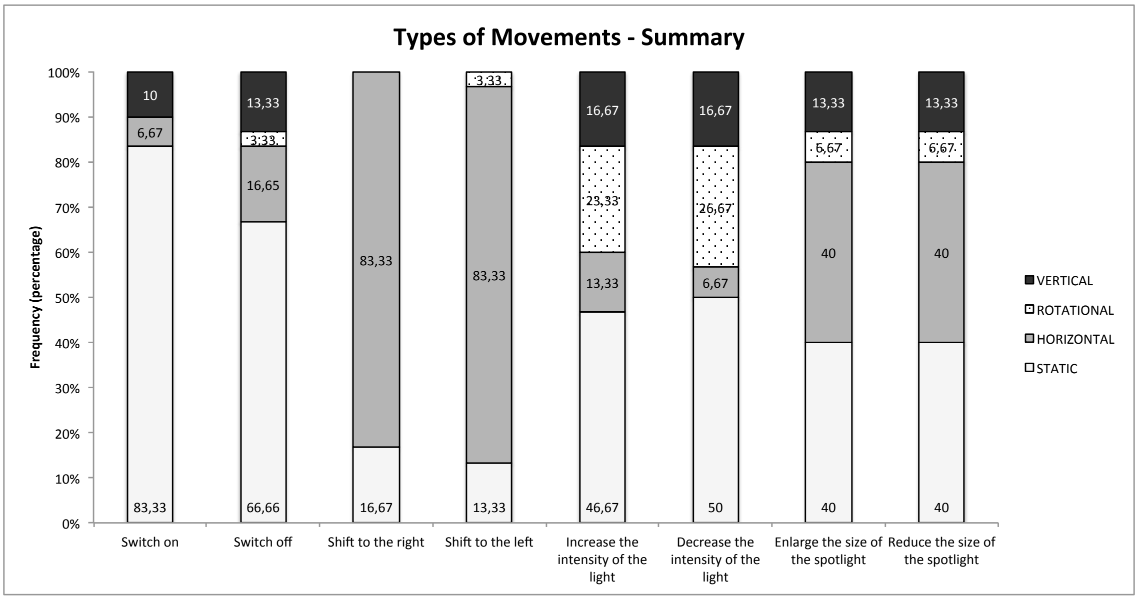

Figure 29 A summary of the type of movements emerged for all the functions considered

An agreement score was calculated also for the types of movement emerged for each task proposed, following the approach illustrated above. Interestingly, consensus emerged among participants for the task requiring the shift of the spotlight within the area of the hood, where movements along the horizontal plane prevailed (0.77 for shifting leftward and 0.72 for shifting rightward). A certain degree of consensus was found also for switching the light on (0.71), but, surprisingly, it dropped for switching the light off (0.49). For the functions related to the modulation of some property of the spotlight (i.e., the brightness and the size), a consensus was not reached among participants, as the agreement scores were quite low.

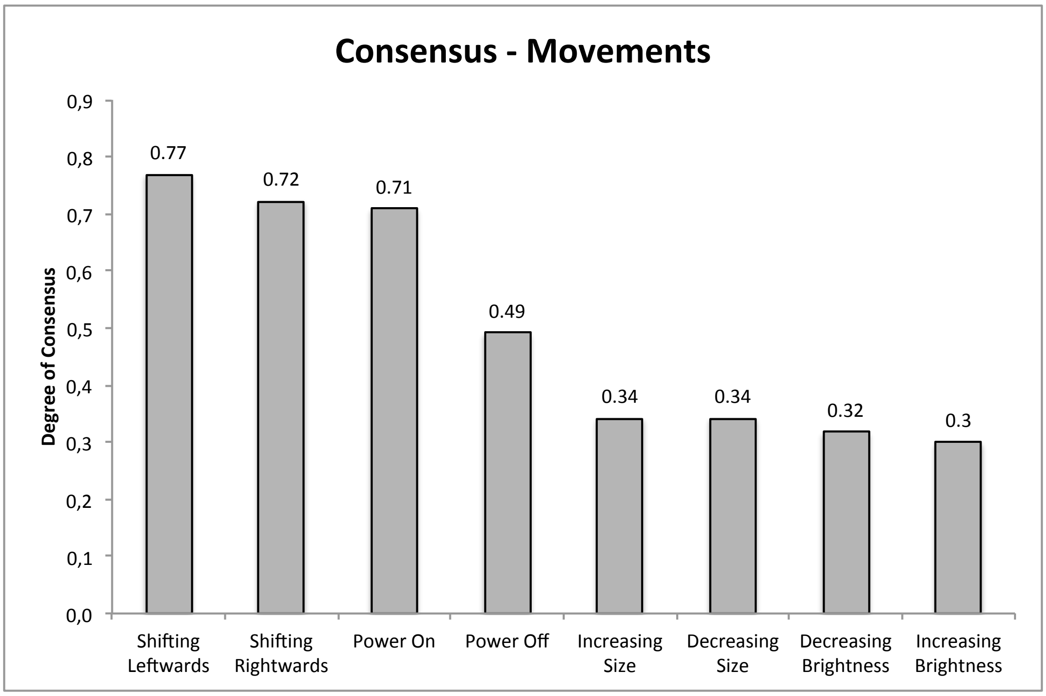

Figure 30 The consensus scores for type of movements emerged for all the functions examined

8.4 Additional observations

Before the test began participants received clear instructions about the hood's operating modalities: they were told that the system was to be controlled by means of hand gestures and that the hands had to be at a maximum distance of 10 cm for the command to be successfully received by the system. Despite those instructions, we noticed that some participants input the commands by touching the dark bar representing the bar for the gesture detection. Other users made the gestural commands with the hand(s) at a distance greater that than 10 cm, a situation that wouldn't have caused the starting of any function with a real hood. This kind of behavior was therefore scored as an error. In general, participants were able to interact with the interface successfully, however they tended to touch the detection bar.

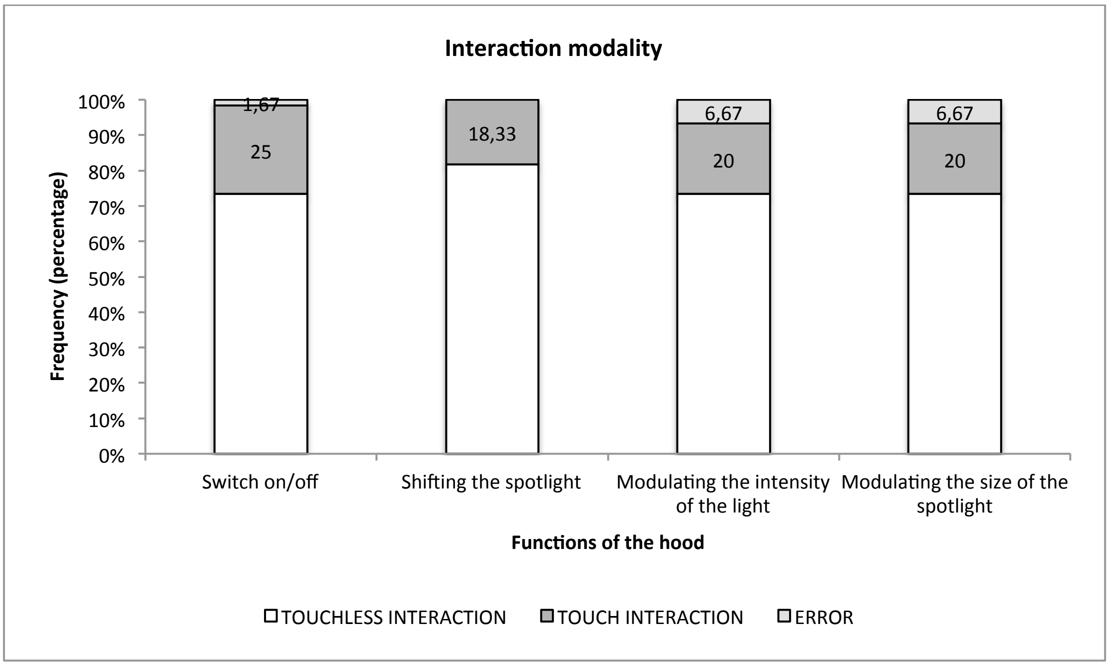

Figure 31 The frequency of the interaction modalities emerged for the functions examined

Participants tended to use their right hand for prompting the commands, which is not surprising given that almost all of them were right-handed. However, what it is interestingly to note that when the request was to shift the spotlight leftwards a consistent part of the sample used their left hand. Also, it is worth noticing that for modulating the size of the spotlight, a considerable amount of participants prompted a gesture with both hands.

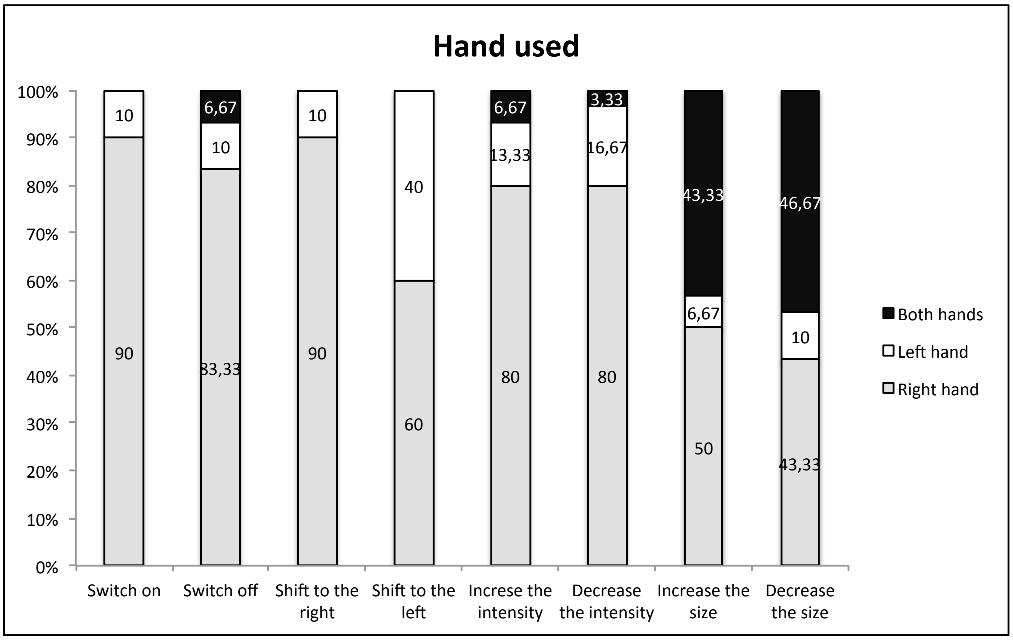

Figure 32 The frequency with which participants prompted the gesture with the right or left hand or with both.

9. Discussion

In the experiment reported in the present chapter we aimed at developing a gestural vocabulary for controlling the light embedded in a hood starting from users' spontaneous motions. The methodology devised involved having a sample representative of the target users interacting with a virtual simulation of the prototype in real size. Participants were shown the expected outcome of their action and then they were asked which gesture they would prompt to produce that particular outcome without any constraints. Participants' spontaneous gestures were categorized according to the shape of the hands and the direction of the movements. For what concerns the shapes of the hands, no clear consensus emerged for all the functions considered, despite the majority of participants tended to either point their finger or to face their palm toward the detection bar. The variability of the hand postures emerged suggests that participants found no clear and evident affordance leading them to prompt a specific gesture, which is a well-known limitation of gestural interfaces (Marquardt and Greenberg 2012). In addition, it seems that when participants cannot rely on visible affordances, they tend to make a pointing-like gesture, as found by Voida, Podlaseck, Kjeldsen and Pinhanez (2005). Interestingly, reference to pre-existing practice seemed to raise for all the functions explored and was also reported by users as an influence from their everyday interactions with their own kitchens and digital devices. This finding supports Jacob et al.'s (2008) claim according to which building interactions over users' pre-existing knowledge of the everyday world, would facilitate the use of the system. Regarding the direction of the movements, consensus emerged among participants for the tasks requiring to shift the spotlight and to power the light on and off. In this case the shape of the detection bar seemed to provide a valuable affordance for directing users' gesture along the horizontal plane. For switching on and off the light, participants seemed not to rely on any specific features on the hood, rather they seemed to mimic the simple pressure of a button.

To sum up, this experiment suggests that presenting the user with the expected outcome of their action and then requiring her to prompt a gesture for activating that specific function is a valuable method for producing spontaneous gestural commands helping to uncover affordances in the interface. However, proving no constraints on the gestures they can perform may result in a high variability of postures and movements.

Chapter 3

Exploring urban landmarks: comparing a haptic-audio and a audio-visual interface

1. Motivation and aim

Nowadays there are a large number of mobile applications and services, e.g., Google Maps, which allow a user to explore autonomously places that are new to her. Such services rely mainly on visual and audio displays, providing the user with maps, pictures and textual and audio description. However, visual displays are often inappropriate in many mobile scenarios, as they are subject to contextual factors, for instance, in strong sunlight screen reflections impacts on content visibility (Jylhä et al., 2015; Wei et al., 2014). Even if audio can overcome those problems, auditory contents tend to raise privacy issues in public spaces and are affected by ambient noise and the use of headphones is likely to isolate the user (Ahamaniemi et al., 2009; Jacob 2011). Furthermore, the constant flow of information can overwhelm users and distract them from the main task, hampering the experience (Hornecker 2011). Ultimately, the user cannot get a simultaneous picture of the object in the real world and its representation on the map, and for some users it may be challenging to relate the two (Ahamaniemi et al., 2009). Tactile feedback has proved to be a useful and direct means for delivering spatial information to users, as it does not interfere with other tasks nor it requires the users' visual attention resources (Tsukada et al., 2004; Asif et al., 2010). The majority of the systems have been developed to guide the user along a predefined route (Jilhä et al., 2015), however in some contexts, e.g., in tourism, a causal discovery of landmarks is an integral part of the experience (Hoenecker et al., 2011). With this respect, a growing number of systems are deploying non-visual stimuli for the serenpitous discovery of point of interests in the surroundings (see the Releted Work section for an overview).

In the present experiment we compared a non-visual interaction technique based on a vibrotactile glove with a traditional audio-visual display for the casual discovery of landmarks in the center of a historical city, i.e., Padua. The research question leading the study was twofold: first we aimed to investigate if a non-visual interface could reliably guide users to indicate a target. Secondly, we were interested to compare the experience of use with an audio-haptic interface and a traditional audio-visual interface.

2. Related work

Robison et al. (2009) developed a hand-held mobile system providing the user with directional vibrotactile feedback based on location data. By holding the device in the hand, the user can scan the surroundings to check the presence of points of interest nearby, getting a vibrotactile feedback when she is pointing toward a relevant location. The feedback notifies the user of both the direction and the quantity of information available about that spot. In a test indoors the prototype was compared with a visual-based system. The results were promising, with the haptic system supporting users to find all the targets, even if the walking speed was faster when participants employed the visual interface. With a similar approach Ahamniemi et al. (2009) developed a system for scanning the environment and identifying targets by means of a vibrotactile feedback triggered by a hand-held device. In their laboratory test, Authors found that the target size and distance significantly affected the selection time. Hornecker et al. (2011) presented a serendipitous tour guide based on an Android device making use of vibrotactile cues. In their approach, the system is initialized by the places and attractions the user wishes to see. The system utilizes a proximity model, and when the user approaches a landmark, she receives a vibrotactile notification. They also tested the addition of generic notification sound, but they found them not useful in the initial tests. A 3D tactile guidance integrated into a glove was developed by Stevenson, Riener and Ferscha (2013), who tested the prototype in a laboratory study. The system was found to be intuitive by users with high completion rate for both interaction modes tested. Wei et al. (2014) experimented different values of amplitude, rhythm and interaction modality (audio, haptic and a combination of the two) in a laboratory test to assess the best configuration for a mobile tourist application. When the two interaction modalities were combined the identification rate was higher, even if the difference was non significant.

3. Equipment

In the present section the two systems compared in the experiment, namely a tactile glove and the baseline application, will be described.

3.1 Tactile glove

The system is based on a tactile glove (Figure 33), which is equipped with vibrotactile actuators and a set of sensors, and an Android mobile device. The glove has been constructed of thin elastic fabric and several electronic components. Two Arduino microcontrollers (Arduino Pro Mini) process the sensor signals and commands, and control the vibrotactile feedback. For sensing hand orientation, the glove is equipped with a 9-axis Inertial Measurement Unit (IMU, InvenSense MPU-9150), which consists of a gyroscope, an accelerometer, and a compass. Flexible bend sensors (Spectra Symbol flex sensor) are deployed on three fingers (thumb, index, and middle finger). A set of three vibrotactile actuators (Precision Microdrives 10mm shaftless vibration motor) is mounted for providing notification cues and directional guidance on the thumb, index, and middle fingers. The glove communicates wirelessly over Bluetooth with an Android mobile phone (Samsung S3) running the application logic. Participants carried the smartphone in a neck pouch during the experiment. An over-ear headset is connected to the Android device to allow the user to listen to the audio content.

The glove has been built for right-handed interactions (Jylhä et al., 2015). The placement and the number of the actuators were informed based on previous pilot studies. The performance in a target selection task was compared when the guiding vibrations were delivered on the palm or on the fingertips. We found that the placement of the actuators on fingertips outperformed the placement on the palm in terms of speed of selection and intuitiveness of the signals.

When the user is in the proximity of a POI, she receives a vibrotactile notification (all actuators are active at the same time) and an auditory icon is played. At this point she can either discard or select the POI, by respectively bending the thumb or the index finger. When a POI is selected, the user receives vibrotactile guidance on the fingertips to allow her to point at the POI. In particular when the actuator placed on thumb is active the user has to move the hand leftward, when the one on the middle finger is active, she has to move the hand rightward. When the user correctly points the POI, she receives a vibrotactile notification and the name of the POI together with a brief introductory description is played back. If the user bends the index finger, the full audio description is played. An auditory icon is played at the end of the audio description.

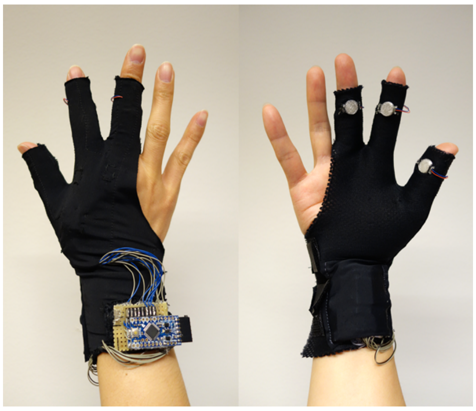

Figure 33 The tactile glove: dorsal view (Left) and ventral view (Right). Figure from Jylhä et al., 2015

3.2 Baseline application

A baseline application was developed *ad hoc* for the present experiment. The application is designed to mimic the most popular mobile apps for the city exploration. The two main features are push notification of a recommended POI, recalling Foursquare 2, and a compass indicating the direction of the POI, imitating Tripadvisor City Guides 3.

When the user is in the proximity of a POI, the application triggers a vibration and a visual notification is shown, suggesting to check out the POI (Figure 34 a). The user can then decide to discard the notification or to access additional information regarding the POI (including text, pictures and audio description), by tapping on the display (Figure 34 b). The application also provides directional indications to spot the POI, however it employs visual feedback – instead of vibrotactile guidance - with an arrow pointing the POI (Figure 34 c).

The application runs on an Apple iPhone 5 connected to an overhear headset, to allow the user listen to the audio description of the POIs.

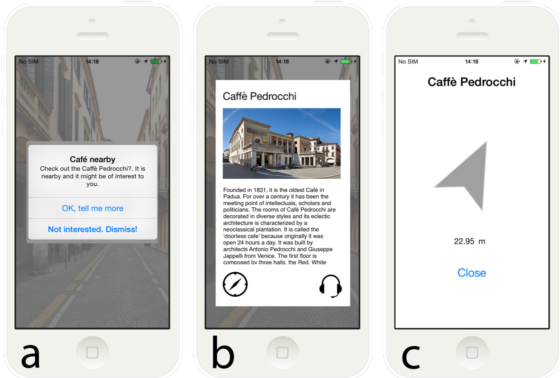

Figure 34 The baseline application. Figure from Jylhä et al., 2015

4. Method

4.1 Experimental design

The experiment followed a between subjects design, with the application as the only independent variable. The primary condition involved using the proposed glove-based interaction technique for exploring the POIs. As a baseline condition we referred to the mobile application described above.

4.2 Setting

The experiment was conducted in the historic center of Padua, in a touristic area with a lot of applicable POIs. Thirteen POIs were chosen within an area of roughly 0.12 km^2. Two extra POIs were selected for training purposes outside the main area (**Errore. L'origine riferimento non è stata trovata.**). The list of POIs comprised different sites, including a church, two cafes, two shops, a statue and facades and towers of historic buildings. For each POI, a tour guide-like content related to the POI in both written and spoken format was prepared. The content included the name of the POI, a short introduction (lasting approximately 10 seconds) to the POI, and a longer (lasting approximately 30 seconds) description. Each description was associated with the corresponding auditory icon representing the category of the POI. To ensure that all the contents were equally accessible to all participants and to avoid them going outside the GPS coverage, we defined a route that the participant had to follow within the area.

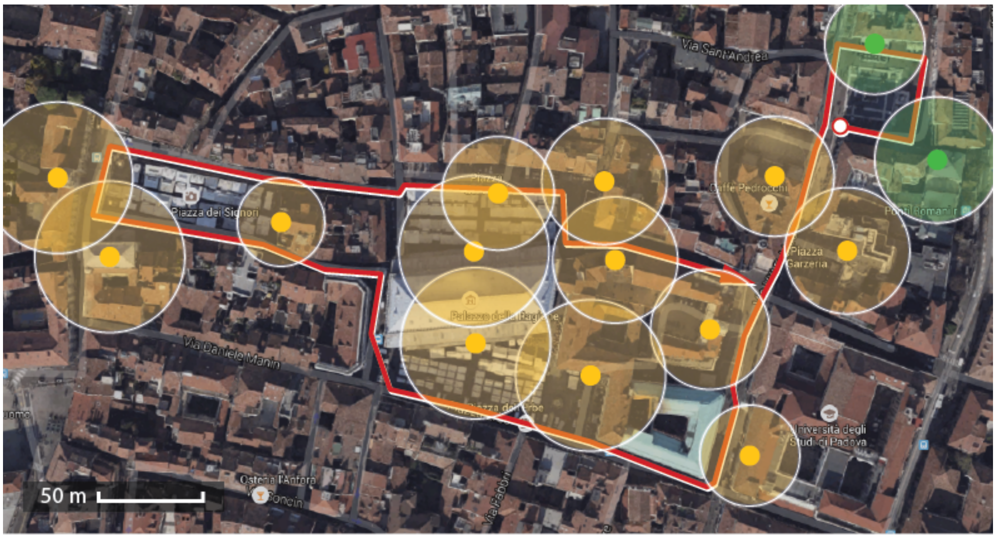

Figure 35 The area in the city center of Padua where the experiment took place. The itinerary is highlighted by a red line. Orange circles represent the experimental POIs. Green circles represent the POIs for the training phase. Figure from Jylhä et al. 2015

4.3 Sample

In total 20 tourists (7 women) participated in the study. All of them but three were Italian, two were from Spain and one was from the United States. The mean age of the sample was 33 years old (*SD*= 15.6). Participants were recruited in the field; none of them lived in Padua and had a high level of familiarity with the city. Nine participants reported that they were used to go sightseeing using paper book guides, nine reported to use applications for smartphone and two Google Maps. Participants were recruited in the field.

4.4 Procedure

Participants first gave informed consent to participate in the study and then filled in a short questionnaire collecting background information. Next, the experimenter explained how the device, either the tactile glove or the baseline application, depending on the experimental assignment, worked. When the user confirmed that the system functioning was clear, she received the device and the training phase began. This stage had the aim of familiarizing the user with the device and the interaction modalities involved. The training consisted of a guided exploration of a small square, on which there were two POIs that the participant had to explore. Participants were free to repeat the training phase if needed. When the user communicated to the experimenter to be ready, the experimental trial started. During the trial, the participant was instructed to follow the experimenter along the pre-defined route (approximately 800 meters in length). After the audio content was concluded, the experimenter asked the participant to which building the narrative referred, to make sure the user was able to relate correctly the description and the object. During the experimental session, the participant was asked to avoid any interaction with the experimenter, with the exception of reporting which building or monument they perceived to be attending at. Participants using the iPhone received no instruction on how to hold it and were left free to keep it in their hand or in their pocket. When the participant reached the end of the route, she was asked to fill in a short questionnaire and to answer a brief semi-structured interview.

4.5 Materials

The measures recorded during the test include behavioral observations, questionnaires, logged data and correct identification of the buildings.

In total participants filled in two questionnaires, one before the experimental session and one after it. The pre-experiment questionnaire collected background information: provenance, age, education, level of knowledge of the city, and how participants were used to sightseeing a city as tourists. The questionnaire completed after the trial explored their impressions of the experience. The survey assessed participants' general feeling of naturalness and pleasantness in using the devices (3 items) and their impressions regarding the vibrotactile feedback (5 items). Finally 3 items explored the fear of appearing strange in the eyes of other people while using the device. Participants using the tactile glove were asked to complete a longer version of the questionnaire, in which they were also asked to evaluate the audio signals they received (5 items) and the physical sensations they had while wearing the tactile glove (5 items). Participants were asked to mark their level of agreement with each statement on a 5-point Likert scale, where 1 indicated a total disagreement and 5 a total agreement with the statement.

The final phase consisted of a brief semi-structured interview, in which participants were asked (a) to freely comment on their use of the devices, (b) what they found easy or difficult, (c) what they liked or disliked, what they thought of the (d) audio and (e) vibrotactile signals and (f) if they felt confident of the indications they received.

Both systems logged timestamped data that contained users' behavior including time spent in each stage, decisions of accepting or rejecting the notification, and distance to the POI, as well as environmental data like GPS accuracy.

5. Results

5.1 Performance

Since not all the POIs were correctly delivered due to poor GPS coverage, the rate of POIs correctly identified by participants was computed as a percentage index. The percentage of POIs correctly identified with the tactile glove and the baseline application was then compared by means of a t-test for two independent samples[9]. The analysis revealed no effect of the application in use on the correct identification rate of the POIs, t_{18}= .394 p= .69; the average identification rate per participant was 79.23% (*SD*= 14.97) with the tactile glove and 76.92% (*SD*= 10.87) with the baseline application.

Considering the rate of correct identification per POI, the average values are very high for both systems, with 90.85% (*SD*= 17.34) of correct identification for the tactile glove and 97.17% (*SD*= 5.44) for the baseline application. It is worth noticing that, for the tactile glove, only two POIs had an identification rate lower than 100%: in both cases users received the notification of a POI in the proximity when the landmark was out of their sight, confounding them.

[9] According to a Kolmorov-Smirnov test the data followed a normal distribution: Success rate for the glove *D*= .24, *p*=.106; baseline application *D*= 1.26 *p*=.053.

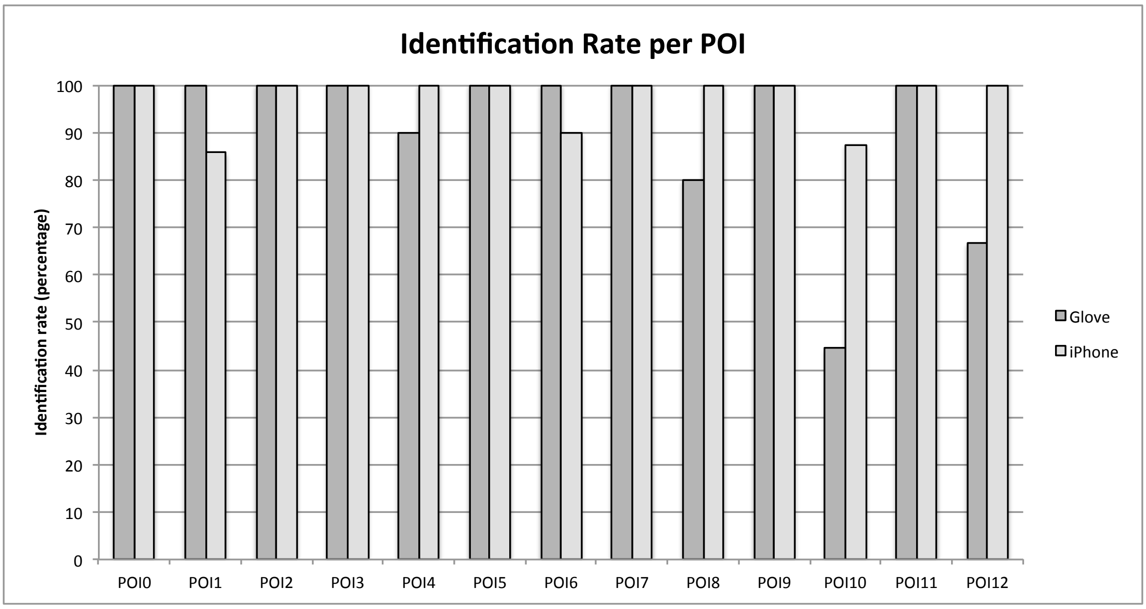

Figure 36 Identification rate per each POI

The time to select or discard the POI was compared between the two devices. The distribution of the data violated the normality assumption[10], according to a Kolmogov-Smirnov test (*D*= .258 *p*= .001). Therefore, non-parametric statistical test was applied for the comparison. The analysis revealed no statically significant difference in the time needed to select the POI with the baseline application (*M*= 3.18 seconds *SD*= 1.47 *Mdn*= 3) and the glove (*M*= 6.21 seconds *SD*= 5.9 *Mdn*= 3.17), *U*= 31 *p*= .17. It seems therefore that the two applications were equally attention getting in an urban exploration.

The time required to identify the POI was also compared between the devices. Again the variables did not follow a normal distribution (Kolomogorov-Smirnov test *D*= .26 *p*= .001) and thus a Mann-Whitney test was applied. The analysis revealed a significant difference (*U*= 10 *p*= .002) in the time to identify the landmark using the glove (*M*= 5.66 *SD*= 2.71 *Mdn*= 4.89) and the baseline application (*M*= 14.30 *SD*= 10.55 *Mdn*= 9). A tactile guidance delivered on fingertips seemed to guide users to identify the target landmark in shorter time as compared to conventional visual indications.

[10] A detailed inspection of the data revealed the presence of one outlier in the data. The exclusion of the data from that participant did not correct the shape of the distribution, therefore analysis were conducted on the whole sample.

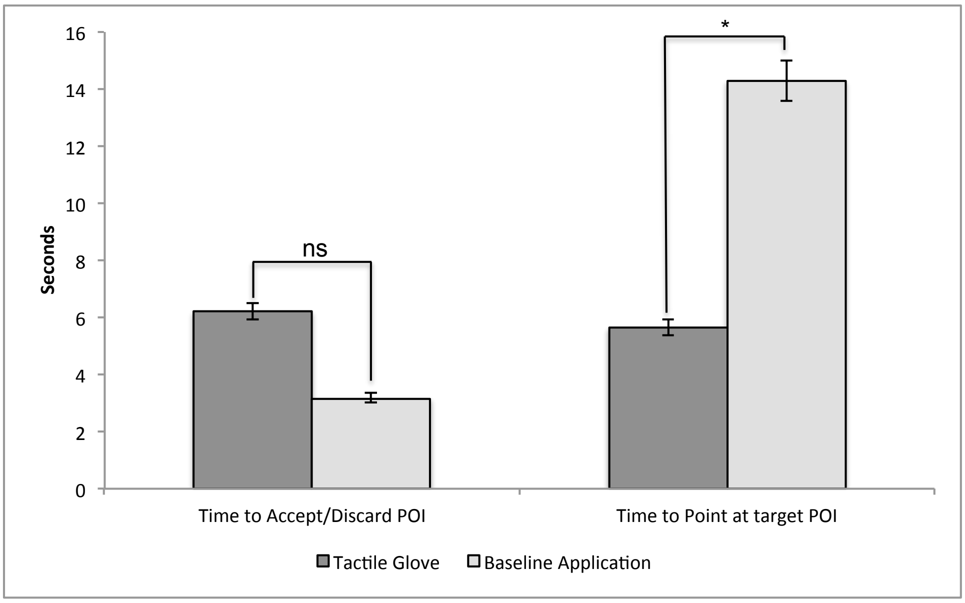

Figure 37 Average time for accepting or discarding the notification of the POIs

5.2 Questionnaires

A consolidated score was computed for each dimension investigated in the post-test questionnaire by averaging the mean score of each item belonging to that specific dimension. The score of the items with a negative phrasing was reversed.[11] To assess if the device in use had an effect on the average evaluation participant gave, the consolidate scores were compared by means of a Mann-Whitney test. The analysis revealed no statistically significant difference between the scores regarding the quality of the experience, *U*= 43.5 *p*= .63, with an average score of 4.22 (*SD*= .44, *Mdn*= 4.12) for the tactile glove and an average score of 4.02 (*SD*= .71, *Mdn*= 4.37) for the baseline application. Similarly the vibrotactile signals were equally clear for both the tactile glove (*M*= 3.9, *SD*= .67, *Mdn*= 4) and the baseline application (*M*= 4.13, *SD*= .32, *Mdn*= 4), *U*= 47 *p*= .85. Regarding the pleasantness of the vibrotactile signals, again no statistically significant difference emerged, *U*= 43 *p*= .63, with an average evaluation of 4.05 (*SD*= .55, *Mdn*= 4) for the tactile glove and 4.14 (*SD*= .33, *Mdn*= 4) for the baseline application. Interestingly, participants were not more concerned of appearing strange to bystanders when using the tactile glove (*M*= 2.9, *SD*= .83, *Mdn*= 2.83) as compared to the baseline application (*M*= 2.4, *SD*= 1.05, *Mdn*= 2.83), *U*= 32 *p*= .19.

[11] The score of items 2.1.3, 2.1.4 and 2.2.4 were reversed because of negative phrasing.

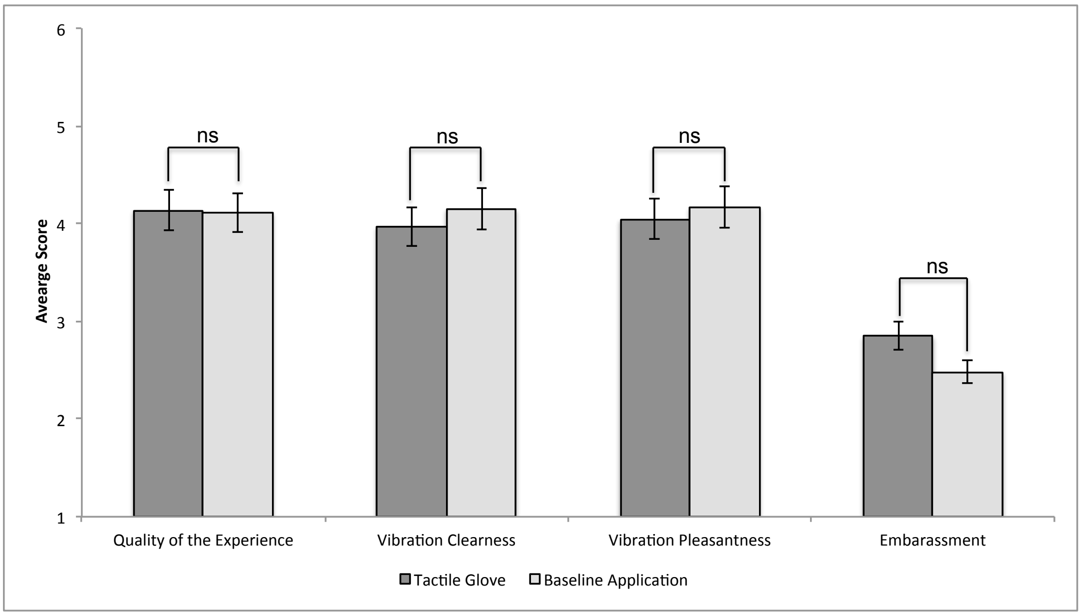

Figure 38 Average scores for each of the dimensions considered in the post-experience questionnaire

5.3 Interviews

The recordings of the interviews were transcribed and participants' answers were categorized to determine the more frequent comments.

Regarding the vibrotactile guidance delivered by the tactile glove, participants generally agreed in saying that they felt confident about the indications received (5 respondents) and that it was simple to locate the target landmark (3 respondents). Furthermore, commenting on the vibrations, they reckoned them precise (5 respondents) and clearly perceivable (4 respondents). None of the participants highlighted any particular difficulty in using the system. In particular, it was praised as easy to use by all the respondents. However two of them highlighted that it was a little complicated to figure out what was the landmark, when it was not in their view (this happened especially for one POI):

"Some places were quite hidden and I couldn't understand where they were"

(F, 24)

Interestingly, two participants reported that the system had made them feel a little disconnected from the environment. Conversely, three respondents said they did not had the feeling to be disconnected from the surroundings and other three participants commented that the system had made them feel even more engaged with the environment.

"Disconnected? No, it was almost like to be in a museum with a guide"

(M, 18)

Finally, two respondents reported that at some point they had the feeling of being stared at by bystanders.

"From time to time I had the feeling that someone was looking at me… but I just kept on walking."

(M, 19)

A general positive attitude emerged also regarding the baseline application. More than half of the respondents (6) affirmed that they felt confident about the guidance received to locate the landmarks. Additionally, all participants praised the vibrational indications as interesting (6 respondents) or attention getting (3 respondents). The system was considered easy to use by the majority of the subsample using the baseline application; five of them reported no difficulties at all, whilst three of them complained that it was troublesome to locate two POIs when they were very close to each other.

"[It was difficult to] find some places because they were too close to each other"

(F, 25)

Regarding the feeling of being detached from the surrounding participants shifted. Half of the sample reported that they did not disengaged from the environment. Conversely, 3 respondents admitted that they focused mainly on the display than on the environment; two users said they were mainly absorbed by the system at the beginning of the trail, but that they got used of it, being able after a while to attend to the surroundings.

6. Discussion

In the present experiment two different modalities for the serentipitous exploration of the point of interests in a city were compared. We have shown that the identification rate was equally high both the guidance of an audio-visual and a haptic-audio interface.

A vibrotactile interface is not only a valuable means for guiding the navigation in an environment (e.g., Tsukada and Yusumura 2004; Heute, Henze, Boll and Pielot 2008), but also to make users identifying landmarks by pointing at them, supporting a more natural interaction modality (Hsieh et al. submitted).

Furthermore the system was found easy to operate and the reported experience of use to be positive and engaging. Data from the interviews highlight a limitation of the current setup of the system. When the target landmark was out of participants' view, they struggled for finding it. This suggests the need of integrating a feature that can support the user to conjure up what she is supposed to look for. To say in other words, participants would have benefit from a higher multimodality of the system.

Consistently with what Szymczak et al. (2012) reported, we found that integrating the touristic exploration with tactile guidance and auditory narration was effective in allowing the user to discover relevant landmarks in the surroundings, while at the same time enhancing the experience of the use. This experiment has expanded the work in this field, as, at least to the writer's best knowledge, no previous study had investigated the use of vibrotactile stimuli delivered by a wearable device to guide the identification of objects in the real-world.

Chapter 4

Following the vibes: a comparison between two tactile displays in a navigational task in the field

1. Motivation and aim

Current navigation systems deliver directional indications mainly visually, i.e., via maps, or acoustically, i.e., via audio instructions, putting a considerable burden on the user's visual and acoustic attention (Asif, Heuten and Boll 2010). Using a navigational system while walking poses additional difficulties related to the constant need to stop and switch the gaze between the environment and the device for either checking the virtual information against the real world and for avoiding obstacles (Pielot, Poppinga, Heuten and Boll 2011; Jacob, Mooney and Winstanley 2011; Heuten, Henze, Boll and Pielot 2008). Furthermore, because of the constant tilting caused by the user's steps, it is challenging to read what is written on the display (Rümelin, Rukzio and Hardy 2011). Also factors related to the context can make the interaction with mobile navigation apps difficult, e.g., due to screen reflection in bright sunlight or just to the need to hold an umbrella (Jacob, Mooney and Winstanley 2011). Even if auditory indications relieve the user from the need of checking the display, they seem in general obtrusive and not efficient. Loudspeakers are in fact unreliable outdoors and can attract undesired attention from bystanders, while headphones tend to isolate the user from the surroundings (Heuten, Henze, Boll and Pielot 2008).

Tactile displays seem to present several advantages over the traditional audio-visual systems. First, they can be useful when the auditory and visual attention is limited or affected by ambient interferences, moreover they have the potential to deliver precise spatial information to the user in an unobtrusive manner (Srikulwong and O'Neill 2011; Weber, Schätzle, Hulin, Preusche and Deml 2011). Secondly, since the tactile channel can be stimulated over 360°, a tactile display may also compensate for the lack of peripheral view (Van Erp 2001). When using a tactile interface the user can focus on the environment, being able to enjoy fully the experience (Pielot, Poppinga, Heuten and Boll 2011). A more general advantage of the tactile directions is that they do not require the user to match the real world with the indications on the map, which can be challenging for some users (Tsukada and Yasumura 2004). Finally, since the tactile indications are delivered directly on the user's body, they overcome cultural barriers and are adequate to be used universally with different targets of users (Jacob, Mooney and Winstanley 2011).

The research question guiding the present study was twofold. First, we aimed to address if vibrotactiles directional cues were effective stimuli to guide a user in a navigational task in an urban setting. Secondly, we were interested in whether any differences could be observed in terms of performance and experience of use when the directional indications are delivered on the fingertips or on the torso.

2. Related work

Vibrotactile guidance has been experimented mainly in three application domains: to direct the user's attention, to guide her movements and in navigational tasks (Weber, Schätzle, Hulin, Preusche and Deml 2011). As *attentional guidance,* the tactile stimuli can direct the user's visual attention toward the location where critical events are about to occur (e.g., Ahmaniemi and Lantz 2009). Also, tactile stimuli have been employed in the educational context for *guiding the movements* of the learner in order to improve the training (e.g., Holland 2010). Finally, vibrations have been used to provide *spatial guidance* and thus to give navigation support, e.g., Tsukada and Yasumura 2004. In this respect several form factors (including both mobile applications and wearables) and different application domains have been experimented. Tactile belts have been widely used to convey information about distance to car drivers (Asif, Heuten and Bolt 2010), to guide the acquisition of waypoints in the military (Van Erp, Van Heen and Jansen 2005) and civilian domains (Srikulwong and O'Neill 2011) and to direct pedestrians (Tsukada and Yusumura 2004; Heuten, Henze, Boll and Pielot 2008). A tactile vest has been tested for delivering turning indications to motorcyclists (Prasad, Taele, Goldberg and Hammond 2014).

Many navigation applications have been developed on hand-held devices to convey information on direction and distance. Pielot, Popping and Boll (2011) proposed

PocketNavigator, a tactile compass running on a smartphone that provides the user with directional cues all around the user. NaviRadar provides the user with tactile turn-by-turn instructions (Rumelin, Ruzkio and Hardy 2011). TactiGuide is a smartphone-based application that vibrates when the user points towards the target destination (Komninos, Astrantzi and Stefanis and Garofalakis 2015). Coupling a smartphone to a smartwatch, Lim, Cho, Rhee and Suh (2015) transmitted tactile guidance indoors.
Researchers have also attempted to direct users by developing small wearables, namely an armband (Weber, Schätzle, Hulin, Preusche and Deml 2011), a wristband (Bosman et al., 2003) and a glove (Bial, Kern, Alt and Schmidt 2011).

3. Equipment

The two tactile display compared in the present experiment are described in the present section. First the tactile glove is presented, then a description of the tactile vest is provided.

3.1 Tactile glove

The system is based on a tactile glove, which is equipped with vibrotactile actuators and a set of sensors, and an Android mobile device. The glove has been constructed of thin elastic fabric and several electronic components. Two Arduino microcontrollers (Arduino Pro Mini) process the sensor signals and commands, and control the vibrotactile feedback. For sensing hand orientation, the glove is equipped with a 9-axis Inertial Measurement Unit (IMU, InvenSense MPU-9150), which consists of a gyroscope, an accelerometer, and a compass. Flexible bend sensors (Spectra Symbol flex sensor) are deployed on three fingers (thumb, index, and middle finger). A set of three vibrotactile actuators (Precision Microdrives 10mm shaftless vibration motor) is mounted for providing notification cues and directional guidance on the thumb, index, and middle fingers. A vibration on the middle finger indicates the user to turn on the right, conversely a vibration on the thumb means she has to turn left. Finally, when the index finger vibrates she has to walk straight on. The glove communicates wirelessly over Bluetooth with an Android mobile phone (Samsung S3) running the application software. The glove has been built for right-handed interactions (Jythä et al., 2015). The placement and the number of the actuators were informed based on previous pilot experiments[12].

For the present experiment a dedicated application running on an Android mobile phone (Samsung S3) was developed to trigger the vibrotactile stimuli.

3.2 Tactile vest

The tactile vest is made of two layers, the inner layer consists of the actual vest on which the actuators are placed and the outer layer is a stretchable vest. The inner garment, is a “one-size-fits all” that holds 29 actuators, moveable in order to ensure they are located on the correct location points for each different shaped body. The outer snuggly-fitting vest is a stretchable vest designed to keep the eccentric rotating mass actuators tightly placed against the bowed-curved areas of the body, such as the lower back and chest, to ensure the vibrations are evenly felt in all areas of the body. To ensure that the actuators were as close as possible to the users’ skin, participants were asked to wear the vest under their own jacket during the experimental trails.

A dedicated application running on an Android tablet (Nvidia Shield Tablet K1) was developed to control the different vibrotactile patterns.

The system can produce 26 different vibrotactile patterns. For the present experiment 3 vibrotactile patterns were selected, in order to provide directional indications comparable with the tactile glove. In particular, to communicate to the participant to turn the actuator on the left or on the right shoulder vibrated. When the participant had to walk straight a set of vibrators on the abdomen vibrated.

---

[12] The performance in a target selection task was compared when the guiding vibrations were delivered on the palm or on the fingertips. We found that the placement of the actuators on fingertips outperformed the placement on the palm in terms of speed of selection and intuitiveness of the signals.

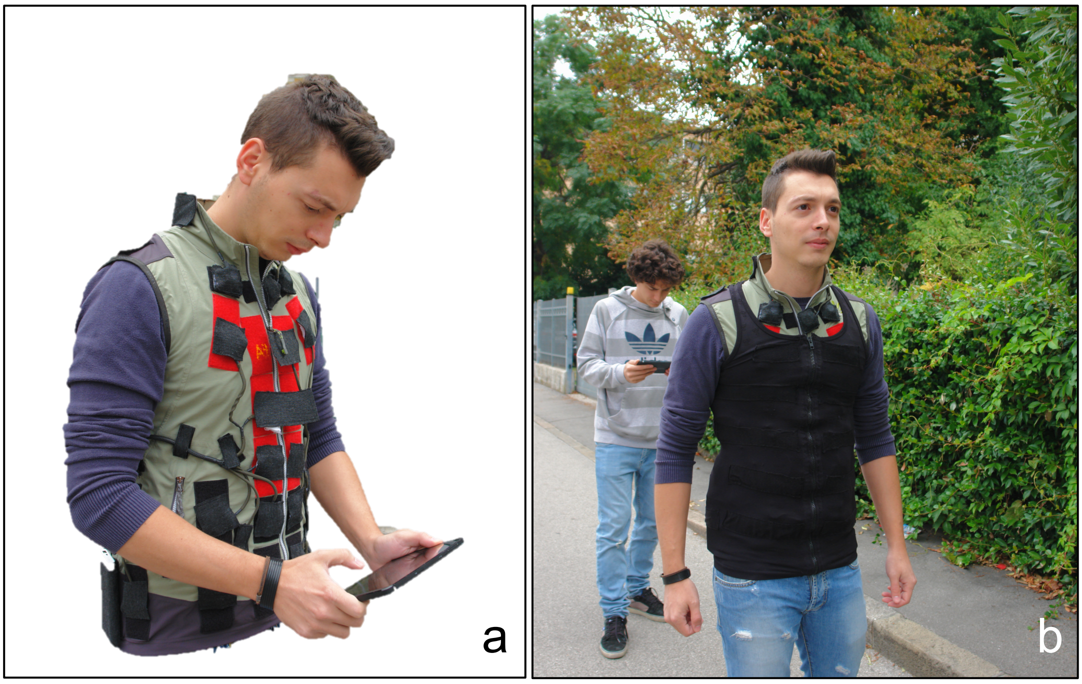

Figure 39. The inner layer of the vest with the actuators placed on the torso (a). An overview of the complete vest with the outer layer on (b)

4. Method

4.1 Experimental Design

The experiment followed a within subjects design, with the device as independent variable.

The order with which participants were required to complete the two itineraries and the order of use of the two devices was counter balanced across participants.

4.2 Setting

The experiment took place in the center of Padua. Two different, yet analogous, itineraries were identified. Both Itinerary 1 and 2 had the same starting and ending point and included 4 turns to the right and 4 to the left. Itinerary 1 required participants to pass through Piazza delle Erbe, whilst Itinerary 2 required them to pass through Piazza della Frutta, as depicted in Figure 40.

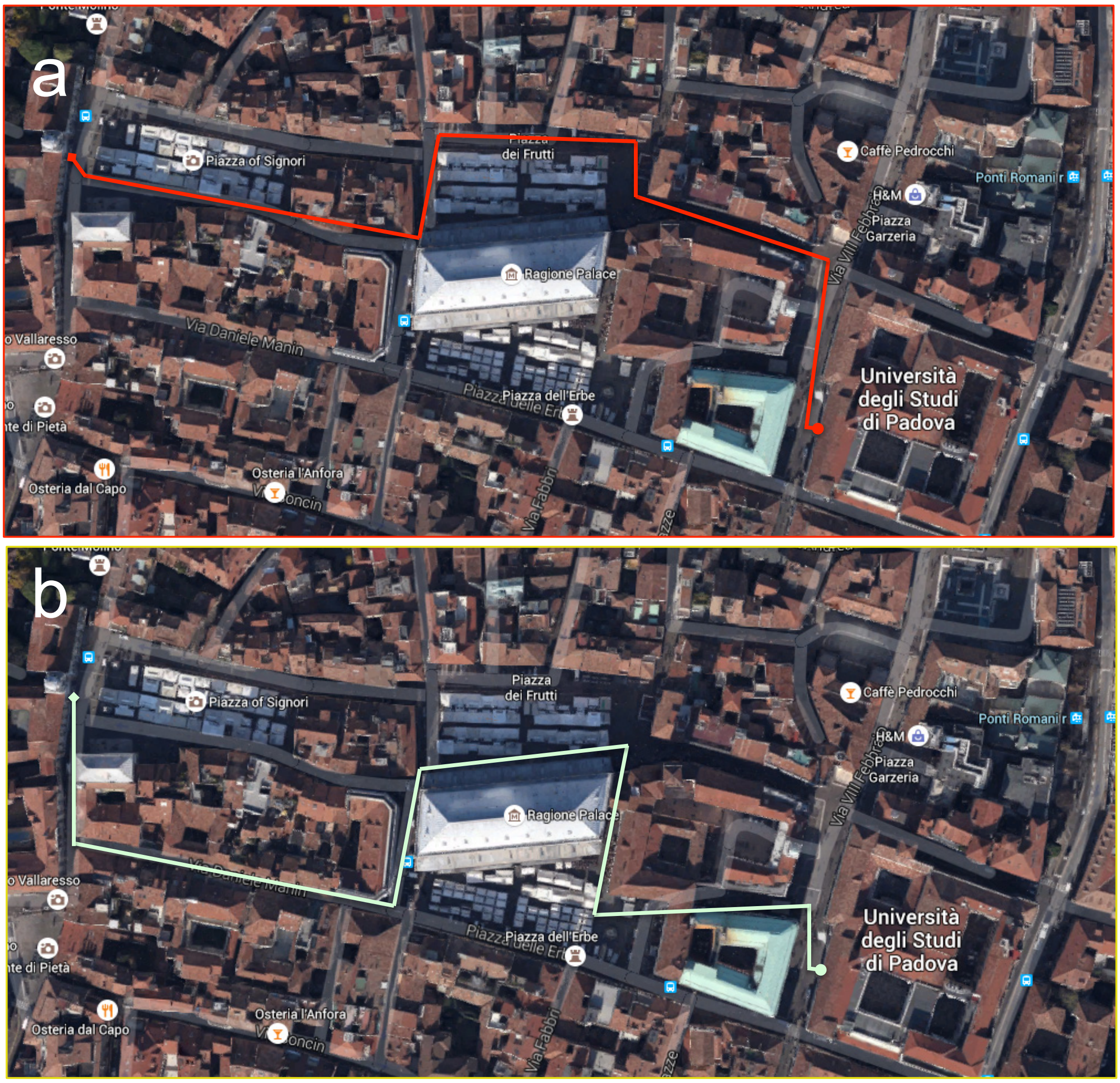

Figure 40 The area in the center of Padua where the experiment took place. The two alternative itineraries are highlighted with a red line (a) and a green line (b)

4.3 Sample

In total, 24 participants volunteered in the present experiment. Nine of them were men and fifteen were women. The mean age of the sample was 24.37 (*SD*= 2.24). Twenty-two participants were Italian, one was Spanish and one was Macedonian. All the participants were right-handed but two. All of them had a fairly good knowledge of the city. Participants were used to navigate in foreign cities by means of applications running on smartphones (22) or paper maps (2).

Participants were recruited by posting advertisement on dedicated pages on social networks, by word of mouth and in the field.

4.4. Experimental Procedure

Participants were met in the city center and were first debriefed on the overall experimental procedure and scopes. After they agreed to partake in the experiment they received, read and signed the informed consent.

Participants wore the device (either the glove or the vest, according to the experimental condition) and tried the vibrotactile patterns they would have received during the test. This phase lasted about three minutes. After that, they filled in the pre-experience questionnaire. At this point participants were told the meaning of the different vibrotactile patterns and were trained to follow them along a short predefined itinerary in front of Palazzo del Bo, for an approximate duration of three minutes. In this phase participants received the three signals available, so to contextualize the meaning of the vibrations. If participants confirmed that all the signals were clear, the actual experimental session began. When participants reached the end of the itinerary, they were asked to complete the post-experience questionnaire. After that, the experimenter and the participant walked back to starting point and the participant wore the other device, completed the pre-experience questionnaire, received a training and finally started the alternative itinerary. Again when she reached the end of the itinerary she filled in a post-experience questionnaire. Next, participant was asked to answer to the semi-structured interview.

In order to make the experimental procedure consistent across all participants, the experimenter provided the vibrations at fixed spots. Furthermore, in case participants did not perceive a vibrational stimulus, namely an "omission", the experimenter provided immediately the same stimulus again. In case of error, intended as a misunderstanding of the meaning of the vibrations, the experimenter led the participant to the destination using a set of vibrational stimuli alternative to the predefined set. In such cases, the alternative route and indications received were tracked by the experimenter.

4.5 Materials

In order to assess the user experience with the devices two questionnaires were created *ad hoc* and administered. The first questionnaire investigated the impressions participants had of the devices before the actual experience and consisted of eight items.

Three items explored the perceived usefulness of the wearable, that is to say the extent to which the user believes the product can enable her to achieve her goals (Davis 1989). The perceived ease of use, defined as the extent to which a user thinks the device will not be tiring or difficult to use (Davis 1989), was investigated by two items. Two items investigated the embarrassment of using the wearables in public, with specific reference to the concern of appearing weird in the eyes of bystanders (Venkatesh and Morris 2000; Buenaflor and Kim 2013). Finally, one item aimed at exploring the comfort of wearing the device, namely wearability (Gemperle et al., 1998).
In the post-experience questionnaire, consisting of 18 items, the same dimensions plus the quality of the experience and the satisfaction of use were assessed. In particular, three items assessed the usefulness and three the ease of use of the device. The embarrassment experienced while using the device in public was assessed by 2 items and the comfort of wearing the device by 4 items. The quality of the experience, that is the set of affective states deriving from the use of a device (Hassenzahl 2004), was investigated by 4 items. Finally, the satisfaction of use, meaning the extent to which the user considers the interaction with the device pleasing (Davis 1989), was investigated through 2 items. Respondents were asked to mark their level of agreement on a 5-pont Likert scale, where 1 indicated total disagreement and 5 total agreement.
Participants answered to a semi-structured interview (3 questions), inquiring (a) their general impression regarding the interaction and the devices themselves; (b) if they found the devices useful for navigating and (c) if they had issues during the task. They were also asked if they found the use of wearable computers in a navigation task useful and if they had any issues during the interaction.

5. Results

5.1 Performance

In the 100% of the cases participant were able to complete the two itineraries successfully using both devices. Furthermore, along the paths all participants made no errors.

Omissions occurred twice, one with the glove and one with the vest. In both cases, after the experimenter had played again the stimulus participants followed it correctly.

To test if the two itineraries selected were effectively analogous the travel times were compared. The Kolmorov-Smirnov test revealed that the normality assumption was violated for both the travel times, respectively $z= .249$ $p= .001$ for Itinerary 1 and $z= .285$ $p< .001$ for Itinerary 2. The Wilcoxon signed-rank test was then employed for the comparison. The analysis confirmed that the two itineraries were similar for travel time, $z= 1.444$ $p= .149$. Itinerary 1 was completed on average in 7.12 minutes (*SD*= 1.15) and the average time to complete Itinerary 2 was 8.33 minutes (*SD*= 3.8).

The effect of the device used was then tested comparing the travel time in the two conditions. Again the Kolmogorov-Smirnov test revealed that the normality assumption was violated for both the variables considered, respectively $z= .266$ $p< .001$ for completing the path with the glove and $z= .341$ $p< .001$ for completing the path with the vest. Again the Wilcoxon signed-rank test highlighted no statistically significant difference between the variables, $z= 1.544$ $p= .123$ (Figure 41). On average it took 7.75 minutes (*SD*= 1.26) to complete the task supported by the glove and 7.76 minutes (*SD*= 3.9) with the guidance of the vest.

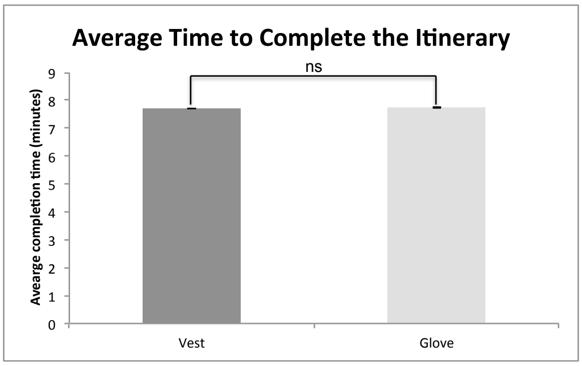

Figure 41 Average time to complete the itinerary with the two devices

Finally, the effect of gender was also tested, as Ahmad, Goldiez and Hancock (2005) found that gender affected the completion times in a navigational task guided by tactile stimuli. The Mann-Whitney test revealed no statistically significant difference between men and women neither for the glove ($z= .158$ $p= .874$) nor for the vest ($z= .522$ $p= .602$). In particular, regarding the men, the average completion time was 8 minutes

(*SD*= 1.58) with the glove and 8 minutes (*SD*= 1.56) with the vest. For women, the average completion time was 7.6 minutes (*SD*= 1.00) with the glove and 8.2 minutes (*SD*= 1.56) with the vest (Figure 42).

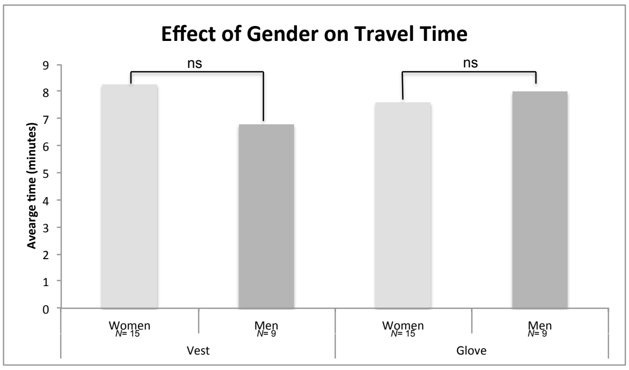

Figure 42 Average time to complete the two devices for men and women

5.2 Questionnaire

A consolidated score was computed for all the dimensions investigated in the pre- and post experience questionnaires. Such consolidated score was computed by averaging the scores of all the items belonging to a specific dimension. The average scores of the dimensions tested in the pre-experience questionnaire were compared with the average scores of the post-experience questionnaire by means of the Wilcoxon test[13]. The analysis was conducted with the aim to assess if, for the dimensions investigated, the expectations participants had were satisfied during the experience. Furthermore, the average scores of the dimensions of the post-experience questionnaire were compared between the two devices by means of a Wilcoxon test, to assess if the experience changed when participants used the glove or the vest. Finally, to assess if participants' opinions were positive or negative, a one-sample t-test was run against the mid-point of the response scale (i.e., 3), which indicates a neutral attitude for all the dimensions considered in the post-experience questionnaire.

Regarding the glove, the analysis revealed that the usefulness of the device significantly increased after participants had tried the device, $z= 2.143$ $p= .032$, the average score before the test was 3.73 (*SD*= 0.47) and after the test it raised to 4 (*SD*= 0.48). Interestingly, the concern of appearing strange to bystanders because of the glove

[13] The scores of the items phrased negatively were reversed.

reduce significantly after using the glove, *z*= 2.233 *p*= .008. The average score recorded before the experience was 2.95 (*SD*= 0.35) and the one after the experience was 3.62 (*SD*= 1.05). The average score regarding the comfort of wearing the glove after the test was in line with the expectations of comfort, *z*= .216 *p*= .82, the average score before the test was 3.76 (*SD*= .84) and the average score after the test was 3.7 (*SD*= 1.16). Also for the ease of use, participants' evaluations did not change after the test, *z*= .000 *p*= 1.00, with an average score before the test of 3.20 (*SD*= 0.42) and 3.20 (*SD*= 0.41) after it.

For what concerns the vest, the usefulness significantly increased after the test *z*= 3.233 *p*= .001, with an average score before the test of 3.62 (*SD*= 0.65) and 4.13 (*SD*= 0.52) after it. On the contrary, the concern of appearing weird to bystanders significantly reduced after the experience, *z*= 2.52 *p*= .012, the average scores were respectively 2.91 (*SD*= 0.28) before the test and 3.62 (*SD*= 1.05) after it. The scores of the comfort of wearing the device did not significantly change after the use of the device, *z*= 2.010 *p*= .44, with an average score of 2.95 (*SD*= 1.08) before the experience and 3.25 (*SD*= 1.29) after it. Finally, the ease of use of the device after the test was in line with the expectations participants' had before it, *z*= 1.395 *p*= .163; respectively, the average score was 4 (*SD*= 2.93) before the test and 3.25 (*SD*= 0.44) after it.

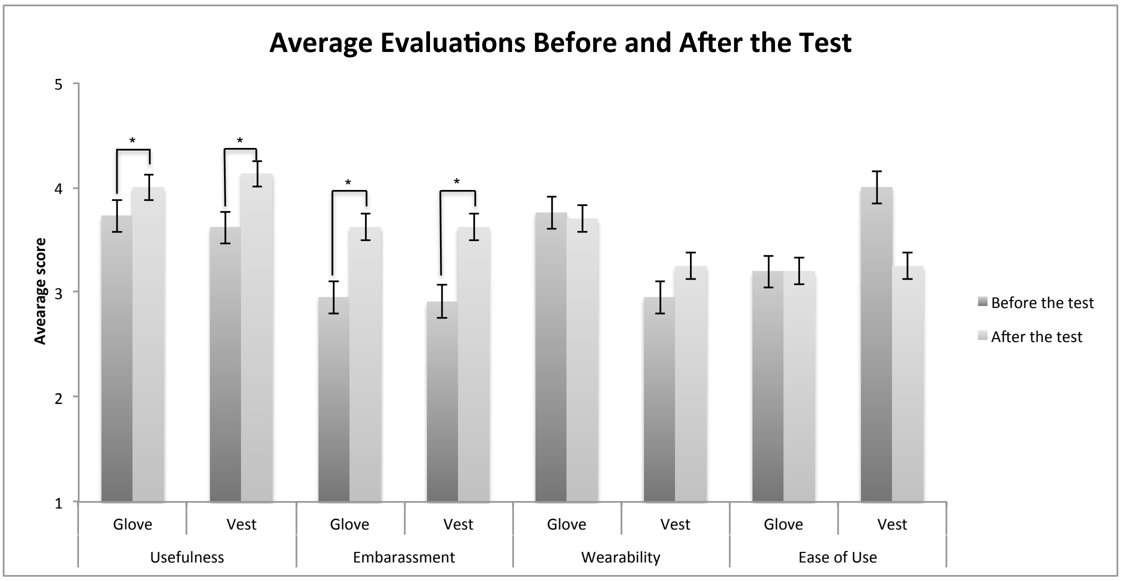

Figure 43 Comparison between the average scores before and after the experience at the different dimensions tackled in questionnaire

Regarding the comparison between the two devices of the average scores recorded after the experience, a significant difference emerged for the comfort of wearing the device, $z= 2.765$ $p= .006$, with the glove being more comfortable ($M= 3.96$; $SD= .53$) compared to the vest ($M= 3.55$; $SD= .75$). The two devices were considered equally useful, $z=1.448$ $p= .148$, with an average score for the glove of 4 ($SD= 0.48$) and of 4.13 ($SD= 0.52$) for the vest. Similarly, they were both perceived easy to use, $z= .573$ $p= .567$ (average score for the glove $M= 3.4$; $SD= .47$ and 3.5 $SD= 0.38$ for the vest). The concern of appearing weird to other people again did not differ between the two devices, $z=1.677$ $p= .094$ ($M= 3.97$ $SD= .85$ for the glove and $M= 3.62$ and $SD= 1.05$ for the vest). The quality of the experience was positive for both the glove ($M= 4.06$ $SD= .40$) and the vest ($M= 3.55$ $SD= .75$), $z= .000$ $p= 1.00$. Finally, the scores regarding the satisfaction of use did not differ ($z=1.629$ $p= .103$) between the glove ($M= 4.06$ $SD= 0.57$) and the vest ($M= 3.83$ $SD= 0.68$).

For what concerns participants' opinions, in general they had a positive attitude with respect to all the dimensions investigated. Regarding usefulness, both the glove and the vest received positive scores ($t_{23}= 10.17$ $p< 0.001$, $M= 4$ $SD= 0.48$ for the glove; $t_{23.}= 10.55$ $p<0,001$, $M= 4.13$ $SD= 0.52$ for the vest). Similarly they were both easy to use ($t_{23}= 4.55$ $p< 0,001$, $M= 3.44$ SD= 0.47 for the glove; $t_{23}= 7.21$ $p< 0,001$, $M= 3.56$ $SD= 0.38$ for the vest). The possibility of appearing strange to bystanders seems not to concern them for the glove ($t_{23}= 5.62$ $p< 0.001$, $M= 3.97$ $SD= 0.85$) neither for the vest ($t_{23}= 2.91$ $p< 0.001$, $M= 3.62$ $SD= 1.05$). Both devices were comfortable to use ($t_{23}= 8.81$ $p< 0.001$, $M= 3.96$ $SD= 0.53$ for the glove; $t_{23} = 3.58$ $p< 0.001$, $M= 3.55$ $SD= 0.75$ for the vest). Participants seemed satisfied with the use of both the glove ($t_{23}= 5.94$ $p<0.001$, $M= 4.06$ $SD= 0.57$) and the vest ($t_{23}= 5.94$ $p< 0.001$, $M= 3.83$ $SD= 0.68$). Finally, the quality of the experience seemed good for both the glove ($t_{23}= 10.73$ $p< 0.001$, $M= 4.06$ $SD= 0.48$) and the vest ($t_{23}= 10.73$ $p< 0.001$, $M= 4.06$ $SD= 0.48$).

Dimensions	Glove		Vest	
	t	*M(SD)*	*t*	*M(SD)*
Usefulness	10.17**	4(.48)	10.55**	4.13(.52)

Ease of Use	4.55**	3.44(.47)	7.21**	3.56(.38)
Embarrassment	5.62**	3.97(.85)	2.91**	3.62(1.05)
Wearability	8.81**	3.96(.53)	3.58**	3.55(.75)
Quality of the Experience	5.94**	4.06(.57)	5.94**	3.83(.68)
Satisfaction	10.73**	4.06(.48)	10.73**	4.06(.48)

Table 1 Values of the one-sample t test with means and standard deviations for the different dimensions tackled

5.3 Interviews

The recordings of the interviews were transcribed and participants' answers were categorized to determine the more frequent comments.

According to the interviews, participants seemed generally satisfied with the overall experience. Regarding the vest, participants affirmed that the vibrotactile patterns were clearer (9 participants) more intuitive (9 participants) and easier to perceive (9 participants) compared to the vibrations emitted by the glove. One of the responded commented:

"I was more doubtful regarding the vest, but in the end it was easy to reach the target. It was light weighted and the signals were very clear"

(F, 23)

However, five respondents disagreed, reporting greater difficulties in perceiving the signals. Regarding the comfort, 11 participants affirmed that the vest was bulky and uncomfortable to wear, three said it limited their movements. Three respondents were concerned about the difficulty of carrying it if not worn. However, according to two participants, the vest allowed them a greater freedom to move. Finally, two participants reported that they had a stronger feeling of being observed while they were using the vest compared to the glove.

Regarding the glove, 12 respondents praised it as more light weighed and more comfortable to wear. It was also considered easier to carry when not in use (12 participants). Additionally, 12 participants appreciate the glove because less noticeable compared to the vest. However some participants (8) reported difficulties in distinguishing the different vibrations between the index and middle finger, and some of them (3) even suggested to move the actuator currently placed on the middle finger to the little finger:

"The vibrations between the index and middle finger are not easy to distinguish. Maybe it would be easier to insert vibrations on the thumb, index and little finger"

(M, 26)

In addition, 5 respondents seemed concerned by the fact that they were not able to use the hand wearing the glove to do anything else.

In general, all participants define both devices useful, because they are effective in supporting the navigation and in leading the user to her target. One of them commented:

"The devices are very useful, they provided me clear indications regarding the streets to and they led me to the destination"

(M, 30)

6. Discussion

In the present experiment we have found that tactile displays are a valuable means for navigating users in an urban environment. In particular, we showed that vibrotactile guidance is equally meaningful in leading users to a target destination either when the directional indications are delivered through a vest or though a glove, thus expanding previous work in this field. Pedestrians benefit of vibrotactile guidance sent on the shoulders, consistently with what Prasad, Taele, Goldber and Hammond (2014) demonstrated for motorcyclists. Additionally, we have successfully navigated users by means of vibrotactile cues on the fingertips. The two wearables led users to the destination within a similar amount of time.

Subjective data show that participants changed their mind on the devices after the use. In particular, both devices were considered more useful after the test than before; moreover users were less concerned of appearing strange to bystanders after the test than what they were expecting beforehand.

The appearance of the form factor confirmed to be very relevant to users (Heuten, Henze, Boll and Pielot 2008). The majority of the participants in fact found the indications provided by the vest more intuitive and clear, however they preferred the glove for being unbulky and light-weighted. This finding is consistent with what Pielot, Poppinga, Heuten and Boll (2011) envisioned for a tactile belt, that a user would probably be unlikely to interact with such a device for a long time.

The results of the present experiment suggest a series of aspects to consider in designing a vibrotactile interface. First, actuators should be reduced to a minimum, consistently with Srikulung (2011). Secondly, actuators should be placed at a reasonable distance to one another, so that the user can easily discriminate among them. Finally, the overall size and shape of the interface should not be obtrusive and showy.

Augmenting the Reality: a comparison between a smartphone and a pair of smartglasses

1. Motivation and aim

Augmented reality (AR) refers to the supplement of the real-world surroundings with computer generated digital elements that are interactive in real time and recorded in 3D (Azuma, 1997; Olsson et al., 2012). AR applications have recently gained popularity thanks to the technological advancements of mobile devices, e.g., smartphones (Olsson et al., 2012; Haugstvedt and Krogstie 2012). In addition, computational devices that the user can wear on the head are becoming widespread for daily use (Serrano, Ens and Irani 2014) and ultimately they have been experimented with mobile AR applications. Head-worn viewers, also referred to as smartglasses, have the big potential to provide the user a large amount of always-available information in a subtle manner, thus preventing distractions (McAtamney and Parker 2006, Costanza et al., 2006), which is extremely desirable in a mobile context. AR applications can be useful in many different application domains (e.g., education, shopping), however the touristic contexts has gained a lot of attention from developers, since there are a variety of information that can be augmented, offering a rich field for experimentations (Ganapathy 2013).
Given their novelty, there is a scarce number of studies investigating the usability and the user experience of smartglasses (e.g, Brusie 2015), and none of them addresses the touristic application context, which is one of the main application domain of AR applications. In addition, despite the growing interest of the scholars, there is still paucity of works specifically addressing the usability and the user experience of AR apps running with hand-held devices (Ko et al. 2013).

The present work has the aim to test usability aspects and the experience of use with a touristic AR app running on two different devices available to the general public: a smartphone and a head-worn AR display.

2. Related work

A wide body of research has involved the development of mobile AR, however the focus is typically on technical challenges and constraints (Düsner et al. 2007; Kourouthanassis et al. 2013). The lack of research on user interfaces and interaction in mobile AR applications from a user centric perspective was first highlighted by Swann and Gabbard (2005). Recently the interest toward the usability and the user experience of such systems is starting to grow thanks to the diffusion of end-user services, however specific and solid evaluation methods have yet to be established (da Sà and Churchill 2013). In this framework, several works have drawn guidelines and heuristics for evaluating the usability of AR applications. Düsner et al. (2007) applied the main principles of HCI to the field AR and derived eight design principles regarding (a) the affordances offered by the application, (c) the cognitive effort required to the user, (d) the comfort of interacting with the application, (e) the intuitiveness of the interaction, (f) the user experience, (g) the flexibility of the application, (h) the responsiveness and the (i) tolerance to errors of the system. Based on the existing research Ko et al. (2013) developed 22 guidelines related to (a) information organization (user-information), (b) cognitive effort required to the user (user-cognitive), (c) supportiveness of the application (user-support), (d) interaction modalities (user-interaction) and (e) basic operations with the system (user-usage). Finally, Kourouthanassis et al. 2013 proposed five principles for the development of mobile AR applications. In sum, they recommend to provide context-related contents, to favor the development of practice of use, in order to reduce the memory load and to provide clear feedback to user's behavior. In addition, they suggest being transparent regarding users' data processing and storage. Such principles and guidelines are useful for developers in the early stage of the project, however they are too broad to be used as for evaluating the usability of the applications developed (Düsner et al. 2007; Kourouthanassis et al. 2013). When investigating the usability and the user experience of AR apps in the outdoor touristic domain, researches have mainly adopted a mix-method approach employing both questionnaires and interviews (Linaza et al., 2012; Suh et al., 2011). Seo et al. (2011) developed an AR tour guide running on a smartphone that overlays historic characters over the cultural heritage sites and evaluated in involving visitors and simply posing them five close questions. Similarly, Lee et al. (2012) asked participants to try CityViewAR, a mobile AR app superimposing original buildings over a city devastated after an earthquake, and then to fill in a short usability questionnaire. No analysis was conducted on participants' performance.

Probably due to the recent release on the market, so far no previous work has completed a usability evaluation of an AR application running on smartglasses with end users. Brusie at al (2015) attempted to compare the ease of use of Google Glass and Vuzix M100 in a surgical application domain. However, they were not able to involve users because they found the devices still immature for a proper usability test.

3. Equipment

In one condition users wore Epson Moverio BT-200 (developer version). Epson Moverio is a lightweight (240 g) see-through AR viewer, in which each lens is endowed with a display allowing 3D vision of the objects overlaid onto the surroundings. The user can input commands by touching a hand-held pad. The device works on Android™ 4.0 System. In the other condition participants used a 4.3" smartphone by Samsung (S4 Mini), which works on Android™ 4.2 System. In the present experiment a commercial AR app was employed as test-bed. The app chosen is Junaio AR Browser. When running, the app shows the points of interest (POIs) in the surroundings, as digital labels superimposed onto the real environment visualized via the camera of the device (Figure 44). Junaio can be downloaded for free from the Google Play Store[14] and from the Moverio Official Market[15].

The entire experimental session was video recorded by means of a Sony Handycam HDR-SR1E videocamera, to allow subsequent off-line videoanalysis.

[14] https://play.google.com/store/apps/details?id=com.metaio.junaio&hl=it

[15] http://www.epson.it/it/it/viewcon/corporatesite/products/mainunits/overview/12411/apps

Figure 44 A screenshot of Junaio

4. Method

4.1 Experimental design

The experiment followed a 2 (Devices) x 3 (Tasks) within subjects design. To prevent order effects, the order of presentation of the devices and the tasks were counterbalanced. Furthermore, to prevent practice effect, for each task two different, yet analogous, assignments were identified. One of the tasks selected required to identify and reach a target destination corresponding to a specific point of interest, which was out of participant's view from the starting location. The target POI was placed at an approximate distance of 200 m. In another task the user had to select the three nearest POIs belonging to a specific category and to report their names to the experimenter. Finally one task required users to access and read aloud the information regarding a specific POI.

4.2 Setting

The data collection took place in center of Padua, in particular in the area comprised within the Dome and Piazza dei Signori. When engaged with the tasks requiring to find and access additional information using the app, the user was standing on a pre-fixed spot in front of the Dome. While, when she was reaching the target POI, she walked from the pre-fixed spot thru via Monte di Pietà and then crossed Piazza dei Signori to reach the target destination.

Figure 45 The area where the data collection took place, highlighted in green

4.3 Sample

Twenty people took part in the study. Technical issues regarding the GPS signal occurred during 2 trails, therefore data from 2 participants were excluded. We had a final sample of 18 people (7 of them were women). The mean age of the sample was 23.3 years old (*SD*= 11.2). Participants were recruited by posting advertisements on dedicated pages on social networks, by word of mouth and in the field.

4.4 Experimental Procedure

Once agreed to participate in the experiment, participants were first debriefed on the experimental aims and procedure. Then they signed the informed consent. After that the test began with a brief training. The training was very similar for both devices and consisted in finding and reading the information associated to the label "Dome Of Padua". The purpose of this phase was to make the participant familiarize with the application design and functionalities. When the participant felt she was able to operate the app, she filled in the pre-experience questionnaire. Then participant was given the device, either smartphone or the headworn viewer, according to the experimental condition, and was asked to perform the three tasks. During the trail the experimenter annotated if the user accomplished the task. Next, the participant filled in the post-experience questionnaire. The second part followed the same procedure as the first one, the only difference was that the participant used the other device. Finally, the brief interview took place.

4.5 Materials

For investigating the experience of use of the two devices a questionnaire was build *ad hoc*. Such questionnaire is composed of two parts: one collecting respondents' expectations to be filled before the experience and one assessing their views after the use of the device. The part preceding the experience consists in total of 6 items, exploring three dimensions:

- **Expected ease use** (3 items) refers to the extent to which the user foresees that the use of the device will be simple and effortless (Davis, 1989);
- **Expected Usefulness** (1 item) refers to the extent to which the user thinks the device will be an effective support for achieving her goals (Davis, 1989);
- **Expected Self Impression** (2 items) refers to the possible concerns the user regarding they way she appears to other people's eyes while using the device (Buenaflor and Kim 2013);
- **Expected Comfort** (2 items) refers to the extent to which users expected to feel at ease with the device while using it. For the headworn viewer, this means to investigate if the device was obtrusive on the user's body. While for the handheld device, expected comfort related to the possible burden placed by holding the smartphone in the hand.

The questionnaire administered after the experience consisted of 14 items exploring 6 dimensions, four of which were the same as in the pre-experience questionnaire. The items belonging to the dimensions already assessed in the pre-experience questionnaire were slightly rephrased to prevent superficial answers and to make the statements consistent with the time of the experience. For instance, one item assessing the ease of use in the pre-experience questionnaire was "I expect the device to be easy to operate", whilst the item on the post-experience questionnaire was "It was ease to operate the device". One item was added in the Usefulness dimension.

The two additional dimensions assessed were:

- **Visual comfort** (2 items) refers to the extent to which the user perceives that the information is displayed clearly and in a proper amount;
- **Satisfaction** (3 items) explores the extent to which the user considers the experience of using the device pleasant.

Two versions of the pre- and post-experience questionnaires were built, in each one the items were phrased in accordance with the device in consideration. See Table 2 for an illustrative example.

Device	Pre-experience questionnaire	Post-experience questionnaire
Handheld device	I fear the smartphone may fall during the usage	During the use, I was concerned that the smartphone could fall
Headworn viewer	I fear the glasses may fall during the usage	During the use I was concerned that the glasses could fall

Table 2 A comparative example of how the items were worded for the two devices in the pre- and post-experience questionnaires.

Two additional questionnaires were administered after the participant had used the headworn device, one aimed at investigating the user's willingness to adopt such device and the other one assessed if the device caused hassles while wearing it on the move. The questionnaire assessing the user's willingness to use the headworn viewer in the future was adopted from Spagnolli et al. (2014) and was composed by the following dimensions:

- **Attitude Toward Technology** (6 items) explores the individual's reaction toward the use of a technology (Venkatesh, Morris, Davis and Davis 2003).
- **Technology Anxiety** (5 items) refers to the feelings uneasiness and worry the individual may have when using a technology or when the use of a technological device is prospected or even fear during the use of technology (Meuter, Ostrom, Bitnera and Roundtree 2003).
- **Facilitating Conditions** (3 items) refers to the extent to which the user perceives her environment to support her in the adoption of the device (Venkatesh, Morris, Davis and Davis 2003).
- **Perceived Usefulness** (4 items) refers to the extent to which the user thinks that the use of a certain device may improve her performance (Davis 1985)

- **Effort Expectancy** (2 items) refers to the extent to which the user thinks the device will be tiring to use (Venkatesh, Morris, Davis and Davis 2003).
- **Behavioral Intentions** (4 items) refers to the user's willingness to adopt and use a device how much a user would use a particular device (Davis 1985).
- **Psychological Attachment** (5 items) refers to the extent to which users would be inclined to adopt a certain technology in response to other people's incentive (Malhotra and Galletta 1999).
- **Perceived Enjoyment** (4 items) refers to the extent to which users expects the activity of using a device as amusing, regardless of the results of the use (Venkatesh and Davis 2000).
- **Perceived Comfort** (7 items) refers to the expected comfort of using a wearable device (Kaplan and Okur 2008).
- **Perceived Privacy** (5 items) refers to the extent to which the user perceives that the data recorded by the system are processed and stored in a safely (Perera, Holbrook, Thabane, Foster and Willison 2011).

A short questionnaire, aiming at exploring the physical sensations the user may experience while wearing a headworn viewer while on the move, namely wearability, was also administered. Such questionnaire consisted of 14 statements. The items explored the physical sensations and possible hassles caused by the components of the eyeglasses in direct contact with the user's head: the sidepieces (6 items), the bridge (4 items). Additionally, four general statements regarding the overall feelings of wearing the device (weight and balance) were included. All statements were enriched with images highlighting the device component to which each item referred.

For all the questionnaires presented respondents were asked to mark their level of agreement with the statement on a 5-point Likert scale, with 1 indicating a full disagreement and 5 indicating full agreement with the statement.

A short interview was conducted at the end of the experimental session, for an approximate duration of three minutes. The interview was video recorded and transcribed later by researchers. The purpose of the interviews was to gather the information about the experiences characterizing the use of the augmented reality app with the head-worn display and the handheld device, such as the difficulties encountered during the tasks and the perception about any mistakes that cannot be explored with closed questions in questionnaires.

4.6 Data collected

Besides exploring users' opinions on the experience of using the devices by means of questionnaires, the rate of accomplishment of the tasks required were recorded by the experimenter, as an indicator of the effectiveness of the performance. Furthermore the experimental trails were video-recorded to allow off-line video analysis. Through the video-analysis the experimenter could record the duration of each task, as a measure of the performance efficiency. Finally, via the video-analysis the way in which users used and interacted with the devices was systematically scored.

5. Results

5.1 Performance

In general the success rate for each task was high (**Errore. L'origine riferimento non è stata trovata.**). The 88.9% of the users found and reached the target POI when using the headworn viewer and the 86.1% of the users succeeded when using the smartphone. When requested to find the three nearest POIs, the 58.33% of participants accomplished the task with the headworn display, whilst the success rate is of 64.81% for those using the smartphone. The 86.11% of the participants was able to access and read the additional information requested and so did the 97.22% of those using the smartphone.

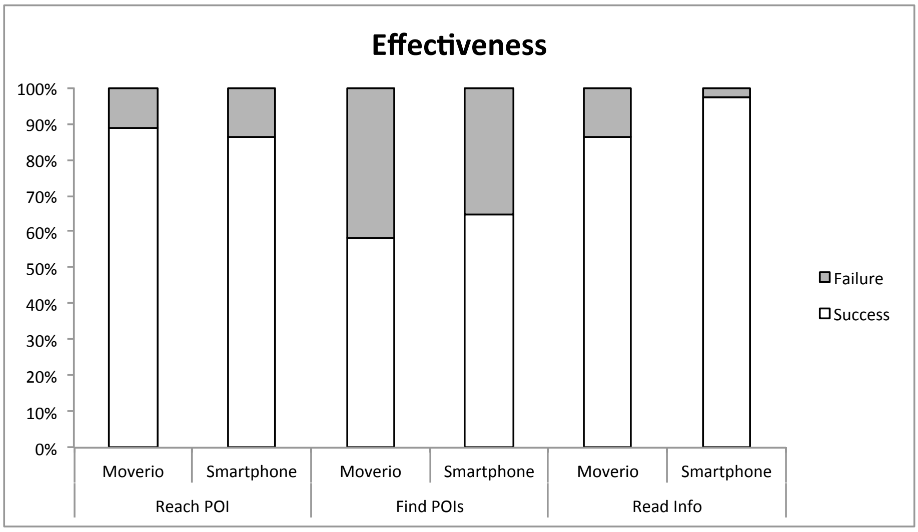

Figure 46 Success rate for the tasks examined for the two devices

A chi-squared test revealed that the success rate was not affected by the device in use in any of the tasks considered: for the request of reaching the target POI the value of the test is X^2= .127 p= .72. For the task of finding the closest POIs the test is X^2= .95 with p= .32. Finally for the task of reading the value of the test was X^2= 2.9 p= .088.

As previous research has identified a significant effect of the gender on the performance in a navigation task in an AR environment (Ahamad, Goldiez and Hancock 2005), the possible impact of the gender on the success rate was investigated by means of a chi-squared test. The analysis highlighted no significant difference in the number of tasks successfully accomplished between men and women (Figure 47). For the task of reaching the target POI the value of the test was X^2= .148 p= .7; for the task of listing the nearest POIs X^2= .618 p= .43, finally for the task of reading the additional information X^2= .00 p=.98. Consistently, no difference emerged considering the three tasks together: X^2= .55 p= .45.

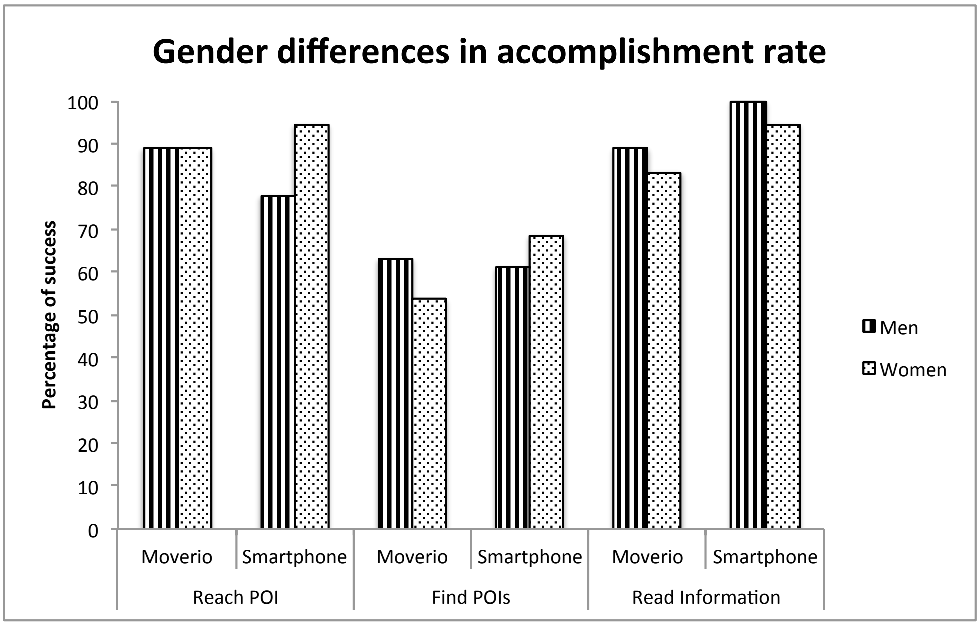

Figure 47 Success rate for men and women in the three tasks

The impact of the device on the time needed to complete the task was also considered. The comparison of the task durations was run on the subsample of users who successfully accomplished the task with both devices (n=28 for the task requiring to reach the target POI; n=30 for the task required to read information; Figure 48). For the task requiring to find and reach the target POI the Wilcoxon test showed no significant difference (*Z*= .364 *p*= .71) in the time to complete the tasks with the headworn viewer (*M*= 271.67 seconds, *SD*= 158.94) and the smartphone (*M*= 247.42 seconds, *SD*= 105.63. Since only one user completed the task of finding the three nearest POIs successfully with both devices, it was not possible to run a statistical comparison. However the average time to complete the task was in the general similar using both the headworn viewer (*M*= 58 seconds *SD*= 18) and the headworn viewer (*M*= 72 seconds *SD*= 46). Regarding the time required to access and read the information of a given POI, the Wilcoxon signed-rank test showed that the smartphone led to faster completion time compared to the headworn viewer *Z*=3.04 *p*= .002, with an average task duration of 56.36 seconds (*SD*= 55.01) using the smartphone and 94.76 seconds using the Moverio eyeglasses (*SD*= 55.01).

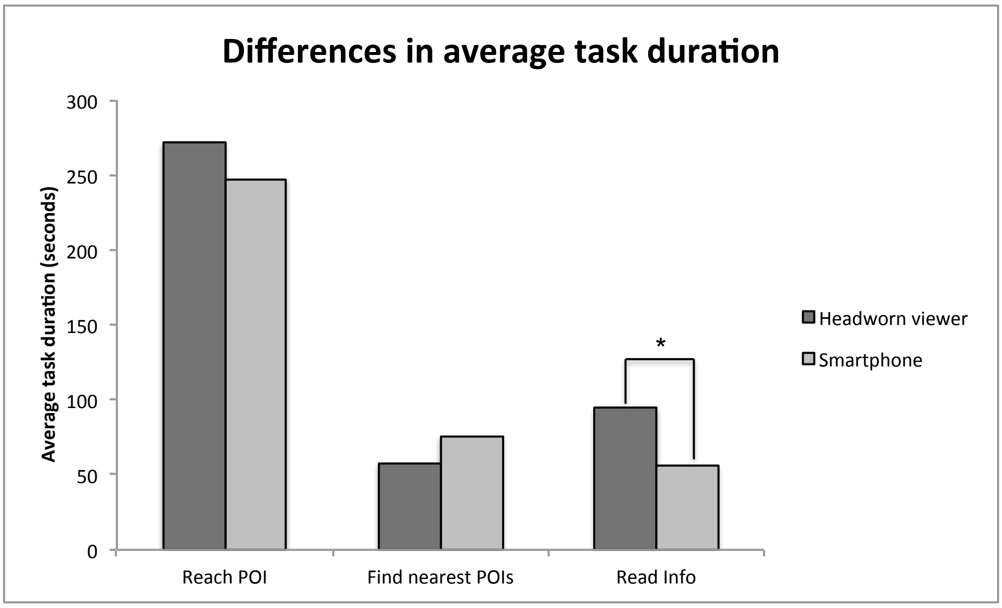

Figure 48 Average duration of the tasks

5.2 Interaction behaviors

A systematic video-analysis of the actions users performed while busy in solving the tasks was conducted on a subsample of 12 participants (5 women). The mean age of the sub-sample was 25 (*SD*= 4.18).

During the video-analysis, conducted with the software The Observer XT 11 by Noldus, the experimenter recorded the number of times specific behaviors occurred together with the duration of such events. The events considered were:

- the strategy the user realized to find and reach the target POI, i.e. how many times and for how long they needed to stop;
- the direction of the user's gaze while using the smartphone;
- the exploration of the environment the user realized in 360°;
- the extent the user engaged her arms during the task completion.

The analysis on the strategy employed by users to find and reach the target POI was conducted on 11 observations, as one user was excluded because was an outlier.

When busy at reaching the target POI, the number of stops users made did not differ statistically between the two conditions, according to a Wilcoxon signed-rank test Z= 1.866, p= .062. However the number of stops made using the headworn viewer was smaller (M= 3.5 SD= 3.21 Mdn= 2.5) than the number of stops made using the smartphone (*Mdn*= 4 *M*= 5.3 *SD*= 3.8). The number of times users restarted walking was higher for the headworn viewer (*Mdn*= 2, *M*= 2.8 *SD*= 3.01) than for the smartphone (*Mdn*= 3.5 *M*= 4.7 *SD*= 3.49), again according to a Wilcoxon signed rank test, *Z*= 2.047 *p*= .041. Taken together these findings suggest that the interaction flow was smoother when using the headworn viewer, as users made a smaller number of interruptions while reaching the target POI.

Similarly, the overall duration of the interruptions was comparable (Wilcoxon test *Z* = .764, *p* = .445) between the two conditions, with an average time spent still of 63.72 seconds (*SD*=51.13; *Mdn*= 60.88) for the smartglasses and an average duration of 63.88 (*SD*=41.80; *Mdn*= 60.88) with the smartphone.

Only for the use of smartphone, a further analysis was run on the task requiring to reach the target POI. First, the percentage of time users spent looking at the device or at the surroundings was computed with respect to the total duration of the task together with the percentage of time participants spend standing still or walking. Then a repeated measures 2 (direction of gaze) x 2 (movements) ANOVA was run. The analysis highlighted a main effect of the movement $F_{1,10}$= 28.62 *p*< .001 η^2_p= .72, with participants spending significantly more time standing still (*M*= 29.4% *ES*= 1.045) than walking (*M*= 19.67% *ES*= .94). A main effect of the direction of the gaze emerged also $F_{1,10}$= 32.54 *p*< .001 η^2_p= .74, with participants looking more at the device (*M*= 37.52% *ES*= 2.35) than at the surroundings (*M*= 11.55% *ES*= 2.26). The interaction was also significant $F_{1,10}$= 11.85 *p*= .005 η^2_p= .51. This suggests that the amount of time users spent looking at the environment did not depended on whether they were walking or standing still (t_1= 1.801 *p*= .099 M_{walk}= 9.73% *SD*= 7.36; M_{still}= 13.36% *SD*= 9.68), however users spent significantly more time attending at the smartphone while they were still (M_{still}= 45.42% *SD*=8.73) than walking ($M_{walking}$= 29.62% *SD*= 10.46), t_1= 5.538 *p*<.001. Please refer to Figure 49 for an overview of the results.

These findings suggest that in order to check the smartphone, users were forced to stop, thus interrupting their action flow.

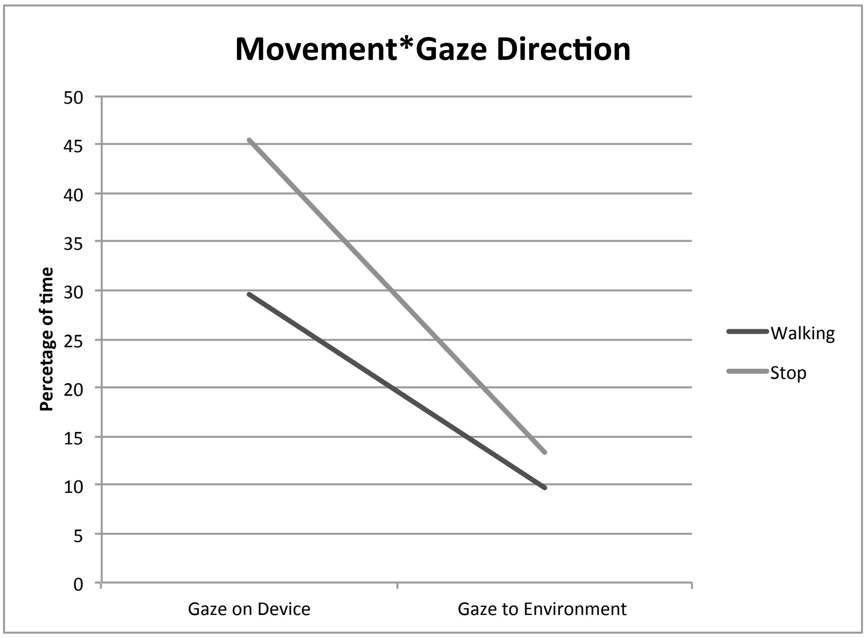

Figure 49 Percentage of time spent gazing on the device or to the environment when walking or stopping by

During the experimental trails we noticed that not all the participants were equally explorative of the environment, in particular we observed that some of them had the tendency to make small movements around, whilst others tended to make wider movements. We hypothesized a relation between the tendency of exploring the environment with wide movements and the concern of looking strange at other people's eyes, measured as Self Impression in the questionnaire. No significant correlation emerged between the time spent by the user exploring the environments with the smartphone and the consolidated score assessing Self Impression at the post-test questionnaire, r= -.81 p= .79. A significant negative correlation[16] emerged between the duration of the exploratory behaviors realized with the headworn viewer and the consolidated score assessing Self Impression at the post-test questionnaire, r= -.81 p= .004. The higher was the concern of looking strange, i.e., the score at Self Impression, the shorter the duration of the explorative behaviors.

[16] Two participants were excluded from the analysis because they were outliers.

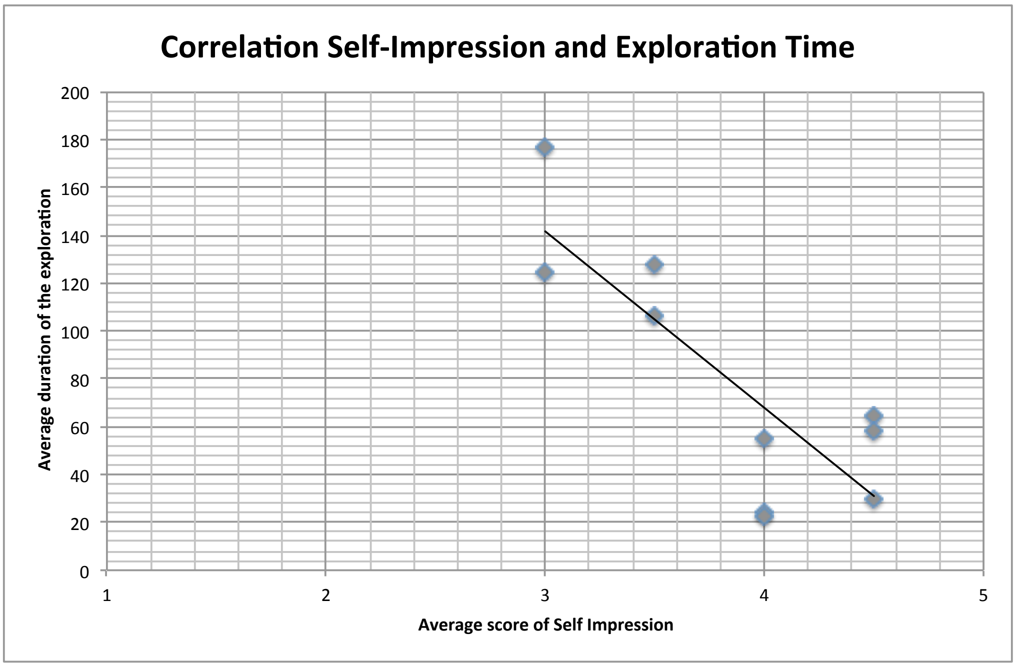

Figure 50 Correlation between the average score of Self Impression and the average duration of the exploration

During the data collection, we observed that, when using the headworn viewer, participants tended to move also their arms and hands, even if it was not needed for the purpose of accomplishing the tasks. We categorized two kinds of behaviors involving the upper limbs: movements realized to prevent light reflection and actions performed to adjust the position of the device on the head. Regarding the need to avoid bright light, 7 participants out of 12 either shield the glasses with the hand or put the hand on the side of the glass to produce shadow. This happened above all in the tasks involving reading, that is the one requiring to find the nearest POIs and the one requiring to read the information associated with a certain POI. For what concerns the need to adjust the position of the glasses or to hold them on the head, this was more frequent in the task involving the movement, i.e., the one requiring to reach the target POI, in which 8 users (out of 12) accommodated the glasses. Interestingly 4 participants did not touch the glasses at all during this task. Regarding the other two tasks, some users also showed the need to hold or adjust the device, yet to a lesser extent.

5.3 Questionnaires

Comparing users' expectations with their actual opinions

We were interested to assess what users expected from an AR app running on a headworn display or on a smartphone. Moreover we were interested to understand if their expectations matched the actual experience of use. Therefore a consolidated score was calculated for the dimensions explored by averaging the average score of the items belonging to each dimension. The consolidated scores derived from the pre- and post-test questionnaires were then compared.

For what concerns the headworn display, we found that participants expected the device to be useful to a higher extent before trying it than after, according to a Wilcoxon signed-rank test $Z = 4.517$, $p < .001$; the mean score before using the device was 4.02 (*SD*= .73, *Mdn*= 4) and 3.12 (*SD*= .75, *Mdn*=3.3) after using it.

Regarding the concerns of being stared at by other people, respondents seemed less worried after the use than before it, again according to a Wilcoxon test, $Z = 2.664$, $p = .008$, the mean score was 3.91 (*SD*= .48, *Mdn*= 4) before the test and 3.59 (*SD*= .68, *Mdn*= 3.67) after it. The Wilcoxon test also highlighted that users expected the device to be easier to use with respect to what actually was while using it, $Z = 3.171$, $p = .002$; with a mean score before the experiment of 3.23 (*SD*=.89, *Mdn*= 3.5) and a mean score after the test of 2.80 (*SD*=.81, *Mdn*= 3). The expected comfort of wearing the glasses was confirmed by the actual experience (see Table 1).

In general the expectations users had with the respect to the use of the handheld device were confirmed by the actual use of the device. This happened for the self impression and the comfort. The expected usefulness was higher before the experience, than after it $Z = 4.527$, $p < .001$, the mean score before the test was 4.11 (*SD*=.52, *Mdn*= 4) and dropped to 3.43 (*SD*=.65, *Mdn*= 3.33) after the use. Also the ease of use decreased after the use $Z = 2.45$, $p = .014$, with a mean score before the test of 4.08 (*SD*= .30, *Mdn*= 3.33) and a mean score after the test of 2.92 (*SD*= .67 *Mdn*= 3).

Dimension	Headworn viewer					Smartphone				
	Z	Mean(SD)		Median		*Z*	Mean(SD)		Median	
		Pre	Post	Pre	Post		Pre	Post	Pre	Post
Usefulness	4.517**	4.02(.73)	3.12(.75)	4	3.33	4.527**	4.11(.52)	3.43(.65)	4	3.33
Ease Of Use	2.664*	3.91(.48)	3.59(.68)	4	3.67	2.45*	4.08(.30)	2.92(.67)	3.33	3
Self Impr.	3.171*	3.23(.89)	2.80(.81)	3.5	3	1.401	2.01(.76)	2.17(.77)	2	2
Comfort	1.902	3.04(.81)	2.81(.94)	3	3	.572	2.40(.73)	2.32(.73)	2.25	2.5

Table 1 Results from the Wicoxon's test comparing the scores at the questionnaires before and after the use of the headworn viewer (on the left) and the smartphone (on the right). The single asterisk indicates p<.05, the double asterisk indicates *p*<.001.

Comparing users' impressions of the two devices

Another point of interest in the present study was to compare the experience of use between the two devices. We found that in general the evaluation of the experience of use was more positive for the smartphone, than for the headworn viewer. A significant difference emerged for the visual comfort, regarding which the smartphone was judged more positively compared to the eyeglasses, according to a Wilcoxon signed-rank test *Z*= 3.356, *p*= .001, with an average score for the former being 2.67 (*SD*= .70, *Mdn*= 2.5)[17] and for the latter 3.28 (*SD*= .99, *Mdn*= 3). Similarly users were less concerned of what other people may have thought of them while using the handheld device (*M*= 2.17, *SD*= .77, *Mdn*= 2) compared to the headworn viewer (*M*=2.80, *SD*= .81, *Mdn*= 3), *Z*= 4.125, *p* < .001. Even if not significant it is worth noticing that users were more satisfied of their use of the headworn viewer compared to the smartphone.

Dimension	Z	Mean(SD)		Median	
		HWD	Smart	HWD	Smart
Usefulness	2.219	3.12(.75)	3.43(.64)	3.33	3.33
Ease Of Use	.755	3.59(.68)	2.92(.67)	3	3
Self Impression	4.125**	2.80(.81)	2.17(.77)	3	2
Comfort	2.387	2.81(.94)	2.32(.73)	3	2.5
Visual Comfort	3.356*	3.28(.99)	2.67(.70)	3	2.5
Satisfaction	1.158	3.77(.85)	3.65(.82)	4	4

Table 2. Results from the Wilcoxon's test comparing the scores in the post-experience questionnaire between the two devices. The single asterisk indicates p<.05, the double asterisk indicates p<.001.

[17] The wording of the statements was negative, thus a low value indicates a positive evaluation.

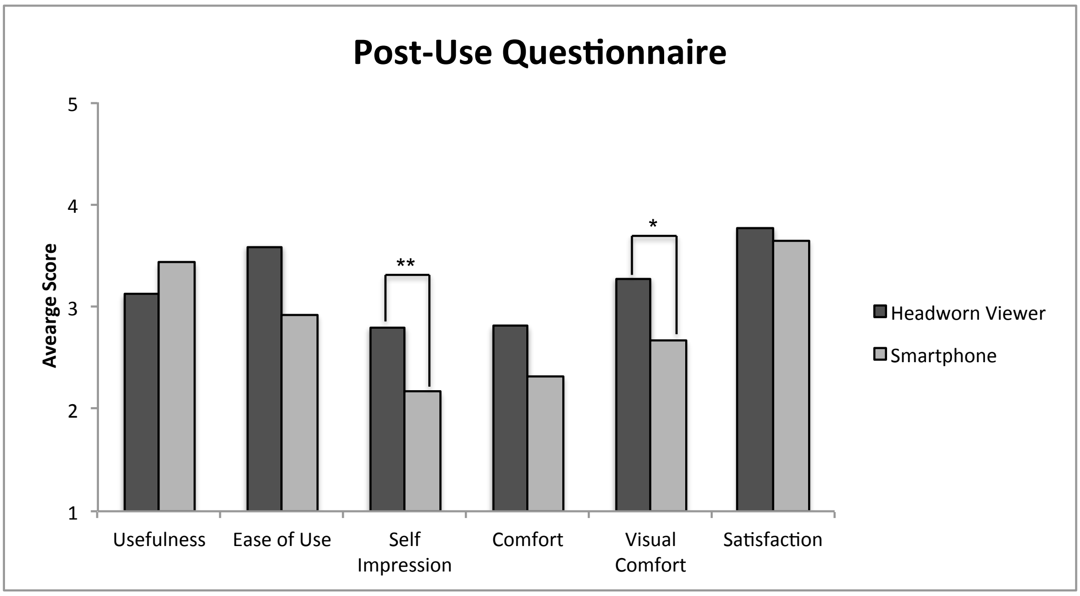

Figure 51 Comparison of the average scores for the two device

Finally we compared separately the scores of the items assessing how satisfactory the use between the two devices was (Figure 52). The Wilcoxon signed rank test revealed that, participants found more amusing to use the headworn device than the smartphone, $Z= 2.66$ $p= .008$, respectively the average score for the headworn display was 4 (SD=.92), and the average score for the smartphone was 3.5 (SD=.91). The other two comparisons, investigating the willingness to use the application again and the pleasantness of use of the application, did not differ significantly between the two devices. For the pleasantness of use the Wilcoxon signed rank test was $Z= .23$ $p= .81$, with a mean score of 3.78 (SD= .79, *Mdn*= 4) for the headworn display and 3.72 (SD= .91, *Mdn*= 4) for the smartphone. Regarding the willingness to use the application again, the Wilcoxon signed rank test was $Z= 1.34$ $p= .17$, with $M= 3.53$ (SD= 1.23, *Mdn*= 4) for the eyeglasses and 3.75 (SD= .90, *Mdn*= 4) for the smartphone.

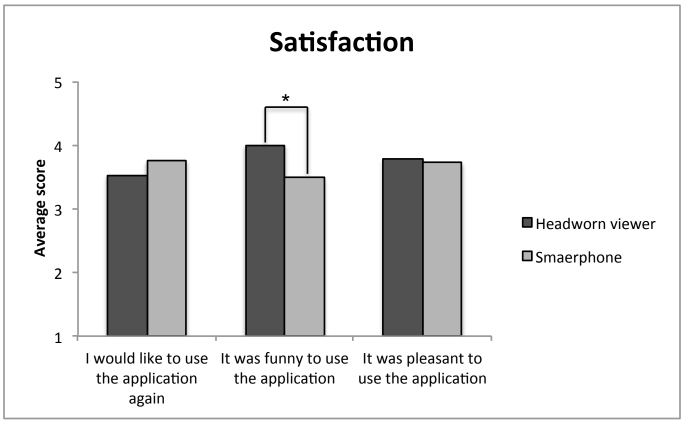

Figure 52 Comparison of the itmes assessing the satisfaction

Comfort of wearing a headworn viewer

This questionnaire was introduced during the data-collection, so a sub-sample 13 participants filled it in, 8 of them were women. The mean age was 26.7 (*SD*=9.15).

The items investigating the comfort of wearing the eyeglasses can be grouped in three categories: items exploring the physical sensations on the nose, that is the bridge of the eyeglasses, items exploring the physical sensations on the head, that is the sidepieces of the glasses, and finally items exploring the general feeling of wearing the device. A one sample t-test against the central value of the response scale, i.e., 3, was run to assess the direction of the users' opinions for each of the abovementioned dimensions. We found that respondents thought that the components of the eyeglasses at direct contact with their nose were not annoying, t_{12}= 3.36, *p*= .006 *M*= 2.37 (*SD*= .68). A similar evaluation emerged for the sidepieces, t_{12}= 4.768, *p*< .001, with a mean score of 1.85 (*SD*= .87). The positive trend was confirmed also with respect to the items investigating general aspects of comfort, t_{12} = 4.07, *p*= .002 (*M*= 2.1, *SD*= .80).

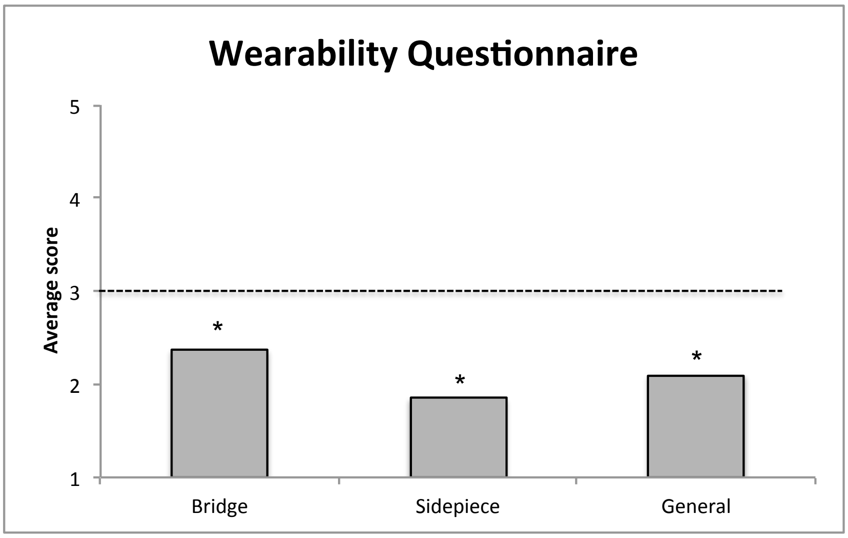

Figure 53 Average scores for the comfort of the different components of the headworn viewer

Users' acceptance of the headworn viewer

To assess which of the dimensions could predict the users' acceptance, a linear regression model was run with the Behavioral Intention as dependent variable and the dimensions assessed by the questionnaire and the age as predictors.

The most satisfactory linear regression model (Table 3) included three dimensions assessed by the questionnaire: Perceived Usefulness, Effort Expectancy and Psychological Attachment. In this model, explaining the 81% of the variance, Perceived Usefulness and Effort Expectancy were strong predictors of Behavioral Intention, respectively β= .429 p< .001 and β= .4 p< .001. Psychological Attachment was also a good predictor of the intention of use β= .329 p= .001.

Predictors of Behavioral intention			
Variables	**Standardized Coefficients**	**Behavioral intention Model 1**	
	β	**t**	**p**
(constant)		-3.687	.001*
Perceived Usefulness	**.429**	4.248	.000**
Effort Expectancy	**.400**	4.185	.000**
Psychological Attachment	**.329**	3.663	.001*

Table 3 The linear regression model

5.4 Interviews

The most frequent comment emerged during the interviews regarded the amount of information displayed, which was considered too abundant by one third of those who answered the questions. On respondent in particular mentioned the clutter produced by so many labels on the screen:

"There were so many labels in front of my view that I was worried not to see people and to bump into them"

(F, 25)

Other respondents claimed that the size of the text was too small and it was difficult for them to read it. They also imagined to be in an elderly person shoes, for whom it would be impossible to read the texts:

"In the glasses the texts were too small. If I was my mum, she wouldn't see anything, because it's too small."

(F, 27)

A couple of participants complained that the augmented labels were floating and overlapping to each other, making it difficult to see clearly. Consistently with what emerged from the questionnaires, the glasses were more amusing to use, and in some cases this was directly compared to a higher usefulness of the smartphone:

"I found the glasses funnier, but the smartphone was more handy"

(M, 24)

Nevertheless for some users the glasses outperformed the smartphone in terms of functionality. One of them commented that with glasses "it was easier to see where I was going" (M, 18) and another one stated that the labels were overlapping much less with the glasses.

6. Discussion

To our knowledge, no previous studies have addressed the usability and the user experience of a headworn viewer for AR compared to a smartphone. Here, we investigated if a light weighted head mounted display is usable in a touristic scenario to a similar extent of a more common AR support device, that is a smartphone, and if the experience of use was positive for the users.

A headworn viewer seems to be a valuable tool for supporting the exploration of an urban environment. It resulted comfortable on the head to wear and efficient for supporting the user in the execution of the requested tasks, as the rate of success was equal to that of the handheld device. Interestingly, the smartphone seemed to force users to stop and to check the display of the device, thus interfering with the action flow. This behavior is not different from the one a user would make to check information on a paper map. Thus, the advantages of AR seem not to be fully exploited and the device seems to obtrusively mediate the experience.

Users were more concerned about their appearance with the headworn viewer before the use than after, consistently with Denning et al., (2014), who reported that users are more likely to change their mind after they have actively tried a technology. Furthermore none of them commented of feeling observed or complained of feeling embarrassed when wearing the headworn viewer. This finding is in line with previous research, according to which participants were more incline to consider smartglasses striking when worn by other than by themselves Koelle et al., (2015).

Reading the text was found to be an issue, which emerged consistently from both objective performance data, i.e., the duration of the task, and subjective report from users. Interestingly, users' complains referred above all to visual the large quantity of labels superimposed onto the environment, which made difficult to identify relevant information, which is a well-known critical issue in mobile AR applications (Tatzgern et al., 2016; Li and Duh 2013). As expected light reflections bothered users while using the smartphone (Jacob 2011), but the headworn viewer seemed to present the same issue.

The extent to which participants praised the usefulness significantly decreased for both the devices after the trail, which suggests that this evaluation is to be attributed to the application itself. As da Sà and Churchill (2013) pointed out, for many users is hard to envision a long-term value in mobile AR apps. So far, in fact the majority of the AR applications developed tend to emulate already existing services, failing to fully exploit the advantages of augmented real-world objects.

Chapter 6

General Discussion

The present work has explored different aspects of the interaction with advanced systems, in the attempt to unveil how users relate with such systems in order to contribute to the creation of new standards of interaction between the human and advanced machines. We have explored gestural interfaces in the development of a set interaction commands. A second theme of investigation involved wearables: a conventional audio-visual system based on a hand-held device was compared with a haptic-audio interface paired with a haptic glove. In addition, the feasibility of delivering navigational directions by means of vibrotactile stimuli was compared for two different wearable devices (i.e., a tactile vest and a tactile glove). Finally, the usability and the experience of use with a commercial mobile AR application were compared between a wearable and a hand-held device, namely a pair of smartglasses and a smartphone.

In the first experiment we have developed a virtual simulation of a prototype representing an interactive hood and involved a group end-users for defining a set of natural gestures aiming to control the lighting system of the hood. Despite the fact that design of the appliance was minimal and plain, it was observed that participants could find some hints that consistently attracted their gestures. This is slightly in contrast with the hypothesized lack of affordances that interfaces embedded in the environment would have according to Marquardt and Greenberg (2012). In addition a clear reference to the pre-existing practices emerged for the control of specific functions (i.e., the modulation of the size and the brightness of the spotlight), participants tended in fact to operate on the interface the same way they would do on a touch-screen device (with a pinch gesture) or on the lens of a camera (rotating the lens). This tendency suggests that the interaction modalities learnt are generalized by the user with respect to the interface to which they have been association in the first place. The need of building new standards advocated by Norman (2010) seems to be at least in part mitigated by the spontaneous reference to the already known interaction gestures. With this we do not mean that new interaction norms for gestural interfaces are not desirable, rather that exploiting the conventions already existing for other interfaces can be a fruitful starting point. Furthermore, what has been observed can be interpreted as an extension of the classical usability principle that encourages developers to favor the use of elements that are already known to users rather than make them learn new ones (Nielsen 1993). All these elements together call for an active involvement of users in the design not only of the gestural vocabulary (Wobbrock, Morris and Wilson 2009; Nielsen, Störring, Moeslund and Granum 2008) but also in the design of the interface itself: by carefully observing their spontaneous interpretation and use of the different elements of the interface, developers would have a meaningful view of the actions each element provokes. Virtual prototyping, that is to say the use of virtual reality to create digital prototypes (Wang 2002), was already employed for testing the apparent usability of home appliances (e.g., Bordegoni, Ferrise and Lizaranzu 2011). Here we show that this method could elicit consistent reactions in a group of users, thus supporting an extension in its adoption.

The second experiment opposed two systems in an exploration task that took place in an urban setting. The first system consisted of a haptic-audio interface in which the interaction was mediated by hand gestures input by means of a tactile glove, while the comparison system was a conventional audio-visual interface based on a smartphone. We found that the haptic-audio interface made participants identify the landmarks with the same degree of accuracy and even faster than the audio-visual interface. In addition, despite its novelty, the haptic-audio interface received positive appraisals from users, who showed low concern of appearing weird at bystanders' eyes. Considered the positive reactions of users, these findings support the potential of vibrotactile interfaces to be integrated into wearables also for more general application domains than the navigation one. Furthermore, it supports the informative value of vibrotactile messages (Brown, Brewster and Purchase 2006). Despite the high degree of accuracy, some difficulties were observed in identifying the landmark following the tactile guidance when the target was out of participants' view. This highlights how much every single interaction mode can be more or less appropriate given the contextual circumstances. The more a system is complex the more is relevant for an effective interaction to include more than a single interaction modality, to allow the user to select the mode that better fits a given situation (Oviatt and Cohen 2000; Oviatt 2003).

In the third experiment we have compared the efficacy and efficiency of two vibrotactile interfaces in a navigational task. One interface was embedded into a vest, namely a tactile vest, while the other one was shaped into a glove, namely a tactile glove. Regardless of the shape of the device, both vibrotactile displays proved to be an efficient tool for navigating users in an urban environment and were both well received by them. Also in this case, users seemed not to worry about their appearance while wearing the devices, even if a general concern emerged regarding the bulkiness of the tactile vest as compared to the glove, in line with previous claims in the literature (Gemperle et al., 1998). In addition, the need for a considerable distance between the actuators emerged, as users sometimes struggled in clearly perceiving which actuator was active on the tactile glove. Finally, results from this experiment support the need to reduce the number of actuators, in order to provide the user with the minimal amount of signals, instead of overloading her with multiple vibrations (Srikulung 2011).

The fourth experiment dealt with the comparative evaluation of the usability and experience of use with two different devices, that is to say a smartphone and a pair of smartglasses, both supporting a mobile AR application. The comparison comprised three tasks that are common in the touristic domain: finding and reaching a given destination, selecting a certain number of points of interest and reading textual information regarding a landmark. The data collected showed that the smartphone outperformed the smartglasses in the reading task, but the two were equally effective in the other two assignments. Qualitative observations suggest that the wearable system supported a smoother action flow, as users tended to stop less often compared to the smartphone. This finding encourages the employment of wearable devices for supporting computational systems unobtrusively integrated in our daily routines (Cho, Lee and Cho 2009; Baber 2001). Furthermore, interpreting observational data together with the self-reported measures, it emerged an association between the concern of appearing weird at other people's eyes and the exploration users did of the surroundings. In particular, the less were users concerned, the more they tended to turn around. This result is in line with previous findings referring to willingness to perform gestural commands in public spaces (Rico and Brewster 2010). Nevertheless, it also expands the existing literature, as for the first time – at least to the writer's best knowledge – the users' actual behavior it associated with their attitudes, while previous studies had examined the matter via self-reported metrics (e.g., Rico and Brewster 2010), or indirectly by means of video fragments (e.g., Profita et al., 2013) and group discussions (e.g., Suero Montero, Alexander, Marshall and Subramanian 2010).

Considering the findings altogether we can draw a set of general recommendations that can be useful to consider when designing advanced systems:

1. **Give users the chance to act,** to uncover affordances in gestural interfaces observe users' spontaneous reactions to the elements of the interface first and after that associate the features emerged to specific functions.

   The direct involvement of users in the design of systems is something advocated from participatory design techniques since a long while (Litosseliti 2007). Such methods typically aim at either understanding users' habits and opinions regarding a specific topic to inform the design or at receiving early feedback regarding the design. Here we advice to extend users' involvement. We recommend engaging users not only to share their views with designers, but also, and more importantly, to observe their behaviors and the actions that the interface design naturally prompts.

2. **Take advantage of established interaction practices**
   a. **Favor recognition,** the knowledge on how to operate a particular interface is easily exported and applied to other systems. Leverage users' pre-existing knowledge by inserting elements that directly evoke already known actions.
   b. **Know your users,** identify the target user group, observe them and identify their daily practices.

Nielsen and Molich (1993) were the first to recommend developers to favor the recognition of elements of the interface to relieve the user's memory load. Their recommendation referred to the need of making all the available options *visible* for the user. Here we encourage developers to exploit the interactions users already master for controlling other systems, regardless of the presence of overt hints on the interface, following Jacob (2011). To say it in other words, the user should know that for zooming in she has to make a pinch-out gesture, exactly the same way she would operate on her smartphone.
For this "cross reference" to be effective, developers need to have a deep understanding of who is expected to use the system. Only the comprehension of the end users' habits and knowledge can provide developers with the proper instruments for building an intuitive set of commands.

3. **Try it virtual first,** advanced systems can be very complex being constituted by the integration of different components. Having users approach a virtual prototype can provide valuable insight for informing the design.

   Using a virtual version of the GUI is a common practice especially for evaluating the usability in the early stages of the project. Here we encourage developers to produce virtual versions of the prototypes pertaining not only graphical interfaces, rather virtual versions of the systems involving different types of interactive modes. A virtual representation is in fact vivid enough to trigger spontaneous and consistent reactions among users.

4. **Favor multimodality,** the appropriateness of every interaction modality varies according to the contextual characteristics. The user should be able to chose the interaction mode that best fits the situation.

   Advanced systems should be developed so that the user can interact with them over long period of time, ideally they should be always available for the user (Weiser and Brown 1996; Baber 2001). This implies that they should be able adapt to different contexts of use (e.g., static versus mobile) and to different environmental characteristics (e.g., bright sunshine versus artificial lighting). To realize an efficient and satisfactory interaction, the user should be able to choose the interaction mode that best fits the contextual characteristics (Oviatt 2003; Oviatt and Cohen 2000).

5. **Allow adaptability,** the user should be able to adjust the settings of the features (e.g., size of the text, volume of audio) in every modality, in order to better respond to the environmental properties.

   Visual, hearing, manual and cognitive abilities greatly vary among users. Advanced systems should allow users to customize the characteristics of the different interaction modes enabled, so that the user can interact at any time with the minimum effort.

6. **Make them try,** the attitude toward advanced systems tend to become more positive after the use. Make users approach the system at their own pace and in a natural context.

   We, as practitioners, should not take for granted that users will receive enthusiastically a system just because of its innovativeness. Rather, users' appreciation of the system seems to increase as they actually use it. Allowing users to approach the system gently and following their own learning pace, help the user feel the system within her reach. Furthermore this gradual approach allows them to progressively discover and master the functionalities of the systems they desire and need.

7. **Put it on,** wearable devices perform better in mobility compared to handheld devices and should be preferred.

   Wearable computers enhance unobtrusively users' abilities, as users can perform the tasks without interrupting other course of actions. This advantage becomes particularly valuable in mobility. For this reason, wearables should be favored to traditional hand-held devices.

8. **Balance it,** the form factor of the device should be the most unbulky and unobtrusive possible, yet allowing the user to easily input commands and perceive the feedbacks from the system.

Despite the intrinsic advantage of using wearables described above, a satisfactory balance is needed between the device's functionalities and its aspect. In fact, the appearance of wearable device is known to affect user's willingness to adopt and use the system (Spagnolli et al., 2014; Gemperle 1998) and also the perceived comfort plays a relevant role in this (Knight et al., 2006). Nevertheless, the overall size of the weareable is related to the amount and type of information he system can deliver, with larger wearables potentially supporting richer interactions (O'Malley and Gupta 2008).

Chapter 7

Conclusions

Today computers pervade our world. The majority of the tools that are in our houses and offices include some kind of computational systems. Furthermore the processing power of portable devices has significantly increased, enabling us to interact with digital devices almost constantly and everywhere. The context in which we interact with computational systems has left the static setting bounded to the desktop and has become increasingly mobile, going through the continuously changing environmental characteristics. As we have shown, wearable devices have proved to perform well in such situations, because they allow users to interact without interrupting the other courses of actions while mobile. Furthermore users tend to have a positive attitude toward wearables. Additionally, such devices can deliver information via alternative modalities than the conventional GUIs, that is to say via vibrations. An audio-haptic interface has proved to support users equally well compared to a traditional audio-visual interface in discovering relevant information in their surroundings. At the same time, we have found that vibrotactile guidance alone may be not sufficient under certain circumstances. This finding reflects the importance of delivering information through more than one single sensory channel and let the user to select the more appropriate modality under the current contextual characteristics. The ability of accessing information in mobility is affected also by how the wearable itself is shaped, as a balance between the clarity in perceiving the signals and the overall bulkiness of the device. Finally, we showed that it is worthy investigating the spontaneous reactions users have in front of a gestural interface. Despite the recognized lack of evident affordances, consistent tendencies in users' actions emerged and, more interestingly, they clearly relied on the interaction modalities they already knew.

In the present work we have treated a variety of applications and interaction techniques. Nevertheless an underlying remark strongly links all the different experiences undertaken. In both the early and advanced stages of the design life cycle, the development of a system should be drawn from the contextual characteristics to which the final system is addressed, rather than being focused on the technology itself. Likewise, the evaluation of such systems cannot overlook the characteristics of the environment in which the system is going to be employed and those of the end users. In that, the testing should be conducted resembling as much as possible a realistic situation.

References

Abowd, G. D. (2012, September). What next, ubicomp?: celebrating an intellectual disappearing act. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 31-40). ACM.

Ahmaniemi, T. T., & Lantz, V. T. (2009, November). Augmented reality target finding based on tactile cues. In *Proceedings of the 2009 international conference on Multimodal interfaces* (pp. 335-342). ACM.

Asif, A., Heuten, W., & Boll, S. (2010, October). Exploring distance encodings with a tactile display to convey turn by turn information in automobiles. In Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries (pp. 32-41). ACM.

Azuma, R.T.: A survey of augmented reality. Presence: Teleoperators and Virtual Environments 6(4), 355–385 (1997)

Baber, C. (2001). Wearable computers: a human factors review. International Journal of Human-Computer Interaction, 13(2), 123-145.

Ballendat, T., Marquardt, N., & Greenberg, S. (2010, November). Proxemic interaction: designing for a proximity and orientation-aware environment. In ACM International Conference on Interactive Tabletops and Surfaces (pp. 121-130). ACM.

Barfield, W., Mann, S., Baird, K., Gemperle, F., Kasabach, C., Stivoric, J., ... & Cho, G. (2001). Computational clothing and accessories. Fundamentals of Wearable Computers and Augmented Reality, Lawrence Erlbaum Associates, 471-509.

Bial, D., Kern, D., Alt, F., & Schmidt, A. (2011, May). Enhancing outdoor navigation systems through vibrotactile feedback. In CHI'11 Extended Abstracts on Human Factors in Computing Systems (pp. 1273-1278). ACM.

Bobeth, J., Schrammel, J., Deutsch, S., Klein, M., Drobics, M., Hochleitner, C., & Tscheligi, M. (2014, June). Tablet, gestures, remote control?: influence of age on performance and user experience with iTV applications. In Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video (pp. 139-146). ACM.

Bolt, R. A. (1980). Put-that-there: Voice and gesture at the graphics interface. Computer Graphics, 14(3), 262–270.

Bordegoni, M., Ferrise, F., & Lizaranzu, J. (2011, March). Use of interactive virtual prototypes to define product design specifications: A pilot study on consumer products. In VR Innovation (ISVRI), 2011 IEEE International Symposium on (pp. 11-18). IEEE.

Bosman, S., Groenendaal, B., Findlater, J. W., Visser, T., de Graaf, M., & Markopoulos, P. (2003). Gentleguide: An exploration of haptic output for indoors pedestrian guidance. In Human-computer interaction with mobile devices and services (pp. 358-362). Springer Berlin Heidelberg.

Brereton, M., Bidwell, N., Donovan, J., Campbell, B., & Buur, J. (2003, February). Work at hand: an exploration of gesture in the context of work and everyday life to inform the design of gestural input devices. In Proceedings of the Fourth Australasian user interface conference on User interfaces 2003-Volume 18 (pp. 1-10). Australian Computer Society, Inc..

Brown, L. M., Brewster, S. A., & Purchase, H. C. (2006, September). Multidimensional tactons for non-visual information presentation in mobile devices. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 231-238). ACM.

Brusie, T., Fijal, T., Keller, A., Lauff, C., Barker, K., Schwinck, J., ... & Guerlain, S. (2015, April). Usability evaluation of two smart glass systems. In Systems and Information Engineering Design Symposium (SIEDS), 2015 (pp. 336-341). IEEE.

Buxton, W. (1997). Living in Augmented Reality: Ubiquitous Media and Reactive Environments. In K. Finn, A. Sellen & S. Wilber (Eds.). Video Mediated Communication. Hillsdale, N.J.: Erlbaum, 363-384.

Cabral, D., Orso, V., El-khouri, Y., Belio, M., Gamberini, L., & Jacucci, G. (2014, October). The role of location-based event browsers in collaborative behaviors: an explorative study. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (pp. 951-954). ACM.

Carton, A., & Dunne, L. E. (2013, March). Tactile distance feedback for firefighters: design and preliminary evaluation of a sensory augmentation glove. In *Proceedings of the 4th Augmented Human International Conference* (pp. 58-64). ACM.

Cho, G., Lee, S., & Cho, J. (2009). Review and reappraisal of smart clothing. International journal of human-computer interaction, 25(6), 582-617.

Chouvardas, V. G., Miliou, A. N., & Hatalis, M. K. (2008). Tactile displays: Overview and recent advances. *Displays*, *29*(3), 185-194.

Costanza, E., Inverso, S. A., Pavlov, E., Allen, R., & Maes, P. (2006, September). Eye-q: Eyeglass peripheral display for subtle intimate notifications. In Proceedings of the 8th conference on Human-computer interaction with mobile devices and services (pp. 211-218). ACM.

de Sá, M., & Churchill, E. F. (2013). Mobile augmented reality: A design perspective. In Human Factors in Augmented Reality Environments (pp. 139-164). Springer New York.

Dix, A. (2009). Human-computer interaction (pp. 1327-1331). Springer US.

Do, J. H., Jang, H., Hoon Jung, S., Jung, J., & Bien, Z. (2005, August). Soft remote control system in the intelligent sweet home. In Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on (pp. 3984-3989). IEEE.

Duchowski, A. (2007). Eye tracking methodology: Theory and practice (Vol. 373). Springer Science & Business Media.

E. Hornecker, S. Swindells, and M. Dunlop. A mobile guide for serendipitous exploration of cities. In Proc. MobileHCI, pages 557–562, 2011.

Epps, J., Lichman, S., & Wu, M. (2006, April). A study of hand shape use in tabletop gesture interaction. In CHI'06 extended abstracts on human factors in computing systems (pp. 748-753). ACM.

Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., & Sharit, J. (2009). Designing for older adults: Principles and creative human factors approaches. CRC press.

Freeman, E., Brewster, S., & Lantz, V. (2014, November). Tactile Feedback for Above-Device Gesture Interfaces: Adding Touch to Touchless Interactions. In *Proceedings of the 16th International Conference on Multimodal Interaction* (pp. 419-426). ACM.

Ganapathy, S. (2013). Design guidelines for mobile augmented reality: User experience. In Human Factors in Augmented Reality Environments (pp. 165-180). Springer New York.

Gandy, M., Ross, D., & Starner, T. E. (2003). Universal design: Lessons for wearable computing. Pervasive Computing, IEEE, 2(3), 19-23.

Garzotto, F., & Valoriani, M. (2013, September). Touchless gestural interaction with small displays: a case study. In Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI (p. 26). ACM.

Gemperle, F., Kasabach, C., Stivoric, J., Bauer, M., & Martin, R. (1998, October). Design for wearability. In Wearable Computers, 1998. Digest of Papers. Second International Symposium on (pp. 116-122). IEEE.

Go, K., & Carroll, J. M. (2004). Scenario-based task analysis. The handbook of task analysis for human-computer interaction, 117-133.

Gorlenko, L., & Merrick, R. (2003). No wires attached: Usability challenges in the connected mobile world. IBM Systems Journal, 42(4), 639-651.

Greenberg, S., Boring, S., Vermeulen, J., & Dostal, J. (2014, June). Dark patterns in proxemic interactions: a critical perspective. In Proceedings of the 2014 conference on Designing interactive systems (pp. 523-532). ACM.

Greenfield, A. (2010). Everyware: The dawning age of ubiquitous computing.

Hall, E. T. (1966). The hidden dimension .

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in psychology, 52, 139-183.

Haugstvedt, A. C., & Krogstie, J. (2012, November). Mobile augmented reality for cultural heritage: A technology acceptance study. In Mixed and Augmented Reality

(ISMAR), 2012 IEEE International Symposium on (pp. 247-255). IEEE.

Heath, C., Hindmarsh, J., & Luff, P. (2010). Video in qualitative research. Sage Publications.

Heuten, W., Henze, N., Boll, S., & Pielot, M. (2008, October). Tactile wayfinder: a non-visual support system for wayfinding. In Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges (pp. 172-181). ACM.

Hoggan, E., & Brewster, S. (2006, October). Crossmodal spatial location: initial experiments. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles (pp. 469-472). ACM.

Hoggan, E., Anwar, S., & Brewster, S. A. (2007). Mobile multi-actuator tactile displays. In *Haptic and audio interaction design* (pp. 22-33). Springer Berlin Heidelberg.

Huang, W., Alem, L., & Livingston, M. (2012). Human factors in augmented reality environments. Springer Science & Business Media.

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008, April). Reality-based interaction: a framework for post-WIMP interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 201-210). ACM.

Jacob, R., Mooney, P., & Winstanley, A. C. (2011, September). Guided by touch: Tactile pedestrian navigation. In Proceedings of the 1st international workshop on Mobile location-based service (pp. 11-20). ACM.

Jaimes, A., & Sebe, N. (2007). Multimodal human–computer interaction: A survey. Computer vision and image understanding, 108(1), 116-134.

Jylhä, A., Hsieh, Y. T., Orso, V., Andolina, S., Gamberini, L., & Jacucci, G. (2015, November). A Wearable Multimodal Interface for Exploring Urban Points of Interest. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (pp. 175-182). ACM.

Kjeldskov, J., & Stage, J. (2004). New techniques for usability evaluation of mobile systems. International journal of human-computer studies, 60(5), 599-620.

Ko, S. M., Chang, W. S., & Ji, Y. G. (2013). Usability principles for augmented reality applications in a smartphone environment. International Journal of Human-Computer Interaction, 29(8), 501-515.

Koelle, M., Kranz, M., & Möller, A. Don't look at me that way!–Understanding User Attitudes Towards Data Glasses Usage.

Li, N., & Duh, H. B. L. (2013). Cognitive Issues in Mobile Augmented Reality: An Embodied Perspective. In Human Factors in Augmented Reality Environments (pp. 109-135). Springer New York.

Lim, H., Cho, Y., Rhee, W., & Suh, B. (2015, April). Vi-Bros: Tactile Feedback for Indoor Navigation with a Smartphone and a Smartwatch. In Proceedings of the 33rd

Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (pp. 2115-2120). ACM.

Linaza, M. T., Marimón, D., Carrasco, P., Álvarez, R., Montesa, J., Aguilar, S. R., & Diez, G. (2012). Evaluation of mobile augmented reality applications for tourism destinations. na.

Litosseliti, L. (2003). Using focus groups in research. A&C Black.

López-Cózar, R., & Callejas, Z. (2010). Multimodal dialogue for ambient intelligence and smart environments. In Handbook of ambient intelligence and smart environments (pp. 559-579). Springer US.

Manhartsberger, M., & Zellhofer, N. (2005). Eye tracking in usability research: What users really see. In Usability Symposium (Vol. 198, No. 2, pp. 141-152).

Marquardt, N., & Greenberg, S. (2012). Informing the design of proxemic interactions. IEEE Pervasive Computing, (2), 14-23.

McAtamney, G., & Parker, C. (2006, April). An examination of the effects of a wearable display on informal face-to-face communication. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 45-54). ACM.

Montero, C. S., Alexander, J., Marshall, M. T., & Subramanian, S. (2010, September). Would you do that?: understanding social acceptance of gestural interfaces. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (pp. 275-278). ACM.

Morris, M. R., Huang, A., Paepcke, A., & Winograd, T. (2006, April). Cooperative gestures: multi-user gestural interactions for co-located groupware. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 1201-1210). ACM.

Nielsen, M., Störring, M., Moeslund, T. B., & Granum, E. (2004). A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In Gesture-Based Communication in Human-Computer Interaction (pp. 409-420). Springer Berlin Heidelberg.

Norman, D. A. (2010). Natural user interfaces are not natural. interactions, 17(3), 6-10.

O'hara, K., Harper, R., Mentis, H., Sellen, A., & Taylor, A. (2013). On the naturalness of touchless: putting the “interaction” back into NUI. ACM Transactions on Computer-Human Interaction (TOCHI), 20(1), 5.

O’Malley, M. K., & Gupta, A. (2008). Haptic interfaces. *HCI Beyond the GUI: Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces”, Edited by Kortum, P., Elsevier Inc., USA*, *1*.

Obrist, M., Bernhaupt, R., & Tscheligi, M. (2008). Interactive TV for the home: An ethnographic study on users' requirements and experiences. Intl. Journal of Human–Computer Interaction, 24(2), 174-196.

Olsson, T., Kärkkäinen, T., Lagerstam, E., & Ventä-Olkkonen, L. (2012). User evaluation of mobile augmented reality scenarios. Journal of Ambient Intelligence and Smart Environments, 4(1), 29-47.

Orso, V., Spagnolli, A., Gamberini, L., Ibañez, F., & Fabregat, M. E. (2015, September). Involving Older Adults in Designing Interactive Technology: The Case of SeniorChannel. In Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter (pp. 102-109). ACM.

Oviatt, S. (2003). Multimodal interfaces. The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications, 14, 286-304.

Oviatt, S., & Cohen, P. (2000). Perceptual user interfaces: multimodal interfaces that process what comes naturally. Communications of the ACM, 43(3), 45-53.

Ozok, A. A. (2009). Survey design and implementation in HCI. Human-Computer Interaction: Development Process, 253.

Pacchierotti, C., Tirmizi, A., & Prattichizzo, D. (2014). Improving transparency in teleoperation by means of cutaneous tactile force feedback. *ACM Transactions on Applied Perception (TAP)*, *11*(1), 4.

Panchanathan, S., McDaniel, T., & Balasubramanian, V. (2012, November). Person-centered accessible technologies: Improved usability and adaptation through inspirations from disability research. In *Proceedings of the 2012 ACM workshop on User experience in e-learning and augmented technologies in education* (pp. 1-6). ACM.

Park, T., Lee, J., Hwang, I., Yoo, C., Nachman, L., & Song, J. (2011, November). E-gesture: a collaborative architecture for energy-efficient gesture recognition with hand-worn sensor and mobile devices. In Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (pp. 260-273). ACM.

Perry, M., Dowdall, A., Lines, L., & Hone, K. (2004). Multimodal and ubiquitous computing systems: Supporting independent-living older users. Information Technology in Biomedicine, IEEE Transactions on, 8(3), 258-270.

Pielot, M., Poppinga, B., Heuten, W., & Boll, S. (2011). A tactile compass for eyes-free pedestrian navigation. In Human-Computer Interaction–INTERACT 2011 (pp. 640-656). Springer Berlin Heidelberg.

Prante, T., Röcker, C., Streitz, N., Stenzel, R., Magerkurth, C., Van Alphen, D., & Plewe, D. (2003, October). Hello. wall–beyond ambient displays. In Adjunct Proceedings of Ubicomp (pp. 277-278).

Prasad, M., Taele, P., Goldberg, D., & Hammond, T. A. (2014, April). Haptimoto: Turn-by-turn haptic route guidance interface for motorcyclists. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3597-3606). ACM.

Profita, H. P., Clawson, J., Gilliland, S., Zeagler, C., Starner, T., Budd, J., & Do, E. Y. L. (2013, September). Don't mind me touching my wrist: a case study of interacting with on-body technology in public. In Proceedings of the 2013 International Symposium on Wearable Computers (pp. 89-96). ACM.

Rekimoto, J. (2002, April). SmartSkin: an infrastructure for freehand manipulation on interactive surfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 113-120). ACM.

Rhodes, B. J., Minar, N., & Weaver, J. (1999, October). Wearable computing meets ubiquitous computing: Reaping the best of both worlds. In Wearable Computers, 1999. Digest of Papers. The Third International Symposium on (pp. 141-149). IEEE.

Rubin, J., & Chisnell, D. (2008). Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley & Sons.

Ruiz, J., Li, Y., & Lank, E. (2011, May). User-defined motion gestures for mobile interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 197-206). ACM.

Rümelin, S., Rukzio, E., & Hardy, R. (2011, October). NaviRadar: a novel tactile information display for pedestrian navigation. In Proceedings of the 24th annual ACM symposium on User interface software and technology (pp. 293-302). ACM.

S. Robinson, P. Eslambolchilar, and M. Jones. Evaluating haptics for information discovery while walking. In Proc. British HCI, pages 93–102, 2009

S. Wei, G. Ren, and E. O'Neill. Haptic and audio displays for augmented reality tourism applications. In Proc. HAPTICS, pages 485–488, 2014.

Schmalstieg, D., Reitmayr, G. (2007). The World as a User Interface: Augmented Reality for Ubiquitous Computing. In Gartner, G., Cartwright, W., Peterson, M. Eds. Location Based Services and TeleCartography (pp. 369-391) Springer Berlin Heidelberg.

Serrano, M., Ens, B. M., & Irani, P. P. (2014, April). Exploring the use of hand-to-face input for interacting with head-worn displays. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 3181-3190). ACM.

Shoemaker, G., Tsukitani, T., Kitamura, Y., & Booth, K. S. (2010, October). Body-centric interaction techniques for very large wall displays. In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (pp. 463-472). ACM.

Srikulwong, M., & O'Neill, E. (2011, May). A comparative study of tactile representation techniques for landmarks on a wearable device. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2029-2038). ACM.

Stern, H. I., Wachs, J. P., & Edan, Y. (2008). Designing hand gesture vocabularies for natural interaction by combining psycho-physiological and recognition factors. International Journal of Semantic Computing, 2(01), 137-160.

Stevenson, G., Riener, A., & Ferscha, A. (2013). "TactiGlove"–A Guidance System to Effectively Find Hidden Spots in 3D Space. In Mobile Computing, Applications, and Services (pp. 80-99). Springer Berlin Heidelberg.

Suero Montero, C. A., Alexander, J. M., Marshall, M. T., & Subramanian, S. (2010). Would You Do That? – Understanding Social Acceptance of Gestural Interfaces. In MobileHCI 2010, Lisbon, Portugal. (pp. 275 - 278). ACM. 10.1145/1851600.1851647

Tsukada, K., & Yasumura, M. (2004). Activebelt: Belt-type wearable tactile display for directional navigation. In UbiComp 2004: Ubiquitous Computing (pp. 384-399). Springer Berlin Heidelberg.

Turunen, M., Hakulinen, J., Melto, A., Heimonen, T., Laivo, T., & Hella, J. (2009, January). SUXES-user experience evaluation method for spoken and multimodal interaction. In INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association, Brighton, United Kingdom, September 6-10, 2009 (pp. 2567-2570).

Van Erp, J. B., Van Veen, H. A., Jansen, C., & Dobbins, T. (2005). Waypoint navigation with a vibrotactile waist belt. ACM Transactions on Applied Perception (TAP), 2(2), 106-117.

Voida, S., Podlaseck, M., Kjeldsen, R., & Pinhanez, C. (2005, April). A study on the manipulation of 2D objects in a projector/camera-based augmented reality environment. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 611-620). ACM.

Wachs, J. P., Kölsch, M., Stern, H., & Edan, Y. (2011). Vision-based hand-gesture applications. Communications of the ACM, 54(2), 60-71.

Wagner, D., Pintaric, T., Ledermann, F., & Schmalstieg, D. (2005). Towards massively multi-user augmented reality on handheld devices (pp. 208-219). Springer Berlin Heidelberg.

Wang, G. G. (2002). Definition and review of virtual prototyping. Journal of Computing and Information Science in engineering, 2(3), 232-236.

Watanabe, J., & Ando, H. (2010). Pace-sync shoes: intuitive walking-pace guidance based on cyclic vibro-tactile stimulation for the foot. *Virtual reality*, *14*(3), 213-219.

Weber, B., Schätzle, S., Hulin, T., Preusche, C., & Deml, B. (2011, June). Evaluation of a vibrotactile feedback device for spatial guidance. In World Haptics Conference (WHC), 2011 IEEE (pp. 349-354). IEEE.

Wei, S., Ren, G., & O'Neill, E. (2014, February). Haptic and audio displays for augmented reality tourism applications. In *Haptics Symposium (HAPTICS), 2014 IEEE* (pp. 485-488). IEEE.

Weiser, M. (1991, September). The computer for the 21st century. Scientific American, 265(3), 933–940.

Weiser, M., & Brown, J. S. (1997). The coming age of calm technology. In Beyond calculation (pp. 75-85). Springer New York.

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009, April). User-defined gestures for surface computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1083-1092). ACM.