



Ubiquitous interaction: mobile, multi-display and freeform interfaces

Marcos Serrano

► To cite this version:

Marcos Serrano. Ubiquitous interaction: mobile, multi-display and freeform interfaces. Human-Computer Interaction [cs.HC]. Université Paul Sabatier - Toulouse 3, 2018. tel-01959335v2

HAL Id: tel-01959335

<https://hal.science/tel-01959335v2>

Submitted on 7 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PAUL SABATIER - TOULOUSE 3

HABILITATION À DIRIGER DES RECHERCHES

Présentée et soutenue publiquement par

Marcos Serrano

le 15 Novembre 2018

Spécialité: Informatique - Interaction Homme-Machine

Ubiquitous interaction: mobile, multi-display and freeform interfaces

JURY

Caroline APPERT	CNRS - LRI	Présidente
Michel BEAUDOUIN-LAFON	Université Paris-Sud	Rapporteur
Stephen BREWSTER	University of Glasgow	Rapporteur
Andy COCKBURN	University of Canterbury	Rapporteur
Emmanuel DUBOIS	Université Toulouse 3	Examineur
Antonio KRUGER	Saarland University	Examineur

Habilitation à Diriger des Recherches préparée au sein de l'Institut de Recherche
en Informatique de Toulouse

Acknowledgements

First of all, I want to thank Michel Beaudouin-Lafon, Stephen Brewster and Andy Cockburn for accepting to review this manuscript. I also want to thank the members of my jury, Caroline Appert and Antonio Kruger.

Special thanks to Emmanuel Dubois for his help in preparing this HDR, and more generally for his support since I arrived to the University of Toulouse. It has been a fruitful collaboration so far.

The work presented in this manuscript results from multiple collaborations and I want to acknowledge my co-authors involvement, especially Pourang Irani, Eric Lecolinet, Yves Guiard, Christophe Jouffrais, Anne Roudaut, Barrett Ens, Xing-Dong Yang, Louis-Pierre Bergé, Andrés Lucero and Mathieu Raynal.

Some of the presented contributions have been carried by three PhD students that I co-supervised at the University of Toulouse, Sandra Bardot, Housseem Saidi and Gary Perelman. It has been a pleasure working together and I believe I have learned as much as them during this journey.

Thanks to the ELIPSE group members, with whom I have shared many discussions that surely have an impact on the work presented here, especially Bernard Oriola, Marc Macé and Antonio Serpa.

Finally, thank you Anne-Cécile, Félix and Carmen for filling my life with love and joy.

Abstract

My research activities have focused on solving some of the main challenges of three types of ubiquitous displays: mobile, multi-display, and freeform interfaces.

By 2017, the number of smartphone users in the world was beyond 2 billions, i.e. one third of the world population. Not only the number of smartphones has dramatically increased, but also their computing capabilities and hence the number of available applications and complex data to manipulate. This raises numerous challenges on how to easily and rapidly interact with such growing quantity of data on mobile platforms. We present four major contributions in this context: we created a novel type of gestures called Bezel-Tap Gestures to solve the problem of rapidly launching commands from sleep mode; we proposed using on-body gestures on the face, i.e. hand-to-face input, to facilitate navigating spatial data on head-worn displays; we explored mobile true-3D displays to facilitate mobile interaction with volumetric data; and finally we studied using the smartwatch output capabilities to facilitate non-visual exploration of spatial data for visually-impaired people.

Multi-display environments (MDEs) are common on desktop environments and more and more common on professional and public environments. However they also bring new challenges due to the complexity of the overall system, which can be composed of multiple and heterogeneous devices, arranged in dynamic spatial topologies. We addressed two major challenges in MDEs: the need for fluid interactions in MDEs and complex data exploration using multiple monitors. To tackle the first challenge we adopted two approaches: either using an existing device, in our case a head-worn display interface that we named Gluey, or creating a novel dedicated device, namely TDome. To tackle the second challenge, i.e. facilitate the exploration of complex data on MDEs, we studied the use of around-device gestures to manipulate 3D data on public display and extended the overview+detail interface paradigm with multiple detailed views to explore multiple regions of the data simultaneously.

Finally, to fulfill the adoption of pervasive displays, we need to facilitate a seamless integration of displays into existing environments, ranging from in-vehicle and wearable displays to public displays. Traditional rectangular displays are not well-suited for these applications, as they can not be easily integrated. Emerging technologies allow for the creation of non-rectangular displays with unlimited constraints in shapes. With this eminent adoption comes the urgent challenge of rethinking the way we present content on non-rectangular displays. Our approach was threefold: first, we carried focus groups to gather both concrete usage scenarios and display shapes; then, we studied text content only, being the fundamental brick of any UI; finally, we explored more complex content layouts combining text and icons.

To address the previously presented research challenges, we had to face the methodological challenges of designing, prototyping and evaluating interaction. Traditional design and evaluation methods quite often failed at properly answering the previous research challenges. As a result, every new project came up with the additional challenge of rethinking our approach and research methods at multiple levels: design, prototyping and evaluation. For this reason, I deem of importance to include a chapter on research methods.

Contents

Acknowledgements	i
Abstract	iii
Contents	v
1 Introduction	1
1.1 Towards ubiquitous interaction	1
1.1.1 The advent of mobile computing	1
1.1.2 Multi-display environments	2
1.1.3 Displays everywhere	2
1.2 Our contributions	2
1.2.1 Mobile interactions with complex data	3
1.2.2 Multi-display tools and interfaces	4
1.2.3 FreeForm interfaces	5
1.3 Methodological challenges	6
1.4 Manuscript overview	6
2 Mobile interactions with large data spaces	9
2.1 Research challenges	9
2.1.1 Gestural interaction as mobile input	9
2.1.2 Offscreen and eyes-free mobile output	10
2.1.3 Summary	11
2.2 Command selection from sleep mode on tablets through Bezel-Tap gestures	11
2.2.1 Bezel-Tap Gestures	12
2.2.2 Bezel-Tap Menus	13
2.2.3 Summary	15
2.3 Exploring spatial data on head-worn displays using hand-to-face input	15
2.3.1 Hand-to-face gestures	15
2.3.2 Spatial navigation with hand-to-face input	16
2.3.3 Social acceptability of hand-to-face input	17
2.3.4 Summary	18
2.4 Mobile True-3D Displays	18
2.4.1 Display properties	18
2.4.2 Spatial configuration	19
2.4.3 Direct touch interaction	19
2.4.4 Summary	20
2.5 Mobile map exploration for visually impaired people	20
2.5.1 Interaction principles	20
2.5.2 Non-visual map exploration techniques	21
2.5.3 Experimental results	22
2.5.4 Summary	22
2.6 Chapter conclusion	23

3	Multi-Display tools and interfaces	25
3.1	Research challenges	25
3.1.1	Fluid interaction in MDEs	25
3.1.2	Exploring complex data on multiple displays	26
3.1.3	Summary	28
3.2	Gluey: fluid interaction in MDEs using a head-worn display	28
3.2.1	Usage Scenario	28
3.2.2	Gluey key components	29
3.2.3	Gluey implementation	30
3.2.4	Summary	31
3.3	TDome: fluid interaction in MDEs using a dedicated multi-DoF device	31
3.3.1	Usage Scenario	31
3.3.2	TDome key components	32
3.3.3	Evaluation	33
3.3.4	Summary	34
3.4	Interaction with Overview+Detail Interfaces on 3D Public Displays	34
3.4.1	Interaction techniques for translating the detail view in 3D	35
3.4.2	Experimental results	35
3.4.3	Summary	36
3.5	Splitting Detailed Views in Overview+Detail Interfaces	36
3.5.1	Interface design	36
3.5.2	Experimental results	38
3.5.3	Summary	38
3.6	Chapter conclusion	39
4	Towards Displays Everywhere: FreeForm Interfaces	41
4.1	Research challenges	41
4.2	Collecting usage scenarios	43
4.2.1	Data collection	43
4.2.2	Display shape properties	43
4.3	Displaying text on freeform displays	44
4.3.1	Mapping framework	44
4.3.2	Quantitative studies on reading text	45
4.4	Laying out content on freeform displays	46
4.4.1	Elicitation study	46
4.4.2	Composition principles	47
4.4.3	Experimental studies	48
4.5	Chapter conclusion	50
5	Research methods for ubiquitous interaction	53
5.1	Design methods	53
5.1.1	In-the-wild observations	53
5.1.2	Research through design by device composition (*contribution)	55
5.1.3	Elicitation studies: exploiting similarity over agreement (*contribution)	55
5.1.4	Design probes	57
5.2	Prototyping methods	58
5.2.1	Rapid prototyping tools (*contribution)	58
5.2.2	Proof-of-concept prototypes	58
5.3	Evaluation methods	62
5.3.1	Component-based Wizard-of-Oz (*contribution)	62
5.3.2	Pairwise comparisons	63
5.3.3	Conclusion	64
6	Perspectives	65

<i>Contents</i>	vii
6.1 From digital to physical ubiquitous displays	65
6.2 Immersive data analytics to explore spatio-temporal data	67
6.3 Pervasive freeform interfaces	68
Bibliography	69
Selected publications	79

Chapter 1

Introduction

Chapter content

1.1	Towards ubiquitous interaction	1
1.2	Our contributions	2
1.3	Methodological challenges	6
1.4	Manuscript overview	6

1.1 Towards ubiquitous interaction

Ubiquitous computing was a term coined by Mark Weiser around 1988 and presented in his 1991 paper "The computer for the 21st Century" [Weiser, 1991]. In this paper, Weiser exposes his vision that personal computers should evolve to make computing "an integral, invisible part of people's lives". While we are not yet there, recent advances in computing are making this vision come true.

1.1.1 The advent of mobile computing

One of the major evolutions of personal computing in the last decade has been the one leading **from desktop computers to mobile devices**. I started my PhD in 2007 with the goal to develop an approach to prototype the combination of multiple interaction modalities (such as touch, gestures, etc.), i.e. what is called Multimodal Interaction. It soon became evident that multimodal interaction was not limited to desktop environments: that same year the arrival of two major innovations, iPhone and Android, generalized the use of multitouch and gestural mobile interaction. By 2017, the number of smartphone users in the world was beyond 2 billion, i.e. one third of the world population. Not only the number of smartphones has dramatically increased, but also their computing capabilities and hence the number of available applications. A 2012 survey showed that Android users have an average of 177 applications on their phones, using 73 of them occasionally and 32 very often [Shin *et al.*, 2012]. This raises numerous challenges on how to easily and rapidly interact with such growing quantity of data.

Interestingly, Weiser stated that "ubiquitous computing [...] does not mean just computers that can be carried to the beach, jungle or airport." [Weiser, 1991]. In other words, the ubiquitous computing vision can not be limited to the use of smartphones in a mobile context. The next step in the evolution towards ubiquitous computing is the emergence of **wearable devices**, namely smartwatches, head-mounted displays or jewellery. These devices are slowly becoming available for widespread daily use through lighter form factors.

They leverage mobility and always-available interaction, which are necessary to put in effect the concept of ubiquitous interaction. However, despite their interesting form factors, wearables devices have limited interaction capabilities, which makes it difficult using them for non trivial interactions, such as panning and zooming on a map.

1.1.2 Multi-display environments

Multi-display environments (MDEs) are common on desktop environments, where usually two or three displays are positioned alongside to extend the overall display space. Beyond this now ordinary setup, MDEs are also common on some professional environments, such as control rooms. More recently, multi-displays environments have become available on public environments, such as train stations, airports or malls, through the use of chained displays [Ten Koppel *et al.*, 2012].

When combined, the mobile and multi-display phenomenon result in even more complex environments, which have been called Mobile Multi-Display Environments (MMDEs) [Cauchard *et al.*, 2011]. These environments provide unique opportunities, as mobile devices can be employed to interact with these larger displays. However they also bring new challenges due to the complexity of the overall system, which can be composed of multiple and heterogeneous devices, arranged in dynamic spatial topologies (i.e. some devices/displays can move while others remain static).

1.1.3 Displays everywhere

The pervasive need to access and make decisions based on data is ushering another evolution in ubiquitous computing: displaying interactive dynamic content where and when needed, a vision termed as displays everywhere [Pinhanez, 2001]. This vision covers a broad set of use cases, as these **pervasive displays** can be leveraged to show energy consumption data in-situ in smart environments; facilitate collaborative pedagogy in university classes; leverage interactive presentations in museums, showrooms or shop fronts; or facilitate planning and decision-making during professional meetings.



Figure 1.1: The three types of ubiquitous interfaces that we explored in our research projects.

1.2 Our contributions

My research activities have focused on solving some of the main challenges of the three previous types of ubiquitous interfaces (Figure 1.1): mobile, multi-display, and pervasive freeform interfaces.

1.2.1 Mobile interactions with complex data

Our work on mobile interaction addressed both the challenges of how to extend the limited input capabilities of mobile devices, particularly concerning gestural interaction, and how to access information beyond the display space, i.e. offscreen or eyes-free output (Figure 1.2).

Novel gestures for mobile input

During my 1-year post-doctoral fellowship at Telecom ParisTech (Paris, France), we addressed the issue of how to facilitate micro-interactions with a very large number of commands when the mobile device is asleep. Since mobile devices constantly switch to sleep mode to save energy, interaction is hampered by the need to reactivate them whenever they have gone to sleep, typically by pressing a physical button and sliding a widget on the screen. We addressed this problem through the development of a novel interaction technique, Bezel Tap [Serrano *et al.*, 2013], allowing micro-interactions on mobile devices. Not only can the technique serve to open an application and launch a favorite command rapidly, it can also wake the device if asleep.

During my 1-year post-doctoral fellowship at the University of Manitoba (Winnipeg, Canada), we addressed the issue of how to access commands and facilitate 2D interaction on head-worn displays. Currently, HWDs provide onboard microphones and small capacitive sensors for user input. Voice recognition is useful for command-based tasks such as for search queries but has limited use in certain settings (i.e. noisy environments). The capacitive surface on the temple of HWDs presents a viable on-device method for input, but it has limited input space. We proposed hand-to-face input as a novel, alternative method for interacting with HWDs [Serrano *et al.*, 2014a]. We define hand-to-face input as any gesture that involves touching, rubbing, scratching or caressing the face.

Offscreen and eyes-free mobile output

During my post-doctoral year at the University of Manitoba we also explored the vision of mobile volumetric (or true-3D) displays, i.e. displaying volumetric 3D content in mid-air around a mobile phone. This can be accomplished through different means: stereoscopic displays with head-tracking, optical illusions, moving parts or augmented reality glasses. The challenge is that we possess limited knowledge on both the numerous constraints imposed on viewing and interacting with mobile true-3D interfaces and the usage scenarios suitable for such displays. To this end, we studied the various factors that can potentially influence the effective deployment of true-3D on mobile devices in an emulated environment [Serrano *et al.*, 2014b].

While mobile content is difficult to access for sighted people, they are inherently inaccessible to visually impaired (VI) users. Visually impaired users wanting to use mobile or pervasive displays have to deal with the lack of any haptic cue on the touchscreen and the lack of visual feedback. However, previous work has shown the interest for VI people to use wearables [Ye *et al.*, 2014] and particularly smartwatches: they are small, easily accessible and unobtrusive to wear, and can improve information access and social interactions. Our work in this area has the goal of overcoming the inherent limitations of digital content by providing a visually impaired user with a way to independently explore and/or reconstruct data. To this end, we explored how smartwatches can be used to access digital geospatial data [Bardot *et al.*, 2016, Bardot *et al.*, 2017].

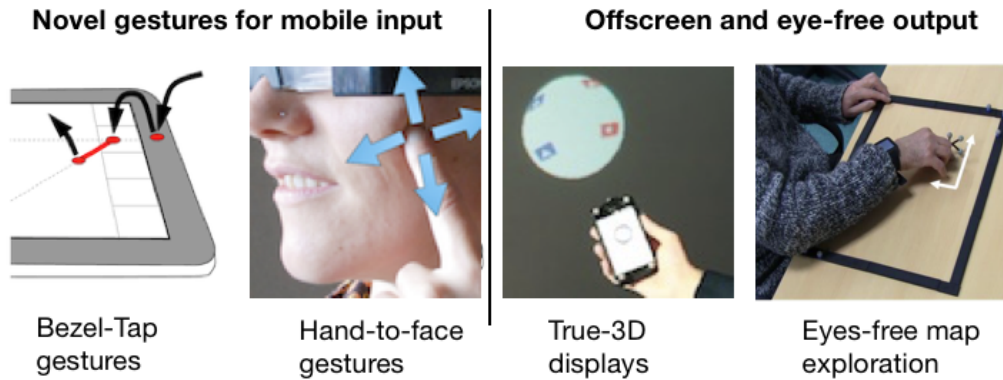


Figure 1.2: We proposed two types of contributions to mobile interaction: novel gestures for mobile input (left), and data exploration techniques beyond the display space, i.e. offscreen or eyes-free output (right).

1.2.2 Multi-display tools and interfaces

Multi-displays environments (MDEs) offer numerous advantages for organizing information across displays, for enhancing individual and group work, and for extending the interaction space. Our contributions to interacting with MDEs can be classified along two main axes (Figure 1.3): novel tools for fluid interaction in MDEs (1) and data exploration on MDEs using overview + detail interfaces (2).

Tools for fluid interaction in MDEs

The incidental emergence of MDEs has resulted in a device vacuum: no device has been specifically implemented for optimizing interactions in such spaces. To this end, we developed two systems to facilitate the overall interaction on MDEs. First, we introduced Gluey [Serrano *et al.*, 2015], a head-worn software interface that acts as a glue for bonding information across the various displays in a distributed environment. Second, we presented a novel touch-enabled device, TDome [Saidi *et al.*, 2017], designed to facilitate interactions and address a range of tasks in MDEs. TDome is the combination of a touchscreen, a dome-like Mouse [Perelman *et al.*, 2015] providing 6 DOF, and a camera that can sense the environment. TDome thus inherits properties of other existing mouse-like devices but includes many novel features to tackle the needs of common MDE tasks.

Data exploration on MDEs using overview + detail interfaces

We also worked on overview + detail interfaces, which are very common on multi-display environments: in such setups, the mobile display shows a detailed view of a section of the large display (the overview). While this technique is well known for 2D overviews with one detail view, its use in 3D environments or with multiple detail views had not yet been fully explored. To this end, first we focused on how to move the detail view (displayed on the smartphone) on a 3D environment [Bergé *et al.*, 2014]. In particular, we explored the use of mid-air gestures for controlling the translation of the Detail view, by moving the hand around the mobile device or by moving the phone itself. Second, we studied the use of different number of detailed views to interact with very large graphs in an overview + detail setting composed of a large screen and a mobile device [Saidi *et al.*, 2016]. Our

goal was to find the optimal number of detailed views to perform tasks involving multiple graph nodes

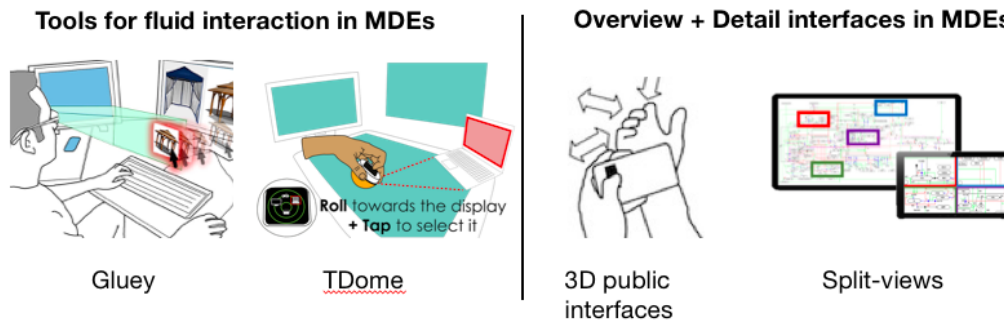


Figure 1.3: Our contributions to interacting with MDEs can be classified along two main axis: novel tools for fluid interaction in MDEs (left) and data exploration on MDEs using overview + detail interfaces (right).

1.2.3 FreeForm interfaces

To fulfill the adoption of pervasive displays, we need to facilitate a seamless integration of displays into existing environments, ranging from in-vehicle and wearable displays to public displays. Traditional rectangular displays are not well-suited for these applications, as they can not be easily integrated. Emerging technologies allow for the creation of nonrectangular displays with unlimited constraints in shapes. For instance, a single non-rectangular display can replace current instrument panels on car dashboards, in-between other on-board instruments; non-rectangular displays will also facilitate inserting displays on nonrectangular objects, furniture and/or urban architecture. In the context of mobile computing, non-rectangular displays can adopt shapes which will better fit wearable devices or replicate existing jewellery. With this eminent adoption comes the urgent challenge of rethinking the way we present content on non-rectangular displays (Figure 1.4).



Figure 1.4: Examples of 2D non-rectangular interfaces.

I started exploring this research challenge in 2016 in collaboration with Pourang Irani (University of Manitoba, Canada) and Anne Roudaut (University of Bristol, UK). We have published two preliminary explorations for displaying text and laying out content on non-rectangular interfaces [Serrano *et al.*, 2016, Serrano *et al.*, 2017]. This initial exploration suggests that the research question is timely, promising and interesting. The first work revealed

that text legibility is affected by shape factors such as text layout and alignment. The second work showed that content layout properties (grid type, type of symmetry, etc.) affect user perception differently according to the display shape. Overall, these studies reveal that different approaches apply to different content (such as text or icons layout).

1.3 Methodological challenges

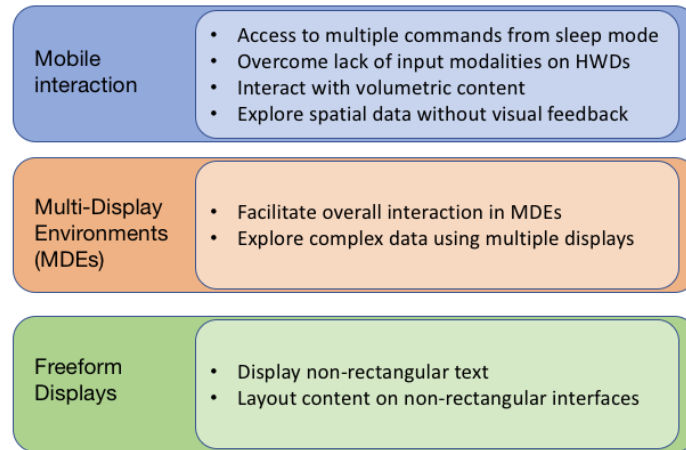


Figure 1.5: Research challenges we addressed for mobile, multi-display and freeform interfaces.

To address the previously presented research challenges, summarized in Figure 1.5, we had to face the methodological challenges of designing, prototyping and evaluating interaction. Traditional design and evaluation methods quite often failed at properly answering the previous research challenges. As a result, every new project came up with the additional challenge of rethinking our approach and research methods at multiple levels: design, prototyping and evaluation. For this reason, I deem of importance to include a chapter on research methods for ubiquitous interfaces (chapter 5). While the ultimate goal of our work was not to address these methodological challenges, we proposed novel methods or adapted existing ones.

Concerning design methods for HCI, we detail our experiences carrying in-the-wild observations [Lucero *et al.*, 2016], user-elicitation studies [Serrano *et al.*, 2014a] and design probes [Serrano *et al.*, 2017]. We also proposed a novel design process to combine existing interaction devices, DECO [Perelman *et al.*, 2016]. Prototyping advanced interactions can be difficult: we report our solutions for developing proof-of-concept prototypes for ubiquitous interaction, such as on-body input [Serrano *et al.*, 2014a] or mobile holographic (or true-3D) displays [Serrano *et al.*, 2014b]. Finally, we discuss evaluation methods with a focus on two key aspects: wizard-of-oz studies [Serrano and Nigay, 2010] and pairwise comparison experiments [Serrano *et al.*, 2017].

1.4 Manuscript overview

This manuscript is organized as follows. One chapter is dedicated to each one of the three types of ubiquitous interfaces explored in our research work: Mobile interfaces (Chapter 2), Multi-Display environments (Chapter 3) and Freeform interfaces (Chapter 4). These chapters start with a section detailing the research challenges addressed in our work, followed

by one section for each major contribution in the domain. Then the Chapter 5 presents our work on the research methods in three main sections: Design, Prototyping and Evaluation methods. Finally, Chapter 6 presents the conclusion and perspectives to our work. Figure 1.6 sums up the organization of this manuscript.

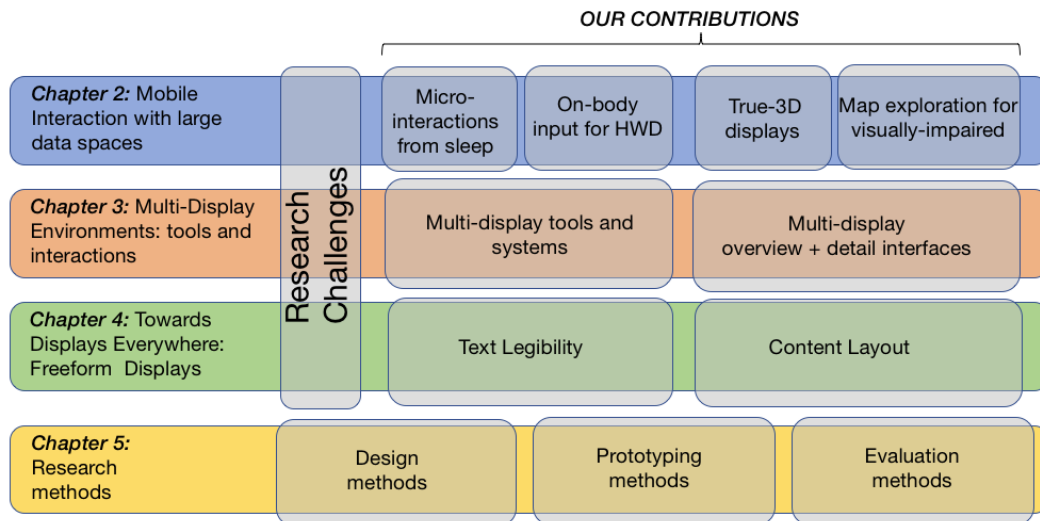


Figure 1.6: Outline of the manuscript organization.

Chapter 2

Mobile interactions with large data spaces

Chapter content

2.1	Research challenges	9
2.2	Command selection from sleep mode on tablets through Bezel-Tap gestures	11
2.3	Exploring spatial data on head-worn displays using hand-to-face input .	15
2.4	Mobile True-3D Displays	18
2.5	Mobile map exploration for visually impaired people	20
2.6	Chapter conclusion	23

With the evolution in mobile computing technologies, smartphone, tablets and wearable devices are more powerful than ever. They can now easily process large volumes of data, such as multiple applications and commands, geospatial information or volumetric data. The bottleneck to fully exploit such data is the interaction capabilities of mobile devices, both in terms of input and output modalities. In this chapter I present our work on facilitating mobile interaction with large data spaces.

2.1 Research challenges

As said in the introduction, our work on mobile interaction addressed both the challenges of how to extend the limited input capabilities of mobile devices, particularly concerning gestural interaction, and how to access information beyond the display space, i.e. offscreen or non-visual output.

2.1.1 Gestural interaction as mobile input

Our work on extending current input on mobile devices concentrated on two main research questions that we detail below: how to rapidly access multiple commands anytime, i.e. from the device sleep mode, and how to overcome the lack of input modalities to explore spatial data on head-worn displays.

Rapid access to numerous commands from sleep mode

A limitation of mobile devices is that they provide little support for quick commands. Micro-interactions, which rest on unconditional availability and fast access [Ashbrook, 2010], are especially desirable for frequent actions such as, for instance, controlling a media player, checking emails and SMS, calling friends and family, getting the local map or the weather forecast, and more basic actions like copying, pasting, and application switching. This is specially difficult due to the large set of mobile applications.

A 2012 survey showed that Android users have an average of 177 applications on their phones, using 73 of them occasionally and 32 very often [Shin *et al.*, 2012].

Since mobile devices constantly switch to sleep mode to save energy, interaction is hampered by the need to reactivate them whenever they have gone to sleep, typically by pressing a physical button and sliding a widget on the screen. This problem is exacerbated when mobile devices are used to control multimedia devices (TV, set-top box, etc.) and home equipment (home automation, domotics systems). In this scenario remote commands expand the large set of mobile applications. The challenge, then, is to allow always available and rapid access to a relatively large number of commands.

Gestural interaction provides an efficient means for activating commands rapidly. Marking menus [Kurtenbach and Buxton, 1991] are a well-known technique relying on 2D gestures. They have inspired many HCI studies, some of them dedicated to mobile devices [Kin *et al.*, 2011]. One major merit of these techniques is to make it easy to discover and learn the gestures: all commands are visible in novice mode and gestures are learned incidentally from repetitive use. These techniques are well suited for triggering a few dedicated commands at the application level, but can not be used as a global shortcut mechanism, i.e. without interfering with application shortcuts. Finally, gestural interaction does not work with the device in sleep mode, and unfortunately that is often the case due to the high power consumption of the touchscreen. In summary, there is a need for a novel type of gesture interaction on handheld devices that can be employed whether the device is alive or in sleep mode.

Exploring spatial data on head-worn displays

Head-mounted devices are becoming widely available for daily use through lighter form factors and with transparent displays. We refer to these modern accessories as head-worn displays (HWDs). As consumers get affordable access to HWDs, ways in which they interact with content on such devices is being actively investigated. These devices are particularly useful for way-finding, and therefore it is important to facilitate map-related tasks, such as pan and zoom. However, interacting with such spatial information is particularly difficult given the limited input capabilities of HWDs.

Currently, HWDs provide onboard microphones and small capacitive sensors for user input. Voice recognition is useful for command-based tasks such as for search queries but has limitations in certain settings (i.e. noisy environments). The capacitive surface on the temple of HWDs presents a viable on-device method for input, but it has limited input space. Other less self-contained options, such as a wearable device or an auxiliary smartphone, can also allow for input. However, these require carrying them and may be occluded by the HWD content.

Mid-air and on-body gestural interaction can overcome the above limitations. However, mid-air input [Bailly *et al.*, 2012] suffers from the lack of tactile feedback. On-body interaction offers an input surface, the human skin, with the advantage of leveraging human proprioception as an extra feedback mechanism. This can overcome some of the limitations with mid-air interactions. A large body of work has considered appropriating the body as an interaction surface. Much of this prior work has considered coupling on-body input with on-body projection using wearable pico-projectors [Harrison *et al.*, 2010]. An open question then is which body area could better fit the specific needs of interacting with HWDs.

2.1.2 Offscreen and eyes-free mobile output

The second main challenge, i.e. exploring information beyond the display space, evolved around two different questions detailed below: how to interact with volumetric content around mobile devices (offscreen space); and how to facilitate the non-visual exploration of spatial data for visually-impaired people.

Interacting with volumetric content

Benefiting from 3D on mobile devices is pertinent, beyond video games, for mobile scenarios such as 3D interior design or 3D map exploration. Mobile devices (smartphones and gaming consoles) incorporate at present auto-stereoscopic displays as a first step toward achieving true-3D content. True-3D [Kimura *et al.*, 2011], i.e. displaying volumetric 3D content in mid-air, can be accomplished through different means: stereoscopic displays with head-tracking [Nii *et al.*, 2012], optical illusions [Hilliges *et al.*, 2012], moving parts or augmented reality glasses.

Integrating true-3D on mobile devices, apart from facing hardware challenges, presents a number of unresolved human factors questions concerning its use. We possess limited knowledge on both the numerous constraints imposed on viewing and interacting with mobile true-3D interfaces and the usage scenarios suitable for such displays. These include knowing about the ideal angular position, size and distance of the volumetric projection, relative to the mobile device, the projection limits on visual search and direct interaction and how to coordinate the mobile and true-3D views. Answers to such questions will equip manufacturers and designers with tools to begin exploring a range of technologies for true-3D input on mobile devices.

Mobile exploration of spatial data without visual feedback for visually-impaired people

Visually impaired (VI) people need regular access to geospatial information during education but also during everyday life. Previous work has already explored how VI people can access virtual maps. Besides static solutions based on using the keyboard or haptic devices (e.g. mouse or phantom), the main approach for exploring digital maps in mobile context is to use a touchscreen (on a smartphone or a tablet). For instance, [Su *et al.*, 2010] investigated the sonification of simple navigation maps on handheld touch screens. The evaluation showed that users of the proposed prototype recognized a few shapes in a matching to sample protocol, but can also develop a rough understanding of an indoor floor plan. However, an obvious drawback of handheld devices is that they only provide a limited surface for exploration, and hence require recurrent panning and zooming operations that are very difficult to perform by VI users. The challenge then is how to provide VI people with mobile techniques to explore spatial data that do not rely on the use of a touchscreen.

2.1.3 Summary

In this chapter we present our solutions to address the aforementioned research questions (see Figure 2.1): to rapidly launch commands from sleep mode, we created a novel type of gestures called Bezel-Tap Gestures (section 2.2); to facilitate navigating spatial data on head-worn displays, we proposed using on-body gestures on the face, i.e. hand-to-face input (section 2.3); to facilitate the visualization of volumetric data, we explored mobile true-3D displays (section 2.4); finally, to facilitate non-visual exploration of spatial data for visually-impaired people, we studied the smartwatch output capabilities (section 2.5).

2.2 Command selection from sleep mode on tablets through Bezel-Tap gestures

During my post-doctoral fellowship at Telecom Paristech, I worked with Eric Lecolinet and Yves Guiard on solving the challenge of enabling micro-interactions from sleep mode on mobile devices. We proposed Bezel-Tap Gestures, a novel interaction technique that not only serves to open an application and launch a favorite command rapidly, but can also wake the device if asleep. This work led to one publication at CHI 2013 [Serrano *et al.*, 2013].

Challenges	Gestural input		Offscreen output	
Research question	Launch commands from sleep mode	Navigate spatial data	Interact with mobile true-3D data	Perform non-visual exploration
Data	Multiple commands	Spatial data	Volumetric data	Spatial data
Devices	Smartphone/Tablets	Wearables: HMD	Smartphone/Tablets	Wearables: watch
Chapter section:	2.2	2.3	2.4	2.5
Contributions	Bezel-Tap gestures	Hand-to-face input	Evaluation of mobile true-3D	Watch-based map exploration

Figure 2.1: Research challenges and our contributions to the domain of mobile interaction with large data spaces.

2.2.1 Bezel-Tap Gestures

In its basic form a Bezel-Tap gesture involves two successive events recorded in different input channels, a tap on the bezel immediately followed by a tap (or a slide) on the screen (Figure 2.2). These gestures were primarily designed for tablets, which have quite large bezels, and do not interfere with common touch-based interaction techniques. Most importantly, the technique makes it possible to both wake up the device and activate a command without the risk of draining the battery. This property is useful even if the device is password protected as many commands are not security threatening, this being especially true when the device serves to control home equipment.

Gesture detection

A Bezel-Tap gesture involves a first tap on the bezel, detected by accelerometers, and a subsequent tap on the touchscreen that must occur within a short time interval (Figure 2.2). So we have two temporal constraints: 1) the time delay between the two events must be greater than a few milliseconds (50ms with our hardware) so that a Bezel-Tap gesture cannot start with a screen tap (this could otherwise happen in the case of a double tap on the screen if the user taps hard); and 2) the delay must not exceed 600ms to avoid taking unrelated events into account (i.e., a screen contact occurring long after a tap).

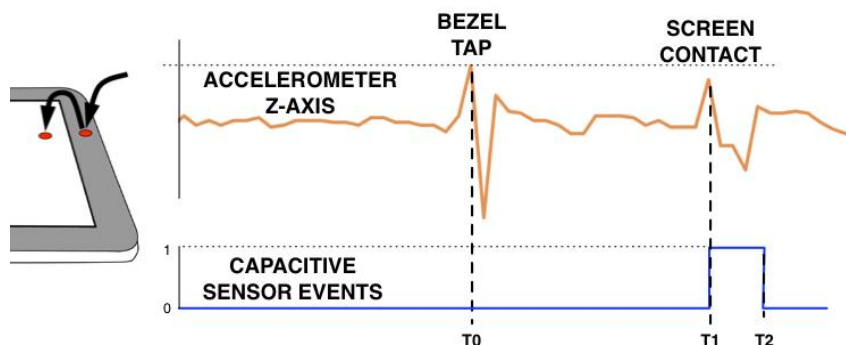


Figure 2.2: A Bezel-Tap gesture involves a first tap on the bezel, detected by accelerometers, and a subsequent tap on the touchscreen.

Incidentally, it is worth noticing that Bezel-Tap gestures allows identifying a location on the device (that serves to select a command) but only rely on the touchscreen to do so. While accelerometers could theoretically serve to detect tap locations, hardware currently available on mobile devices would make this hardly feasible. Using two input modalities (the accelerometer and the touchscreen) solves this problem: the touchscreen provides information the accelerometer is unable to provide.

Power considerations

Bezel-Tap Gestures make it possible to trigger commands quickly, even with the device in sleep mode. This reactivation feature requires the accelerometer to remain permanently powered. The small power consumption of accelerometers makes this feasible without considerably reducing the battery life. Our estimations were that our prototype would only reduce the battery life of about three quarters of an hour every 45h of battery life.

We also investigated the effect of carrying the device in mobile context to see how many taps would be detected by the accelerometer just because of the motion of the device. A high number of detections would drain the battery because, using our technique, taps reactivate the capacitive sensor of the screen. We hence conducted an experiment carrying a tablet inside a backpack during subway, bus and bike journeys. We collected 8 hours of tablet's internal accelerometer log in mobile context. On average 6 taps per hour were detected: 9 taps/hour on bus, 5 taps/hour on subway and 4 taps/hour on bike. Therefore the added power consumption is negligible.

Inadvertent activations

The fast succession of two events from different input channels (a tap followed within a short time interval by a screen touch) is a low probability event that can serve as a unique signature. For instance, no Bezel-Tap gesture is recognized if the user double taps the screen or double taps the bezel. In order to evaluate the probability of inadvertent activations of Bezel-Tap gestures, we completed a field study where we gave 12 participants a tablet for 24 hours, asking them to use a web browser implementing Bezel-Tap Gestures for at least one hour in total. The results show that a few taps were detected by the accelerometer (on average 7.7 taps per hour of use) but none of them were followed by a screen contact in less than 600ms, so that no false positive was detected. This field study suggests that our technique is robust to common usage.

Target selection performance

Selection performance is fast and accurate up to a fairly large number of commands, even in the absence of any screen feedback. We carried an experiment to evaluate the performance of Bezel-Tap gestures in terms of speed and precision depending on the size of the command set (4, 5, 6 or 7 items per bezel region, see Figure 2.3). Additionally, we also wanted to compare performance for the different regions of the bezel (top, bottom, left, right). The results of this experiment confirmed the usability of the Bezel-Tap technique in all four regions of the device in sleep mode. Selection times, on the order of 1.5 sec (with the reaction time included), were compatible with the micro-interaction concept [Ashbrook, 2010]. Performance accuracy was as good for $N = 5$ (96.5%) than for $N = 4$ (96.9%), but decreased for $N=6$. One practical suggestion that arises from these results is that a set size of five items is optimal for the technique and a good solution to select shortcut commands.

2.2.2 Bezel-Tap Menus

Bezel-Tap Gestures can be extended to allow selecting a large number of items. We proposed two new techniques that rest on a hierarchical organization (Figure 2.4). Bezel-Tap Slide (BTSlide) involves a tap on the bezel followed by a slide on the screen. The starting



Figure 2.3: We evaluated the target selection performance depending on the size of the command set: 4, 5, 6 or 7 items.

and ending points of the slide selects a group of items and an item in this group respectively. Bezel-Tap3 (BT3) involves three taps: one on the bezel and two on the screen. The second tap selects a group of items and the third tap an item in this group.

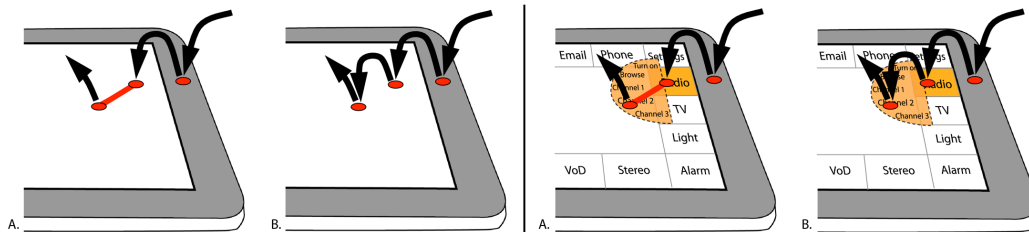


Figure 2.4: Bezel-Tap Menus in expert (left) or novice mode (right). Bezel-Tap Slide (A) involves a tap on the bezel followed by a slide on the screen, while Bezel-Tap3 (B) involves three taps.

These Bezel-Tap menus are hierarchical. Their first level consists of five rectangular items for all four bezel regions (Figure 2.5). According to our experiment results we decided to expand even items and reduce corner items by 15% in order to increase even-numbered items success rate. The second level of the menu rests on 180° radial menus. The second tap (resp. the starting point of a slide) selects one of these radial menu, for instance the "Radio" menu in Figure 2.5. The third tap (resp. the ending point of a slide) selects an item in this menu. Menus only appear if the user waits more than 600ms between the second and the third tap. A complete novice user just needs to tap on the bezel then anywhere on the screen and wait for 600ms. An expert user will bypass these stages by performing all taps (or slides) without any visual feedback. To make interaction easier, we did not assign menus to corner items (which can serve as one-step shortcuts for very frequent commands). This design allows for a total of 12 radial menus (3 per region), and each radial menu offers five items. A Bezel-Tap menu can thus comprise a total of 64 items: 4 corner shortcuts and 60 menu items.

Bezel-Tap Menu performance

We carried a study to evaluate the usability of Bezel-Tap Menus for command selection in expert mode. Participants were asked to perform the full gestural sequence illustrated in Figure 2.4, using either BT3, BTSlide or an extension of Bezel Gestures [Bragdon *et al.*, 2011] that allows selecting more commands than the original technique. The results show that after little training all three techniques allow accurate selection in a two-level hierarchical menu. Bezel-Tap techniques (BT3 and BTSlide) were more accurate than Bezel Gestures (error rates were 5.2%, 4.5% and 8.7%, respectively), but slower. Anyway, the speed was pretty good for Bezel-Tap techniques (1.6s on average), this making them appropriate for

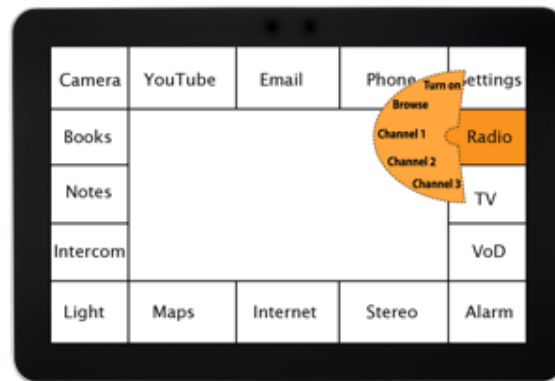


Figure 2.5: Illustration of the content of a two-level Bezel-Tap Menu.

micro-interactions. This result is even more interesting considering that Bezel-Tap gestures can not only select a command but also reactivate the device in this short amount of time.

2.2.3 Summary

Bezel-Tap Gestures is a technique that allows selecting commands quickly on mobile devices. Bezel-Tap Gestures do not interfere with common interaction techniques when the device is in active mode. They also work in sleep mode, both waking up the device and triggering a command. Our work explored the use of a simple but effective method to extend input interaction on mobile devices, i.e. tap gestures. Since the publication of our work on Bezel-Tap gestures in 2013, the use of tap gestures has been applied to back-of-device interaction on smartphones [Granell and Leiva, 2017] and for rapid command selection on smartwatches [Oakley *et al.*, 2015].

2.3 Exploring spatial data on head-worn displays using hand-to-face input

During my post-doctoral fellowship at the University of Manitoba, I worked with Pourang Irani and Barrett Ens on solving the challenge of facilitating input interaction, spatial navigation in particular, on head-worn displays (HWDs). We proposed hand-to-face input as a novel, alternative method for interacting with HWDs. This work led to one publication at CHI 2014 [Serrano *et al.*, 2014a].

2.3.1 Hand-to-face gestures

We define hand-to-face input as any gesture that involves touching, rubbing, scratching or caressing the face (Figure 2.6). This approach is especially well-suited for interaction with HWDs for many compelling reasons: (i) the face is often touched making it a promising area for subtle interactions; (ii) it offers a relatively large surface area for interaction, but not normally clothed as are other areas; (iii) it facilitates eyesfree, single-handed input, which can be invaluable in mobile settings (e.g. riding a bike, holding on in a bus); and (iv) is in close proximity to the HWD, making it likely to accommodate device-borne sensors and creating a natural extension of the device temple.



Figure 2.6: Hand-to-face input for navigation. a) Panning, b) Pinch zooming, c) Cyclo zooming, d) Rotation zooming. Our studies show that Cyclo was not socially acceptable while Rotation was not efficient.

Suitable areas for interaction

To explore the breadth of potential hand-to-face gestures and their mapping to interactive tasks, we elicited user input through a guessability study [Wobbrock *et al.*, 2009]. For a set of common mobile tasks [Ruiz *et al.*, 2011], divided into action or navigation tasks, we asked participants to suggest suitable gestures on the face (above the neck) and on the HWD.

On the face, participants produced gestures for a total of 11 different areas, such as the cheek, forehead, ear, chin or jaw. The results reveal a distribution (Figure 2.7) with gestures concentrated on the cheek and then on the forehead. Other areas saw an equal distribution of gestures: jaw, ear, temple, and chin. Areas such as the eyes, nose, lips, neck and hair were least used. On the HWD, participants used 5 different interaction areas: temple, hinge, frame, bridge and glass. Most of the gestures were situated on the temple.

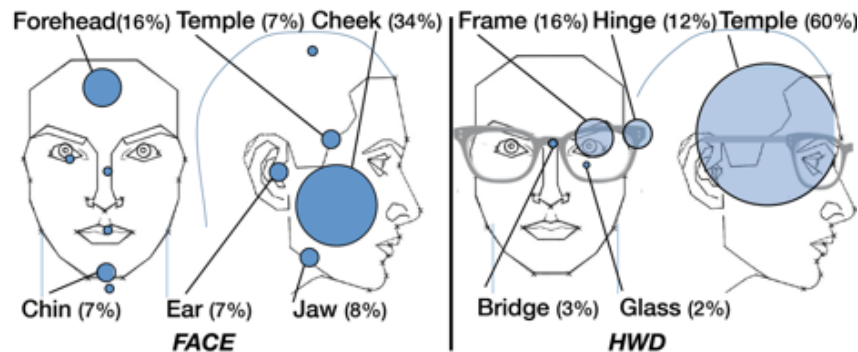


Figure 2.7: Main areas identified by participants as suitable for input on the face (left) and on the HWD (right). The circles' size is proportional to the percentage of gestures (in brackets).

Overall users preferred interaction on the face for the navigation tasks, while opinions were mixed for the actions tasks. Users particularly preferred using the face for panning and zooming. Users indicated that "the face provides a larger area", which is perceived as a benefit for panning and zooming. This is particularly true when using the cheek, since it is "the best part to interact with the face" and it is "like a touchpad".

2.3.2 Spatial navigation with hand-to-face input

We carried two studies to evaluate the performance, physical effort and user preference of hand-to-face gestures for panning (study 1) and zooming (study 2).

For the first study we implemented three panning interactions: displacement-based (Pan-D), flick (Pan-F) and rate-based (Pan-R). With Pan-D, finger movement is directly mapped to the movement of the map. In Pan-F, the user flicks to pan, mimicking the iOS flick behaviour. In Pan-R, the distance panned from the initial touch position is mapped to the finger velocity movement. Participants used four different interactive surfaces, two on the face (Cheek and Forehead) and two on the HWD (Oversized and Regular temple). We include two different temple sizes in order to study the impact of its size on navigation.

This first study revealed that the best facial area for input is the Cheek. The Forehead and the Regular HWD Temple not only showed worse performance, but also result in higher fatigue. Overall there was no difference between the Cheek and the Oversized temple, but both were favored over the Regular temple. The Oversized temple, however, is far larger than most HWDs, suggesting that the Cheek is a preferred interaction region.

For the second study, we selected three different zooming techniques, based on a combination of prior known methods and from the guessability study: Linear, Rotation and Cyclo. Linear zooming, by pinching with two fingers, is the classical touchscreen technique. Circular zooming with two fingers (using the angle of rotation) is based on the metaphor of adjusting an optical lens. Cyclo is a one finger zooming technique proposed as a way to avoid clutching when zooming [Malacria *et al.*, 2010]. It consists of doing a circular movement with one finger. The orientation of the rotation is mapped to the zoom direction (in or out). From the previous study we dismissed the Forehead area due to its low overall results.

The results of the second study extend further our exploration of Study 1, providing insight into hand-to-face interaction for document navigation. The main finding is that the Cheek is more efficient than both the Oversized and Regular temples for zooming. While the Oversized temple was efficient in Study 1 for one finger panning, it becomes inefficient with a two-finger gesture. Both the classical Pinch and the single-finger Cyclo are equally efficient in our study.

2.3.3 Social acceptability of hand-to-face input

While we demonstrated that hand-to-face interaction techniques improve navigation performance, we know little on how comfortable users would feel in different social contexts. We therefore carried a controlled exploration of the social acceptance [Rico and Brewster, 2010] of our hand-to-face gestures. Participants watched a video of an actor performing panning and zooming gestures in front of a wall and then performed themselves the same gestures 3 times. Participants were asked to rank on a 5-point Likert scale the social acceptability of each hand-to-face gesture.

We found no difference in social acceptability between Face and HWD, but with a trend showing better acceptance for interaction on the HWD. The acceptance rate for both face and HWD gestures in any social context is above 50%. Results were rather homogeneous on the HWD, with a constant 10%-12% of disagreement for all social contexts except in front of strangers, where this value is 18%. We found more differences on the Face, with no disagreement when at Home or Alone, but with 31% disagreement in Public places and 25% in front of strangers. Comments from participants also show that most of them don't mind using the face: "I don't think it would disturb me to do the gesture either on the skin or on the temple.". One female participant indicated the problem of dirty hands on the face: "the face can be affected when perspiring".

Concerning the acceptability of the different hand-to-face gestures, we found a significant difference between the panning techniques, Pan-D being better perceived than Flick. We also found a difference between the zooming techniques, Linear being better perceived than Cyclo. Participants commented that Cyclo might be perceived as insulting, as it could signal that "you are crazy" in many cultures. This gesture seemed also more visible: "I feel all the gestures are quite subtle except Cyclo which might attract attention".

2.3.4 Summary

We explored hand-to-face gestures for HWDs without emphasizing the technology that would ultimately support this style of input. Interestingly, in 2017 researchers at Keio University (Japan) proposed a technology that senses touch gestures on the cheek by detecting skin deformation [Yamashita *et al.*, 2017]. Our work on face input was the first to propose the use of such an area for interaction, and has been succeeded by other interesting explorations of head-based input for mobile computing involving the ear [Kikuchi *et al.*, 2017] or the nose [Lee *et al.*, 2017].

One of the remaining challengers of on-body interaction is to consider the Midas-touch problem [Hinckley, 2003], i.e. how to differentiate casual and explicit on-body gestures. Two obvious but cumbersome solutions include touching the HWD to initiate the face detection or using a voice command. Another solution is to use gestures that are very different from casual ones.

2.4 Mobile True-3D Displays

During my post-doctoral fellowship at the University of Manitoba, I investigated with Pourang Irani (University of Manitoba, Canada) and Sriram Subramanian (University of Bristol, UK) how to explore volumetric content on mobile true-3D displays. As said earlier, true-3D refers to any 3D digital display capable of producing mid-air, full-depth-cue (or volumetric), multi-angle and/or multi-user images without the need for user instrumentation. This work led to one publication at MobileHCI 2014 [Serrano *et al.*, 2014b].

2.4.1 Display properties

We first studied the various factors that can potentially influence the effective deployment of true-3D on mobile devices (Figure 2.8). We focus on mobile-mounted 3D projection, which means that the true-3D projection moves with the mobile device as if both were attached.

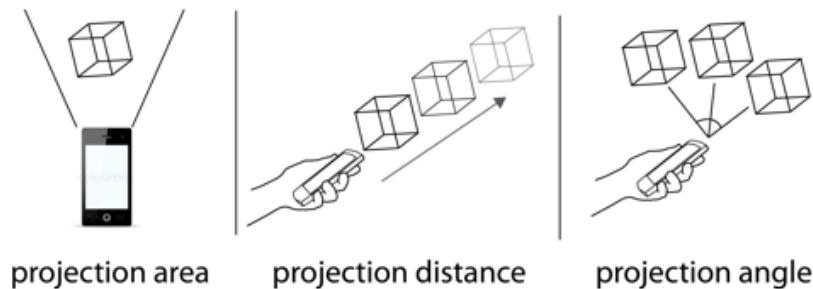


Figure 2.8: Mobile true-3D properties: projection area, distance and angle.

Projection area: While pico-projectors need a projection surface, true-3D projectors may display an image in mid-air around the mobile device. Prior work has generally kept the mobile projection pointing downward, straight or steerable (using a motor to direct the projector) [Cauchard *et al.*, 2012]. These solutions provide significant flexibility in finding a suitable projection surface. A true-3D mobile display needs not be constrained by the position of the projection throw. Therefore after considering the potential projection areas around the smartphone, we decided to focus on the top area of the phone (Figure 2.8-left). This area always remains visible when the user rotates the phone to inspect the 360° true-3D image.

Projection distance to the 3D object: The distance between the mid-air 3D projection area and the smartphone (Figure 2.8-center) may have an impact on users' visual perception and direct input. If the projection is far from the device, it may affect the continuity of the visual search but even further limit direct-touch interaction with the true-3D and require indirect forms of interaction.

Projection angle: We define the projection angle as the angle between the phone's y-axis and the 3D object (Figure 2.8-right). Traditional depictions of mobile true-3D envision the 3D content at a 90° angle relative to the phone's plane or displayed directly over the touchscreen. These depictions assume the best projection extends the mobile display into a 3D volume. However, this vision is limited as it considers the true-3D simply as an extension of the touchscreen instead of viewing it as a secondary display that can extend the mobile phone's capabilities.

Beyond these three main properties, we also considered the *volume* of the projection and the user's *point of view*. The display *volume* may affect visual search as well as direct-touch interaction. The visual exploration (hence the user's *point of view*) of a mobile true-3D display relies on wrist rotation dexterity to avoid complex interactions for rotating it. Given the restrictions in wrist rotation angles, we expect limited accessibility to occluded areas on the true-3D projection.

2.4.2 Spatial configuration

We carried an experiment to identify the best spatial configuration for the projection of a mobile true-3D display to ensure effective visual search. We explore the case of natural user interaction, i.e. wrist rotation for search without interface support, but do not explore any type of user input. We focus on the properties described in the previous section (Figure 2.9): projection's angle to the phone plane, distance to the phone, volume and pattern position on the true-3D (point-of-view). Participants were required to identify the location of a graphical pattern on a 3D opaque sphere on the true-3D display. The sphere was separated into eight parts, each one containing a unique pattern (Figure 2.9-right).

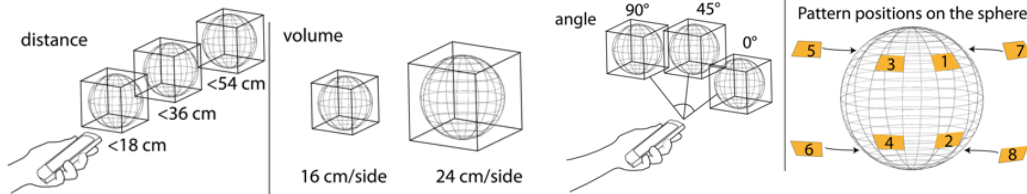


Figure 2.9: Mobile true-3D properties evaluated in our first experiment.

The results of this first study demonstrates that visual search is easier when the 3D object is positioned at an angle of 0° or 45° and at a distance of less than 36 cm. We also found that the region in the back and opposite to the hand holding the device is the weakest for object search. This is primarily due to the wrist dexterity as observed during the experiment and from participant feedback. Wrist dexterity also affects objects located further away, i.e. these become hard to inspect under all angles. Thus, our results recommend shorter distances if the device is to solely rely on wrist rotation for viewing the display.

2.4.3 Direct touch interaction

We carried a second experiment to investigate the effect of volume size on visual search and direct input. Participants were required to identify the location of a graphical pattern on the true-3D display and to select it with direct mid-air touch. We added one larger

volume size than in experiment 1 to further investigate this factor. The results of this second experiment informed us on the suitable values for the projection volume for direct interaction. Overall, completion time increased with the volume of the display. Projections smaller than 24cm/side improved efficiency. The slight cost in accuracy at smaller volumes suggests that target sizes need to be considered carefully for such displays.

2.4.4 Summary

In HCI we find numerous examples of novel technologies whose adoption, from discovery to commercial use, take decades. Buxton refers to this process as the Long Nose of Innovation [Buxton, 2008]. Our work was motivated by our will to reduce the long nose for mobile true-3D. Extensive research is taking place to engineer mobile true-3D [Nii *et al.*, 2012] and following our work, researchers at Queen's University (Canada) presented a 3D flexible smartphone rendering "holographic" images providing motion parallax and stereoscopy to multiple users without glasses [Gotsch *et al.*, 2016]. Our contribution was to identify application guidelines, limitations and scenarios to help future adoption of this technology.

2.5 Mobile map exploration for visually impaired people

In this section we present the work carried by Sandra Bardot, a PhD student at the University of Toulouse that I co-supervised with Christophe Jouffrais (CNRS - IRT Lab, Toulouse). We designed a mobile technique, based on hand tracking and a smartwatch, in order to provide pervasive access to virtual maps for visually impaired (VI) people. This work led to one publication at the conference MobileHCI 2016 [Bardot *et al.*, 2016].

2.5.1 Interaction principles

Our work focused on the spatial exploration of geographic maps with associated data (e.g. demographic or weather maps). We focused on the complementary objectives of improving the spatial component of exploration, while providing the user with large surfaces and collocated feedback. To this end, we first established the principles of such an interaction in terms of which data should be explored, the tracking to use and the interaction modalities to employ.

Map content for VI people

We first analyzed the layout and content of regular raised-line maps, which are the most commonly used tools by VI people for exploring maps. Tactile raised-line maps have two main advantages: information is tactile, and spatial exploration is direct. They are made according to guidelines (simplification of contours, reduction of the number of elements, legends, etc.). The most important elements of these maps are the contour of areas, the points of Interest (POI) and a Braille legend describing each area and POI (Figure 2.10). Finally these maps present data associated to each area or POI, for instance the population of a region. This information is usually written outside the map with Braille.

In addition, tactile exploration of novel spaces relies on behavioural strategies such as scanning the image from left to right and top to bottom. We aimed to preserve these strategies during spatial exploration of virtual maps, therefore using hand tracking to locate the hands and render the information that is under the hand.

Hand tracking for ubiquitous map exploration

Spatial exploration by using hand tracking instead of touch input offers several advantages for VI people. First, VI people tend to put multiple fingers and hand palm down on the surface, which, in absence of visual feedback, generates unexpected events. Instead, hand

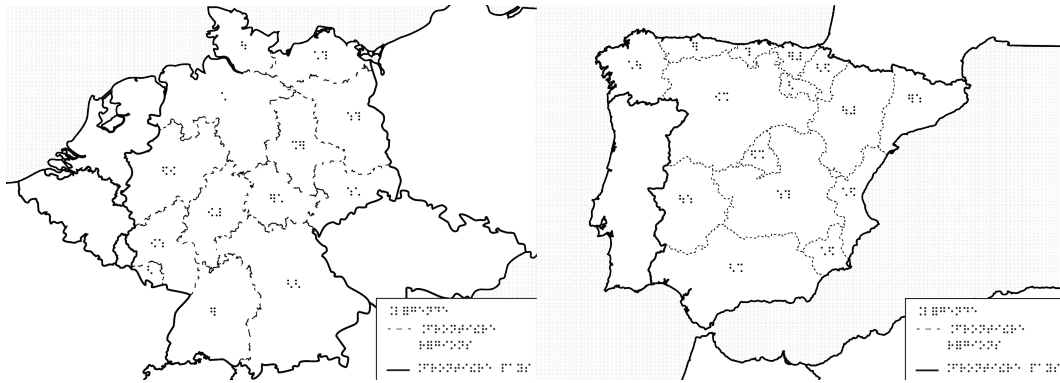


Figure 2.10: Raised-line maps of Germany (left) and Spain (right) showing regions, points of interest and legends.

tracking can simply associate one point with each hand. Second, hand tracking allows performing mid-air gestures, for instance to change the information level of the map when raising the hand. Although mid-air exploration may seem difficult for VI people, we explored its use to perform coarse-grained spatial exploration. Finally, mobile and low-cost hand tracking solutions have recently been proposed, which could leverage making map exploration possible in different contexts and on different surfaces, such as a regular desk, without the need for an underlying touch sensitive surface. Coupling hand tracking with a wearable device for input and feedback (e.g. a smartwatch) makes it possible for VI people to explore virtual maps in many places such as at school, in a mall or at home.

Using a smartwatch for localized feedback

Previous work has shown the interest for VI people to use wearables [Ye *et al.*, 2014] and particularly smartwatches: they are small, easily accessible and unobtrusive to wear, and can improve information access and social interactions. Current smartwatches have the advantage of including speakers and vibratory feedback. We decided to use a smartwatch to provide hands-free map exploration. We used the smartwatch both as input and output. As input, the device's touchscreen is used to filter or brush data by performing simple horizontal or vertical swipe gestures. As output, the device is used to render localized Text to Speech (TTS), for instance the name of regions. The vibratory feedback is also used to render information, such as the border between regions.

2.5.2 Non-visual map exploration techniques

We identified different mappings between the smartwatch input/output modalities and the map exploration task, resulting in three different exploration techniques (Figure 2.11).



Figure 2.11: We investigated three non-visual map exploration techniques: plain (left), filter (center) and grid-filter explorations (left).

Plain exploration

Plain exploration is the equivalent to the exploration performed on a raised-line map: each element on the map is rendered. The smartwatch is only used as an output for this technique. We combined auditory and vibratory feedback. TTS reads out information underneath the hand, such as the name of the region and its population. A 100 ms vibration notifies the transition from one region to another one. A continuous vibration means that the hand is outside the map.

Filter exploration

Filtering data before exploration allows reducing the amount of information to render through TTS, and thus reduces the user's cognitive load. The filtering allows selecting a sub-range of values, for instance regions with more than a hundred thousand residents. To perform the filtering, users make swipe gestures on the smartwatch (Figure 2.11-center). After selection, only the data that corresponds to the selected filter is read out, similar to the Plain exploration technique.

Grid-Filter exploration

With the previous techniques it can be difficult to find certain regions in a map especially if they are small. To get a full glance of a map without missing any region, one solution consists in using a 3x3 grid, i.e. reading out the information concerning all the regions contained in each cell of the grid [Zhao *et al.*, 2008].

However, when gathered within a grid, the spatial relationships between regions are masked. To overcome this limitation, we combined the Filter exploration mode with a Grid-based exploration mode. The user can use one or the other interaction level according to hand height above the map. When the hand is lying onto the table, the user explores the map in Filter mode (i.e. only the data that corresponds to the selected filter is read out). When the hand is moving over the table, the user explores the map in grid mode (i.e. reading out the information concerning all the regions contained in each cell of the grid).

2.5.3 Experimental results

We carried a study to compare the effectiveness of our virtual map exploration techniques against exploring a raised line printed version (Figure 2.12). The task was to explore a map, and answer a question as fast as possible. Twelve visually impaired participants volunteered for this experiment.

Overall, Grid-Filter was faster than the other techniques: on average, answering a question with the Grid-Filter technique took 40 s, with Filter 83 s, with Raised-line 127 s and with Plain 172 s. Response times were significantly longer with Raised-Line and Plain. This is due to the fact that users had to thoroughly explore the map in order to find the targeted region and the associated data to answer a question. On contrary, Filter and Grid-Filter renderings quickly provide access to the answer. Interestingly, the Grid-Filter technique was the more efficient but not the preferred one. Most participants ranked the Filter technique first on user preference criteria. These results can be explained by the use of mid-air gestures in the Grid-Filter technique: many participants reported that it is tiring, and that it is difficult to build a mental representation of the map when their hand is moving above the map.

2.5.4 Summary

Our work focused on the usability of virtual maps, including different exploration techniques (Plain, Filter and Grid-Filter), as opposed to the regular raised-line maps. Overall, the results suggest that VI people are able to explore geo-spatial data in virtual maps.

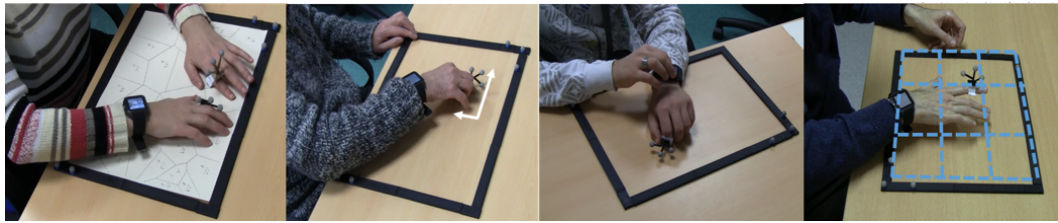


Figure 2.12: From left to right: exploration of a raised-line map with an infrared tracker on the finger; plain exploration of a virtual map; left/right swipes on the smartwatch to filter the data (filter exploration); and raising hands to gather the data within a virtual grid layout (grid-filter exploration).

More precisely, they show that when filtering functions are added to virtual maps, which is impossible to achieve with tactile maps, they provide the user with an efficient mean to retrieve specific answers about the elements of the map. These results could be of interest to design eyes-free map exploration for sighted users, as a way to overcome the limited display space of smartwatches. To this end, the performance of our non-visual techniques would need to be evaluated against current visual techniques for map exploration on watches.

2.6 Chapter conclusion

In this chapter we presented our contributions to mobile interaction with large data spaces, which evolved around gestural input and offscreen data exploration. We proposed two novel types of input gestures: Bezel-Tap gestures, for selecting discrete commands from sleep mode, and hand-to-face input, a type of on-body gestures which proved to be valuable for continuous exploration of spatial data on head-worn displays. Concerning offscreen data exploration, we investigated how to access volumetric content around the smartphone's displays using true-3D, and how to explore spatial data without visual feedback using the smartwatch output capabilities.

Chapter 3

Multi-Display tools and interfaces

Chapter content

3.1	Research challenges	25
3.2	Gluey: fluid interaction in MDEs using a head-worn display	28
3.3	TDome: fluid interaction in MDEs using a dedicated multi-DoF device	31
3.4	Interaction with Overview+Detail Interfaces on 3D Public Displays	34
3.5	Splitting Detailed Views in Overview+Detail Interfaces	36
3.6	Chapter conclusion	39

Multi-displays environments (MDEs) have shown significant value for interacting with heterogeneous data sources and in multiple contexts such as collaborative analysis [Bortolaso *et al.*, 2013], crisis management [Chokshi *et al.*, 2014] and scientific data visualisation [Wang Baldonado *et al.*, 2000]. MDEs allow organizing information across displays, enhancing individual and group work, providing support to peripheral information and extending the interaction space. However, interaction in such environments is challenging due to the heterogeneity of existing solutions, and the benefits of extending the display space have been insufficiently studied for complex data (such as 3D content or very large graphs).

3.1 Research challenges

Our contributions to interacting with MDEs address two main research challenges: novel tools for fluid interaction in MDEs (1) and complex data exploration on MDEs (2).

3.1.1 Fluid interaction in MDEs

The emergence of MDEs has resulted in a device vacuum: to our knowledge no device has been specifically implemented for optimizing interactions in such spaces. Researchers have mainly proposed adapting existing devices such as the mouse for multi-monitor pointing [Benko and Feiner, 2005], or smartphones for cross-display data-transfer [Boring *et al.*, 2010] or distant pointing [Nancel *et al.*, 2013].

However such adaptations can result in undesirable side effects: mice do not support well input on large displays and smartphones held in mid-air can be tiring and cumbersome for long interactions. Recent research has demonstrated the use of wearable devices to perform cross-device interactions [Houben and Marquardt, 2015]. However, current wearables lack proper input mechanisms and mainly serve private purposes. If MDEs are to become the office of the future (see Figure 3.1), as envisioned by [Raskar *et al.*, 1998],

can we design a device specifically tuned for such an environment? Adopting a unique device would avoid the homing effect when switching from one device to another, enhance privacy in such environments through personal data control and visualization, lead to a coherent set of interactions with the varied MDE applications, and ultimately contribute to a more fluid interaction [Bederson, 2004]. The design of such an unifying device could be based on two complementary approaches: using an already existing device, to facilitate the spread of the solution, or creating a new tailored device to replace existing ones.



Figure 3.1: Office of the future as envisioned by [Raskar *et al.*, 1998].

3.1.2 Exploring complex data on multiple displays

To explore large information spaces such as web pages or maps, research on visualization proposes three solutions to separate focused and contextual views: Overview+Detail (spatial separation), Zooming (temporal separation) and Focus+Context (seamless focus in context) [Cockburn *et al.*, 2009]. Overview+detail configurations fit particularly well MDEs, since overview and detail (O+D) views can be displayed on different screens. In our work we addressed two under-explored research questions around the use of multi-display O+D to explore complex data (3D content and very large graphs): interacting with 3D data on public displays, and exploring multiple regions of very large graphs simultaneously. We present and motivate the importance of these two research questions in the following paragraphs.

Interaction with 3D data on public displays

Public displays allow pedestrians to interact with 2D content such as city maps or tourism information. In this context, existing systems have used smartphones as personal views (Detail) of the public display (Overview), leveraging multi-user access to one display [Cockburn *et al.*, 2009]. Visualisation of 3D content on public displays is emerging to visualize scientific 3D data [Song *et al.*, 2011]; to explore culture heritage 3D scanned objects; to play public 3D games or to navigate a city 3D map.

Most of these examples already include the use of a personal device to interact with the 3D content (to orient and position a slice plane or to navigate in the 3D environment) but few consider how to apply the Overview and Detail (O+D) paradigm using the smartphone. Using the O+D on mobile devices will provide the user with the ability to pri-

vately visualize details of the 3D environment while still taking advantage of the mobile device input to interact with the 3D content. However, the public context imposes certain constraints in terms of user's profiles (mainly beginners) and appropriate interaction techniques (which need to be easy to understand and perform). One of the main challenges then is to provide an easy to perform technique to control the position of the Detail view in a 3D Overview.

Exploring multiple regions of very large data spaces simultaneously

As said earlier, MDEs are well suited for scientific data visualization. We identified the following challenge during a collaboration with biologists carrying research on cancer. These biologists archive knowledge in graphs called molecular interaction maps (MIM [Kohn, 1999], see Figure 3.2), which contain several types of nodes (molecules, protein, etc.) and connections. As research on cancer progresses, results are added to existing MIM maps, which grow extremely large (the Alzheimer MIM map contains 1347 nodes [Kohn, 1999]) making them difficult to read and edit using the traditional panning+zooming interactions. Moreover, another consequence of this growth is that connected nodes can be located far apart from each other, thus, requiring even larger surfaces to visualize the data.



Figure 3.2: Illustration of a MIM graph containing a total of 880 nodes and 732 connections [Kaizu *et al.*, 2010].

Despite the advantages offered by O+D interfaces when working on large datasets (like graphs), these interfaces reach their limits when it comes to work on multiple regions of the overview simultaneously. An example from the previous context would be connecting distant nodes of very large graphs for example or to create a link between 2, 3 or 4 nodes. These types of multi-node links are usual in large graphs such as MIMs. Moving the detailed view repeatedly from one region to another is tedious and interaction complexity increases with the number of regions to work on.

To address this situation, several techniques have been designed in single or multi-display configurations to support the use of more than one detailed view simultaneously [Elmqvist *et al.*, 2008] [Javed *et al.*, 2012]. Earlier work had also established a set of rules for working with multiple views [Wang Baldonado *et al.*, 2000]: the "rule of diversity" recommends the use of one view per information type and the "rule of parsimony" suggests using multiple views minimally. However none of these works has investigated the optimal number of detailed views to use, most existing techniques using 2 or 4 views. Finding such optimal number is thus still an open question.

3.1.3 Summary

In this chapter we present our contributions to the aforementioned challenges and related research questions (Figure 3.3). To address the need for fluid interactions in MDEs, we adopted two approaches: using an existing device, in our case a head-worn display interface that we named Gluey (section 3.2), or creating a novel dedicated device, namely TDome (section 3.3). Concerning the exploration of 3D data on public displays, we studied the use of around-device gestures to manipulate the detail view on a 3D public display (section 3.4). To explore multiple regions of graphs simultaneously, we studied the optimal number of views on an overview+detail interface (section 3.5).

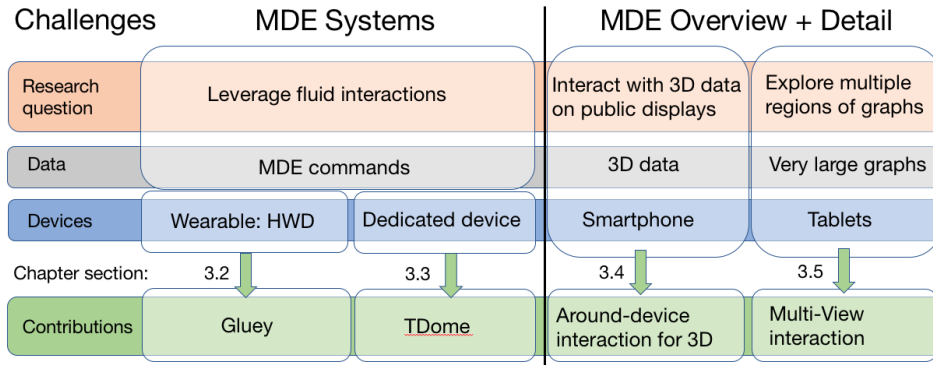


Figure 3.3: Research questions and our corresponding contributions to the domain of Multi-Display Environments.

3.2 Gluey: fluid interaction in MDEs using a head-worn display

To address the challenge of creating fluid interaction in MDEs, we consider what unique role can Head-Worn Displays (HWDs) play beyond simply being an additional display medium. Our contribution, named Gluey, is a HWD user interface that acts as a "glue" to facilitate seamless information flow and input redirection across multiple devices. This work carried in collaboration with Pourang Irani and Barrett Ens (University of Manitoba, Canada), and Xing-Dong Yang (Darmouth College, USA), was presented at MobileHCI 2015 [Serrano *et al.*, 2015].

3.2.1 Usage Scenario

We present a scenario depicting our vision of Gluey (Figure 3.4). John, an architect, relies on numerous digital devices while juggling between ad-hoc tasks in his daily work (Fig. a - b). Most of his drawing takes place on a desktop computer attached to two monitors and a printer. John uses a tablet for sketching and for showing his drawings to clients. He uses a smartphone for his communication needs.

John completes a plan for a gazebo on his desktop computer and prepares to use his tablet for presenting it to a client. After saving his presentation, John moves it with his mouse to a clipboard on his HWD (Fig. c), so that he can later copy it onto any other device. He then glances at his tablet and uses his mouse to grab the presentation from the HWD clipboard and transfers it to the device that is in the HWD's direct view, i.e. the tablet screen (Fig. d). Meanwhile, his business partner James sends a text message to his smartphone, noting the urgency of their soon to begin meeting. He is able to quickly reply to James, without opening his phone's soft keyboard. He instead uses the desktop

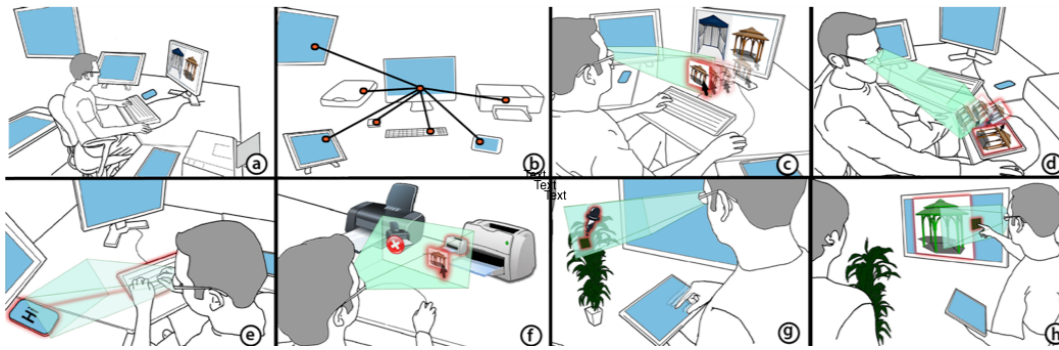


Figure 3.4: Usage scenario of Gluey.

keyboard, already near his hands, to type his reply. He does this by turning to look at his phone, which is detected by Gluey's head-mounted camera. As he types, the text gets redirected to the smartphone (Fig. e). Before leaving to meet his client, John prints a copy of the same presentation he had on the HWD clipboard by dragging its document icon to a printer icon, which comes in view when he looks in the direction of the printer (Fig. f).

At his client's office, he is led to a boardroom equipped with a large wall display for presentations. Since John has previously registered this display within Gluey's spatial configuration, he can immediately use his tablet as a trackpad to drag the presentation of his gazebo drawing from his HWD onto the large display. The client discusses different options, and suggests a different color for the gazebo, which matches the leaves of a nearby plant. John glances over to the plant, and selects its deep green color with Gluey (Fig. g), then fills in the gazebo using a finger gesture (Fig. h).

This scenario captures how John seamlessly interleaves many tasks in an ad hoc manner. He minimizes his need to switch input devices and can seamlessly migrate content to where it is needed by glancing at devices. While the actual data transfer is happening in the "cloud", John need only concern himself with simple, intuitive gestures for manipulating abstract representations of objects that suit his immediate needs. Furthermore, he can extract content from his physical surroundings (such as color or captured images) which he copies onto his digital display.

3.2.2 Gluey key components

To implement the previous scenario, we exploit the unique features of HWDs, such as view-fixed displays, cameras and inclusion of spatial sensors. We describe the three key components of our system using the terms Glueboard, Glueview and Gluon (Figure 3.5) as these define concepts in our implemented system but have slightly different meanings than what might traditionally be considered a "clipboard", "head-gaze input" and "input redirection".

Glueview

Many HWDs now include built-in sensors from which we can determine a user's head position in relation to the environment and parse details about what they are viewing. Because head motion dominates over eye movement when switching targets [Han Kim *et al.*, 2010], we can gain information based on a user's general head orientation, for instance which device they are looking at. We use the term Glueview to denote the use of Field of view (FoV) tracking (i.e. head orientation as a proxy for gaze tracking) for enabling implicit device registration: for example by simply having the smartphone in the FoV, the user can link his desktop keyboard and the smartphone (Figure 3.4 - e).

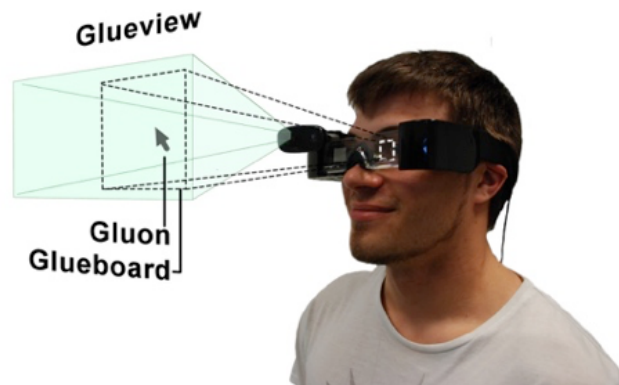


Figure 3.5: Three key components of Gluey: Glueview (field of view tracking), Glueboard (always-available feedback) and Gluon (input redirection). Together, these components enable several novel interaction techniques.

Glueboard

Rather than treat the HWD as just another information display, we can use the device's view-fixed display as an always-available "canvas" for displaying visual feedback about interactions in the MDE. For example, this display can show text next to a keyboard, augment devices with useful status information or provide a visible clipboard space to store multiple data objects in transit between copy/paste operations to multiple destinations. We call this combined, always-available feedback and multi-object-clipboard space the Glueboard.

Gluon

A main limitation of multi-device use is the need to change input modes when switching between devices. Since HWDs do not need to be held, the Gluey user can use any input device at hand, whether it be a keyboard, a mouse, a mobile touchscreen or mid-air finger gestures to control multiple displays. The Gluon represents the concept of pairing all available input devices with displays in the MDE to provide a unified interaction experience, independent of input mode or display type. For instance a user can use a desktop keyboard to enter text on a smartphone SMS (Figure 3.4 - e).

3.2.3 Gluey implementation

We created a proof-of-concept prototype using a HWD equipped with a webcam. We implemented all the features presented in our scenario: migration techniques across multiple devices; input redirection across multiple contexts; always-available visual feedback and environment's spatial model. All the implementation details can be found in the paper. We gathered preliminary user feedback by having 12 participants use Gluey for roughly 5 minutes for each technique. Overall, participants were positive at the concept of Gluey. They gave mostly favorable comments on content redirection features (copy and paste on the Glueboard) and input redirection. Seamlessly taking a snapshot and pasting it later on a device was another highly appreciated interaction. These results were encouraging taking into account the limitations of the prototype, such as its weight or the limited display FoV.

3.2.4 Summary

We introduced Gluey, a head-worn software interface that acts as a "glue" for bonding information across the various displays in a MDEs. Gluey exploits three key components: Glueview (i.e. using users' head-direction as a proxy to detect which device they are looking at), GlueBoard (i.e. using the device's view-fixed display as an always-available "canvas" for displaying visual feedback about interactions) and Gluon (i.e. using the input device at hand seamlessly between displays). Some challenges need still to be addressed to improve the concept of Gluey, such as helping users transition from their prior multi-display use (a mix of USB memory sticks, email and cloud services) to Gluey; avoiding unwanted device activations due to head orientation; and reduce the visual overload due to the always-available visual feedback.

3.3 TDome: fluid interaction in MDEs using a dedicated multi-DoF device

In this section, we present TDome, a novel touch-enabled 6DOF input and output device that facilitates interactions in MDEs. This project was carried by Houssein Saidi, a PhD student that I co-supervised with Emmanuel Dubois (IRIT - University of Toulouse), in collaboration with Pourang Irani (University of Manitoba, Canada). This work led to one publication at CHI 2017 [Saidi *et al.*, 2017].

3.3.1 Usage Scenario

We present an illustrative usage scenario with TDome prior to presenting its features. This scenario is inspired by our implication in neOCampus, a smart campus project at the University of Toulouse. Harry is an engineer working on a smart campus project that monitors data collected by multiple sensors on the university. To visualize and interact with the large datasets of energy consumption, the university has set up a multi-display environment composed of several displays, a large projection wall and two TDome devices resting on a tabletop.

As Harry enters the room to start his daily supervision of the energy data, he grabs one TDome and uses it to initialize the multi-display environment by simply pointing at each active display (Figure 3.6-A). He then selects the wall projection by rolling the device toward the wall (Figure 3.6-B). Harry decides to spread the data visualisation across two displays: he selects both displays with TDome and transfers the visualizations from one to the other with a TDome gesture (Figure 3.6-C). As he wants to look closer at information on the second display, he grabs TDome and walks towards the display, using the device in mid-air to perform a zoom on the data for a closer look.

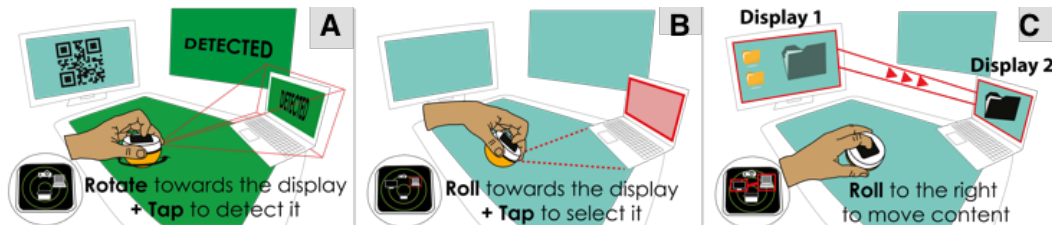


Figure 3.6: TDome facilitates several common tasks in MDEs, such as d) display registration using its embedded camera, e) device selection and f) cross-display data transfer.

Later that day, Mary enters the room and grabs the second TDome. They have a meeting to explore the university map to mark points of interest. Harry and Mary take their

personal smartphones and couple them with each TDome to benefit from personal interactions. Each smartphone shows a personal view of the projected map, which allows them to add and access personal annotations. Before ending, Harry wants to log onto the campus website and upload his annotations: he rolls TDome towards himself to display a virtual keyboard on the device's touchscreen and enter his personal password discreetly on the login page, displayed on the tabletop.

This scenario illustrates how TDome allows users to detect surroundings displays arrangement, select one display, move content between displays, reach to content at distant displays and perform personal interactions on TDome.

3.3.2 TDome key components

TDome results from the composition of a touchscreen with the Roly-Poly Mouse [Perelman *et al.*, 2015]. The Roly-Poly Mouse is a dome-like mouse providing rock, roll, translation and lift-up motions, initially developed at the University of Toulouse by Gary Perelman, a PhD student that I co-supervised with Emmanuel Dubois. TDome also includes a camera that can sense the environment. Regarding the touchscreens, we implemented both a Small and Large versions (Figure 3.7): to create the Small version, we enclosed a smartwatch touchscreen into TDome; for the Large version, the user can employ his personal smartphone. To support device modularity, the interchange of both touchscreens is easy and quick (Figure 3.7-right).

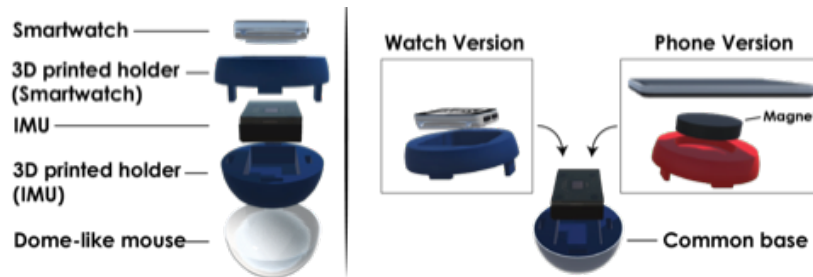


Figure 3.7: Arrangement of TDome elements.

TDome key components (in terms of spatial sensing, input and output interaction, mid-air capabilities and form factor) suit the major MDE interaction requirements [Boring, 2011]: input redirection (i.e. redirect input channels to different displays), output redirection (i.e. move content between displays), physical relationship (i.e. possess high-level information on the spatial layout of the displays), reachability (i.e. interact with a distant display) and personal data management (i.e. personal input and output interaction). We describe hereafter how each TDome component suits these requirements.

Spatial sensing

TDome physical manipulations allow performing 3D pointing in the surrounding space. Combined with the on-board camera, it allows sensing the environment. This can be used to detect and locate nearby displays, creating a spatial layout of the MDE displays represented through a radar-view, as illustrated in Figure 3.7 (physical relationship).

Input interaction

TDome allows up to 3 types of 2D pointing: by moving the device, by rolling it or by interacting with the touchscreen. These ranges of positioning facilitate input redirection.

This also offers input that best suits a given display, such as a cursor for precise tasks, or touch input for coarser input.

Output interaction

The touchscreen display can be used as a visual buffer to move data among displays in MDEs (output redirection). It may also be useful to display a zoomed-in version of a selected area on a distant display (reachability). The built-in vibratory capabilities are an alternative to discretely provide the user with private information (personal data management). Through the easy interchange of the Small and Large TDome versions, the user can adopt the most appropriate display for each task; e.g., to visualize large graphs, the user can choose the Large version, but to display the compact radar-view (i.e. a view of the MDE spatial layout), a smaller display is more appropriate (output redirection).

Mid-air interaction

Two of TDome's physical manipulations (roll and rotate) can be used in mid-air, thus facilitating physical displacements to interact with distant displays (reachability). It also offers more flexibility to the user to ensure the privacy for some of its tasks (personal data management).

Form factor

TDome's tilting capabilities facilitate orienting the device towards oneself for private input and output interaction (personal data management); and attaching their personal smart-phone to TDome's base allows users to access their personal applications and data (personal data management).

3.3.3 Evaluation

In our work we addressed two major challenges for applying TDome in MDEs: first, the device's usability, which demands the user to coordinate a physical manipulation with a touch gesture (we refer to as combined gestures - see Figure 3.8); second, the mapping between TDome gestures and MDE tasks.

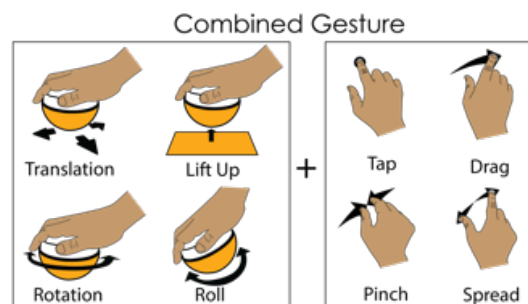


Figure 3.8: TDome combined gestures.

We explored combined gestures involving a physical manipulation (Translation, Roll, Rotation or Lift-Up) followed by a touch gesture (Tap, Drag, Pinch or Spread) through a 2-steps process. First, an exploratory study focusing on comfort established that 60 combined gestures could be comfortably performed. Second, a controlled experiment evaluated the user's performance as well as the subjective perceived difficulty. Results revealed that the

number of gestures that can be precisely and easily performed is 17 with the Small version, and 54 with the Large version.

Finally, a user survey explored the mappings between these gestures and MDE tasks. Results show that some combined gestures are more prone to be used in specific tasks than other. For instance, translation for panning or for moving a focus; roll for private pincode input; pinch and spread were preferred for zooming, drag for sending content from the tabletop to other displays and tap for display selection. In general, we found participants are able to match TDome features to MDE tasks.

3.3.4 Summary

We presented TDome, a device designed for interactions in MDEs. We designed two TDome prototypes: a Small version with an integrated touchscreen and a Large version based on attaching a smartphone. TDome allows versatile interactions that address major MDE tasks, which we illustrated through various proof-of-concept implementations: detect surrounding displays, select one display, transfer data across displays, reach distant displays and perform private interactions.

TDome interaction techniques still need to be fine tuned and future work should compare their performance with a baseline for each MDE task. Theoretically, since TDome integrates the same capabilities as existing MDE devices, we hypothesise that it can perform similarly for each individual MDE task. TDome should however improve the overall performance by reducing homing transition times and promoting the interaction flow. Therefore, beyond individual controlled comparisons, it would be interesting to carry a longitudinal study.

3.4 Interaction with Overview+Detail Interfaces on 3D Public Displays

In this section we present our work on mobile-based interaction with Overview+Detail (O+D) interfaces on 3D public displays. This work results from a collaboration with Louis-Pierre Bergé, a PhD student at the University of Toulouse supervised by Emmanuel Dubois, leading to one publication at MobileHCI 2014 [Bergé *et al.*, 2014].

In this work we focused on the translation task, i.e. how to move the Detail view (displayed on the smartphone) on a 3D environment, the Overview (displayed on the public display), as illustrated in Figure 3.9-left. Generally the user controls 3 degrees of freedom (DOF) to translate the point of view and 3-DOF to rotate the point of view. In our work, we limit user control of the Detail view to a 3-DOF translation. This task is sufficient to explore public 3D content such as museum objects, and it simplifies the task in a public setting, where interaction needs to be intuitive and straightforward.

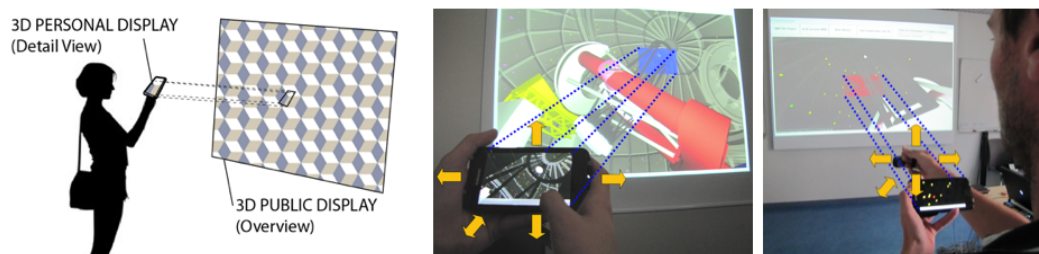


Figure 3.9: a) General setting of smartphone-based Overview+Detail interface on a 3D Public Display. We used two mid-air navigation techniques in a public installation to explore a 3D telescope visualization: b) Mid-Air Phone and c) Mid-Air Hand.

3.4.1 Interaction techniques for translating the detail view in 3D

We explored three different approaches to control the position of the Detail view based on previous work on mobile-based interaction with 3D content: moving the hand around the device [Kratz *et al.*, 2012], moving the device [Boring *et al.*, 2009] or using a touchscreen [Hachet *et al.*, 2009] :

- Mid-Air Hand (Figure 3.10 - a): the position of smartphone serves as a spatial reference. The position of the hand in this referential is mapped to the virtual position of the Detail view. We constrain the movement of the hand to the area behind the mobile phone. A virtual button on the mobile screen (de)activates this navigation mode.
- Mid-Air Phone (Figure 3.10 -b): translations applied to the mobile phone translate the virtual position of the Detail view. As for Mid-Air Hand, a virtual button on the mobile display (de)activates this navigation mode.
- Touchscreen (Figure 3.10 -c): inspired by commercial mobile 3D games, we use two rate-based joysticks to control the virtual position of the Detail view. The left circular joystick controls the 2D translation along the X and Y axis. The right cylinder joystick controls the 1D translation along the Z axis. Both pads can be used at the same time to control the 3-DOF navigation.

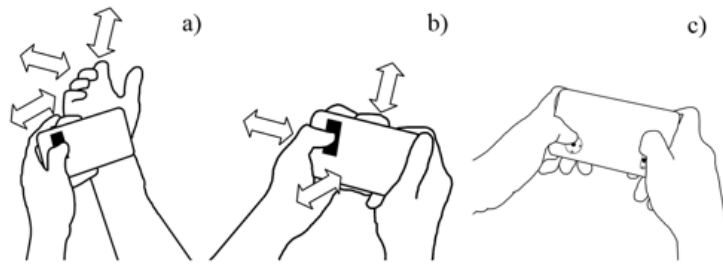


Figure 3.10: Three types of techniques we studied to translate a Detail view on a 3D Overview: a) Mid-Air Hand, b) Mid-Air Phone and c) Touchscreen..

3.4.2 Experimental results

We conducted a first experiment to evaluate the comparative performance of the three techniques presented in the previous section for the 3D translation of the Detail view. We asked 12 users to reach a target on a 3D scene using the Detail view. Our study revealed that techniques based on direct mapping (Mid-Air Hand and Phone) are better than those based on indirect mapping (Touchscreen) for controlling the 3D translation of a Detail view. The study also reveals Mid-Air Hand scores better in terms of attractiveness and user preference, although there is no significant difference concerning SUS score. An interesting result of our study is that Touchscreen, i.e. the most common technique, is the worst in terms of performance, of perceived attractiveness and of user preference.

These results were very encouraging and led us to further explore the two mid-air techniques in a second experiment: the goal was to evaluate the difficulty of performing the two mid-air techniques in usual public context, i.e. without training and without human explanation. Overall, results confirm that mid-air gestures can effectively be used to interact with O+D interface on 3D public display: 91.7% of participants have successfully used the Mid-Air Hand technique and 95.8% the Mid-Air Phone without any training or human explanation. These percentages would probably rise in a public context since users would

be able to imitate other participants as observed in [Walter *et al.*, 2013]. A surprising outcome of our study is that the Mid-Air Hand gesture is more difficult to understand and to perform at first than the Mid-Air Phone. Not only the success rate is higher for the Mid-Air Phone, but it also allows a faster interaction during the first trial. However, our results also reveal that after the first trial, both techniques are comparable in terms of task completion time. Interestingly our study shows that despite the initial difficulty, participants preferred the Mid-Air Hand technique.

3.4.3 Summary

We explored the design space of mobile-based interaction with Overview+Detail (O+D) interfaces on 3D public displays. We evaluated three mobile-based techniques from the literature to navigate in a 3D scene. One of the main findings of our experiments is that mid-air gestures are more efficient and preferred than touchscreen input for 3D interaction. Previous works on mid-air interaction with 2D content showed mixed results: some found mid-air interaction to perform as well as touchscreen [Jones *et al.*, 2012], while others found touchscreen input to perform better [Nancel *et al.*, 2011].

We employed the two mid-air techniques (Mid-Air Hand and Mid-Air Phone) in a public deployment at the University of Toulouse (Figure 3.9). The 3D scene, projected on a public display in the local university hall, represented a large telescope. The goal was to explore the different parts of the telescope and understand how it works. During two days, a large and varied audience (approx. 100 visitors composed of students, teachers and external public) explored the virtual dome of the telescope using these techniques. This in-situ installation revealed that selecting an object in the 3D scene with the hand handling the mobile phone can sometimes be difficult (due to the limited reach of the thumb). Designing alternative selection procedures should be considered in the future.

3.5 Splitting Detailed Views in Overview+Detail Interfaces

In this work we compared the use of different number of detailed views to interact with very large graphs in an overview + detail (O+D) setting composed of a large screen and a mobile device (tablet). This project was carried by Houssein Saidi, a PhD student that I co-supervised with Emmanuel Dubois at the University of Toulouse. This work led to one publication at MobileHCI 2016 [Saidi *et al.*, 2016].

3.5.1 Interface design

We designed and implemented an O+D visualization interface that consists of a large screen to display the contextual information and a tablet to show a magnified version of selected region(s) of the large space. We describe the three main views of our interface (overview, split views and translation view) and we analyze our design with the 8 rules for multiple views defined by Baldonado [Wang Baldonado *et al.*, 2000]: diversity (1), complementarity (2), decomposition (3), parsimony (4), space/time resource optimization (5), self-evidence (6), consistency (7) and attention management (8). All the details concerning each rule can be found in the original paper, we only describe here how our interface design suits these rules.

Split views

Our technique allows the user to have up to four independent split views at the same time, offering a detailed view on a graph region to support tasks requiring focusing on different places of the overview. We implemented three configurations for the multiple views on the tablet: 1-view, 2-views and 4-views (Figure 3.11). Using split views allows to decompose (rule 3) the complex graph rendering.

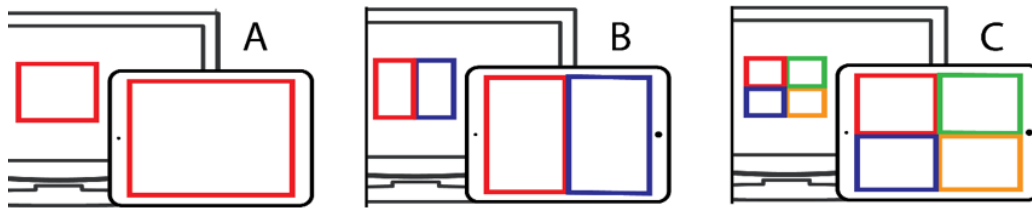


Figure 3.11: We studied splitting the detail view in 1, 2, and 4 split-views.

With the 1-view technique, the split view occupies the entire tablet display; with 2-views, each view occupies half; and with 4-views a quarter. For all of them, the zoom level is always the same, which means that as the number of views augment, the information displayed by each view decreases. This design conforms to the rule of consistency (rule 7) as the overall detailed area size is consistent over the 3 versions of our technique and when several focus are displayed their relative size is consistent as well. It also presents different conditions of space/time resource allocation (rule 5): sequential for 1-view, and side-by-side for 2-views and 4-views.

Swipe gesture inside one of the split views moves the underlying graph in the same direction: this behavior is consistent (rule 7) with regular map interactions on mobile devices. Finally, when the user selects a node in one of the split views, appropriate feedback is provided so that user's attention (rule 8) is focused on the appropriate view.

Overview

The overview displays the entire graph on a large display (Figure 3.12). A contour color is applied to the split views on the tablet and to its representation on the overview to help the user establish the relationship between the points of view (rule 6).

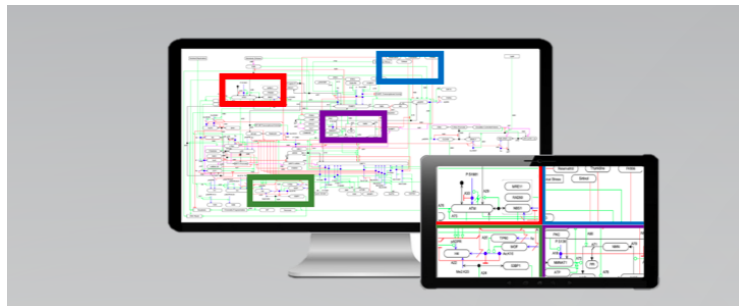


Figure 3.12: Illustration of the overview on a large display and the 4-views configuration on the tablet.

Translation view

Positioning the split views relies on the use of the translation view on the tablet, which is activated when the user presses the "switch" button displayed on the tablet (Figure 3.13). The translation view provides a representation of the position of the 1, 2 or 4 split views on the overview. In the translation view, each split view position is represented using a view icon. Given the density of the graphs, displaying a miniature of it on the tablet would be useless. Therefore, the view icons are displayed on an empty background. By looking at

the overview, the user can use multiple (rule 1) view icons in complementarity (rule 2) for selecting multiple nodes.

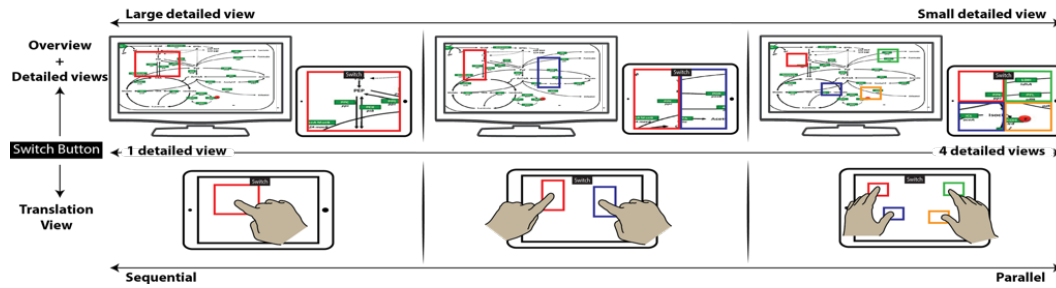


Figure 3.13: Arrangement of the 1, 2, 4 split-views configurations (top) and expected input control of the position of the views (bottom).

The user can adjust the position of one or several view icons simultaneously by direct touch manipulation. Using two hands and the multi-touch screen, the user can theoretically translate 4 view icons at the same time. Closing the translation view restores the split views. In our configuration, no zoom is allowed: this ensures a higher consistency over the split views (rule 7).

3.5.2 Experimental results

We conducted a controlled experiment to evaluate the effect of using multiple detailed views (1, 2 or 4) when connecting various number of nodes (2, 3 or 4) situated on different areas of large graphs. Participants were asked to create a connection between 2, 3 or 4 nodes. Overall, results show that using two or more split views is significantly faster than using only one detailed view. Results reveal that using 4 split views is only better than 2 split views for working on more than 2 regions of the graph.

An interesting finding of our experiment is that, when using 4 split views, users did not take full benefit of bimanual multitouch interaction to translate several view icons at the same time. Most of them (77%) used a sequential approach, first using one finger of each hand to move two icons, and then moving the two remaining view icons. While previous work on symmetric bimanual interaction (where each hand is assigned an identical role) has already highlighted its benefit in some settings [Balakrishnan and Hinckley, 2000] [Moscovich and Hughes, 2008], we are only aware of one work [Geyer *et al.*, 2012] exploring symmetric bimanual multitouch interaction (each finger performs a pointing gesture on a different target). In this previous work, up to 47% of the trials for some tasks were performed using multiple fingers in a bimanual setting. In contrast, our results indicate that symmetric bimanual multi-touch input is hard to perform. We believe these results are dependent on the task and we need to further explore the factors influencing symmetric bimanual multi-touch interaction.

3.5.3 Summary

Our work demonstrated that using two or more split views is significantly faster than using only one detailed view. Some challenges need still to be addressed to improve multi-view interaction. First, how to improve bimanual multitouch interaction to facilitate the translation of several split views at the same time. One idea could be to study combinations of fingers that can be moved synchronously and to help the user in employing these fingers. Second, as most participants used only one finger of each hand, we could consider other

potential uses of the remaining fingers: for example additional fingers might act as modifiers to bring split views together, or to move views to specific positions such as corners.

3.6 Chapter conclusion

In this chapter we presented our contributions to the field of Multi-Display Environments through two main axes. First, proposing novel tools to facilitate the overall interaction on MDEs. We adopted two complementary approaches: using a head-worn display, or creating a dedicated multi-DoF device. Second, we explored Overview + Detail interfaces in MDEs with a focus on two specific contexts: 3D environments on public displays and very large graphs.

Chapter 4

Towards Displays Everywhere: FreeForm Interfaces

Chapter content

4.1	Research challenges	41
4.2	Collecting usage scenarios	43
4.3	Displaying text on freeform displays	44
4.4	Laying out content on freeform displays	46
4.5	Chapter conclusion	50

It is commonly accepted that interactive devices should have rectangular screens and, by proxy, rectilinear interfaces. However, recent breakthroughs in display technologies are enabling the design of non-rectangular displays and interfaces. Such displays are particularly adapted to fulfill the vision of pervasively displaying information anywhere. To this end, our goal is to facilitate the adoption of freeform interfaces that challenge many of the fundamental HCI principles. I started exploring this research challenge in 2016 in collaboration with Pourang Irani (University of Manitoba, Canada) and Anne Roudaut (University of Bristol, UK). The following work led to two major publications at CHI 2016 [Serrano *et al.*, 2016] and CHI 2017 [Serrano *et al.*, 2017].

4.1 Research challenges

The pervasive need to access and make decisions based on data is ushering the next evolution in ubiquitous computing: displaying interactive dynamic content where and when needed, a vision termed as displays everywhere [Pinhanez, 2001]. This vision covers a broad set of application areas, such as smart cities, in-vehicle displays or wearable devices. Displays everywhere can be leveraged to show energy consumption data in-situ in smart environments; facilitate collaborative pedagogy in university classes; leverage interactive presentations in museums, showrooms or shop fronts; or facilitate planning and decision-making during professional meetings.

Displays everywhere can benefit from recent advances in display technologies which allow creating non-rectangular displays. We can divide these technologies into three large groups: electronic systems, multifaceted systems and projection systems. In the field of electronic systems, Sharp recently introduced technologies to design arbitrary 2D display shapes (Figure 4.1-left). Multifaceted systems [Poupyrev *et al.*, 2006] use display primitives to compose larger displays (Figure 4.1-center). Finally, projections can be used to create

nonrectangular displays and can take place on arbitrary surfaces or surfaces with pre-computed geometries (Figure 4.1-right). All these technologies share the common property of allowing the creation of non-rectangular or freeform displays.

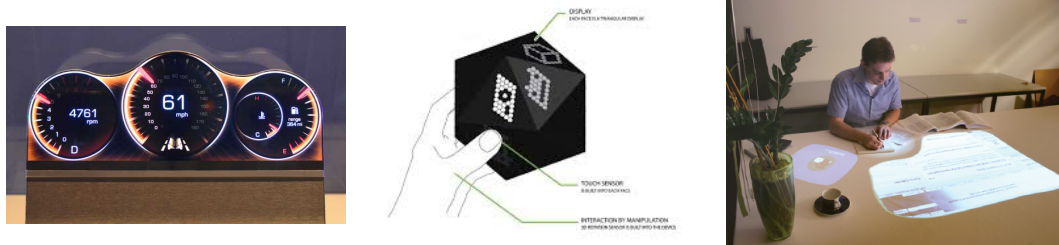


Figure 4.1: Different examples of non-rectangular technologies: Sharp electronic display (left), multifaceted system [Poupyrev *et al.*, 2006] (center) and projection-based display [Cotting and Gross, 2006] (right).

These non-rectangular displays can meet the needs of pervasive displays for which traditional displays are not well suited. For instance, a single non-rectangular display can replace current instrument panels on car dashboards. Non-rectangular displays will also facilitate inserting displays on non-rectangular objects, furniture and/or urban architecture, such as public road signs. In the context of mobile computing, non-rectangular displays can adopt shapes which could better fit wearable devices, allow more ergonomic hand grips or replicate existing jewellery, such as bracelets, pocket mirrors or round smart-watches, which are already commercially available.

However, such novel form factors challenge many of the fundamental Human Computer Interaction (HCI) principles and guidelines that have been accumulated over the past decades for presenting and interacting with content [Dix *et al.*, 2003]. Traditional WIMP interfaces are based on presentation principles such as the use of rectangular text areas and windows; the linear organization of menu bars; or the diagonal UI balance (title at the top-left, and buttons at the bottom-right). Similar principles apply to classical input interaction techniques such as rectangular selection by click and drag, or vertical text scrolling. Many of these principles will not still be valid on novel non-rectangular displays.

Previous work did only partially address the problem of interacting with non-rectangular interfaces. Specific graphical widgets have been developed to solve issues concerning occlusion on display surfaces (e.g. content occluded by the arm, or a physical object), such as nonrectilinear menus or circular layouts around physical objects. However, all these works considered very simple grid layouts (mostly circular) and used rectangular windows, which is space inefficient in a non-rectilinear surface. Cotting *et al* [Cotting and Gross, 2006] proposed mapping content that was originally designed for conventional rectangular displays into freeform bubbles (Figure 4.1-right). This technique is space-efficient but has not been evaluated to confirm its usability. In summary, previous work has not explored how presenting content on non-rectangular interfaces affects reading and viewing usability.

I started exploring the question of how to present content on non-rectangular displays in 2016 in collaboration with Anne Roudaut (University of Bristol, UK) and Pourang Irani (University of Manitoba, Canada). Our first challenge was to convene on the fundamental questions to investigate, since we were facing a vast research space resulting from the combination of multiple types of content and display shapes. Our approach was threefold: first, we carried focus groups to gather both concrete usage scenarios and display shapes; then, we studied text content only, being the fundamental brick of any UI; finally; we explored more complex content layouts combining text and icons.

4.2 Collecting usage scenarios

We decided to start our exploration by collecting usage scenarios of free-form displays in order to generate display shape properties that would inform our choices of shape categories in further quantitative studies [Serrano *et al.*, 2016].

4.2.1 Data collection

We first brainstormed with hci students to capture a subset of compelling shapes in terms of displaying and interacting with content. We ran two focus groups with 20 participants in 2 countries (France and the UK) to maximize the diversity of scenarios we could collect, and to avoid cultural biases (albeit, both countries are dominated by Western culture). We collected 62 ideas depicting 41 shapes once redundancies were eliminated. Most were 2D, and 3D ones were represented using 2D likenesses that corresponded to the user's point of view (e.g. a circle for a sphere).

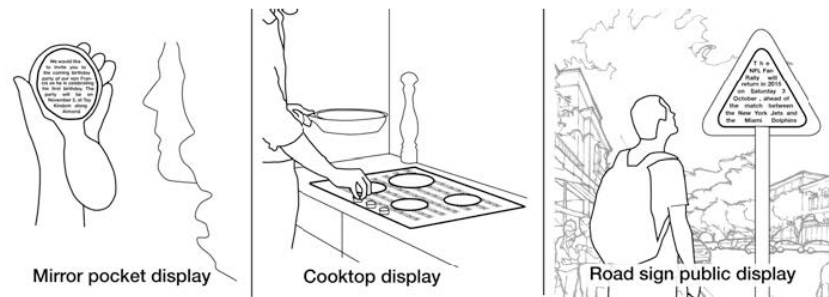


Figure 4.2: Freeform display usage scenarios collected during our focus groups.

The usage scenarios collected during the two focus groups illustrate the diversity of shapes that can hold content, such as circular mirrors for private notifications, shapes with holes such as a cooktop displays for recipes or the back of triangular road signs as public displays, as illustrated in Figure 4.2. Interestingly, in most cases, existing artifacts having nonrectangular features were suggested for content augmentation. Some examples included placing text on road signs, kitchen cooktops, pocket mirrors, puzzle pieces, bike handles, shoes, drink cans, and electric plugs, among others.

4.2.2 Display shape properties

To analyze the collected shapes, we used a clustering algorithm similar to the one proposed in [Roudaut *et al.*, 2014] to create groups of shapes and extract shape properties. From this analysis we observed a set of display shape properties:

- Symmetry: overall there are slightly more symmetrical shapes than non-symmetrical.
- Curvature: most of the shapes are ovoid in nature, for example car side mirrors, purses, sinks, or oval tables. Some shapes have sharp boundaries such as a triangle (road signs), miniature house shape, a tee shirt, or a cooktop.
- Porosity: we found several shapes with holes, such as bathroom elements, electrical plugs, glasses, and cooktops.
- Length-Width ratio: a good number of samples included long and thin shapes such as pencils, faucets, chair arms, or belts.

- Orientation: we observed that some displays had particular orientations, for instance the handle of a frying pan.

In addition to these display properties, we also noted additional interaction observations. For instance, many scenarios involve a display content that needs to be scrolled to access more information, such as the case of the cooking jar or the umbrella.

4.3 Displaying text on freeform displays

After this initial exploration of the usage scenarios and display shapes, we tackled the core concern of how to display text on a screen that is non-rectangular. Reading text is fundamental to many tasks including visually scanning, flicking through a document for specific content or displaying icons. However, running a large study comparing text legibility on multiple shapes is difficult because of the high dimensionality of possible topologies. To address this issue, we propose a text mapping framework.

4.3.1 Mapping framework

We proposed a framework that aims at presenting different mappings of text content onto arbitrary shapes. The framework describes three axes with increasing levels of abstraction. This list is non exhaustive as we only considered text mappings that relate to readability, e.g. we dropped cases with upside down text.

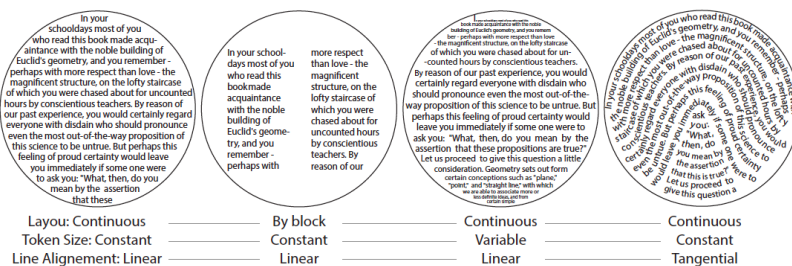


Figure 4.3: Illustrative examples on a circle of different types of text mappings according to the three axes of our mapping framework.

We proposed the three following axes, illustrated in Figure 4.3 :

- **Layout:** this axis describes the general text layout, which can be continuous or by block. For example, the CHI Proceedings layout is in blocks (formatted on two columns). We could have also considered the case where the layout is not continuous (e.g. random), but this would clearly disturb text readability.
- **Token size:** this axis describes the size of the tokens, which can be constant or variable. E.g. the fisheye menu illustrates the case variable.
- **Line alignment:** this axis describes the line alignments in which the text fits. It could be linear, i.e. horizontal, or oriented parallel lines, or what we call tangential, i.e. following the shape. More precisely, text could follow a vector field around the shape boundary. This is typically the case in calligrams.

We then drew on relevant text legibility work and formulated 10 hypotheses to predict how the mappings affect legibility when displayed with different display shape properties. These hypotheses relied on existing knowledge on text legibility, but also extended it

as we were unaware of any study investigating text legibility on non-rectangular shapes. For instance, because we are familiar with reading text that is aligned to the left, we can assume that return sweeps will be more difficult when the text is not left aligned. We then carried a set of quantitative studies on different display shapes to validate or invalidate our predictions. We provide an overview of the studies and the major findings in the following section.

4.3.2 Quantitative studies on reading text

Overview

To examine our 10 hypothesis on how text mapping affects legibility on non-rectangular shapes, we carried four experiments in total (Figure 4.4):

- Experiment 1: We compared shapes with different left or right text alignments.
- Experiment 2: We compared different text layout on various shapes with or without a hole.
- Experiment 3: We compared different token sizes on different shapes. We also wanted to compare the impact of continuous scrolling vs. page scrolling on text legibility.
- Experiment 4: We compared different line alignments on different shapes.



Figure 4.4: Illustrations of some of our experiment conditions. Left, different left alignments (study 1). Center, text layouts on shapes with or without a hole (study 2). Right, tangential line alignments (study 4).

Studies design

These four experiments were based on the same task and procedure. Reading tasks need to be carefully designed so they bear resemblance on how we commonly read. Two primary task options exist. In one case, the post-reading comprehension of users is evaluated using procedures such as the Nelson-Denning reading test. However, this test is designed primarily for gauging reading deficiencies. A second approach consists of seeking spelling mistakes or finding specific words. Such tasks promote skimming. We adopted a task similar to that of Jankowski et al. [Jankowski *et al.*, 2010], which introduces word substitution errors, forcing participants to read and comprehend sentences. Incomprehensible sentences need to be flagged for errors and subjects must read the entire passage to recognize all substituted words. The new words are common words that are placed grossly out of context.

As in prior work [Jankowski *et al.*, 2010], we measured text legibility by both examining reading time and reading accuracy. We focused on short text (150 to 170 words) as a result of our brainstorming sessions. Using longer texts may have shown more differences in results, but small passages are ecologically valid and in line with the scenarios we gathered. A total of 37 people (8 female) with normal or corrected to normal vision took part in our experiments.

Resulting design guidelines

From the results of our studies we can provide a set of design guidelines for optimizing text legibility on non-rectangular displays:

- Both left and right irregular alignments should be avoided, as text in these are perceived to be difficult to read and overall not aesthetic. Instead, symmetric shapes are preferred.
- Shapes with circular or sharp alignments are acceptable for presenting text: they are perceived to be easy to read, and overly clean, beautiful and interesting.
- If the shape contains a hole, text should be displayed using a broken layout with two columns around the hole to prevent any impact on reading performance.
- Shapes without holes are perceived to be less interesting than with holes. Thus, using holes in freeform shapes is not only a solution to context requirements (such as the cockpit), but also an aesthetical feature to explore.
- To use dynamic scrolling on non-rectangular shapes, text should be resized so that each line contains the same amount of text. Otherwise, use page scroll with constant text size.
- While resizing text for dynamically scrolling is perceived as beautiful and clean, resizing text with page scrolling raises mixed results. Some users disliked it because of display space loss and of varying interline spacing. Thus, resizing text should be limited to dynamic scrolling.
- Shapes with continuous line alignment where lines are cut by the shape curvature should be avoided as they are perceived to be difficult to read and non aesthetic. This is similar to the effect of holes on continuous text. Even though tangential alignment does not affect reading performance on linear shapes, continuous text should be preferred as it reduces the perceived difficulty.
- Text on very sharp shapes should be avoided, as text on these is harder to read than on linear shapes. If used, such shapes should be filled with continuous text rather than tangential that impacts reading performance.

4.4 Laying out content on freeform displays

After studying how to map text onto non-rectangular displays, we extended our investigations to more ecological content, i.e. combining text and images, which lead us to address the fundamental question of how to layout content. Our goal was thus to investigate if the established composition principles for traditional displays generalize to non-rectangular displays, and if not how they can be adapted for designers to create the layout of freeform display content.

4.4.1 Elicitation study

We first performed a qualitative study that consisted in asking graphic designers to map traditional web content onto non-rectangular shapes. We recruited 5 professional graphic designers with expertise in print or web design. We gave them a webpage (the home page of The Guardian) with all associated content. We asked them to fill this content into four shapes: a circle and a triangle, with or without hole (Figure 4.5).

All designers found that shapes with hole were more difficult to fill and were particularly not satisfied with designs on the triangle with hole. Designers agreed that the easiest shape was the circle without hole. Besides collecting the designers subjective feedback on

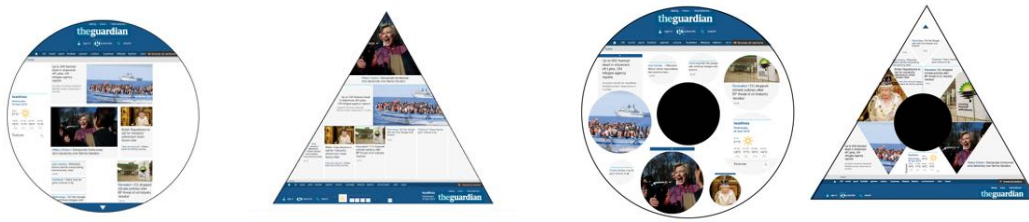


Figure 4.5: Examples of graphical designs collected during our elicitation study.

satisfaction and difficulty, we analyzed the resulting productions using previously existing design composition principles.

4.4.2 Composition principles

Presenting content on rectangular displays is reasonably well understood and there exists numerous composition guidelines. In our work, we analyzed the composition principles proposed by Galitz [Galitz, 2007]: these principles consist in aesthetic composition guidelines extracted from tacit knowledge that visual designers have accumulated over years of experience (e.g. balance, proportion or unity).

Using the probes collected during our elicitation study with graphic designers, we examined how the initial definition of the composition principles could be adapted to non-rectangular displays. As in our previous work on text mapping, we suggested hypotheses on how these existing composition principles would generalize to freeform interfaces, and we evaluated them in a set of studies. The analysis of the results showed that graphic designers' inner sense for composing layouts matches existing composition principles (simplicity, sequentiality, economy, proportion and unity) but that some revisions (balance, regularity, predictability) are needed:

Balance

In the original definition, *balance* means providing an equal weight of elements on each side of the horizontal or vertical axis. There are two aspects that can change the way we define Balance when moving to non-rectangular displays: (1) the symmetrical axes of the display and (2) the definition of a regular shape. For the symmetrical axes, examples produced by our designers suggest that balance should follow the vertical axis. Most element shapes in the productions were rectangular except for the designs shown in Figure 4.5-right. In these two designs the elements shapes were directly related to the shape of the display (the designed cut the elements in circle or triangle).

Regularity

Regularity means providing consistently spaced horizontal and vertical alignment points and spacing, as well as using similar element sizes, shapes and colors overall. Similar to Balance, there are two aspects that can potentially change the way we define Regularity when moving to non-rectangular displays: (1) the alignment axes which can be more than just horizontal/vertical and (2) the definition of a regular shape. Concerning the alignment axes, it is possible to imagine different layouts and to deviate from the rectangular grid: in fact only 10/20 productions used a rectangular grid, 4 used a radial alignment and 6 used a tangential alignment (aligned with one or more edges of the screen). Concerning the regularity of element shapes we observed that designers reshaped the elements for two main purposes: to fit the shape and for aesthetics.

Predictability

Predictability means providing conventional orders or arrangements to ensure that viewing a screen enables one to predict how another will look. Predictability is also enhanced through design Regularity. This guideline rather links the way several pages are designed and thus we have no reason to think that this should change with the shape of the device. To a large extent we can also couple this definition with some of Nielsen's guidelines of "consistency across platform", e.g. the fact that headers and menus are always at the top of a webpage. While all designers decided to keep the regular menu position at the top of the shape for the Circles, most of them inversed the position for the Triangles. This change is rather surprising, given that it goes against traditional web page layouts.

Other composition principles

Our analysis suggested that the other composition principles of Galitz should not change on freeform displays: *Proportion* (using aesthetically pleasing proportions for components of the screen); *Simplicity* (optimizing the number of elements on a screen as well as minimizing the alignment points); *Sequentiality* (arranging elements to guide the eye through the screen); *Economy* (providing few styles to deliver the message as simply as possible) and *Unity* (using similar sizes or shapes for related information). We had no reasons to believe these would change so we did not investigate these principles further.

4.4.3 Experimental studies

As in our previous work on text mapping, we suggested hypotheses on how these three existing composition principles (balance, regularity and predictability) would generalize to freeform interfaces, and we evaluated them in a set of online surveys.

Surveys design overview

To evaluate the differences between the composition principles we opted for paired comparison experiments which consist in asking participants to choose between two conditions, here two layout visualizations. The experiment is designed so that each participant rates each pair of visualizations. Pairwise comparison ratings have been proven to produce more realistic results than asking for individual rankings (e.g. using a Likert scale). We asked participants to compare pairs of layout visualizations and say which one was nicer (i.e. visually pleasing), clearer (i.e. not confusing) and more symmetric (aesthetics terms proposed in [Lavie and Tractinsky, 2004]).

We used the same shapes as in our previous studies, i.e. a circle and triangle, with and without a hole, and we compared visual compositions among shapes, but not between shapes. In a follow-up experiment, we systematically explored shapes with increasing number of edges: triangle (3), trapezoid (4), pentagon (5) and hexagon (6). We also included an inversed triangle to see if the orientation of the shape had any effects. We gave four surveys to the participants matching different hypotheses:

- Survey 1: Balance and symmetry (Figure 4.6). We studied 3 symmetry axes (vertical, shape and all) and 2 element shapes (rectangular or matching display shape).
- Survey 2: Regularity (Figure 4.7). We studied four grid layouts (regular, radial, oriented and random).
- Survey 3: Regularity (Figure 4.8). We tested whether it was better to follow the regularity but have elements cut by the display shape or to break the regularity by having the elements fit the shape. We tested 2 conditions (elements out or in).
- Survey 4: Predictability (Figure 4.9). We changed the position of the menu. We tested 3 positions for the menu (top, bottom and following the shape).

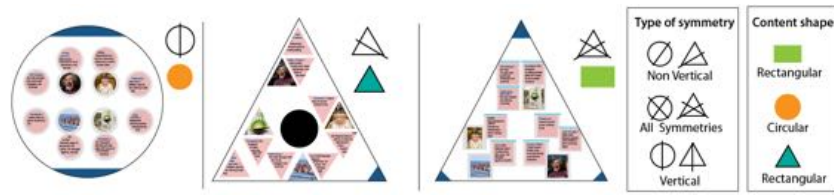


Figure 4.6: Survey 1 conditions.

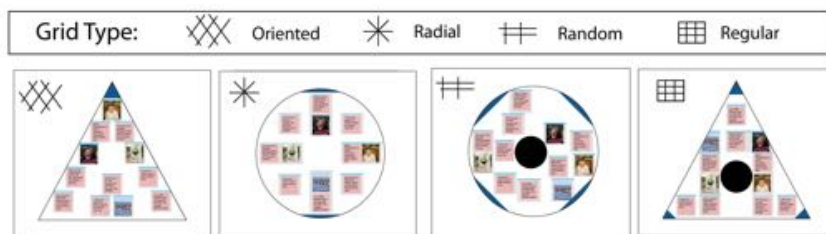


Figure 4.7: Survey 2 conditions.



Figure 4.8: Survey 3 conditions.



Figure 4.9: Survey 4 conditions.

Summary of studies results

Based on our findings we propose a set of guidelines for laying out content on non-rectangular displays. Some of these design guidelines contradict current conventions on rectangular displays.

- Symmetry axis: The symmetry axis should be vertical to ensure that the final design is nice, clear and symmetric.
- Content shape: Instead of using the traditional rectangular boxes for text or images, designers can reshape the content to fit the display (circular on circle, triangular on triangles, etc.). This reshaping will have different effects depending on the display shape: it will look nicer with circular content, or more symmetric with triangular content. However designers should be aware that sometimes reshaping content might make it appear less clear (such as in our triangle condition).
- Grid layout: While using the traditional regular grid works well for certain shapes (regular and inversed triangles), using a grid with the same shape as the display shape can make the overall design look pleasing, clear and symmetric (as with radial grids in circle, pentagon and hexagon displays). A non-regular grid can benefit from non-rectangular content, as it better fits the shape of the grid (triangular content in oriented grid for instance).
- Breaking content: To solve the problem of content not fitting exactly on the display, designers should favor breaking the regularity of the grid and making all content fit, rather than cutting elements by trimming the edges.
- Menu position: While placing the menu at the traditional position on top of the interface works best for triangle and hexagon displays, designers could place it at the bottom in certain cases: this is a position that is nicer, clear and symmetric for a circular display, and that is equivalent to the top position for certain shapes (pentagon and trapezoid).

Generalization to other shapes

Since this work is the first exploration on how visual composition principles apply to non-rectangular displays, we decided to adopt a context-independent approach. We chose to study the generic properties of layout design instead of focusing on a given interface for a given application. The reason is that we wanted to provide generalizable findings rather than narrow in on specific guidelines that would be only valid for a specific case. Our choice of shapes was based on the usage scenarios envisioned for non-rectangular displays. Some of our results seem to be consistent across shapes, such as the fact that shape-like content looks better than a rectangular content, suggesting that they are probably valid for other shapes. Other results seem to depend on the display shape, such as the layout grid: while a radial grid is best for most display shapes (circle, hexagon, pentagon and trapezoid displays), a regular grid is better on triangular displays.

4.5 Chapter conclusion

In this chapter we presented our contributions to tackle the challenge of how to display information on freeform interfaces. The introduction of such displays creates unprecedented challenges for designers who have to rethink new ways of creating user interfaces. The foremost concern is how to legibly present textual content. Our results agree with and extend upon other findings in the existing literature on text legibility, but they also uncover unique instances in which different rules need to be applied for non-rectangular displays. In a follow-up exploration, we studied how traditional content layouts can be adapted to fit

different non-rectangular displays. Based on our findings we propose a set of guidelines, some of which contradict current conventions on rectangular displays. Our work is a first step toward defining new guidelines for the design of free-form displays. We detail some perspectives to this research challenge in the last chapter of this manuscript.

Chapter 5

Research methods for ubiquitous interaction

Chapter content

5.1	Design methods	53
5.2	Prototyping methods	58
5.3	Evaluation methods	62

The most common approach for building UIs is to adopt a user-centered iterative design process. This process involves four main activities [Preece *et al.*, 2001]: identifying needs and requirements; developing alternative designs; building interactive versions of the designs; and evaluating these interactive versions with users. While this process relies on relatively well known methods when building desktop-based UIs, the characteristics of ubiquitous interaction call for rethinking these traditional methods. Without being exhaustive, in this chapter we sum up the methods we used to create the ubiquitous interactions presented in the previous chapters. For some of them, we proposed novel approaches which represented a contribution by themselves (marked with an * in the title of the section).

5.1 Design methods

The first two activities of interaction design, i.e. gathering needs and developing design alternatives, can be challenging when considering ubiquitous interaction. First, the activity takes place in external environments, requiring **in-the-wild observations (1)**. Second, the set of potential interactions and mappings to interactive tasks is usually quite large, leading to the use of novel design methods, such as **design through device composition (3)** or **user-elicitation (or guessability) studies (3)**. Finally, some ubiquitous interaction scenarios are prospective, i.e. exploring future possible usages, making it difficult to ground the design on existing practices. In this case, it can be interesting to gather **design probes (4)** from designers to generate novel usage scenarios. Hereafter we describe how we applied each one of these three methods in our work.

5.1.1 In-the-wild observations

Observing users early in the design process can help designers to understand the users' context, task and goals [Preece *et al.*, 2001]. Observations may take place in controlled environments, such as a laboratory, or in the field, where people are observed performing

their tasks in a natural setting. Such observations can have different degrees of participation, from a passive observation ('outsider') to an active observation ('insider'), where the observer is part of the group under study. As mobile technologies are now common in public spaces, it becomes easier to carry in the wild 'outsider' observations, which can be very valuable to unveil unanticipated usages, difficult to observe in a more active observation. Such unpredicted usages are important to consider early in the design process of ubiquitous interactions, as they usually reveal how people get around current technology, opening new research paths.

To illustrate how in-the-wild observations can be used to unveil unexpected usages, we report on informal observations made during a workshop on Interaction Techniques for Mobile Collocation that I co-organized at MobileHCI 2016 [Lucero *et al.*, 2016], on September 6, 2016 in Florence, Italy. During the workshop we discussed F-formations: the term F-formation (or facing formation) was originally coined to describe the spatial arrangement of people in social encounters. Later, proxemics prototypes have been developed that exploit F-formations to support mobile collocated interactions. Inspired by the work presented at the workshop on dynamic F-formations in non-traditional settings, the workshops organizers together with the workshop participants decided to go out and make exploratory F-formation observations in the wild.

During this in-the-wild observation session, participants (n=12) were to observe and make annotations of anything that would seem unusual or that had not been previously reported in previous studies of F-formations in controlled settings, some of which were discussed during the workshop. Such observations could include information on group sizes, how groups move in an open space, physical distance between people, or their potential use of devices. Participants were split into three groups of four and were asked to observe formations of tourists around the Dome of Florence Cathedral (i.e., Il Duomo) pedestrian area.

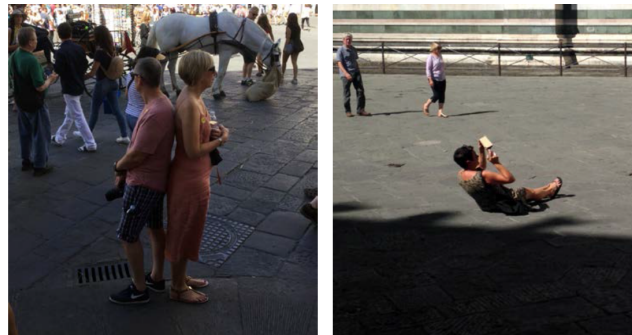


Figure 5.1: Examples of an unusual F-formation (left, back to back) and phone usage (right) captured during in-situ observations.

The results are fully reported in [Lucero and Serrano, 2017]. We observed a mix of tourists and locals doing different activities in this space: tourists walking alone or as small groups, shoppers carrying bags, persons resting by sitting on or lying down on benches, families carrying suitcases or with strollers, people riding or walking next to their bikes, persons walking dogs, mobile street artists selling their work. While we observed activities that are closely related to tourism and thus one would expect to encounter in such a context, we also observed some unusual formations that have not been previously reported in F-formation studies conducted in controlled settings, for instance a standing back-to-back formation (Figure 5.1-left). Another unusual situation consisted of a lady who broke away from a group to capture a picture of the Duomo (Figure 5.1-right). As she wanted to capture as much of the Duomo as possible, or perhaps from a particular angle, she decided

to sit on the ground to take a low-angle shot (Figure 6, right). While this picture does not allow us to say much about F-formations themselves, it does again help us make a point about what we gained by going into the wild to make such observations, as these have not been reported in F-formation studies in controlled settings.

5.1.2 Research through design by device composition (*contribution)

Ubiquitous interaction often implies rethinking existing interaction devices. When coming up with a novel interaction device, HCI practitioners often follow an empirical design approach resulting in the development of ad-hoc solutions, usually combining existing devices, without relying on a systematic or structured process.

To overcome this limitation, we investigated the concept of device composition to promote the potential of combining existing devices to create new ones, and to leverage performing this combination in a more systematic manner. Device composition consists in physically putting together several existing devices to create a new one, hereafter referred to as a compound device. To develop this concept, we defined, illustrated and evaluated a design space for the physical combination of interaction devices, DECO [Perelman *et al.*, 2016].

Our design space, DECO (Figure 5.2), is structured along two dimensions: 1) the physical arrangement, i.e. how the different devices are physically composed and 2) the physical manipulation, i.e. the way a user will manipulate each element of the compound device.

		Physical arrangement			
Topology		Enclosed	Glued		Separated
Fusion type			A-B	A<-B	A<->B
Temporal usage	Sequential	GlobeMouse	Hexagonal bezel-less ★	Trackball Keyboard	Ink222
			FlipKeyboard	LensIMouse	
	Parallel	Lexip 3D Inflatable Mouse	Codex	PadMouse	Siftable
			Razer naga Dualscreen HybridTouch	CubicMouse Mouse 2.0 RPM2 ★	HoverPad RPM2 ★

★ = Dynamic compound device

Figure 5.2: Different compound devices from the literature classified using DECO.

To validate our design space, we used the approach proposed by Beaudouin-Lafon [Beaudouin-Lafon, 2004] to evaluate design models. This approach is based on three properties characterizing the ability of a model to describe existing solutions (descriptive power), compare existing solutions (evaluative power), and generate novel solutions (generative power). Using DECO, we classified and compared existing devices to illustrate how our design space helps in describing and comparing different solutions. We also used this design space to elaborate a novel compound device supporting multi-dimensional interaction: this device results from the combination of a spherical mouse, the Roly-Poly Mouse [Perelman *et al.*, 2015] with a traditional laser mouse. Through our 3-part exploration, we demonstrated that DECO is a useful design space that can be used to describe, compare and generate novel compound devices.

5.1.3 Elicitation studies: exploiting similarity over agreement (*contribution)

To explore the breadth of potential interaction gestures and their mapping to interactive tasks, one approach is to elicit user input through an elicitation (or guessability) study. For

exploring potentially rich and vast gesture sets, user elicitation or guessability studies have shown favourable results. Wobbrock et al. found that eliciting gestures from users resulted in over 40% more gestures than if asked by expert designers [Wobbrock et al., 2009]. This motivated the use of such an approach to identify gestures for multitouch tabletops, mobile motion gestures or for foot interaction among others.

During our exploration of Hand-To-Face gestures [Serrano et al., 2014a], we asked participants to suggest suitable gestures on the face (above the neck) and on the HWD. We put aside any recognizer issues and asked users to perform gestures at their will without worrying about the underlying sensing technology. We asked participants to include gestures on the entire face, i.e. any region on or above the neck. This allows for a larger set of potential gestures. We also assessed users' preference for interacting with either the face or areas of the HWD, for each of the tasks.

We analyzed the agreement between participants for the set of gestures produced for each task using Wobbrock's approach [Wobbrock et al., 2009]. The agreement value ranges between 0 (no agreement) and 1 (total agreement) and indicates whether users agreed on using a specific gesture for a given task. We group gestures which are of the same type (swipe, tap, etc.) and occur in the same area (cheek, chin, etc.). The mean value for the agreement score was 0.14 (SD 0.06), with 36% of the tasks having an agreement value higher than 0.2%. While this score seems low, it is on par with that from other previous guessability studies. This low agreement score is mainly due to the variety of areas used for zooming (ear, hair, nose, neck, mouth, jaw, forehead, chin and cheek). Swiping the cheek is considered different than swiping the forehead. To solve this issue, we suggested **using similarity instead of agreement**.

To this end, we first proposed a taxonomy to describe hand-to-face gesture properties to evaluate gesture similarity. Our taxonomy includes five properties to describe the gesture mapping and its physical characteristics (nature, temporal, pose, number of fingers and area). Based on the previous taxonomy, we defined a formula to calculate the similarity score, which indicates whether different gestures share common properties. The similarity score St of a task t is the average of the agreement for every property Pi of our taxonomy, from the set of properties Pt . To calculate this value we use the formula of the agreement score: Gi is the subset of gestures with identical value for the property Pi from the set of gestures Gt .

$$St = \frac{\sum_{pi} (\sum_{Gi} (\frac{Gi}{Gt})^2)}{Pt} \quad (5.1)$$

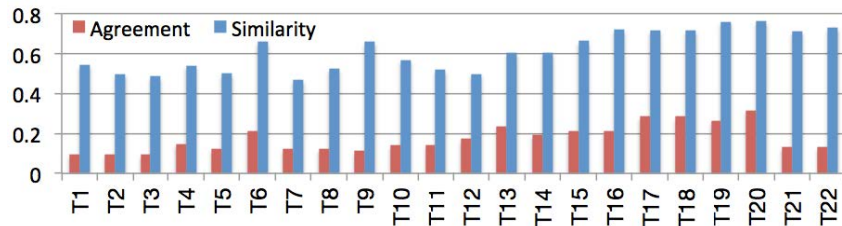


Figure 5.3: Comparison between agreement and similarity scores.

Using this formula, we gathered some more informative results on the proposed gestures. The mean value for the overall similarity score (figure 5.3) is 0.61 (SD 0.1). The Nature of zooming gestures (0.86) is mainly based on the metaphor of pinching with two fingers. The Pose for zooming is always dynamic (1.0) and most subjects used two fingers for zooming (0.80). Actually, our similarity analysis allows also to describe gestures more precisely: we can describe panning gestures using our taxonomy as <abstract, continuous, dynamic,

one finger, cheek/forehead> and zooming gestures as <abstract discrete/continuous, dynamic, two-fingers, cheek/chin/forehead>.

Our approach only addressed one issue of guessability studies, i.e. the limitation of using the agreement as unique score. This problem was also pointed out in a recent work [Tsandilas, 2018] that identifies other problems of the analyses methods for these studies (such as the interpretation of the agreement values or the statistical inferences) and suggests some alternative solutions. As practitioners in HCI, i.e. a relatively recent research field, it is our responsibility to question and update existing research methods, particularly when they are as widely employed as guessability studies.

5.1.4 Design probes

The first activity in interaction design is to identify the needs and establish requirements. However, one of the key challenges when designing prospective ubiquitous interfaces is to identify the needs that will emerge from implementation of novel ubiquitous services. In such cases, working with designers can help generating novel usage scenarios, leading to design probes that will feed the analysis on needs and requirements.

We adopted this approach when working on non-rectangular displays [Serrano *et al.*, 2017]. To begin our exploration on the visual composition of graphical elements on non-rectangular displays, we first captured how graphical designers tacitly organize visual layouts. We gathered qualitative probes that we could use to generate new hypotheses for the visual composition of elements on free-form displays.

We asked five graphic designers to compose webpages on non-rectangular shapes (Figure 5.4). This task relies heavily on creativity and thus we designed this study to be in the form of homework. Designers had a week to do the task wherever they wanted, thus avoiding us to interfere with any creative processes that might emerge from their environment. We then analyzed the productions to better understand which choices they made. We analyzed how their designs can be generalized to non-rectangular displays and proposed a set of hypotheses, later evaluated through a user study.



Figure 5.4: Design probes on non-rectangular interfaces.

The previous design probe was driven by a concrete research question: designers were given a precise task to complete. Another approach when conducting prospective explorations is to give designer the freedom to propose novel usage scenarios and designs based on a general concept. We adopted this approach during a 1-week workshop in January 2018 with design students from the Design Master DTCT (Design Transdisciplinaire, Culture et Territoires) of the University Toulouse Jean Jaures. The workshop, entitled "#Stranger-Screens", was preceded by a one-day seminar, where researchers and practitioners from different fields (HCI, design, hardware) presented their work, as an inspiration for the design students. Then, the students were given three days to generate ideas and craft one design proposal (Figure 5.5). These designs were presented the last day of the week, opening different perspectives on the usage scenarios, contexts and shapes of non-rectangular displays. The quality, diversity and depth of the proposed scenarios were beyond what

usual brainstormings would produce, showing the value of carrying design probes with designers.

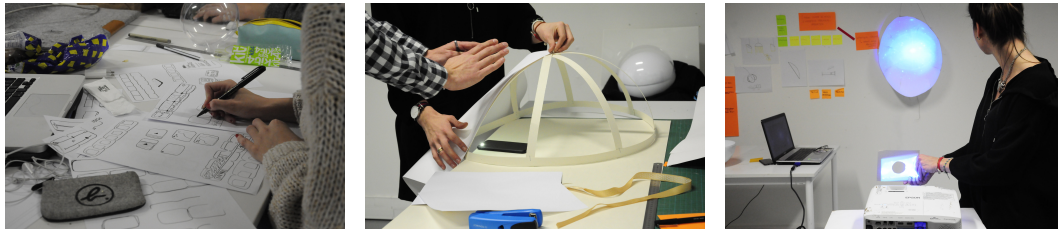


Figure 5.5: Design workshop on non-rectangular interfaces.

5.2 Prototyping methods

Prototyping ubiquitous interaction often requires combining novel interaction modalities, which can be difficult in a rapid iterative design process. Rapidly designing and creating new devices is also complex, in particular due to the lack of proper design and ideation methods. Finally, research on ubiquitous interaction often requires to experimentally validate interactions which can not be easily implemented at the current time (such as on-body gestures for instance). To solve these three issues, in this section we present our work on rapid prototyping tools and the methods we used to develop proof-of-concept pervasive prototypes.

5.2.1 Rapid prototyping tools (*contribution)

During my PhD, we were specifically interested in multimodal interactive systems that include several interaction modalities used by the user (input modalities) in an independent or combined way. My doctoral research was dedicated to conceptual and software tools for efficiently exploring the huge set of possibilities during the design phase of input multimodal interfaces [Serrano, 2010]. We presented a component-based approach for designing and prototyping functional and/or simulated input multimodal interfaces. Our approach relied on the data-flow, from input devices to the user-tasks, depicted by an assembly of components.

Our work was both conceptual and practical. Our conceptual model consisted of a characterization space for describing components. This space helps interaction designers to define component assemblies and to consequently consider different design alternatives for input multimodality. Our software tool is a component-based prototyping tool that implements the characterization space (i.e., our conceptual model). We implemented several multimodal prototypes using this tool, involving a large variety of interaction modalities including 2D gestures, 3D pointing techniques, speech commands, head and body movements, geolocalization techniques as well as tangible techniques.

5.2.2 Proof-of-concept prototypes

A proof-of-concept (PoC) is the realization of a certain method or idea to demonstrate its feasibility. In HCI, PoCs are used to fill an interactive technology gap, with the idea that if the PoC validates the concept (for instance in terms of usability), future work will ultimately fill the gap. In this section we describe how we implemented and applied proof-of-concept prototypes to evaluate ubiquitous interactions that could not be easily implemented otherwise.

On-body input

We explored hand-to-face gestures for HWDs without emphasizing the technology that would ultimately support this style of input. Several experimental options exist to implement this on-body input, such as a camera mounted on the HWD, body-implanted sensors or instruments worn on the finger.

To evaluate our interaction (see Chapter 2.3), we implemented a proof-of-concept prototype using a Vicon T20 infrared optical tracking system with six cameras positioned around the user (front, front-right and right side at different heights). We placed infrared markers on the participant's index finger (Figure 5.6). To detect skin contact, we used a proximity sensor connected to a micro-controller. The sensor set was connected through a USB cable to a desktop computer. To make the sensor set unobtrusive, we integrated it into a glove worn by the user. During our studies we had a negligible number of tracking errors (0.15% of all trials). The system had no perceivable latency: all input was merged to the same program and sent to the HWD through USB. Optical tracking ran at 690 Hz; contact sensor at 600 Hz; merged data were sent to the HWD at a measured rate of 142 Hz. While this solution would obviously be impractical for real use in mobile context, our solution was sufficient for carrying in-the-lab studies and demonstrating the utility of on-body interaction.

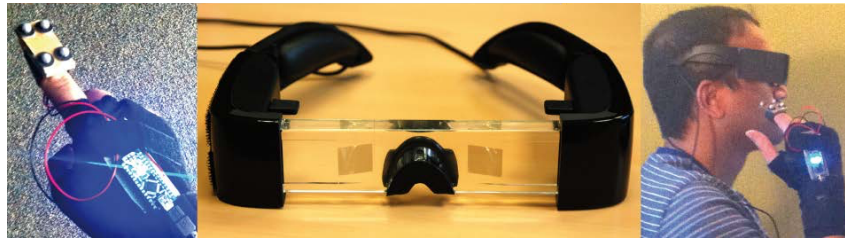


Figure 5.6: Experimental implementation of hand-to-face gestures: IR markers and micro-controller on hand; Epson Moverio HWD and subject wearing HWD.

Finger identification

Detecting fingers individually is a difficult task, and no touch technology efficiently supports finger identification yet. During our work on how visually impaired (VI) people explore raised-line diagrams [Bardot *et al.*, 2017], we needed to accurately track hands and fingers movements, since previous studies showed that VI people use both hands and different fingers. Our approach needed to be unobtrusive, since we did not want to limit or influence how participants used their hands and fingers. We adopted a colour tracking camera-based approach: we used a Logitech C270 webcam (1280x720 px) located above the touch screen in order to track the ten fingers according to coloured markers placed on each nail (Figure 5.7). The acquisition rates were 50 Hz for the camera and 100 Hz for the touchscreen. While this approach has some limitations, such as the number of possible colors that can be used without reducing the tracking accuracy, it proved to be sufficient for our study in a light-controlled environment.

Holographic displays

Holographic or true-3D refers to any 3D digital display capable of producing mid-air, full-depth-cue (or volumetric), multi-angle and/or multi-user images without the need for user instrumentation. During our exploration of true-3D interaction around a mobile device [Serrano *et al.*, 2014b], we needed a way to emulate such a display for a single user, which

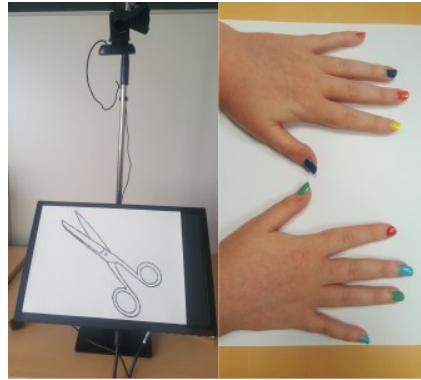


Figure 5.7: Experimental setup: a raised-line drawing is placed over the touchscreen. A camera located above the drawing tracks fingers movements.

could be instrumented for our study purposes. Our implementation was based on a stereoscopic display coupled with head tracking, on the VisCube platform, an immersive environment composed of a projection wall and floor (Figure 5.8). In this system the user has to wear polarized glasses with IR markers to allow visual head-tracking. 3D content was developed using GLUT. The position of the mobile device in the environment was tracked using a Vicon IR motion tracking system and IR markers.

The results of our experiments are influenced by the technology we used to emulate the true-3D displays. The obvious differences between this technology and the final true-3D display, in terms of color, brightness or 3D perception, may alter the results from our experiments. However, most of those results are strongly influenced by human physiological limitations on wrist-based rotation and arm reach. Thus we think these technical differences do not have a fundamental impact on our findings. Moreover, researchers have used such platforms for developing and testing novel technologies.



Figure 5.8: We built our true-3D prototype in a VisCube 3D immersive system, tracking the mobile device using optical tracking.

5D Mouse prototype

Building a working version of our spherical 5D mouse, the Roly-Poly Mouse (RPM) [Perelman *et al.*, 2015], presented several technical challenges to properly track its position, orientation, and user input (i.e. a selection mechanism). To track RPM **position** several embedded and non-embedded solutions are possible. Embedded solutions include using magnetic sensors, although these are not precise enough to allow mouse-like pointing.

Not-embedded solutions include using IR cameras (similar to the Vicon system used for on-body input) or an underlying sensitive surface, similar to Wacom's tablets. All of these solutions would allow detecting the z-dimension at a certain level. To track the **orientation** of the device, we could use an embedded inertial measurement unit (IMU).

There are two main options to integrate an always-available **selection mechanism** invariant to rotation: an all-around ring button and a capacitive surface. The ring button, situated all around the device, permits multi-touch input with several fingers. A more elegant solution would be to use a multi-touch surface. However, this solution would lack of haptic feedback, like on a regular button. This solution has proved to be useful for extending a traditional mouse such as Apple's Magic Mouse in which it is combined with a mechanical switch. The main challenge with this solution would be to distinguish a finger touch from a palm contact or with a grasping/squeezing gesture.

For our initial exploration of the Roly-Poly Mouse (RPM) gestures [Perelman *et al.*, 2015], we used infrared optical markers tracked by 12 OptiTrack cameras (1mm precision) to track the translation, rotation and roll of the device. The system senses the position (x, y, z) and orientation (yaw, pitch, roll) of RPM at 100Hz. We placed IR markers on the RPM to allow the cameras detect the device without impeding the user's ability to grab the device with different hand postures. Informal tests had also confirmed that the marker did not limit the amplitude of comfortable rolls: the maximum possible roll of RPM given these physical markers was 70° in the marker support directions. Our tracking setup did not register contact with the underneath surface, thus we could not use clutching in our experiments. To demonstrate the feasibility of such a device, we created a working an integrated version using a Polhemus Patriot Wireless tracker, which is however less precise and robust than the infrared solution.

In our next evolution of RPM, called TDome [Saidi *et al.*, 2017], the device holds an x-IMU of x-io Technologies to detect the Roll and Rotation of the device in 3D. The IMU is composed of a triple-axis gyroscope, accelerometer and magnetometer. The refresh rate of the sensors goes up to 512Hz and we used Bluetooth to connect the IMU with the computer. The IMU offered an angular precision of 1°. We 3D printed a holder to fit the IMU in a horizontal position inside the TDome. To detect the displacement of the device, we used an infrared bezel that generated TUIO events. We implemented a filtering process to discard touch events that were detected when fingers touched the surface around the device.

In our latest work, which explored the use of RPM to facilitate selecting toolbar items [Dubois *et al.*, 2018], we enhanced the original RPM. As the original one, the version consists of a sphere with a diameter of 8 cm, which includes a Bluetooth enabled Inertial Measurement Unit (the same xIMU by xIO Tech). In comparison to the original RPM, our version was placed on a Wacom Intuos 3D tablet (216x135 mm, resolution: 2540 lpi). As the tablet is multitouch, it can detect the translation of RPM and finger taps. We therefore covered the RPM surface with a graphite lacquer to give the device a conductive coating. We tried to insert various forms of button on top of RPM. However our pre-tests showed that using a physical button on RPM altered the device handling gesture and brought a number of technical issues (button position, etc.). Instead, we considered the use of a tactile surface underneath RPM to detect a user's finger tap: the user can employ any finger of the same hand that manipulates RPM to tap on the surface, although participants seemed to prefer the thumb. An algorithm associates the first touch on the tablet to the RPM position, and triggers a tap event only when detecting a second touch. Alternatively, the user can press a key on the keyboard with the non-dominant hand: as the user's main task is probably involving keyboard input, this bimanual setting offers a fluid interaction compatible with regular keyboard input (the keyboard is only used as a validation).

In conclusion, we implemented various versions of the Roly-Poly Mouse (Figure 5.9), which illustrates the interest of using in-the-lab prototypes to evaluate interactions. Each iteration demonstrated the benefits of the device for a different task or context, which motivated us to pursue our research agenda on this device and to improve its implementation.

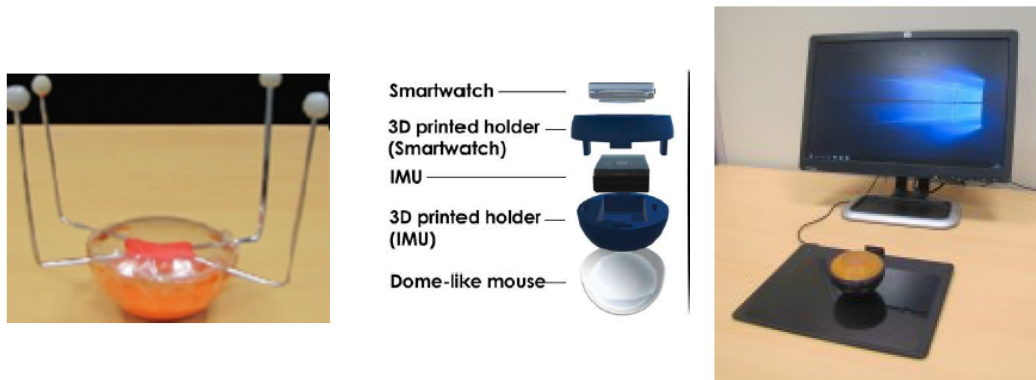


Figure 5.9: From left to right: first version of RPM using optical tracking; second version integrating an IMU to track rotation and orientation; last version additionally used a Wacom tablet to track its position.

5.3 Evaluation methods

In this section we sum up two evaluation approaches that we carried. The first one concerns the evaluation of interaction modalities that are still not implemented using a component-based wizard-of-oz approach. The second consists on a type of evaluation that we believe has not yet got full attention from the HCI community, i.e. subjective feedback gathering through pairwise comparisons.

5.3.1 Component-based Wizard-of-Oz (*contribution)

To accelerate the prototype-test cycle at the early stage of the design, one solution is to adopt a Wizard-of-Oz approach. As explained in [Davis *et al.*, 2007], "Wizard of Oz prototype is an incomplete system that a designer can simulate behind curtain while observing the reactions of real end users". WoZ studies have been shown to allow fast prototyping of interactive applications by allowing the evaluator (wizard) to simulate missing functions.

Following our work on the OpenInterface framework, while the rapid development of multimodal prototypes is possible, testing such prototypes raised further technical and experimental issues. To overcome these limitations, we proposed OpenWizard [Serrano and Nigay, 2010], a component-based approach for the rapid prototyping and testing of input multimodal interaction. OpenWizard is based on a simple idea: the designer and the developer should be able to work both on the multimodal prototype and on the WoZ experiment at the same time.

We illustrate this approach using the puzzle metaphor (Figure 5.10 - left). A multimodal interaction prototype can be seen as a puzzle, where each piece corresponds to a physical device, a transformation algorithm, a fusion algorithm or a task. The designer and the developer assemble those puzzle pieces in order to create the complete puzzle, their multimodal application prototype, that can then be evaluated. The aim is to avoid losing time and effort in waiting for missing pieces and in creating and implementing pieces that will then be discarded during the evaluation.

If we translate this idea into our component-based approach (Figure 5.10 - right), we simply have to replace puzzle pieces by software components. The designer and the developer can then build multimodal prototypes easily by assembling functional components (devices, transformation, composition, task) with WoZ (non-functional) components. To do so, we define WoZ components as parts of the component-based approach for rapidly developing multimodal prototypes. WoZ components are characterized according to the

roles that a WoZ component can play in the data-flow of input multimodal interaction, from devices to tasks.

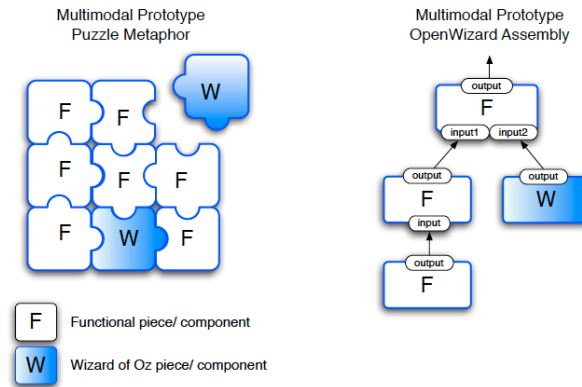


Figure 5.10: Definition of a multimodal prototype using the component-based WoZ approach. Left: concept illustration through the puzzle metaphor. Right: its implementation on our component-based approach.

5.3.2 Pairwise comparisons

Quantitative controlled experiments are assumed to be the best tool to demonstrate UI efficiency. But they are only useful when it is possible to test counterbalanced variables without introducing confounds. In some situations, it is not possible to find an experimental setup following this rule. For instance, in readability studies, the issue is that text presentations affect readability. Thus any effects observed could also be a result of the text presentations that change according to the conditions (confound variables).

A way to get participants' input is to use subjective judgement. Estimating preferences based on subjective judgements is a critical step in psychological experiments with applications in many fields such as marketing, environmental sciences and health economics. In particular pairwise experiments have been widely used. In such studies, two conditions are presented to participants who then indicate one alternative over the other. Pairwise comparison ratings have been proven to produce more realistic results than asking for individual rankings (e.g. using a Likert scale).

During our exploration of visual layouts on non-rectangular interfaces [Serrano *et al.*, 2017], we asked 57 participants to compare pairs of layout visualizations and say which one was nicer (i.e. visually pleasing), clearer (i.e. not confusing) and more symmetric. Participants could give three answers for each question: Visualization-1, Visualization-2 or Both. Our analysis of the results consisted in three steps:

Step 1: Individual consistency checking. We computed the Transitivity Satisfaction Rate (TSR), which quantifies the consistency of a participant's judgments over multiple questions. E.g. if A is found more restrictive than B, and B more than C, then we should have A more restrictive than C. We removed 4 participants whose TSR was below 0.8. The mean TSR for all other users was 0.92 (SD = 0.05) and at least over 0.8 for all of them, thus denoting that they paid full attention to the study.

Step 2: Overall consistency checking. To test the overall consistency across participants we checked the stochastic transitivity properties or computed Kendall's-coefficient. For each participant, we computed a list of rankings of visualisations and used the kendall-tau Python library to produce a coefficient for each pair of participants, computed as a per-

centage (100% means all participants perfectly agreed, 0% they perfectly disagreed). Our results show that the mean Kendall's-coefficient is above 50% for all conditions.

Step 3: Model the data. The individual and overall consistencies were confirmed, so we proceeded to model the data. We used the Bradley-Terry-Luce model, which associates an ability metric to each condition that have been paired-compared as well as the p-value for each pair comparisons. Note that the Bradley-Terry-Luce model computes a p-value that express how the visualizations compare to one specific visualization only, which serves as reference and is a parameter of the formula. We thus performed several tests to compute the significant level for each comparison. To counteract the problem of making multiple comparisons tests we used a Bonferroni correction for each result.

To report the results, we represented the metric of each visualization computed via the Bradley-Terry-Luce model. The metric gives a value between 0 and 1, where the lowest condition equals 0 and the highest 1. We also indicated the standard error values given by the model.

In conclusion, using a pairwise comparison study allowed us to evaluate the subjective value of layout visualizations, which would have been very difficult otherwise. This shows the importance of looking for evaluation methods from other fields (here in psychological experiments). The main challenges in this case are to find the appropriate methods, apply them properly and explain to HCI practitioners and reviewers the value and validity of such methods.

5.3.3 Conclusion

Our work on designing novel interactions for ubiquitous computing led to some advances on design, prototyping or evaluation methods. Each project implied not only thinking about our contribution to HCI, but also about our research procedures. Research on HCI is a relatively young domain, at the intersection of different fields which have well known methods (such as AB experiment design on psychology). As HCI matures, some of our methods will become established and hopefully applied more methodically, improving the overall quality of our research.

Chapter 6

Perspectives

Chapter content

6.1	From digital to physical ubiquitous displays	65
6.2	Immersive data analytics to explore spatio-temporal data	67
6.3	Pervasive freeform interfaces	68

Our work has explored different types of ubiquitous interfaces (mobile, multi-display and freeform) with the general goal to facilitate the exploration and manipulation of complex data (large number of commands, spatial, 3D and volumetric data). As these types of interfaces slowly populate our surroundings, they reshape our city environments. While the term "smart city" englobes a variety of concepts, technology is one of its core factors. According to [Chourabi *et al.*, 2012], "smart computing refers to a new generation of integrated technologies that provide IT systems with real-time awareness of the real world and advanced analytics to help people make more intelligent decisions about alternatives and actions".

As these cities get richer in terms of collected data, it becomes crucial to better interlace interactive data exploration in our everyday spatial environment. This general idea drives our three perspectives. To this end, we propose to investigate three complementary approaches: tying data to physical objects (1), facilitating immersive data analytics (2) and supporting the rise of pervasive freeform interfaces (3).

6.1 From digital to physical ubiquitous displays

In our work we have investigated the exploration of virtual content, such as spatial data, 3D environments or very large graphs. The next evolution in ubiquitous data analytics could consist in taking advantage of data physicalization. Data physicalization is an emerging area of research [Jansen *et al.*, 2015], which can have obvious advantages for visually impaired users, but could also leverage embedded data analysis, i.e. accessing data in the spatial environment where it belongs or where it was collected.

Physical interactive maps for visually impaired users

Providing a visually impaired user with a way to independently explore or construct physical representations of visual graphics could be invaluable. Tangible interfaces are, in this sense, particularly relevant, as they provide a way to interact with digital information via the manipulation of phicons (i.e. physical icons [Ishii and Ullmer, 1997]), and therefore also provide a way to translate digital information into a physical form.

To this end, previous work has mostly focused on rendering punctual symbols tangible (i.e. physical and associated to a digital content). For instance, [McGookin *et al.*, 2010] developed a tangible interface for the exploration of line graphs and bar charts by visually impaired people: phicons are placed in a physical grid to represent the top of a bar or the turning point of a linear function. Authors observed that the objects were regularly knocked over during the exploration and hence provided a few recommendations concerning the design of phicons for visually impaired users, which need to be stable.

Such an approach is limited for the construction of most graphics that include lines and not only points. As a first step to overcome this limitation, our group at the University of Toulouse designed and constructed a novel type of physical icon called Tangible Reels, presented at CHI 2016 [Ducasse *et al.*, 2016]. A Tangible Reel is composed of a sucker pad that ensures stability, and of a retractable reel that is used to create physical links between phicons. We also developed a non-visual tabletop interface that enables a visually impaired user to build a tangible map using Tangible Reels and to interact with it.

This contribution is a first step and still has limitations: it only allows to render graphics composed of a limited number of points and lines. There is a need to design an interface that is accessible to visually impaired users without any assistance, and that allows them to render tangible any digital graphical content that is composed of points, lines, surfaces and volumes, as well as dynamically explore and annotate them. As fabrication methods (3D printing, laser cutting, etc.) become more affordable and easier to put into service, they could represent an interesting perspective to fill this gap. The main challenge then is how to make these fabricated graphics interactive and accessible for the visually impaired. Two major approaches seem to emerge: either integrating the interactive mechanisms in the physical object itself, or have users wear the interaction devices, as in our previous work with smartwatches.

Public interactive physical models

Besides the interest for visually impaired people, embedding data into physical objects and into our surroundings is becoming more important as we move toward an interlaced reliance on data for our daily activities [Piper *et al.*, 2002] [Raskar *et al.*, 2004]. Such embeddings allow for in-place or in-situ analytics [Elmqvist and Irani, 2013], but also constitute effective communication tools when presenting information to non-expert users [Jansen *et al.*, 2013]. While technical capabilities, including spatial augmented reality [Bimber and Raskar, 2005] or on-object interactions [Zhang *et al.*, 2017] are maturing, there exists little discussion on the granularity levels for interacting with such datasets in public places.

We can illustrate this need for embedded data interaction with a specific use-case: architectural models. Digital architectural models are growing in complexity as architects use these for recording detailed building specifications (material types, 3D representations, structural monitoring sensors), all of which are contained in standardized documents, such as the Building Information Model (BIM). Such large volumes of data often require specialized expertise to comprehend and are inaccessible to non-experts.

An alternative for exploring this data is to embed a selected subset of relevant information only, into their physical counterpart in which unnecessary details are hidden. It thus involves a degraded version of the data and digital model enclosed into a traditional BIM approach. Such physical models are typically on display to showcase a newly constructed building or at a museum as a cultural exhibit. Within this use-case, there is a need to propose and explore embedded data interaction with such architectural models.

There exist numerous challenges to embed interactive data into physical models. First, constructing the physical model needs to be quick to support flexibility and allow for a rapid update on the dataset; second, linking the physical model to its associated digital information must be easy to allow a range of end-users to use such a system; third, interacting with the embedded information needs to account for the environmental constraints

(for example, models that are encased in a vitrine should still be interactive); finally, the presentation of the information should be easily visible and should clearly link the data to the physical model. Our approach to tackle this challenge will be first to focus on the tools for the design and prototyping of interactive physical models; then we will explore interaction techniques to access this physical data in a public setting, with a specific focus on tactile glass cases. This work will be carried by a PhD starting in 2018, funded by the neOCampus project and the Occitanie region.

6.2 Immersive data analytics to explore spatio-temporal data

A complementary approach to the physicalization of data consists in exploring virtual data in an augmented reality environment, a paradigm called "immersive data analytics" [Chandler *et al.*, 2015]. Recent technological advancements in augmented reality devices open up new opportunities for users to interact in their environments. Systems like the HoloLens, MetaVision or Moverio allow the user to display numerical data and visualisations directly on the physical world by attaching them to a fixed physical anchor. These technologies offer the user interaction opportunities that are to this day insufficiently explored. We do not have an implicit design rules to guide the developer towards solutions for these environments. This results in a compilation of partially satisfactory solutions for interaction.

There is a need to provide users of augmented reality environments with rich solutions to explore complex data. Previous solutions for mixed reality mostly considered a limited number of degrees of freedom. For instance, in our work on spatial data exploration using hand-to-face input, we only used 2 degrees of freedom to pan or zoom. This approach no longer stands when exploring complex data such as spatio-temporal datasets. Spatio-temporal datasets are generated daily, whether to collect disease movements, migration patterns, or other movement data. Their visualization usually includes animated representations of space-time cubes [Bach *et al.*, 2014]. When facing such rich and complex data environments, the exploration and interpretation of such data is crucial for acting on the systems, such as, for example, to limit the spread of a disease or to study traffic patterns. For such three-dimensional datasets, finding a coherent and efficient interaction to perform a set of operations to better explore these data is a major challenge.

Solutions explored so far include the mouse, tangible user interfaces and the mid-air interaction [Cordeil *et al.*, 2017]. Using a 2D device, such as the mouse, for a 3D task greatly limits the interaction. Although tangible interfaces add physicality, a lack of feedback raises some issues in terms of the number of physical bricks. Finally, the mid-air interaction supports direct data manipulation, but may be limiting in terms of gesture detection and 3D perception (related to depth detection). Our goal to maintain the freedom of movement of mid-air interactions, the degrees of freedom of tangible interactions and the accuracy of the mouse to provide a flexible and precise solution for interaction with immersive visualizations.

A perspective to solve this challenge is to use our multi-DoF mouse (Roly-Poly Mouse) in this new interaction context. We will study the concept of on-body tangible interactions using the user as a physical support for interaction to give him the freedom of movement necessary to explore immersive visualization anchored in the real world. This solution will be based on the combination of 1) the use of a multi-DOF mouse-type wireless device, combining the precision inherent to a mouse, the flexibility of tangibles and the large control capabilities of multi-degrees of freedom mice and 2) the use of the body to guide the physical manipulations of the device by exploiting the proprioception of the user (the body's natural capacity to sense its own body parts and perceive their movements and localizations) while limiting muscle fatigue inherent to mid-air interactions.

6.3 Pervasive freeform interfaces

Our work on freeform interfaces started to address some of the challenges of these types of pervasive interfaces, in particular in terms of displaying text and laying out content. However, our initial work did not look into the cognitive aspects of looking for information on non-rectangular displays and only considered a limited set of simple shapes (such as circles and triangles). Hence our perspectives include studying these cognitive aspects and generalizing our approach to design content for freeform interfaces using computer graphics methods.

Cognitive aspects of non-rectangular interfaces

Although our previous work showed essential differences between rectangular and non-rectangular interfaces when reading text, these prior studies have not examined visual search strategies employed by users. Revealing these patterns is key in identifying how best to place and structure content [Buscher *et al.*, 2009] on non-rectangular displays. Designers could use such information to place relevant information at strategic locations for rapid access, key knowledge for identifying where to place menus or attention-grabbing banners on websites [Buscher *et al.*, 2009].

Our future work will fill this gap by investigating gaze patterns when visually searching information on non-rectangular interfaces. We will build on previous work and explore this question using the visual layout and display shapes investigated in our previous studies. Since gaze tracking has been widely used to understand how users search information on traditional screens, we will use it to log the gaze patterns and highlight particular strategies on non-rectangular interfaces.

Mapping content and input interaction on freeform interfaces

We will develop an interaction paradigm for non-rectangular user interfaces with a focus on two main objectives: exploring alternate mappings of content onto non-rectangular surfaces (1) and proposing novel input interaction techniques for non-rectangular displays (2).

Mapping content onto non-rectangular surfaces poses organisation, adaptation and deformation challenges. Organisation consists in geometrically arranging items on the shape to fit the available space (for instance, laying out icons). Adaptation consists in defining how the content changes when it is displaced (for instance, when text is scrolled). Deformation occurs when the display surface shape changes over time (for instance when an occluding physical object, such as a cup, is displaced on the surface). We will propose and study different mappings of content onto freeform displays to organize, adapt and deform content. The first step will be to geometrically define the coordinate system of the non-rectangular surface: for this, we will use computer graphics methods. We propose to rely on Computer Aided Design trimming technics in which free-form parametric shapes are defined over a support surface (in our case the full projection). Distortions will be controlled with local parametrizations and implicit shape representation may be considered for eventual in/out display shape tests. We will then first focus on simple content, such as text, menus, windows and/or GUI items among others. We will finally move to more complex visual content, such as time series or geospatial data visualisation.

We will explore two approaches to interact with such displays, either direct touch on the surface itself or indirect input on a distant touch surface (e.g. on a handheld touchscreen, a common approach to interact with public displays). For both input types, our objective is to provide pointing and content manipulation techniques (such as scrolling or 2D panning). For indirect input, we will also design and evaluate different mappings between the rectangular input coordinates (touchscreen) and the non-rectangular projection.

Bibliography

- [Ashbrook, 2010] Daniel L Ashbrook. *Enabling mobile microinteractions*. Georgia Institute of Technology, 2010.
- [Bach *et al.*, 2014] Benjamin Bach, Pierre Dragicevic, Daniel Archambault, Christophe Hurter, and Sheelagh Carpendale. *A Review of Temporal Data Visualizations Based on Space-Time Cube Operations*, pages 23–41. Eurographics Conference on Visualization (EuroVis 2014), 2014.
- [Bailly *et al.*, 2012] Gilles Bailly, Jörg Müller, Michael Rohs, Daniel Wigdor, and Sven Kratz. Shoesense: A new perspective on gestural interaction and wearable applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1239–1248, New York, NY, USA, 2012. ACM.
- [Balakrishnan and Hinckley, 2000] Ravin Balakrishnan and Ken Hinckley. Symmetric bi-manual interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, pages 33–40, New York, NY, USA, 2000. ACM.
- [Bardot *et al.*, 2016] Sandra Bardot, Marcos Serrano, and Christophe Jouffrais. From tactile to virtual: Using a smartwatch to improve spatial map exploration for visually impaired users. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '16, pages 100–111, New York, NY, USA, 2016. ACM.
- [Bardot *et al.*, 2017] Sandra Bardot, Marcos Serrano, Bernard Oriola, and Christophe Jouffrais. Identifying how visually impaired people explore raised-line diagrams to improve the design of touch interfaces. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 550–555, New York, NY, USA, 2017. ACM.
- [Beaudouin-Lafon, 2004] Michel Beaudouin-Lafon. Designing interaction, not interfaces. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 15–22, New York, NY, USA, 2004. ACM.
- [Bederson, 2004] Benjamin B. Bederson. Interfaces for staying in the flow. *Ubiquity*, 2004(September):1–1, September 2004.
- [Benko and Feiner, 2005] Hrvoje Benko and Steven Feiner. Multi-monitor mouse. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1208–1211, New York, NY, USA, 2005. ACM.
- [Bergé *et al.*, 2014] Louis-Pierre Bergé, Marcos Serrano, Gary Perelman, and Emmanuel Dubois. Exploring smartphone-based interaction with overview+detail interfaces on 3d public displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, MobileHCI '14, pages 125–134, New York, NY, USA, 2014. ACM.
- [Bimber and Raskar, 2005] Oliver Bimber and Ramesh Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA, 2005.

- [Boring *et al.*, 2009] Sebastian Boring, Marko Jurmu, and Andreas Butz. Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, OZCHI '09, pages 161–168, New York, NY, USA, 2009. ACM.
- [Boring *et al.*, 2010] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: Mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296, New York, NY, USA, 2010. ACM.
- [Boring, 2011] Sebastian Boring. Interacting in multi-display environments. *DC Ubicomp*, 05 2011.
- [Bortolaso *et al.*, 2013] Christophe Bortolaso, Matthew Oskamp, T.C. Nicholas Graham, and Doug Brown. Ormis: A tabletop interface for simulation-based training. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '13, pages 145–154, New York, NY, USA, 2013. ACM.
- [Bragdon *et al.*, 2011] Andrew Bragdon, Eugene Nelson, Yang Li, and Ken Hinckley. Experimental analysis of touch-screen gesture designs in mobile environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 403–412, New York, NY, USA, 2011. ACM.
- [Buscher *et al.*, 2009] Georg Buscher, Edward Cutrell, and Meredith Ringel Morris. What do you see when you're surfing?: Using eye tracking to predict salient regions of web pages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 21–30, New York, NY, USA, 2009. ACM.
- [Buxton, 2008] Bill Buxton. The long nose of innovation. *BusinessWeek*, January 2008.
- [Cauchard *et al.*, 2011] Jessica R. Cauchard, Markus Löchtfeld, Pourang Irani, Johannes Schoening, Antonio Krüger, Mike Fraser, and Sriram Subramanian. Visual separation in mobile multi-display environments. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 451–460, New York, NY, USA, 2011. ACM.
- [Cauchard *et al.*, 2012] Jessica R. Cauchard, Mike Fraser, Teng Han, and Sriram Subramanian. Steerable projection: Exploring alignment in interactive mobile displays. *Personal Ubiquitous Comput.*, 16(1):27–37, January 2012.
- [Chandler *et al.*, 2015] T. Chandler, M. Cordeil, T. Czauderna, T. Dwyer, J. Glowacki, C. Goncu, M. Klapperstueck, K. Klein, K. Marriott, F. Schreiber, and E. Wilson. Immersive analytics. In *2015 Big Data Visual Analytics (BDVA)*, pages 1–8, Sept 2015.
- [Chokshi *et al.*, 2014] Apoorve Chokshi, Teddy Seyed, Francisco Marinho Rodrigues, and Frank Maurer. eplan multi-surface: A multi-surface environment for emergency response planning exercises. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ITS '14, pages 219–228, New York, NY, USA, 2014. ACM.
- [Chourabi *et al.*, 2012] Hamed Chourabi, Taewoo Nam, Shawn Walker, J. Ramon Gil-Garcia, Sehl Mellouli, Karine Nahon, Theresa Pardo, and Hans Scholl. Understanding smart cities: An integrative framework. *45th Hawaii International Conference on System Sciences*, pages 2289–2297, 01 2012.
- [Cockburn *et al.*, 2009] Andy Cockburn, Amy Karlson, and Benjamin B. Bederson. A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv.*, 41(1):2:1–2:31, January 2009.

- [Cordeil *et al.*, 2017] Maxime Cordeil, Benjamin Bach, Yongchao Li, Elliott Wilson, and Tim Dwyer. *A Design Space for Spatio-Data Coordination: Tangible Interaction Devices for Immersive Information Visualisation*, pages 46–50. IEEE Computer Society Press, 9 2017.
- [Cotting and Gross, 2006] Daniel Cotting and Markus Gross. Interactive environment-aware display bubbles. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, UIST '06, pages 245–254, New York, NY, USA, 2006. ACM.
- [Davis *et al.*, 2007] Richard C. Davis, T. Scott Saponas, Michael Shilman, and James A. Landay. Sketchwizard: Wizard of oz prototyping of pen-based user interfaces. In *Proceedings of UIST'07*, pages 119–128. ACM Press, 2007.
- [Dix *et al.*, 2003] Alan Dix, Janet E. Finlay, Gregory D. Abowd, and Russell Beale. *Human-Computer Interaction (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2003.
- [Dubois *et al.*, 2018] Emmanuel Dubois, Marcos Serrano, and Mathieu Raynal. Rolling-menu: Rapid command selection in toolbars using roll gestures with a multi-dof mouse. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 367:1–367:12, New York, NY, USA, 2018. ACM.
- [Ducasse *et al.*, 2016] Julie Ducasse, Marc J-M Macé, Marcos Serrano, and Christophe Jouffrais. Tangible reels: Construction and exploration of tangible maps by visually impaired users. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2186–2197, New York, NY, USA, 2016. ACM.
- [Elmqvist and Irani, 2013] Niklas Elmqvist and Pourang Irani. Ubiquitous analytics: Interacting with big data anywhere, anytime. *Computer*, 46(4):86–89, April 2013.
- [Elmqvist *et al.*, 2008] Niklas Elmqvist, Nathalie Henry, Yann Riche, and Jean-Daniel Fekete. Melange: Space folding for multi-focus interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1333–1342, New York, NY, USA, 2008. ACM.
- [Galitz, 2007] Wilbert O. Galitz. *The Essential Guide to User Interface Design: An Introduction to GUI Design Principles and Techniques*. John Wiley & Sons, Inc., New York, NY, USA, 2007.
- [Geyer *et al.*, 2012] Florian Geyer, Anita Höchtl, and Harald Reiterer. Harnessing the benefits of bimanual and multi-finger input for supporting grouping tasks on interactive tabletops. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, NordiCHI '12, pages 496–499, New York, NY, USA, 2012. ACM.
- [Gotsch *et al.*, 2016] Daniel Gotsch, Xujing Zhang, Jesse Burstyn, and Roel Vertegaal. Holoflex: A flexible holographic smartphone with bend input. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, pages 3675–3678, New York, NY, USA, 2016. ACM.
- [Granell and Leiva, 2017] Emilio Granell and Luis A. Leiva. btap: Back-of-device tap input with built-in sensors. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, pages 52:1–52:6, New York, NY, USA, 2017. ACM.
- [Hachet *et al.*, 2009] Martin Hachet, Fabrice Decle, Sebastian Knödel, and Pascal Guitton. Navidget for 3d interaction: Camera positioning and further uses. *Int. J. Hum.-Comput. Stud.*, 67(3):225–236, March 2009.
- [Han Kim *et al.*, 2010] K Han Kim, Matthew Reed, and Bernard J Martin. A model of head movement contribution for gaze transitions. *Ergonomics*, 53:447–57, 04 2010.

- [Harrison *et al.*, 2010] Chris Harrison, Desney Tan, and Dan Morris. Skinput: Appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 453–462, New York, NY, USA, 2010. ACM.
- [Hilliges *et al.*, 2012] Otmar Hilliges, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson. Holodesk: Direct 3d interactions with a situated see-through display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2421–2430, New York, NY, USA, 2012. ACM.
- [Hinckley, 2003] Ken Hinckley. The human-computer interaction handbook. In Julie A. Jacko and Andrew Sears, editors, *The Human-computer Interaction Handbook*, chapter Input Technologies and Techniques, pages 151–168. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 2003.
- [Houben and Marquardt, 2015] Steven Houben and Nicolai Marquardt. Watchconnect: A toolkit for prototyping smartwatch-centric cross-device applications. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1247–1256, New York, NY, USA, 2015. ACM.
- [Ishii and Ullmer, 1997] Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, pages 234–241, New York, NY, USA, 1997. ACM.
- [Jankowski *et al.*, 2010] Jacek Jankowski, Krystian Samp, Izabela Irzynska, Marek Jozwicz, and Stefan Decker. Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1321–1330, New York, NY, USA, 2010. ACM.
- [Jansen *et al.*, 2013] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. Evaluating the efficiency of physical visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2593–2602, New York, NY, USA, 2013. ACM.
- [Jansen *et al.*, 2015] Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. Opportunities and challenges for data physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3227–3236, New York, NY, USA, 2015. ACM.
- [Javed *et al.*, 2012] Waqas Javed, Sohaib Ghani, and Niklas Elmqvist. Polyzoom: Multiscale and multifocus exploration in 2d visual spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 287–296, New York, NY, USA, 2012. ACM.
- [Jones *et al.*, 2012] Brett Jones, Rajinder Sodhi, David Forsyth, Brian Bailey, and Giuliano Maciocci. Around device interaction for multiscale navigation. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 83–92, New York, NY, USA, 2012. ACM.
- [Kaizu *et al.*, 2010] Kazunari Kaizu, Samik Ghosh, Yukiko Matsuoka, Hisao Moriya, Yuki Shimizu-Yoshida, and Hiroaki Kitano. A comprehensive molecular interaction map of the budding yeast cell cycle. *Molecular systems biology*, 6:415, 09 2010.
- [Kikuchi *et al.*, 2017] Takashi Kikuchi, Yuta Sugiura, Katsutoshi Masai, Maki Sugimoto, and Bruce H. Thomas. Eartouch: Turning the ear into an input surface. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, pages 27:1–27:6, New York, NY, USA, 2017. ACM.

- [Kimura *et al.*, 2011] Hidei Kimura, Akira Asano, Issei Fujishiro, Ayaka Nakatani, and Hayato Watanabe. True 3d display. In *ACM SIGGRAPH 2011 Emerging Technologies*, SIGGRAPH '11, pages 20:1–20:1, New York, NY, USA, 2011. ACM.
- [Kin *et al.*, 2011] Kenrick Kin, Björn Hartmann, and Maneesh Agrawala. Two-handed marking menus for multitouch devices. *ACM Trans. Comput.-Hum. Interact.*, 18(3):16:1–16:23, August 2011.
- [Kohn, 1999] Kurt W. Kohn. Molecular interaction map of the mammalian cell cycle control and dna repair systems. *Molecular biology of the cell*, 10 8:2703–34, 1999.
- [Kratz *et al.*, 2012] Sven Kratz, Michael Rohs, Dennis Guse, Jörg Müller, Gilles Bailly, and Michael Nischt. PalmSpace: Continuous around-device gestures vs. multitouch for 3d rotation tasks on mobile devices. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, pages 181–188, New York, NY, USA, 2012. ACM.
- [Kurtenbach and Buxton, 1991] Gordon Kurtenbach and William Buxton. Issues in combining marking and direct manipulation techniques. In *Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology*, UIST '91, pages 137–144, New York, NY, USA, 1991. ACM.
- [Lavie and Tractinsky, 2004] Talia Lavie and Noam Tractinsky. Assessing dimensions of perceived visual aesthetics of web sites. *Int. J. Hum.-Comput. Stud.*, 60(3):269–298, March 2004.
- [Lee *et al.*, 2017] Juyoung Lee, Hui-Shyong Yeo, Murtaza Dhuliawala, Jedidiah Akano, Junichi Shimizu, Thad Starner, Aaron Quigley, Woontack Woo, and Kai Kunze. Itchy nose: Discreet gesture interaction using eeg sensors in smart eyewear. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, ISWC '17, pages 94–97, New York, NY, USA, 2017. ACM.
- [Lucero and Serrano, 2017] Andres Lucero and Marcos Serrano. Towards proxemic mobile collocated interactions. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 9(4):10, 2017.
- [Lucero *et al.*, 2016] Andrés Lucero, Aaron Quigley, Jun Rekimoto, Anne Roudaut, Martin Porcheron, and Marcos Serrano. Interaction techniques for mobile collocation. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, MobileHCI '16, pages 1117–1120, New York, NY, USA, 2016. ACM.
- [Malacria *et al.*, 2010] Sylvain Malacria, Eric Lecolinet, and Yves Guiard. Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces: The cyclostar approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2615–2624, New York, NY, USA, 2010. ACM.
- [McGookin *et al.*, 2010] David McGookin, Euan Robertson, and Stephen Brewster. Clutching at straws: Using tangible interaction to provide non-visual access to graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1715–1724, New York, NY, USA, 2010. ACM.
- [Moscovich and Hughes, 2008] Tomer Moscovich and John F. Hughes. Indirect mappings of multi-touch input using one and two hands. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1275–1284, New York, NY, USA, 2008. ACM.

- [Nancel *et al.*, 2011] Mathieu Nancel, Julie Wagner, Emmanuel Pietriga, Olivier Chapuis, and Wendy Mackay. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 177–186, New York, NY, USA, 2011. ACM.
- [Nancel *et al.*, 2013] Mathieu Nancel, Olivier Chapuis, Emmanuel Pietriga, Xing-Dong Yang, Pourang P. Irani, and Michel Beaudouin-Lafon. High-precision pointing on large wall displays using small handheld devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 831–840, New York, NY, USA, 2013. ACM.
- [Nii *et al.*, 2012] Hideaki Nii, Kening Zhu, Hiromi Yoshikawa, Nyan Lin Htat, Roland Aigner, and Ryohei Nakatsu. Fuwa-vision: An auto-stereoscopic floating-image display. In *SIGGRAPH Asia 2012 Emerging Technologies*, SA '12, pages 13:1–13:4, New York, NY, USA, 2012. ACM.
- [Oakley *et al.*, 2015] Ian Oakley, DoYoung Lee, MD. Rasel Islam, and Augusto Esteves. Beats: Tapping gestures for smart watches. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 1237–1246, New York, NY, USA, 2015. ACM.
- [Perelman *et al.*, 2015] Gary Perelman, Marcos Serrano, Mathieu Raynal, Celia Picard, Mustapha Derras, and Emmanuel Dubois. The roly-poly mouse: Designing a rolling input device unifying 2d and 3d interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 327–336, New York, NY, USA, 2015. ACM.
- [Perelman *et al.*, 2016] Gary Perelman, Marcos Serrano, Mathieu Raynal, Celia Picard, Mustapha Derras, and Emmanuel Dubois. Deco: A design space for device composition. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, DIS '16, pages 435–446, New York, NY, USA, 2016. ACM.
- [Pinhanez, 2001] Claudio S. Pinhanez. The everywhere displays projector: A device to create ubiquitous graphical interfaces. In *Proceedings of the 3rd International Conference on Ubiquitous Computing*, UbiComp '01, pages 315–331, London, UK, UK, 2001. Springer-Verlag.
- [Piper *et al.*, 2002] Ben Piper, Carlo Ratti, and Hiroshi Ishii. Illuminating clay: A 3-d tangible interface for landscape analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 355–362, New York, NY, USA, 2002. ACM.
- [Poupyrev *et al.*, 2006] Ivan Poupyrev, Henry Newton-Dunn, and Olivier Bau. D20: Interaction with multifaceted display devices. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, pages 1241–1246, New York, NY, USA, 2006. ACM.
- [Preece *et al.*, 2001] Jenny Preece, Yvonne Rogers, and Helen Sharp. *Beyond Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, Inc., New York, NY, USA, 2001.
- [Raskar *et al.*, 1998] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '98, pages 179–188, New York, NY, USA, 1998. ACM.
- [Raskar *et al.*, 2004] Ramesh Raskar, Paul Beardsley, Jeroen van Baar, Yao Wang, Paul Dietz, Johnny Lee, Darren Leigh, and Thomas Willwacher. Rfig lamps: Interacting with a self-describing world via photosensing wireless tags and projectors. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH '04, pages 406–415, New York, NY, USA, 2004. ACM.

- [Rico and Brewster, 2010] Julie Rico and Stephen Brewster. Usable gestures for mobile interfaces: Evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 887–896, New York, NY, USA, 2010. ACM.
- [Roudaut *et al.*, 2014] Anne Roudaut, Rebecca Reed, Tianbo Hao, and Sriram Subramanian. Changibles: Analyzing and designing shape changing constructive assembly. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2593–2596, New York, NY, USA, 2014. ACM.
- [Ruiz *et al.*, 2011] Jaime Ruiz, Yang Li, and Edward Lank. User-defined motion gestures for mobile interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 197–206, New York, NY, USA, 2011. ACM.
- [Saidi *et al.*, 2016] Houssem Saidi, Marcos Serrano, and Emmanuel Dubois. Investigating the effects of splitting detailed views in overview+detail interfaces. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '16, pages 180–184, New York, NY, USA, 2016. ACM.
- [Saidi *et al.*, 2017] Houssem Saidi, Marcos Serrano, Pourang Irani, and Emmanuel Dubois. Tdome: A touch-enabled 6dof interactive device for multi-display environments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5892–5904, New York, NY, USA, 2017. ACM.
- [Serrano and Nigay, 2010] Marcos Serrano and Laurence Nigay. A wizard of oz component-based approach for rapidly prototyping and testing input multimodal interfaces. *Journal on Multimodal User Interfaces*, 3(3):215–225, Apr 2010.
- [Serrano *et al.*, 2013] Marcos Serrano, Eric Lecolinet, and Yves Guiard. Bezel-tap gestures: Quick activation of commands from sleep mode on tablets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3027–3036, New York, NY, USA, 2013. ACM.
- [Serrano *et al.*, 2014a] Marcos Serrano, Barrett M. Ens, and Pourang P. Irani. Exploring the use of hand-to-face input for interacting with head-worn displays. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 3181–3190, New York, NY, USA, 2014. ACM.
- [Serrano *et al.*, 2014b] Marcos Serrano, Dale Hildebrandt, Sriram Subramanian, and Pourang Irani. Identifying suitable projection parameters and display configurations for mobile true-3d displays. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, MobileHCI '14, pages 135–143, New York, NY, USA, 2014. ACM.
- [Serrano *et al.*, 2015] Marcos Serrano, Barrett Ens, Xing-Dong Yang, and Pourang Irani. Gluey: Developing a head-worn display interface to unify the interaction experience in distributed display environments. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, pages 161–171, New York, NY, USA, 2015. ACM.
- [Serrano *et al.*, 2016] Marcos Serrano, Anne Roudaut, and Pourang Irani. Investigating text legibility on non-rectangular displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 498–508, New York, NY, USA, 2016. ACM.
- [Serrano *et al.*, 2017] Marcos Serrano, Anne Roudaut, and Pourang Irani. Visual composition of graphical elements on non-rectangular displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 4405–4416, New York, NY, USA, 2017. ACM.

- [Serrano, 2010] Marcos Serrano. *Interaction multimodale en entrée : Conception et Prototypage*. Thèses et habilitations, University of Grenoble, 2010. Thèse de doctorat Informatique préparée au Laboratoire d’Informatique de Grenoble (LIG), Université Joseph Fourier. 283 pages.
- [Shin *et al.*, 2012] Choonsung Shin, Jin-Hyuk Hong, and Anind K. Dey. Understanding and prediction of mobile application usage for smart phones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp ’12*, pages 173–182, New York, NY, USA, 2012. ACM.
- [Song *et al.*, 2011] Peng Song, Wooi Boon Goh, Chi-Wing Fu, Qiang Meng, and Pheng-Ann Heng. Wysiwyf: Exploring and annotating volume data with a tangible handheld device. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’11*, pages 1333–1342, New York, NY, USA, 2011. ACM.
- [Su *et al.*, 2010] Jing Su, Alyssa Rosenzweig, Ashvin Goel, Eyal de Lara, and Khai N. Truong. Timbremap: Enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI ’10*, pages 17–26, New York, NY, USA, 2010. ACM.
- [Ten Koppel *et al.*, 2012] Maurice Ten Koppel, Gilles Bailly, Jörg Müller, and Robert Walter. Chained displays: Configurations of public displays can be used to influence actor-, audience-, and passer-by behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’12*, pages 317–326, New York, NY, USA, 2012. ACM.
- [Tsandilas, 2018] Theophanis Tsandilas. Fallacies of agreement: A critical review of consensus assessment methods for gesture elicitation. *ACM Trans. Comput.-Hum. Interact.*, 25(3):18:1–18:49, June 2018.
- [Walter *et al.*, 2013] Robert Walter, Gilles Bailly, and Jörg Müller. Strikeapose: Revealing mid-air gestures on public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’13*, pages 841–850, New York, NY, USA, 2013. ACM.
- [Wang Baldonado *et al.*, 2000] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces, AVI ’00*, pages 110–119, New York, NY, USA, 2000. ACM.
- [Weiser, 1991] Mark Weiser. The computer for the 21st century. *Scientific American*, 265(3):94–104, 1991.
- [Wobbrock *et al.*, 2009] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’09*, pages 1083–1092, New York, NY, USA, 2009. ACM.
- [Yamashita *et al.*, 2017] Koki Yamashita, Takashi Kikuchi, Katsutoshi Masai, Maki Sugimoto, Bruce H. Thomas, and Yuta Sugiura. Cheekinput: Turning your cheek into an input surface by embedded optical sensors on a head-mounted display. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, VRST ’17*, pages 19:1–19:8, New York, NY, USA, 2017. ACM.
- [Ye *et al.*, 2014] Hanlu Ye, Meethu Malu, Uran Oh, and Leah Findlater. Current and future mobile and wearable device use by people with visual impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’14*, pages 3123–3132, New York, NY, USA, 2014. ACM.

- [Zhang *et al.*, 2017] Yang Zhang, Gierad Laput, and Chris Harrison. Electrick: Low-cost touch sensing using electric field tomography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 1–14, New York, NY, USA, 2017. ACM.
- [Zhao *et al.*, 2008] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Jonathan Lazar. Data sonification for users with visual impairment: A case study with georeferenced data. *ACM Trans. Comput.-Hum. Interact.*, 15(1):4:1–4:28, May 2008.