Ordre: 42373

Université Lille 1 / INRIA Lille Nord Europe
Centre de Recherche en Informatique, Signal et Automatique de Lille
EDSPI Université Lille Nord-de-France

Hanaë Rateau

# Exploring interactive sub-spaces for gestural midair interaction.

## Exploration de l'usage de sous-espaces pour l'interaction à geste dans l'air.

17 Mai 2017
Thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Science.

**Committee:**

| | |
|---|---|
| Kris Luyten, University of Hasselt | Reviewer |
| Philippe Palanque, Université Toulouse 3 | Reviewer |
| Caroline Appert, Université Paris Sud | Examiner |
| Fanny Chevalier, Inria Lille - Nord Europe | Examiner |
| Laurence Duchien, Université Lille 1 | Examiner |
| Laurent Grisoni, Université Lille 1 | Supervisor |

© 2017 Tous droits réservés. lilliad.univ-lille.fr

© 2017 Tous droits réservés. lilliad.univ-lille.fr

# Abstract

The multiplication of computerized devices in our everyday life is not to prove. This new computerized environment is remarkably diverse in terms of size, interaction modalities, available actions, types of feedbacks…The expanding heterogeneity of our daily life environments requires fluid interaction in order to make easier and more fluid the transition between the different devices and interaction contexts.

This dissertation focuses on how to exploit gestural midair interaction to extend the possibilities of existing devices by using interactive spaces. The starting point is in the nonverbal communication theory of proxemics introduced by Eward T. Hall in [48] who stated that our perception of space is dynamic. From this, I argue that we could apply this dynamic understanding of space to interactive spaces. I propose a novel concept of interaction and an associated design framework for interactive spaces : Mimetic Interaction Space (MIS). To show the prospects MIS gives for midair interaction, I propose three instantiations of the concept that uses it in different ways. The *first one* is the use of MISs as a standalone interface the control of a remote display where I present an eliciting study for creation and deletion gestures of MIS in this context, and from the findings propose a prototype for indirect pointing. The *second instantiation* is the use of one or several MIS tied up to the tablet in two ways. First by cutting out the MIS in multiple ones. I present two applications, one for MRI visualization and one for tools selection in a drawing application. The second way of using a MIS linked to the tablet is by considering it as a continuation of the tablet screen around it. From this I present a new interaction technique, a study and a prototype. The *third instantiation* is in the context of interaction on wall displays where a MIS is placed right in front of the screen and has the role of a transition space from touch to midair interaction. This MIS allows for a continuous transition between the physical and direct nature of touch interaction, and the more abstract nature of midair interaction. I finally conclude by discussing the future of interfaces regarding midair gestures. I also discuss a facet of MIS that opens a novel way to think about MIS interaction.

© 2017 Tous droits réservés. 

Thèse de Hanaë Rateau, Lille 1, 2017



# Résumé

La multiplication des appareils numériques dans notre vie quotidienne n'est plus à prouver. L'ensemble de ces nouveaux appareils forment un nouvel environnement hétéroclite ou les interfaces diffèrent en taille, en modalité d'interaction, en actions possibles, en type de retour,...Cette expansion requiert alors une interaction fluide afin de pouvoir passer d'un appareil à l'autre facilement et rapidement ou d'un contexte à l'autre.

Cette dissertation s'intéresse à comment utiliser les gestes dans l'air pour enrichir l'interaction Homme-Machine en utilisant des espaces interactifs. Cette thèse s'inspire d'un concept de la communication non verbale : la proxémie. Cette théorie, introduite par Edward T. Hall dans [48], affirme entre autres que notre perception de l'espace est dynamique. Et s'accorde à l'environnement que nous percevons. En m'inspirant de cette théorie, je présente ici un nouveau concept d'interaction accompagné de son framework de design : Mimetic Interaction Space (MIS). Afin de montrer ce que le concept peut apporter à l'interaction, en plus d'une relecture de la littérature sur l'interaction dans les airs, je propose trois instanciations de ce concept autour des trois types d'utilisation du concept. La *première instanciation* est pour le contrôle indirect sur un écran distant en utilisant un MIS comme une interface à part entière. Je présente une étude d'élicitation de geste pour la création et la suppression de MIS ainsi qu'un prototype appliquant les résultats de l'étude. La *seconde instanciation* d'enrichir l'interaction sur tablette en utilisant un ou des MISs l'entourant. Deux propositions d'utilisation sont faites. Une première en subdivisant le MIS en plusieurs autour de la tablette pour laquelle je propose deux applications. Puis une seconde utilisation du MIS comme étant la continuité de l'écran de la tablette pour laquelle je présente une nouvelle technique d'interaction ainsi qu'une étude et un prototype pour la navigation. La *troisième instanciation* se fait dans le contexte de l'interaction sur très grands écrans tactiles. Ici, un MIS a pour rôle de faire la transition continue entre l'interaction tactile et l'interaction dans les airs. Pour finir, j'introduis quelques pistes de développement pour l'avenir des MIS et je propose une réflexion sur une facette du concept des MIS qui ouvre d'importantes questions sur l'interaction basée MIS.

© 2017 Tous droits réservés. lilliad.univ-lille.fr

# Contents





© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017





© 2017 Tous droits réservés. lilliad.univ-lille.fr

# List of Figures





© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017





© 2017 Tous droits réservés. lilliad.univ-lille.fr

# One

## Introduction

*“We must begin seeing man as an interlocutor with his environment."*

— Edward T. Hall, *The Hidden Dimension.*

Looking at the definition of the word "interaction", we find in the Oxford Dictionary, *"Reciprocal action or influence."* Then, when thinking about human-computer interaction, we should take into account the environment in which the user might perform by observing how the conditions of interactions could influence the user’s actions. So that the actions of the designed interfaces fit the constraint of both user and environment.

The multiplication of computerized devices in our everyday life is not to prove. This multiplication transforms our environments. In Mark Wieser’s vision [114] of the 21st century, he describes an environment where several computers would be interconnected and seamlessly integrated to our world. Though they would be invisible to users, the services brought by these computers would clearly be observable. This transformation is already underway. This is what we can call now, smart cities, smart homes…Those smart environments propose among others things to control easily, quickly and with the desired degree of precision the devices composing our environments. Due to their inherent heterogeneity, those environments are contexts in which fluid interaction has a key role to play. We need always-available, (ideally) low-instrumented, interaction techniques, that would permit users interacting with several devices; we also need interaction techniques that allow collaboration in the same room for a given task as well as private tasks. Then, we need to think about how human-



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



computer interaction can contribute to the evolution of such environments. Probably, our most used tools to act on our environment are our hands and arms. Sometimes, with the help of external tools, and sometimes we act with our bare hands. Being so since we are born, it gives us an extraordinary skill set to interact with the physic world. With this learning, comes spatial awareness of our own body and the surrounding world. Tapping into this almost natural skills can benefit to human-computer interaction [87]. It is also what Can Liu argues in her thesis [74]. I investigate in this document the use of midair gestures for Human-Computer Interaction (HCI).

Additionally, communication is the heart of interaction. The speech is the first thing we think of when speaking about communication. But speech is not the only way. Gestures are a major communication channel in human-human interaction. Communication and functional gestures have been learned since the early days of a life and before speech. In this dissertation, I then explore the possibilities of using midair gestures for HCI by taking inspiration from what we know about us. In particular, I am focusing on midair interaction using delimited interactive spaces. They are sub-spaces defined either by the user, or the interaction designer for a particular use and context of use. Previous work has proposed to investigate the use of sub-spaces for mid-air interaction to address the continuous gesture problem of midair gestures. Wigdor and Wixon refer to this as the "live mic" problem in [116] due to the always-on nature of gestures.

## 1.1 Thesis statement

In this document, I explore the potential new approaches gestural midair interaction brings to HCI as a complementary modality of interaction and on its own. My dissertation focuses on how to exploit gestural midair interaction to extend the interaction possibilities of existing interactive devices by using interactive spaces. My starting point is in the nonverbal communication theory of proxemics introduced by Eward T. Hall in [48] who stated that our perception of space is dynamic, defined by what can be done in this space. From this dynamic understanding of space to interactive spaces, I propose a novel concept of interaction and an associated design framework for interactive spaces : Mimetic Interaction Space (MIS). To show the prospects MIS gives for midair interaction, I propose three instantiations of it in different contexts of use.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



## 1.2 Contributions

The major contribution of this thesis is the introduction of the MIS concept and the design framework. The remaining of the dissertation gives several arguments to show that this concept is interesting for HCI.

After the definition of the concept, I give some leads for the design space exploration supported by the design framework as well as examples of interaction scenarios that use MIS-based interaction.

The MIS concept also gives the possibility of a re-reading of previous work on midair interaction using interactive spaces. This re-reading shows how MIS can unify and include this past work.

Finally, I propose *three* instantiations of the concept to show the prospect the MIS concept brings. The *first one* is the use of an independent MIS for remote display interaction. I also present an eliciting study for creation and deletion gestures of MIS in this context, and from the findings propose a prototype allowing distant mouse control. The *second instantiation* is the use of MIS around the tablet to extend it and enrich the direct interaction inherent to such touch devices. The MIS device is here part of the MIS and vice-versa. I present two ways of using the surrounding space. Either by dividing it in several MIS or by considering the MIS as a continuous extension of the tablet . In each case I present some applications, first an application for MRI visualization and one for tools selection in a drawing application, then an application for large content navigation such as maps. The *third instantiation* is in the context of interaction on wall displays where a MIS is placed right in front of the screen. It has the role of a transition space from touch to midair interaction allowing seamless transition between those two modalities. In this last contribution, the MIS is fused to the interactive system.

## 1.3 Structure of the dissertation

The chapter 2 present previous work related to midair interaction using interactive spaces. The three remaining chapters present instantiations of the MIS concept in different contexts. Those instantiations are made according to the presented MIS framework by varying one or multiple components. The chapter 3 introduces the major contribution of this dissertation, which is Mimetic Inter-



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



action Spaces (MIS). They are basically dynamic delimited spaces for interaction. The concept comes with a framework that formally defines a MIS with four components. Then, chapter 4 present a study to understand how users might interact with a planar MIS in a remote control mouse cursor context. The study allowed us to develop a proof of concept applying the different results found. Then, in chapter 5, we explore the possibilities of gestural interaction in space in different interactive contexts with a tablet. Here, the MIS are attached to a tablet in order to improve or enrich interaction with it. Finally, in the last chapter 6, we investigate how midair interaction allows to interact with large tactile displays. I will conclude this dissertation talking about global future work with MIS concept. One part being further study on how we might use MIS in different contexts of use and with what kind of data and interfaces. Finally, I will talk in more details about what I call the reflexive nature of MISs in section 3.5 and try to start a reflexion on the subject.

In the end of the dissertation, I not only would like the reader to understand what MIS is but also to see the prospects of this novel concept thanks to the re-reading of the literature and the three different instantiations of the concept applied to varied contexts.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# Two

## Interactive Spaces

*"The past, like the future, is indefinite and exists only as a spectrum of possibilities."*

— Stephen Hawking

### Contents





© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



As said in the introduction, I propose in this dissertation a novel concept for midair interaction that relies on the use of interactive spaces. Previous work has explored the use of spaces to extend or improve interaction. In this chapter, I present this work. In some cases, the use of interactive spaces is not mention explicitly but is in fact relying on it. The large amount of work proposing interaction in specific spaces makes the classification of such work challenging. I chose to refer to the nature of the space and its context of use as discriminants.

Each section present previous work using interactive spaces of similar nature or purpose. First, I present papers that use invisible interactive spaces that are interfaces. Then, work that enables device surroundings interaction will be detailed, then visible interactive spaces, and finally the particular case of tabletop and display extensions will be discussed. In each of these sections, the work is regrouped so that it shares the same context of use.

## 2.1 Standalone Invisible Spaces

I call standalone invisible space, interactive spaces that can be seen as independent interfaces that do not rely on other input to work. Of course, they have to be linked to a computer to indeed control it.

In [46], Gustafson et al. propose a system with no visual feedback. Screen is replaced by short term memory. The user defines dynamically the space in which he wants to interact with a non-dominant hand posture as a reference point. Interactions start with a posture and stop when the user releases the pose. Three studies show that the more time spent, the more degraded memory. But using the non dominant hand as a reference point improves performance. In the different example of use, the imaginary space is linked to the non-dominant hand and its lifetime is define by whether or not a specific hand posture is performed. Its position and orientation are also determined by the hand respective properties. In the paper, there is no mention of the space dimension. The hand only defines a plane. The gestures used in the space as we could use a pen (*i.e.*, draw, write, annotate,...). Following this work, in the paper presenting the imaginary phone [47], an implementation of imaginary interfaces concept, the authors show that knowledge from a physical to an imaginary interface can be transfered and that palm interaction is precise enough for standard mobile phone interaction. Building on the previous work of Gustafson et al. [46,47], the proposed concept in [25] wants to leverage the palm as an interactive surface for TV control. In



© 2017 Tous droits réservés. lilliad.univ-lille.fr

these works, the invisible space is linked to the palm of the non-dominant hand and the interaction is restricted to pointing gestures to select items or push buttons.

The work in [72] introduces a virtual touch panel named AirTouch Panel. The user has to form an L-shape with his left hand to define a virtual panel and then can interact with an AirTouch panel-compatible smart-devices. In this work, the panel has a pre-defined size the user cannot control but he can define the position and the orientation of the panel at the creation with the L-shape pose. Once created, the panel is anchored to its original user-defined position and the user can re-anchored it by performing the corresponding L-shape.

In [111], the authors present an interaction system for "creative expression of 3D shapes". Along with a framework for intelligent shape generation, the authors introduce a virtual slab parallel to the designer's (x-y) plane in which the designer create the shapes. The slab moves with the designer and, in the prototype, is defined at a distance of 75% of the total arm-length of the designer. This system try to integrate the designer in the 3D modeling process. Through the gestures, the designer keeps expressiveness of sketch while beginning the 3D modeling. Doing so, this reduces the gap between the early sketches and the final CAD-based design.

With Shoesense [4], Bailly et al. present a wearable system where the space above one shoe is interactive. The user can then invoke commands by gesturing. The authors propose a set of three example gestures that can be performed and exploited to invoke command. The triangle gestures are a set of gestures the users make with their two arms and torso. Sliding one hand along the other arms vary the area of the triangle or the angles. Thus, a mapping between poses and commands can be made as well as continuous commands such as volume tuning. The 3D radial menus are an extension of 2D radial menus or marking menus more generally. When users trigger a pinch pose, 3D radial menu is activated and the users may invoke a specific command. Finally, the finger count gestures are static gestures inspired by [3] where the number of fingers is mapped to a specific command in a particular context. Here, this is the space above the shoe that is a define as an interactive space and this feature presents several advantages: performing frequent gestures without reaching the phone (for mobile interaction), discreet actions when in a social context, alleviate some accessibility problems (*i.e.*, fat finger problem). In this work, the authors also propose to use the gestures in a discrete manner along with their continuous nature (*i.e.*, pose, number of fingers,...)



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



The authors in [65] present a wrist-worn sensor to track and recognize hand posture along with a real-time tracking pipeline, new kinematic models, use case scenarios and an evaluation of the system. This system leverages on our full dexterity to enrich and expand interaction in mobile use context with discrete as well as continuous hand gestures. In the different proposed scenarios, the fingers gestures in the palm space can be either performed when interacting with an on-screen application or combined with touch or with no visual feedback.

PUB [73] (Point Upon Body) uses the forearm as an input interface. Two studies were ran to determined (1) how many points could be discriminate by users, (2) how users would tap and (3) how proprioception and haptic feedback of the skin affect the accuracy of the tap. Results show that with 7 points aligned along the forearm users managed to correctly tap on it without making too many errors. Plus, the touch feeling when touching the skin improves the accuracy. The prototype was used to control a music player on a smartphone allowing eye-free interaction. Users associate the functions to one of the 6 area of the forearm and could switch to the previous/next song by sliding. Plus, the authors implemented an application to interact with a hierarchical GUI on a remote display. Here the interaction space is directly attached and in contact with a body part to leverage on haptic and tactile feedback.

Virtual Shelves [71] is an interaction technique allowing mobile phone users to trigger shortcuts depending on the direction pointed with the phone within a hemisphere centered on users bodies and in front of them. The hemisphere is divided in equal angular width and height shelves where one item is stored. The content of shelves is context dependent. It is clear here that the interaction volume is a hemisphere that accompanies the users all along during the interaction with their mobile phones. The gesture used is a simple pointing gesture, as clicking is triggered by key press and release. Similarly, in [18], Cauchard et al. propose three phone-based prototype for quick workspace switching depending on the mobile phone orientation and position around the user. One of them, mSpace, simply assigns one area to one application. In order to switch to another workspace(*i.e.*, application), users have to move the phone into a different area. Here, the space around the user is only use to switch between applications. The interaction is still done on the phone. The space allwos only to display different information. Chen et al. [19] use on and around the body area as anchors to data or actions. Bringing the mobile phone in one of this area triggers the associated action (*e.g.*, call a specific contact) or displays the stored data (*e.g.*, bookmark). From the work of Hall [48] and neuropsychologists [58] that we have multiple spaces centered on us, the authors focus on the exploitation of the pericutaneous space (*i.e.*, on-body space) and the peripersonal space (*i.e.*, space within arms



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



reach). The use of these space is made by measuring position and orientation of the device in these spaces discretely or continuously. For example, on-body discrete position (*e.g.*, left wrist) can be associated to a certain shortcut (*e.g.*, give the next event from the calendar). In this work, though the space on and around the body can only be used as a storage, it can also be an interactive space in which the interaction is performed.

In [81], the authors propose to project information on real world surfaces while interacting in the viewing frustum volume of a head-worn camera. From flight ticket, to a wall, or even the wrist, the embedded pico-projector allows to display information as well as standard GUI. To interact with these content, conventional fingers(thumbs and index) gestures and postures are tracked thanks to color markers as well as specific postures. It is important to note that though there is projected content, interaction space is separated from visual feedback space. One planar space is dedicated to display information (*i.e.*, displaying space) and the frustum volume is for gestures. In the prototype, displaying spaces are not always there. For example, a frame gesture in front of the camera triggers a photo. So here, the interaction space is invisible though visual feedback of the interfaces is possible. Similarly, Gunslinger [75] is a midair interaction technique relying on arms-down hands postures and gestures focused on large tactile displays. The midair gestures are performed near tights in the viewing frustum of the LeapMotion device. The touch interaction and midair interaction vocabularies are similar and coherent so that all of the actions the user can accomplish with touch can be accomplished in midair. During tactile interaction with one hand, the other hand is used to complement it (*i.e.*, tactile manipulation with dominant hand, midair navigation with non dominant hand).

This work shares the fact that the interactive spaces are not materialized, not visible or visually represented, and that are attached to an high-level entity (e.g. world or user's body part). In the next section, I present previous work where spaces are attached to an existing device.

## 2.2 Invisible Spaces Around the Device

We are interested here by related work using the space around a device. A lot of them are mobile phones but some work investigated around the desktop and laptop space.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



### Around Mobile Device

Mobile devices such as smartphones ans smartwatches offer anywhere-anytime interaction to users. One major drawback of interaction with mobile devices is the lack of interaction space and the occlusion because of touch interaction. These issues yield to propositions using the space around the device as an interactive plane or volume with midair gestures and postures.

In order to alleviate the occlusion issue on mobile devices, Lv et al. [76] propose to use the integrated camera of smartphones for in-air interaction with the fingers. The device is attached to the forearm so the camera is oriented toward the fingers. The interaction volume is then located on the users hand. In the prototype, the authors propose to use extension and flexion gestures of the fingers to interact with a circular menu showing several applications. In the same vein, but compatible with typical touch interaction, the authors in [103] propose to use the space around the mobile devices such as smartphones or smartwaches to extend their interaction space and enrich the interaction using gestures. They present a new in-air static gestures real-time recognizer that only uses the RGB camera of the devices allowing to use in-air gestures without having to modify the devices. Several scenarios are presented to illustrate how and in what contexts leveraging in-air gestures around devices extend, enrich and ease interaction. As the algorithm process camera RGB stream, interactive spaces are then in front of the cameras. Gestures can complement tactile interaction allowing to modify several parameters at once to switch mode while interacting on screen or invoke menus and select an item or else highlighting text by touch while invoking a magnifier lens and controlling it via gestures. Doing so, the authors argue it reduces effort made while interacting on the device.

This previous work essentially uses the integrated RGB camera of the devices. Doing so, the interaction volume is then behind it. Others exploit the devices sides as an extension. Abracadabra [50] is magnetically-based input technique for small mobile devices like smartwatches to extend their interaction volume. Users index finger is equipped with a small magnet to override the Earth's magnetic field and the magnetometers are mounted on the device. Then, finger tracking is possible in a range of about 10 cm. Here, radial control is presented as well as pointer control. Interaction takes place in the plane defined by the screen. Tap is performed by flicking the finger up and down through the plane. Ad-binning [53] is an interactive method designed to use the space around the device as a storage for virtual content such as web pages, bookmarks, notes,...The space around is a half-circle around the device split in five sectors



© 2017 Tous droits réservés.
lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



which are then split in several concentric bins. The authors also explored design factors such as bin size, bin location and organization around the device, binning and selection methods. Evaluation study showed that AD-Binning reduces browsing time compared to on-screen interaction. The authors identify three reasons for this: AD-Binning is in browsing mode by default; due to larger target sizes in space than those on-screen; AD-Binning leverages on spatial memory and proprioception to memorize where the best item was found. The authors recommend to concentrate binned content on dominant hand side to avoid screen occlusion when browsing along with size guidelines for bins and space. In this work, the interactive space is around the device is used as a virtual content storage and so is discretized. It is not the continuity of space that is exploited but users spatial memory. Each bin can be associated with a content either automatically like a stack or manually by the users allowing a better learning of items location. The gestures used in this space are then essentially pointing and selection gesture. In [20] , the authors present a new class of interaction gestures using the complementarity of touch and air gestures. Touch events help delimiting midair gestures and midair gestures add expressiveness to touch events. A vocabulary is proposed along with a prototype using some Air+Touch gestures on smartphones. Here, midair gestures occurring before, between or after touch have different meaning. The proposed techniques offers a fluid interaction by giving a specific meaning (*i.e.*, vocabulary) to midair gestures segmented by touch. The interaction in the 2D space of the touchscreen is used to delimit the midair gestures made above the device.

The work presented in [67] proposes to use both behind and beside spaces for 3D manipulation on mobiles devices. In the prototype, a depth space camera attached to the device is used to capture the hand and the recognition algorithm used estimates the rotational parameters of the users flat hand gesture. Those parameters are then applied to the 3D object in the virtual scene. According to the user study, using either behind or beside space is faster than touch trackball method to rotate 3D objects. The interaction volume here is dependent of the camera viewing frustum and users arms length and directly linked to its position and orientation. Though the studied gesture in this work is flat hand posture for rotational control, the authors mention that with better hardware, other gestures could be exploited such as palm bending or finger gesturing,…to control all of the 6DOF for a complete 3D objects manipulation. 3D manipulation using around the device interaction have also been investigate in [12] but in the context of interacting with a remote display. The authors propose to use the cubic space around the mobile device as a space for translation and rotation of 3D objects. The mapping between hand movements and objects displacements is relative. Clutch is made by touching the screen. The device screen defines



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



the frame of reference. Rotation (resp. translation) is made by measuring the hand offset around the input reference frame and applied to the 3D object. The authors compared this technique to a tactile one [110] and to a tangible one using the phone as if it was the object still relatively mapped. Results of comparative study show that the tangible and the around-device interaction techniques outperform the tactile one while no significant differences were found between the two of them. Overall, users preferred performing around-the device for 3D manipulation. Avrahami et al. present in [2] a system for tangible interaction on and around a tablet. Thanks to two cameras placed above the screen filming the surroundings and to a visual recognizer for tangible objects. Several proof-of-concept applications are described. Such as a tic-tac-toe, an galactic arcade game, a penalty shootout and *Tabletgotchi* which is a game inspired by the Tamagotchi virtual pets. In the example of the penalty shootout, users throw a small soccer ball toward the tablet. Once the ball touch the tablet, the virtual avatar of the ball continues its trajectory in the goal. In *Tabletgotchi*, a physical toy zebra can eat, drink or sleep according its position around the tablet and its posture.

Another study has been made by Jones et al. in [59] to explore the design space of around the device interaction for multiscale navigation (*i.e.*, pan and zoom task). They evaluated four interaction techniques around the device against touch. All of them were different in clutch, speed control and simultaneity of pan and zoom tasks. They also analyzed the volume utilization around the device in order to draw users preference in this context. The results show that the best interaction technique for pan-and-zoom free-space interaction is what they call the 1Button Simultaneous technique. It consist of panning on the plane given by the back of the device (xy-plane) while zooming using the z-axis (*i.e.*, device front). The button serves to clutch. Without surprise, the analysis of volume utilization shows the users(all right-handed) naturally prefer to interact on the space at the right of the device. In this paper, the interaction volume preferred by right-handed users is the cuboid at the right of the device. It is defined by the plane in which the back of the device is, and the normal of it corresponding to the z-axis of the mobile. It is not delimited in dimensions. The space is used as a pan-and-zoom modality. Considering the best interaction technique (*i.e.*, One Button Simultaneous), it is the continuous nature of gesture and the different actions are mapped to directions and amount of displacement.

The presented work in [15] is a prototype of a mobile phone with proximity sensors that allows multitouch interaction with the surface around the device when rested on a flat surface. Once on a flat surface, typical gestures such as pan,zoom and rotate are handled to interact with the content on the mobile device. This work transforms the inert surface around the mobile device into



© 2017 Tous droits réservés. lilliad.univ-lille.fr

an interactive one. Other work investigates the use of inert surface particularly around the desktop and the laptop.

#### Around the Desktop and the Laptop

Though workspace is larger on desktop and laptop than mobile devices, using the space around it has been investigated by researchers to address other issues. In the following, we will see several propositions that use spaces around desktop and laptop in different ways.

Mouseless [80] is an invisible mouse for computer. It works with a infrared laser and camera. The laser creates an invisible layer on a surface in which users fingertips are then illuminated and seen by the infrared camera. Hand position, right and left click can then be recognized and translated as mouse events. Other gestures can be recognized by Mouseless such as curling fingers toward the palm for scroll, or zooming gesture touch-style with the thumb and index. In the paper, the authors also propose a multilayer prototype where above-the-surface gestures can be performed such as drag and drop by pinching on the surface, lifting the hand, moving to the desired location lowering and releasing the pinch do drop. With Mouseless, the authors created an interactive space next to the computer to emulated an invisible mouse. In this space, continuous gestures as well as discrete gestures are part of the interaction. Here, hand gestures are to replace mouse interaction.

Using the space around the desktop has been proposed in [56]. The authors argue that working on computers separates objects in several workspaces: the primary workspace, the secondary workspace and the off-screen workspace. Spatially, these workspaces have different importance. The primary workspace takes almost all the screen room, the secondary space contains artifacts related to the current task such as menus or tools and the off-screen workspace holds the remaining concerning the current task or not is not visible by the users. Hence, when users have to access off-screen entities, they have to explicitly bring the object of interest into the primary space even if it is for a short time. Moving secondary and/or off-screen workspace on the desk allows to make space on the screen for the primary workspace, and eases the access to off-screen artifacts. The authors take stance on not altering the desk as much as possible. They propose to add an depth sensing camera above the desk to track users hands gestures around the desktop. The use of the desk surface (which is the interactive space) is restricted to placing items (*e.g.*, applications, tools,...). Thus



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

interaction gestures are mostly hovering and tapping. Along with this prototype, the authors ran two studies to understand how users would use such a system. In the first study, the authors show that people arrange items in grids on the desk. The the results of the second study tells us that targets less than 10cm wide lead to more errors in retrieving items. In both studies, having no feedback on which item users retrieved decreased retrieval time but might cause more errors. Results suggest that the number of items should be limited to 10 items on the desk. Though this work has the particularity of storing shortcuts or command around the desktop by leveraging on memory an proprioception, interaction in these last three papers is quite standard (the hand replaces the mouse and typical touch interaction).

Though the work presented in [123] does not only focus on around the desktop interaction, the system aims to enable typical mouse, keyboard and 3D control thanks to a simple sheet of paper on a near computer. The paper sheet is tracked in the space along with users hands. Users fingertips are tracked so that mouse and keyboard input can be emulated. For the keyboard, a keyboard layout is printed on the paper sheet for ease of use. The mouse emulation is done as if the paper sheet was a regular trackpad. The 3D controller acts like a proxy of a 3D model. Users can rotate the paper sheet to examine a model. Here, an interaction space is attached to a paper sheet that then represents the space. The space itself is invisible and is only supported by the sheet of paper. In the paper, the authors propose to use the space as a keyboard and propose to print the layout to ease the task. Having to print the layout gives a hint on the necessity in some cases to get a visual feedback or a materialization of the interactive space. In the next section, we will see propositions and prototypes that give to users this feedback.

## 2.3 Visible Interactive Spaces

Having extra interaction space brings new possibilities,it can improve interaction and/or enrich it. Invisible interaction spaces works because the effect of acting in it can be perceived. In the following, we will see work that uses the concept of interactive spaces with a visual materialization of them. These visible spaces can be seen on the body, around a device, in the air, on a flat surface,…



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

### Peephole Displays

Inspired by prior research [33,34] that showed the value of spatially tracked displays acting like windows on a virtual world of information, one kind of materialization of virtual interactive planes are peepholes displays. A virtual interactive information plane existing in space is revealed by moving around a tracked handheld touch display such as a smartphone or tablet [40,78,108,121]. The same principle is used in [64] but coupled with a pico-projector attached to the mobile phone. The viewed space is then larger than on a mobile device.

The concept of window the virtual world can also be applied to 3D virtual world like in T(ether) [69]. It is a spatially-aware interactive system for 3D manipulation and animation of objects. The handheld tablet is a window into the virtual space in which the users walk and move. Like a dimensional portal. In this work, the tablet is a reference frame to help exploiting proprioception above, on and behind the device. The space above is used to control global parameters of animation such as time and keyframes. The 2D surface and GUI of the tablet are used for freehand drawing with a plane constraint and to give feedback on the animation timeline (*i.e.*, ), granularity and current keyframe). Finally, the space behind the tablet is the direct manipulation space of the virtual world's 3D objects. The main advantages of the system is the multi-user feature for basic 3D modeling tasks thanks to an untethered device for each user. Here, regarding the MIS framework, there is one big world-anchored virtual space in which users moves and interact. Plus, three interactive spaces device-anchored are used for specific commands. In every spaces, the continuous nature of gestures is used as well as their well-known expressiveness.

### On-body and Around-body Visible Spaces

Using the body to display information as well as interacting with it avoids having to reach for a specific device and provides a larger area for interaction and visualization. Plus, it leverages on proprioception, kinesthetic memory, haptic and bilateral tactile feedback.

Skinput [52] uses acoustic signal to allow skin input (taps) on the arm. An armband embedding two arrays of five sensors sends the recorded acoustic waves to a main application that segments and classifies the taps. The authors focused on using users forearm as it offers an easily accessible large surface, the



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



method could be used on other part. In order to find the layout maximizing tap classification (*i.e.*, tap location) and test the system possibilities, the authors compared three layouts of tap-area: one on each finger, five and ten distributed on the forearm and the hand. The classification accuracy of taps is higher in the five-buttons condition (95% classification accuracy). Overall, the classification is 87% accurate across the three conditions. Though using the forearm space as an input device as an invisible space is possible, the developed proof of concept couples Skinput with a pico-projector directed to the forearm. Doing so, it facilitates interaction as learning buttons placement demands time to users.

While Skinput explores using the forearm of the user and simple tap gestures, the work in [51] investigates how the arms and hands can expand on-body-interaction design space. Body is also used as input and output but uses arms and hand gestures. Here are some examples of use. Whether the palms are directed downward or upward means for example to answer a yes/no question of an application, each side space of the hand having a particular meaning. The forearm space is associated to a linear menu. By positioning spaces in front of the user, mode switch is done by simply moving the hand in the right area and the GUI will then be projected onto the hand. This paper clearly propose a large number of use of on-body interaction. The important thing is that in any case, one or several interactive spaces are declared with an associated function and gestures or postures are defined to interact in these spaces.

OmniTouch [49] follows up and extend the concept using everyday surfaces as well as body parts thanks to a pico-projector and a depth camera shoulder worn by users. It enables graphical interactive multitouch input (*e.g.*, pan, zoom, rotation). The system tracks finger and surfaces (*e.g.*, planar surfaces, organic surfaces such as hands or arms,...). Hence, it allows typical interaction with conventional interfaces/widgets like projected keyboard but also multi-surfaces interaction (*e.g.*, using the wall and the palm). The authors also propose to use the surface properties to infer some context parameters. For example, a curved hand instead of a completely flat one would mean a private context interaction. Whereas a vertical surface (*e.g.*, on a wall) would mean a public context. This work allows to create an interactive plane wherever users want and to interact with it in a standard manner. Thus, touch interaction knowledge can be transferred to any everyday surface. Moreover, the authors propose to use the geometric properties of the space to adapt the interaction.

The previous work is presented as wearable interactive systems that uses the body as an input and output device. Instead of using the body itself as a device, other work builds an virtual interactive space around the body to use it



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



as a data or app storing space. Allowing quick switches between applications. One could say that this work is the visible derivation of Virtual Shelves [71].

We saw in the previous section the mSpace concept in [18] where the phone position around the user was associated with a specific application allowing fast switching. In the remaining of the article, the authors also present two additional concepts : pSpace and m+pSpace. With pSpace, a pico-projector is fixed to the phone displaying the current workspace while the mobile screen displays the workspaces spatial arrangement. The phone position is represented by a white dot so that users always know where they are in the space. With m+pSpace, the main workspace is on the phone while the pico-projector shows a selected secondary workspace according to the phone position. In this work, for pSpace, the space where workspaces are stored is displayed on the phone while in m+pSpace, it is to be learned by exploration. In a study comparing the different concepts, participants unanimously said the visual feedback of the workspace was not useful as spatial memory was enough to remember the spatial arrangement. We can see each workspace as planes with a specific position and orientation around the user. Once the phone in this plane, the stored content is displayed.

The main drawback of the last presented papers are that visual feedback is done in public whether the interaction is private or public. HWDs can solve this problem of intimacy during interaction.

The work of Ens et al. [31] is a solution that enables multitasking on HWDs. The sphere around users is divided into virtual windows with which users can directly interact. The authors present the results of their studies for designing a personal cockpit. Three studies were made for the design of the personal cockpit. The first to determine the best window angular size relative to field of view and the best apparent distance from the user. The second to determine the best input parameters which are, the spatial frame reference (world-fixed, body-fixed and view-fixed), the distance of the window and window location. The third one compares different layouts. From these studies, the authors design the best personal cockpit with a world-fixed reference frame, with the windows 50cm from the user's right shoulder and a layout of 4x4 windows. The authors explored the design parameters for a HWD-based multitasking system. Their results give some guidelines for future similar solutions like MIS-inspired ones.

The early work of Feiner et al. in [32] presents also a concept allowing 2D windows in 3D augmented reality. Three different types of windows are described depending on where there are fixed. Their can be fixed to a position in the users



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



| Group | Dimension | Values | | |
|---|---|---|---|---|
| ***Reference Frame*** | *Perspective* | egocentric | exocentric | |
| | *Movability* | movable | fixed | |
| ***Spatial Manipulation*** | *Proximity* | far | near | on-body |
| | *Input mode* | direct | indirect | |
| | *Tangibility* | tangible | intangible | |
| ***Spatial Composition*** | *Visibility* | high | intermediate | low |
| | *Discretization* | continuous | discrete | |

Figure 2.1: Ethereal Planes Framework.

surroundings, display-fixed windows are windows that are attached to a display whatever the users position is, and finally world-fixed windows. The interaction is made thanks to the mouse to specify a 3D location of a windows or interact with them.

Others proposed taxonomy to better explore derivations of such systems. The framework proposed in [29] aim to support the design of HMDs interfaces. The authors focus on designs of virtual 2D spaces. From the literature, the authors devised a taxonomy that is the basis of the framework. It is composed of seven dimensions grouped in three distinct categories. We see in Figure 3.2 the seven dimensions and their possible values as well as the three groups.

The perspective dimension delimits egocentric and exocentric reference frames. Egocentric reference frames denotes that the virtual 2D interface would be linked to the users bodies. In exocentric reference frames are relative to a real-world point. In the movability dimension, refers to whether or not the workspaces are movable. The proximity of an information space indicates the distance with the user. The possible values are on-body, near and far (*i.e.*, beyond arm's reach). Input mode is to differentiate indirect in put from direct input. Where indirect input includes cursors, ray-casting,…and direct input refers to direct touch by hand, touch or stylus. The dimension of tangibility informs if the workspace is touchable or not. Visibility designate the amount of available information on the workspace. High visibility means that the workspace is largely visible. Intermediate visibility stands for viewing constraint during the interaction. Finally, low visibility means that few information or visual feedback is available. Discretization dimension defines whether the space is continuous or composed of units. The work presented in [30] also focuses on 2D interfaces for HMDs as it is built on the Ethereal Planes framework. It introduces



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



a layout manager that proposes a layout for applications windows depending on the surroundings of the users and that satisfies three constraints: (1) conform to surface structure (*e.g.*, boundaries occluding objects); (2) Maintain layout consistency relative to the user; (3) Preserve background information to avoid interferences. These last two articles focuses on 2D interfaces for HMDs and is presented as a core framework for design and classification. It is actually very close to what I propose in this document. Here, the authors focus on visible interactive spaces and detail the equivalent of the Interaction Attributes and the Reference Frame component for such spaces and context of use.

### Extension of Mobile Devices

In this previous section, we saw different propositions of how exploit interactive spaces with visual feedback on-body or around the body with projected feedback or thanks to HWDs. In all these papers, interactive spaces were attached to the user or to a body part. Interactive spaces can also extend the visual feedback range of mobile devices as well as the input space.

In [41], Grubert et al. take advantage of both head-worn displays and smart-devices (*e.g.*, smartwatches) to overcome their inherent drawbacks. Which are poor interaction with head-mounted displays and small display area for the smart-devices. They enlarge the display space of the smart-devices thanks to the head-worn display and allow interacting with the latter thanks to the former. They propose a platform, named MultiFi, to implement multiple-display compatible interface widgets. This platform allows different modes of alignment between devices: (1) in the body-aligned mode, devices share an information space attached to users body. The touchscreen device provides detailed information about the area in which it is while the HMD displays the remaining of the space; (2) the device-aligned mode attaches the space to the touchscreen. The HMD extends the screen of the touchscreen; (3) with the side-by-side mode, the devices are separated. The combination of both touchscreen device and HMD allows spatial navigation by moving the touch device in addition to standard multitouch gestures. Here, depending on the alignment, the interaction should be done either via touch or midair gestures. In any case, the interactive space is used to display more data.

Similarly, the authors in [77] propose a design space to explore interaction techniques combining smartwatches and head-worn displays. The design space is organized according to which input and output devices are used. For example,



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



We can use both the HMD and the smartwatch for input. This way when the smartwatch is in front of the user, it can be started automatically. If the HMD is used as output only, the user can see who is calling and just blind-tap the watch for rapid answering.

In the previous presented work, we saw that spaces are used to access and interact with virtual content. In [102], the content is real on a sheet of paper and augmented by a mouse equipped with a pico-projector. The virtual clone of the paper-content can be edited thanks to a digital pen. The projector-mouse can be moved across the digital paper to augment the region of interest. The projected space is an extension of the mouse and allow to display additional data to the one on the paper sheet. So the space is physical and virtual at the same time. The system interprets pen strokes to interact with the virtual content projected on top of the paper.

### Desktop and Worspace Augmentation

The use of extra interactive space has also been used to augment desktop and physical workspaces. Doing so, users have access to specific and contextual content while taking advantage of multitouch interaction on the desktop, for instance.

Relative to what we saw in the previous sections, augmenting the desktop thanks to HMDs is proposed in [100]. Here, the authors propose a derivation of the Gluey prototype [101] for desktop environment. The Gluey prototype unifies input across display devices such as tablets, desktop, laptop thanks to HMDs. For example, users can copy a figure on the desktop screen and by moving their head toward a tablet to paste the figure on it. In Desktop-Gluey, HMDs allow to extend beyond the physical desktop. Users can add extra virtual displays, spatially arrange them, keep this layout far from the desktop and still interact with it, and even share their view with other users in a collaborative context. Serrano et al. propose to use planar surfaces as displays to extend the physical desktop and to make it mobile. In order to interact with these virtual displays, the Gluey prototype allows for diverse means of communication depending on what is available at the time of the interaction. Users are free to use either the mouse and keyboard or, the tablet or event midair gestures. Thus, the input and the output interaction volumes are distinct. It is important to note that few is said about the available interaction means.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



Rekimoto et al. augment the desktop and laptop workspace in [94]. Users can move content from they laptop to the table or a wall to share it or better visualize it. They call these extra spaces *environmental computers* and describe them as an extended desktop. The interaction with such spaces is made only with classical mouse and keyboard. These extended desktop are then more like extra displays. In the case of multiple users, they can interchange data by simply dragging the object of interest from their laptop to the wall/table. In the prototype, users can also bind data to physical objects. A mobile version of this work is PlayAnywhere [118]. It is a compact system combining a camera, a projector and an infrared(IR) illuminant. It can be placed on a table to make it interactive. Users can then manipulate content such as images or other type of data. The system differentiates hover from touch and recognizes visual codes allowing for tangible interaction. Similarly, Kane et al. present Bonfire [61], a laptop-mounted system that projects displays on the sides of a laptop. The projected displays are interactive thanks to cameras allowing for gestures and objects recognition and tracking. For one display, the system consists of one laser projector and one RGB camera. Bonfire allows to use the surface around the laptop as an extra interactive displays. It allows tapping, dragging, flicking and interacting with external interactive or inert objects in the laptop surroundings. In any case, Bonfire enrich or augment the interaction with or between devices/objects by using extra interactive display planes around the laptop. Interaction here is limited to standard touch interaction.

In [13] the surroundings of the desktop are used to enrich the desktop interaction. The authors choose to study four interactive surfaces of the desktop. Comparing one-handed versus two-handed interaction in each zone shows that top and bottom zones relative to the keyboard are well suited for two-handed interaction whereas left and right zones fit better one-handed gestures. From the findings, the authors present a prototype combining multitouch and desktop interaction. They attribute each of the four zones to specific tasks. The top zone (*i.e.*, above the keyboard) is used to display an abstract version of a chosen window dragged from the screen to the zone. The bottom zone is a dedicated to window management and task bar. Thumbnails of the opened windows are displayed with the same spatial organization than on the monitor. Plus, users can resize them, minimize, maximize and restore them directly from the zones thanks to multitouch gestures (zoom, double tap, flick). The mouse area is also augmented with a digital mouse pad where contextual menu commands are always visible such as copy, cut and paste,...Finally, the left zone next to the keyboard is a *Multi-Functional Touch Pad*. The authors implemented several functions for this area: degree of freedom control, control-distance gain adjustment of the mouse, secondary cursor control, custom tool palette. These functions



© 2017 Tous droits réservés. lilliad.univ-lille.fr

can be either used in parallel with the mouse or sequentially. In this work, four interactive planes have been studied and integrated to the desktop to enrich and improve the experience. Guided by the study, the top and bottom zones are dedicated to two-handed multitouch interaction and side zones are designed for one-handed multitouch interaction.

### Whole Room Augmentation

In the previous section, we saw work that focused on augmenting the desktop or the workspace with extra displays and extra interaction volumes or surfaces. This augmentation can also apply to larger physical contexts such as a whole room/ a house or an office. In the following, we will see such work.

In [119], Wilson et al. present a room equipped with depth cameras and projectors that enables any surface augmentation (*e.g.*, table, wall, display, user's body,…). The system can simulate interactive surface, manipulate virtual object in space (*e.g.*, picking up objects), use their body for menu display and navigation and to transfer data from one interactive surface to another by touching the data with one hand and the destination surface with the other. Gugenheimer et al. propose in [42] to explore in-house augmentation in different domestic environment. After conducting in-situ user study and exploring use case, the authors implemented a portable projector-camera system for domestic use. The system features a projector, a depth camera and two servomotors. It allows to transform any surface in an interactive space. In [66], it is the user's palm that stands for a display not surrounding surfaces. Depending on where the user is the displayed information might change. The room in which act the user is equipped with projectors and multiple depth cameras. This way users have always access to contextual information at the same location (*i.e.*, their palm). The navigation of projected information on the palm being done by moving it in a vertical plane.

In this previous work we saw that the extra virtual spaces are mostly planar either in thin air or on surfaces. The input methods being direct midair manipulation or multitouch, this work does not explore enough the potential of midair gestures by reducing it to standard pointing. In the following we will see the use of midair gestures above tabletops or large displays as it is the environment in which most of the work on gestures in interaction volumes has been explored.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



## 2.4 Extension Spaces for Tabletop and Large Displays.

In order to reach distant area of the display, we introduce midair interaction to the tactile interaction. Some previous work uses midair and tactile to enrich interaction. In the following, all the interactive volumes are cubic and above the tabletop or display. We will focus on how gestures are used or what are the specific interaction parameters.

In the early work [98] of Schmalstieg et al. use a transparent pad and pen above a stereographic tabletop. The proposed uses are palette tool, window tool (*i.e.*, magic lens), through-the-plane tool and volumetric manipulation tool. This setup enables bimanual interaction in a stereographic environment, having 2D widgets in a 3D environment and continuous or discrete controls. The volumetric tool uses the pad as a fish net to select objects. In this work, the props are used in a mimicking way like the fish net and the palette tool. Using tangibles in interaction volumes has been done for different uses in other papers. In order to better view 3D virtual content above tabletops, in [106], the authors propose to use handheld magic lenses. A tracked sheet of paper is used as the magic lens on which is projected the corresponding content according to its position and orientation. The implemented magic lens are, volumetric information (*e.g.*, MRI visualization), zoom and time control(*e.g.*, videos navigation). Here the focus is on data exploration. The input is restricted to the sole manipulation of the sheet of paper. Whereas in [104], the presented system uses the tabletop to display a 3D scene and interacting with the content is made thanks to tracked mobile multitouch screens and head location. The mobile screens are like windows on the scene displayed by the tabletop. They allow magic lens based interaction(*e.g.*, wireframe view on the mobile screen, mesh view on the tabletop), manipulation of the scene objects and visualization. Users select an object by aiming at it, touching the selection button on the mobile screen. Their can also observe them by aiming and double-clicking. Then, observation is done by simply manipulating the tangible mobile display as if it was the object. When holding an object, getting out of the tabletop space deleted it. In this work, the space above the tabletop is used for tangible interaction with a mobile display that allows 6DOF manipulation of the scene objects and advanced visualization.

Hilliges et al. [57] propose a static extension above the 2D display that allows the user to perform 3D gestures above the screen. In that case, the interaction volume is static, always active and of a predefined size (here the screen size). There is a direct mapping between the hand above the surface and the output (shadows displayed). As long as the system can detect the user's hands, the



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



user can manipulate objects of the 3D scene with 6DOF. TractorBeam [85] is a hybrid pointing technique allowing seamless switch between touch and remote pointing on tabletop with a stylus. The pointed position is calculated thanks to the position and orientation of the stylus. To select an object, the user has to press the button on-top of the stylus even when the stylus is touching the surface. There is no difference between tactile and midair but the tactile interaction is then similar to mouse interaction. The interaction volume above the tabletop allows here to continuously interact on the displayed content on the tabletop but do not take advantage of midair gesture expressiveness as interaction is still done with the stylus. On the contrary, in [11], Benko et al. propose to seamlessly transition data from a 2D multitouch display to the 3D space above. For example, users can grab the picture of an object on the display with a grab gesture, lift the hand to bring it in 3D above the screen (*i.e.*, Pull) and then with the reverse gesture put it back on the screen as a picture (*i.e.*, Push). Users can also connect the 2D representation with the 3D model to examine both of them at the same time. The manipulation and transformation of the 3D model is made by pinning it to the table, and manipulating it via 2D touch-based interaction and context-sensitive menus. This work does not take advantage of midair gesture for the direct manipulation of objects as the authors argue that the lack of haptic feedback makes the tasks cumbersome and prefer to use the passive haptic feedback of tactile interaction. Here, midair gestures are more used for their punctual metaphoric meaning such as pulling and pushing objects between 2D and 3D spaces.

In the context of ambient display interaction, authors have proposed to segment the space surrounding the display in different ones in order to better adapt interaction attributes to the current situation. In [112], the authors split the space in front of an ambient display in four spaces from close to the display to far from it. The act of moving from one space to another modifies the degree of how personal the interaction is. This action then smoothly modifies the displayed data on the screen as well as the possible interaction. According to the space. If users get closer in the personal space, more detailed data will be presented and touch screen input will be possible. On the contrary, when in the ambient space (farthest space), only coarse data are displayed and no command invocation is possible. Moving from one space to another means that users want more or less details or interaction possibilities. The authors In [122] proposed a two-modes interaction technique depending on the distance between the user and the wall-sized display : the 'Near-Mode' and the 'Far-Mode'. Near-Mode is dedicated to touch interaction whereas Far-Mode is for midair interaction. The switch between those modes is determined by whether or not the user is tracked by the Kinect placed just under the screen. Then the space in front



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



of the screen commands the mode of interaction. To switch between them, users have to move their bodies in the appropriate space. Similarly, in [26] Dingler et al. propose to split the space in front a display in four sub-spaces allowing from touch to coarse midair interaction. Though they argue to provide seamless gesture interaction and transition between those sub-spaces, they do not study this seamless transition. One of their study elicit manipulation gestures in the different spaces. They argue that assigning different sets of commands to different spaces will ease interaction. This mapping between spaces and commands should be done according to users preferences. Then, those interactive spaces would be define by the actions they allow and the user. This is one of the MIS concept motivation but here, through a preliminary study, the authors fixed the dimensions of the spaces around the screen.

In the past years, interaction volume above tabletops like interactive workbench have been investigated for 3D modeling. As researchers wanted to take advantage of our ability to represent forms with our hands, midair gestures for 3D modeling was studied and integrated to 3D modeling systems.

To extend 2D sketch application, Forsberg et al. in [36] propose a modeling system allowing 2D pen input, midair gestures and vocal input. Users work on and above a modified Responsive Workbench [68] called ActiveDesk. 2D pen input is used to sketch the 3D model in the viewing plane while 3D input is for manipulative tasks such as translation, rotation. Users have to grab a tracked prop on the edge of the ActiveDesk and use it as a proxy of the virtual model. Bimanual interaction can be used while drawing or examining the 3D model. The non-dominant hand is used to either control the camera view or the 3D model orientation when annotating it.

In [97], the authors introduce a system to create *organic shapes* by using either hands or tangible tools on the semi-immersive virtual environment Responsive Workbench. Users draw with their hands surface like we would do to describe a vase with our hands, by bending our hands and drawing a stroke with them. The hand shape give the cross-section of the surface. In order to manipulate the created 3D models, tongs are provided. One tong allows to move the model, a pair of them is used to scale the model. To erase a part of the model, users take the eraser (a silicone tool) and the magnet tool allows deformation of the model.

The modeler presented in [115] proposes designers drawing direct 3D curves or 2D curves in a restricted plane with a tool on the Responsive Workbench. Doing so, The authors want to facilitate 3D modeling during the creative process. The curves can then be modified as wanted, smoothed, copied, mirrored,. . . The



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



curves can then be filled to create a surface. Switching between the different tools and function is made through a menu attached to the pen and triggered when users touch the 3D model with the tip of the pen. The menu is displayed at the tip of the pen, then users select the function of interest and press the pen button. The non-dominant hand controls the position and orientation of the 3D model. Once finished, the model can be transferred to a CAD software for further elaboration.

In [24,79], midair gestures are not segmented by touch. The surface of a tabletop and the space above it are seen as a continuum. In this configuration, several kind of gestures are proposed. In particular, *extended continuous gesture* results in an action that begins in a certain space and moves or finishes into another. The transition between touch and midair does not alter the current action. It is the same action. The authors take advantage of continuous midair gestures for CAD design as it is easier to sculpt an object with our hands than drawing it.

## 2.5 Conclusion

In this chapter, I presented previous work related to interactive spaces. We saw that there can be used in many different ways and in combination with a wide variety of existing devices. In some case, they can be used as proper interfaces, they can contain data or widgets, they can be used to enrich the interaction with a smartwatch or a tablet and even tabletops. This diversity of interactive spaces that has been proposed in Human Computer Interaction calls for structure and a way to describe and define each of these very different interactive spaces. This is what I intend to do by introducing the concept of MIS and its design framework.

In the following, I will present my contributions, beginning by introducing the MIS concept in chapter 3. After that, I will present work that materializes the concept in particular ways addressing different issues. The chapter 4 study the use of an invisible world-fixed MIS for remote mouse cursor control and present a prototype of remote cursor control on distant displays. The second contribution explores the MIS concept around the tablet in chapter 5. Finally, chapter 6 presents the use of MIS as an extension of wall displays to ease distant interaction while acting locally on the screen.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# Three

# Presentation of Mimetic Interaction Spaces

*"Man senses distance as other animals do. His perception of space is dynamic because it is related to action — what can be done in a given space — rather than what is seen by passive viewing."*

— Edward T. Hall, *The Hidden Dimension.*

## Contents





© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



Grasping the mouse, or touching the pad, is currently, by far, the most common way to start interacting with an application. Such paradigm imply both proximity between user and interactive system. For interaction situations in which distance between user and screen can not be avoided(e.g distant screen), and instrumented interaction may be difficult to deploy (public displays), or limiting (family in front of connected TV, work meetings, etc …), mid-air gestural interaction appears to have great potential for such contexts.

Midair interaction still has several drawbacks that are not overcome yet; moreover it is still poorly understood, quite apart from elementary tasks [95]. A common (wrong) approach is to think about mid-air gestures as "touch at a distance", as stated in [117]. This sentence refers to the fact that midair gestures are continuous. Gesture segmentation is then more complex compared to touch interaction where one begins a gesture with a touch and stop it by simply releasing the touch. Midair interaction is also a common channel in human-human communication. Gestures in nonverbal communication have been well studied and a robust classification emerged from it [27] . This classification is presented in section 4.1. Those gestures are performed in the space around us which is structured according to the current context. This arrangement of space in human-human communication is how space is socially and personally structured. Proxemics is introduced in [48] as "the term I[the author] have coined for the interrelated observations and theories of man's use of space as a specialized elaboration of culture". Those spaces are defined and perceived by what can be achieved in them. Hall defines several metrics to characterize those spaces. The size and position of spaces depend on several factors: the environment, the cultural context, the degree of acquaintance with the interlocutor, the hierarchy and the senses implied in the interaction. All of these factors make our perception of space and interaction with others completely dynamic.

Some work has proposed to use static space to enable midair interaction and to facilitate the discrimination between intended command gestures to the system and the ones that are not. By leveraging on our dynamic vision of spaces, we extend these past propositions. Instead of interacting in a pre-defined static space, users create and delete their own interaction space at any time and place thanks to a simple gesture that mimics the interaction space. This chapter presents the concept of MIS published in the first part of [91], gives the design framework of a MIS decomposed into 4 components and discuss the state of the art of chapter 2 through the lens of the proposed design framework. I will then explore the design space of MIS and describe several interaction scenarios to illustrate how the concept of MIS apply to different contexts of use. Finally, I



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



will discuss an interesting side of the MIS concept which is its modularity and its reflective nature.

## 3.1 The Concept of Mimetic Interaction Spaces

We present here the concept of Mimetic Interaction Spaces (MIS) and *MIS gestures*. A MIS is a delimited sub-space of the user's space, used to perform interaction gestures. It can be of arbitrary dimension i.e. a 1D curve, a 2D shape or a finite volume depending on the application. The chosen sub-space is simple enough so that it is possible to evaluate whether or not user's hand is within this sub-space and if gestures shall be taken into account for interaction or not. The *MIS gestures* are defined as the set of user's gestures which can be performed within such space, to interact with it, as well as to create or delete it.

It may relate (but not necessarily) to a physical object, or an imaginary representation of it. By gesturing on or in the MIS, the user can interact with a distant screen, e.g control a mouse cursor on an invisible touchpad (planar MIS). We think this concept is interesting because it is more specific than the standard understanding of mid-air interaction, while obviously leaving quite an interesting design space. Once a MIS has been specified, user can use it to clutch easily delimiting its interaction gesture. A simple inclusion test (in the case of MIS volumes), or proximity (in the case of MIS shapes or curves) can be used by a gesture acquisition system to know if the gesture shall be taken into account for interaction or not.

In the case where the area is user-defined, an interaction gesture may be used for specifying an interaction area through all its geometrical (position/orientation/dimensions/...) parameters, hand posture, or ad-hoc shapes, may either be used for such a command. An application may be set up using some specific class of MIS area (e.g. a cube), but may leave user(s) with the ability to activate several areas of various dimensions at the same time (whichever the interaction purpose).

### Design Framework and MIS-based Interaction Technique

As seen in the previous section, our interactions with others are structured into spaces around us. Spaces have a specific use, dimension, shape, etc...and are



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



dynamic as they change according to the situation and intentions of users. Since the MIS concept build on this observation, we first wanted to define what a MIS is. The definition must cover the numerous characteristics and diversity of the possible spaces. We then proposed a design framework that covers the different aspects of MIS.

A MIS is a virtual interactive sub-space with four characteristic components detailed as follows: *geometric definition (GD), input reference frame (IRF), action reference frame (ARF), interaction attributes (IA)*. Each of these components are described below. We define a MIS-*based interaction technique* as a particular combination of these four components. A MIS can be created either dynamically or not, by the user or the interaction designer, from a gesture, or an event. The components are described in the following and summed up in Figure 3.1.

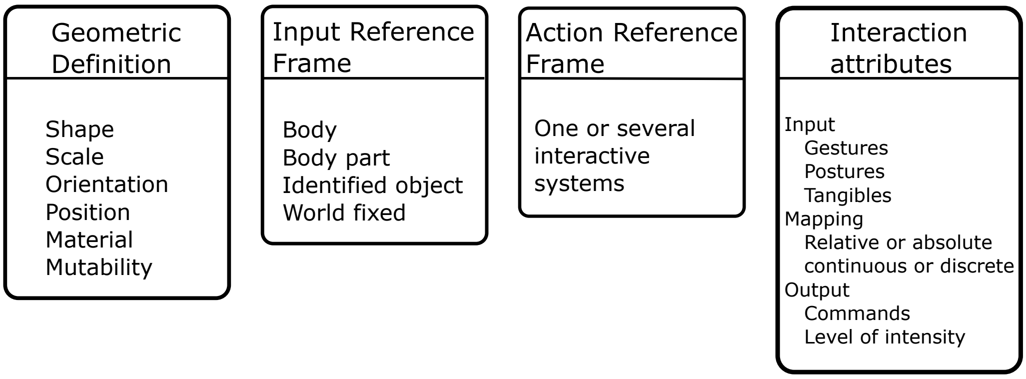



Figure 3.1: Short version of MIS Framework

#### Geometric Definition (GD)

We defined here the elementary geometric aspects of a MIS: shape, orientation, scale, position. They are expressed relative to the input frame of reference of the MIS they describe. The material of the MIS expresses different aesthetic aspects of the MIS like its opacity (if it is visible or not), its texture (both physical and visual).

#### Input Reference Frame (IRF)

This is the reference frame that links the MIS to the physical world in which the user evolves. In the general case, a MIS can be anchored to an *entity* of the real



© 2017 Tous droits réservés.
lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



world, possible entities being user's body or a part of it (e.g hand, head,...), or any identified object or the world (fixed position). If this entity moves, then the MIS moves as well. A MIS may have multiple IRFs. Then, a main IRF must be declared for the primary properties. Plus, it can be changed during interaction, using specific command gesture associated to the MIS.

#### Action Reference Frame (ARF)

This is the frame of reference that links MIS to the display with which the user is willing to interact. A MIS can have multiple ARFs. A default ARF is defined, that may be changed during interaction.

#### Interaction Attributes (IA)

The interaction attributes gather all properties that may be necessary to define the interaction technique based on the MIS defined by a set (GD, IRF, ARF). They may relate to human factors, data acquisition specificity, or any additional element that needs to be taken into account to define the interaction technique. Such attributes may vary in numbers, types and values, depending on the interaction techniques we target. Interaction attributes consist in three main parts : Input, Mapping and Output attributes. By defining them, one can fix how the MIS will be used and what will it be used for. Those parameters can be seen as semantics attributes as their give to the MIS its function and meaning.

The **Input** parameter includes the means users have to act in a MIS. Which are gestures, postures and tangibles. For each input means, the measured values that will affect interaction should be defined (*e.g.*, velocity of a gesture performance, distance between particular points of interest,...). The three means are described as follow:

- Gestures : the recognized gestures in the MIS. The gestures may be performed by a particular body part or if performed with the hands, bimanual or not. For each gestures, one has to fix the measure taken when the gesture is performed.
- Postures : the recognized poses of the body or other parts of it. This parameter details what are the poses and what is measured once recognized.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

- Tangible objects : in some cases, interaction might take advantage of tangible interaction. The tangible parameter defines what are the interactive tangible objects and what users can do with as input.

The **Mapping** parameter is the relation between what is performed and measured from the input and how it affects the output. Mapping can be relative or absolute, continuous or discrete.

The **Output** parameter describes the different possible triggered commands and their possible intensity level.

In this section, I described the MIS design framework and detailed all its components. This framework allows for a fine tuning of a MIS before implementation. Additionally, the framework offers a systemic way to explore the given possibilities of MIS interaction. In the next section, I propose a re-reading of the literature through the lens of the MIS design framework.

## 3.2 Re-reading of Previous Work Through the Lens of MIS

Now that the concept of MIS has been presented and defined, I will discuss, in this section, the papers presented in the previous chapter with regard to the MIS concept. I intend to show how the concept gives a mean to unify the previous work related to midair interaction and interactive volumes. In the following, I will insist on interesting and relevant features of the papers according to the MIS concept. Only few work propose MIS-like design framework and are mainly specific to particular contexts. In this latter case, I will discuss to what extent those are derivations of the MIS framework.

In the presentation of these papers, I do not use the word MIS but interactive spaces or volumes as I don't want to modify the first intent of the authors when they wrote their papers. But by describing the work, I will name attributes and components that are in the MIS framework to discuss the work. I will also quickly reintroduce the papers to ease the remembrance.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

### Standalone Invisible Spaces

All the interactive spaces presented in this section shares the particularity to be immutable non-user defined spaces. In some cases, users can control the position and orientation but not the shape, and dimensions(Fixed GD). The other shared features are the invisibility of those spaces (Transparent material in GD), and their standalone nature since the spaces are either attached to the user or to location in the world (IRF).

The concept of imaginary space in [46], is linked to the non-dominant hand and its lifetime is define by whether or not a specific hand posture is performed. Its position and orientation are also determined by the hand respective properties. In the paper, there is no mention of the space dimension. Thus, in the MIS framework, imaginary interfaces are plane shaped, hand posture attached, acting on mobile phones, allowing pantomimic gestures. The imaginary phone in [47], and the remote control in [25] are planes attached to the NDH palm allowing for pointing gestures. The work AirTouch Panel in [72] is also a planar space, anchored to user or world position. It acts on a smart TV and input is made with pointing and sliding gestures to control volume or channels.

In [111], the authors present an interaction system for "creative expression of 3D shapes". The vertical virtual slab is attached to the user at around 75% of arm length. Input is triggered by a pointing pose of the hand. The trace of the hand-arm motion draws the contours of the 3D model (Interaction Attributes).

With Shoesense [4], the space above the shoe (attached to it) has the shape of the camera view cone. The input is made thanks to three types of gestures. The finger count, the pinch and the arm triangle. The first two are discrete and mapped to either the number or the location. The arm triangle is continuous and mapped in absolute.

The wrist-worn sensor for freehand 3D interaction in [65] proposes a cuboid volume directly above the palm. The fingers gestures in the palm space can be either performed when interacting with an on-screen application or combined with touch or with no visual feedback. The major input gestures are hand postures since the rest of the input are given by the accelerometer.

PUB [73] (Point Upon Body) uses the forearm as an input interface. Pointing gestures and slides enable quick command triggering (next song, volume tuning).The interactive space is a surface that fits the forearm of the user.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



The three papers [18,19,71] are storage volumes where interaction is made only to retrieve information or data. All proposes interactive volume around the user. The exploration of the data is made via pointing gestures to select the data of interest. An external device can be used to visualize the data. The volume is a sphere or curved surface around the user's body. In [19] the body parts can also be assigned to specific storage.

In WUW [81], the authors propose to project information on real world surfaces while interacting in the viewing frustum volume of a head-worn camera. The interaction volume is the viewing cone of the camera and then attached to it. The user is wearing the camera, the volume is bound to the user's position. It is invisible and available gestures are postures (*e.g.*, to take a picture) or standard mobile interaction hand gestures (*e.g.*, to zoom). In Gunslinger [75] the volume for interaction is the viewing frustum of the Leap Motion and again, postures and hand motion are used for input.

The MIS design framework I propose allows to classify and describe the previous work. These papers all share the fact that the interactive spaces are not materialized, not visible or visually represented. Whereas the anchored entities (ARF and IRF), geometry properties (GD) and the rest of the interaction attributes (IA) are all different. And are attached to an high-level entity (e.g. world or user's body part). In the next section, I present previous work where spaces are attached to an existing device.

#### Invisible Spaces Around the Device

In this section, all the mentioned work share the following features: invisible space(GD), device attached (IRF) and acts on the device(ARF).

##### Around Mobile Device

The work presented in [76,103] uses the RGB camera which that implies that the interactive volume is directly linked to the camera's viewing frustum. The available gestures are either standard touch gestures but performed in midair or simple flexion and extension of fingers to control an attached menu.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

In Abracadabra [50] and Ad-binning [53] the interactive space is around the device and is used as either a rotational control and a conventional cursor control on a smartwatch or a virtual content storage. The space is plane extending the devices screens and the used gestures are essentially pointing and tap. The space is also discrete since divided into content or buttons except for [50] in the cursor control mode. In [20] where touch gestures are used to segment midair gestures, the space is the cuboid right above the screen of the smartphone and gestures are the designed vocabulary defined by the authors. Mostly, the gestures are performed by one finger to allow one-handed use.

The work presented in [67] and [12] authors propose to use the cubic space around the mobile device as a space for translation and rotation control of a 3D object. In both, it is the hand position and rotation that defines the object respective properties. In [12] however, the mapping is relative as it uses the offset of the hand initial position and the clutch is made with the NDH touching the screen.

#### Around the Desktop and the Laptop

The work presented in the corresponding section in the previous chapter is quite heterogeneous. So here I explain how they can be described with report to the MIS design framework components separately.

With Mouseless [80] the interactive space is the plane right next to the laptop and the available gestures are pointing, clicking, and scrolling as if the hand was holding a mouse. A zoom gesture is also available via the pinch gesture. Since the first goal of the prototype is to replace a mouse, the mapping is then relative.

The Unadorned Desk [56] uses the planar surface space of the desk as a shortcuts storage. Applications and other commands can be triggered form the space and are arranged in a grid layout of ten items. As it compares to buttons, the user has to perform point and tap gestures to select a command.

Tangible interaction is also comprised in the MIS framework. Such as the work presented in [88] by Pohl and Rohs. They propose to exploit potential proximity sensors of a mobile phone in order to scan the surroundings to use the near objects as input devices. The notion of space is here defined as the surrounding volume of the mobile device. Since it is the surrounding objects that enables interaction, the available gestures are the manipulation gestures of



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



such objects. The mapping is defined by both the inherent nature of the object, the associated command and the goal of the user. This paper is also relevant to the issue I mention in section 3.5 about the reflective nature of MIS-based interfaces.

In [2], the work implements among other tangible enables applications *Tablet-gotchi*. The surrounding space is split into several ones mostly corresponding to the extension of the tablet sides. The interaction is tangible and specific to the application. For example, in the case of the penalty shootout, it is the ball velocity and angle that participate in the interaction.

#### Visible Interactive Spaces

Like I said in the presentation of MIS, a MIS can have a material. This material can be transparent (making the MIS invisible), or not. In the latter case, I call this the materialization of the MIS. It is the actual MIS that is shown to the user, not just the content or output. Indeed, the MIS can be materialized by the content it displays or any other kind of materialization. In this section, the work have this feature in common.

#### Peephole Displays

Peephole displays-like work is a materialization example of MIS in the sense of the virtual world (the displayed content) is the MIS and the device used for interaction is part of the input (IA). The position of the device gives a constrained context and set of variables. It is true for both 2D spaces like in [40,64,78,108,121] and 3D MIS like in T(ether) [69]. But also, in T(ether), the authors use two MIS : one behind the tablet for direct manipulation, and one above the tablet for spatial control of global parameters.

#### On-body and Around-body Visible Spaces

Since I discuss here on and around body spaces, it is obvious that they are attached to the user or a user's body part. In some cases the output and the input are the same spaces.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

| Group | Dimension | Values | | |
|---|---|---|---|---|
| ***Reference Frame*** | *Perspective* | egocentric | exocentric | |
| | *Movability* | movable | fixed | |
| ***Spatial Manipulation*** | *Proximity* | far | near | on-body |
| | *Input mode* | direct | indirect | |
| | *Tangibility* | tangible | intangible | |
| ***Spatial Composition*** | *Visibility* | high | intermediate | low |
| | *Discretization* | continuous | discrete | |

Figure 3.2: Ethereal Planes Framework.

Skinput [52] and its follow-ups [49,51] uses the body as an interactive space or objects in personal space. The interface is displayed and mostly consists in buttons and sliders or menus. Drawing is also possible so the input is taps, swipes and slides, and strokes. The action can have consequences on either the space itself (*e.g.*, drawing mode) or the smartphone.

The pSpace and m+pSpace presented in [18] or conceptually near the Virtual Shelves where content is stored around the user in a sphere where the pico-projector allows to see them and so to see the sphere around. It must be noted that for m+pSpace uses two MIS. The invisible one similar to the mSpace, and the visible one similar to pSpace.

The work of Ens et al. in [31] is also similar to [18] as it implements a personal cockpit around the user. Here the cockpit is entirely displayed in front of the user. The displayed content consists in windows that can handle any standard content (map, calendar, messages…).

The framework proposed in [29] aims to support the design of HMDs interfaces (see Figure 3.2). The follow-up work [30] implements it in a reflective way since it proposed a layout manager for windows in HMDs systems.

The case of the Ethereal Plane Framework is interesting as it is the closest design framework to the MIS. It is a specification of it for planar, visible, and HMDs attached MISs. As we can see, many of the dimensions and values of Ethereal Planes are present in the MIS framework.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



### Extension of Mobile Devices

In order to make virtual spaces visible in the context of mobile interaction, the one solution is to use HMDs. In this section, all the work I talk about is based on HMDs devices attached to mobile devices. These extension spaces are ones of the first applications of MIS I thought of when I first designed the framework. The described interactive spaces are device-attached and are visible by the user.

In [41], Grubert et al. explicitly mention the creation of a space, an information space, that is displayed and with which we can interact via the touchscreen. Depending on the alignment mode, users interact with the information space by gesturing in it or near it. In the body-aligned mode, there is one user dependent information space, in the device-aligned mode, this information space is an extension of the touchscreen device and finally there are two separated space in the side-by-side mode.

The design space proposed in [77] explore interaction techniques combining smartwatches and head-worn displays. The authors propose here different distributions of input spaces and output spaces attached to the HMD or the smartwatch. In the case of HMD input, the interaction volume is the viewing frustum of the HMD camera, the space can be an extension of the smartwatch, or else completely independent displays. This work fits the MIS concept. The difference here is that the point of view is centered on the devices and how they are used. No explicit mention of interactive spaces is made. The design space is for devices only. So in the different cases, MIS concept is used differently. However, the authors do not detail enough for an analysis of the work with the MIS framework.

In the previous work, we saw spaces are used to access and interact with virtual content. In [102], the content is real on a sheet of paper and augmented by a mouse equipped with a pico-projector. Like I said in the previous chapter, the interactive space is both physical and virtual. With the non dominant hand, users augment a specific area while the dominant hand interact with both the physical and the virtual content through pen-based interaction.

### Desktop and Workspace Augmentation

The work presented in this section is also related to MIS and can fit the design framework but not fully as it proposes extra information space instead of extra



© 2017 Tous droits réservés.
lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



*interactive* space. Desktop-Gluey [100], HMDs allow to use extra virtual displays, spatially arrange them, keep this layout far from the desktop and still interact with it, and even share their view with other users in a collaborative context. In the MIS framework, Desktop-Gluey uses visible information spaces. There is no direct input attribute for these spaces. The interaction is made in other spaces or via other devices. Then, Desktop-Gluey is not clearly defined as MIS since we do not have enough information about how interaction works. It extends the workspace with other virtual spaces, it has geometric attributes, an action reference frame (the Operating System it extends), an interaction reference frame, but the interaction attributes are not well defined. It is not clearly said how the displayed data can be manipulated or modified. Without any information about the interaction in the space, Desktop-Gluey could be seen as a degenerated form of it. The only input is the manipulation of the virtual displays (creation, repositioning, etc…). But again, no information is given about those manipulation gestures. It is the same for the work of Rekimoto et al.presenting augmented desktop [94]. The *environmental computers* are more like extra displays. But here the said interactive space is represented by physical objects like a table or a wall. What it is visible here is the shared data on the space. With PlayAnywhere [118], the authors propose a tangible MIS based interactive system. It explicitly creates a dynamic interactive space where the user wants it to be. Input can be made via touch or tangible interaction to manipulate content or to modify it. The space is directly attached to the world coordinate where the system is setup. The same is true with Bonfire [61]. The projected displays are interactive thanks to cameras allowing for touch gestures and objects recognition and tracking. Except here, the spaces are attached to the laptop.

The work in [13] is really interesting in regard of the MIS concept and design framework. The authors studied the use of four zones around the desktop and apply they findings in a prototype where each space has its set of best suited gestures and commands. For example, the left and right zones of the keyboard are best suited for one-handed gestures. The mouse space is also augmented with contextual commands. Here each MIS around the desktop has its own purpose and component values (essentially interaction attributes values). For example, bottom zone is two-handed gestures enabled whereas right zone is only for one-handed gestures.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



#### Whole Room Augmentation

In this section, though the presented work shares the room augmentation context, in the MIS framework, they are quite different.

Indeed, in [119], Wilson et al. present a room that enables any surface augmentation (*e.g.*, table, wall, display, user's body,…). It is completely dynamic and interaction allows picking up virtual objects, transferring data from one surface to the other…Several MIS are used. Most of the presented interactions use world fixed MIS but others are user's body dependent. Concerning interaction modalities, multitouch interaction is implemented for MIS attached to surfaces while hand position is used for other kind of MIS (*e.g.* the vertical layered menu). In the work of Gugenheimer et al. [42] for in-house augmentation, the surfaces are made touch enabled by a projector-camera system that creates a MIS where users want. Though the majority of the paper investigates what are the requirement of such system, I focus here on the final system proposed in the paper. MIS are planar, attached to the system (that users can move) and minimal touch interaction is implemented. In [66], it is the user's palm that stands for a display. And interaction is made by moving the palm in space, contact between the palm and objects or also midair gestures for a specific command depending on which space is displayed on the palm (*i.e.*, context).

In this work, we see the application of concepts we saw previously. The use of body parts, around devices spaces, workspace augmentation (*i.e.*, kitchen augmentation). Here, only varies the scale of the interaction context.

#### Extension Spaces for Tabletop and Large Displays.

Here, all the work has in common the cubic interactive space above the tabletop and acts on the displayed data(GD, IRF and ARF). The only difference is the interaction attributes that I already detailed previously in chapter 2.

To summarize, the first three papers [98,104,106] propose tangible interaction with respectively a transparent pad and pen, a paper sheet and mobile touch displays. Mostly, they act as magic lenses on the displayed data.

In [11,57,85] it is the seamless transition between the 2D world of the display to the 3D world above it that is explored. For example by pulling object out of



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



the display. The gestures are mainly manipulation gestures of objects either 2D or 3D.

In the context of ambient display interaction, the three papers [26,112,122] propose to split the space in front of an ambient display in different zones allowing different kind of interaction according the distance to the screen. The level of detail displayed or the grain of interaction gestures and commands is mapped to the distance to screen.

I also discussed work related to virtual workbenches where our midair gestures skills are often exploited to form 3D models. They use iconic gestures which provide object properties like its shape or size (*e.g.*, shaping a vase with our hands) and pantomimic gestures which mimic a real world gesture.

The particularity of [36,115] is that they rely on our sketching skills as users draw curves and shapes thanks to a pen. Surface Drawing [97] uses hand shape and sculpture proxy tools to create and modify organic shapes. Whereas the papers [24,79] rely on CAD knowledge of users while freehand drawing is still possible.

#### A complete MIS description of previous work

In this section, I reviewed the previous work presented in chapter 2 but through the lens of the MIS concept and showed that very heterogeneous work can be unified thanks to MIS. **A detailed table in Appendix A** is available. For each previous work I discussed in this section, its complete MIS description is available in the table.

#### The case of Proxemic interaction

Though I presented the MIS concept with reference to proxemic theory, I talked few about proxemic interaction work. The work on proxemic has been largely presented to be close to ubiquitous computing and even the future of it [39].

Even though researchers often mention spaces around devices or people in which interaction take place, proxemic interaction is much larger than the MIS concept since it is the *relationship and arrangement themselves* between such



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



spaces in a room or a workplace that is explored and studied in most of the work on proxemics. For example, in the work presented in [5] the authors "imagine proxemic interaction as devices with fine-grained knowledge of nearby people and other devices" and talk about using proxemic interaction as part of an "ecology of multiple devices and objects". As well, in [70], the authors introduce *proxemics-aware controls*. By exploiting spatial relationship between a handheld device and appliances around, the authors propose to create *dynamic control interface*. Doing so, the authors argue that discovery, selection and control of appliances in ubiquitous environment is made easier. Basically, users have access to more or less controls over equipment according to their position and orientation in the environment.

In the remaining of this chapter I will first show how the MIS framework allows to explore the design space by looking at each components of the framework separately. Then, I will give some examples of interaction scenarios that use MISs in different contexts to show the adaptability of the concept to diverse contexts. Finally, I will discuss the modularity and the reflection of the MIS concept.

## 3.3 Design Space Exploration

From the MIS design framework, I explore in this section the design space according the four components defining a MIS. The mentioned variations can be combined to provide a large and flexible set of MIS-based interaction techniques. Those MIS instantiations are ideas of what MIS design space can offer in short since actual developed instantiations are presented in the next three chapters.

### On Geometric Definition

One specific shape could represent one specific range of possible actions. A plane may refer to a 2D control of a cursor, whereas a sphere, for example, may suggest a rotation control of virtual objects. As well, a particular orientation may refer to a particular action. Different dimensions could allow more or less accuracy.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



### On Input Reference Frame

Attaching a MIS to the world as a reference frame of input links it to the world. Even if the user moves, the MIS will not. If the MIS is associated and linked to the user, the latter can move around the environment keeping the MIS next to him at the same position regarding his position. The MIS could also be attached to a physical object. The MIS will remain attached to the object and then can be shared in a collaboration context.

### On Action Reference Frame

As explained in the section 3.1, Action Reference Frame links MIS to the display/system it controls. It can be associated to a static display or the ARF can also be associated to moving display. In this latter configuration, whatever the position of the display is, the MIS still controls it. The ARF may be re-affected in a multiple displays configuration.

### On Interaction Attributes

In this section, only very few "properties" of the MIS are addressed. These attributes may enable bimanual gestures, tuning of the sensitivity of the MIS, relative or absolute mapping, 3D or 2D input …

Using bimanual gestures can be done either using one MIS per hand or one MIS for both. Used bimanually, MIS accuracy might be improved if needed. The non-dominant hand (NDH) would provide a coarse control to specify a work area whereas the dominant hand (DH) would act accurately in the work area defined by the NDH. The bimanual property could also be used to allow mode switching. The NDH, when in the MIS would serve to select a, action mode for the DH.

Another property of interest, is the sensitivity of the MIS according to the user movements. As we saw in the Section 3.3 about the primary properties, the dimensions of the MIS can strongly influence accuracy depending on the mapping. But for comfort reasons, a MIS shall not be too large. Enabling the tuning of sensitivity of the MIS would provide more control. With the same MIS, several sensitivities may be available. If during the interaction user needs more



© 2017 Tous droits réservés. lilliad.univ-lille.fr

accuracy, he can lower the sensitivity and in return increase it, or create another MIS with a different sensitivity.

In the case where the MIS shape represents the entity we wish to interact with, the mapping can be either relative or absolute. For example, controlling a cursor thanks to a planar MIS, functions for relative or absolute mapping can be defined, inspired by functions involved in using a touchpad or a graphic tablet. This property should be consistent with the goal of the user.

The creation gesture can communicate the primary properties of the MIS. But it could also provide other properties. For example, a gesture ending by a cross might create a MIS that controls the cross section of a 3D model. A circle gesture will create a sphere shaped MIS. To create a volumetric MIS, a gesture representing a profile curve and ending with a loop could define a solid of revolution that would allow virtual 3D sculpture.

## 3.4 Examples of Interaction Scenarios

Considering the MIS design space described in the previous section, I present three possible scenarios varying for each of them both the number of users and distant displays which are currently being controlled by a user. Each scenarios illustrate different ways to use MISs and MIS gestures. Regarding multiplicity of the user and the screens "controlled", three general cases can be spotted. To these cases, multiple way of use of the MIS can be observed. In the following, I chose one way of use for each use case to give an example of the possibilities that gives the concept of MIS gestures. Those scenarios are not exhaustive. Here, I focus on the occasional use of one MIS per controlled screen.

### 1 User, 1 Screen

One user interacts with one display. When the user wants to interact with the distant screen, he defines a MIS with arbitrary primary properties, attached to the world, for example, and acting on the distant screen, with default controls corresponding to the system controlled. In that case, one scenario could be the distant use of a media center where the user from his sofa, wants to choose a movie to watch. He create the MIS and then can navigate through the media



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

center by controlling a cursor. The user's hand position in the MIS is associated with the mouse cursor.

#### 1 User, n Screens

Nowadays it is more and more usual to evolve in a space (either public or private) that gathers several displays. In such contexts, any use may interact with these screens. Then the attribution of which system the MIS controls must be done before any use. This affectation can then be made either by the user thanks to a gesture, or user body orientation. For example, at home, a user can control at the same time, his laptop and the TV, watching a movie on the TV and listening to the music on his laptop while cooking. If he wants to switch on the TV volume, he defines a MIS to control the laptop, turn off the volume, and then re-affects the MIS action frame to the TV to increase the volume.

#### n Users, 1 Screen

This case mainly refers to collaboration contexts. In the case of meetings where interactions between people are strong and participants study a single problem (focus on a single white board or screen), like brainstorming sessions, MIS can be seen as an interesting tool. For example, working on a new design of car, the designers may interact with the shared sketch. By defining a MIS attached to the designer who is speaking, each suggestion can then be attached to the right designer. At the end of the brainstorming meeting, the team can see easily each suggestions of each designer and then can choose which propositions are kept or not.

#### n Users, m Screens

In this configuration, several cases emerge. The case in which n=m, the one where n<m and finally where n>m.

**Case n=m** In this case, one can think that every user work on their own screen. We are then in the first configuration *(1 user, 1 screen)*.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



**Case n<m** Here, we have less users than we have screens. So they can all interact with a screen which is the *(1 user, 1 screen)* configuration. At some point, we could consider that a user might want to switch between the screens and fall in the *(1 user, n screens)* configuration since there are more screens than users.

**Case n>m** With more users than screens, we can have at the same time *(1 user, 1 screen)* and *(n users, 1 screen)* configurations.

In all of these possibilities, what stands out is that in the context of *(n users, m screens)*, we can have all the three previous configuration at the same time and more importantly, users can perform *transitions* between those configurations. We can imagine that in a collaboration context, multiple users with one screen each (*i.e.*a workstation) can then regroup around one screen. Then we fall in the situation of *(n users, 1 screen)* where we can imagine the same kind of interactions except that they can happen at the same time on different screens or from one screen to another. This possible transition between different screens and thus different collaboration contexts is greatly interesting. For example, one designer comes at another colleague's desk to discuss a feature. They both start using MIS around the coworker's screen and then have to go to the designer desk to visualize other data. One could imagine that at the coworker's desk, his MIS would have the upper hand and when moving to the designer's desk it is the designer's MIS that have the upper hand for control.

### 3.5 Discussion on Modularity and Reflective Nature of MIS

While working with MIS, one interesting point about MIS we came by is what we call modularity and reflection of MISs.

The modularity of MISs is inherent to it since an MIS is a sub-space we can arrange with others to form a more complex and global interface. An MIS-based interface is made either with one or several MISs. Each MISs having their specificities, when the interaction context does not fit the current interface, changing the MIS properties or replacing it with a more suited one is part of the concept. To reach this modularity and enrich MIS use, MISs should enable self-transformation. That is to say, users should be able to change the MIS properties thanks to interaction with the MIS itself. Which bring us to the reflective nature of MISs. MISs reflective nature means that by performing specific actions in a MIS or on its frontiers we can modify the properties of the MIS. This part of MIS is really interesting as it opens novel big questions and need



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



to be explored. The first set of questions is about MIS concept itself : how users could modify MIS? What does this imply for interaction, users, and acceptance of MIS interaction? The other interesting question is about the logical paradigm of interactive systems. The actual major paradigm (WIMP) might not be always best suited for MIS-based interaction. An other type of interfaces that is both compatible with the WIMP paradigm and better support MIS-based interaction. I will get back to these questions in the conclusion 7.1 of this document.

The point to remember here is that MIS interaction is reflective. An MIS can be modified by gesturing in the MIS itself.

## 3.6 Conclusion

In this chapter, I have detailed my proposition of the MIS concept and the design framework. I also gave a re-reading of the literature helped by the MIS design framework, some leads for the exploration of the design space and some use case scenarios. These last two sections are not to be exhaustive but illustrate how the design framework can be used.

Now that the MIS concept has been presented, the remaining of the thesis is structured with the goal of showing the use and the adaptability of the MIS concept to diverse contexts. To do so, in the next three parts, I will present my contributions. Those contributions instantiate the MIS concept. In those parts, I will also discuss more specific work that is related to the presented contribution. The first contribution studies the use (*i.e.*, creation and deletion) of dynamic MISs to interact with a remote display and presents a prototype for remote mouse cursor control on multiple screen. The second contribution proposes gesture control around the tablet in different contexts of use and different MISs. The third contribution presents an hybrid touch-midair interaction technique for wall displays allowing for continuous drag and drop in the space directly in front of the display. In each of this chapters, one or multiple MISs have been used in a different way. In particular, in the next chapter, the MIS is an input entity that is well distinct from the system it controls or acts on.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# Four

## MIS as a standalone interface

*"Interactive nonverbal behavior encompasses those acts which meet this last criterion: They are acts by one person in an interaction which clearly modify or influence the interactive behavior of the other person(s)."*

— Paul Ekman and Wallace Friesen, *The Repertoire of Nonverbal Behavior.*

### Contents





© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



This chapter proposes an instantiation of the MIS concept to help users in situations where distance between the user and the screen can not be avoided or limited. This first instantiation of the MIS concept I present here is one of the simplest but still the most representative: a dynamic invisible world-fixed MIS allowing to control a mouse cursor on remote displays. This chapter first presents the second part of the published paper [91] which includes a user study that provides some elements of knowledge about how users, in a participative design approach, would potentially use such systems to control a mouse cursor on different displays and then present a corresponding prototype.

In our results, we show that users validate the idea of a planar MIS, and that most users that run the experiment instinctively state that plane position is user-defined and dynamic. Interestingly, we show that users have a good spatial perception of created MIS and that most of user-defined creation gestures may be used also for calculating position of plane (including orientation), as well as interaction frontiers. We also show that users easily integrate mental representation of interaction MIS, since user-defined deletion gestures take plane location into account. Finally, we provide guidelines for MIS gestures in mid-air interaction techniques. We also describe the design space associated to the presented concept, and describe the proof of concept of MIS interaction that illustrates two key scenarios.

## 4.1 Creating and Deleting a MIS for Remote Display Interaction

To our knowledge, there is no existing studies on how users may create or delete interaction sub-spaces. However, some work on multitouch interaction can give some hints.

In [120] several participants conceived imaginary areas around the screen with particular properties, as clipboard or a trash can. Similarly, some of them also imagine invisible widgets and reused them. The mental representation of invisible interfaces is not unnatural or too much exotic to users. In this same study, participants mostly preferred one-hand gestures as in [82] for the efficiency/simplicity and energy saving.

In [72] the authors also conducted two studies. The first is related to what kind of click gesture will be more appropriate. Results showed that, considering the average miss-clicks, the tapping gesture is the worst, the left hand click is the more tiring and a specific gesture, which is stretching the thumb away from



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



the index, has the highest satisfaction rate. Interestingly, in [17], the air tap is the preferred gesture to click in mid-air. The second study investigates the more appropriate size of panel to avoid miss click and satisfy user's comfort. The 24" panel was the more appropriate size. Concerning the size, in [63], Kattinakere et al. study and model a steering law for 3D gestures in above-the-surface layers, resting the hand on the surface. Results suggest that a layer should be at least 2 cm thick and that steering along more than 35 cm generates more errors. Considering that the hand is resting on the surface, the minimal thickness should be more that 2 cm for complete mid-air interaction. A similar work [107] studies the geometric properties of layers above a a tabletop for tangible displays interaction as well as interaction parameters such as vertical or horizontal gestures for search. Concerning the thickness of a layer, results show that for a holding tasks (*i.e.*, keeping the hand still), a minimum 1cm is required. For vertical gestures, 1cm is also the minimum thickness required to reduce the errors whereas a minimum of 4cm for horizontal search (*i.e.*, similar to steering) is needed.

Concerning the visual feedback, as said in the introduction of MIS concept, a spatial acquisition study in [21] shows that large movements in a 2D plane are rapid and accurate while raycasting is rapid and not accurate, and 3D gestures are slow but expressive as more DOF are available.

#### Methodology for Eliciting Gestures

We chose to carry out a gesture elicitation study, as in several prior work, in order to see how potential users could use the MISs, and what they could expect.

The methodology proposed by Nielsen et al. in [84] consists in identifying "the functions that will be evoked" by the gestures, which are in our work a creation, click and deletion functions, then, finding "the most appropriate gesture for each of those functions" by an analysis phase of the gestures performed by the users. In [120], Wobbrock et al. conducted a similar study in the context of gesture-based surface computing. They identified 27 common commands and the participants had to choose a gesture for each of these. In [82], which is the follow up of [120], the authors concluded "that participatory design methodologies [...] should be applied to gesture design".

Our user study was designed to collect gesture data that could be used to define MISs and questions the users on what they could expect of MIS inter-



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



action. In order to perform such study without suggesting any solution, we decided to simulate distant screens using a large curved screen (5.96 meters by 2.43 meters).Using such environment, we are able to project images of displays at different locations and of different sizes. By doing so, we expected to represent daily scenarios in an abstract way such as using a computer screen or a television at home or in collaborative working sessions.... The remaining of the section describes our experimental protocol and how it relates to our concept of MIS interaction. With their agreement, all participant sessions have been videotaped using a Kinect camera in front of the user and a video camera on the side recording a different point of view and sound as shown by Figure 4.1.

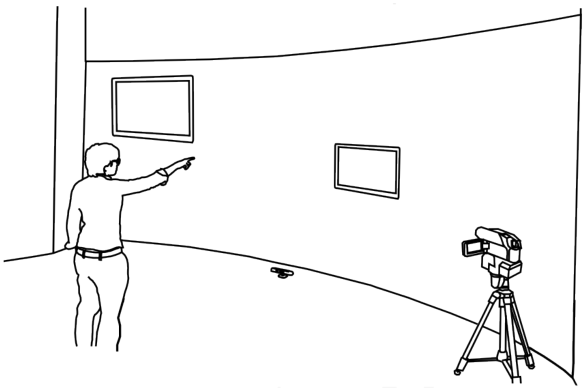


Figure 4.1: Our apparatus for the study: participants are in front of a large scale display to simulate screens of different sizes at different locations. The sessions are videotaped using a camera on the right side and a Kinect in front of the user.

#### Protocol

The experiment was composed of four phases:

- **Phase 1** was a questionnaire based on Likert scales (1 for never, 7 for everyday) to retrieve users habits and expertise about computer, tactile devices and video gaming frequencies of use.
- **Phase 2** was a brainstorming phase. Participants were asked to tell by what means they would control a distant display. After that, we explained quickly the concept of MIS interaction as an interaction area to control



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



displays at a distance without suggesting any dimension or predefined shape. Then they were asked different gestures to define such interactive area.

- In **Phase 3**, participants had to define 90 areas corresponding to projected virtual screens of two different sizes : 32 inches and 55 inches. Each virtual screen were displayed 3 times at 15 different positions on the large screen. They could take a break every 10 trials to avoid fatigue. For each trial, participants had to define, by a gesture or a posture, an area they thought was the most relevant and comfortable to control the shown virtual screen. Then they had to touch it as if they were interacting with the virtual screen. They were told the virtual screen could be either a computer screen or a television. The only constraint was that they were not allowed to walk but they could turn around. After the repetitive trials, they were asked to tell which gesture they preferred during the experiment. Then they had to imagine a gesture they would perform to delete an area they have previously defined.
- **Phase 4** was a feedback questionnaire. In a first part, users were questioned using Likert scales (1 for "I totally agree" to 7 for "I completely disagree") about the concept of MISs.

#### Participants

18 participants volunteered for the study (4 female). 8 participants worked in HCI. They were between the ages of 22 and 43 (mean: 27.6). Two participants were left-handed and one was ambidextrous. All participants used a PC and 39 % of them used tactile devices almost everyday (mostly smartphones).However, only 28 % of the participants played video games regularly. Even if they were not gamers, all of them had already tried and knew 3D gestures using the Wiimote, the Kinect, the Eyetoy or the PS Move.

#### Gesture Classification

In order to analyze participants gestures, we looked for existing classifications of gestures in the literature. Several have been proposed and are described in the following. We first describe the famous classification of Ekman and Friesen of nonverbal behaviors, then focus on proposed classifications for the field of interaction.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

### NONVERBAL BEHAVIORS GESTURE CLASSIFICATION

In [27], Ekman and Friesen describe a categorical scheme to analyze and study nonverbal behaviors. From past work, they propose three fundamental aspects of nonverbal behaviors which are *Usage*, *Coding*, and *Origin*. Those three considerations about behaviors allow to formally describe the process that leads to the nonverbal behavior. In the following, the word *Act* is defined by the authors as “movement of the hands and arms, legs and feet, shoulders, or total posture".

In the rest of the paper, they propose a classification of five categories of nonverbal behavior:

- **Emblems** are acts that can be replaced by one or two words. They have a verbal translation or definition. They are usually common to a certain cultural group.
- **Illustrators** are movements that support what is said verbally. They illustrate the speech. Ekman and Friesen decompose illustrators in six types:
  - Batons: emphasize a particular word or phrase
  - Ideographs: sketch of a path or direction of thoughts
  - Deictic: pointing movement
  - Spatial: identify a space or a spatial relationship
  - Kinetographs: depict a physical action
  - Pictographs: draw a picture
- **Affect Displays** communicate feelings and emotions
- **Regulators** are acts which control the conversational flow between two interactants
- **Adaptors** are behavioral adoptions to satisfy bodily needs, manage emotions or communicate (intentionally or not) a message to other

Though this classification may be well-suited for psychological studies and gestures oriented, it is too large to classify interaction gestures as interactive gestures are little investigated. Indeed, many of the described types are often unaware gestures. Whereas interactive gestures are intended to the other interactant. Which leads us to more specific classifications for Human-Computer Interaction



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

#### Gesture classification for Human-Computer interaction

Cadoz [16] suggested a classification regarding the function of the gestures which are complementary and dependent : semiotics (communication), ergotic (action) and epistemic (perception). This classification is proposed in the context of the study of instrumental gestures – which are semiotic gestures. Though this classification is more precise than previous one, classifying gestures from an elicitation study with this classification might be problematic as almost all gestures will be semiotic or ergotic. Then, we won't be able to formally classify the participants gestures.

Karam and Schraefel proposed a classification adapted to HCI based on gesture styles : deictic, manipulative, semaphoric(with a cultural meaning like thumb up for OK), gesticulation(conversational gesture), sign language, multiple(combined) gestures styles. Aigner et al. presented in [1] a modified taxonomy of Karam and schraefel [62] adapted to gesture elicitation study in mid-air without speech command or sign language. Thus, guaranteeing precise analysis.

To classify the different gestures performed by the participants, we used the gesture taxonomy proposed by the Aigner et al. [1] and depicted in Figure 4.2. This taxonomy proposes four different classes of gestures: pointing, semaphoric, pantomimic and iconic.

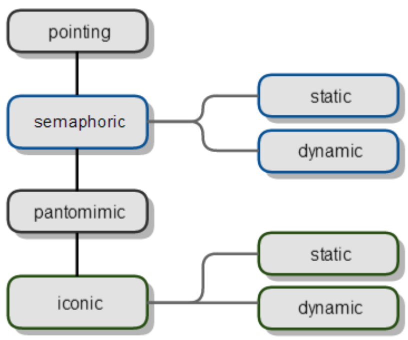



Figure 4.2: Classification used to analyse the gestures made in the user study.

While *pointing* gestures are mostly used to name an object or a direction, *semaphoric* gestures are gestures that are meaningful. There are *static* semaphoric gestures like the thumb-up posture that means "OK", and *dynamic* semaphoric gesture like waving the index finger sidewards to mean "no". Note that these meanings are strongly dependent of the cultural background and experience



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



of the user. *Pantomimic* gestures refer to gestures used to mimic an action like grabbing an imaginary object and rotating it. Finally *iconic* gestures represent informative gestures. They inform about the properties of an object like specifying a size or a shape. There are static iconic gestures and dynamic gestures. Unlike semaphoric gestures, no common knowledge of the user's past experience is needed to understand these kind of gestures.

### Results

This section presents the results and observations of our study. We decouple our analysis into three parts related to the *MIS interaction* basic steps which are: the gestures to create it, how users can interact with it and finally how participants propose to delete it.

#### Preliminary questionnaire

We started the brainstorming by asking to the participants "By what means would you control a distant display?" if they could not reach the display physically. From the 18 participants, 50% answered they would use hand gestures or use classical input devices such as a mouse (5 participants) or a remote control(4 participants). 4 users also refer they would use voice commands to control the display content.

#### Interaction space creation gesture

We analyzed the video of each participant and described each gesture performed along the 90 trials of the experiment using the gesture taxonomy presented by Figure 4.2 and complemented with the information about which hands were used, hand postures and the relationship between the location of the gesture and the user field of view or any significant body part. We choose to discard any isolated gesture performed or slightly different variants from the same gesture.

Looking to the set of the 33 gestures performed by all users, 71 % of them describes an area that can be assimilated to a plane. We noticed that 89 % of users performed iconic dynamic gestures, representing 60 % of all the gestures. They mostly represent rectangular shapes (66 %) or opening gesture (28 %) along



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

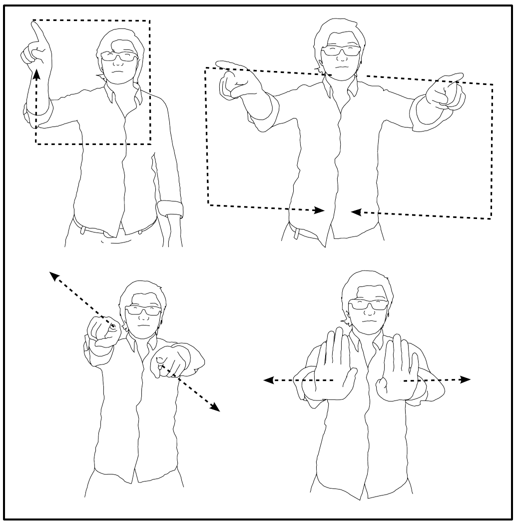


Figure 4.3: Frequent creation gestures proposed by the user: defining a rectangular area using one or both hands (top) and using an opening gesture in its field of view with diagonal or horizontal symmetric gesture (bottom).

a line or diagonal delimiting the size of a frame as depicted by Figure 4.3. Circular motions such as circles and waving in front or around the user were less common (9 %).

Regarding hand usage, we noticed that 33 % of them exclusively defined gestures using one hand, 33 % using both hands and 33 % mixing both approaches while performing the several trials. While all unimanual gestures were mainly done using the dominant hand, most of bimanual gestures described symmetrical movements or poses. Only three users presented gestures following the asymmetric bimanual Guiard model [43]. While performing the gestures, we noticed that most of participants used a reduced set of hand poses shown in Figure 4.4. Index finger pointing to the screen, and mimic of a pencil were prominent among participants (77 %) compared to both L shape (27 %) and open flat hand postures (33 %).

About display position influence, we noticed that most of the participant aligned their field of view prior to start the gesture by rotating both the head and body. However, 39 % of the users depicted gestures in a fixed position regarding their body. The preferred approach (61 % of users) was to create vertical planes aligned with the field of view or the projected screen by drawing rectangles or defining static frames. In the case of horizontal or oblique planes independently



© 2017 Tous droits réservés. lilliad.univ-lille.fr

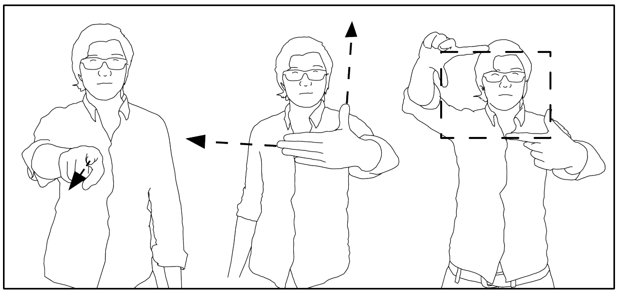


Figure 4.4: The 3 main hand postures. From left to right: pointing to a given direction, flat hand posture defining a spatial reference, two L hand postures delimiting an area.

of the screen position or field of view user was never looking at his hands while performing the gesture.

#### Interacting on a MIS

For each trial, we asked the participants to touch or interact on the previously defined interaction area. They mainly simulated drawing or small push actions close to the area defined as shows Figure 4.5. Users touched the imaginary space using their dominant hand, except one with both hands. We noticed three different major hand poses: pointing using the index finger, pointing using a flat hand and pushing using an open hand with a percentage of 56, 22 and 17 respectively. People using an open or a flat posture tend to push, grab or swipe close to the MIS definition. While participants using their index finger tried to mimic drawing short scribbles or push small imaginary buttons. These behaviors showed a strong materialization of the MIS as a physical tool.

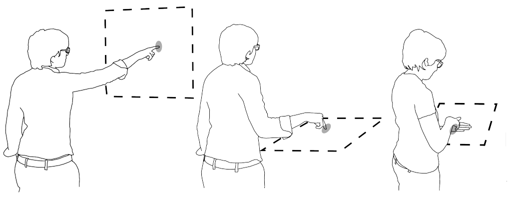


Figure 4.5: Common touch gestures proposed by the subjects: pointing on a vertical or horizontal imaginary area and touching the non dominant hand as a reference.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

### Deleting a MIS

At the end of experiment, we asked participants to propose a delete gesture considering that their interaction zone creation was persistent. Looking to the 23 gestures collected, we noticed a strong usage of pantomimic gestures since most of users materialized the interaction MIS. 23 % of the proposals do not fit in this classification such as leaving the interactive area, waiting for it to disappear, drawing a cross or using the inverse of creation movement. For users that used non dominant hand as a support to interact, the area shall disappear just by removing the hand. Figure 4.6 illustrates the main proposed gestures.

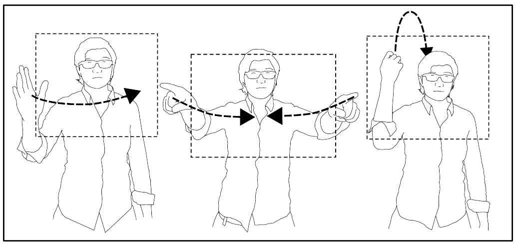


Figure 4.6: Participants delete gesture proposals: pushing the area with one hand, closing the MIS using both hand or throwing it away to a given location.

### Observations

From the current user study, we can highlight the following observations and remarks to implement MIS based applications and better take advantage of the design space offered by such concept.

**Make MIS planar, and dynamic :** most of users spontaneously create planar MISs, and take for granted that they can specify them in arbitrary position, without any experience.

**User tends to turn in the direction of the screen :** in that case, MIS tends to be vertical, and directly relates to the field of view of user. In case where users do not orientate themselves in the direction of the screen, MIS is created horizontally, for indirect interaction.

**Gesture for creating and deleting MISs can be parameterized gestures**: for most users, these gestures specify both a *command* (e.g create subspace)



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



and some *parameters* of the command (e.g some geometric features such as MIS location for creation), in the same gesture.

**User has proper mental perception of MISs he/she creates** Since all users provided delete gestures that start in a location devoted to the MIS that was previously created. The MIS became real.

## 4.2 Proof of Concept

Following the observations resulting from our user study, we devised an application as a proof of concept to let one or more users interact with one or more distant displays.

Several key scenarios were possible to implement regarding both the number of users and the number of screens. The one user interacting with one screen scenario, the one user with multiple screens scenario and the multiple users with one screen scenario. For the proof of concept, the application consisted of providing to two users the capacity to control and share the mouse cursor between several displays allowing to interact with any content displayed by the screens. We chose to implement a planar MIS solution defined by rectangular gestures since such gestures were the most common among our user study. The application was implemented as a daemon sending mouse inputs directly to the operating system (Microsoft Windows 7).

To track the user's gestures, we chose to rely on a wireless magnetic based tracking system i.e. Liberty LATUS system from Polhemus complemented with a button to emulate the mouse click as depicted in Figure 4.7. At first, such solution was preferred to non intrusive tracking solutions such as the Microsoft Kinect depth sensor, in order to obtain reliable positions and orientations of the user's hand. However, our MIS concept could be used in a more pervasive environment using several cameras to track users in a non-intrusive way. And later, we used a Kinect-based prototype in order to have a less intrusive and a more movable setup. All input data were streamed to our software daemon using a TUIO client approach.

The details of the implementation are discussed in the following chronologically from creation gesture to deletion gesture.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



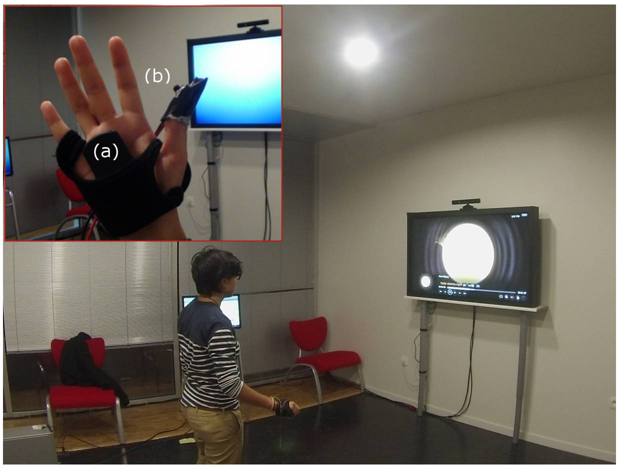



Figure 4.7: The user is only equipped with (a) a tracker and (b) a wired button.

### Application

The detection of a MIS creation gesture is made through 3 steps analyzing the user's hand motion. **First**, both the beginning and the end of a gesture are triggered based on threshold values over the hand acceleration. All the positions and orientations retrieved in between these two events are recorded tracking user gestures. **The second step** is the computation of the plane thanks to PCA. We then define the origin, the normal and construct the reference frame of the plane from the average of the orientation vectors of the user's hand during the gesture to get the "up direction" (i.e y-axis) and the "right direction" (i.e x-axis) as depicted by Figure 4.8. The dimensions are computed by projecting the gesture points on the newly defined plane and computing the aligned bounding box on its reference frame. **Finally** to detect rectangular shape creation gesture, we use the 1$ recognizer on the 2D path corresponding to the projection of the 3D hand positions on the pre-computed plane. A pop-up on the screen informs the user the MIS is created if the gesture is rectangular.

Once the MIS is created, each 3D position received is then treated regarding the MIS. When the hand is near enough from the MIS, we allow the user to control the mouse cursor with his hand. The mapping between the hand position in the MIS and the mouse cursor position on the screen is absolute. While the accuracy of mid-air gestures cannot match the one achieved by a mouse device, preliminary results show that such approach is usable. Moreover, interacting with a distant display must consider the eye resolution of the user according to



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



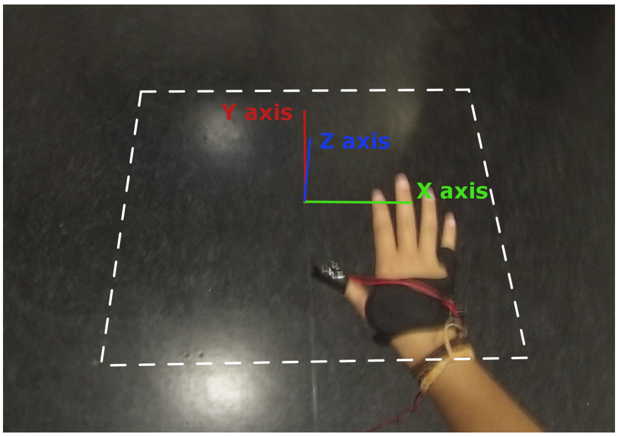



Figure 4.8: The frame of reference of a MIS

the distance. As we are at a distance from the screen, we do not need has much precision as if we were in a desktop context.

Finally, our application allows multiple users to define and use their own MISs as well as attaching them to a given display. The current implementation is aware of screen locations in the room and use directional swipe gestures to the screen in order to attach a MIS to an existing distant display as described in our user scenarios. Currently this proof of concept was defined to track two users max and interact with two screens. When the MIS is created by a user, it is automatically attached to the closer screen regarding the user's position. The directional swipe gesture allows to change such default binding. This directional swipe gesture and this reassignment of what screen we control is a very simple and naive expression of the reflective nature of MIS I talked about in section 3.5.

To delete such space, we choose to detect horizontal swipe gestures starting within the MIS and finishing out of it with a given velocity and along the x-axis of the plane.

### 4.3 MIS DESCRIPTION

This application of the MIS concept can be describe by the design framework like the previous work. Here, the planar MISs are defined in position, size and orientation by the user. And the available input are a rectangle stroke made by the hand to create a MIS, A fast swipe to the left in the MIS to delete it, a pointing



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



gesture from the MIS toward a screen to switch the controlled screen, and the hand position in the MIS controls the mouse cursor. The complete description is presented in Table 4.1.

| Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|
| **Shape :** Plane<br>**Dimension :** user defined<br>**Position :** user defined<br>**Orientation :** user defined<br>**Material :** Invisible | World Fixed | Two screens. One at a time | **Input :** Rectangle stroke ; Pointing gesture ; Fast left swipe ; Hand position in MIS<br>**Output :** Creation ; Switch screen ; Deletion ; Mouse cursor displacement<br>**Mapping :** Continuous and Absolute |

Table 4.1: MIS description of the application

## 4.4 Conclusion

We presented elements of knowledge about midair interaction with distant displays. We introduced the concept of *MIS gestures*, that we think is a flexible approach to midair interaction within pervasive environments. We showed that MIS gestures are, to the highest acceptability, planar and dynamic when interacting with screens. Though it has some limitations in the current state of the prototype, the application developed allows to see few interesting possibilities among all of possible MIS-based interaction techniques. Here, the MIS is an entity apart of the controlled system (*i.e* remote screen). In the next chapter, I will present an application of MIS in combination with a tablet. where the tablet is in the MIS, it is a part of it. From this, two uses of the MIS are presented. In the first one, I will describe the use of midair gestures for MRI visualization,



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



the use of fingers identification and touch for menu selection. All of this will be made by cutting out the MIS in several MIS around the tablet. The second use of the MIS that contains the tablet is an interaction technique that where the MIS is a continuum of the tablet screen.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

# Five

# Attaching the Tablet and MISs

*"Everything you can imagine is real."*

— Pablo Picasso

## Contents





© 2017 Tous droits réservés.

lilliad.univ-lille.fr

In the previous chapter, I presented a prototype and a study of an invisible world-attached MIS instantiation. In this chapter I will present and discuss the use of MIS that surrounds a MIS. Having more space to interact with a tablet offers the use of midair gestures with the hands and the fingers. With this work I intend to take advantage from both touch and midair interaction by relying on their own strengths. Tablets allow for direct interaction with the displayed objects. Using the space around with the same directness of tablets could benefit the user experience on such devices. Interacting around mobile devices have been explored in the literature as we saw in the Related Work section 3.2 like in [50,53,59]. In particular, in [59] the authors argue that using free-space around the device makes the control more fluid and avoid screen occlusion. The best technique is the one that uses the plane defined by the back of the phone to pan and the z-axis of it to zoom. Users have a button to clutch. The study shows that for pan-and-zoom task, the right side space of a smartphone is preferred by users (right-handed). Though this work has been applied to smartwatches and smartphones, they show the interest and advantages one could benefit from using the space around devices. The MIS concept gives us a mean to design such interactive space in.

I explored these new possibilities in two ways. First by cutting out the space into sub spaces and using it in the context of MRI visualization published in [90] and a drawing application. Second by using the surrounding space as a continuum in Section 5.2.

## 5.1 Tabslab: Workspaces Around the Tablet

The core concept of TABSLAB is to broaden the interaction with the tablet by using the space around it. We cut out this space into eight areas denominated as slabs (see Figure 5.1). These eight slabs can then be used as spaces for gestural interaction close to the device. Those slabs are attached to the tablet and act on it as well as potential remote systems. The slabs are rectangular with or without thickness according to what kind of interaction we want to perform. 2D AND 3D hand gestures are then possible as well as postures. Since the tablet is held by the NDH, performing gestures with the DH is consistent with Guiard's kinematic chain [44] where the NDH defines a frame of reference for the DH to act within.

Since the user is not required to use the tablet to interact with the displayed data, it opens new possible options to use these new inputs and control image attributes. For example, the right slab can be devoted to manipulating the



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



detailed view on the tablet. The top slab could be used to interact with the distant overview. And the top right slab could be used as a view filter selector.

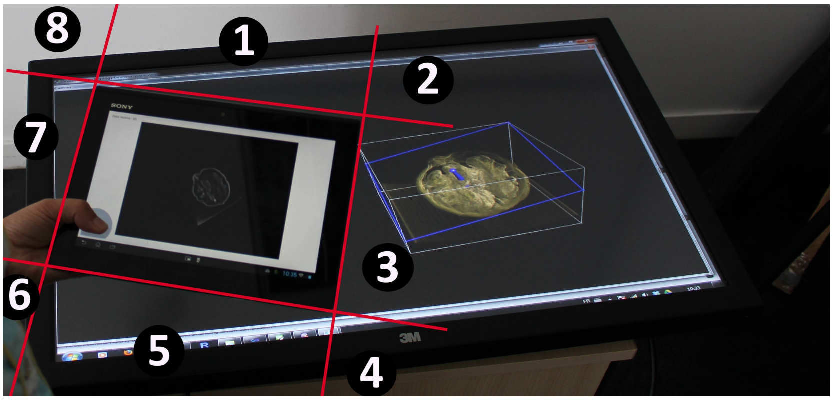



Figure 5.1: Slabs arrangement around the tablet and MRI prototype setup

In the following two sections, I present two applications of TABSLAB. One for MRI visualization via 3D gestures and the use of two slabs. The other application is a drawing application relying on four slabs and finger identification for tools selection.

### MRI Visualization

Medical imaging is essential to support most diagnosis. It often requires visualizing individual 2D slices from 3D volumetric datasets and switching between both representations. Combining an overview with a detailed view of the data [22] enables to keep the user in context when looking in detail at a slice. Given both their mobility and their adequacy to support direct manipulation, tablets are attractive devices to ease imaging analysis tasks. In [89], the authors investigate how different configurations of input and output across displays affect performance, subjective workload and preferences in map, text and photo search tasks.Their results show that a mobile device-controlled large display configuration performs at least equal or best than distributed information on several displays. They have been successfully combined with tabletops [105], allowing new ways to explore volumetric data. However, while touch allows for a more direct manipulation, it suffers from the well-known fat finger problem which can interfere with the display, making it hard to understand subtle visual changes. To overcome this problem, we propose to explore the space around



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



tablet devices. Such approach has been used for displays [56] to separate several workspaces of the desktop.

Here, we use around tablet spaces to *invoke commands* that are not required to be performed on the tablet. The user sees the whole volume on a distant display and the tablet displays the slice. I developed a simple prototype (see Figure 5.1) that uses the right slab of the tablet to allow the arbitrary slicing of a volume (space number 3 on Figure 5.1). Here, the hand position determines the slice position along the orientation axis given by the tablet orientation (tablet orientation gives the slice orientation). In the top slab (space number 1 on Figure 5.1), users can tune the contrast of the 2D slice displayed on the tablet. The horizontal hand position gives the level of contrast (like a slider). Once the user is on an the slice of interest, and position can be locked by pressing with the thumb of the handling hand on the side of the tablet screen (see the bottom left of the screen on Figure 5.1). Once, locked, tablet orientation and hand position do not have influence anymore. Then, users can annotate the slice if needed by touch. The plane of slicing is displayed and updated simultaneously on the remote display so that users can have a global feedback of what they are watching on the tablet.

#### MIS description

Following the description of the previous work according to the MIS framework, this presented implementation can also be described in the same way. Here, I use two MISs around the tablet. A cube attached to the right side of the tablet, and a plane attached to the top of it. In both of them, it is the hand position that is used as input but in different ways. In the cube, it is the hand position along the tablet normal in the cube while in the plane, I use the hand position according the tablet top side. It controls the slice position and the contrast value, respectively. You can see the detailed description in Table 5.1 below.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

| Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|
| **Shape :** Cube ; Plane **Dimension :** Tablet right side ; Tablet top side **Position :** Tablet right side ; Tablet top side **Orientation :** Tablet orientation ; Same **Material :** Invisible ; Same | Tablet Fixed ; Same | Tablet view and Context view on screen ; Tablet view | **Input :** Cube : Hand position with report to tablet normal ; Plane : Hand position with report to tablet top side **Output :** Cube : Slice position along the slicing normal ; Plane : Contrast value **Mapping :** Continuous and Absolute ; Same |

Table 5.1: MIS Description of the prototype for MRI visualization

This work, published in the proceedings of SUI'14 [90], was motivated by leveraging on human learned manipulation gestures of physical objects (equivalent of "adaptors" in the framework à Ekman and Friesen) to ease visualization of virtual 3D models. However, MRI visualization is not a typical/common tablet application, I wanted to investigate TABSLAB compatibility with more standard interaction context.

### Tool selection for drawing application

In this section, I will present a preliminary work that uses TABSLAB as tool selection interface for a drawing application on a tablet. This work takes its inspiration from finger identification papers like [38,109] where modes, tools, commands or even data are assigned to specific fingers (*e.g.*, copy and paste respectively assigned to ring and middle fingers). In [38], the authors argue that interaction on large multitouch takes little advantage of what shortcuts like



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



WIMP interfaces do and propose fingers chords as equivalents. The NDH acts as a menu storage, where a certain chord calls an associated menu and the DH selects the tool of interest by beginning the action with the finger associated with the wanted tool/command/data (seef Figure 5.2 for an example of use).

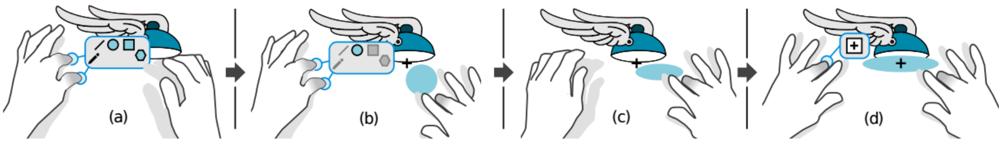



Figure 5.2: Illustration of chords from [38]. (a) a chord displays the commands (b) middle finger selects ellipse tool (c) the DH draws the ellipse (d) an additional chord constrains the ellipse.

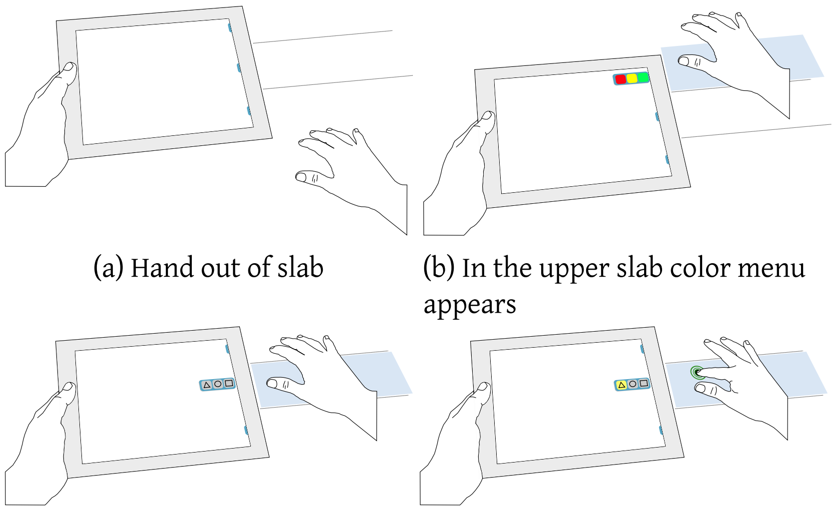
In what we called DrawingSlab, the different drawing tools are selected through hovering the one of the four slabs right to the tablet and selecting the tool with the corresponding finger. A command is a defined combination of a slab and a finger midair tap (see Figure 5.3). We developed a basic drawing application allowing different forms, colors, line styles and thicknesses. For each of these tools, four values were proposed. In order to ease learning, when an hand was detected in a slab, after a one second threshold the associated menu is displayed to the user in the order of fingers starting from the thumb for the left item. This way, no displayed menu is required when expert users draw with the application. For example, in the middle slab assigned to forms, line, rectangle, ellipse and triangle were available. Following this example, in order to select the ellipse tool, users had to place their hand in the middle slab and tap with the middle finger. Once the tool or parameter selected, users just have to draw on the tablet. We tracked the fingers and the tablet positions with a Leap Motion. The tap was detected with the Leap Motion API.

### Discussion About Tabslab

In the previous sections, we saw two different uses of TABSLABS. MRI uses 3D midair gestures whereas TABSLAB borrow from WIMP interfaces and is only based on discrete actions. We could easily imagine a combination of both those discrete and continuous use of TABSLAB applications. For example, a color picker in DrawingTabslab, a slider for line thickness. Or even, applications where users perform 3D gestures and invoke specific commands via discrete gestures. With those two prototypes, arise two major questions :



© 2017 Tous droits réservés. lilliad.univ-lille.fr



(a) Hand out of slab

(b) In the upper slab color menu appears

(c) Mid slab is for from selection

(d) Index tap selects the triangle form as it is the first item of the menu

Figure 5.3: Here is an example of use of the drawing application.

- For manipulative 3D midair gestures, is the precision of manipulation is enough? What kind of one-handed gestures is better suited in slabs to explore and manipulates 3D models? Could haptic feedback improve precision?
- For WIMP-like interfaces, what kind of menu or interface paradigm is better suited for such invisible midair interfaces? Or more specifically, are WIMP interfaces suited for TABSLAB?

But those questions are part of a bigger question about hybrid touch-midair interaction. Indeed, the specificity of TABSLAB is its inherent nature of hybrid touch-midair interaction since it is to be used around a mobile touch device. Hence, in order to develop interesting applications using TABSLAB to content users, it is crucial to well understand in what situations and contexts touch interaction could benefit from midair interaction and vice-versa. This issue should not be approached by separating touch studies and midair studies. But it is the combination of the two that has to be studied and adapted. Because we have to take into account the device itself held in one hand or not, as well as the potential cognitive processes that could emerge from the union of the two.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



In this section, I presented a proposition of that binds the tablet several MISs in order to expand interactive possibilities. In the next section, I propose to use the tablet and the MIS attached to it as a continuum.

### 5.2 Ghost Tap: an invisible input extension

In this section, I present a new interaction technique called Ghost Tap using a MIS that extends the tablet screen. The Ghost Tap concept is using the surrounding space of a tablet as an input area that extends the tablet screen. Users can interact with the invisible content around the tablet as if it was displayed on the tablet. No visual feedback of where users are performing is given. For example, when visualizing a map, users can tap right to the tablet to pan the map to the location of the tap. It allows to interact with the content beyond the tablet's viewport. As the interaction relies on blind taps, we wanted to explore how well we would perform in this context.

There have been numerous investigations into understanding and evaluating the performance of off-screen interaction. Research has provided more understanding into the off-screen midair pointing in static visualization conditions. For instance, Hasan et al. [54] exhibit the advantage of off-screen pointing in terms of navigation time when comparing off-screen pointing with Peephole and standard Flick & Pinch for map navigation with or without the help of an overview. Gustafson et al. [46] measured the performance of "coordinate-based imaginary pointing" when using the non-dominant hand as a reference. Their findings showed that the performance deteriorated significantly when increasing distance between the target and the reference hand. Ens et al. [28] investigated off-screen pointing performance for Fitt's task with dynamic and continuous feedback. Two novel pointing techniques were proposed namely,*launche-adjust* and *glide*. In both of these techniques, a continuous feedback on the target position and the finger position is provided by using either the *wedge* technique [45] or an overview of the scene. Their findings indicated that further targets should be larger to minimize the error. Another interesting result is that the direction has a significant effect on the movement time. In particular, diagonal direction increases the movement time. and that users tend to undershoot the target. Markussen et al. [78] investigated off-screen midair horizontal pointing performance on a distant display. They showed that the perceived space around the screen is modified depending on how far is the intended target position from the screen. Users tend to guess the intended position closer than it is from the screen. Hasan et al. [53] found that for midair target selection a



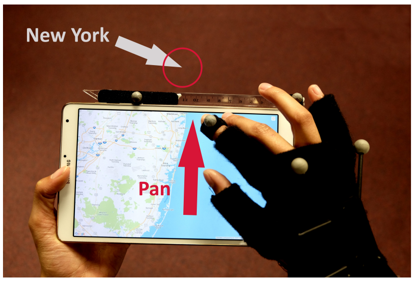
© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



(a) Initial state




(b) After panning, New York is in the upper area of the tablet


(c) A Ghost Tap is performed a bit north to New York

(d) After the Ghost Tap, New York is visible again on the screen

Figure 5.4: Example of Ghost Tap use for map navigation.

*DownUp* gesture decreases the error rate but increases the trial time. In contrast, *LiftOff* gesture provides the best compromise between error rate and trial time. Their findings also indicated that the radial bins should be bigger than 4 cm wide and the input range should not extend beyond 40 cm around the device. In this work, we are interested in understanding the effect of how the target is moved to its position affects the target selection performance.

### Demonstration application

We applied Ghost Tap to Google Maps navigation. We implemented a NodeJS server that provides a Google Maps page (thanks to the Maps API) where tapping around the tablet brings the tapped position at the center of the screen (see Figure 5.4). This demo leverages on our geographical knowledge and our sense of direction.

During navigation, the seen data is moved out of the screen either by panning or zooming gestures. After performing those gestures, estimating where



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



the last seen data is done based on what we remember from our last gestures. Additionally, from what we know of the displayed data, we can also predict its not displayed surroundings. For example, while navigating in the New York area, we know that Philadelphia is South-West. Depending on the scaling we might also estimate the distance to New York. From this observation, we can identify three cases of position estimation: after a pas gesture, a zoom, or based on what we see on the screen (and recreating the overview of the arrangement in our mind).

#### Experiment: Target selection

We conducted an experiment to compare the impact of different visualization techniques when selecting an off-screen target. The goal of this experiment was to evaluate how the visualization technique affects the minimum target size and position users can successfully acquire.

#### Apparatus and Participants

The experiment ran on a 8.4-inch Samsung Galaxy Tab S held horizontally in the left hand. To track the tablet, right hand and right index position, we used the OptiTrack system. One rigid body was attached to a glove we equipped the participants with. Another rigid body was attached to the tablet. For the index, we just attached one marker to the tip of it.

We implemented a simple tap recognizer that analyzed the Optitrack data. When a tap was performed, it was send to a NodeJS server with its coordinates in the reference frame of the tablet. The server provided the experiment web page and after each trial compared the tap coordinate to the target coordinate and logged the measures. The implemented tap builds on the results discussed in [53]. It is a fast down and up gesture.

Ten participants (one female and nine males) volunteered to take part into our experiment. The range of participants age was between 24 and 32 years (mean=27.6, s.d=2.67). All of them were right-handed. So they held the tablet in the left hand while performing the taps with the right hand.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



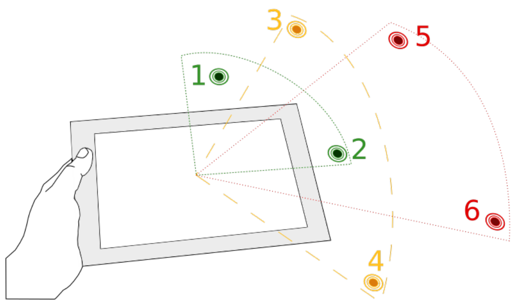



Figure 5.5: Targets positions where green targets are the near targets, the orange are the middle and the red ones are far ones.

#### Design

Dependent measures are analyzed using a $3 \times 3 \times 6$ repeated measures within-subjects analysis of variance for the factors: *visualization* technique (*overview*, *pan* and *zoom*, where *visualization* techniques relate to the most common gestures/actions used in mobile device interaction when acquiring a target as it will be described later), target *size* (*small*: 3 cm, *medium*: 6 cm and *large*: 10 cm, where target *size* corresponds to the target diameter) and target *position* (1-6, with 1: 11 cm; 2: 15 cm, 3: 20 cm; 4: 22 cm, 5: 25 cm and 6: 28 cm) (See Figure 5.5).

#### Task & Procedure

The task required participant to tap on an off-screen target as quickly and accurately as possible. Targets were displayed randomly one at a time within the ghost space. Each trial began after the previous target was selected and ended with the selection of the current target. In the *overview* condition, a gray rectangle representing the tablet surrounded by the ghost space that contains the actual target are displayed on the screen (see Figure 5.6a). To start the trial, participant had to press the "start" button. In the *pan* condition, first the target with its actual size (4 cm; 6 cm or 10 cm) and the background are displayed on the screen. Then participant had to press the "start" button. Both the target and background are then moved until the target reaches its corresponding position in the ghost space (see Figure 5.6b). In the *zoom* condition, first the target with a diameter of 2 cm and the background are displayed on the screen. Then participant had to press the "start" button. The target is then simultaneously scaled and translated until it gets its corresponding size (4 cm; 6 cm or 10 cm) and position in the ghost space (see Figure 5.6c). It is important to note that the

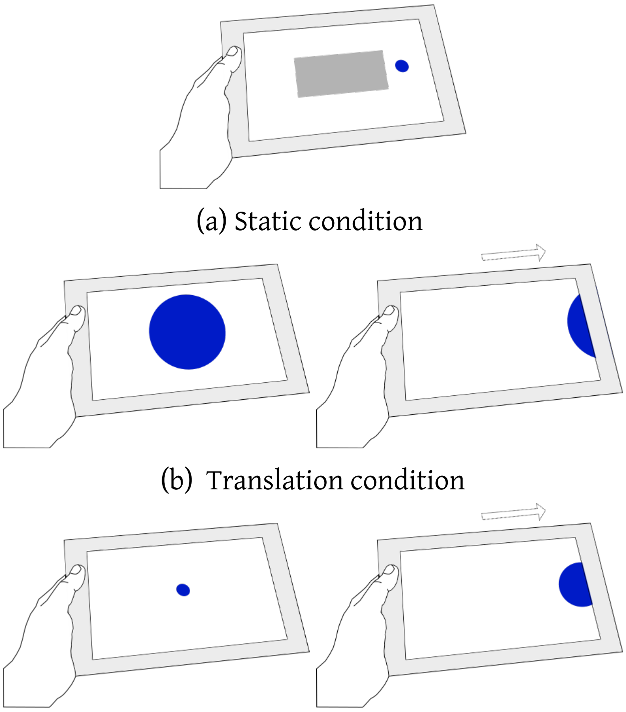


© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



target and the background are attached together. By this respect, each animation (zoom-in/translation) applied to the target is also applied to the background at the same time. At the end of the visualization, participant had to tap on the target by touching the ghost space at the corresponding location, lifting the finger, and then pressing the "stop" button to validate the tap. If the tap was inside the target area, a green circle appears on the tablet to confirm the successful trial. If finger is raised off the target area then, a red circle appears on the tablet. Finally, participant hit the "next" button for the next trial.



(a) Static condition

(b) Translation condition

(c) Zoom condition

Figure 5.6: The three visualization conditions for target number 2. N.B: The arrow was not displayed to participants.

The order of *visualization* techniques was counterbalanced across participants. Inside each *technique*, *sizes* were experimented separably and presented in a random order. For each *size*, participants completed 24 trials that varied target positions with four repetitions for a given position. Overall, we hence have a total of 3 *visualizations* $\times$ 3 *sizes* $\times$ 6 *positions* $\times$ 4 *repetitions*= 216 trials per participant. Before each technique, participant could practice the different conditions.

After each technique, participants responded to 5-point Likert-scale questions (strongly disagree to strongly agree): i) I performed well, ii) I accomplished



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

the task rapidly, iii) I needed a lot effort to finish the task, iv) I needed to concentrate to accomplish the task; v) I felt frustrated/stressed/irritated/annoyed, vi) I felt confident in my ability to hit the target, vii) I enjoyed interacting with the technique. At the end of the experiment, participants were asked to rank each technique according to their preferences. The average duration of the experiment was 45 minutes.

### Results

The dependent measures are *number of errors* (targets that were not correctly selected were marked as errors), *target distance* (distance between the finger tap position and the center of the target) and *trial time* (trial time is measured from the target apparition, to target successfully selected). We have also analyzed subjective responses. All analyses are multi-way ANOVA. Tukey tests are used post-hoc when significant effects are found. In the following, we report the results for each of the dependent variables.

### Error rate and distance to target

There were significant main effect of *size* ($F_{2,18} = 122.29$, $p < .0001$) on *error rate*, with the *large* size providing the most accurate performance (mean=34.18, s.d=5.43), then *medium* (mean=57.58, s.d=5.84) and then *small* (mean=72.86, s.d=5.05). Post-hoc tests revealed that the performance deteriorated more significantly across decreasing target sizes ($p < .05$).

We also found significant main effect of *position* ($F_{2,45} = 2.54$, $p = .04$) on *error rate*, but there was also a significant *visualization* × *position* ($F_{10,90} = 2.64$, $p = .007$) interaction. Post-hoc tests revealed that when pointing at target 4, the *overview* visualization reduce significantly the *error rate* as compared to *pan* visualization. We also found that, when using *overview* visualization, the performance is significantly better when pointing at target 4 than when pointing at target 5.

Concerning the distance to target, there was significant *visualization* × *position* ($F_{10,90} = 2.14$, $p = .02$) interaction. Post-hoc tests revealed that when pointing at target T3, the *overview* visualization reduces significantly the *target distance* as compared to *pan* visualization.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

| | ***Zoom*** | | ***Pan*** | | ***Overview*** | | ***Friedman*** |
|---|---|---|---|---|---|---|---|
| | mean | s.d | mean | s.d | mean | s.d | $\chi^2$ |
| Performance | 33 | .92 | 2.3 | .58 | 3.2 | .81 | 3.81 |
| Time | 3.3 | .82 | 3.3 | .77 | 3.9 | .89 | .69 |
| Physical | 2.1 | .61 | 1.9 | .61 | 2 | .65 | 1 |
| Concentration | 4 | .58 | 3.6 | .52 | 3.1 | .84 | 4.9 |
| Frustration | 2.5 | .83 | 3.3 | .71 | 2.4 | .83 | 2.4 |
| Confidence | 3.3 | .77 | 3 | .50 | 3.2 | .56 | 2.24 |
| Enjoyment | 4.1 | .45 | 3.7 | .51 | 4 | .29 | 3.26 |

Table 5.2: Mean and s.d questionnaire responses, with 1=strongly disagree, and 5 = strongly agree.

### Trial time

There was significant main effect of *visualization* ($F_{2,18}$ = $12.50$, p = $.0004$) on *trial time* with *overview* (mean=4360, sd=333 ms) significantly slower than both *pan* (mean=3074, sd=329 ms) and *zoom* (mean=2862, sd=188 ms). Interestingly, we found that there was no significant *visualization* $\times$ *size* (p=$.55$) nor *visualization* $\times$ *position* (p=$.75$) interaction, suggesting that the inconvenience of *overview* are consistent across the different *sizes* and *positions*.

### Subjective results

We recall that participants were asked to rank the three visualization techniques after completing the experiment. Overall, the *overview* technique was ranked 60% first, 30% second and 10% third. While, *zoom* technique was ranked 30% first, 40% second and 30% third. The *pan* technique was ranked only 10% first, 30% second and 60% third.

Participants were also asked to rate each visualization technique (see Table 6.1). Friedman tests revealed that there were no significant differences between the three visualization techniques in term of task performance, the time, the physical effort and the concentration needed to accomplish the task, the feeling of frustration and confidence in order to hit the target and the enjoyment of using the visualization technique.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



We correlate these findings with comments from participants. No consensus has been reached across participants for each technique. On one hand, four out of ten participants (P4, P9, P10 and P11) found the *overview* technique difficult because of the scaling they had to performed. But on the other hand, the rest of the participants explicitly said that this technique was easier because scaling is easy to do. Again, four participants (P4, P6, P9 and P11) found that the *zoom* technique gave better information about the angle and the distance of the target. Whereas the rest found it was hard to estimate the target position. The *pan* technique frustrated five participants (P2, P7, P9, P10 and P11) because it seemed easier than it was to them. Interestingly, some participants (P3, P5, P6) had the feeling that with the *pan* technique, the targets went further than they actually went. As a strategy, those participants tapped a little bit before where they thought the target would be. Still, all participants stated that the task was harder than it seemed when presented. It is the only unanimous remark.

**Summary.** The key finding of our experiment is that the larger the target is, the more accurate the user is. These performance benefits were consistent across the different visualization techniques and target positions. We, also, found that *zoom* visualization led to the fastest target tapping followed by *pan* visualization and then the *overview* visualization. These performance benefits were consistent across the different target positions and sizes. These findings reinforce our belief that with the three visualization techniques, participants are able to tap on 'ghost' targets. In addition, participants did not make any preference to a specific visualization technique. This makes the three visualization techniques available for such a task.

We introduced GHOST TAP in this section and a small study that evaluates the impact of visualization on blind tap performance around the tablet. We found that the *overview* visualization is slower than the *pan* and *zoom* conditions. But significantly reduces *target distance* and for some targets, reduces error rate. GHOST TAP defines the plane around the tablet as *one* continuous extension of the screen. In the next section I describe the prototype presented in Section 5.2 by using the MIS design framework.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



## 5.3 MIS description of GhostTap

In GhostTap, the MIS is a plane attached to the tablet and centered on it. In this prototype, the only input is a tap gesture that centered the map on the tapped point. In the context of a tap on a widget, the tap would invoke the corresponding command. You can see the complete description in Table 5.3.

| Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|
| **Shape :** Plane<br>**Dimension :** 25cm x 25cm<br>**Position :** Tablet<br>**Orientation :** Tablet<br>**Material :** Invisible | Tablet fixed | Tablet | **Input :** Tap<br>**Output :** Invoke corresponding command<br>**Mapping :** Discrete and Absolute |

Table 5.3: MIS description of the GhostTap prototype

## 5.4 Conclusion

In this chapter, I first introduced Tabslab. A concept derived from MIS that defines midair slabs around the tablet. Two prototypes have been proposed in different contexts and with different uses of midair gestures. The MRI Visualization setup leverages on our learned capacity of manipulation gestures of objects while the drawing software explores WIMP interfaces for such slabs with the use of finger identification. Those applications raised majors questions about hybrid touch-midair interaction. Future work involves the study of what kind of compatible menu paradigm for tabslab interaction as well as further study and prototyping to integrate mobile touch interaction with tabslab. I then presented Ghost Tap that uses the plane around the tablet as an extension of what is displayed on the tablet. The planar that surrounds the MIS is a continuation of the tablet. The MIS and the tablet are part of the same input entity.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



In the next chapter, I further investigate the use of the MIS as a space fused to the system. The MIS has the role of a transition space from touch to midair interaction. I will present an interaction technique that uses the space right in front of a wall display to improve interaction with objects out of reach. As well as a study that compares the interaction technique with a touch-based technique. The proposed interaction technique, named Talaria, is a hybrid one allowing continuous transition between touch and midair and proves to be faster than touch for out of reach objects. This next chapter is quite different from the previous ones as it offers another vision of the concept due to the fact that the MIS is a transition space.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

© 2017 Tous droits réservés.
lilliad.univ-lille.fr

# Six

# Fusion of a MIS and a Wall Display

*"Life is pleasant. Death is peaceful. It's the transition that's troublesome."*

— Isaac Asimov

## Contents





© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



In the previous chapters, we saw two instantiations of the MIS concept that we could call standard applications of the concept. The first was the instantiation of dynamic invisible interaction spaces at arbitrary position for cursor control, the second was the use of the surrounding space of a tablet. In both instantiations, the MISs used are perceptible by the user. In this last contribution, I present the use of the space right in front of a wall display as a transition one for continuous transition between the physicality of touch and the symbolism of midair interaction. The MIS is part of the screen, so that touch and midair interaction are unified as one. This is why in the following, I will use the term of continuous interaction. Talaria is completely different from the last two instantiations since here the user is not conscious of the MIS attached to the screen. The MIS is here to extend the capacity of the touchscreen but is not here from the users' point of view. The MIS concept helped during the design process of the interaction technique and is not part of the user interface. In the following sections I will introduce Talaria and present a study that compares this technique to standard touch interaction, published in [92].

Large tactile displays are becoming ever more functional and affordable. This makes them increasingly adopted for public installations [86,99], as well as in small and medium-scale collaborative settings for a variety of tasks [96]. This is because the large display surface makes large quantities of information readily and visually accessible and easy to manipulate in natural ways by small groups.

However, the basic interactions currently afforded by large tactile displays are mostly limited to direct interaction [35]. This leads to major issues when manipulating information on such surfaces, that have been well revealed in research literature: (1) reaching content beyond arms' length is not easy. Accessing corners requires people either to squat/lean forward or to stand on tiptoe. And moving beyond arms reach requires users to walk along the display; (2) Interacting with other users on the same display. Indeed, when interacting with content on the display, people nearby can get on the way.

To alleviate these issues, Forlines et al [35] proposed a direct interaction technique allowing users to switch between absolute and relative actions. However, relative mapping poses other problems such as clutching that might prove cumbersome on very large displays, including movement discontinuities when switching modes, or adding external devices such as a tablet [83] to control indirect activation.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



We propose to keep continuous interaction active throughout an operation without requiring users to move when their physical limits are reached, in order to preserve the naturalness of direct interaction on large displays. This is because in human-to-human interaction, gestural pointing arises naturally as it is one of the first gestures learned in life to point at objects out of reach [48]. We introduce TALARIA to leverage on natural deictic human ability and the directness of tactile interaction.

TALARIA is an interaction technique combining Touch actions and Midair pointing that enables accessing unreachable content on a large touch display without resorting to walk alongside its surface. The core idea is to start a Touch gesture on the display surface and to finish it by Midair pointing to push content away or inversely to retrieve out-of-reach content. This has two key advantages. First the transition between Touch and Midair is continuous. Therefore, users do not have to explicitly switch between the two modalities. Second, the semantics of pointing are well understood in human-to-human interaction [27]. These, coupled with the semantics of proximal relations and deixis make the proposed technique very powerful as it leverages on well-understood human-human interaction modalities [48]. Indeed, our technique leverages on contextual information given by proximity relations to the display as well as explicit spatial relations afforded by deixis to provide implicit arguments to most commands [6].

Compared to the MIS framework, TALARIA is an interactive space attached to the wall display directly on the screen. The space acts directly on the mouse cursor of the computer screen. Pointing gestures are used in it. As said before, the transition between the space and the touch screen is continuous when wanted. More details on how the technique has been implemented will be found in the remaining.

## 6.1 Large Displays and Touch-Midair Interaction

In this section, we review previous work on Touch interaction on large displays. We also present hybrid interaction techniques where Touch and Midair interaction appear intermixed.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



### Touch Interaction on Large Displays

Touch interaction on large displays suffers from the size factor. Reaching a target on a large surface can be easy if the area is in the user's physical reach. However, acquiring it can be tiring if the target position is distant from one's current position. In this case, users have to walk/jump/bend down, and in some cases it can be impossible for them to reach the target if someone or something obstructs the movement (*e.g.*, other users) or the target is simply out of reach (*e.g.*, too high). This has led to orthogonal techniques being developed. For example, some have proposed to switch between absolute and relative interaction by counting the number of fingers in contact with a surface [83] or by using a paper sheet [113] or by directly emulating a pad when needed through a multi-touch gesture [37] or by clicking on a widget with a pen [35]. However, those techniques require explicit switching between distinct interaction modes which can be quite frustrating. Others advocated for providing users with a miniature desktop to directly bring distant objects within reach or dragging closed objects to distant targets [23,55] or temporarily bring distant targets within arm's reach to interact with them [10]. However, those techniques did not allow users to explore and navigate the whole display and require knowledge about the objects of interest. In parallel, others have proposed to use direct pen gesture to throw objects on a distant target [93] or by using foot gestures to bring distant bottom objects instead of bending down to reach them with ones' hands [60]. However, those techniques only allowed users to coarsely push away objects or bring bottom objects within reach.

### Hybrid Interaction: Touch and Midair Interaction

Several works proposed to mix Midair with Touch actions to enrich interaction. Specific approaches adopt Midair interaction for users situated far away from the display and Touch interaction for those close to the display [8,75,122]. Other methods adopt a vocabulary mixing Midair gestures and Touch input devices [20]. However, those techniques, do not support continuity in gestures when switching from one mode to another. To deal with this limitation, in TractorBeam [85], pen-based interaction is the same whether operating on screen or above the tabletop. The cursor is determined by a raycast of the pen on the tabletop for Touch and Midair interaction. In [24,79] authors have proposed to continue the interaction when switching from surface to Midair manipulations. The surface of a tabletop and the space above it are thus considered as a continuum. In this



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



configuration, several gestures are proposed. In particular, *extended continuous gestures*. The gesture begins in a certain space and continues or finishes in another. The transition between Touch and Midair does not alter the current action avoiding discontinuities while affording a more fluid operation. TALARIA builds upon this previous work to insure continuous control of the cursor.

The technique we propose here rely solely on absolute mapping to preserve the natural directness of large touch displays. The continuity of the gesture is the key in Talaria. There is no interruption of the user's action to switch between techniques. The user is able to retrieve and move away content.

## 6.2 Talaria: Continuous Interaction for Wall Displays

TALARIA is designed to overcome the limitations of direct touch interaction with large display surfaces. We were inspired by "Talaria", the winged sandals of Hermes that allowed the god to fly as a bird in Greek mythology. Hermes also is the god of transitions and boundaries. TALARIA allows one to reach past the boundaries of physical space and body abilities to enable two main actions: (1) pushing away an object and (2) retrieving a (possibly distant) object. In each interaction, we define two modalities: one-handed and two-handed.

In order to PUSH-AWAY an object when using the one-handed modality, users may start dragging the object by directly touching it and then continuously dragging the object when switching to midair interaction (see Figure 6.1.a). We name the transition from touch to midair TAKE-OFF. To discriminate between TAKE-OFF and a finger release, we defined a velocity threshold. The velocity threshold was determined from preliminary tests conducted with three people. When the finger is lifted off the display, if its velocity is above the threshold, then TAKE-OFF is activated and the object is then controlled in midair. By doing so, TALARIA does not affect the standard touch interaction. In both touch and midair modes, the mapping between the user's hand and the controlled object is absolute as we wanted to keep the directness of touch interaction on large displays. In the two-handed modality, when users touch directly an object moving for example their non-dominant hand (NDH) and then make a TAKE-OFF of the object by their NDH while point with their dominant hand (DH) in midair, the selected object is immediately dragged to the pointed-at position (see Figure 6.1.b).



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



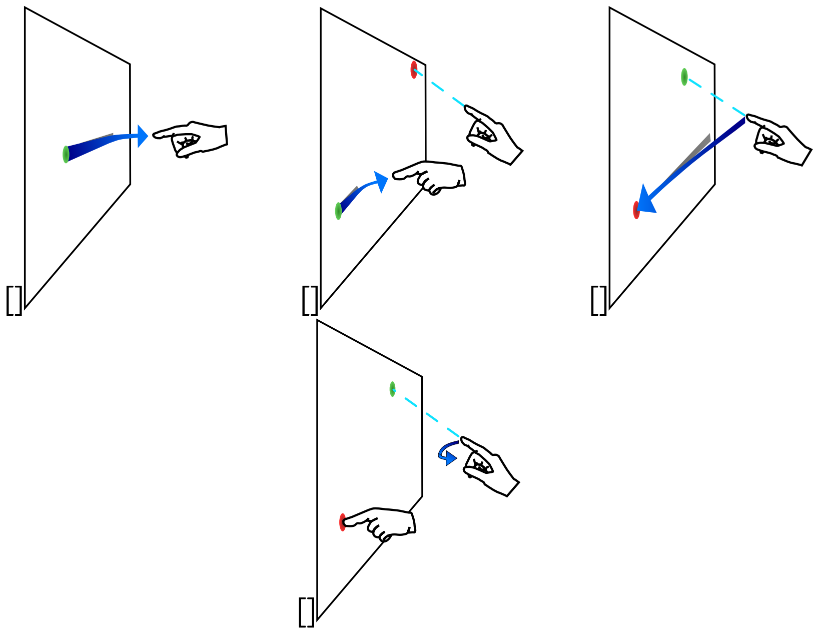



Figure 6.1: Push away ((a) one-handed and (b) two-handed) and retrieving ((c) one-handed and (d) two-handed) actions.

As for retrieving objects, in the one-handed modality, the user has to select in midair the object of interest, drag it in front of him/her and then touch it directly when the object is close to her/him (see Figure 6.1.c). Once in touch interaction, the user can perform any standard touch manipulation on the object. In the two-handed modality, when retrieving an object, users may select an object by pointing at it using their DH for example and then make a flick gesture with the DH while touching the display with their NDH, the selected object is immediately dragged to the touched position (see Figure 6.1.d).

**Proof of concept.** In order to test Talaria in a real scenario, we simulate mouse input to integrate Talaria in MS/Windows. We then added the two actions. As in midair interaction, if we do not detect a click event then we generate a click event when pointing with one hand while the other hand is touching the display. By doing so, at the end of a Push-Away action, users have to click on the screen to drop the object being dragged (see Figure 6.2). Conversely, at the beginning of the Retrieving action, after pointing at an object, users have to click on the screen to start dragging it.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

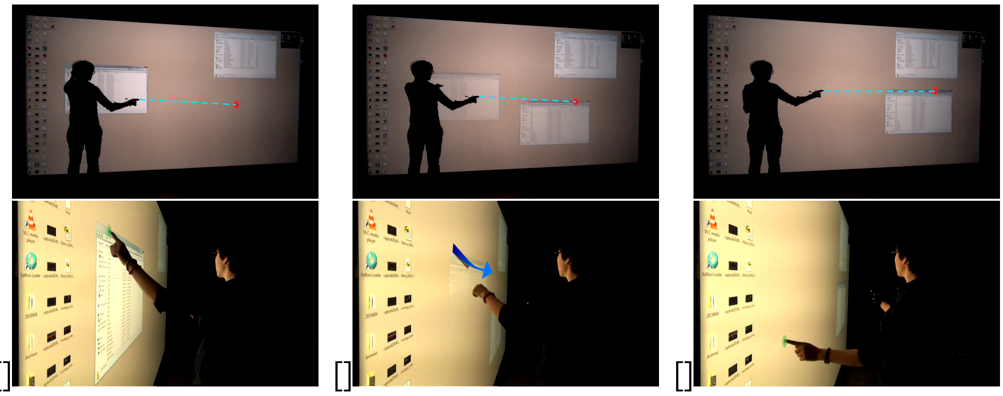


Figure 6.2: Two-Hand Push Away interaction example in use: (a) Touch the window with NDH while pointing at the screen, (b) Take-off happens; then the window moves to the pointed-at position, (c) Positioning the window and touching the screen to drop it.

## 6.3 Experiment: Dragging an Object

We conducted an experiment to compare performance between touch and Talaria techniques. Overall, we hypothesize:

**H1.** The selection time will be lower for Talaria than for touch. Since, users can interact from the same place over all the task when using Talaria.

**H2.** Talaria will reduce movement time compared to touch. Since, contrary to touch, users not have to move a lot when using Talaria.

**H3.** touch will be more accurate than Talaria. The direct touch interaction maximizes the opportunity to be the more accurate as touch interaction is more familiar.

**H4.** Talaria will reduce physical effort and increase the enjoyment compared to touch. Since the body movement can be optimized when using Talaria.

### Participants

10 participants (4 females) volunteered to take part in our experiment. Participants' ages varied between 24 and 32 years (mean age 26.7, SD=2.71 years). All participants were right-handed. All participants were regular users of smart



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



phones and tablet devices with multi-touch displays, and 3 participants were regular users of kinect games.

### Method

The experiment was conducted on a 4 m × 2 m multi-touch display starting from the ground. An infrared based touch frame sent touches to the operating system using TUIO protocol. In order to track the participant hand and forearm for Talaria, we used an infrared motion capture system. We setup six cameras above and around the display allowing us to track the participants interacting on and far from the display (up to 1.5 m). One constellation of markers was strapped to the forearm and another one to a glove participants had to wear for Talaria.

For Talaria, detecting a take-off was made when a touch release occurred. If at the release, the touch velocity was higher than a defined threshold, a raycast (forearm-hand) against the display yielded the cursor position and the touch release did not generate an event to the operating system. From that moment on, participants interacted with the control area in mid-air. There was no multi-touch support during the experiment. If multiple touches were detected, an error was triggered.

### Task, Procedure & Design

Participants were instructed to perform a sequence of object dragging as quickly and accurately as possible. In Talaria technique, participant was only informed by the one-handed modality and the click action is activated after 0.2s of holding on the object. Participant was then given the exact procedure to follow for each trial:

**State 1.** A blue circular control area with a diameter of 5 cm and a red circular target area appear on the display.

**State 2.** Touch the control area and hold it for 0.2 s to free it. Then after, the control area is free to move.

**State 3.** Drag control area over target area (see Figure 6.3). After holding for 0.2s, and if the center of mass of the control area was inside the target



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



area, the target turns green to confirm the successful trial and the next trial started. If finger is raised off the control area or hand is pointing outside the display, during the dragging task, then an error was counted, the target flashes orange and the trial is repeated.

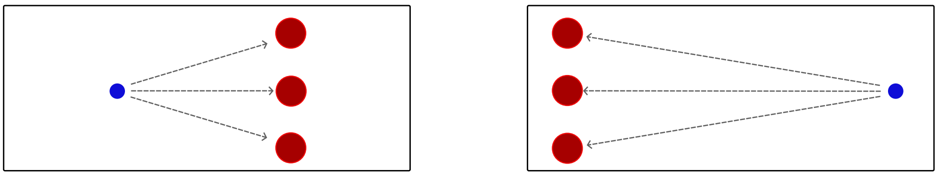


Figure 6.3: Target acquisition task in (a) the LEFT-TO-RIGHT direction with the SHORTEST AMPLITUDE and (b) the RIGHT-TO-LEFT direction with the LARGEST AMPLITUDE with the different target positions.

Dependent measures are analyzed using a $2 \times 2 \times 3 \times 2 \times 3$ repeated measures within-subjects analysis of variance for the factors: TECHNIQUE ( TOUCH, and TALARIA), AMPLITUDE (SHORTEST :1.5m, and LONGEST: 3m, where AMPLITUDE corresponds to the distance between the center of the control area to the center of the target area), TOLERANCE (S:10 cm, M: 20 cm and L: 30 cm where TOLERANCE corresponds to target diameter), DIRECTION (LEFT-TO-RIGHT and RIGHT-TO-LEFT) and POSITION ( TOP, MIDDLE, and BOTTOM, where POSITION corresponds to target position).

In the experiment phase, the order of TECHNIQUE, AMPLITUDE, TOLERANCES and DIRECTION was counterbalanced across participants. The experimental trials were then administered as 24 blocks of 15 trials, each block sharing a technique, an amplitude, a tolerance and a direction. Inside each block, the 15 trials (3 POSITION $\times$ 5 repetitions) were randomly presented to the participant – for a total of 360 trials per participant.

After each technique, participants responded to 5-point Likert-scale questions (strongly disagree to strongly agree): i) I performed well, ii) I accomplished the task rapidly, iii) I needed a lot effort to finish the task, iv) I needed to concentrate to accomplish the task; v) I felt frustrated/stressed/irritated/annoyed, vi) I felt confident in my ability to hit the target, vii) I enjoyed interacting with the device(s). "At the end of the experiment, participants were asked to rank each technique according to their preferences. Experiments took on average 45 minutes.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



## 6.4 Results

The dependent measures are Reaction Time, Movement Time, Error rate, and Number of Failed Attempts. We also analyzed subjective responses. All analyses are multi-way ANOVA. Tukey tests are used post-hoc when significant effects are found.

Due to a technical issue, the data of two participants were not completely logged. In the following we report results for each of the dependent variables for eight participants.

### Reaction Time

Reaction Time is the total control selection time, from the start of the trial, to the control area is successfully freed. Technique ($F_{1,7} = 114.4$, $p < .0001$) affected reaction time: Talaria was significantly faster (mean 1258 ms, s.d. 65) than touch (mean 2133, s.d. 102) by 40%.

As anticipated, there were significant main effects of Amplitude ($F_{1,7}$= 88.3, p<.0001), Tolerance ($F_{2,14} = 8.50$, p < .0001), Direction ($F_{1,7} = 9.58$, p = .017), and Position ($F_{2,14} = 6.16$, p = .001) on Reaction Time, but there was also a significant Technique × Amplitude ($F_{1,7} = 20.48$, p < .01), Technique × Direction ($F_{1,7} = 6.82$, p = .034), and Direction × Position ($F_{2,14} = 4.27$, p = .035) interaction. Post-hoc tests revealed that reaction time was significantly lower for Talaria than for touch ($p < .05$) with the Longest Amplitude with no significant difference for the Shortest amplitude. We also found that reaction time was significantly lower for the shortest Amplitude than for the longest one when using touch ($p < .05$). We correlate these findings with participants behavior: all our participants stayed between the control area and the target area when using touch technique for the shortest Amplitude to minimize their body movement and consequently reducing the reaction time. Reaction time was also found significantly lower with Talaria than with touch for both Direction (p< .05). However, while there was no significant difference between the Direction for touch, we found that reaction time was significantly lower when moving from Left-to-Right than the inverse direction when using Talaria. We correlate this finding with technical issues: regardless of the movement direction, our participants must always use their dominant hand (right) to drag the control area which promotes movements from left



© 2017 Tous droits réservés. lilliad.univ-lille.fr

to right. Interestingly, we found that there was no significant TECHNIQUE × TOLERANCE (p=.07) nor TECHNIQUE × POSITION (p=.07) interaction, suggesting that the benefits of TALARIA are consistent across the different TOLERANCES and POSITIONS. These results partially support **H1**.

### MOVEMENT TIME

MOVEMENT TIME is measured from the first control area movement, to target successfully selected. TECHNIQUE ($F_{1,7} = 5.87$, $p = .045$) significantly affected movement time: TALARIA was significantly faster (mean 3256 ms, s.d. 128) than TOUCH (mean 4193, s.d. 322) by 22%.

We also found main effects of AMPLITUDE ($F_{1,7}$= 42.37, p<.0001), and TOLERANCE ($F_{2,14} = 7.44$, $p < .01$) on MOVEMENT TIME and a significant TECHNIQUE × AMPLITUDE ($F_{1,7} = 6.57$, $p = .037$) and TOLERANCE × DIRECTION × POSITION ($F_{4,28} = 2.72$, $p = .045$) interaction. Post-hoc tests revealed that movement time was significantly lower for TALARIA (mean 3725 ms , s.d. 185) than for TOUCH (mean 5665 ms , s.d. 530) ($p < .05$) with the longest AMPLITUDE. Without surprise, we found that participants were significantly faster when using shorter AMPLITUDE than with the longer one when using TOUCH technique. Again we correlate this finding with participants positions. Interestingly, we found that there was no significant TECHNIQUE × TOLERANCE ($p = .79$) nor TECHNIQUE × DIRECTION ($p = .56$) or TECHNIQUE × POSITION ($p = .84$), suggesting that the benefits of TALARIA are consistent across the different TOLERANCES, DIRECTIONS and POSITIONS. These results partially support **H2**.

### ERROR RATE AND NUMBER OF FAILED ATTEMPTS

Targets that were not selected on first attempt were marked as errors. Surprisingly, while TALARIA (mean 12%, s.d 1) is more accurate than TOUCH (mean 16%, s.d 2), there was not a significant effect of TECHNIQUE ($F_{2,14} = 3.83$, $p = .09$) on ERROR RATE. However, there were significant AMPLITUDE × POSITION ($F_{2,14} = 4.18$, $p = .03$) and DIRECTION × POSITION ($F_{2,14} = 6.64$, $p < .01$) interactions.

Similarly to ERROR RATE, while we found that TALARIA (mean 15%, s.d 2) reduced the NUMBER OF FAILED ATTEMPTS as compared to TOUCH (mean 22%,



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



| | touch | | Talaria | | Wilcoxon | | |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Z | Sig | r |
| Performance | 3.37 | .82 | 3.25 | .61 | .21 | 1 | - |
| Time | 2.5 | .74 | 3.75 | .88 | -1.49 | .05 | .37 |
| Physical | 2.63 | .90 | 1.63 | .73 | -1.34 | .05 | .33 |
| Concentration | 2.38 | .97 | 2.13 | .57 | -.35 | .84 | |
| Frustration | 2.25 | 1.03 | 1.25 | .80 | -2.37 | .03 | .59 |
| Confidence | 3.5 | 1.28 | 3.75 | 1.03 | -.15 | 1 | - |
| Enjoyment | 2.37 | .90 | 4 | .74 | -1.98 | .04 | .49 |

Note: Wilcoxon-Signed-Rank tests are reported at $p$=.05 (*) significance levels. The significatif tests are highlighted.

Table 6.1: Mean and SD questionnaire responses, with 1=strongly disagree, and 5 = strongly agree.

s.d 3) by 31%, there was not a significant main effect of Technique ($F_{2,14} = 26.8$, $p = .09$) on the number of failed attempts. However, there was a clear effect of Amplitude ($F_{1,7} = 6.75$, $p = .03$) on the number of failed attempts with significant Direction × Position ($F_{2,14} = 4.50$, and Tolerance × Direction × Position ($F_{4,28} = 2.77$, $p = .04$) $p = .03$) interaction. These results lead us to reject **H3**.

### Subjective Results and Observations

We recall that participants were asked to rank the two techniques conditions after completing the experiment. Overall, Talaria technique was ranked 88% first and 12% second.

Participants were also asked to rate each technique condition. Overall, participants found that Talaria was faster, demands less physical effort, implies less concentration and less frustration, while being more confident and more enjoyable than touch technique. However, Wilcoxon-Signed-Rank tests showed that there were significant differences between the two Technique conditions only for time, physical effort, frustration and enjoyment (see Table 6.1), supporting **H4**.

We correlate these findings with comments from participants that felt that touch technique was cumbersome and required more effort and time. Some quotes are: *"It is really tiring to move across the display"*, *"They should put those kind of big displays in gyms!"*. Additionally, in order to reduce effort, some participants



© 2017 Tous droits réservés. lilliad.univ-lille.fr

were observed changing their hand as well as their fingers. Three participants were observed using their right hand when the DIRECTION was RIGHT-TO-LEFT and inversely using their left hand when the DIRECTION was LEFT-TO-RIGHT to reduce arm movement. Interestingly, all our participants tried to reduce their body movement (*e.g.*, walking). For instance, all participants stayed in the middle between the control area and the target area when the AMPLITUDE was 1.5 m and moved only their arm to accomplish the task. However, for the 3 m condition, all participants tried to minimize their walk by stretching their arms to select the control area or the target. One participant moved very slowly to avoid losing the control area and said "*this technique is the most frustrating, so I prefer to move slowly and be accurate to not repeat the trial!*". Surprisingly, two participants decide to run alongside the display, to achieve faster execution. However, after a couple of tries they stopped running as they found that the task became too tiring. Another participant stepped back after selecting the target to save time.

In contrast, participants found that TALARIA was both easier and faster than direct interaction. All participants were enthusiastic to touch the surface and continue interacting when they switched to midair interaction, and witnessed they "*feel more free*", "*having super powers*", "*are super heroes*", while one participant said : "*when I take off, continuing to control the control area feels like a dream*". Surprisingly, while no time or distance constraints were given for touch and midair interaction (the only condition was to stay in the 1.5m area before the display when switching to midair), all participants freed the control area and after few seconds switched to midair interaction which limited their touch distance. For instance, we found that the touch covered distance was in average equal to 18.53 cm (s.d .76 cm) for the shortest AMPLITUDE and 36.07 cm (s.d 4.22 cm) for the longest AMPLITUDE. Wilcoxon-Signed-Rank tests showed that there was a significant difference between the two AMPLITUDES ($Z=-16.52$, $p<.0001$). Additionally, all participants have limited their body movements and stayed close to the start position of the control area (*i.e*, the viewing distance is similar to AMPLITUDE value and the viewing angle quite sharp) which affects the visual appearance of both control and target shapes. Consequently, participants felt that when in midair, they needed to concentrate. Some quotes: "*it is simple to select the control area, but the dragging task was difficult in some cases as I didn't see clearly the other end of the display*", "*as I am on the opposite edge of the display, I need to concentrate to correctly select the target, but this technique is funnier and easier.*" To have a better view, two main strategies were used. For instance, four participants took a step backward or leaned back while the rest stayed at arm's length distance from the screen through the session.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



## 6.5 Discussion and Design Guidelines

Our key finding is that Talaria technique improved both the selection and movement times, increases the enjoyment and decreases the physical effort over conventional touch techniques, without compromising accuracy. The performance benefits were consistent across different Tolerances and target Positions. Our analysis suggest also that Talaria is best combined with longer Amplitude and Left-to-Right movement direction without decreasing performance on shorterAmplitude and Right-to-Left movement direction. Additionally, our findings indicate that touch interaction on large displays is more appropriate when it occurs in front of the user. However, from the moment users must move along the display to complete tasks, touch interactions became unsuitable and even boring in some cases.

Informed by our experimental findings and discussion, we outline relevant guidelines for designing interaction techniques on large displays:

- touch interaction on large displays works best in a restricted space: in front of the user and targets must lie within users arms' reach. Indeed our participants often expressed dissatisfaction when making distant target selections requiring longer selection times and featuring lower accuracy on selection tasks.
- Midair interaction should be preferred for distant interactions (*i.e.*, beyond arms' length). Our findings indicate that beyond arms' length target selection, Talaria outperforms touch interaction.
- Design for flexible input by allowing users to combine touch and midair interaction. Our participants prefer Talaria as it supports both modalities synergistically.
- Provide continuous transition between touch and midair interaction as our participants insisted on the fun brought about by this transition while reducing frustration.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



## 6.6 MIS description of Talaria

Again, TALARIA can also be described thanks to the MIS framework. With this technique, we use a Rectangular MIS that covers the entire screen with a small depth. There are two inputs allowed in this MIS. The first one is the take off of the finger from the screen an that goes through the MIS in order to switch to midair pointing. The second input is the opposite gesture which is the landing of the finger on the touch screen going through the MIS. The complete description is presented in Table 6.2 below.

| Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|
| **Shape :** Rectangular<br>**Dimension :** Screen<br>**Position :** Screen<br>**Orientation :** Vertical<br>**Material :** Invisible | Screen Fixed | Screen | **Input :** Finger take off ; Finger landing<br>**Output :** Switch to midair ; Switch to touch interaction<br>**Mapping :** Discrete and Absolute |

Table 6.2: MIS description of TALARIA

## 6.7 Conclusion and Future Work

We presented Talaria, a novel interaction technique on large displays that combines touch interaction and midair pointing to access out-of-reach content. This is done by using a MIS as the space that allows a seamless transition between touch and midair interaction.

We then conducted an experiment to evaluate and compare Talaria with touch interaction. Our findings indicate that Talaria improved the selection time and the enjoyment over touch, without compromising accuracy. Finally, we hope that this work will advance our knowledge for direct dragging on large displays and that Talaria technique will prove useful by adding to the growing



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



toolkit of large display interaction techniques as it is seemingly well-suited to perform both casual and group collaborative tasks in both natural and appealing ways.

Future work would look at adopting multiple fingers simultaneously for touch interaction with Talaria as well as using the technique by multiple users in the same time. Finally, one potential usability issue of our technique is that, with Talaria, when switching to the midair modality, participants have a distorted view of both control and target shapes due to a sharp vision angle. To visually help users, the technique could add a magic distortion lens [14]. Future work will study the effect of adding this kind of lens on the distortion around the manipulated object when using Talaria. This should not detract from extending proxemics to other large-scale display interactions a trend that we hope to have furthered with the present work.

This was the last chapter presenting my contributions and instantiations of the MIS concept. In the following section, I will conclude this dissertation by discussing the future for MIS-based techniques and interesting aspects of the MIS concept in further depth.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# Seven

# Conclusion

*"It is not the strongest of the species that survives, not the most intelligent that survives. It is the one that is the most adaptable to change."*

— Charles Darwin

In this dissertation the novel concept, MIS, and its corresponding framework have been presented. I gave some leads for the design space exploration supported by the framework as well as examples of interaction scenarios that use MIS-based interaction. The other interesting part of MIS is the re-reading of past work made in chapter 3 that shows the unifying power of MIS while leaving huge design possibilities to be explored and raising important questions on our use of spaces in HCI. In the last three chapters, we saw *three* instantiations of the concept. Each of them proposed a different way to integrate MIS to interactive systems. The *first one* was in the context of remote display where I presented an eliciting study for creation and deletion gestures of MIS in this context, and from the findings proposed a prototype. Here the MIS is a distinct interface for users to control a mouse cursor on the remote display. This work and the MIS concept have been published in the IUI proceedings [91]. The *second instantiation* was the use of MISs for around-the-tablet interaction. The MIS and the tablet are linked together and can be seen as one interactive system unlike the first instantiation. I presented two applications: one with several MISs to invoke specific commands for MRI visualization [90] and for a drawing application, and the second where the MIS is a continuation of the tablet screen. The *third instantiation* was in the context of interaction on wall displays where a MIS is placed right in front of the screen [92] and has the role of the transition space



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



allowing touch and midair interaction hybridization. In the following, I discuss future work around MIS and aspects of MIS that need to be studied.

## 7.1 Future Work Around MIS

In this section, I discuss future work and perspectives on MIS. This future of interfaces should extend and make evolve the typical WIMP paradigm we currently use in our everyday human-computer interaction life. Indeed, without a new paradigm compatible with midair interaction, acceptance of such techniques by end users will never happen. Finally, I talk about a side of MIS I mentioned little in the previous chapters which is what I call *reflection* of MIS.

But first, in order to design good MIS-based interaction techniques, we must know better about how we would interact with them. Future work should investigate what is the best way for volumic MIS creation (*e.g.*, sphere, cube,...), complex surfaces, and what kind of gestures are possible and acceptable in those different MISs. These MISs must also be studied in diverse contexts of use and with different kind of data. Arranged in a certain way, are 2D MIS enough for 3D modeling? And for huge dataset visualization? In which context do we need 3D MIS? From these studies or in parallel, new instantiations of MIS-based interaction techniques would emerge. Along with these considerations, work has also to be done on interfaces with which we could interact via MIS.

### Novel Interfaces for MIS

Most of our interfaces today are based on the WIMP paradigm. Though it provides effective and easy interface navigation, command invocations, etc..., it has been designed with point and click logic in order to be best suited for mouse-based interaction. Using gestures is no longer only point and click, though it can include it. The spectrum of expression is much larger. The major strength of WIMP interfaces is that all of the commands are displayed to the users and organized. By exploring the menus and the buttons, users know what commands are available. This should not be lost in future interfaces. Feedforward seems then to be necessary for commands explorations. One can think of OctoPocus [9] presented by Bau et al. as a dynamic guide for learning gesture-based commands. This was presented for 2D gestures but a midair version of it could be interesting to investigate. This raises the question of feedback. This work gives continuous



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



feedback and feedforward to users as they perform the gesture. Feedback and feedforward might depend on the context of use of the MIS-based interaction technique. Maybe, we should turn to games where midair gestures were first use for general public and study both success and failures, rejected and adopted propositions. Because WIMP paradigm is little embodied in games, users have more objective or less biased point of view concerning midair-gesture-based interfaces due to little legacy in comparison to other public. More generally, novel interfaces should be adaptable to interaction like a MIS is. This plasticity of interfaces has been investigated in [7]. The author argues that plasticity relies on distributed systems, dynamically reconfigurable and composed of heterogeneous logic blocks. He then proposes a model for dynamic components. This kind of work definitely support the modularity of MIS concept and can enable the development of such interaction techniques.

#### Reflection of MIS

MIS reflective nature means that by performing specific actions in a MIS or on its frontiers we can modify the properties of the MIS. In other words, MIS parameterization is done thanks to MIS interaction. In order to explore this side of MISs, we can make use of the framework to avoid scattering. I propose in the following a basic and naive exploration of MIS parametrization.

First as we interact with a MIS, it has its own attributes. The important difference here is that the MIS has two Action Reference Frames. One is a specific device and the second is itself. Then it is in the Interaction attributes that this latter option is detailed and split into two parts (the one for interacting with the device and the other for the MIS itself). To avoid accidental modifications of the MIS, it would be best to design a gestural language to modify a MIS. For this, I take my inspiration from the work of Chen et al. [20] that uses touch to segment midair gestures and proposes a new vocabulary for one-finger commands. Here, I would use postures and gestures for this parametrization vocabulary. In the following I propose diverse kind of gestures to modify each component of the framework. Those propositions are only hints and should be further developed via elicitation studies, prototypes and validation studies. From the study in chapter 4, we know that participants where manipulating the MIS like it was a real physical object. Then gestures for parametrization should take from our daily gesture to guarantee a coherent vocabulary, a better comprehension of it and then an easier learning.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

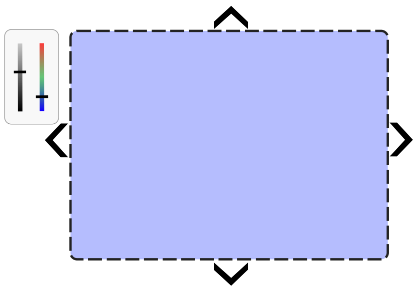


Figure 7.1: Example for parameterization of material component of a MIS

The Geometric Definition component(GD), since it represents the shape and material of the MIS should be modified by starting with a posture that communicates the act of modifying a shape. The Geometric component encompasses the position, the dimensions, the shape and the material. For the position, I suggest a spread fingers hand posture followed by a grab like we were grabbing the MIS to start moving it. To set the new position, users should then release the grab. The position is change relatively to the Interaction Reference Frame (IRF). Concerning the dimensions, and the shape, re-performing the defining gesture by starting in the MIS. The material is strongly linked to the hardware and might not be completely parameterizable. In the case of visualization of the MIS, tuning the material could be possible via a displayed menu. It is important to note here, that if visualization of MIS is made possible by interaction hardware, feedback for *every* component changing must be given to users.

Concerning the Action Reference Frame component (ARF), the gesture must express *"This* MIS *now acts on this entity"*. In chapter 4, changing the screen on which we controlled the mouse cursor was made by a simple directional swipe gesture toward the screen of interest starting in the MIS and finishing out of it. This gesture is simple and can easily be associated with the meaning of *"acting on this screen"*. In contexts of high density of interactive setups, such gesture could yield to wrong re-affectation. Disambiguation could then be done by allowing a small dwell time at the end of the gesture. During this time, the potential devices should give feedback on the current affectation. The question of amplitude of the gesture is also important depending on the context of interaction as it can have different meanings and determine the relevance of the gestures for users.

The case of Input Reference Frame (IRF) is different since affectation is possible on objects but also world coordinates and user's body parts (or relative to it). The meaning of the gesture should express the anchorage of the MIS. For this, I suggest almost the same gesture than for changing position. A spread



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



fingers hand posture followed by a grab like we were grabbing the MIS to start the re-affectation of the anchor. To affect the new IRF, users should then release the grab directly followed with another spread fingers posture and a fast small top-bottom gesture of the hand as if we were sticking the MIS to something.

The Interaction Attributes component(IA) is a special component as it is the one which is in direct relation with the controlled interface. Parametrization of this component is then difficult and not advised as it can completely change the interaction. The only parameterizable part would be the mapping one. One could imagine changing the gain constant of a relative mapping. But again, for interaction designer, giving such power to the user should be done wisely as it can make the interaction more difficult for non-experts users.

To conclude, like said in the last paragraph, the reflection of MIS raises the question of accessibility. Can we give access to reflective nature of MIS to end-users? Is it too difficult? To what extent users can be designers? The amount of possible parameterization of MISs has to be studied according to each component values and for different expertise levels. The other question is, can this power of reflection be dynamically adapted according to users expertise during interaction? All these questions are part of the future work on MIS that has to be done for better suited MIS-based interaction.

The issues of novel interface paradigm and the reflective nature of MIS yield to other interesting questions. This last section about future work with MIS are only leads for further research in very different fields. Designing and thinking MIS concept open up a large field of research ahead even though I presented it as a way to overcome the inherent difficulties of midair interaction that prevents a larger acceptance by end-users. Thanks to its diverse dimensions combined with our early learning and perception of space, I think midair interaction in spaces has the potential to extend and profoundly transform human-computer interaction. I hope that what has been presented in this dissertation will benefit and inspire future research due to the novel questions the MIS concept raises.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# Appendices



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

© 2017 Tous droits réservés.

lilliad.univ-lille.fr

© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# A

# MIS description of previous work

This Appendix gathers the MIS description of every previous work on interactive spaces I discussed in Chapter 2 and Section 3.2.

When several spaces are used in the work, the spaces attributes of the Geometric definition, the Input Reference Frame and the Action Reference Frame components are separated by a semicolon. In the Interaction Attributes components, the space is specified for the Input. The Output and the Mapping attributes are then in the same order than in the Input section.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [47] | **Shape :** Plane<br>**Dimension :** Palm dimension<br>**Position :** Palm fixed<br>**Orientation :** Palm fixed<br>**Material :** Invisible | Smarthpone | Palm | **Input :** Pointing<br>**Output :** Push buttons<br>**Mapping :** Discrete and absolute |
| [46] | **Shape :** Plane<br>**Dimension :** Arm reach<br>**Position :** User's Non Dominant Hand<br>**Orientation :** User's Non Dominant Hand<br>**Material :** Invisible | Smartphone | User's hand | **Input :** Hand position; Hand gestures<br>**Output :** Tap; Strokes<br>**Mapping :** Discrete/ Absolute; Continuous/ Absolute |
| [25] | **Shape :** Plane<br>**Dimension :** User's hand<br>**Position :** User's hand<br>**Orientation :** User's hand<br>**Material :** Invisible | TV | User's hand | **Input :** Tap<br>**Output :** Push Button<br>**Mapping :** Discrete/ Absolute |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [72] | **Shape :** Plane<br>**Dimension :** Fixed by designer<br>**Position :** Hand position<br>**Orientation :** Hand Orientation<br>**Material :** Invisible | TV | User's hand | **Input :** L-shape; Tap<br>**Output :** Creation/ re-anchoring; Push Button<br>**Mapping :** Discrete/ Absolute; Discrete/ Absolute |
| [4] | **Shape :** Camera view cone<br>**Dimension :** Camera view cone<br>**Position :** Shoe position<br>**Orientation :** Shoe orientation<br>**Material :** Invisible | Smartphone | User's shoe | **Input :** Finger count; Pinch gesture and position; Arm triangle<br>**Output :** Number; Pick up the phone; Volume level<br>**Mapping :** Discrete; Discrete or continuous; Continuous and absolute |
| [65] | **Shape :** Cuboid volume<br>**Dimension :** Palm dimension<br>**Position :** Above palm<br>**Orientation :** Palm orientation<br>**Material :** Invisible | Any device MIS-enable | Palm | **Input :** Hand pose<br>**Output :** TBD<br>**Mapping :** Discrete/ Absolute |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [73] | **Shape :** Forearm surface<br>**Dimension :** Forearm<br>**Position :** Forearm<br>**Orientation :** Forearm<br>**Material :** Invisible | Any compatible device | User's forearm | **Input :** Tap<br>**Output :** Push button<br>**Mapping :** Discrete/ Absolute |
| [71] | **Shape :** Spherical<br>**Dimension :** Arm length<br>**Position :** User centered<br>**Orientation :** Around user<br>**Material :** Invisible | Mobile device | User's body | **Input :** Angles on the sphere<br>**Output :** Select associated command<br>**Mapping :** Discrete/ Absolute |
| [18] | **Shape :** Tore<br>**Dimension :** Arms reach<br>**Position :** User centered<br>**Orientation :** Around user's body<br>**Material :** Invisible | Mobile phone | User's body | **Input :** Angle on the tore<br>**Output :** Workspace switching<br>**Mapping :** Discrete/ Absolute |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [19] | **Shape :** Surfaces<br>**Dimension :** Designer fixed<br>**Position :** Around or on the user - Designer fixed<br>**Orientation :** Designer fixed<br>**Material :** Invisible | Smartphone | User's body | **Input :** Smartphone position<br>**Output :** Associated command<br>**Mapping :** Discrete/ Absolute |
| [81] | **Shape :** Camera viewing cone<br>**Dimension :** Camera viewing cone<br>**Position :** User's position<br>**Orientation :** Camera orientation<br>**Material :** Invisible | Smartphone | User's body | **Input :** Hand pose ; Standard touch interaction<br>**Output :** Associated command ; pan/ zoom/ tap<br>**Mapping :** Discrete/ Absolute ; Continuous/ Discrete/ Absolute |
| [75] | **Shape :** Leap Motion viewing cone<br>**Dimension :** Leap Motion viewing cone<br>**Position :** Leap Motion position<br>**Orientation :** Leap Motion orientation (along the tigh)<br>**Material :** invisible | Remote display | User's tigh | **Input :** Hand postures ; Hand gestures<br>**Output :** Associated commands<br>**Mapping :** Discrete/ Continuous/ Absolute/ Relative |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [76] | **Shape :** Camera viewing cone<br>**Dimension :** Camera viewing cone<br>**Position :** Camera position<br>**Orientation :** Camera orientation<br>**Material :** Invisible | Smartphone | Smartphone | **Input :** Flexion+extension ; Fingers positions along x axis<br>**Output :** Selection ; Rotate menu<br>**Mapping :** Discrete/ Absolute ; Continuous/ Relative to center |
| [103] | **Shape :** Camera viewing cone<br>**Dimension :** Camera viewing cone<br>**Position :** Camera position<br>**Orientation :** Camera Orientation<br>**Material :** Invisible | Smartphone | Camera position | **Input :** Hand pose ; Hand gestures<br>**Output :** Associated commands<br>**Mapping :** Discrete/ Absolute ; Continuous/ Relative or Absolute |
| [50] | **Shape :** Round Plane<br>**Dimension :** 10 cm diameter<br>**Position :** Smartwatch position<br>**Orientation :** Smartwatch orientation<br>**Material :** Invisible | Smartwatch | Smartwatch | **Input :** Finger Position ; Tap<br>**Output :** Pointer position (angular or xy) ; Tap<br>**Mapping :** Continuous/ Absolute or Relative ; Discrete/ Absolute |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [53] | **Shape :** Circle plane<br>**Dimension :** 40 cm<br>**Position :** Smartphone position<br>**Orientation :** Smartphone orientation<br>**Material :** Invisible | Smartwatch | Smartwatch | **Input :** Angle + distance from smartwatch<br>**Output :** Select associated item<br>**Mapping :** Discrete/ Absolute |
| [20] | **Shape :** Cube<br>**Dimension :** Smartphone screen dimension<br>**Position :** Above smartphone screen<br>**Orientation :** Smartphone orientation<br>**Material :** Invisible | Smartphone | Smartphone | **Input :** One finger gesture starting and/or ending with a tap<br>**Output :** Associated command (e.g zooming/ contextual menu/ etc...)<br>**Mapping :** Discrete/continuous/ Absolute/relative |
| [67] | **Shape :** Camera viewing cone<br>**Dimension :** Camera viewing cone<br>**Position :** Camera position<br>**Orientation :** Camera orientation<br>**Material :** Invisible | Smartphone | Smartphone | **Input :** Palm orientation (xy axis)<br>**Output :** 3D model orientation<br>**Mapping :** Continuous/ Absolute |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [12] | **Shape :** Cube<br>**Dimension :** Arm reach<br>**Position :** Smartphone position<br>**Orientation :** Smartphone Orientation<br>**Material :** Invisible | Distant display | Smartphone | **Input :** Hand position ; Button on smartphone<br>**Output :** Translation/ Rotation along 3 axis; Current task (translation or rotation mode)<br>**Mapping :** Continuous/ relative 1/1; Discrete/ absolute |
| [80] | **Shape :** Plane<br>**Dimension :** Camera intersection plane with the surface<br>**Position :** Right next to the laptop<br>**Orientation :** Paralell to surface<br>**Material :** Invisible | Laptop | Laptop | **Input :** Hand position ; Index tap ; Middle tap ; All fingers scroll ; Pinch ; Pickup<br>**Output :** Cursor moving ; Right click ; Left click ; Scroll ; Zoom in and out ; Drag and drop<br>**Mapping :** Continuous/ relative ; Discrete/ Absolute ; same ; Continuous/ relative ; same ; same |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [56] | **Shape :** Plane<br>**Dimension :** 40 cm x 36 xm<br>**Position :** Next to keyboard<br>**Orientation :** On surface aligned with desk's edges<br>**Material :** Invisible | Desktop computer | Desk | **Input :** Hand position ; Tap<br>**Output :** Item selection ; Trigger command<br>**Mapping :** Continuous/ Absolute ; Discrete/ Absolute |
| [88] | **Shape :** Sensing volume<br>**Dimension :** Sensing volume<br>**Position :** Phone position<br>**Orientation :** Sensors orientation<br>**Material :** Invisible | Any compatible device | Smartphone | **Input :** Tangible gestures ; Midair gestures ; Touch gestures<br>**Output :** Associated command<br>**Mapping :** Any |
| [2] | **Shape :** Planes<br>**Dimension :** Camera viewing plane on surface<br>**Position :** Around the tablet<br>**Orientation :** Surface<br>**Material :** Invisible | Tablet | Tablet | **Input :** Objects manipulation<br>**Output :** Associated Command<br>**Mapping :** Any |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [121] | **Shape :** Planes<br>**Dimension :** Arm reach<br>**Position :** Phone position at start<br>**Orientation :** Phone orientation at start<br>**Material :** Displayed Content | Smartphone | Smartphone | **Input :** Smartphone position and orientation ; Pen position ; Pen tap<br>**Output :** Creation of the plane ; Standard touch action ; Tap<br>**Mapping :** Discrete/ absolute ; Continuous/ Relative ; Discrete/ Absolute |
| [40] | **Shape :** Plane<br>**Dimension :** Poster dimension<br>**Position :** Poster position<br>**Orientation :** Poster orientation<br>**Material :** Poster content + additional digital content | Itself (poster content) | Poster | **Input :** Smartphone position ; Standard touch interaction<br>**Output :** Displays associated content ; Standard touch interaction<br>**Mapping :** Continuous/ Absolute ; Discrete/Continuous/ Absolute/Relative |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [78] | **Shape :** Plane<br>**Dimension :** No more than 5 times screen dimension<br>**Position :** Centered on the screen<br>**Orientation :** Vertical<br>**Material :** Displayed content | Itself (displayed content) | Display | **Input :** Pointing ; Grab<br>**Output :** Moving cursor ; Click<br>**Mapping :** Continuous/ Absolute ; Discrete/ Absolute |
| [108] | **Shape :** Plane<br>**Dimension :** Content wise<br>**Position :** Smartphone position at the start<br>**Orientation :** Smartphone Orientation<br>**Material :** Displayed content | Itself | Smartphone | **Input :** Smartphone position ; Smartphone orientation ; Finger down<br>**Output :** Pan and zoom of MIS ; Moves MIS orientation ; Clutch<br>**Mapping :** Conitnuous/ relative ; same ; Discrete/ relative |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [64] | **Shape :** Plane<br>**Dimension :** Defined by designer<br>**Position :** Projection wall position<br>**Orientation :** Vertical<br>**Material :** Displayed content | Itself | Wall | **Input :** Smartphone pointing position ; Tap<br>**Output :** Displays associated content/ draws a stroke ; Triggers command<br>**Mapping :** Continuous/ absolute; Discrete/ relative; |
| [69] | **Shape :** Cube ; Cube<br>**Dimension :** Sensor volume (4mx3.6mx2.7m) ; Tablet dimension<br>**Position :** World fixed (sensor wise) ; Above Tablet<br>**Orientation :** Sensor wise ; Tablet orientation<br>**Material :** Displayed content (3D objects) ; Invisible | Itself ; Tablet | World fixed ; Tablet | **Input :** Behind: middle pinch ; Behind: index pinch ; Behind: ring pinch; Behind: hand position ; Above: Index pinch ; Above: Hand z-axis position ; Above: Hand position x-axis<br>**Output :** Create object ; Manipulate object ; Delete Object ; Select Object or move it; Adjust time granularity ; Time cursor position<br>**Mapping :** Discrete/ Absolute ; same ; same ; Continuous/ Absolute ; Continuous/ relative ; Continuous/ relative |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [52] | **Shape :** Forearm/Palm Surface<br>**Dimension :** Forearm/Palm Surface<br>**Position :** Forearm/Palm<br>**Orientation :** Forearm Orientation<br>**Material :** UI | Compatible device | Forearm | **Input :** Tap ; Swipe ; Stroke<br>**Output :** Associated commands<br>**Mapping :** Discrete/ Absolute |
| [51] | **Shape :** Forearm/Palm Surface<br>**Dimension :** Forearm/Palm Surface<br>**Position :** Forearm/Palm<br>**Orientation :** Forearm Orientation<br>**Material :** UI | Compatible device | Forearm | **Input :** Tap ; Swipe ; Stroke<br>**Output :** Associated commands<br>**Mapping :** Discrete/ Absolute |
| [49] | **Shape :** Forearm/Palm Surface<br>**Dimension :** Forearm/Palm Surface<br>**Position :** Forearm/Palm<br>**Orientation :** Forearm Orientation<br>**Material :** UI | Compatible device | Forearm | **Input :** Tap ; Swipe ; Stroke<br>**Output :** Associated commands<br>**Mapping :** Discrete/ Absolute |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [18] | **Shape :** Cylinder<br>**Dimension :** Radius = Distance between user and projection wall<br>**Position :** User centered and distance to projection wall<br>**Orientation :** Around user's body<br>**Material :** Workspace | Mobile phone | User's body | **Input :** Angle on the cylinder<br>**Output :** Workspace switching<br>**Mapping :** Discrete/ Absolute |
| [18] | **Shape :** Tore ; Cylinder<br>**Dimension :** Arms reach<br>**Position :** User centered<br>**Orientation :** Around user ; vertical<br>**Material :** Invisible ; Workspaces | Mobile phone ; Itself | User's body | **Input :** Angle on the tore ; Angle on the cylinder<br>**Output :** Displays workspace ; Displays other workspaces<br>**Mapping :** Discrete/ Absolute ; Discrete/ Absolute |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [31] | **Shape :** Sphere ; Planes<br>**Dimension :** radius = 50 cm ; 22cm<br>**Position :** User centered ; User defined<br>**Orientation :** Around user ; User defined on the sphere<br>**Material :** MIS-Windows ; Content | Itself | User's body/ palm or world fixed ; Sphere | **Input :** Sphere:Tap on HWD ; Windows : Grab an app ; Windows: Pinch ; Windows : Tap;<br>**Output :** Sphere : Change IRF ; Windows : Create new window of app ; Windows : resize window and rearrange ; Windows : Tap<br>**Mapping :** Discrete/ Absolute ; Discrete and Absolute ; Continuous/ Absolute; Discrete/ Absolute |
| [29] | **Shape :** Plane<br>**Dimension :** User defined<br>**Position :** User defined<br>**Orientation :** User defined<br>**Material :** Displayed content | Itself | User'body or world fixed | **Input :** Tangible or not ; on-body or far ;<br>**Output :** Associated command<br>**Mapping :** Discrete or Continuous/ Absolute or Relative |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [41] | **Shape :** Plane<br>**Dimension :** Designer defined<br>**Position :** Smartwatch centered<br>**Orientation :** Smartwatch Orientation<br>**Material :** UI elements | Itself and smartwatch | Smartwatch | **Input :** Smartwatch position ; Touch interaction<br>**Output :** Selection ; Associated command<br>**Mapping :** Continuous/ Absolute or relative ; Discrete or continuous/ Absolute or relative |
| [102] | **Shape :** Plane<br>**Dimension :** Sheet of paper<br>**Position :** Sheet of paper<br>**Orientation :** Sheet of paper<br>**Material :** Both augmented content and physical content | Itself and sheet of paper | Sheet of paper | **Input :** MouseLight position ; Pen tap ; Pen stroke ;<br>**Output :** Change displayed content ; Click for selection or copy paste ; Associated command (e.g. search)<br>**Mapping :** Continuous/ Absolute ; Discrete/ Absolute or relative |
| [100] | **Shape :** Planes<br>**Dimension :** User's defined<br>**Position :** User's defined<br>**Orientation :** User's defines<br>**Material :** Display | Itself | World fixe/ view fixed | **Input :** N/A (not enough information)<br>**Output :** N/A (not enough information)<br>**Mapping :** N/A (not enough information) |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [118] | **Shape :** Plane<br>**Dimension :** 40 inch diagonal<br>**Position :** Projector position<br>**Orientation :** Horizontal (on a table)<br>**Material :** UI/ tangible content/ augmented content | Itself | Projector | **Input :** Tangible position ; Tangible orientation ; Fingers positions ; Zoom gesture ;<br>**Output :** Content position ; Content orientation ; Pan ; Zoom<br>**Mapping :** Continuous/ Absolute ; same ; same ; same |
| [61] | **Shape :** Plane<br>**Dimension :** Viewing plane<br>**Position :** Sensor wise<br>**Orientation :** Horizontal<br>**Material :** UI/ digital content | Itself and Laptop | Laptop | **Input :** Put tangible ; Touch gestures<br>**Output :** Augmented UI feedback/ associated command ; standard touch actions<br>**Mapping :** Discrete/ Absolute; Continuous/ Absolute; |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [13] | **Shape :** 4 Planes<br>**Dimension :** 44cmx33cm ; 44cmx33cm ; 44cmx17cm ; 44cmx33cm<br>**Position :** Top ; Right ; Bottom ; Left<br>**Orientation :** Horizontal<br>**Material :** UI elements | Desktop/ itself | Desktop | **Input :** Multitouch gestures ; Mouse position ; Keyboard position<br>**Output :** Standard multitouch commands ; Contextual menu Position ; Adapt MIS arrangement : away from body = enlarge top MIS/ putting back = default arrangement<br>**Mapping :** Continuous or discrete/ Absolute or relative; Continuous/ Absolute ; Discrete/ Absolute |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [119] | **Shape :** Planes ; Cubes ; Circles **Dimension :** Designer defined **Position :** Designer defined ; same ; User's hand **Orientation :** Suface orientation ; Designer defined ; Horizontal **Material :** Displayed Content ; Menu items ; Colored and contained | Themselves | World fixed ; world fixed ; User's body | **Input :** Planes : Multitouch gestures ; Planes : touching two planes ; Planes : drag content out of it in hand ; Cube : Hand position vertical axis ; Cube : 2 seconds dwell ; Circle : Hand position in space ; Circle : Place hand on plane **Output :** Planes : standard commands ; Planes : Transferring content from one plane to other ; Planes:create new Circle MIS encapsulating the content ; Cube : displays associated menu item ; Cube : select menu item ; Circle : move in space; Circle : delete MIS and transfer encapsulated content to the plane **Mapping :** Any ; Discrete/ Absolute; same; same ; same; continuous/ absolute; Discrete/ absolute |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [42] | **Shape :** Planes<br>**Dimension :** User defined (depending on system position and projection surface)<br>**Position :** User defined (by placing the system)<br>**Orientation :** User defined (by placing the system)<br>**Material :** Widgets | itself | projector-camera System | **Input :** touch down ; long touch ; move ; touch release<br>**Output :** Standard touch commands<br>**Mapping :** Discrete/ Absolute ; same; continuous/ absolute ; Discrete/ Absolute |
| [66] | **Shape :** Planes<br>**Dimension :** Palm size<br>**Position :** User or designer defined<br>**Orientation :** Palm orientation<br>**Material :** Displayed content | Associated device | World fixed/ user's body | **Input :** Hand position ; Touch a device<br>**Output :** Change contextual data on palm ; Store or transfer data to/from body<br>**Mapping :** Discrete/ absolute |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [98] | **Shape :** Cube ; Plane ; Planes **Dimension :** Tabletop dimension and around 1m high ; Pad dimension ; Pad dimension **Position :** Tabletop ; Pad position ; user defined position **Orientation :** Horizontal ; Pad orientation ; User defined **Material :** Invisible ; UI elements/ scene/ modified scene ; Scene | Tabletop and pad ; Tabletop/ itslef ; Itself | Tabletop ; Pad ; Tabletop | **Input :** Cube : sweeping path ; Pad : Pad orientation + position ; Pad : pen stroke; Pad : pen tap on UI ; Pad : Pen input on objects ; Pad : turning over ; Pad : decouple snapshot ; Pad : eye position; Window : recouple snapshot; **Output :** Cube: select scene objects on the pad ; Pad : change view on pad ; Pad : Select objects in the scene ; Pad : associated command ; Pad: change objects properties ; Pad : XRay view of the scene ; Pad: create new MIS as a new viewport on the scene from a snapshot taken before(Window) ; Pad : change scene view on the pad; Window: fuse window to pad ; **Mapping :** Discrete/ Absolute ; Continuous/ Absolute ; Discrete/ relative ; Discrete/ absolute; Any ; Discrete/ Absolute; Discrete/ absolute; Continuous/ Absolute; Discrete/ Absolute |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [106] | **Shape :** Cube ;<br>**Dimension :** Tabletop dimensions x 1m high<br>**Position :** Tabletop position<br>**Orientation :** Horizontal<br>**Material :** Information to be explored | Tabletop/ paper sheet | Tabletop | **Input :** Paper sheet Orientation ; Paper sheet XY-position ; Papersheet Z-position<br>**Output :** associated slice for volumetric space; Associated data of the information space ; associated keyframe for temporal space/ associated layer for layered space/ associated data for volumetric space<br>**Mapping :** Continuous/ Absolute; continuous or discrete/ absolute; same |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [104] | **Shape :** Cube ; Planes<br>**Dimension :** Tabletop ; Display<br>**Position :** Tabletop position; Display position<br>**Orientation :** Horizontal ; User defined display orientation<br>**Material :** Invisible ; Scene viewport | Tabletop (global scene) ; Itself | Tabletop ; Display | **Input :** Cube: Display position and rotation ; Plane : Slight press ; Plane : hard press on object ; Plane hard press on scene background ; Plane : double tap on object; Plane : double tap on scene background ; Plane : Orientation ;<br>**Output :** Cube: Change viewport on display for WIM and moving mode or change object position in object manipulation; Plane: display the selection stick or active clutching if in fishtank or WIM mode; Plane: create a ghost of the object to move or copy; Plane: pan scene mode; Plane: fishtank mode of object ; Plane: WIM mode; Plane:Change object orientation on the display in fishtank mode;<br>**Mapping :** Continuous/ absolute or relative if clutch active; |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [57] | **Shape :** Cube<br>**Dimension :** Tabletop<br>**Position :** Above tabletop<br>**Orientation :** Horizontal<br>**Material :** Invisible | Tabletop | Tabletop | **Input :** Pinch gesture; Hand position<br>**Output :** Grab an object; Move the virtual hand in the scene<br>**Mapping :** Discrete/ absolute ; Continuous/ Absolute |
| [85] | **Shape :** Cube<br>**Dimension :** Tabletop<br>**Position :** Above tabletop<br>**Orientation :** Horizontal<br>**Material :** Invisible | Tabletop | Tabletop | **Input :** Pen position and orientation above surface; Pen tabletop touch ; Pen button click while pointing ;<br>**Output :** Change pointer position on tabletop ; Click ; Click<br>**Mapping :** Continuous/ Absolute ; Discrete/ absolute ; Discrete/ Absolute |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [11] | **Shape :** Cube<br>**Dimension :** Tabletop<br>**Position :** Above tabletop<br>**Orientation :** Horizontal<br>**Material :** Inspected 3D objects | Tabletop and itself | Tabletop | **Input :** Grab from tabletop to space ; Flat hand above 3D object pushing towards tabletop ; Grab from tabletop to space then tap while grabbing ; Hand position and orientation in space while grabbing ; Placing 3D Object close to the tabletop and tapping ; Placing a vertical hand on the tabletop<br>**Output :** Pull out 3D object in space ; Push 3D object back to 2d on the tabletop ; Bind 2D and 3D object ; Move and rotate object ; Pin 3D object to the tapped position (trackball style touch rotation is then available) ; Activate privacy mode for previous gestures<br>**Mapping :** Discrete/ Absolute ; same; same ; same ; continuous/ absolute ; discrete/ absolute; Discrete Absolute |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [112] | **Shape :** Cube ; Cube; Cube ; Cube<br>**Dimension :** Display dismension x 60 cm depth<br>**Position :** In front of the display ; 60 cm away from the display ; 120 cm away from the display ;180 cm away from the display<br>**Orientation :** Veritcal<br>**Material :** Invisible | Display | Display | **Input :** Personal : Entering ; Personal : touch event calendar; Personal : Exiting; Subtle Interaction : Entering ; Subtle Interaction : Exiting ; Subtle Interaction : Move along the screen ; Subtle Interaction : Palm Up an upward flick ; Subtle Interaction : Palm Vertical continuous horizontal gesture ; Subtle Interaction: Palm Down continuous vertical gesture ; Subtle Interaction : Palm down a downward flick gesture ; Subtle Interaction: Palm down Left or right flick ; Implicit Interaction : Entering ; Implicit Interaction : Implicit Interaction : user's body location ; Implicit Interaction : body orientation ; Implicit Interaction : head orientation ; Implicit Interaction : Palm Away posture ; Implicit Interaction : Palm Facing posture ; Ambient Interaction : user's distance to screen ;<br>**Output :** Personal : display with small font personal data ; Personal : open event; Personal : Hide personal data ; Subtle Interaction : Display personal and shared data ; Subtle Interaction : hide data ; Subtle Interaction : navigate information ; Subtle Interaction : returns to the OVERVIEW; Subtle Interaction : adjusts the selection point's position ; Subtle Interaction: highlight items; Subtle Interaction: select item ; Subtle Interaction: cancel ; Implicit Inter- |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [122] | **Shape :** Cube ; Cube ; **Dimension :** Display x 1.2m depth ; Display x 2.3m depth **Position :** Front of the display ; 1.2m away from display **Orientation :** Vertical **Material :** Invisible | Display | Display | **Input :** Near space : Standard touch gestures ; Far space : Pick gesture : Far space : dwell ; Far space : Wave gesture ; Far space : Hand position; **Output :** Near space : Associated commands ; Far space : display large-sized menu and activate hand shaped cursor; Far space : click; Far space : 3D navigation mode ; Far space : controls either hand shaped cursor or navigation; **Mapping :** Discrete or continuous/ absolute; Discrete/ absolute; Discrete/ absolute; Discrete/ absolute; Continuous/ absolute; |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [26] | **Shape :** 4 Zones<br>**Dimension :** 1. Touch zone: Direct screen interaction. 2. Fine-grained gesture zone: up to 0.5 m in front of screen. 3. General gesture zone: between 0.5 m and 2 m. 4. Coarse gesture zone: more than 2 m afar<br>**Position :** 1. Touch zone: Direct screen interaction. 2. Fine-grained gesture zone: up to 0.5 m in front of screen. 3. General gesture zone: between 0.5 m and 2 m. 4. Coarse gesture zone: more than 2 m afar<br>**Orientation :** N/A<br>**Material :** Invisible | Display | Display | **Input :** N/A (not enough information)<br>**Output :** N/A<br>**Mapping :** N/A |

Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [36] | **Shape :** Cube<br>**Dimension :** ActiveDesk dimension<br>**Position :** ActiveDesk position<br>**Orientation :** ActiveDesk orientation<br>**Material :** Invisible | ActiveDesk | ActiveDesk | **Input :** 2D pen : stroke+draw button pressed; 2D pen : stroke+manipulation button pressed; 2D pen: stroke+camera button pressed; 3D input: pickup prop; 3D input: translating and rotating prop ; Trackball : input ; Trackball : input + 2D pen ; Speech input : command alone ; Speech input : in context;<br>**Output :** 2D pen : draw; 2D pen : move object; 2D pen: move camera; 3D input: switch to stereoscopic view + attached object to end of prop; 3D input: translate and rotate object ; Trackball : Camera control ;Trackball : Camera control with respect to the pen; Speech input : said command action ; Speech input : same;<br>**Mapping :** Continuous/ Absolute; same; same; Discrete/ Absolute; continuous/ Absolute; same; same; Discrete/ Absolute; same; |



Thèse de Hanaë Rateau, Lille 1, 2017

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [115] | **Shape :** Cube<br>**Dimension :** Table<br>**Position :** Table<br>**Orientation :** Horizontal<br>**Material :** Invisible | Table | Table | **Input :** Pen stroke ; Pointing a a surface; Marking menu stroke;<br>**Output :** Draw curve/ Change an existing one/ Soothing/ sharpening/ sculpting; Fill surface; Choose tool;<br>**Mapping :** Continuous/ Absolute; Discrete/ Absolute; Discrete/ Absolute; |
| [97] | **Shape :** Cube<br>**Dimension :** Workbench Dimension<br>**Position :** Workbench Position<br>**Orientation :** Horizontal<br>**Material :** Invisible | Workbench | Workbench | **Input :** Hand stroke ; Kitchen tongs press+move+release ; Eraser movement ; Magnet movement;<br>**Output :** Draw surface according hand posture along the stroke; Grab/ move or rotate and release ; Remove a region of the drawing ; Attracts the nearest surface to modify it;<br>**Mapping :** Continuous/ Absolute; |

| Ref. | Geometric Definition | Input Reference Frame | Action Reference Frame | Interaction Attributes |
|---|---|---|---|---|
| [24] | **Shape :** Cube<br>**Dimension :** Tabletop Dimension x 1m high<br>**Position :** Tabletop Position<br>**Orientation :** Horizontal<br>**Material :** Invisible | Tabletop | Tabletop | **Input :** NDH Hover ; Pinch object in space / Touch object an surface; DH stroke while pinching ; NDH pinch + DH pinch + Hand movements; DH pinch on surface + DH stroke above surface;<br>**Output :** Display contextual menu ; Select object ; Draw curve or surface; scale/ translate/ rotate; Extrusion in extrusion mode;<br>**Mapping :** Discrete/ Absolute; Discrete/ Absolute; Continuous/ Absolute; Continuous/ Absolute |



Thèse de Hanaë Rateau, Lille 1, 2017

© 2017 Tous droits réservés.

lilliad.univ-lille.fr

# Bibliography

[1] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, and J. T. K. V. Koh. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for hci. Technical Report MSR-TR-2012-11, Redmond, WA, USA, Nov 2012.

[2] D. Avrahami, J. O. Wobbrock, and S. Izadi. Portico: Tangible interaction on and around a tablet. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 347–356, New York, NY, USA, 2011. ACM.

[3] G. Bailly, E. Lecolinet, and Y. Guiard. Finger-count & radial-stroke shortcuts: 2 techniques for augmenting linear menus on multi-touch surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 591–594, New York, NY, USA, 2010. ACM.

[4] G. Bailly, J. Müller, M. Rohs, D. Wigdor, and S. 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.

[5] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: Designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 121–130, New York, NY, USA, 2010. ACM.

[6] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: Designing for a proximity and orientation-aware environment. In *Proc. of ITS*, 2010.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017




[7] L. Balme. *Interfaces homme-machine plastiques : une approche par composants dynamiques*. Theses, Université Joseph-Fourier - Grenoble I, June 2008.

[8] A. Banerjee, J. Burstyn, A. Girouard, and R. Vertegaal. Pointable: An in-air pointing technique to manipulate out-of-reach targets on tabletops. In *Proc. of ITS*, 2011.

[9] O. Bau and W. E. Mackay. Octopocus: A dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, pages 37–46, New York, NY, USA, 2008. ACM.

[10] P. Baudisch, E. Cutrell, D. Robbins, M. Czerwinski, P. Tandler, B. Bederson, A. Zierlinger, et al. Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch-and pen-operated systems. In *Proc. of INTERACT*, 2003.

[11] H. Benko, E. W. Ishak, and S. Feiner. Cross-dimensional gestural interaction techniques for hybrid immersive environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 209–216, March 2005.

[12] L.-P. Bergé, E. Dubois, and M. Raynal. Design and evaluation of an "around the smartphone" technique for 3d manipulations on distant display. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, SUI '15, pages 69–78, New York, NY, USA, 2015. ACM.

[13] X. Bi, T. Grossman, J. Matejka, and G. Fitzmaurice. Magic desk: Bringing multi-touch surfaces into desktop work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2511–2520, New York, NY, USA, 2011. ACM.

[14] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: The see-through interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pages 73–80, New York, NY, USA, 1993. ACM.

[15] A. Butler, S. Izadi, and S. Hodges. Sidesight: Multi-"touch" interaction around small devices. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, UIST '08, pages 201–204, New York, NY, USA, 2008. ACM.

[16] C. Cadoz. Le geste canal de communication homme/machine: la communication instrumentale. *TSI. Technique et science informatiques*, 13(1):31–61, 1994.




© 2017 Tous droits réservés.

lilliad.univ-lille.fr

[17] F. Camp, A. Schick, and R. Stiefelhagen. How to click in mid-air. In *Proc. of HCII 2013*, pages 78–86, July 2013.

[18] J. Cauchard, M. Löchtefeld, M. Fraser, A. Krüger, and S. Subramanian. M+pspaces: Virtual workspaces in the spatially-aware mobile environment. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 171–180, New York, NY, USA, 2012. ACM.

[19] X. A. Chen, N. Marquardt, A. Tang, S. Boring, and S. Greenberg. Extending a mobile device's interaction space through body-centric interaction. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '12, pages 151–160, New York, NY, USA, 2012. ACM.

[20] X. A. Chen, J. Schwarz, C. Harrison, J. Mankoff, and S. E. Hudson. Air+touch: Interweaving touch & in-air gestures. In *Proc. of UIST*, 2014.

[21] A. Cockburn, P. Quinn, C. Gutwin, G. Ramos, and J. Looser. Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *Int. J. Hum.-Comput. Stud.*, 69(6):401–414, June 2011.

[22] D. Coffey, N. Malbraaten, T. B. Le, I. Borazjani, F. Sotiropoulos, A. Erdman, and D. F. Keefe. Interactive slice wim: Navigating and interrogating volume data sets using a multisurface, multitouch vr interface. *IEEE TVCG*, 18(10):1614–1626, Oct 2012.

[23] M. Collomb, M. Hascoët, P. Baudisch, and B. Lee. Improving drag-and-drop on wall-size displays. In *Proc. of GI*, 2005.

[24] B. De Araujo, G. Casiez, J. Jorge, and M. Hachet. Mockup Builder: 3D Modeling On and Above the Surface. *Computers & Graphics*, 37(3), 2013.

[25] N. Dezfuli, M. Khalilbeigi, J. Huber, F. Müller, and M. Mühlhäuser. Leveraging the palm surface as an eyes-free tv remote control. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, pages 2483–2488, New York, NY, USA, 2012. ACM.

[26] T. Dingler, M. Funk, and F. Alt. Interaction proxemics: Combining physical spaces for seamless gesture interaction. In *Proceedings of the 4th International Symposium on Pervasive Displays*, PerDis '15, pages 107–114, New York, NY, USA, 2015. ACM.

[27] P. Ekman and W. V. Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1):49–98, 1969.




© 2017 Tous droits réservés. 

[28] B. Ens, D. Ahlström, A. Cockburn, and P. Irani. Characterizing user performance with assisted direct off-screen pointing. MobileHCI '11, 2011.

[29] B. Ens, J. D. Hincapié-Ramos, and P. Irani. Ethereal planes: A design framework for 2d information space in 3d mixed reality environments. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, SUI '14, pages 2–12, New York, NY, USA, 2014. ACM.

[30] B. Ens, E. Ofek, N. Bruce, and P. Irani. Spatial constancy of surface-embedded layouts across multiple environments. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, SUI '15, pages 65–68, New York, NY, USA, 2015. ACM.

[31] B. M. Ens, R. Finnegan, and P. P. Irani. The personal cockpit: A spatial interface for effective task switching on head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 3171–3180, New York, NY, USA, 2014. ACM.

[32] S. Feiner, B. MacIntyre, M. Haupt, and E. Solomon. Windows on the world: 2d windows for 3d augmented reality. In *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, UIST '93, pages 145–155, New York, NY, USA, 1993. ACM.

[33] G. W. Fitzmaurice. Situated information spaces and spatially aware palmtop computers. *Commun. ACM*, 36(7):39–49, July 1993.

[34] G. W. Fitzmaurice, S. Zhai, and M. H. Chignell. Virtual reality for palmtop computers. *ACM Trans. Inf. Syst.*, 11(3):197–218, July 1993.

[35] C. Forlines, D. Vogel, and R. Balakrishnan. Hybridpointing: Fluid switching between absolute and relative pointing with a direct input device. In *Proc. of UIST*, 2006.

[36] A. S. Forsberg, J. J. L. Jr., and R. C. Zeleznik. Ergodesk: A framework for two- and three-dimensional interaction at the activedesk. In *In Proceedings of the Second International Immersive Projection Technology Workshop*, pages 11–12, 1998.

[37] J. Gilliot, G. Casiez, and N. Roussel. Direct and Indirect Multi-Touch Interaction on a Wall Display. In *Proc. of IHM*, 2014.

[38] A. Goguey, G. Casiez, T. Pietrzak, D. Vogel, and N. Roussel. Adoiraccourcix: Multi-touch command selection using finger identification. In *Proceedings of the 26th Conference on L'Interaction Homme-Machine*, IHM '14, pages 28–37, New York, NY, USA, 2014. ACM.




© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017




[39] S. Greenberg, N. Marquardt, T. Ballendat, R. Diaz-Marino, and M. Wang. Proxemic interactions: The new ubicomp? *interactions*, 18(1):42–50, Jan. 2011.

[40] J. Grubert, R. Grasset, and G. Reitmayr. Exploring the design of hybrid interfaces for augmented posters in public spaces. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, NordiCHI '12, pages 238–246, New York, NY, USA, 2012. ACM.

[41] J. Grubert, M. Heinisch, A. Quigley, and D. Schmalstieg. Multifi: Multi fidelity interaction with displays on and around the body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3933–3942, New York, NY, USA, 2015. ACM.

[42] J. Gugenheimer, P. Knierim, J. Seifert, and E. Rukzio. Ubibeam: An interactive projector-camera system for domestic deployment. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ITS '14, pages 305–310, New York, NY, USA, 2014. ACM.

[43] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model, 1987.

[44] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, 19(4):486–517, 1987.

[45] S. Gustafson, P. Baudisch, C. Gutwin, and P. Irani. Wedge: Clutter-free visualization of off-screen locations. CHI '08, 2008.

[46] S. Gustafson, D. Bierwirth, and P. Baudisch. Imaginary interfaces: spatial interaction with empty hands and without visual feedback. In *Proceedings of UIST '10*, pages 3–12, NY, USA, 2010. ACM.

[47] S. Gustafson, C. Holz, and P. Baudisch. Imaginary phone: Learning imaginary interfaces by transferring spatial memory from a familiar device. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 283–292, New York, NY, USA, 2011. ACM.

[48] E. T. Hall. The hidden dimension . 1966.

[49] C. Harrison, H. Benko, and A. D. Wilson. Omnitouch: Wearable multitouch interaction everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 441–450, New York, NY, USA, 2011. ACM.




© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



[50] C. Harrison and S. E. Hudson. Abracadabra: Wireless, high-precision, and unpowered finger input for very small mobile devices. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pages 121–124, New York, NY, USA, 2009. ACM.

[51] C. Harrison, S. Ramamurthy, and S. E. Hudson. On-body interaction: Armed and dangerous. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, TEI '12, pages 69–76, New York, NY, USA, 2012. ACM.

[52] C. Harrison, D. Tan, and D. 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.

[53] K. Hasan, D. Ahlström, and P. Irani. Ad-binning: Leveraging around device space for storing, browsing and retrieving mobile device content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 899–908, New York, NY, USA, 2013. ACM.

[54] K. Hasan, D. Ahlström, and P. P. Irani. Comparing direct off-screen pointing, peephole, and flick & pinch interaction for map navigation. SUI '15, 2015.

[55] M. Hascoët. Throwing Models for Large Displays. In *Proc. of HCI*, 2003.

[56] D. Hausen, S. Boring, and S. Greenberg. The unadorned desk: Exploiting the physical space around a display as an input canvas. In *Proc. of Interact 2013*, Cape Town, South Africa, Sep 2013.

[57] O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza, and A. Butz. Interactions in the air: adding further depth to interactive tabletops. In *Proc. of UIST '09*, pages 139–148, NY, USA, 2009. ACM.

[58] S. C. Holmes NP. The body schema and the multisensory representation(s) of peripersonal space. (2015SACLS207), 2015.

[59] B. Jones, R. Sodhi, D. Forsyth, B. Bailey, and G. 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.

[60] R. Jota, P. Lopes, D. Wigdor, and J. Jorge. Let's kick it: How to stop wasting the bottom third of your large screen display. In *Proc. of CHI*, 2014.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

[61] S. K. Kane, D. Avrahami, J. O. Wobbrock, B. Harrison, A. D. Rea, M. Philipose, and A. LaMarca. Bonfire: A nomadic system for hybrid laptop-tabletop interaction. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pages 129–138, New York, NY, USA, 2009. ACM.

[62] M. Karam and m. c. schraefel. A taxonomy of gestures in human computer interactions. Technical report, University of Southampton, 2005.

[63] R. S. Kattinakere, T. Grossman, and S. Subramanian. Modeling steering within above-the-surface interaction layers. In *Proceedings of CHI '07*, pages 317–326, New York, NY, USA, 2007. ACM.

[64] B. Kaufmann and M. Hitz. X-large virtual workspaces for projector phones through peephole interaction. In *Proceedings of the 20th ACM International Conference on Multimedia*, MM '12, pages 1279–1280, New York, NY, USA, 2012. ACM.

[65] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier. Digits: Freehand 3d interactions anywhere using a wrist-worn gloveless sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, UIST '12, pages 167–176, New York, NY, USA, 2012. ACM.

[66] S. Kim, S. Takahashi, and J. Tanaka. A location-sensitive visual interface on the palm: Interacting with common objects in an augmented space. *Personal Ubiquitous Comput.*, 19(1):175–187, Jan. 2015.

[67] S. Kratz, M. Rohs, D. Guse, J. Müller, G. Bailly, and M. 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.

[68] W. Krüger, C.-A. Bohn, B. Fröhlich, H. Schüth, W. Strauss, and G. Wesche. The responsive workbench: A virtual work environment. *Computer*, 28(7):42–48, July 1995.

[69] D. Lakatos, M. Blackshaw, A. Olwal, Z. Barryte, K. Perlin, and H. Ishii. T(ether): Spatially-aware handhelds, gestures and proprioception for multi-user 3d modeling and animation. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, SUI '14, pages 90–93, New York, NY, USA, 2014. ACM.




© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



[70] D. Ledo, S. Greenberg, N. Marquardt, and S. Boring. Proxemic-aware controls: Designing remote controls for ubiquitous computing ecologies. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, pages 187–198, New York, NY, USA, 2015. ACM.

[71] F. C. Y. Li, D. Dearman, and K. N. Truong. Virtual shelves: Interactions with orientation aware devices. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pages 125–128, New York, NY, USA, 2009. ACM.

[72] S.-Y. Lin, C.-K. Shie, S.-C. Chen, and Y.-P. Hung. Airtouch panel: A re-anchorable virtual touch panel. In *Proceedings of ACM Multimedia 2013 (ACM MM)*, pages 625–628. ACM, october 2013.

[73] S.-Y. Lin, C.-H. Su, K.-Y. Cheng, R.-H. Liang, T.-H. Kuo, and B.-Y. Chen. Pub - point upon body: Exploring eyes-free interaction and methods on an arm. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 481–488, New York, NY, USA, 2011. ACM.

[74] C. Liu. *Embodied Interaction for Data Manipulation Tasks on Wall-sized Displays*. Theses, Université Paris-Saclay, Dec. 2015.

[75] M. Liu, M. Nancel, and D. Vogel. Gunslinger: Subtle arms-down mid-air interaction. In *Proc. of UIST*, 2015.

[76] Z. Lv, A. Halawani, M. S. Lal Khan, S. U. Réhman, and H. Li. Finger in air: Touch-less interaction on smartphone. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*, MUM '13, pages 16:1–16:4, New York, NY, USA, 2013. ACM.

[77] S. Marcos, H. Khalad, E. Barrett, X.-D. Yang, and I. Pourang. Smartwatches + head-worn displays: the 'new' smartphone. In *Workshop on Mobile Collocated Interactions: From Smartphones to Wearables*, CHI'15, pages 65–68, New York, NY, USA, 2015. ACM.

[78] A. Markussen, S. Boring, M. R. Jakobsen, and K. Hornbæk. Off-limits: Interacting beyond the boundaries of large displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5862–5873, New York, NY, USA, 2016. ACM.

[79] N. Marquardt, R. Jota, S. Greenberg, and J. A. Jorge. The continuous interaction space: Interaction techniques unifying touch and gesture on and above a digital surface. In *Proc. of INTERACT*, 2011.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

[80] P. Mistry and P. Maes. Mouseless: A computer mouse as small as invisible. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1099–1104, New York, NY, USA, 2011. ACM.

[81] P. Mistry, P. Maes, and L. Chang. Wuw - wear ur world: A wearable gestural interface. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, pages 4111–4116, New York, NY, USA, 2009. ACM.

[82] M. Morris, J. Wobbrock, and A. Wilson. Understanding users' preferences for surface gestures. In *Proceedings of GI '10*, pages 261–268, Toronto, Canada, 2010.

[83] M. Nancel, O. Chapuis, E. Pietriga, X.-D. Yang, P. P. Irani, and M. Beaudouin-Lafon. High-precision pointing on large wall displays using small handheld devices. In *Proc. of CHI*, 2013.

[84] M. Nielsen, T. Moeslund, M. Störring, and E. Granum. A procedure for developing intuitive and ergonomic gesture interfaces for hci. In *Proc. of the 5th Internation Gesture Workshop*, GW 2003, 2003.

[85] J. K. Parker, R. L. Mandryk, and K. M. Inkpen. Integrating point and touch for interaction with digital tabletop displays. *Proc. of CGA*, 2006.

[86] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko. It's mine, don't touch!: Interactions at a large multi-touch display in a city centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1285–1294, New York, NY, USA, 2008. ACM.

[87] K. Perlin and D. Fox. Pad: An alternative approach to the computer interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pages 57–64, New York, NY, USA, 1993. ACM.

[88] H. Pohl and M. Rohs. Around-device devices: My coffee mug is a volume dial. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, MobileHCI '14, pages 81–90, New York, NY, USA, 2014. ACM.

[89] U. Rashid, M. A. Nacenta, and A. Quigley. The cost of display switching: A comparison of mobile, large display and hybrid ui configurations. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, pages 99–106, New York, NY, USA, 2012. ACM.



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



[90] H. Rateau, L. Grisoni, and B. Araujo. Exploring tablet surrounding interaction spaces for medical imaging. In *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, SUI '14, pages 150–150, New York, NY, USA, 2014. ACM.

[91] H. Rateau, L. Grisoni, and B. De Araujo. Mimetic interaction spaces: Controlling distant displays in pervasive environments. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*, IUI '14, pages 89–94, New York, NY, USA, 2014. ACM.

[92] H. Rateau, Y. Rekik, L. Grisoni, and J. Jorge. Talaria: Continuous Drag & Drop on a Wall Display. In *ISS'16*, Niagara Falls, Canada, Nov. 2016.

[93] A. Reetz, C. Gutwin, T. Stach, M. Nacenta, and S. Subramanian. Superflick: A natural and efficient technique for long-distance object placement on digital tables. In *Proc. of GI*, 2006.

[94] J. Rekimoto and M. Saitoh. Augmented surfaces: A spatially continuous work space for hybrid computing environments. In *Proc. of CHI*, 1999.

[95] G. Ren and E. O'Neill. 3d selection with freehand gesture. *Computers & Graphics*, 37(3):101 – 120, 2013.

[96] Y. Rogers and S. Lindley. Collaborating around vertical and horizontal large interactive displays: which way is best? *Interacting with Computers*, 16(6):1133–1152, 2004.

[97] S. Schkolne, M. Pruett, and P. Schröder. Surface drawing: Creating organic 3d shapes with the hand and tangible tools. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 261–268, New York, NY, USA, 2001. ACM.

[98] D. Schmalstieg, L. M. Encarnação, and Z. Szalavári. Using transparent props for interaction with the virtual table. In *Proceedings of the 1999 Symposium on Interactive 3D Graphics*, I3D '99, pages 147–153, New York, NY, USA, 1999. ACM.

[99] C. Schmidt, J. Müller, and G. Bailly. Screenfinity: extending the perception area of content on very large public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1719–1728. ACM, 2013.

[100] M. Serrano, B. Ens, X.-D. Yang, and P. Irani. Desktop-gluey: Augmenting desktop environments with wearable devices. In *Proceedings of the 17th*



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



*International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, MobileHCI '15, pages 1175–1178, New York, NY, USA, 2015. ACM.

[101] M. Serrano, B. Ens, X.-D. Yang, and P. 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.

[102] H. Song, F. Guimbretiere, T. Grossman, and G. Fitzmaurice. Mouselight: Bimanual interactions on digital paper using a pen and a spatially-aware mobile projector. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2451–2460, New York, NY, USA, 2010. ACM.

[103] J. Song, G. Sörös, F. Pece, S. R. Fanello, S. Izadi, C. Keskin, and O. Hilliges. In-air gestures around unmodified mobile devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pages 319–329, New York, NY, USA, 2014. ACM.

[104] M. Spindler, W. Büschel, and R. Dachselt. Use your head: Tangible windows for 3d information spaces in a tabletop environment. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, pages 245–254, New York, NY, USA, 2012. ACM.

[105] M. Spindler, W. Büschel, C. Winkler, and R. Dachselt. Tangible displays for the masses: spatial interaction with handheld displays by using consumer depth cameras. *Personal and Ubiquitous Computing*, 18 (5):1213–1225, Jun 2014.

[106] M. Spindler and R. Dachselt. Paperlens: Advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, pages 7:1–7:1, New York, NY, USA, 2009. ACM.

[107] M. Spindler, M. Martsch, and R. Dachselt. Going beyond the surface: studying multi-layer interaction above the tabletop. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1277–1286. ACM, 2012.

[108] M. Spindler, M. Schuessler, M. Martsch, and R. Dachselt. Pinch-drag-flick vs. spatial input: Rethinking zoom & pan on mobile displays. In



© 2017 Tous droits réservés. lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 1113–1122, New York, NY, USA, 2014. ACM.

[109] A. Sugiura and Y. Koseki. A user interface using fingerprint recognition: Holding commands and data objects on fingers. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology*, UIST '98, pages 71–79, New York, NY, USA, 1998. ACM.

[110] C. Telkenaroglu and T. Capin. Dual-finger 3d interaction techniques for mobile devices. *Personal Ubiquitous Comput.*, 17(7):1551–1572, Oct. 2013.

[111] Vinayak, S. Murugappan, H. Liu, and K. Ramani. Shape-it-up: Hand gesture based creative expression of 3d shapes using intelligent generalized cylinders. *Comput. Aided Des.*, 45(2):277–287, Feb. 2013.

[112] D. Vogel and R. Balakrishnan. Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, pages 137–146, New York, NY, USA, 2004. ACM.

[113] N. Weibel, A. M. Piper, and J. D. Hollan. Hiperpaper: Introducing pen and paper interfaces for ultra-scale wall displays. In *Proc. of UIST*, 2010.

[114] M. Weiser. The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11, July 1999.

[115] G. Wesche and H.-P. Seidel. Freedrawer: A free-form sketching system on the responsive workbench. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '01, pages 167–174, New York, NY, USA, 2001. ACM.

[116] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2011.

[117] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2011.

[118] A. D. Wilson. Playanywhere: A compact interactive tabletop projection-vision system. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, UIST '05, pages 83–92, New York, NY, USA, 2005. ACM.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr

Thèse de Hanaë Rateau, Lille 1, 2017



[119] A. D. Wilson and H. Benko. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 273–282, New York, NY, USA, 2010. ACM.

[120] J. Wobbrock, M. Morris, and A. Wilson. User-defined gestures for surface computing. In *Proc. of CHI '09*, pages 1083–1092, NY, USA, 2009. ACM.

[121] K.-P. Yee. Peephole displays: Pen interaction on spatially aware handheld computers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 1–8, New York, NY, USA, 2003. ACM.

[122] Y. Zhai, G. Zhao, T. Alatalo, J. Heikkilä, T. Ojala, and X. Huang. Gesture interaction for wall-sized touchscreen display. In *Proc. of UbiComp*, 2013.

[123] Z. Zhang, Y. Wu, Y. Shan, and S. Shafer. Visual panel: Virtual mouse, keyboard and 3d controller with an ordinary piece of paper. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces*, PUI '01, pages 1–8, New York, NY, USA, 2001. ACM.



© 2017 Tous droits réservés.

lilliad.univ-lille.fr